WebSphere® Information Integration

IBM

Version 9

Administration Guide for Federated Systems

SC19-1020-00

WebSphere® Information Integration

IBM

Version 9

Administration Guide for Federated Systems

SC19-1020-00

Note

Before using this information and the product that it supports, be sure to read the general information under "Notices and trademarks" on page 353.

Contents

Chapter 1. Overview of a federated system

Federated systems

A *federated system* is a special type of distributed database management system (DBMS). A federated system consists of a DB2® instance that operates as a federated server, a database that acts as the federated database, one or more data sources, and clients (users and applications) that access the database and data sources.

With a federated system, you can send distributed requests to multiple data sources within a single SQL statement. For example, you can join data that is located in a DB2 table, an Oracle table, and an XML tagged file in a single SQL statement. The following figure shows the components of a federated system and a sample of the data sources you can access.

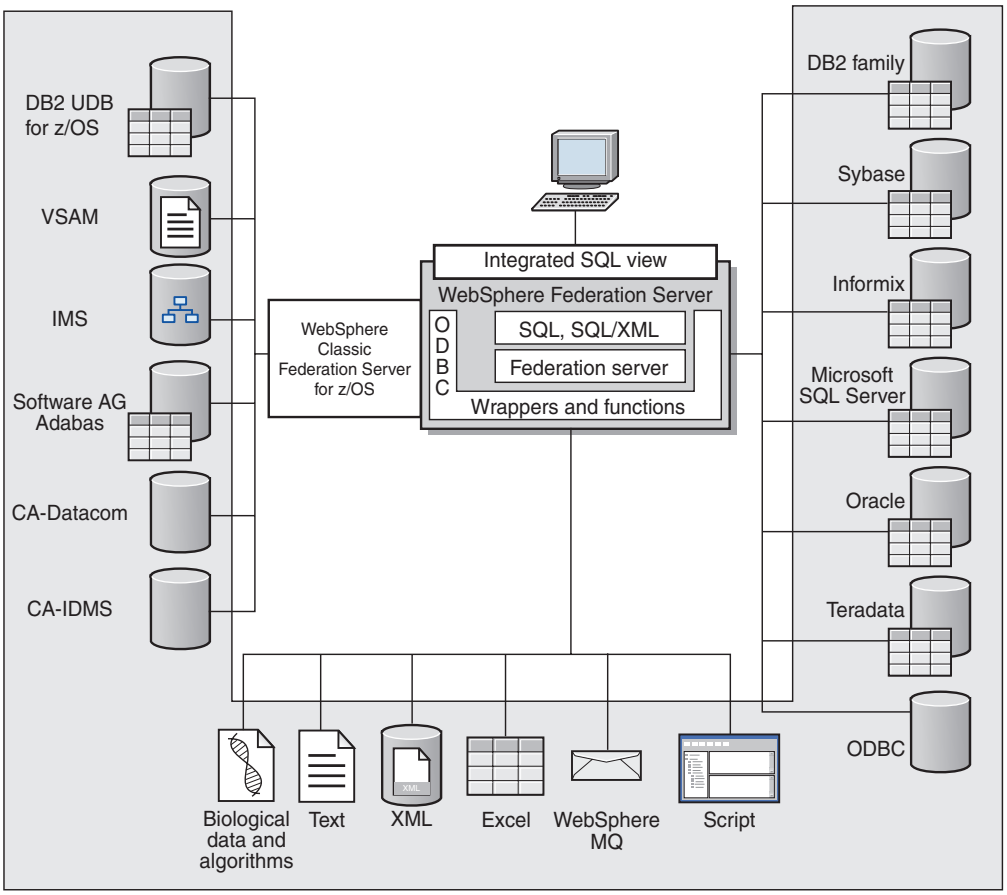


Figure 1. The components of a federated system

The power of a federated system is in its ability to:

- Correlate data from local tables and remote data sources, as if all the data is stored locally in the federated database
- Update data in relational data sources, as if the data is stored in the federated database

- Move data to and from relational data sources
- Take advantage of the data source processing strengths, by sending requests to the data sources for processing
- Compensate for SQL limitations at the data source by processing parts of a distributed request at the federated server

The federated server

The DB2® server in a federated system is referred to as the federated server. Any number of DB2 instances can be configured to function as federated servers. You can use existing DB2 instances as your federated servers, or you can create new ones specifically for the federated system.

The DB2 instance that manages the federated system is called a server because it responds to requests from end users and client applications. The federated server often sends parts of the requests it receives to the data sources for processing. A pushdown operation is an operation that is processed remotely. The DB2 instance that manages the federated system is referred to as the federated server, even though it acts as a client when it pushes down requests to the data sources.

Like any other application server, the federated server is a database manager instance. Application processes connect and submit requests to the database within the federated server. However, two main features distinguish it from other application servers:

- A federated server is configured to receive requests that might be partially or entirely intended for data sources. The federated server distributes these requests to the data sources.
- Like other application servers, a federated server uses DRDA® communication protocols (over TCP/IP) to communicate with DB2 family instances. However, unlike other application servers, a federated server uses the native client of the data source to access the data source. For example, a federated server uses the Sybase Open Client to access Sybase data sources and an Microsoft® SQL Server ODBC Driver to access Microsoft SQL Server data sources.

What is a data source?

In a federated system, a *data source* can be a relational database (such as Oracle or Sybase) or a nonrelational data source (such as a BLAST search algorithm or an XML tagged file).

Through some data sources you can access other data sources. For example, with the ODBC wrapper you can access WebSphere® Classic Federation server for z/OS® data sources such as DB2 UDB for z/OS, IMS™, CA-IDMS, CA-Datacom, Software AG Adabas, and VSAM.

The method, or protocol, used to access a data source depends on the type of data source. For example, DRDA® is used to access DB2® for z/OS™ data sources.

Data sources are autonomous. For example, the federated server can send queries to Oracle data sources at the same time that Oracle applications can access these data sources. A federated system does not monopolize or restrict access to the other data sources, beyond integrity and locking constraints.

The federated database

To end users and client applications, data sources appear as a single collective database in DB2®. Users and applications interface with the *federated database* that is managed by the federated server.

The federated database contains a system catalog that stores information about data. The federated database system catalog contains entries that identify data sources and their characteristics. The federated server consults the information stored in the federated database system catalog and the data source wrapper to determine the best plan for processing SQL statements.

The federated system processes SQL statements as if the data from the data sources were ordinary relational tables or views within the federated database. As a result:

- The federated system can correlate relational data with data in nonrelational formats. This is true even when the data sources use different SQL dialects, or do not support SQL at all.
- The characteristics of the federated database take precedence when there are differences between the characteristics of the federated database and the characteristics of the data sources. Query results conform to DB2 semantics, even if data from other non-DB2 data sources is used to compute the query result.

  Examples:

  – The code page that the federated server uses is different than the code page used that the data source uses. In this case, character data from the data source is converted based on the code page used by the federated database, when that data is returned to a federated user.
  – The collating sequence that the federated server uses is different than the collating sequence that the data source uses. In this case, any sort operations on character data are performed at the federated server instead of at the data source.

Wrappers and wrapper modules

Wrappers are mechanisms by which the federated database interacts with data sources. The federated database uses routines stored in a library called a *wrapper module* to implement a wrapper.

These routines allow the federated database to perform operations such as connecting to a data source and retrieving data from it iteratively. Typically, the federated instance owner uses the CREATE WRAPPER statement to register a wrapper in the federated database. You can register a wrapper as fenced or trusted using the DB2_FENCED wrapper option.

You create one wrapper for each type of data source that you want to access. For example, you want to access three DB2 for z/OS™ database tables, one DB2 for iSeries™ table, two Informix® tables, and one Informix view. In this case, you need to create one wrapper for the DB2 data source objects and one wrapper for the Informix data source objects. After these wrappers are registered in the federated database, you can use these wrappers to access other objects from those data sources. For example, you can use the DRDA® wrapper with all DB2 family data source objects—DB2 Version 9.1 for Linux, UNIX®, and Windows®, DB2 for z/OS, DB2 for iSeries, and DB2 Server for VM and VSE.

You use the server definitions and nicknames to identify the specifics (name, location, and so forth) of each data source object.

A wrapper performs many tasks. Some of these tasks are:

- It connects to the data source. The wrapper uses the standard connection API of the data source.
- It submits queries to the data source.
  - For data sources that support SQL, the query is submitted in SQL.
  - For data sources that do not support SQL, the query is translated into the native query language of the source or into a series of source API calls.
- It receives results sets from the data source. The wrapper uses the data source standard APIs for receiving results set.
- It responds to federated database queries about the default data type mappings for a data source. The wrapper contains the default type mappings that are used when nicknames are created for a data source object. For relational wrappers, data type mappings that you create override the default data type mappings. User-defined data type mappings are stored in the global catalog.
- It responds to federated database queries about the default function mappings for a data source. The federated database needs data type mapping information for query planning purposes. The wrapper contains information that the federated database needs to determine if DB2 functions are mapped to functions of the data source, and how the functions are mapped. This information is used by the SQL Compiler to determine if the data source is able to perform the query operations. For relational wrappers, function mappings that you create override the default function type mappings. User-defined function mappings are stored in the global catalog.

Wrapper options are used to configure the wrapper or to define how WebSphere Federation Server uses the wrapper.

How you interact with a federated system

Because the federated database is a DB2® family database, you can interact with it in a variety of ways.

You can interact with a federated system using any one of these methods:

- The DB2 command line processor (CLP)
- The DB2 Command Center GUI
- The DB2 Control Center GUI
- Application programs
- DB2 family tools
- Web services providers

The steps in the federated documentation provide the commands and SQL statements that can be entered in the DB2 command line processor or the DB2 Command Center GUI. The documentation indicates when tasks can be performed through the DB2 Control Center GUI. Since the DB2 Control Center GUI is intuitive, the steps to perform these tasks through the DB2 Control Center are not included in this documentation.

DB2 command line processor (CLP)

You can perform most of the tasks necessary to setup, configure, tune, and maintain the federated system through the DB2 command line processor. In some cases you can use either the DB2 command line processor or the DB2 Command Center.

For example, you can use the DB2 command line processor or the DB2 Command Center to perform the following tasks:

- Create, alter, or drop user-defined data type mappings
- Create, alter, or drop user-defined function mappings

DB2 Command Center

Through the DB2 Command Center, you can create and run distributed requests without having to manually type out lengthy SQL statements.

Use the DB2 Command Center when you are tuning the performance of the federated system. The DB2 Command Center is a convenient way to use the DB2 Explain functionality to look at the access plans for distributed requests. The DB2 Command Center can also be used to work with the SQL Assistant tool.

DB2 Control Center

The DB2 Control Center GUI allows you to perform most of the tasks necessary to setup, configure, and modify the federated system. The DB2 Control Center uses panels—dialog boxes and wizards—to guide you through a task.

The DB2 Control Center panels contain interactive help when your mouse hovers over a control such as a list box or command button. Additionally, each panel has a help button that provides information about the panel task, and links to related concepts and reference information.

You can either use a wizard to create the federated objects, or you can create each object individually.

Use the DB2 Control Center to configure access to Web services, WebSphere® Business Integration, and XML data sources. The features built into the DB2 Control Center simplify the steps that are required for you to configure the federated server to access these data sources.

The DB2 Control Center GUI is the easiest way to perform the essential data source configuration tasks:

- Create the wrappers and set the wrapper options
- Specify the environment variables for your data source
- Create the server definitions and set the server options
- Create the user mappings and set the user options
- Create the nicknames and set the nickname options or column options

After you configure the federated server to access your data sources, you can use the DB2 Control Center to:

- Modify the data source configuration
- Monitor the status of the nicknames and servers
- Maintain current statistics for your nicknames
- Create and modify cache tables
- Specify informational constraints on nicknames
- Create remote tables through WebSphere Federation Server using transparent DDL

Application programs

Applications do not require any special coding to work with federated data. Applications access the system just like any other DB2 client application. Applications interface with the federated database that is within the federated server.

To obtain data from data sources, applications submit queries in DB2 SQL to the federated database. The federated database then distributes the queries to the appropriate data sources, collects the requested data, and returns this data to the applications. However, since the federated database interacts with the data sources through nicknames, you need to be aware of:

- The SQL restrictions you have when working with nicknames
- How to perform operations on nicknamed objects

DB2 family tools

You can interact with a federated database using host and midrange tools.

Host and midrange tools such that you can use to interact with a federated database include:

- DB2 SPUFI on DB2 for z/OS™
- Interactive SQL (STRSQL) on DB2 for iSeries™

Rational Data Architect

You can use the metadata tools that are provided with Rational® Data Architect to enhance and extend the information that you access with WebSphere Federation Server federated functions.

With Rational Data Architect, you can find and map relationships among a variety of data sources, build scripts that represent those relationships, and then deploy scripts to federated, non-federated, local, or remote servers.

By using Rational Data Architect with WebSphere Federation Server, you can share the metadata between the heterogeneous products. Rational Data Architect provides tools that support the management, impact analysis, search, and reporting across all forms of metadata. You can then map your heterogeneous data and structures and deploy to a server.

Web services providers

You can interact with a federated database through web services providers.

Within federated systems, a Web services wrapper is available to access Web services with SQL statements on nicknames and views that invoke Web services. You can create a Web services wrapper and nicknames that specify input to the Web service and access the output from the Web service with SELECT statements.

Supported data sources

There are many data sources that you can access using a federated system.

The following table lists the supported data sources:

Table 1. Supported data source versions and access methods.

Data source	Supported versions	Access methods and requirements
BLAST	2.2.3 through 2.2.8 fixpacks supported	BLAST daemon (supplied with the wrapper)
BioRS	5.2.x.x	HTTP
DB2 Database for Linux®, UNIX®, and Windows®	8.1, 8.2, 9.1	DRDA®
DB2 Universal Database™ for z/OS	6.1, 7.1 with the following APARs applied: • PQ62695 • PQ55393 • PQ56616 • PQ54605 • PQ46183 • PQ62139 8.1	DRDA
DB2 Universal Database for iSeries™	5.2 with the following APARs and PTFs applied: • APAR SE06003, PTF SI04582 • APAR SE07533, PTF SI05991 • APAR SE08416, PTF SI07135 • APAR II13348, PTFs SF99502, SI11626, SI11378 5.3 5.4 with APAR SE23546, PTF SI21661 applied.	DRDA
DB2 Server for VM and VSE	7.1 (or later) with fixes for APARs for schema functions applied.	DRDA
Entrez	Supported	HTTP.
GenBank	Supported	HTTP. Connection to the NCBI through the Web. Use the Entrez wrapper to access this data source.
HMMER	2.2g, 2.3	HMMER daemon (supplied with the wrapper)

Table 1. Supported data source versions and access methods. (continued)

Data source	Supported versions	Access methods and requirements
Informix®	7.31, 8.4, 8.5, 9.4, 10.0	Informix Client SDK V2.81 (or later) On Solaris, the Informix client SDK version 2.81.FC2 is not supported. If you are using the Informix client version 2.81.FC2, update the client to version 2.81.FC2R1 or later. On Windows, the Informix client SDK version 2.81.TC2 or later. On the 64-bit mode zLinux operating system, the Informix client version 2.81.FC1 or 2.81.FC2 is not supported. If you are using one of those client versions, update the client to version 2.81.FC3 or later.
KEGG	KEGG API 3.2	User-defined functions for KEGG
Microsoft® Excel	97, 2000, 2002, 2003	Excel 97, 2000, 2002, or 2003 installed on the federated server
Microsoft SQL Server	2000 SP3 and later service packs on that release, 2005	On Windows, the Microsoft SQL Server Client ODBC 3.0 (or later) driver. On UNIX: • DataDirect Technologies (formerly MERANT) Connect ODBC 4.2 (or later) driver. • Microsoft SQL Server wrapper with a UTF-8 database requires DataDirect Connect for ODBC 4.2 Service Pack 2 or later.
ODBC	3.x	ODBC driver for the data source. ODBC driver access to Redbrick and ODBC driver access to WebSphere Classic Federation Server for z/OS data sources, such as IMS, VSAM, CA-Datacom, CA-IDMS, and Software AG Adabas.
OLE DB	2.7, 2.8	OLE DB 2.0 (or later)

Table 1. Supported data source versions and access methods. (continued)

Data source	Supported versions	Access methods and requirements
OMIM	Supported	HTTP. Connection to the NCBI through the Web and the OMIM query.fcgi utility. Use the Entrez wrapper to access this data source.
Oracle	8.1.7, 9.0, 9.1, 9.2, 9i, 10g	Oracle net client or NET8 client software
PeopleSoft	8.x	IBM® WebSphere Business Integration Adapter for PeopleSoft v2.3.1, 2.4. Requires WebSphere MQ Series.
PubMed	Supported	HTTP. Connection to the NCBI through the Web. Use the Entrez wrapper to access this data source.
SAP	3.x, 4.x	IBM WebSphere Business Integration Adapter for mySAP.com v2.3.1, 2.4. Requires WebSphere MQ Series.
Script		Script daemon (supplied with the wrapper)
Siebel	7, 7.5, 2000	IBM WebSphere Business Integration Adapter for Siebel eBusiness Applications v2.3.1, 2.4. Requires WebSphere MQ Series.
Sybase	12.0, 12.5	Sybase Open Client ctlib interface
Table-structured files		None
Teradata	V2R4, V2R5, V2R6	Teradata Call-Level Interface, Version 2 (CLIv2) Release 04.06 (or later) On Windows, the Teradata client TTU 7.0 or later and the Teradata API library CLIv2 4.7.0 or later on the federated server.
Web services	SOAP 1.0., 1.1, WSDL 1.0, 1.1 specifications	HTTP, HTTPS. SOAP user-defined functions consume Web services.
WebSphere MQ	Server edition 6.0	WebSphere MQ user-defined functions in schemas DB2MQ, DB2MQ1C, and DB2MQT.
XML	1.0 specification	None

The federated database system catalog

The federated database system catalog contains information about the objects in the federated database and information about objects at the data sources.

The catalog in a federated database is called the global catalog because it contains information about the entire federated system. DB2® query optimizer uses the information in the global catalog and the data source wrapper to plan the best way to process SQL statements. The information stored in the global catalog includes remote and local information, such as column names, column data types, column default values, index information, and statistics information.

Remote catalog information is the information or name used by the data source. Local catalog information is the information or name used by the federated database. For example, suppose a remote table includes a column with the name of *EMPNO*. The global catalog would store the remote column name as *EMPNO*. Unless you designate a different name, the local column name will be stored as *EMPNO*. You can change the local column name to *Employee_Number*. Users submitting queries which include this column will use *Employee_Number* in their queries instead of *EMPNO*. You use the ALTER NICKNAME statement to change the local name of the data source columns.

For relational and nonrelational data sources, the information stored in the global catalog includes both remote and local information.

To see the data source table information that is stored in the global catalog, query the SYSCAT.TABLES, SYSCAT.NICKNAMES, SYSCAT.TABOPTIONS, SYSCAT.INDEXES, SYSCAT.INDEXOPTIONS, SYSCAT.COLUMNS, and SYSCAT.COLOPTIONS catalog views in the federated database.

The global catalog also includes other information about the data sources. For example, the global catalog includes information that the federated server uses to connect to the data source and map the federated user authorizations to the data source user authorizations. The global catalog contains attributes about the data source that you explicitly set, such as server options.

The SQL compiler

The DB2 SQL compiler gathers information to help it process queries.

To obtain data from data sources, users and applications submit queries in SQL to the federated database. When a query is submitted, the DB2 SQL compiler consults information in the global catalog and the data source wrapper to help it process the query. This includes information about connecting to the data source, server information, mappings, index information, and processing statistics.

The query optimizer

As part of the SQL compiler process, the query optimizer analyzes a query. The compiler develops alternative strategies, called access plans, for processing the query.

Access plans might call for the query to be:

- Processed by the data sources
- Processed by the federated server

- Processed partly by the data sources and partly by the federated server

The query optimizer evaluates the access plans primarily on the basis of information about the data source capabilities and the data. The wrapper and the global catalog contain this information. The query optimizer decomposes the query into segments that are called query fragments. Typically it is more efficient to pushdown a query fragment to a data source, if the data source can process the fragment. However, the query optimizer takes into account other factors such as:

- The amount of data that needs to be processed
- The processing speed of the data source
- The amount of data that the fragment will return
- The communication bandwidth
- Whether there is a usable materialized query table on the federated server that represents the same query result

The query optimizer generates access plan alternatives for processing a query fragment. The plan alternatives perform varying amounts of work locally on the federated server and on the remote data sources. Because the query optimizer is cost-based, it assigns resource consumption costs to the access plan alternatives. The query optimizer then chooses the plan that will process the query with the least resource consumption cost.

If any of the fragments are to be processed by data sources, the federated database submits these fragments to the data sources. After the data sources process the fragments, the results are retrieved and returned to the federated database. If the federated database performed any part of the processing, it combines its results with the results retrieved from the data source. The federated database then returns all results to the client.

Compensation

The ability by WebSphere Federation Server to process SQL that is not supported by a data source is called *compensation*.

The federated database does not push down a query fragment if the data source cannot process it, or if the federated server can process it faster than the data source can process it. For example, suppose that the SQL dialect of a data source does not support a CUBE grouping in the GROUP BY clause. A query that contains a CUBE grouping and references a table in that data source is submitted to the federated server. The federated database does not pushdown the CUBE grouping to the data source, but processes the CUBE itself.

The federated database compensates for lack of functionality at the data source in two ways:

- It can request that the data source use one or more operations that are equivalent to the DB2 function stated in the query. For example, a data source does not support the cotangent (COT(x)) function, but supports the tangent (TAN(x)) function. The federated database can request that the data source perform the calculation (1/TAN(x)), which is equivalent to the cotangent (COT(x)) function.
- It can return the set of data to the federated server, and perform the function locally.

For relational data sources, each type of RDBMS supports a subset of the international SQL standard. In addition, some types of RDBMSs support SQL constructs that exceed this standard. An *SQL dialect*, is the totality of SQL that a type of RDBMS supports. If an SQL construct is found in the DB2 SQL dialect, but not in the relational data source dialect, the federated server can implement this construct on behalf of the data source.

The federated database can compensate for differences in SQL dialects. An example of this ability is the common-table-expression clause. DB2 SQL includes the clause common-table-expression. In this clause, a name can be specified by which all FROM clauses in a fullselect can reference a result set. The federated server will process a common-table-expression for a data source, even though the SQL dialect used by the data source does not include common-table-expression.

With compensation, the federated database can support the full DB2 SQL dialect for queries of data sources. Even data sources with weak SQL support or no SQL support will benefit from compensation. You must use the DB2 SQL dialect with a federated system, except in a pass-through session.

Pass-through sessions

You can submit SQL statements directly to data sources by using a special mode called *pass-through*.

You submit SQL statements in the SQL dialect used by the data source. Use a pass-through session when you want to perform an operation that is not possible with the DB2® SQL/API. For example, use a pass-through session to create a procedure, create an index, or perform queries in the native dialect of the data source.

Currently, the data sources that support pass-through, support pass-through using SQL. In the future, it is possible that data sources will support pass-though using a data source language other than SQL.

Similarly, you can use a pass-through session to perform actions that are not supported by SQL, such as certain administrative tasks. However, you cannot use a pass-through session to perform all administrative tasks. For example, you can create or drop tables at the data source, but you cannot start or stop the remote database.

You can use both static and dynamic SQL in a pass-through session.

The federated server provides the following SQL statements to manage pass-through sessions:

SET PASSTHRU
: Opens a pass-through session. When you issue another SET PASSTHRU statement to start a new pass-through session, the current pass-through session is terminated.

SET PASSTHRU RESET
: Terminates the current pass-through session.

GRANT (Server Privileges)
: Grants a user, group, list of authorization IDs, or PUBLIC the authority to initiate pass-through sessions to a specific data source.

REVOKE (Server Privileges)
Revokes the authority to initiate pass-through sessions.

Default wrapper names

There are wrappers for each supported data source. Some wrappers have default wrapper names.

When you use the default name to create the wrapper, the federated server automatically picks up the data source library associated with the wrapper.

Table 2. Data sources with default wrapper names.

Data source	Default wrapper names
DB2 Version 9.1 for Linux, UNIX and Windows®	DRDA
DB2 Universal Database for z/OS	DRDA
DB2 Universal Database for iSeries	DRDA
DB2 Server for VM and VSE	DRDA
Informix	INFORMIX
Microsoft® SQL Server	MSSQLODBC3
ODBC	ODBC
OLE DB	OLEDB
Oracle	NET8
Sybase	CTLIB
Teradata	TERADATA

Server definitions and server options

After wrappers are created for the data sources, the federated instance owner defines the data sources to the federated database.

The instance owner supplies a name to identify the data source, and other information that pertains to the data source. This information includes:

- The type and version of the data source
- The database name for the data source (RDBMS only)
- Metadata that is specific to the data source

For example, a DB2® family data source can have multiple databases. The definition must specify which database the federated server can connect to. In contrast, an Oracle data source has one database, and the federated server can connect to the database without knowing its name. The database name is not included in the federated server definition of an Oracle data source.

The name and other information that the instance owner supplies to the federated server are collectively called a *server definition*. Data sources answer requests for data and are servers in their own right.

The CREATE SERVER and ALTER SERVER statements are used to create and modify a server definition.

Some of the information within a server definition is stored as *server options*. When you create server definitions, it is important to understand the options that you can specify about the server.

Server options can be set to persist over successive connections to the data source, or set for the duration of a single connection.

User mappings

Typically, you need to define an association between the federated server and a data source.

When a federated server needs to pushdown a request to a data source, the server must first establish a connection to the data source. For most data sources, the federated server does this by using a valid user ID and password to that data source. When a user ID and password is required to connect to a data source, you can define an association between the federated server authorization ID and the data source user ID and password. This association can be created for each user ID that will be using the federated system to send distributed requests. This association is called a *user mapping*.

In some cases, you do not need to create a user mapping if the user ID and password that you use to connect to the federated database are the same as those that you use to access the remote data source. You can create and store the user mappings in the federated database, or you can store the user mappings in an external repository, such as LDAP.

Nicknames and data source objects

After you create the server definitions and user mappings, the federated instance owner creates the nicknames. A nickname is an identifier that is used to reference the object located at the data sources that you want to access. The objects that nicknames identify are referred to as data source objects.

Nicknames are not alternative names for data source objects in the same way that aliases are alternative names. They are pointers by which the federated server references these objects. Nicknames are typically defined with the CREATE NICKNAME statement along with specific nickname column options and nickname options.

When an end user or a client application submits a distributed request to the federated server, the request does not need to specify the data sources. Instead, the request references the data source objects by their nicknames. The nicknames are mapped to specific objects at the data source. These mappings eliminate the need to qualify the nicknames by data source names. The location of the data source objects is transparent to the end user or the client application.

Suppose that you define the nickname *DEPT* to represent an Informix® database table called *NFX1.PERSON*. The statement SELECT * FROM *DEPT* is allowed from the federated server. However, the statement SELECT * FROM *NFX1.PERSON* is not allowed from the federated server (except in a pass-through session) unless there is a local table on the federated server named *NFX1.PERSON*.

When you create a nickname for a data source object, metadata about the object is added to the global catalog. The query optimizer uses this metadata, and the information in the wrapper, to facilitate access to the data source object. For

example, if the nickname is for a table that has an index, the global catalog contains information about the index. The wrapper contains the mappings between the DB2® data types and the data source data types.

Nickname data based on data source objects that use a labeling security system are not ordinarily cached, so data in the object remains secure. For example with the Oracle Net8 wrapper, if you create a nickname on an Oracle table that uses Oracle Label Security, the table is automatically identified as secure. The resulting nickname data cannot be cached which means that materialized query tables cannot be created on it. This setting ensures that the information is viewed only by users with the appropriate authority on the Oracle system. Nicknames that were created on data source objects that use Oracle Label Security before this feature was available need to be changed to disallow caching. You can use the ALTER NICKNAME statement to allow or disallow caching on a nickname. All other nicknames based on wrappers other than Oracle Net8 must use the ALTER NICKNAME statement to disable data caching on the federated server.

Currently, you cannot execute some DB2 utility operations on nicknames, such as RUNSTATS.

You cannot use the Cross Loader utility to cross load into a nickname.

Valid data source objects

Nicknames identify objects at the data source that you want to access. The following table lists the types of objects that you can create a nickname for in a federated system.

Table 3. Valid data source objects

Data source	Valid objects
BioRS	BioRS databanks
BLAST	FASTA files indexed for BLAST search algorithms
DB2 Version 9.1 for Linux, UNIX, and Windows	Nicknames, materialized query tables, tables, views
DB2 for iSeries	Tables, views, physical and logical files
DB2 for VM and VSE	Tables, views
DB2 for z/OS	Tables, views
Entrez	Entrez databases
HMMER	HMM database files (libraries of Hierarchical Markov Models, such as PFAM), that can be searched by HMMER's hmmpfam or hmmsearch programs.
Informix	Tables, views, synonyms
Microsoft Excel	.xls files (only the first sheet in the workbook is accessed)
Microsoft SQL Server	Tables, views
ODBC	Tables, views
Oracle	Tables, views, synonyms
Script	Scripts
Sybase	Tables, views

Table 3. Valid data source objects (continued)

Data source	Valid objects
Teradata	Tables, views
Table-structured files	Text files that meet a specific format.
Websphere Business Integration adapters	Websphere Business Integration business objects that map to BAPIs in SAP, business components in Siebel, and component interfaces in PeopleSoft
Web Services	Operations in a Web services description language file
XML-tagged files	Sets of items in an XML document

Nickname column options

You can supply the global catalog with additional metadata information about the nicknamed object. This metadata describes values in certain columns of the data source object. You assign this metadata to parameters that are called *nickname column options*.

The nickname column options tell the wrapper to handle the data in a column differently than it normally would handle it. The SQL complier and query optimizer use the metadata to develop better plans for accessing the data.

Nickname column options are used to provide other information to the wrapper as well. For example for XML data sources, a nickname column option is used to tell the wrapper the XPath expression to use when the wrapper parses the column out of the XML document.

With federation, the DB2® server treats the data source object that a nickname references as if it is a local DB2 table. As a result, you can set nickname column options for any data source object that you create a nickname for. Some nickname column options are designed for specific types of data sources and can be applied only to those data sources.

Suppose that a data source has a collating sequence that differs from the federated database collating sequence. The federated server typically would not sort any columns containing character data at the data source. It would return the data to the federated database and perform the sort locally. However, suppose that the column is a character data type (CHAR or VARCHAR) and contains only numeric characters ('0','1',...,'9'). You can indicate this by assigning a value of 'Y' to the NUMERIC_STRING nickname column option. This gives the DB2 query optimizer the option of performing the sort at the data source. If the sort is performed remotely, you can avoid the overhead of porting the data to the federated server and performing the sort locally.

You can define nickname column options for relational nicknames using the ALTER NICKNAME statements. You can define nickname column options for nonrelational nicknames using the CREATE NICKNAME and ALTER NICKNAME statements.

Data type mappings

The data types at the data source must map to corresponding DB2® data types so that the federated server can retrieve data from data sources.

Some examples of default data type mappings are:

- The Oracle type FLOAT maps to the DB2 type DOUBLE
- The Oracle type DATE maps to the DB2 type TIMESTAMP
- The DB2 for z/OS™ type DATE maps to the DB2 type DATE

For most data sources, the default type mappings are in the wrappers. The default type mappings for DB2 data sources are in the DRDA® wrapper. The default type mappings for Informix® are in the INFORMIX wrapper, and so forth.

For some nonrelational data sources, you must specify data type information in the CREATE NICKNAME statement. The corresponding DB2 data types must be specified for each column of the data source object when the nickname is created. Each column must be mapped to a particular field or column in the data source object.

For relational data sources, you can override the default data type mappings. For example, by default the Informix INTEGER data type maps to the DB2 INTEGER data type. You could override the default mappings and map Informix's INTEGER data type to DB2 DECIMAL(10,0) data type.

Function mappings

For the federated server to recognize a data source function, the function must be mapped to an existing counterpart function in DB2® Version 9.1 for Linux, UNIX® and Windows®.

WebSphere Federation Server supplies default mappings between existing data source functions and DB2 counterpart functions. For most data sources, the default function mappings are in the wrappers. The default function mappings to DB2 for z/OS™ functions are in the DRDA® wrapper. The default function mappings to Sybase functions are in the CTLIB wrapper, and so forth.

For relational data sources, you can create a function mapping when you want to use a data source function that the federated server does not recognize. The mapping that you create is between the data source function and a DB2 counterpart function at the federated database. Function mappings are typically used when a new built-in function or a new user-defined function become available at the data source. Function mappings are also used when a DB2 counterpart function does not exist. In this case, you must also create a function template.

Index specifications

When you create a nickname for a data source table, information about any indexes that the data source table has is added to the global catalog. The query optimizer uses this information to expedite the processing of distributed requests. The catalog information about a data source index is a set of metadata, and is called an index specification.

The query optimizer uses this information to expedite the processing of distributed requests.

A federated server does not create an index specification when you create a nickname for:

- A table that has no indexes
- A view, which typically does not have any index information stored in the remote catalog
- A data source object that does not have a remote catalog from which the federated server can obtain the index information

Suppose that a table acquires a new index, in addition to the ones it had when the nickname was created. Because index information is supplied to the global catalog at the time the nickname is created, the federated server is unaware of the new index. Similarly, when a nickname is created for a view, the federated server is unaware of the underlying table (and its indexes) from which the view was generated. In these circumstances, you can supply the necessary index information to the global catalog. You can create an index specification for tables that have no indexes. The index specification tells the query optimizer which column or columns in the table to search on to find data quickly.

Federated stored procedures

Federated procedure access enables users of federated systems to access remote stored procedures at remote data sources.

A federated stored procedure is a local stored procedure that is mapped to a stored procedure at a data source. You use the CREATE PROCEDURE (Sourced) statement to register a federated stored procedure.

Collating sequences

The order in which character data is sorted in a database depends on the structure of the data and the collating sequence defined for the database.

Suppose that the data in a database is all uppercase letters and does not contain any numeric or special characters. A sort of the data should result in the same output, regardless of whether the data is sorted at the data source or at the federated database. The collating sequence used by each database should not impact the sort results. Likewise, if the data in the database is all lowercase letters or all numeric characters, a sort of the data should produce the same results regardless of where the sort actually is performed.

If the data consists of any of the following structures:

- A combination of letters and numeric characters
- Both uppercase and lowercase letters
- Special characters such as @, #, €

Sorting this data can result in different outputs, if the federated database and the data source use different collating sequences.

In general terms, a collating sequence is a defined ordering for character data that determines whether a particular character sorts higher, lower, or the same as another character.

How collating sequences determine sort orders

A collating sequence determines the sort order of the characters in a coded character set.

A character set is the aggregate of characters that are used in a computer system or programming language. In a coded character set, each character is assigned to a different number within the range of 0 to 255 (or the hexadecimal equivalent thereof). The numbers are called code points; the assignments of numbers to characters in a set are collectively called a code page.

In addition to being assigned to a character, a code point can be mapped to the character's position in a sort order. In technical terms, then, a collating sequence is the collective mapping of a character set's code points to the sort order positions of the set's characters. A character's position is represented by a number; this number is called the weight of the character. In the simplest collating sequence, called an identity sequence, the weights are identical to the code points.

Suppose that database ALPHA uses the default collating sequence of the EBCDIC code page, and that database BETA uses the default collating sequence of the ASCII code page. Sort orders for character strings at these two databases would differ, as shown in the following example:

```
SELECT.....
  ORDER BY COL2

EBCDIC-Based Sort           ASCII-Based Sort

COL2                        COL2
----                        ----
V1G                         7AB
Y2W                         V1G
7AB                         Y2W
```

Similarly, character comparisons in a database depend on the collating sequence defined for that database. In this example, database ALPHA uses the default collating sequence of the EBCDIC code page. Database BETA uses the default collating sequence of the ASCII code page. Character comparisons at these two databases would yield different results, as shown in the following example:

```
SELECT.....
  WHERE COL2 > 'TT3'

EBCDIC-Based Results     ASCII-Based Results

COL2                     COL2
----                     ----
TW4                      TW4
X82                      X82
39G
```

Setting the local collating sequence to optimize queries

Administrators can create federated databases with a particular collating sequence that matches a data source collating sequence.

Then for each data source server definition, the COLLATING_SEQUENCE server option is set to 'Y'. This setting tells the federated database that the collating sequences of the federated database and the data source match.

You set the federated database collating sequence as part of the CREATE DATABASE API. Through this API, you can specify one of the following sequences:

- An identity sequence
- A system sequence (the sequence used by the operating system that supports the database)
- A customized sequence (a predefined sequence that DB2 supplies or that you define yourself)

Suppose that the data source is DB2 for z/OS. Sorts that are defined in an ORDER BY clause are implemented by a collating sequence based on an EBCDIC code page. To retrieve DB2 for z/OS data sorted in accordance with ORDER BY clauses, configure the federated database so that it uses the predefined collating sequence based on the appropriate EBCDIC code page.

Chapter 2. Modifying data source configurations

Altering a wrapper (DB2 Control Center)

After you configure a wrapper, you can use the ALTER WRAPPER statement to modify the configuration based on your system requirements. You can alter a wrapper by using the DB2 Control Center or by using the DB2 command line.

Before you begin

The authorization ID associated with the statement must have SYSADM or DBADM authority.

Restrictions

You cannot drop the DB2_FENCED wrapper option.

The federated server cannot process an ALTER WRAPPER statement within a given unit of work if the unit of work already includes one of the following statements:

- A SELECT statement that references a nickname for a table or view at the data source that the wrapper includes
- An open cursor on a nickname for a table or view at the data source that the wrapper includes
- An insert, delete, or update issued against a nickname for a table or view at the data source that the wrapper includes

About this task

This task describes how to alter a wrapper from the DB2 Control Center.

Procedure

To alter a wrapper from the DB2 Control Center:

1. Expand the Federated Database Objects folder. The wrapper objects are displayed in the contents pane of the DB2 Control Center window.
2. Right-click on the wrapper that you want to change and click **Alter** from the list of actions. The Alter Wrapper notebook opens.
   a. On the Settings page, make the changes.
   b. Click **Set Variables** to set the data source environment variables for the wrapper. Environment variables are not required for all wrappers.
3. Click **OK** to alter the wrapper and close the Alter Wrapper notebook.

Altering a wrapper - examples

This topic provides examples of modifying wrapper options with the ALTER WRAPPER statement.

To change the DB2_FENCED option to 'Y' for the wrapper named drda, issue the following statement:

```
ALTER WRAPPER drda OPTIONS (SET DB2_FENCED 'Y');
```

To change the MODULE option to '/opt/odbc/lib/libodbc.a(odbc.so)' for the wrapper named odbc , issue the following statement:

```
ALTER WRAPPER odbc OPTIONS (SET MODULE '/opt/odbc/lib/libodbc.a(odbc.so)');
```

Altering a wrapper (DB2 command line)

After you configure a wrapper, you can use the ALTER WRAPPER statement to modify the configuration based on your system requirements. You can alter a wrapper by using the DB2 Control Center or by using the DB2 command line.

Before you begin

The authorization ID associated with the statement must have SYSADM or DBADM authority.

Restrictions

You cannot drop the DB2_FENCED wrapper option.

The federated server cannot process an ALTER WRAPPER statement within a given unit of work if the unit of work already includes one of the following statements:

- A SELECT statement that references a nickname for a table or view at the data source that the wrapper includes
- An open cursor on a nickname for a table or view at the data source that the wrapper includes
- An insert, delete, or update issued against a nickname for a table or view at the data source that the wrapper includes

About this task

This task describes how to alter a wrapper from the DB2 command line. You can use the ALTER WRAPPER statement to add, set, or drop one or more wrapper options.

Procedure

To alter a wrapper from the DB2 command line, issue the ALTER WRAPPER statement.

Altering server definitions and server options

You use the ALTER SERVER statement to modify a server definition. Some of the information within a server definition is stored as server options. When you modify a server definition, it is important to understand the options that you can specify about the server.

A server definition identifies a data source to the federated database. The server definition consists of a local name and other information about that data source server. The server definition is used by the wrapper when SQL statements that use nicknames are submitted to the federated database.

For relational data sources, server options can be set temporarily using the SET SERVER OPTION statement. This statement overrides the server option value in the server definition for the duration of a single connection to the federated database.

Modify a server definition when:

- You upgrade to a new version of the data source
- You want to make the same change to all of the server definitions for a specific data source type
- You want to add or change a server option on an existing server definition

Restrictions on altering server definitions

You need to be aware of several restrictions when you alter server definitions.

The following restrictions apply to altering server definitions:

- You cannot specify a wrapper in the ALTER SERVER statement that is not registered with the federated server.
- The federated server cannot process an ALTER SERVER statement within a given unit of work (UOW) under either of the following conditions:
  - The statement references a single data source, and the UOW already includes one of the following statements:
    - A SELECT statement that references a nickname for a table or view within the data source
    - An open cursor on a nickname for a table or view within the data source
    - An insert, delete or update issued against a nickname for a table or view within the data source
  - The statement references a category of data sources (for example, all data sources of a specific type and version), and the UOW already includes one of the following statements:
    - A SELECT statement that references a nickname for a table or view within one of the data sources
    - An open cursor on a nickname for a table or view within one of the data sources
    - An insert, delete or update issued against a nickname for a table or view within one of the data sources
- The federated server does not verify whether the server version that you specify matches the remote server version. Specifying an incorrect server version can result in SQL errors when you access nicknames that belong to the DB2 server definition. These SQL errors are most likely to occur when you specify a server version that is later than the actual version of the remote server. In that case, when you access nicknames that belong to the server definition, the federated server might send SQL to the remote server that the remote server does not recognize. On the ALTER SERVER statement, this situation applies only to the form of the statement that alters the server version (*server-name* VERSION server-version).
- You must specify the full server option name. For example, you cannot specify the abbreviation DB2_TWO_PHASE. Instead, you need to specify the full server option name DB2_TWO_PHASE_COMMIT.

Altering the data source version in a server definition (DB2 Control Center)

You can alter an existing server definition to change the version of the data source that the remote server uses. You can alter a server definition from the DB2 Control Center or from the DB2 command line.

Before you begin

The authorization ID issuing the ALTER SERVER statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

See "Restrictions on altering server definitions" on page 23.

Procedure

To alter a server definition from the DB2 Control Center:

1. Expand the Federated Database Objects folder. The server definition objects are displayed in the contents pane of the DB2 Control Center window.
2. Right-click on the server definition that you want to change and click **Alter** from the list of actions. The Alter Server Definition notebook opens.
3. On the Server page, click the **Version** arrow to specify a different version for the data source.
4. Click **OK** to alter the server definition and close the Alter Server Definition notebook.

Altering the data source version in a server definition (DB2 command line)

You can alter an existing server definition to change the version of the data source that the remote server uses. You can alter a server definition from the DB2 Control Center or from the DB2 command line.

Before you begin

The authorization ID issuing the ALTER SERVER statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

See "Restrictions on altering server definitions" on page 23.

Procedure

To alter a server definition from the DB2 command line, issue the ALTER SERVER statement.

Example: You are working with a server definition for a Microsoft SQL Server Version 6.5 data source. The name that you assigned the server in the CREATE SERVER statement is SQLSVR_ASIA. If the Microsoft SQL Server is upgraded to Version 7.0, the following statement alters the server definition:

```
ALTER SERVER SQLSVR_ASIA VERSION 7
```

Altering all of the server definitions for a specific data source type

You can alter all of the existing server definitions for a particular type of data source in a single ALTER SERVER statement. Use a single statement when you want the same change to apply to all of the server definitions of the same type.

Before you begin

The authorization ID issuing the ALTER SERVER statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

You can only set or drop server options using the ALTER SERVER statement for an entire type of data sources if the server options were added by a prior ALTER SERVER statement operation.

Procedure

To alter all of the existing server definitions for a particular type of data source, issue a single ALTER SERVER statement.

Example: Five Sybase servers are registered in the global catalog for your Sybase data sources. Whenever a user ID is sent to any of these Sybase servers for authentication, you want the federated server to always fold the user ID to uppercase. In addition, you want to set how long the federated server will wait for a response from these Sybase servers to an SQL statement. You specify the amount of time in seconds. You can modify all five server definitions at the same time the following ALTER SERVER statement:

```
ALTER SERVER TYPE sybase
   OPTIONS (ADD FOLD_ID 'U', ADD TIMEOUT '600')
```

Using server options in server definitions (DB2 Control Center)

There are general server options and server options that are for specific data source types. You can alter a server definition from the DB2 Control Center or from the command line prompt to add or change a server option.

Before you begin

The authorization ID issuing the ALTER SERVER statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

See “Restrictions on altering server definitions” on page 23.

About this task

Server options are set to values that persist over successive connections to the data source. These values are stored in the federated system catalog.

Procedure

To do this task from the DB2 Control Center:

1. Expand the Federated Database Objects folder. The server definition objects are displayed in the contents pane of the DB2 Control Center window.
2. Right-click on the server definition that you want to change and click **Alter** from the list of actions. The Alter Server Definition notebook opens.
3. On the Settings page, select the server option that you want to add or remove.
4. For options that you are adding or changing, specify the value of an option.
5. Click **OK** to alter the server definition and close the Alter Server Definition notebook.

   Some server options are required and cannot be dropped. Other server options cannot be added if specific server options are already set.

Changing server options temporarily for relational data sources

The SET SERVER OPTION statement overrides the server option value in the server definition for the duration of a single connection to the federated database. The overriding value does not get stored in the global catalog.

Procedure

To set a server option value temporarily for a relational data source, use the SET SERVER OPTION statement.

For example:

```
SET SERVER OPTION PLAN_HINTS TO Y' FOR SERVER ORA_SERVER
```

When used with static SQL statements, the SET SERVER OPTION statement will have no effect with the IUD_APP_SVPT_ ENFORCE server option.

The hierarchy of server option settings

When you have the same server option set with one value for a data source type and set with another value on a specific data source server, there is a hierarchy among the settings.

For example, the PLAN_HINTS server option is set to 'Y' for the data source type ORACLE. However, the PLAN_HINTS server option is set to 'N' in the server definition for a specific Oracle data source server PURNELL. The setting for the specific data source server overrides the setting for the data source type. This configuration causes PLAN_HINTS to be enabled at all Oracle data source servers except PURNELL.

Using server options in server definitions (DB2 command line)

There are general server options and server options that are for specific data source types. You can alter a server definition from the DB2 Control Center or from the command line prompt to add or change a server option.

Before you begin

The authorization ID issuing the ALTER SERVER statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

See “Restrictions on altering server definitions” on page 23.

About this task

Server options are set to values that persist over successive connections to the data source. These values are stored in the federated system catalog.

Procedure

To do this task from the command line prompt, issue the ALTER SERVER statement. For example:

- You created a server definition for an Informix server using the server name of INFMX01. You now want to change the DB2_MAXIMAL_PUSHDOWN option to Y. The statement to alter the server definition is:

  ```
  ALTER SERVER INFMX01 OPTIONS (SET DB2_MAXIMAL_PUSHDOWN 'Y')
  ```

- You created a server definition for an Oracle server using the server name of ORCL99. You now want to add the FOLD_ID and FOLD_PW options to the definition. The statement to alter the server definition is:

  ```
  ALTER SERVER ORCL99 OPTIONS (ADD FOLD_ID 'U', FOLD_PW 'U')
  ```

- You want to set the timeout value to the number of seconds the CTLIB wrapper should wait for a response from the Sybase server. You use the TIMEOUT server option to set this value. The statement to alter the server definition is:

  ```
  ALTER SERVER SYBSERVER OPTIONS (ADD TIMEOUT '60')
  ```

Altering a user mapping (DB2 Control Center)

A user mapping is the association between the authorization ID at the federated server and the authorization ID at the data source. User mappings are needed so that distributed requests can be sent to the data source.

Before you begin

If the authorization ID issuing the statement is different than the authorization ID that is mapped to the data source, then the authorization ID issuing the statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

The federated server cannot process an ALTER USER MAPPING statement within a given unit of work (UOW) if the UOW already includes one of the following statements:

- A SELECT statement that references a nickname for a table or view at the data source the mapping includes
- An open cursor on a nickname for a table or view at the data source that the mapping includes
- An insert, delete, or update issued for a nickname of a table or view at the data source the mapping includes

About this task

The ALTER USER MAPPING statement is used to change the authorization ID or password that is used at the data source for a specific federated server authorization ID.

You can alter a user mapping from the DB2 Control Center or the command line prompt.

Procedure

To alter a user mapping from the DB2 Control Center:

1. Expand the Federated Database Objects folder. The user mapping objects are displayed in the contents pane of the DB2 Control Center window.
2. Right-click on the user mapping that you want to change and click **Alter** from the list of actions. The Alter User Mapping window opens.
3. Change the value of the option.
4. Click **OK** to alter the user mapping and close the Alter User Mapping window.

Altering a user mapping (DB2 command line)

A user mapping is the association between the authorization ID at the federated server and the authorization ID at the data source. User mappings are needed so that distributed requests can be sent to the data source.

Before you begin

If the authorization ID issuing the statement is different than the authorization ID that is mapped to the data source, then the authorization ID issuing the statement must include either SYSADM or DBADM authority on the federated database.

Restrictions

The federated server cannot process an ALTER USER MAPPING statement within a given unit of work (UOW) if the UOW already includes one of the following statements:

- A SELECT statement that references a nickname for a table or view at the data source the mapping includes
- An open cursor on a nickname for a table or view at the data source that the mapping includes
- An insert, delete, or update issued for a nickname of a table or view at the data source the mapping includes

About this task

The ALTER USER MAPPING statement is used to change the authorization ID or password that is used at the data source for a specific federated server authorization ID.

You can alter a user mapping from the DB2 Control Center or the command line prompt.

Procedure

To alter a user mapping from the DB2 command line, issue the ALTER USER MAPPING statement:

For example, Jenny uses the federated server to connect to a Sybase server called SYBSERVER. She accesses the federated server with the authorization ID of *jennifer*. The authorization ID *jennifer* is mapped to the authorization ID *jenn* on the Sybase

server. The authorization ID for Jenny on the Sybase server is changed to *jen123*. The ALTER USER MAPPING statement to map *jennifer* to *jen123* is:

```
ALTER USER MAPPING FOR jennifer SERVER SYBSERVER
   OPTIONS (SET REMOTE_AUTHID 'jen123')
```

Tomas uses the federated server to connect to an Oracle server called ORASERVER. He accesses the federated server with the authorization ID of *tomas*. The authorization ID *tomas* is mapped to the authorization ID *tom* on the Oracle server. The password for Tomas on the Oracle server is changed. His new password is day2night. The ALTER USER MAPPING statement to map *tomas* to the new password is:

```
ALTER USER MAPPING FOR tomas SERVER ORASERVER
   OPTIONS (SET REMOTE_PASSWORD 'day2night')
```

The REMOTE_AUTHID and REMOTE_PASSWORD user options are case sensitive unless you set the FOLD_ID and FOLD_PW server options to 'U' or 'L' in the CREATE SERVER statement.

Altering a nickname (DB2 Control Center)

Nicknames are identifiers that are used to reference an object that you want to access at a data source. You can change the data source column names that are stored in the global catalog and set column options by altering the nicknames. You can alter a nickname from the DB2 Control Center or the DB2 command line.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

You might want to alter a nickname to:

- Alter the local column names for the columns of the data source object
- Alter the local data types for the columns of the data source object
- Add, set, or drop nickname and column options
- Add or drop a primary key
- Add or drop one or more unique, referential, or check constraints
- Alter one or more referential, check, or functional dependency constraint attributes
- Prevent nickname caching on the federated server

- Enable nickname caching on the federated server. If cache tables or materialized query tables are associated with a cached nickname, you must drop those tables before you alter the caching option.

Procedure

To alter a nickname from the DB2 Control Center:

1. Select the **Nicknames** folder.
2. Right-click on the nickname that you want to change and click **Alter**. The Alter Nickname notebook opens.
3. On the Nicknames page change the applicable options.
4. On the Keys page, set the referential integrity constraints for the nickname. You can set a primary key, unique key, or foreign key constraint.
5. On the Check Constraints page, set the check constraints or functional dependency constraints for the nickname.
6. On the Settings page, set the nickname options for the nickname.
7. Click **OK** to alter the nickname and close the notebook.

   Some nickname options are required and cannot be dropped. Other nickname options cannot be added if specific nickname options are already set.

When the data source object structure or content changes significantly, you should update the nickname statistics. Significant changes include the addition or removal of multiple rows.

Restrictions on altering nicknames

You need to be aware of several restrictions when you alter a nickname.

Column names

The ALTER NICKNAME statement cannot be used to alter column names for the following data sources. You must drop the nickname and create the nickname again with the correct column names.

- BLAST
- HMMER

Column options

If one of the following options is set on a column, you cannot add any other options to that column:

- SOAPACTIONCOLUMN
- URLCOLUMN
- PRIMARY_KEY
- FOREIGN_KEY

For BioRS

- If you change the element name of a column by using the ELEMENT_NAME option, the new name is not checked to ensure that it is correct. An incorrect option might result in errors when the column is referenced in a query.
- If you make changes to the IS_INDEXED column option, the changes are not verified with the BioRS server. An incorrect option might result in errors when the column is referenced in a query.

Data types

- If you change the data type of a column, the new data type must be compatible with the data type of the corresponding data source column or element. Changing the local data type to a data type that is incompatible with the remote data type might cause unpredictable errors.
- The *local_data_type* cannot be LONG VARCHAR, LONG VARGRAPHIC, XML, or a user-defined data type.
- The *data_source_data_type* cannot be a user-defined type.
- You cannot override the existing local types or create new local types for some of the nonrelational data sources. Check the documentation for the specific data source wrapper for more information on this restriction.
- When the local specification of a column's data type is changed, the federated database manager invalidates any statistics (for example, HIGH2KEY and LOW2KEY) that are gathered for that column.
- The local type is set for the specific data source object when it is accessed with that nickname. The same data source object can have another nickname that uses the default data type mapping.

Indexes

The ALTER NICKNAME statement cannot be used to register a new data source index in the federated database. Use the CREATE INDEX statement with the SPECIFICATION ONLY clause to create an index specification.

LOCAL NAME and LOCAL TYPE parameters

- ALTER NICKNAME statement cannot be used to change the local names or data types for the columns in the nickname if:
  - The nickname is used in a view, SQL method, or SQL function
  - You define an informational constraint on the nickname
- The federated_column_options clause must be specified last if you also need to specify the LOCAL NAME parameter, the LOCAL TYPE parameter, or both in the ALTER NICKNAME statement.

Nicknames

The ALTER NICKNAME statement cannot be used to change the name of the BioRS databank that is referenced by or used in a BioRS nickname. If the name of a BioRS databank changes, you must drop the nickname and create the nickname again.

You cannot use the ALTER NICKNAME statement to disallow caching on a nickname with cache tables or materialized query tables. You must drop the cache tables and the materialized query tables before you disallow caching on the nickname.

Units of work

The federated server cannot process an ALTER NICKNAME statement within a given unit of work under any of the following conditions:

- If the nickname referenced in the ALTER NICKNAME statement has a cursor open on it in the same unit of work.
- If an insert, delete or update is issued in the same unit of work for the nickname that is referenced in the ALTER NICKNAME statement.

Altering nickname column names (DB2 Control Center)

You can alter a nickname to change the column names. You can change column names from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

When you create a nickname, the column names that are associated with the data source object are stored in the federated database. For some data sources, the wrapper specifies the column names. For other data sources, you must specify the column names when you create the nickname.

Procedure

To alter nickname column names from the DB2 Control Center:

1. Select the **Nicknames** folder.
2. Right-click on the nickname that you want to change and click **Alter**. The Alter Nickname notebook opens.
3. On the Nicknames page, select the column that you want to change and click **Change**. The Change Column window opens.
4. Type the column name.
5. Click **OK** to change the column name and close the window.
6. Click **OK** to alter the nickname and close the notebook.

Altering nickname column names (DB2 command line)

You can alter a nickname to change the column names. You can change column names from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See “Restrictions on altering nicknames” on page 30.

About this task

When you create a nickname, the column names that are associated with the data source object are stored in the federated database. For some data sources, the wrapper specifies the column names. For other data sources, you must specify the column names when you create the nickname.

Procedure

To alter nickname column names from the DB2 command line, issue the ALTER NICKNAME statement.

```
ALTER NICKNAME nickname
   ALTER COLUMN current_name
   LOCAL NAME new_name
```

Altering nickname options (DB2 Control Center)

Nickname options are parameters that you specify on the nickname when you issue the CREATE NICKNAME and ALTER NICKNAME statements. You can add, set, or drop nickname options by using the ALTER NICKNAME statement. You can change column names from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See “Restrictions on altering nicknames” on page 30.

About this task

For example, the nickname DRUGDATA1 is created for the table-structured file `drugdata1.txt`. The fully qualified path that was originally defined in the CREATE NICKNAME statement was `/user/pat/drugdata1.txt`.

To change the FILE_PATH nickname option, issue the following statement :

Procedure

To change column names from the DB2 Control Center:

1. Select the **Nicknames** folder.
2. Right-click on the nickname that you want to change and click **Alter**. The Alter Nickname notebook opens.
3. On the Settings page, select the check box next to any option that you want to add or remove. You cannot remove a required option.

4. To specify or change the value of an option, click the **Value** field for the option. Depending on the option, you can either select a value from the list, click to select multiple values, or you can type a new value.
5. Click **OK** to alter the nickname and close the notebook.

Altering nickname options (DB2 command line)

Nickname options are parameters that you specify on the nickname when you issue the CREATE NICKNAME and ALTER NICKNAME statements. You can add, set, or drop nickname options by using the ALTER NICKNAME statement. You can change nickname options from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

Procedure

To change a nickname option from the command line prompt, issue the ALTER NICKNAME statement.

```
ALTER NICKNAME nickname
    OPTIONS (SET option_name 'option_string_value')
```

Example: The nickname DRUGDATA1 is created for the table-structured file `drugdata1.txt`. The fully qualified path that was originally defined in the CREATE NICKNAME statement was `/user/pat/drugdata1.txt`. To change the FILE_PATH nickname option, issue the following statement:

```
ALTER NICKNAME DRUGDATA1 OPTIONS (SET FILE_PATH '/usr/kelly/data/drugdata1.txt')
```

Altering nickname column options (DB2 Control Center)

You can add, set, or drop nickname column options using the ALTER NICKNAME statement. You can change column names from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement

- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

You specify column information in the CREATE NICKNAME and ALTER NICKNAME statements by using parameters called nickname column options.

Procedure

To alter nickname column options from the DB2 Control Center

1. Select the **Nicknames** folder.
2. Right-click on the nickname that you want to change and click **Alter**. The Alter Nickname notebook opens.
3. On the Nicknames page, select the column that you want to change and click **Change**. The Change Column window opens.
4. Select the column option that you want to add or remove.
5. For options that you are adding or changing, specify the value of an option.
6. Click **OK** to change the column option and close the window.
7. Click **OK** to alter the nickname and close the notebook.

Altering nickname column options (DB2 command line)

You can add, set, or drop nickname column options using the ALTER NICKNAME statement. You can change column names from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

You specify column information in the CREATE NICKNAME and ALTER NICKNAME statements by using parameters called nickname column options. You can specify any of these values in either uppercase or lowercase letters.

Procedure

To alter nickname column options from the command line prompt, use the ALTER NICKNAME statement.

Example 1: Specifying the NUMERIC_STRING column option with relational data sources

The NUMERIC_STRING column option applies to character type columns (CHAR and VARCHAR). Suppose that a data source has a collating sequence that differs from the federated database collating sequence. The federated server typically would not sort any columns containing character data at the data source. It would return the data to the federated database and perform the sort locally. However, suppose that the column is a character data type and contains only numeric characters ('0','1',...,'9'). You can indicate this by assigning a value of 'Y' to the NUMERIC_STRING column option. This gives the DB2 UDB query optimizer the option of performing the sort at the data source. If the sort is performed remotely, you can avoid the overhead of sorting the data at the federated server.

The nickname ORA_INDSALES for an Oracle table called INDONESIA_SALES. The table contains the column POSTAL_CODE with the data type of VARCHAR. Originally the column contained only numeric characters, and the NUMERIC_STRING column option was set to 'Y'. However, the column now contains a mixture of numeric and non-numeric characters. To change the NUMERIC_STRING column option to 'N', use this statement:

```
ALTER NICKNAME ORA_INDSALES ALTER COLUMN POSTAL_CODE
   OPTIONS (SET NUMERIC_STRING 'N')
```

Example 2: Specifying the VARCHAR_NO_TRAILING_BLANKS column option with relational data sources

The VARCHAR_NO_TRAILING_BLANKS column option can be used to identity specific columns that contain no trailing blanks. The SQL Compiler will factor in this setting when it checks for all operations (such as comparison operations) performed on columns.

The nickname ORA_INDSALES is for an Oracle table called INDONESIA_SALES. The table contains the column NAME with the data type of VARCHAR. The NAME column does not have trailing blanks. To add the VARCHAR_NO_TRAILING_BLANKS option to the nickname, use this statement:

```
ALTER NICKNAME ORA_INDSALES ALTER COLUMN NAME
   OPTIONS (ADD VARCHAR_NO_TRAILING_BLANKS 'Y')
```

Example 3: Specifying the XPATH column option with nonrelational data sources

The nickname EMPLOYEE is for an XML data source. An XPATH was specified for the *fnmae* column. To set the XPATH column option to a different path, use this statement:

```
ALTER NICKNAME EMPLOYEE ALTER COLUMN fname
   OPTIONS (SET XPATH './@first')
```

Altering a nickname (DB2 command line)

Nicknames are identifiers that are used to reference an object that you want to access at a data source. You can change the data source column names that are stored in the global catalog and set column options by altering the nicknames. You can alter a nickname from the DB2 Control Center or the DB2 command line.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

You might want to alter a nickname to:

- Alter the local column names for the columns of the data source object
- Alter the local data types for the columns of the data source object
- Add, set, or drop nickname and column options
- Add or drop a primary key
- Add or drop one or more unique, referential, or check constraints
- Alter one or more referential, check, or functional dependency constraint attributes

Procedure

To alter a nickname from the DB2 command line, issue the ALTER NICKNAME statement with the appropriate parameters set.

When the data source object structure or content changes significantly, you should update the nickname statistics. Significant changes include the addition or removal of multiple rows.

Dropping a wrapper

There are several reasons why you might want to drop a wrapper.

Before you begin

To issue the DROP WRAPPER statement, you must have SYSADM or DBADM authority.

About this task

Sometimes there is more than one wrapper that you can use to access a data source. The one you choose might depend on the version of the data source client software that you are using. Or it might depend on the operating system that you are using on your federated server. Suppose that you want to access two Oracle tables and one Oracle view. You are using Oracle Version 8, and the operating system on your federated server is Windows NT®. Originally you created the SQLNET wrapper. Since WebSphere Federation Server does not support the SQLNET wrapper, you can drop the SQLNET wrapper and create the NET8 wrapper.

Another reason to drop a wrapper is that you no longer need access to the data source that the wrapper is associated with. For example, suppose that your organization has a requirement to access client information in both Informix and Microsoft SQL server databases. You create one wrapper for the Informix data source and one wrapper for the Microsoft SQL Server data source. Later your organization decides to migrate all of the information from Microsoft SQL Server to Informix. You no longer need the Microsoft SQL Server wrapper and you can drop it.

Important: There are serious consequences when you drop a wrapper. Other objects that you registered with the federated server are impacted:

- All server definitions that are dependent on the dropped wrapper are also dropped.
- All objects that are dependent on the dropped server definitions are also dropped.
- All nicknames that are dependent on the dropped server definitions are also dropped. Dropping the nicknames dependent on the server definition affects the objects dependent on those nicknames:
  - Any index specifications dependent on the dropped nicknames are also dropped.
  - Any views dependent on the dropped nicknames are marked inoperative.
  - Any materialized query tables dependent on the dropped nicknames are also dropped.
- All packages and cached dynamic SQL statements dependent on the dropped nicknames are marked invalid, and remain invalid until the dependent objects are re-created.

Procedure

To drop a wrapper, use the DROP statement.

Example: Drop the Microsoft SQL Server *MSSQLODBC3* wrapper:

```
DROP WRAPPER MSSQLODBC3
```

Dropping a server definition

Dropping a server definition deletes the definition from the global catalog. The data source object that the server definition references is not affected. You can drop a server definition using the DB2 Control Center or using DROP statement from the DB2 command line processor.

Before you begin

You must have SYSADM or DBADM authority to drop a server definition.

Restrictions

The federated server cannot process a DROP SERVER statement within a given unit of work (UOW) under either of the following conditions:

- The statement references a single data source, and the UOW already includes one of the following statements:
  - A SELECT statement that references a nickname for a table or view within the data source
  - An open cursor on a nickname for a table or view within the data source
  - An insert, delete or update issued against a nickname for a table or view within the data source
- The statement references a category of data sources (for example, all data sources of a specific type and version), and the UOW includes one of the following statements:
  - A SELECT statement that references a nickname for a table or view within one of the data sources
  - An open cursor on a nickname for a table or view within one of the data sources
  - An insert, delete or update issued against a nickname for a table or view within one of the data sources

About this task

When you no longer need to access a data source server, drop the server definition from the federated database. When you drop a server definition, other objects that you registered with the federated server are impacted:

- All user-defined function mappings, user-defined data type mappings, and user mappings that are dependent on the dropped server definition are also dropped.
- All nicknames that are dependent on the dropped server definition are also dropped. Dropping the nicknames dependent on the server definition, affects the objects dependent on those nicknames:
  - Any index specifications dependent on the dropped nicknames are also dropped.
  - Any views dependent on the dropped nicknames are marked inoperative.
  - Any materialized query tables dependent on the dropped nicknames are also dropped.
- All packages and cached dynamic SQL statements dependent on the dropped nicknames are marked invalid, and remain invalid until the dependent objects are re-created.

Procedure

To delete a server definition, issue the DROP statement:

```
DROP SERVER server_name
```

where *server_name* identifies the server definition to be dropped.

Example: An Informix server uses the server name INFMX01. The following DROP statement drops the server definition:

```
DROP SERVER INFMX01
```

Dropping a user mapping

When a user no longer needs access to a remote data source, drop the user mapping between the federated server and the remote data source server. If the user is mapped to more than one data source server, you will need to drop each mapping separately.

Before you begin

To issue the DROP USER MAPPING statement, the authorization ID of the DROP statement must have SYSADM or DBADM authority, if this authorization ID is different from the federated database user ID specified in the user mapping. Otherwise, if the authorization ID and the user ID in the user mapping match, no authorities or privileges are required.

Procedure

To drop a user mapping, issue the DROP statement:

```
DROP USER MAPPING FOR user_ID SERVER local_server_name
```

where:

- *user_ID* is the authorization ID for the user on the federated server
- *local_server_name* is the local name that is used to identify the remote data source server in the server definition.

Dropping a nickname

Dropping a nickname deletes the nickname from the global catalog on the federated server. The data source object that the nickname references is not affected.

Before you begin

The nickname must be listed in the catalog.

The privileges that must be held by the authorization ID of the DROP statement when dropping nicknames must include one of the following:

- SYSADM or DBADM authority
- DROPIN privilege on the schema for the nickname
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname
- CONTROL privilege on the nickname

Restrictions

For nicknames to relational data sources, the federated server cannot process the DROP NICKNAME statement within a given unit of work (UOW) under either of the following conditions:

- A nickname referenced in the statement has a cursor open on it in the same UOW.
- A nickname referenced in this statement is already referenced by a SELECT statement in the same UOW.

- An insert, delete or update is issued in the same UOW for the nickname referenced in the statement.

For nicknames to nonrelational data sources, the federated server cannot process the DROP NICKNAME statement within a given unit of work (UOW) under any of the following conditions:

- A nickname referenced in this statement has a cursor open on it in the same UOW.
- A nickname referenced in this statement is already referenced by a SELECT statement in the same UOW.

About this task

When you drop a nickname, other objects that you registered with the federated server are impacted:

- Dropping a nickname affects the objects dependent on those nicknames:
  - Any index specifications dependent on the dropped nicknames are also dropped.
  - Any views dependent on the dropped nicknames are marked inoperative.
  - Any materialized query tables dependent on the dropped nicknames are also dropped.
- All packages and cached dynamic SQL statements dependent on the dropped nickname are marked invalid, and remain invalid until the dependent objects are re-created.

Procedure

To delete a nickname, issue the DROP statement:

```
DROP NICKNAME nickname
```

where *nickname* identifies the nickname to be dropped.

Chapter 3. Data type mappings

Data type mappings in a federated system

The data types at a data source must map to corresponding DB2 data types. This mapping enables the federated server to retrieve data from the data source.

The federated database supplies a set of default data type mappings for some data sources. For other data sources you must provide the data type mappings that you want to use. For nonrelational data sources, you cannot override the existing data type mappings or create new mappings.

Some examples of default data type mappings are:

- The Oracle type FLOAT maps by default to the DB2 type DOUBLE
- The Oracle type DATE maps by default to the DB2 type TIMESTAMP
- The DB2 Universal Database for z/OS type DATE maps by default to the DB2 type DATE

Nicknames that are created after a mapping is changed use the new type mapping. Nicknames that are created before the mapping is changed use the default data type mapping.

If you already created the nicknames, you can update the existing nicknames in one of the following ways:

- You can alter each nickname
- You can drop and re-create each nickname

DB2 federated servers do not support mappings for the following data types:

- The local data type cannot be LONG VARCHAR, LONG VARGRAPHIC, DATALINK, or a user-defined data type.
- The remote data type cannot be a user-defined type.

  However, you can use a cast function to convert the user-defined data type in a view at the remote data source to a built-in or system data type. You can then create a nickname for the view. For most data sources, if you create such views, the views have no statistics or indexes and you cannot update the views.

Data type mappings and the federated database global catalog

Local data type definitions are stored in the SYSCAT.COLUMNS catalog view of the federated database global catalog.

When you write a CREATE NICKNAME statement, you specify a data source object that the nickname represents. In most cases, the federated server defines a DB2-supported data type for each column or field in that data source object. For some nonrelational data sources, you must supply the DB2 data type.

For relational data sources, to determine which local data type to store in the SYSCAT.COLUMNS catalog view, the federated server looks for forward data type mapping information in the wrappers and in the SYSCAT.TYPEMAPPINGS catalog view. Mappings in the SYSCAT.TYPEMAPPINGS catalog view take precedence over the default mappings in the wrappers. If you create alternative mappings to

override the default data type mappings, the federated server uses the alternative mappings. If multiple mappings apply to a column, the federated server uses the most recently created mappings.

For nonrelational data sources, to determine which local data type to store in the SYSCAT.COLUMNS catalog view, the federated server looks for data type mapping information in the wrappers. Depending on the nonrelational data source, the degree to which you can modify the data types defined by the wrapper varies. For some nonrelational data sources, you do not specify any columns. The wrapper defines the data types. For other data sources you can override the data types. And for other data sources, you must specify the column data types on the CREATE NICKNAME statement.

When you write CREATE TABLE transparent DDL for relational data sources, specify DB2 data types in the statement. The federated server checks for information about the reverse data type mappings between the federated database and the data source. The federated server looks for this information in the wrapper and the SYSCAT.TYPEMAPPINGS catalog view.

When values from a data source column are returned to the federated database, the values conform fully to the DB2 data type that the data source column is mapped to. If this mapping is a default mapping, the values also conform fully to the data source type in the mapping. For example, if an Oracle table with a FLOAT column is defined to the federated database, the default mapping of Oracle FLOAT to DB2 DOUBLE automatically applies to that column. The values that are returned from the column conform fully to both the FLOAT and DOUBLE data types.

When to create alternative data type mappings

You can create alternative data type mappings for relational data sources.

You might want to create alternative data type mappings in the following situations:

- To override a default data type mapping

  For some wrappers, you can change the format or length of values that are returned. You can change the format or length by changing the DB2 data type that the values must conform to. For example, the Oracle DATE data type is used as a timestamp and contains the century, year, month, day, hour, minute, and second. By default, the Oracle DATE data type maps to the DB2 TIMESTAMP data type. To return only the hour, minute, and second information, you can override the default data type mapping so that the Oracle DATE data type maps to the DB2 TIME data type. When the Oracle DATE columns are queried, only the time portion of the Oracle timestamp values is returned to the federated server.
- When a default mapping does not exist

  If a default data type mapping is not available for a data source data type, you must create a mapping for the new data type.

You use the CREATE TYPE MAPPING statement to define new data type mappings. Mappings that you create are stored in the SYSCAT.TYPEMAPPINGS catalog view in the federated database.

Data type mappings for nonrelational data sources

For some nonrelational data sources, the data type mappings are not in the wrappers. In some cases, you must specify the local type information in the CREATE NICKNAME statement.

The following example shows how column data types are specified in the CREATE NICKNAME statement for some of the nonrelational data sources:

```
CREATE NICKNAME DRUGDATA1
   (Dcode Integer NOT NULL, Drug CHAR(20), Manufacturer CHAR(20))
   FOR SERVER biochem_lab
   OPTIONS (FILE_PATH '/usr/pat/DRUGDATA1.TXT', COLUMN_DELIMITER ',',
   SORTED 'Y', KEY_COLUMN 'DCODE', VALIDATE_DATA_FILE 'Y')
```

Forward and reverse data type mappings

Forward type mappings and reverse type mappings are the two kinds of mappings between data source data types and federated database data types.

A *forward type mapping* is a mapping from a remote data type to a comparable local data type. Forward type mappings are used when you a create a nickname for a data source object. The comparable local type for each column in the data source object is stored in the global catalog.

A *reverse type mapping* is a mapping from a local data type to a comparable remote data type. Reverse type mapping is used with transparent DDL.

Figure 2 shows forward and reverse data type mapping.

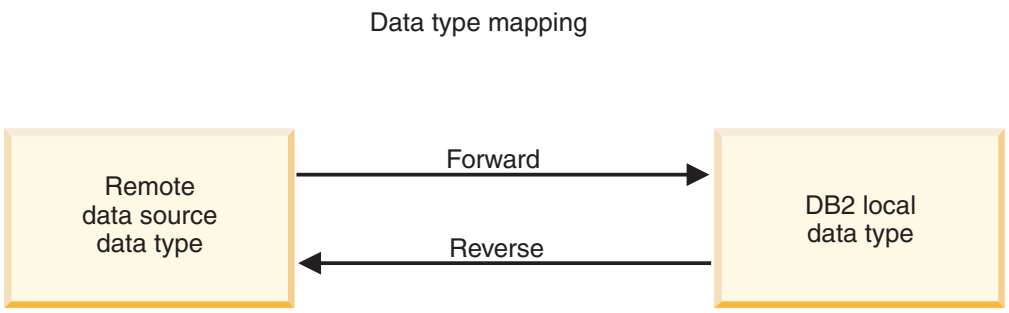


Figure 2. Forward and reverse data type mappings

Creating data type mappings

To create a data type mapping you use the CREATE TYPE MAPPING statement. You can issue the statement from the DB2 Command Center or the command line processor, or include it in an application program. You cannot use the DB2 Control Center to create or modify data type mappings.

Before you begin

The privileges held by the authorization ID associated with the statement must have SYSADM or DBADM authority.

Restrictions

- The *local_data_type* value cannot be LONG VARCHAR, LONG VARGRAPHIC, DATALINK, or a user-defined data type.

- The *data_source_data_type* value cannot be a user-defined type.
- For nonrelational data sources, the degree to which you can override existing data type mappings or create mappings is limited.

Procedure

To create a data type mapping, issue the CREATE TYPE MAPPING statement.

Creating a data type mapping for a data source data type – example

In this example, all Oracle tables and views that use the Oracle NUMBER data type must map to the DB2 DECIMAL(8,2) data type. The Oracle NUMBER data type is mapped by default to the DB2 DOUBLE data type, a floating decimal data type.

Use the ALTER NICKNAME statement to change the local types of existing nicknames. You must modify each nickname separately to change the local data type to DECIMAL(8,2).

If the nicknames do not exist, create a data type mapping that specifies the data source type.

For example, to create the type mapping from Oracle NUMBER data type to the DB2 DECIMAL(8,2) data type, issue the following statement:

```
CREATE TYPE MAPPING MY_ORACLE_DEC FROM SYSIBM.DECIMAL(8,2)
    TO SERVER TYPE ORACLE TYPE NUMBER
```

MY_ORACLE_DEC
: The name that you give to the type mapping. The name cannot duplicate a data type mapping name that already exists in the catalog.

FROM *SYSIBM.DECIMAL(8,2)*
: The local DB2 schema and the local data type. If the length or precision and scale are not specified, then these values are determined from the source data type.

TO SERVER TYPE *ORACLE*
: The type of data source.

TYPE *NUMBER*
: The data source data type that you are mapping to the local data type. User-defined data types are not allowed.

The DB2 DECIMAL(8,2) data type is defined locally for the Oracle columns.

When you create nicknames on Oracle tables and views that contain NUMBER columns, the Oracle NUMBER data type maps to the DB2 DECIMAL(8,2) data type.

Creating a type mapping for a data source data type and version – example

In this example, Oracle tables and views exist on different versions of the Oracle server. For all tables and views on Oracle Version 8.0.3 servers, columns that use the Oracle NUMBER(23,3) data type must map to the DB2 DECIMAL(8,2) data type.

The Oracle NUMBER(23,3) data type is mapped by default to the DB2 DECIMAL(23,3) data type.

Use the ALTER NICKNAME statement to change the local types of existing nicknames. You must modify each nickname separately to change the local data type to DECIMAL(8,2).

If the nicknames do not exist, create a data type mapping that specifies the data source type.

For example, to map the Oracle NUMBER(23,3) data type to the DB2 DECIMAL(8,2) data type for Oracle servers using Version 8.0.3, issue the following statement:

```
CREATE TYPE MAPPING ORA_DEC FROM SYSIBM.DECIMAL(8,2)
    TO SERVER TYPE ORACLE VERSION 8.0.3 TYPE NUMBER(23,3)
```

ORA_DEC
: The name that you give to the type mapping. The name cannot duplicate a data type mapping name that already exists in the catalog.

FROM *SYSIBM.DECIMAL(8,2)*
: The local DB2 schema and the local data type. If the length or precision and scale are not specified, then these values are determined from the source data type.

TO SERVER TYPE *ORACLE*
: The type of data source.

VERSION *8.0.3*
: The version of data source server. You must specify the version. You can also specify the release and the modification of the release, as shown in this example.

TYPE *NUMBER(23,3)*
: The data source data type that you are mapping to the local data type. User-defined data types are not allowed.

The federated database defines the DB2 DECIMAL(8,2) data type locally for the Oracle columns on Version 8.0.3 servers.

When you create nicknames on Oracle tables and views that contain NUMBER columns, the Oracle NUMBER data type maps to the DB2 DECIMAL(8,2) data type.

Oracle tables and views on servers that do not use Version 8.0.3 use the default data type mapping.

Creating a type mapping for all data source objects on a server – example

In this example, the server is defined to the federated database as ORA2SERVER. Each table contains a column with an Oracle DATE data type.

The Oracle DATE data type contains the century, year, month, day, hour, minute, and second. The Oracle DATE data type is mapped by default to the local DB2 TIMESTAMP data type. However, when you query any object on this server, the result set must return only the time information (hour, minute, and second).

Use the ALTER NICKNAME statement to change the local types of existing nicknames. You must modify each nickname separately to change the local data type to TIME.

If the nicknames do not exist, create a data type mapping that specifies the data source type.

To map the Oracle DATE data type to the DB2 TIME data type for the ORA2SERVER, issue the following statement:

```
CREATE TYPE MAPPING ORA2_DATE FROM SYSIBM.TIME
   TO SERVER ORA2SERVER TYPE DATE
```

ORA2_DATE
The name that you give to the type mapping. The name cannot duplicate a data type mapping name that already exists in the catalog.

FROM *SYSIBM.TIME*
The local DB2 schema and the local data type. If the length or precision and scale are not specified, then these values are determined from the source data type.

TO SERVER *ORA2SERVER*
The local name of the data source server.

TYPE *DATE*
The data source data type that you are mapping to the local data type. User-defined data types are not allowed.

The federated database locally defines the DB2 TIME data type for the Oracle columns of data type DATE.

When you create nicknames on Oracle tables and views that contain DATE columns, the Oracle DATE data type maps to the DB2 DECIMAL(8,2) data type.

Data source objects on other Oracle servers are not affected by this data type mapping.

Altering a local type for a data source object (DB2 Control Center)

You use the ALTER NICKNAME statement instead of the CREATE TYPE MAPPING statement to change a local data type. You can change the data type from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname.

Restrictions

See “Restrictions on altering nicknames” on page 30.

About this task

When you create a nickname, the data types that are associated with the data source object are stored in the federated database. For some data sources, the wrapper specifies the data types for you. For other data sources, you must specify the data types when you create the nickname.

You can specify a local type for a column of a specific data source object.

Important: Changing the local data type can result in errors or loss of information if you change the local data type for a column to a type that differs greatly from its remote type.

Procedure

To change the local data type from the DB2 Control Center:

1. Select the **Nicknames** folder.
2. Right-click on the nickname that you want to change and click **Alter**. The Alter Nickname notebook opens.
3. On the Nicknames page, select the column that you want to change and click **Change**. The Change Column window opens.
4. Select the data type.
5. Click **OK** to change the data type and close the window.
6. Click **OK** to alter the nickname and close the notebook.

To treat the contents of a local column that has a character data type as bit (binary) data, use the FOR BIT DATA clause in the ALTER NICKNAME statement. When you use this clause to change the local data type of a column, code page conversions are not performed when data is exchanged with other systems. Comparisons are done in binary, irrespective of the remote database collating sequence.

Altering a local type for a data source object – examples

This topic provides examples that show how to change the data types for a data source object.

Example: A numeric data type mapping

In an Oracle table for employee information, the BONUS column is defined with a data type of NUMBER(32,3). The Oracle data type NUMBER(32,3) is mapped by default to the DB2 data type DOUBLE, a double-precision floating-point number data type. A query that includes the BONUS column might return values that look like this:

```
5.0000000000000E+002
1.0000000000000E+003
```

The scientific notation indicates the number of decimal places and the direction that the decimal point should be moved. In this example +002 signifies that the decimal point should be moved two places to the right, and +003 signifies that the decimal point should be moved three places to the right.

Queries that include the BONUS column can return values that look like dollar amounts. You change the local definition for the BONUS column in the table from the DOUBLE data type to DECIMAL data type. Use a precision and scale that

reflect the format of actual bonuses. For example, if the dollar portion of the bonuses would not exceed six figures, map NUMBER(32,3) to DECIMAL(8,2). Under the constraint of this new local type, queries that include the BONUS column return values like this:

```
 500.00
1000.00
```

The nickname for the Oracle table is ORASALES. To map the BONUS column in the ORASALES table to the DB2 DECIMAL (8,2) data type, issue the following ALTER NICKNAME statement:

```
ALTER NICKNAME ORASALES ALTER COLUMN BONUS
   LOCAL TYPE DECIMAL(8,2)
```

ORASALES
The nickname that you defined for the Oracle table.

ALTER COLUMN *BONUS*
The name of the column that is defined locally in the federated database SYSCAT.COLUMNS catalog view.

LOCAL TYPE *DECIMAL(8,2)*
Identifies the new local type for the column.

This mapping applies only to the BONUS column in the Oracle table that is identified by the nickname ORASALES. All other Oracle data source objects that include the BONUS column use the default data type mapping for the Oracle NUMBER data type.

Example: A date data type mapping

The nickname for an Oracle table named SALES is ORASALES. The SALES table contains a column that is the Oracle DATE data type. The default type mapping for the Oracle DATE data type is to the DB2 TIMESTAMP data type. However, you want to display only the date value when you retrieve data from this column. You can alter the nickname for the SALES table to change the local type to the DB2 DATE data type.

```
ALTER NICKNAME ORASALES ALTER COLUMN ORDER_DATE
   LOCAL TYPE DATE
```

Example: A data type mapping for a nonrelational data source

The nickname for a table-structured file named `drugdata1.txt` is DRUGDATA1. The `drugdata1.txt` file contains a column that lists pharmaceutical drug names. The column name is DRUG. The DRUG column was originally defined as a CHAR(20). The length of the column must be changed to CHAR(30). You can alter the nickname for the `drugdata1.txt` file to change the mapping to the correct length:

```
ALTER NICKNAME DRUGDATA1 ALTER COLUMN DRUG
   LOCAL TYPE CHAR(30)
```

Altering a local type for a data source object (DB2 command line)

You use the ALTER NICKNAME statement instead of the CREATE TYPE MAPPING statement to change a local data type. You can change the data type from the DB2 Control Center or the DB2 command line.

Before you begin

The authorization ID issuing the statement must include at least one of the following privileges:

- SYSADM or DBADM authority
- ALTER privilege on the nickname specified in the statement
- CONTROL privilege on the nickname specified in the statement
- ALTERIN privilege on the schema, if the schema name of the nickname exists
- Definer of the nickname as recorded in the DEFINER column of the catalog view for the nickname.

Restrictions

See "Restrictions on altering nicknames" on page 30.

About this task

When you create a nickname, the data types that are associated with the data source object are stored in the federated database. For some data sources, the wrapper specifies the data types for you. For other data sources, you must specify the data types when you create the nickname.

You can specify a local type for a column of a specific data source object.

Important: Changing the local data type can result in errors or loss of information if you change the local data type for a column to a type that differs greatly from its remote type.

Procedure

To change the local data type from command line prompt, use the ALTER NICKNAME statement. For example:

```
ALTER NICKNAME nickname ALTER COLUMN column_name
   LOCAL TYPE data_type
```

To treat the contents of a local column that has a character data type as bit (binary) data, use the FOR BIT DATA clause in the ALTER NICKNAME statement. When you use this clause to change the local data type of a column, code page conversions are not performed when data is exchanged with other systems. Comparisons are done in binary, irrespective of the remote database collating sequence.

Altering LONG data types to VARCHAR data types

To enable insert and update operations for LONG data types, you can alter the LONG data types to the VARCHAR data type.

Table 4 lists the long data type by data source that you can alter.

Table 4. LONG data types by data source that can be altered to the VARCHAR data type

Data source	Remote data type	Length	Local default data type	ALTER to VARCHAR
DRDA	LONG VARCHAR	1–32672	CLOB	VARCHAR

Table 4. LONG data types by data source that can be altered to the VARCHAR data type (continued)

Data source	Remote data type	Length	Local default data type	ALTER to VARCHAR
	LONG VARCHAR FOR BIT DATA	1–32672	BLOB	VARCHAR FOR BIT DATA
Informix	BYTE	1–32672	BLOB	VARCHAR FOR BIT DATA
	TEXT	1–32672	CLOB	VARCHAR
Microsoft SQL Server	IMAGE	1–32672 host vars; 1–8000 literal	BLOB	VARCHAR FOR BIT DATA
	TEXT	1–32672 host vars; 1–8000 literal	CLOB	VARCHAR
Oracle NET8	LONG	1–32672 host vars; 1–4000 literal	CLOB	VARCHAR
	LONG RAW	1–32672 host vars; 1–4000 literal	BLOB	VARCHAR FOR BIT DATA
Sybase CTLIB	IMAGE	1–32672	BLOB	VARCHAR FOR BIT DATA
	TEXT	1–32672	CLOB	VARCHAR
Teradata	BYTE	32673–64000	BLOB	VARCHAR FOR BIT DATA(32672)
	CHAR	32673–64000	CLOB	VARCHAR(32672)
	VARBYTE	32673–64000	BLOB	VARCHAR FOR BIT DATA(32672)
	VARCHAR	32673–64000	CLOB	VARCHAR(32672)

Chapter 4. Mapping functions and user-defined functions

Function mappings in a federated system

WebSphere Federation Server supplies default mappings between existing data source functions and DB2 counterpart functions.

For the federated server to use a data source function, a mapping is needed from a DB2 function or a function template to the data source function.

The default function mappings are in the wrapper modules.

For nonrelational data sources, you cannot override the existing function mappings or create new mappings.

When to create your own function mappings

When a default function mapping is not available for a data source function, you can create a function mapping.

One reason that a function mapping is not available is that the federated database does not have a function that corresponds to the data source function.

Another reason a mapping is not available is that the data source has a similar function to a DB2 function, but it does not return the same results. If the data source returns slightly different results or different results for certain sets of input data, the wrappers do not normally map to these functions. However, if you do not care about the differences in the result sets, then you can create a mapping between the functions. Creating a mapping might improve performance.

Use functions mappings when:

- A new built-in function becomes available at the data source
- A new user-defined function becomes available at the data source
- A DB2 counterpart function does not exist
- A counterpart function exists but returns slightly different results, which you do not care about

The settings for function mappings are stored in the SYSCAT.FUNCMAPPINGS catalog view.

When you create a function mapping, it is possible that the return values from a function evaluated at the data source will be different than the return values from a compatible function evaluated at the federated database. WebSphere Federation Server uses the function mapping, but it might result in an SQL syntax error or unexpected results.

Why function mappings are important

Function mappings are one of several important inputs to the pushdown analysis performed by the query optimizer.

In deciding on the best query access plan, the query optimizer factors the capabilities of the data source to perform a particular type of SQL function or

operation. If the function does not have a mapping, the function will not be sent to the data source for processing. Functions and other operations that can be pushed down to the data source improve performance.

If the data source has a similar function to a DB2 function, but it returns slightly different results, creating a function mapping might improve performance. For example, the Informix® STDEV (standard deviation) function produces different results than the DB2 STDDEV function for some sets of input data. For this reason the Informix wrapper does not have a default mapping between these two functions. If you do not care about the result set differences, you might improve the performance of queries that access Informix data sources and use the DB2 STDDEV function. By creating a function mapping between the Informix STDEV and the DB2 STDDEV function, you provide the query optimizer the choice of sending the processing of that function down to the data source.

How function mappings work in a federated system

When you submit queries to the federated server that contain one or more functions, the federated server checks for information about the mappings between the DB2® functions and the data source functions.

The federated server checks two places for mapping information:

- The wrapper. The data source wrapper contains the default function mappings.
- The SYSCAT.FUNCMAPPINGS catalog view. This view contains entries you create that override or augment the default function mappings that are in the wrapper. It also contains new mappings that you create when there is no default function mapping. When multiple mappings can be applied to a function, the most recently created one is applied.

Function mapping options specify information about the function and the potential cost of processing a function at the data source. Function mapping options provide information such as:

- Name of the remote data source function
- The estimated number of instructions processed the first and last time that the data source function is invoked.
- Estimated number of I/Os performed the first and last time that the data source function is invoked.
- Estimated number of instructions processed per invocation of the data source function.

When you create a function mapping, you are mapping from a DB2 function or function template to a counterpart function at the data source. When a DB2 counterpart function does not exist, or when you want to force the federated server to use the data source function, you can create a function template to act as the counterpart.

Requirements for mapping user-defined functions (UDFs)

Before you can invoke a data source user-defined function in a federated system, the federated database must associate the data source function with a function specification stored in the global catalog on the federated server.

There are two conditions under which the federated database can associate a function specification with a data source function:

- The federated database has a function whose signature corresponds to that of the signature of the data source function. A *signature* consists of a function name and function input parameters. Signatures *correspond* when both the following conditions are true:
  - They contain the same names and the same number of parameters
  - The data type of each parameter in one signature is the same as (or can be converted to) the data type of the corresponding parameter in the other signature.
- If the federated database does not have a function with the requisite signature, you can define a function template that contains this signature. You then map the function template to the data source function that you want to invoke.

The settings for function mappings are stored in the SYSCAT.FUNCMAPPINGS catalog view.

Function templates

The federated server recognizes a data source function when there is a mapping between the data source function and a DB2® counterpart function at the federated database.

You can create a function template to act as the DB2 counterpart function when no counterpart exists.

A *function template* is a DB2 function that you create for the purpose of forcing the federated server to invoke a data source function. However, unlike a regular function, a function template has no executable code. When the federated server receives queries that specify the function template, the federated server will invoke the data source function.

The function template is created with the CREATE FUNCTION statement using the AS TEMPLATE parameter.

After you create a function template, you must then create the function mapping between the template and the data source function. A function mapping is created using the CREATE FUNCTION MAPPING statement.

Creating function templates

The federated server recognizes a data source function when there is a mapping between the data source function and a counterpart function at the federated database. You can create a function template to act as the counterpart when no counterpart exists.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- IMPICIT_SCHEMA authority on the databaes, if the implicit or explicit schema name of the function does not exist
- CREATEIN privilege on the schema, if the schema name of the function exists

Restrictions

If the data source function has input parameters:

- The DB2 counterpart function must have the same number of input parameters that the data source function has.
- The data types of the input parameters for the DB2 counterpart function must be compatible with the corresponding data types of the input parameters for data source function. The data type cannot be LONG VARCHAR, LONG VARGRAPHIC, DATALINK, or a user-defined type.

If the data source function has no input parameters, the DB2 counterpart function cannot have any input parameters.

Procedure

To create a function template:

1. Use the CREATE FUNCTION statement with the AS TEMPLATE parameter. For example:

   ```
   CREATE FUNCTION BONUS ()
     RETURNS DECIMAL(8,2)
     AS TEMPLATE
     DETERMINISTIC
     NO EXTERNAL ACTION
   ```

   BONUS ()
   The name you give to the function template.

   RETURNS *DECIMAL(8,2)*
   The data type of the output.

   AS TEMPLATE
   Indicates that this is a function template, not a function.

   DETERMINISTIC
   Specifies that the function always returns the same results for a given set of argument values.

   NO EXTERNAL ACTION
   Specifies that the function has no external impact on objects that are not managed by the database manager.

   You must specify the DETERMINISTIC and NO EXTERNAL ACTION clauses according to whether the function itself is deterministic and whether it causes any external action. Otherwise, restrictions will be imposed on the SQL operations that are supported with this function template.

2. After you create a function template, you must then create the function mapping between the template and the data source function. A function mapping is created using the CREATE FUNCTION MAPPING statement. For example:

   ```
   CREATE FUNCTION MAPPING MY_INFORMIX_FUN FOR BONUS()
     SERVER TYPE INFORMIX OPTIONS (REMOTE_NAME 'BONUS()')
   ```

   MY_INFORMIX_FUN
   The name you give to the function mapping. The name cannot duplicate a function mapping name that is already described in the federated database global catalog. It must be unique.

   FOR *BONUS()*
   The local DB2 function template name. Include data type input parameters in parenthesis.

SERVER TYPE *INFORMIX*
Identifies the type of data source which contains the function that you want to map to.

OPTIONS *(REMOTE_NAME 'BONUS()')*
An option that identifies the name of the remote data source function that you are mapping to the local DB2 function template.

Providing function mapping overhead information to the query optimizer

The query optimizer uses overhead information to estimate executing costs.

When a query that contains a function is received by the DB2 SQL Complier, the query optimizer determines if the function can be pushed down to the data source. Suppose that pushdown analysis determines that either the data source or the federated database is able to process the function.

When you create a function mapping, you can provide the federated database with important information about the potential cost, or *overhead*, of executing a data source function at the data source. The query optimizer uses these overhead estimates to compare the estimated cost of executing the data source function with the estimated cost of executing the DB2 function.

This information helps the DB2 query optimizer determine the best strategy for executing the query. When a distributed request is processed, the optimizer evaluates multiple access strategies and estimates the overhead to invoke the DB2 function and the data source function. The strategy that is expected to cost the least amount of overhead is used.

You include estimated overhead statistics in the CREATE FUNCTION MAPPING statement. For example, the statement can specify the estimated number of instructions that would be required to invoke the data source function. It can specify the estimated number of I/Os that would be expended for each byte of the argument set that is passed to this function. These estimates are stored in the global catalog, and appear in the SYSCAT.FUNCMAPOPTIONS view. When a DB2 function is used in the mapping (instead of a data source function or a DB2 function template) the global catalog contains estimates of overhead that would be consumed when the DB2 function is invoked. You can see these estimates in the SYSCAT.ROUTINES view.

Function mapping options that specify function overhead - examples

This topic lists the mapping options that specify function overhead and provides examples.

To specify estimated statistics in the CREATE FUNCTION MAPPING statement, use the function mapping options that specify function overhead. The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

The following table lists the function mapping options that specify function overhead and the default values for these options.

Table 5. Function mapping options that specify function overhead

Option	Valid settings	Default setting
INITIAL_INSTS	Estimated number of instructions processed the first and last time that the data source function is invoked.	'0'
INITIAL_IOS	Estimated number of I/Os performed the first and last time that the data source function is invoked.	'0'
INSTS_PER_ARGBYTE	Estimated number of instructions processed for each byte of the argument set that is passed to the data source function.	'0'
INSTS_PER_INVOC	Estimated number of instructions processed per invocation of the data source function.	'450'
IOS_PER_ARGBYTE	Estimated number of I/Os expended for each byte of the argument set that is passed to the data source function.	'0"
IOS_PER_INVOC	Estimated number of I/Os per invocation of a data source function.	'0'
PERCENT_ARGBYTES	Estimated average percent of input argument bytes that the data source function will actually read.	'100'

Example: PERCENT_ARGBYTES function mapping option

Suppose that you want to map a user-defined function named US_DOLLAR at an Oracle data source to a DB2 user-defined function that you create. The Oracle data source server is called ORACLE2. You decide to name the DB2 user-defined function DOLLAR and to name this function mapping ORACLE_DOLLAR. You want set the PERCENT_ARGBYTES function mapping option to provide the optimizer with more accurate overhead information. Use the following syntax:

```
CREATE FUNCTION MAPPING ORACLE_DOLLAR FOR DOLLAR()
   SERVER ORACLE2
   OPTIONS (REMOTE_NAME 'US_DOLLAR()', PERCENT_ARGBYTES '250')
```

Example: INSTS_PER_INVOC function mapping option

Suppose that you want to map the local function UCASE(CHAR) to an Oracle user-defined function called UPPERCASE. The Oracle function is at a data source called ORACLE2. You want to include the estimated number of instructions per invocation of the Oracle user-defined function. Use the following syntax:

```
CREATE FUNCTION MAPPING MY_ORACLE_FUN4 FOR SYSFUN.UCASE(CHAR)
   SERVER ORACLE2
   OPTIONS (REMOTE_NAME 'UPPERCASE', INSTS_PER_INVOC '1000')
```

Updating overhead information

If the estimates of consumed overhead change, you can record the change in the global catalog.

To record new estimates for the data source function, you must first drop or disable the function mapping. Then re-create the mapping specifying the new estimates in the CREATE FUNCTION MAPPING statement. The new estimates will be added to the SYSCAT.FUNCMAPPINGS catalog view. To record changed estimates for the DB2 function, you can update the SYSSTAT.ROUTINES catalog view directly.

Specifying function names in a function mapping

The values that you enter in the CREATE FUNCTION MAPPING statement depend on whether the functions that you are mapping together have the same name or different names.

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

Mapping functions with the same name

You can create a mapping between two functions (or a DB2 function template and a data source function) that have the same name.

Procedure

To map two functions with the same name, issue the CREATE FUNCTION MAPPING statement.

Example: You want to map a user-defined function named MYFUN at an Informix data source to the DB2 user-defined function named TINA.MYFUN. The Informix data source server is called INFORMIX2. The following statement maps the function:

```
CREATE FUNCTION MAPPING FOR TINA.MYFUN(SYSTEM.INTEGER) SERVER INFORMIX2
```

Mapping functions with different names

You can create a mapping between two functions (or a DB2 function template and a data source function) that have different names.

Procedure

To create a mapping between two functions with different names, issue the CREATE FUNCTION MAPPING statement:

1. Assign the name of the DB2 function or function template to the function_name parameter.
2. Specify a function mapping option called REMOTE_NAME and assign the name of the data source function to this option. The REMOTE_NAME must be less than 255 characters.

Example: You want to map a user-defined function named UPPERCASE at an Oracle data source to the DB2 function UCASE(CHAR). The Oracle data source server is called ORACLE2. You also want to include the estimated number of instructions per invocation of the UPPERCASE function. You decide to name this function mapping ORACLE_UPPER. The syntax would be:

```
CREATE FUNCTION MAPPING ORACLE_UPPER FOR SYSFUN.UCASE(CHAR)
    SERVER ORACLE2 OPTIONS
    (REMOTE_NAME 'UPPERCASE', INSTS_PER_INVOC '1000')
```

How to create function mappings

Use the CREATE FUNCTION MAPPING statement to specify alternative function mappings that override the default function mappings.

When you create alternative function mappings, the entries appear in the SYSCAT.FUNCMAPPINGS catalog view.

You can also use the CREATE FUNCTION MAPPING statement to specify function mapping options. When you specify function mapping options, the information appears in the SYSCAT.FUNCMAPOPTIONS catalog view.

With the CREATE FUNCTION MAPPING statement, you can:

- Create a function mapping for all data sources of a specific type. For example, all Informix® data sources.
- Create a function mapping for all data sources of a specific type and version. For example, all Informix 9 data sources.
- Create a function mapping for a specific server.
- Provide function mapping statistical information to the optimizer
- Disable a default function mapping or a function mapping that you defined.

You can issue the CREATE FUNCTION MAPPING statement in the DB2® Command Center or in the command line processor (CLP). You can also embed the CREATE FUNCTION MAPPING statement in an application program. The DB2 Control Center does not support creating or modifying function mappings.

Creating a function mapping for a specific data source type

You can create a mapping to a function for all data sources of a specific type.

Before you begin

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

Restrictions

You cannot override the existing function mappings or create new mappings for nonrelational data sources.

Procedure

To map a DB2 function template to a data source function, use the CREATE FUNCTION MAPPING statement.

Example: Map a DB2 function template to an Oracle user-defined function for all Oracle data sources

```
CREATE FUNCTION MAPPING MY_ORACLE_FUN1
     FOR NOVA.STATS ( DOUBLE, DOUBLE )
     SERVER TYPE ORACLE
     OPTIONS (REMOTE_NAME 'STAR.STATISTICS')
```

The template is called STATS and belongs to a schema called NOVA. The Oracle user-defined function is called STATISTICS and belongs to a schema called STAR.

Creating a function mapping for a specific data source type and version

You can create a mapping to a function for all data sources that use a specific version of the data source type.

Before you begin

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

Restrictions

You cannot override the existing function mappings or create new mappings for nonrelational data sources.

Procedure

To create a mapping for a specific data source type and version, use the CREATE FUNCTION MAPPING statement.

Example: Map a DB2 function template to a Sybase user-defined function for all Sybase data sources that use Version 12

The template is called SYB_STATS and belongs to a schema called EARTH. The Sybase user-defined function is called STATISTICS and belongs to a schema called MOON. The CREATE FUNCTION MAPPING is:

```
CREATE FUNCTION MAPPING SYBASE_STATS
   FOR EARTH.SYB_STATS ( DOUBLE, DOUBLE )
   SERVER TYPE SYBASE VERSION 12
   OPTIONS (REMOTE_NAME 'MOON.STATISTICS')
```

Creating a function mapping for all data source objects on a specific server

You can create a mapping to a function that is used by all data sources objects on a specific remote server.

Before you begin

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

Restrictions

You cannot override the existing function mappings or create new mappings for nonrelational data sources.

Procedure

To create a function mapping for all data source objects on a specific server, use the CREATE FUNCTION MAPPING statement.

Example: Map a function template called BONUS to a user-defined function called BONUS

You only want the mapping to apply to an Oracle data source server called ORA_SALES. Because the function names are the same, you do not need to specify the REMOTE_NAME function mapping option.

```
CREATE FUNCTION MAPPING BONUS_CALC FOR BONUS()
 SERVER ORA_SALES
```

User-defined functions in applications

Application developers often need to create their own suite of functions specific to their application or domain. They can use user-defined scalar functions for this purpose.

For example, a retail store could define a PRICE data type for tracking the cost of items that it sells. This store might also want to define a SALES_TAX function. This function would take a given price value as input, compute the applicable sales tax, and return this data to the requesting user or application.

These functions can operate over all database types, including large object types and distinct types. User-defined functions allow queries to contain powerful computation and search predicates to filter irrelevant data close to the source of the data, thereby reducing response time. The SQL optimizer treats user-defined functions exactly like built-in functions such as SUBSTR and LENGTH. You can develop applications using different application language environments, such as C, C++, COBOL, and FORTRAN. The applications can share a set of SQL user-defined functions even though they are developed using different application language environments.

User-defined functions can manipulate data and perform actions. For example, you might enable a user-defined function to send an electronic message or to update a flat file.

In DB2®, user-defined functions can include:

- Functions that you define from scratch.
- Functions in the SYSFUN schema. Examples include mathematical functions such as SIN, COS, and TAN; scientific functions such as RADIANS, LOG10, and POWER; and general purpose functions such as LEFT, DIFFERENCE, and UCASE.

Disabling a default function mapping

Default function mappings cannot be dropped. However, you can render them inoperable by disabling them.

Before you begin

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

Procedure

To disable a default function mapping, the CREATE FUNCTION MAPPING statement specifies the name of the DB2 function and sets the DISABLE option to 'Y'.

Example: Disable a default function mapping between the DB2 SIN function and a similar function on Oracle data sources

When a query that requests Oracle data and that references SIN is processed, either function might be invoked. The function invoked depends on which function is estimated by the query optimizer to require less overhead.

To ensure that the DB2 SIN function is invoked and that the Oracle SIN function is not invoked, you must disable the default function mapping. Use the following syntax:

```
CREATE FUNCTION MAPPING FOR SYSFUN.SIN(INT)
TYPE ORACLE OPTIONS (DISABLE 'Y')
```

Dropping a user-defined function mapping

When you no longer require a function mapping that you created, you can delete the function mapping.

Before you begin

The privileges held by the authorization ID of the statement must have SYSADM or DBADM authority.

About this task

If you drop a user-defined function mapping that was created to override a default function mapping, the default function mapping will be used.

User-defined function mappings are listed in the SYSCAT.FUNCMAPPINGS catalog view.

Procedure

To drop a function mapping that you created, use the DROP FUNCTION MAPPING statement.

Example: Drop a function mapping called BONUS_CALC

```
DROP FUNCTION MAPPING BONUS_CALC
```

Chapter 5. Creating index specifications

Index specifications in a federated system

In a federated system, you use the CREATE INDEX statement with a nickname to store information in the global catalog about the availability of an index on the remote object. The query optimizer uses this information to optimize queries.

When you issue a CREATE INDEX statement

- If a nickname is created for a table, the CREATE INDEX statement collects index information about the index that was created on the remote table.
- If a nickname is created for a view, the CREATE INDEX statement references the nickname for the view and contains information about the index on the table underlying the view.

The index specification tells the federated server about the columns and their uniqueness properties that comprise a remote index. It does not tell the federated server about the statistical properties of the index, such as, the number of unique values of the index key.

You do not need to supply index specifications if the remote index was in place at the time that the nickname was created.

A federated server does not create an index specification when you create a nickname for:

- A table that has no indexes
- A view, which typically does not have any index information stored in the remote catalog
- A data source object that does not have a remote catalog from which the federated server can obtain the index information

Suppose that a table acquires a new index, in addition to the ones it had when the nickname was created. Because index information is supplied to the global catalog at the time the nickname is created, the federated server is unaware of the new index. Similarly, when a nickname is created for a view, the federated server is unaware of the underlying table (and its indexes) from which the view was generated. In these circumstances, you can supply the necessary index information to the global catalog. You can create an index specification for tables that have no indexes. The index specification tells the query optimizer which column or columns in the table to search on to find data quickly.

Use index specifications with relational data sources. Creating an index specification for a nonrelational data source will not improve performance.

Creating index specifications for data source objects

When a nickname is created for a data source table, the federated server supplies the global catalog with information about any indexes that the data source table has. The optimizer uses this information to expedite the processing of distributed requests. This information is a set of metadata and is called an *index specification*.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- One of CONTROL privilege on the object or INDEX privilege on the object. And one of IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the index does not exist, or CREATEIN privilege on the schema, if the schema name of the index refers to an existing schema.

Restrictions

There are some restrictions when creating an index specification on a nickname.

- If the bind option DYNAMICRULES BIND applies, the statement cannot be dynamically prepared. Also, you cannot use the INCLUDE, CLUSTER, PCTFREE, MINPCTUSED, DISALLOW REVERSE SCANS, and ALLOW REVERSE SCANS parameters in the CREATE INDEX statement.
- UNIQUE should be specified only if the data for the index key contains unique values for every row of the data source table. The uniqueness will not be checked.
- The sum of the stored lengths of the specified columns must not be greater than 1024.
- No LOB column, DATALINK column, or distinct type column based on a LOB or DATALINK can be used as part of an index. This restriction is enforced even if the length attribute of the column is small enough to fit within the 1024-byte limit.

About this task

The federated server does not create an index specification if:

- A nickname is created for a table that has no index.
- A nickname is created for a data source object that does not contain indexes such as a view or Informix synonym.
- A nickname is created for a nonrelational object, for example, a table-structured file, Excel spreadsheet, BLAST algorithm, or XML tagged file.
- The remote index is on a LOB column.
- The remote index contains a total key length greater than 1024 bytes.
- The maximum number of key parts is more than 16.

In these circumstances the federated server does not store index specifications for the data source objects. However, for the first two items in the previous list you can supply the necessary index information to the global catalog. You can use the CREATE INDEX statement to specify the index information.

Procedure

To create an index, you can embed the CREATE INDEX statement in an application program or issue the statement as a dynamic SQL statement from the Control Center or the command line.

When used with nicknames, the CREATE INDEX statement creates an index specification in the federated global catalog; it does not create an index on the data source table.

Use the following syntax to create an index specification:

```
CREATE INDEX index_name ON nickname
(column_name) SPECIFICATION ONLY
CREATE UNIQUE INDEX index_name ON nickname
(column_name DESC) SPECIFICATION ONLY
```

For an index specification, *column_name* is the name by which the federated server references a column of a data source table.

Creating index specifications on tables that acquire new indexes

For situations in which a table acquires a new index, you should create an index specification on the nickname that corresponds to the table.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- One of CONTROL privilege on the object or INDEX privilege on the object. And one of IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the index does not exist, or CREATEIN privilege on the schema, if the schema name of the index refers to an existing schema.

Restrictions

There are some restrictions when creating an index on a nickname.

- If the bind option DYNAMICRULES BIND applies, the statement cannot be dynamically prepared. Also, you cannot use the INCLUDE, CLUSTER, PCTFREE, MINPCTUSED, DISALLOW REVERSE SCANS, and ALLOW REVERSE SCANS parameters in the CREATE INDEX statement.
- UNIQUE should be specified only if the data for the index key contains unique values for every row of the data source table. The uniqueness will not be checked.
- The sum of the stored lengths of the specified columns must not be greater than 1024.
- No LOB column, DATALINK column, or distinct type column based on a LOB or DATALINK can be used as part of an index. This restriction is enforced even if the length attribute of the column is small enough to fit within the 1024-byte limit.

About this task

There are several situations in which a table acquires a new index:

- You create a nickname for a table that does not have an index, but acquires an index later
- You create a nickname for a table that has an index, but acquires another index later

In these situations, you should create an index specification for the table so that the SQL Complier can use this information when processing queries that reference the table.

Procedure

The following examples describe how to create an index specification for a nickname that corresponds to a table that acquires an index.

Example: A table that has no index, later acquires an index

Suppose that you create the nickname *EMPLOYEE* for a data source table called CURRENT_EMP, which has no indexes. Sometime after this nickname is created, an index was defined on CURRENT_EMP using the WORKDEPT and JOB columns for the index key.

To create an index specification that describes this index, the syntax would be:

```
CREATE UNIQUE INDEX JOB_BY_DEPT ON EMPLOYEE
(WORKDEPT, JOB) SPECIFICATION ONLY
```

where *JOB_BY_DEPT* is the index name.

Example: A table acquires a new index

Suppose that you create the nickname *JP_SALES* for a table called JAPAN_SALES. A new index is later added to the table in addition to the ones it had when the nickname was created. The new index uses the MARKUP column for the index key.

To create an index specification that describes this index, the syntax would be:

```
CREATE UNIQUE INDEX JP_MARKUP ON JP_SALES (MARKUP) SPECIFICATION ONLY
```

where *JP_MARKUP* is the index name.

Creating index specifications on views

When a nickname is created for a view, the federated server is unaware of the underlying table (and its indexes) from which the view was generated. Create an index specification for the view so that the SQL Compiler can use this information when processing queries that reference the view.

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- One of CONTROL privilege on the object or INDEX privilege on the object. And one of IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the index does not exist, or CREATEIN privilege on the schema, if the schema name of the index refers to an existing schema.

Restrictions

There are some restrictions when creating an index on a nickname.

- If the bind option DYNAMICRULES BIND applies, the statement cannot be dynamically prepared. Also, you cannot use the INCLUDE, CLUSTER, PCTFREE, MINPCTUSED, DISALLOW REVERSE SCANS, and ALLOW REVERSE SCANS parameters in the CREATE INDEX statement.
- UNIQUE should be specified only if the data for the index key contains unique values for every row of the data source table. The uniqueness will not be checked.

- The sum of the stored lengths of the specified columns must not be greater than 1024.
- No LOB column, DATALINK column, or distinct type column based on a LOB or DATALINK can be used as part of an index. This restriction is enforced even if the length attribute of the column is small enough to fit within the 1024-byte limit.

Procedure

To create an index specification for a view:

- Ensure that the column or columns that the table index is based on is part of the view.
- If you want to create index specifications for all indexes on the underlying table, each index specification must be created separately.

Example: Create an index specification that describes the REGION index

Suppose that you create the nickname *JP_SALES2003* for a view called JAPAN_SALES2003. The underlying table for this view is the JAPAN_SALES table which contains several indexes: REGION, AMOUNT, SALES_REP. The CREATE INDEX statement you create will reference the nickname for the view and contain information about the index of the underlying table for the view.

```
CREATE UNIQUE INDEX JP_2003_REGION ON JP_SALES2003
(REGION) SPECIFICATION ONLY
```

where *JP_2003_REGION* is the index name, and *JP_SALES2003* is the nickname for the view JAPAN_SALES2003.

Creating index specifications on Informix synonyms

This topic describes the action that the federated server takes for Informix synonyms based on a table or on a view:

Before you begin

The privileges held by the authorization ID of the statement must include at least one of the following:

- SYSADM or DBADM authority
- One of CONTROL privilege on the object or INDEX privilege on the object. And one of IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the index does not exist, or CREATEIN privilege on the schema, if the schema name of the index refers to an existing schema.

Restrictions

There are some restrictions when creating an index on a nickname.

- If the bind option DYNAMICRULES BIND applies, the statement cannot be dynamically prepared. Also, you cannot use the INCLUDE, CLUSTER, PCTFREE, MINPCTUSED, DISALLOW REVERSE SCANS, and ALLOW REVERSE SCANS parameters in the CREATE INDEX statement.
- UNIQUE should be specified only if the data for the index key contains unique values for every row of the data source table. The uniqueness will not be checked.

- The sum of the stored lengths of the specified columns must not be greater than 1024.
- No LOB column, DATALINK column, or distinct type column based on a LOB or DATALINK can be used as part of an index. This restriction is enforced even if the length attribute of the column is small enough to fit within the 1024-byte limit.

About this task

In Informix, you can create a synonym for a table or view. While the DB2 federated server allows you to create nicknames for Informix synonyms, the action that the federated server takes depends on whether the synonym is based on a table or a view:

- Suppose that a nickname is created for a synonym, and the synonym is based on an Informix table. If the federated server determines that the table the synonym refers to has an index, then an index specification is created for the synonym. If the table that the synonym refers to does not have an index, then no index specification is created for the synonym. However you can create an index specification manually, using the CREATE INDEX statement.
- Suppose that a nickname is created for a synonym, and the synonym is based on an Informix view. The federated server can not determine which underlying table or tables the view is based on. Therefore no index specification is created for the synonym. However you can create an index specification manually using the CREATE INDEX statement.

Procedure

The following examples describe how to create an index specification on a nickname that corresponds to an Informix synonym.

Example: A nickname is created on an Informix synonym that is based on a table

When the synonym is based on an Informix table that does not contain an index, you can create an index specification for the synonym to tell the optimizer which column or columns to search on to find data quickly. The statement you create will specify the nickname for the synonym, and you will supply information about the column or columns in the table that the synonym is based on.

In this example, you create the nickname *CONTRACTS* for a synonym called SALES_CONTRACTS. The table that this synonym is based on is called SALES2006_TABLE and contains several indexes: REGION, AMOUNT, SALES_REP. The CREATE INDEX statement you create will reference the nickname for the synonym and contain information about the index of the underlying table for the synonym.

To create an index specification that describes the REGION index, the syntax would be:

```
CREATE UNIQUE INDEX NORTHWEST_2006_REGION ON CONTRACTS (REGION) SPECIFICATION ONLY
```

where *NORTHWEST_2006_REGION* is the index name and *CONTRACTS* is the nickname for the synonym SALES_CONTRACTS.

Example: A nickname is created on an Informix synonym that is based on a view

You create the nickname *JP_SALES2003* for a synonym based on a view called JAPAN_SALES2003. The underlying table for this view is the JAPAN_SALES table which contains several indexes: REGION, AMOUNT, SALES_REP. The CREATE INDEX statement that you create will reference the nickname for the synonym and contain information about the index of the underlying table for the view.

When creating an index specification for a synonym based on a view, make certain that the column or columns the table index is based on, is part of the view. If you want to create index specifications for all indexes on the underlying table, each index specification must be created separately.

To create an index specification that describes the REGION index, the syntax would be:

```
CREATE UNIQUE INDEX JP_2003_REGION ON JP_SALES2003 (REGION) SPECIFICATION ONLY
```

where *JP_2003_REGION* is the index name and *JP_SALES2003* is the nickname for the view JAPAN_SALES2003.

Chapter 6. Developing federated procedures

Federated procedures enable you to invoke procedures at a data source as if the remote procedure is a local procedure.

Federated procedures

A *federated procedure* is a federated database object that references a procedure on a data source. Federated procedures are sometimes called federated stored procedures.

Federated procedures are not alternative names for data source procedures in the same way that aliases are alternative names. A federated procedure is defined at the federated database but calls a data source procedure when the federated procedure is invoked. Because the federated procedure is a federated database object, users or client applications can invoke the data source procedure logic by calling a federated database procedure. The results of the data source procedure, such as the output parameters, are returned by the federated procedure. Using a federated procedure makes the location of the data source procedure transparent to the user or the client application. You call the data source procedure by using the name of the federated procedure.

A federated procedure is to a remote procedure what a nickname is to a remote table. Nicknames and federated procedures are objects on the federated database. A nickname is an object that references an object, such as a table or view, on the data source. With a nickname, you can query a data source object. With a federated procedure, you can call a data source procedure

Use the CREATE PROCEDURE (Sourced) statement to register a federated stored procedure. You can embed the CREATE PROCEDURE (Sourced) statement in an application program or issue the statement with dynamic SQL statements.

When you create a federated procedure, the data types for the parameters for the data source procedure are mapped to the federated data types using the default forward data type mappings.

You can create federated procedures for the following data sources:

- Oracle
- Sybase

For Oracle data sources, you can create federated procedures for Oracle procedures or functions.

Important: You can only create federated procedures when you use a trusted, or unfenced, wrapper to access the data source.

Restrictions on federated procedures

There are incompatibilities with and limitations on the federated procedures.

The restrictions that apply to local procedures also apply to federated procedures. The key restrictions on federated procedures are as follows:

- Federation returns a maximum of one result set when you use federated procedures.
- You can create federated procedures for Oracle and Sybase procedures.
- The federated procedures must be associated with a trusted wrapper.

The following sections highlight the additional restrictions that apply federated procedures.

Result sets

For Oracle procedures, federation returns only one result set. If the data source procedure returns multiple result sets, only the first result set is returned to the user or client application. The other result sets are discarded, and the SQL0464W warning message is returned.

For Sybase procedures that return an output parameter and a result set, the result set is discarded. If the Sybase procedure returns a result set and return value, a return value of 0 is provided, regardless of the actual return value from the data source procedure. You will not receive a warning message in either of these situations.

However, the result sets are discarded and the SQL0464W message is returned when both of the following conditions are true:

- The federated procedure is defined for a Sybase procedure that returns result sets, and
- The federated procedure is invoked inside a trigger or a user-defined function.

Federated procedures do not support WITH HOLD cursors, scrollable cursors, or updatable cursors. If the data source procedure uses these types of cursors, you do not receive a warning or error message. Applications that access result sets with data source procedure that use these types of cursors might behave differently. Typically the cursors are returned without the hold capability and forward read-only cursors.

Sybase data sources

For Sybase server version 12.0, all of the parameters are input parameters. Federated procedures cannot return an output parameter value. This is a Sybase catalog limitation and does not apply to higher versions of Sybase, such as version 12.5 and version 15.

The DB2_RETURN_STATUS value is retrieved for Sybase procedures that return result sets, but always returns a value of zero (0) regardless of the actual return value from the Sybase procedure.

The Sybase wrapper cannot call a procedure in these situations:

- When another procedure is already called
- When another statement is executed during a single connection

To work around these limitations you can define the federated procedure on another server. For example, the following statements succeed if the nickname *syb_nick* and the procedure *syb_proc* are defined on the different servers. If the nickname and the procedure are defined on the same server, the statements fail.

```
DECLARE clientcur CURSOR
FOR SELECT colsml,coldec,colvch,coltsp
FROM syb_nick
OPEN clientcur;
CALL syb_proc();
```

Federated clients

To retrieve data from remote result sets, your federated client must be DB2 Universal Database for Linux, UNIX, and Windows, version 8 or higher, or DB2 Database for z/OS.

Procedure calls and access levels

Procedure calls and access levels have these issues:

- By default, the CALL RESOLUTION IMMEDIATE clause is used with federated procedures when you issue the PRECOMPILE command. The CALL RESOLUTION DEFERRED clause is not supported with federated procedures.
- You cannot not see output when a federated procedure calls a data source procedure that prints to a buffer or the standard output.
- When you call a federated procedure from an external user-defined function, the federated procedure must not have the access level set to READS SQL DATA or MODIFIES SQL DATA. Federated access is blocked inside external functions.

Data types

For the parameters and result sets, you can use all of the data types that are supported for nickname columns except LOB data types. Federated procedures do not support complex data types.

Transactions

Consider the following issues when using federated procedures with transactions:

- The data source procedure that the federated procedure references must not issue a COMMIT or a ROLLBACK statement. Federation does not enforce this restriction and data inconsistencies might occur if the data source procedure issues a COMMIT or ROLLBACK statement.
- Federated procedures with the MODIFIES SQL DATA access level cannot be invoked inside of triggers, dynamic compound statements, SQL scalar, tables, row functions, and methods. After a SAVEPOINT statement is issued, you cannot call a federated procedure with the MODIFIES SQL DATA access level.

DataJoiner® stored procedure syntax

The DataJoiner syntax to create stored procedure nicknames is not supported. DataJoiner stored procedure nicknames that are migrated to WebSphere Federation Server must be dropped. Create federated procedures to replace the DataJoiner stored procedure nicknames by using the CREATE PROCEDURE (Sourced) statement.

How procedures are invoked has also changed. Any applications or scripts that invoke DataJoiner stored procedure nicknames will need to be migrated to WebSphere Federation Server.

Overloaded procedures in federated systems

Overloaded procedures are procedures that have identical names and schemas. Overloaded procedures can have a different number of parameters or different parameter signatures. The purpose of overloading a procedure is to create similar versions of a procedure.

The federated server allows overloaded procedures only if each procedure has a different number of parameters.

Oracle allows overloaded procedures if each procedure has a different number of parameters or if the parameter types are different. You can create federated procedures for Oracle overloaded procedures.

To distinguish procedures that use the identical name, schema name and number of parameters, you must specify the UNIQUE ID when you create the federated procedure.

There are two ways that you can determine the UNIQUE ID, by using the discovery feature in the Control Center and by querying the Oracle catalog.

Use the discover feature in the Control Center

When you create a federated procedure in the Control Center, the discovery feature determines if the Oracle procedure is overloaded. The overloaded value is displayed in the UNIQUE ID field for each Oracle procedure.

Query the Oracle catalog

To determine the UNIQUE ID, you can query the OVERLOAD column in the Oracle SYS.ALL_ARGUMENTS catalog. The UNIQUE ID value is a character literal that contains a number, such as 1. You can use the passthru mode on the federated server or query the Oracle client directly.

For example, to query the Oracle catalog and display the signature and overload column for a procedure that begins with HJZ, use the following SELECT statement:

```
SELECT owner, package_name, object_name, overload, position, argument_name, in_out,
data_type
   FROM all_arguments aa
   WHERE object_name like 'HJZ%'
   ORDER BY owner, package_name, object_name,overload, position;
```

The following output from the above query shows that the HJZ_PACK1 package contains three procedures that use the name HJZTEST1. You determine the number of procedures by looking at the OBJECT_NAME AND OVERLOAD columns. The first procedure has one IN parameter with a number data type. The second procedure has one IN parameter with a character data type. The third procedure has one OUT parameter with a character data type and one IN parameter with a number data type. The output shows that there are two procedures in the HJZ_PACK1 package that use the name HJZTEST3. There is also a procedure with the name HJZTEST1 that is not in the package. This last procedure has one IN parameter that uses a number data type.

```
OWNER    PACKAGE_NAME OBJECT_NAME OVERLOAD POSITION ARGUMENT_NAME IN_OUT DATA_TYPE
-------- ------------ ----------- -------- -------- ------------- ------ ---------
J15USER1 HJZ_PACK1    HJZTEST1    1        1        A             IN     NUMBER
J15USER1 HJZ_PACK1    HJZTEST1    2        1        A             IN     CHAR
J15USER1 HJZ_PACK1    HJZTEST1    3        0        -             OUT    CHAR
J15USER1 HJZ_PACK1    HJZTEST1    3        1        A             IN     NUMBER
```

```
J15USER1 HJZ_PACK1    HJZTEST3    1        1        A             IN     NUMBER
J15USER1 HJZ_PACK1    HJZTEST3    1        2        B             OUT    NUMBER
J15USER1 HJZ_PACK1    HJZTEST3    2        1        A             IN     CHAR
J15USER1 HJZ_PACK1    HJZTEST3    2        2        B             OUT    CHAR
J15USER1 -            HJZTEST1    -        1        A             IN     NUMBER
9 record(s) selected.
```

To create a federated procedure for the second overloaded procedure with an IN parameter of CHAR data type, issue the following CREATE PROCEDURE statement:

```
CREATE PROCEDURE HJZTEST1 SOURCE J15USER1.HJZ_PACK1.HJZTEST1
 NUMBER OF PARAMETERS 1 UNIQUE ID '2'
 FOR SERVER ORA_SERVER WITH RETURN TO CLIENT ALL;
```

Important: In the above example, the NUMBER OF PARAMETERS clause does not uniquely identify the procedure. There are two procedures in the table with the name HJZTEST1 and each procedure has one parameter. You must specify the UNIQUE ID clause to indicate the overloaded procedure that you want to use. Use the value from the OVERLOAD column as the value for the UNIQUE_ID clause. When the UNIQUE ID clause is specified, the NUMBER OF PARAMETERS clause is optional. Use the NUMBER OF PARAMETERS clause to validate that the data source procedure has the number of parameters that you expect.

Creating federated procedures

You must create a federated procedure for each data source procedure that you want to invoke from the federated server.

Before you begin

- The procedure must already exist on the data source.
- Register a wrapper and server definition for the data source that contains the procedures that you want to use.
- For data sources that require user mappings, create the user mappings between the federated server and the data source server.

Restrictions

- The authorization ID of the statement must have at least one of the following:
  - IMPLICIT_SCHEMA privilege on the database, if the schema name of the procedure does not refer to an existing schema
  - CREATEIN privilege on the schema, if the schema name of the procedure refers to an existing schema
  - SYSADM or DBADM authority
- For data sources that require user mappings, the data source authorization ID for the CREATE PROCEDURE statement must include the privilege to select the procedure description from the data source catalog tables.

About this task

In the CREATE PROCEDURE (Sourced) statement, the *source* procedure is the procedure on the data source. The *sourced* procedure is the federated procedure.

For Oracle data sources, you can create federated procedures for Oracle procedures or functions. Oracle functions are similar to Oracle procedures except that Oracle functions use a return value. You can create a federated procedure and a function

mapping for the same Oracle function. Use a function mapping when you need to use the Oracle function in SQL as a scalar function. Use a federated procedure when you need to use the Oracle function in CALL statements. The return value for the federated procedure appears at the beginning of the parameter list as an extra OUT parameter. The name of the parameter is always DEFAULT.

You can create a federated procedure by using the Control Center or from the command line. The Control Center populates the fields and settings that are based on information from the data source procedure.

Procedure

To create a federated procedure, use one of the following methods:

Method	Description
Control Center	1. Right-click the **Federated Stored Procedures** folder and click **Create**. The **Federated Stored Procedures** folder is located under the wrapper and server definition that you registered for the data source. The **Federated Stored Procedures** folder appears only when the wrapper that you registered is a trusted wrapper. 2. In the **Discover** window, generate a list of potential data source procedures. To specify the information for a single procedure, click **Add**. 3. In the **Create Federated Stored Procedure window**, select the check box next to the procedures that you want to create. 4. Click **OK** .

Command line	Issue the CREATE PROCEDURE (Sourced) statement. For example:
	`CREATE PROCEDURE`
	`  federated_procedure_name`
	`  SOURCE remote_schema_name.`
	`remote_procedure_name`
	`  FOR SERVER server_definition_name;`

Use the SYSCAT.ROUTINESFEDERATED catalog view to see a list of the federated procedures that you created.

To make changes to a federated procedure, you must drop the procedure and create the procedure again with the new settings.

Discovering data source procedures

You can retrieve information about the data source procedures by using the **Discover** window in the Control Center.

About this task

To discover data source procedures, the federated server searches the system catalogs on the data source and retrieves information about the data source procedures. The Control Center automatically populates the fields with the required information.

Procedure

To discover data source procedures:

1. Access the **Create Federated Stored Procedure** window. Right-click the **Federated Stored Procedures** folder and click **Create**.
2. In the **Create Federated Stored Procedure** window, click **Discover**.
3. In the **Discover** window, specify criteria that filters the search by schema, package name, or procedure or function name. By default, system objects are excluded from the search.

Table 6. Examples of criteria that you can specify in the Discover window

Criteria	Action
Schema names that start with a range of letters of the alphabet, such as the letters, A to P	Use the BETWEEN operator and specify the value `'A' AND 'P'` for the remote schema. The filter returns information about the data source procedures and functions with schemas like ADMIN1, ADMIN2, ORA1, ORA2, and PDB.
Specific schema names such as ADMIN1 and PDB	Use the IN operator and specify the value `'ADMIN1','PDB'` for the remote schema. The filter returns information about the data source procedures and functions that are only in the ADMIN1 or the PDB schemas.
All schemas except SCHEMA3 or SCHEMA6	Use the NOT IN operator and specify the value `'SCHEMA3' OR 'SCHEMA6'`.
Procedure names that begin with PROC	Use the LIKE operator and specify the value `PROC%`.

Note: When you specify a filter, you must use single quotation marks around the values that you specify in the **Value** fields.

Tip: Use the **Count** button in the **Discover** window to determine how many procedures will be located based on your criteria. If the number of procedures is large, specify additional criteria to limit the search.

To review the information that is retrieved from the data source, select a procedure from the list in the **Create Federated Stored Procedures** window and click **Properties**. Verify the information and ensure that all of the required fields are completed before you create the federated procedures.

Input and output parameters for federated procedures

The data types that are used by the data source procedure are mapped to federated data types. This mapping includes data types for the input and output parameters and data types for the columns in the result set. The federated server uses the default forward data type mappings.

You can use the CREATE TYPE MAPPING statement to override the default type mapping for the data source parameters. However, the type mappings of the result set are not affected by user-defined type mappings.

Each argument in the CALL statement must be compatible with the corresponding parameter in the procedure. Federated procedures follow the same parameter assignment rules as local procedures.

Federation uses three types of parameters: IN, OUT, and INOUT.

Oracle procedures and functions

You can create federated procedures for Oracle functions. Oracle returns only a single value for functions. The return value for the function is included as an output parameter at the beginning of the function argument list. When you specify the NUMBER OF PARAMETERS clause in the CREATE PROCEDURE (Sourced) statement, do not count the return values.

For some Oracle data types, information about the precision, length and scale is not stored in the Oracle catalog when the parameters of a procedure are declared. When a federated procedure is created, information about the Oracle procedure is gathered from the Oracle catalog. Because information about the precision, length and scale is not stored in the Oracle catalog, the federated procedures behave in the following way:

- Uses the maximum length for the parameter data types.
- Maps the Oracle NUMBER data types to the federated DOUBLE data type. You can change this mapping by overriding the default forward data type mapping for Oracle NUMBER.

Tip: Overriding the default forward data type mappings will affect other federated DDL operations, such as CREATE NICKNAME. To avoid this problem, change the type mapping before you create the procedure. Create the procedure for the Oracle procedure with the new type mapping, then DROP the new type mapping. Subsequent nicknames and procedures that you create will use the default type mapping.

Sybase procedures

Sybase procedures use INPUT and OUTPUT parameters. The Sybase wrapper maps a Sybase INPUT parameter to a federated IN parameter and a Sybase OUTPUT parameter to a federated INOUT parameter.

The Sybase wrapper can return either an OUTPUT parameter or a result set. When a Sybase procedure returns both an OUTPUT parameter and a result set, only the parameter is returned. The result set is discarded. You do not receive an error or warning when the result set is discarded.

Although you can use optional procedure parameters in Sybase, you cannot use optional procedure parameters in federated procedures. You must specify all of the parameters including the optional parameters, when you call the federated procedure.

CREATE PROCEDURE (Sourced) statement - examples

Use the CREATE PROCEDURE (Sourced) statement to create federated procedures. The examples show the required parameters, the optional parameters, and parameters for specific data sources.

Required parameters

The following example shows the required parameters when you create a federated procedure:

```
CREATE PROCEDURE PROC1 SOURCE BHATIA.PROC1_SYBASE
   FOR SERVER SYBASE_SERVER;
```

PROC1
Specifies the name of the federated procedure.

SOURCE BHATIA.PROC1_SYBASE
Specifies the name of the schema and procedure on the data source.

FOR SERVER SYBASE_SERVER
Specifies a server definition where the federated procedure is created.

Optional parameters - Sybase procedures

The following example shows the optional parameters when you create a federated procedure for a Sybase procedure:

```
CREATE PROCEDURE PROC1 SOURCE BHATIA.PROC1_SYBASE
   NUMBER OF PARAMETERS 3 FOR SERVER SYBASE_SERVER
   SPECIFIC MYPROC1 WITH RETURN TO CLIENT ALL
   MODIFIES SQL DATA DETERMINISTIC EXTERNAL ACTION;
```

NUMBER OF PARAMETERS 5
Specifies the total number of IN, OUT, and INOUT parameters that the Sybase procedure uses. Use this parameter when you have more than one procedure with the same schema name and procedure name. For example, if your schema is BHATIA and you have a PROC1 procedure with three parameters and another PROC1 procedure with one parameter, the name for both of these procedures is BHATIA.PROC1. The value for the NUMBER OF PARAMETERS in the data source procedure indicates which procedure you refer to in the CREATE PROCEDURE statement.

SOURCE BHATIA.PROC1_SYBASE
Specifies the schema and name for the Sybase procedure. For Sybase procedures, you specify a two-part name in the CREATE PROCEDURE statement. The format for this two-part name is *source_schema_name.source_procedure_name*.

SPECIFIC MYPROC1
Specifies a unique name for the federated procedure that you are creating. This parameter is used only for federated procedures and is not associated with data source procedures. If you do not specify a unique name, a name is generated by the federated database manager.

WITH RETURN TO CLIENT ALL
Specifies that the result set is returned to the client application. Federation returns a maximum of one result set. If this parameter is not specified, the default is WITH RETURN TO CALLER ALL.

MODIFIES SQL DATA
Indicates the level of data access for SQL statements that are included in the federated procedure. If the clause specified does not match the Sybase

procedure, an error message is returned. If you do not specify this clause, the clause for the Sybase procedure is used.

DETERMINISTIC
Specifies if the federated procedure always returns the same results for a given set of argument values. This parameter can improve the performance of the interaction between the federated server and the data source.

EXTERNAL ACTION
Specifies if the federated procedure takes an action that changes the state of an object that is not managed by the database manager.

Optional parameters - Oracle procedures

The following example shows the optional parameters when you create a federated procedure for an Oracle procedure:

```
CREATE PROCEDURE PROC2 SOURCE ZELLER_SCHEMA.ORACLE_PKG9.PROC2
   NUMBER OF PARAMETERS 5 UNIQUE_ID '2' FOR SERVER ORA_SERVER
   SPECIFIC MYPROC1 WITH RETURN TO CLIENT ALL
   MODIFIES SQL DATA DETERMINISTIC NO EXTERNAL ACTION;
```

NUMBER OF PARAMETERS 5
Specifies the total number of IN, OUT, and INOUT parameters that the Oracle procedure uses. Use this parameter when you have more than one procedure with the same schema name and procedure name. For example, if your schema is ZELLER and you have a PROC1 procedure with two parameters and another PROC1 procedure with three parameters, the name for both of these procedures is ZELLER.PROC1. The value for the NUMBER OF PARAMETERS in the data source procedure indicates which procedure you refer to in the CREATE PROCEDURE statement. Oracle REFCURSOR parameters must be included in the NUMBER OF PARAMETERS count.

SOURCE ZELLER_SCHEMA.ORACLE_PKG9.PROC2
Specifies the schema, package, and name for the Oracle procedure or function. If the Oracle procedure or function is in a package, you must specify a three-part name in the CREATE PROCEDURE statement. The format for this three-part name is *source_schema_name.source_package_name.source_procedure_name*. If the Oracle procedure or function is not in a package, you must specify a two-part name in the CREATE PROCEDURE statement. The format for this two-part name is *source_schema_name.source_procedure_name*.

UNIQUE_ID '2'
Specifies the unique identifier for the Oracle procedure. Use the UNIQUE_ID parameter only when the schema name, the procedure name, and the number of parameters do not uniquely identify an Oracle procedure. The UNIQUE ID is the value in the ALL_ARGUMENTS.OVERLOAD column in the Oracle system catalog. If you do not specify the UNIQUE ID parameter, the federated server detects the overloaded procedures and returns an error. Use this option only with Oracle procedures.

SPECIFIC MYPROC1
Specifies a unique name for the federated procedure that you are creating. This parameter is used only for federated procedures and is not associated with data source procedures. If you do not specify a unique name, a name is generated by the federated database manager. This parameter is optional.

WITH RETURN TO CLIENT ALL
Specifies that the result set is returned to the client application. Federation returns a maximum of one result set. If this parameter is not specified, the default is WITH RETURN TO CALLER ALL.

MODIFIES SQL DATA
Indicates the level of data access for SQL statements that are included in the federated procedure. If the clause specified does not match the Oracle procedure, an error message is returned. If you do not specify this clause, the clause for the Oracle procedure is used.

DETERMINISTIC
Specifies if the federated procedure always returns the same results for a given set of argument values. This parameter can improve the performance of the interaction between the federated server and the data source.

NO EXTERNAL ACTION
Specifies if the federated procedure takes an action that changes the state of an object that is not managed by the database manager.

Granting or revoking authorizations to call federated procedures

The administrator of the federated database must grant other users the required authorizations to call the federated procedures.

Before you begin

The user that calls the federated procedure must have a valid user mapping from the federated server to the data source. The remote user ID from the user mapping must have the authorization on the data source that is equivalent to the EXECUTE authorization on the federated server. A user can be granted EXECUTE authorization on the federated procedure. But if the user authorization on the data source is not equivalent to the EXECUTE authorization on the federated server, calls to the data source procedure fail.

The authorization ID for the GRANT statement must have at least one of the following authorities:

- The WITH GRANT OPTION for EXECUTE on the federated procedure
- SYSADM or DBADM authority

Procedure

To grant or revoke the authorization to call federated procedures:

Specify the authorization privileges by using the Control Center or from the command line:

Method	Description
Control Center	1. Right-click the name of the federated procedure and click **Privileges**. 2. Select the user or group that you want to set privileges for. To add a new user, click **Add User**. To add a new group, select the Group tab and click **Add Group**. 3. From the **Privileges: EXECUTE** drop-down box: • Select **Yes** to grant only the EXECUTE privilege. • Select **Grant** to grant the EXECUTE and WITH GRANT OPTION privileges. • Select **No** to remove the EXECUTE privilege from the user or group. 4. You can grant or revoke privileges to multiple users at the same time. Select the users from the list and click **Grant All** or **Revoke All** to grant or revoke the EXECUTE privilege for the selected users. 5. Click **OK**.

Command line	Specify the privileges in the GRANT statement.
	Example 1:
	To grant the EXECUTE privilege on all procedures in the BHATIA schema, including any procedures that are created in the future, to users in the HR_DEPT group, use the following statement:
	`GRANT EXECUTE ON PROCEDURE BHATIA.* TO HR_DEPT`
	Example 2:
	To grant the EXECUTE privilege on the PROC1 procedure to user ZELLER and give the this person the ability to grant the EXECUTE privilege on this procedure to others, use the following statement:
	`GRANT EXECUTE ON PROCEDURE PROC1 TO ZELLER WITH GRANT OPTION`
	Example 3:
	To grant the EXECUTE privilege to user ERFAN on the PROC2 procedure that was created with a specific name of MY_PROC2, use the following statement:
	`GRANT EXECUTE ON SPECIFIC PROCEDURE MY_PROC2 TO ERFAN`

Locating parameter information

After you create the federated procedures, you can find information about the input and output parameters by using the Control Center or by querying the catalog views.

About this task

There are system catalog views that contain information about the federated procedures:

- SYSCAT.ROUTINES
- SYSCAT.ROUTINESFEDERATED
- SYSCAT.ROUTINEOPTIONS
- SYSCAT.ROUTINEPARMS
- SYSCAT.ROUTINEPARMOPTIONS

Use the SYSCAT.ROUTINESFEDERATED view to determine the data source server that the federated procedure is associated with.

Procedure

To locate federated procedure information:

Use one of the following methods:

Method	Description
Control Center	Select the name of the federated procedure. The parameter information is displayed in the details pane on the right side of the Control Center window.
Command line	Query the SYSCAT.ROUTINEPARMS view.

For example, you create a federated procedure BHATIA.FEDPROC1 that is for the data source procedure ZELLER.EMPLOYEE. The data source procedure uses one input and one output parameter. You issue this statement:

```
CREATE PROCEDURE BHATIA.FEDPROC1 SOURCE ZELLER.EMPLOYEE
   NUMBER OF PARAMETERS 2 FOR SERVER S1;
```

The following examples show SELECT statements that find information about the input and output parameters in the SYSCAT.ROUTINEPARMS catalog view:

```
SELECT rowtype, parmname, typename, ordinal
   FROM syscat.routineparms
   WHERE routinename='FEDPROC1' AND routineschema='BHATIA';

SELECT rowtype, char(parmname,30), char(typename,30), ordinal
   FROM syscat.routineparms
   WHERE routinename='FEDPROC1' AND routineschema = 'BHATIA'
   ORDER BY ordinal;
```

The **routineschema** parameter is the schema name of the federated procedure and not the data source procedure.

Calling federated procedures

You call data source procedures the same way that you call local procedures. You specify the federated procedure, which references the data source procedure, in a CALL statement.

Before you begin

- The user that calls the federated procedure must have EXECUTE privilege on the data source and a valid user mapping from the federated server to the data source.
- For data sources that require a user mapping, the data source authorization ID for the statement must have the privilege to call the data source procedure.

Restrictions

There are incompatibilities and limitations on calling federated procedures.

About this task

You can call a federated procedure by using the Command Editor or from the command line.

Procedure

To call a federated procedure:

1. You must know the input and output parameters for the federated procedure before you issue a CALL statement. Use one of the following methods to look up the parameter information:

Method	Description
Control Center	Click the **Federated Stored Procedures** folder. Information about the parameters appears in the properties pane on the lower right side of the Control Center window. You can see the name of each parameter, the parameter data types, and the parameter mode (IN, OUT, INOUT). To access the Command Editor, click **Tools, Command Editor**.

Method	Description

Command line	Issue a SELECT statement to look up the parameter information in the SYSCAT.ROUTINEPARMS view in the federated database catalog. You can look up information, such as the ordinal number, the parameter name, and the parameter data type. For example, if the federated procedure FEDPROC1 is in the federated schema BOB, issue this SELECT statement:
	`SELECT ordinal, char(parmname,30) AS name, rowtype, char(typename,30) AS type FROM syscat.routineparms WHERE routinename='FEDPROC1' AND routineschema = 'BOB' ORDER BY ordinal;`
	The result of the query lists the parameters:
	`ORDINAL NAME ROWTYPE TYPE`
	`------- ---- ------- --------`
	`1 P1 P INTEGER`
	`2 P2 O VARCHAR`
	The P row type indicates an input parameter. The O row type indicates an output parameter. There is also a B row type that indicates both an input and output parameter, or INOUT parameter.

2. Issue the CALL statement either on the command line or in the Command Editor. For example, if the FEDPROC1 federated procedure is defined with one input parameter and one output parameter, you issue this CALL statement:
   ```
   CALL FEDPROC1(10 , ?)
   ```

Authorization to call federated procedures

To call a federated procedure, you must have the correct authorizations on the federated procedure and the data source procedure.

When you call a federated procedure, the user mapping and authorization privileges of the user ID that created the federated procedure are used to access the data source tables.

Example

The user ZELLER creates a federated procedure called FP1. The FP1 procedure references a Sybase procedure that accesses a Sybase table. The remote authorization ID in the user mapping for ZELLER has the privilege to update the Sybase table. The user ZELLER grants the EXECUTE privilege to the user BHATIA on the FP1 procedure. The user BHATIA must have a valid user mapping to a remote authorization ID that has EXECUTE privilege on the Sybase procedure that is referenced by the FP1 procedure. The remote authorization ID that user BHATIA is mapped to does not need to have SELECT privilege on the Sybase procedure. When the user BHATIA calls the FP1 procedure, the user BHATIA can update the table in Sybase.

Altering or dropping federated procedures

You cannot alter a federated procedure directly. To make changes to a federated procedure, you must drop the procedure and create the procedure again with the new settings.

Before you begin

The Authorization ID for the DROP PROCEDURE statement must have one of the following authorities:

- SYSADM or DBADM authority
- DROPIN privilege on the schema for the federated procedure
- Definer of the procedure as recorded in the DEFINER column of the catalog view for the federated procedure
- CONTROL privilege on the federated procedure

About this task

When you drop a federated procedure, the procedure is deleted from the system catalog on the federated database. The data source procedure that the federated procedure references is not affected. When you drop a federated procedure, the applications that are dependent on the dropped procedures are invalidated.

You can drop a federated procedure by using the Control Center or from the command line.

Procedure

To drop a federated procedure, use one of the following methods:

Method	Procedure
Control Center	1. Expand the **Federated Objects** folder, the **Server definitions** folder, and the **Federated Stored Procedures** folder in the object tree. 2. Right-click the federated procedure that you want to drop and click **Drop**.
Command line	Use the DROP statement. For example: `DROP PROCEDURE federated_procedure_name`

Federated procedure troubleshooting

If you encounter problems with federated procedures, there are several ways that you can troubleshoot the problems.

The following queries and diagnostic tools help you to view information about the federated procedures. This information will assist you in resolving problems with the federated procedures.

Verify data source procedure information

If the SQL1253N error is returned when you issue a CREATE PROCEDURE statement, you can issue the following queries against the catalog tables on the

data source to verify information about the data source procedure. The SQL1253N error indicates that the source procedure specified in the CREATE PROCEDURE (Sourced) statement was not found at the data source. You can query the Oracle server directly or use a pass-through session to query the Oracle server.

For Oracle procedures that are in a package:

```
SELECT owner, package_name, object_name, overload, parm_count
   FROM (
      SELECT owner, package_name, object_name, overload,
      SUM(case
         WHEN data_type IS NULL
         THEN 0
         ELSE 1
         END)
      AS parm_count
      FROM sys.all_arguments
      WHERE data_level = 0
      GROUP BY owner, package_name, object_name, overload
      ) aa
   WHERE object_name = '' AND
      package_name = '' AND
      owner = '' AND
      overload = '' AND   <-- optional
      parm_count =;       <-- optional
```

For Oracle procedures that are not in a package:

```
SELECT object_name, object_type, status
   FROM sys.all_objects
   WHERE owner = '' AND
     object_name = '' AND
     object_type IN ('PROCEDURE', 'FUNCTION')
```

For Sybase procedures:

```
SELECT id
   FROM dbo.sysobjects
   WHERE id = object_id('.') AND
     (TYPE = 'P' OR TYPE ='XP')
```

Diagnostic tools

Use the Explain utility, the DESCRIBE command, or the db2audit command to diagnose problems with federated procedures.

For example, the FED_PROC1 procedure has three OUTPUT parameters. To use the DESCRIBE command on the FED_PROC1 procedure, issue the following command:

```
DESCRIBE CALL FED_PROC1(?,?,?);
```

System monitor

The system monitor elements in the federated database contain information about federated procedures. The monitor elements are as follows:

- The Stored Procedure Time monitor element, stored_proc_time, contains the time it has taken the data source to respond to federated procedure statements.
- The Rows Returned by Stored Procedures monitor element, sp_rows_selected, contains the number of rows that are sent from the data source to the federated server. You can use this element to calculate the average number of rows sent to

the federated server from the data source for each federated procedure. Or, you can also calculate the average time to return a row to the federated server from the data source.

- The Stored Procedures monitor element, stored_procs, contains a count of the total number of procedures that the federated server has called from this data source.

SQL error SQL30090 with return code 21

There are several situations in which why the SQL30090 error with return code 21 is returned. One of the most common situations is when a federated procedure is being created using a fenced wrapper. Federated procedures can be created only on trusted wrappers.

Result set not returned

A result set might not be returned to the client or caller for one of the following reasons:

- The clause for returning result sets is not specified correctly in the federated procedure.
- Some data sources do not return sets in the same order each time a procedure is called. Because federated procedures return only the first result set, a different result set might be returned from the data source when the federated procedure is called.

For example, there are two procedures on the data source, PROCEDURE A and PROCEDURE B. PROCEDURE B calls PROCEDURE A. The statements to create these procedures are:

```
CREATE PROCEDURE A ()
BEGIN
   DECLARE cur1 CURSOR WITH RETURN TO CLIENT
   FOR SELECT * FROM t;
   OPEN cur1
END
```

```
CREATE PROCEDURE B (arg1 INT)
BEGIN
   DECLARE cur2 CURSOR WITH RETURN TO CLIENT
   FOR SELECT * FROM t;
   IF arg1<10) THEN
      CALL A();
   END IF;
   OPEN cur2
END;
```

The federated procedure FEDPROC1 references the data source PROCEDURE B. The statement for the FEDPROC1 procedure is:

```
CREATE PROCEDURE FEDPROC1
SOURCE newton.B
FOR SERVER s1
NUMBER OF PARAMETERS 1
WITH RETURN TO CLIENT 1;
```

A local procedure calls the federated procedure FEDPROC1. The statement for the local procedure is:

```
CREATE PROCEDURE local (arg1 INT)
   BEGIN
   CALL FEDPROC1 (arg1)
END;
```

When you issue the `CALL LOCAL(1)` statement, the cur1 result set from PROCEDURE A is returned. The result set cur2 is not returned.

However, if you issue the `CALL LOCAL(20)` statement, the cur2 result set from PROCEDURE B is returned.

Stored procedure nicknames from DataJoiner

DataJoiner stored procedure nicknames are not migrated to IBM WebSphere Federation Server. You must drop the stored procedure nicknames that you used in DataJoiner and create the procedures again using the CREATE PROCEDURE (Sourced) statement.

Federated procedures are called differently than stored procedure nicknames in DataJoiner. You might need to make updates to the applications or scripts that used the DataJoiner stored procedure nicknames.

Pass-through session (Oracle only)

If you create a data source procedure, function, or package in a pass-through session, a successful message is returned even if the object definition has an error. The object is created on the Oracle server but it is marked INVALID. You cannot create federated procedures on INVALID objects. When you attempt to create a federated procedure that references an INVALID Oracle object, the CREATE PROCEDURE (Sourced) statement fails.

Use one of the following methods to determine why an object is not valid:

- Use the SHOW ERRORS command in the SQL*Plus utility from Oracle.
- Query the Oracle sys.all_errors catalog table.

Chapter 7. Transparent DDL

What is transparent DDL

Transparent DDL provides the ability to create and modify remote tables through the federated database without using pass-through sessions.

The SQL statements you use with transparent DDL are CREATE TABLE, ALTER TABLE, and DROP TABLE.

A transparent DDL CREATE TABLE statement creates a remote table at the data source and a nickname for that table at the federated server. It will map the DB2 data types you specify to the remote data types using the default reverse type mappings. In general, the wrappers provide type mappings. You can also create user-defined reverse type mappings to override the default mappings.

The advantage of using transparent DDL is that database administrators can use procedures that they are familiar with to create both local and remote tables. Transparent DDL centralizes table administration and facilitates granting authorizations.

Transparent DDL is supported with the following data sources:

- DB2 for z/OS
- DB2 for iSeries
- DB2 Database for Linux, UNIX, and Windows
- DB2 Server for VM and VSE
- Informix
- Microsoft SQL Server
- ODBC
- Oracle
- Sybase
- Teradata

The database administrator can either use the DB2 Control Center or DDL statements in the DB2 command line processor (CLP) to create the tables. Using transparent DDL avoids the need to learn the different DDL syntax required for each data source.

Before you can create remote tables on a data source through the federated database, you need to configure access to the data source:

- The wrapper for that data source needs to be registered in the global catalog
- The server definition needs to be created for the server where the remote table will be located
- The user mappings, if required, need to be created between the federated server and the data source server

Use the remote table wizard in the DB2 Control Center to create remote tables.

The privileges held by the authorization ID of the transparent DDL statements must include at least one of the following:

- SYSADM or DBADM authority
- CREATETAB authority on the database and USE privilege on the table space as well as one of:
  - IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the table does not exist
  - CREATEIN privilege on the schema, if the schema name of the table refers to an existing schema

To issue transparent DDL statements, your authorization ID must have the necessary privileges on the nickname (for the federated server to accept the request), and the comparable privileges on the remote data source server (for the data source to accept the request).

Remote LOB columns and transparent DDL

You specify the length of a LOB column when using transparent DDL.

Some data sources, such as Oracle and Informix, do not store the lengths of LOB columns in their system catalogs. When you create a nickname on a table, information from the data source system catalog is retrieved including column length. Since no length exists for the LOB columns, the federated database assumes that the length is the maximum length of a LOB column in DB2 Database for Linux, UNIX, and Windows. The federated database stores the maximum length in the federated database catalog as the length of the nickname column.

However, when you create a remote table using transparent DDL you must specify the length of the LOB column. When the federated server creates a nickname on the remote table, it stores the length you specify in the federated database catalog as the length of the nickname column. The maximum length of a LOB column is 2 gigabytes.

Creating remote tables and transparent DDL

When a remote table is created through the federated database using transparent DDL, several other actions occur.

When you create the remote table:

- A nickname is automatically created for the remote table. The nickname has the same name as the table name specified in the CREATE TABLE statement. The remote table has the same name as the table name unless you specify another name using the REMOTE_TABNAME option.
- The schema of the remote table is the nickname schema unless you specify another schema using the REMOTE_SCHEMA option.
- The nickname created using transparent DDL can be used like any other nickname. In addition, you can ALTER and DROP the remote table (something you cannot do with a nickname created using CREATE NICKNAME).
- A row is added in the SYSCAT.TABOPTIONS catalog view with an option name of TRANSPARENT and a value of 'Y'.

Creating new remote tables using transparent DDL

To create a remote table using transparent DDL, you can use either the DB2 Control Center wizard or the CREATE TABLE statement.

Before you begin

Before you create a remote table, you must configure the federated server to access that data source. This configuration includes:

- Creating the wrapper for that data source type
- Supplying the server definition for the server where the remote table will be located
- Creating the user mappings between the federated server and the data source server

To issue transparent DDL statements, your authorization ID must have the necessary privileges on the nickname (for the federated server to accept the request), and the comparable privileges on the remote data source server (for the data source to accept the request).

The privileges held by the authorization ID issuing the transparent DDL statements must include at least one of the following:

- SYSADM or DBADM authority
- CREATETAB authority on the database and USE privilege on the table space as well as one of:
  - IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the table does not exist
  - CREATEIN privilege on the schema, if the schema name of the table refers to an existing schema

Restrictions

The following restrictions apply to creating a remote table using transparent DDL:

- You cannot modify or drop tables that were natively created at the remote data source.
- Materialized query tables cannot be created on remote data sources.
- You can specify basic column information in the table definition, but you will not be able to specify table options or column options. For example, the LOB options (LOGGED and COMPACT) are not supported.
- You cannot specify a comment on a column.
- You cannot generate column contents.
- You can specify a primary key, but you cannot specify a foreign key or check constraints. The columns used for a primary key must be NOT NULL, and cannot include columns containing LOBs.
- You cannot modify the parameters of existing columns, such as the data type or length.
- The DEFAULT clause in the CREATE TABLE statement is not supported.

About this task

Use the remote table wizard in the DB2 Control Center to avoid specifying a parameter or option that is not supported. Through the wizard you can specify columns by selecting from a list of predefined columns, or by specifying the attributes for a new column.

Procedure

To create a remote table from the command line prompt, issue the CREATE TABLE statement with the appropriate parameters set.

To create a remote table in the DB2 Control Center, use the Create Remote Table wizard:

1. Expand the **Federated Database Objects** folder.
2. Expand the wrapper and server definition objects for the data source that you want to create a remote table for.
3. Right-click the **Remote Tables** folder and click **Create**. The Create Remote Table wizard starts.
4. Complete the steps in the wizard.

Creating new remote tables using transparent DDL - examples

The following examples illustrate what to specify to create remote tables using transparent DDL and the use of data type mappings.

When you create remote tables using transparent DDL:

- The remote data source must support the column data types and primary key option in the CREATE TABLE statement.

  Example: The remote data source does not support primary keys. Depending on how the data source responds to requests it does not support, an error might be returned or the request might be ignored.
- The remote server must be specified in the OPTIONS clause. The OPTIONS clause can be used to override the remote name or the remote schema of the table being created. The SQL_SUFFIX option is allowed at the end of the CREATE TABLE statement. You can specify this option for any relational data source to add data source-specific options to the CREATE TABLE statement that is issued at the data source.

Example: You want to create the table EMPLOY on an Oracle server. In the CREATE TABLE statement, use the DB2 data types when you specify each column. Using the CLP, the syntax to create the table is:

```
CREATE TABLE EMPLOY
   ( EMP_NO       CHAR(6) NOT NULL,
     FIRSTNAME    VARCHAR(12) NOT NULL,
     MIDINT       CHAR(1) NOT NULL,
     LASTNAME     VARCHAR(15) NOT NULL,
     HIREDATE     DATE,
     JOB          CHAR(8),
     SALARY       DECIMAL(9,2),
     PRIMARY KEY (EMP_NO) )
     OPTIONS (REMOTE_SERVER 'ORASERVER',
     REMOTE_SCHEMA 'J15USER1', REMOTE_TABNAME 'EMPLOY' )
```

EMPLOY
: The name of the nickname associated with the table.

REMOTE_SERVER *'ORASERVER'*
: The name that you supplied for the server in the CREATE SERVER statement. This value is case-sensitive.

REMOTE_SCHEMA *'J15USER1'*
: The remote schema name. Although this parameter is optional, it is recommended that you specify a schema name. If this parameter is not specified, the nickname schema is used for the remote schema name. This value is case-sensitive.

REMOTE_TABNAME *'EMPLOY'*
: The remote table name. This parameter is optional. If this parameter is not specified, the local table name is used for the remote table name. This

value must be a valid name on the remote data source and cannot be an existing table name. This value is case-sensitive.

In the example above, the federated database uses reverse data type mappings to map the DB2 data types to Oracle data types. On the remote Oracle server, the EMPLOY table is created using Oracle data types. The following table shows the mappings from the DB2 data types to the Oracle data types for the columns specified in the example.

Table 7. An example of reverse data type mappings from the federated database to Oracle

Column	DB2 data type specified in the CREATE TABLE statement	Oracle data type used in the remote table
EMP_NO	CHAR(6) NOT NULL	CHAR(6) NOT NULL
FIRST_NAME	VARCHAR(12) NOT NULL	VARCHAR2(12) NOT NULL
MID_INT	CHAR(1) NOT NULL	CHAR(1) NOT NULL
LAST_NAME	VARCHAR(15) NOT NULL	VARCHAR2(15) NOT NULL
HIRE_DATE	DATE	DATE
JOB	CHAR(8)	CHAR(8)
SALARY	DECIMAL(9,2)	NUMBER(9,2)

Altering remote tables using transparent DDL

You can alter remote data source tables that were created through the federated database using transparent DDL. You cannot alter tables that were created directly at the remote data source.

Before you begin

The privileges held by the authorization ID of the transparent DDL statements must include at least one of the following:

- SYSADM or DBADM authority
- CREATETAB authority on the database and USE privilege on the table space as well as one of:
  - IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the table does not exist
  - CREATEIN privilege on the schema, if the schema name of the table refers to an existing schema

To issue transparent DDL statements, your authorization ID must have the necessary privileges on the nickname (for the federated server to accept the request), and the comparable privileges on the remote data source server (for the data source to accept the request).

Restrictions

The following restrictions apply to altering a remote table using transparent DDL:

- You cannot modify tables that were natively created at the remote data source.
- An existing primary key cannot be altered or dropped in a remote table.
- Altering a remote table invalidates any packages dependent on the nickname associated with the remote table.

- The remote data source must support the changes in the ALTER TABLE statement. For example, suppose that the remote data source does not support primary keys. Depending on how the data source responds to requests it does not support, an error might be returned or the request might be ignored.
- You cannot specify a comment on a column.
- You cannot generate column contents.
- You can specify a primary key, but you cannot specify a foreign key or check constraints. The columns used for a primary key must be NOT NULL, and cannot include columns containing LOBs.
- You cannot modify the parameters of existing columns, such as the data type or length.
- The DEFAULT clause in the ALTER TABLE statement is not supported.

About this task

You can use either the DB2 Control Center or the ALTER TABLE statement to modify tables created through WebSphere Federation Server using transparent DDL. Use the DB2 Control Center to avoid specifying a parameter or option that is not supported. Using the ALTER TABLE statement you can:

- Add new columns
- Add the table primary key

Do not use the ALTER TABLE statement to add or modify column options. Use the ALTER NICKNAME statement instead.

Procedure

To alter a remote table using transparent DDL, issue the ALTER TABLE statement:

Example: You want to add a primary key on a remote table EMPLOYEE that you created using transparent DDL. Using the following ALTER TABLE statement to modify the table:

```
ALTER TABLE EMPLOYEE
   ADD PRIMARY KEY (EMP_NO, WORK_DEPT)
```

The columns used for a primary key must be NOT NULL, and cannot be columns that contain LOBs.

Example: You want to add the columns ORDER_DATE and SHIP_DATE to the remote table SPALTEN that was created using transparent DDL. Using the following ALTER TABLE statement to create the table:

```
ALTER TABLE SPALTEN
   ADD COLUMN ORDER_DATE  DATE
   ADD COLUMN SHIP_DATE  DATE
```

Dropping remote tables using transparent DDL

You can drop remote data source tables that were created through the federated database using transparent DDL. You cannot drop tables that were created directly at the remote data source.

Before you begin

The privileges held by the authorization ID of the transparent DDL statements must include at least one of the following:

- SYSADM or DBADM authority
- CREATETAB authority on the database and USE privilege on the table space as well as one of:
  - IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the table does not exist
  - CREATEIN privilege on the schema, if the schema name of the table refers to an existing schema

To issue transparent DDL statements, your authorization ID must have the necessary privileges on the nickname (for the federated server to accept the request), and the comparable privileges on the remote data source server (for the data source to accept the request).

Restrictions

You cannot drop tables that were natively created at the remote data source.

About this task

To drop a remote table that was created through the federated database using transparent DDL, you can use either the DB2 Control Center or the DROP statement.

Dropping a nickname for a remote table created using transparent DDL merely drops the local nickname for that table. The DROP NICKNAME statement does not drop the remote table. You must use the DROP TABLE statement to drop the remote table.

Dropping a remote table first deletes the table on the data source, then deletes the corresponding nickname for the remote table in the federated database. Deleting the nickname invalidates any packages based on that nickname.

Procedure

To drop a remote table, issue the DROP TABLE statement.

Example: To drop a table named SPALTEN, issue the following DROP statement:

```
DROP TABLE SPALTEN
```

where *SPALTEN* is the local name for the remote table.

Chapter 8. Transaction support in a federated system

Understanding federated system transaction support

Knowledge of transaction processing concepts in a DB2 Database for Linux, UNIX, and Windows distributed environment will help you understand federated system transactions.

To understand federated system transaction processing, you should be familiar with the following distributed transaction processing concepts:

- Unit of work (UOW)
- Remote unit of work (RUOW)
- Distributed unit of work (DUOW)
- Multisite update
- Transaction manager (TM)
- Resource manager (RM)
- Type 1 connection
- Type 2 connection
- One-phase commit
- Two-phase commit

These concepts work identically in both federated and non-federated database systems. However, the scope of each concept changes in a federated system.

For example, a unit of work implicitly begins when any data in a database is read or written. For a unit of work in a federated system, the database can be a federated database or a data source database. For a distributed unit of work in a federated system, you can access both a federated database and a data source database.

An application must end a unit of work by issuing either a COMMIT or a ROLLBACK statement, regardless of the number of databases that are accessed. The COMMIT statement makes all changes within a unit of work permanent. The ROLLBACK statement removes these changes from a database. Changes made by a unit of work become visible to other applications after a successful commit.

Recommendation: Always explicitly commit or roll back units of work in your applications.

In a distributed unit of work that involves updates of multiple databases on multiple sites, data must be consistent. The multisite update or two-phase commit protocol is commonly used to ensure data consistency across multiple databases within a distributed unit of work.

Federated transactions support both one-phase commit protocol and two-phase commit protocol. The DB2_TWO_PHASE_COMMIT server option enables two-phase commit support for the following data sources:

- DB2 family data sources
- Informix
- Oracle

- Sybase
- MS SQL Server

When a data source is declared as a federated two-phase commit data source, that is, the DB2_TWO_PHASE_COMMIT server option is set to "Y", a commit against this data source uses two-phase commit protocol, even if it is a single site update transaction or a multi site update transaction.

When a data source is declared as a federated one-phase commit data source (the default), and it is a single site update transaction, a commit against this data source uses one-phase commit protocol.

In the following example of a one-phase commit operation, Oracle is defined as a one-phase commit data source:

```
SELECT * FROM oracle_nickname
UPDATE oracle_nickname
COMMIT
```

In the following example of a two-phase commit operation, Oracle and DRDA are defined as two-phase commit data sources:

```
SELECT * FROM oracle_nickname
UPDATE oracle_nickname

SELECT * FROM drda_nickname
UPDATE drda_nickname
COMMIT
```

What is an update in a federated system?

In a federated system, an update is not just a transaction that includes an INSERT, UPDATE, or DELETE statement. There are certain operations that are considered updates in a federated system and certain types of updates that are allowed in federated system.

In a federated system, updates can be performed locally or remotely.

- Local site updates are updates to DB2 tables or views that do not reference nicknames
- Remote site updates are updates to objects on a remote data source. Remote data sources include:
  - Another DB2 Database for Linux, UNIX, and Windows database or instance on the federated server
  - Another DB2 Database for Linux, UNIX, and Windows database or instance on another server
  - Data sources other than DB2 Database for Linux, UNIX, and Windows, such as DB2 for iSeries, Informix, Oracle, and Teradata

There are four types of actions that the federated server considers to be update transactions. The following table shows the updates that you can perform on a federated system.

Table 8. Types of updates and the site where the updates are performed

Type of action	Local site	Remote site	Explanation
Local update (DDL and DML)	Y	N	An update on an object in the federated database.

Table 8. Types of updates and the site where the updates are performed (continued)

Type of action	Local site	Remote site	Explanation
Remote update (nickname)	N	Y	An update on a remote data source object that you created a nickname for.
Dynamic SQL in pass-through sessions	N	Y	An update on a remote data source object. You cannot use a pass-through session to update local objects. Even SELECT queries sent in pass-through sessions are considered to be an update action.
Transparent DDL	Y	Y	A pair of transactions that create, alter, or drop remote tables and their corresponding nicknames in a federated database. For example, a pair of transactions that create a remote table on a data source and a nickname on the federated server.

What is an update transaction in a pass-through session?

A federated server treats all dynamic SQL statements sent through pass-through sessions as updates. This behavior ensures data integrity.

If a dynamic SQL statement that is sent through a pass-through session is successful, the transaction is recorded as an update. The SQL can be any type of statement, including SELECT statements.

Data sources that automatically commit DDL statements

Some data sources, such as Oracle, automatically commit the current transaction at their data source sites as part of a DDL statement execution.

If you create a remote table using transparent DDL or in a pass-through session, these data sources cannot rollback the remote table after the table is created. You must delete the remote table manually.

User-defined functions that are pushed down to the data source for processing

If a remote user-defined function performs an update on a data source, the federated server is unaware of the update.

Because the federated server does not treat these user-defined functions as update statements, all the statement level protection that the federated system applies to the update operations is not applicable. As a result, data integrity might be compromised in some situations.

Important: Data integrity cannot be guaranteed when a user-defined function that is pushed down to a data source performs an update.

Two-phase commit for federated transactions

A federated system can use two-phase commit for transactions that access one or more data sources. Two-phase commit uses the industry standard X/Open XA protocol to coordinate the processing of distributed unit of work transactions.

In a two-phase commit operation, commit processing occurs in two phases: the prepare phase and the commit phase. During the prepare phase in a federated system, a federated server polls all of the federated two-phase commit data sources that are involved in a transaction. This polling activity verifies whether each data source is ready to commit or roll back the data. During the commit phase, the federated server instructs each two-phase commit data source to either commit the data or to roll back the transaction.

In a one-phase commit environment, multiple data sources are updated one data source at a time using separate commit operations. This can cause data synchronization problems if some data sources are successfully updated and others are not.

For example, if a transaction withdraws funds from one account and deposits them in another account using one-phase commit, the system might successfully commit the withdraw operation and unsuccessfully commit the deposit operation. The deposit operation can be rolled back, but the withdraw operation cannot because it has already been successfully committed. The result is that the funds are virtually ″lost″.

In a two-phase commit environment, the withdraw and deposit transactions are prepared together and either committed or rolled back together. The result is that the integrity of the fund amounts remains intact.

Related concepts

Multisite Updates

X/Open distributed transaction processing model

“Data source requirements and configuration for federated two-phase commit transactions” on page 116
Before enabling federated two-phase commit for a data source, you must ensure it is a supported data source.

“Recovering from federated two-phase commit problems” on page 123
A federated system can recover from problems during two-phase commit with automatic resynchronization or manual recovery of indoubt transactions.

“Federated two-phase commit performance” on page 127
Data sources that are configured for two-phase commit transactions incur a performance penalty when you compare them to data sources that are configured for one-phase commit transactions.

Related tasks

“Enabling two-phase commit for federated transactions” on page 114
To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

Planning for federated two-phase commit

Federated two-phase commit does not provide benefits in all business environments. Also, there are several factors to consider before you decide to deploy federated two-phase commit.

To use two-phase commit for federated transactions, you must consider the following issues:

- Whether your operating system and data source environment can support two-phase commit for federated transactions.
- Whether your business environment requires two-phase commit for federated transactions.

  To decide whether or not two-phase commit is right for your business environment, you need to understand how two-phase commit for federated transactions works and the problems that it solves.
- How to configure the federated server and compatible data sources to use two-phase commit for federated transactions.

  There are basic federated server and data source requirements to use two-phase commit for federated transactions as well as performance considerations you must consider when deploying two-phase commit.
- To manually resolve indoubt transactions, you need to understand the inner workings of two-phase commit for federated transactions.

  Two-phase commit for federated transactions can resolve problems without intervention. But when there are extended network outages, hardware failures, or an urgent need to free up system resources, you can manually resolve problems through *heuristic processing*.

  Related concepts

  "Manually recovering indoubt transactions" on page 124
  If you cannot wait for resynchronization to automatically resolve indoubt transactions, you can resolve the indoubt transactions manually. This process is sometimes referred to as heuristic processing.

  "Data source requirements and configuration for federated two-phase commit transactions" on page 116
  Before enabling federated two-phase commit for a data source, you must ensure it is a supported data source.

  Related tasks

  "Enabling two-phase commit for federated transactions" on page 114
  To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

  Resolving indoubt transactions manually

Federated architecture for two-phase commit

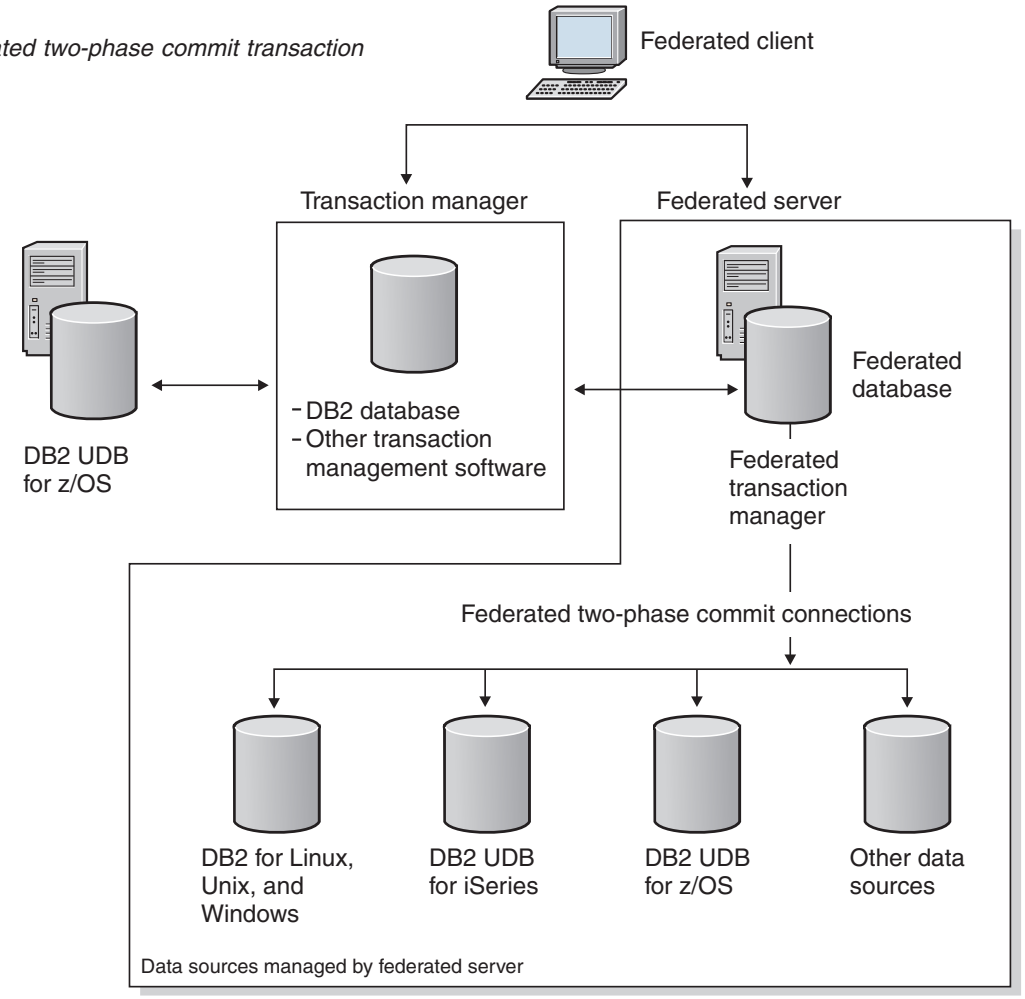
Federated two-phase commit is based on the two-phase commit feature available in DB2. In two-phase commit, the X/Open Distributed Transaction Processing (DTP) model has multiple components: transaction identifiers, transaction managers, and resource managers. In federated systems using federated two-phase commit, another component is added, the federated transaction manager.

A federated server becomes a federated transaction manager if the server coordinates activity for one or more remote data sources that use the two-phase commit protocol. A federated transaction manager performs some transaction management functions on behalf of the transaction manager. The client or application that initiates a distributed unit of work transaction and the transaction manager are unaware of the activity that the federated transaction manager coordinates at the remote data sources. The federated transaction manager communicates with DB2 Universal Database transaction managers using an XA interface. In addition to any X/Open requirements for two-phase commit, the

transaction manager database must be accessible from the federated instance. Resource managers follow the instructions that are provided by a federated transaction manager to commit or roll back a transaction.

The following figure shows an example of a simple two-phase commit transaction in a typical federated system, from client initiation to the data source updates.

Figure 3. Simple federated two-phase commit transaction



In the previous figure, the connection from the client to the transaction manager is a type 2 connection. Each database connection also has its own sync point setting. A sync point is a point in time when all the recoverable data that a program accesses is consistent. Sync point two-phase connections support distributed unit of work transactions with updates to multiple data sources.

When the client connects to the DB2 database, the transaction manager is aware of the transaction, but no additional coordination is required from the federated server. When the federated server connects to the data sources by using the two-phase commit protocol, the federated server becomes the federated transaction manager. The federated server monitors and coordinates the two-phase commits. At this point, the transaction manager is unaware of the two-phase commit transactions with the data sources. The transaction manager only knows that a single transaction is being processed with the federated server.

Data sources are not capable of initiating resynchronization if a failure occurs in a federated system. The federated server initiates the resynchronization process.

Results can be unpredictable if you attempt to access a data source by using multiple paths in the same transaction with federated two-phase commit. For example, if the federated server is a resource manager to an external transaction manager, the data source might be accessed indirectly from the federated server and directly as a resource manager to the transaction manager. In this case, the data source might not be able to tell whether these two paths are from the same global transaction. The data source might create two transaction entries for the same global transaction and treat each transaction as separate from the other, possibly leading to unpredictable results. The data source might also detect that the two paths are from the same global transaction, and reject the second path.

Related concepts

X/Open distributed transaction processing model

DB2 transaction manager

Two-phase commit for federated transactions - examples

A federated system that uses two-phase commit can be configured in several different ways. Configuration choices depend on the required solution.

Configurations can use Type 1 or Type 2 connections.

Type 1 connections are connections in which an application process is connected to an application server according to the rules for remote unit of work.

Type 2 connections are connections in which an application process is connected to an application server and establishes the rules for application-directed distributed unit of work. The application server is then the current server for the process.

The following figure shows a DB2 Type 1 connection with a federated server functioning as the transaction manager.

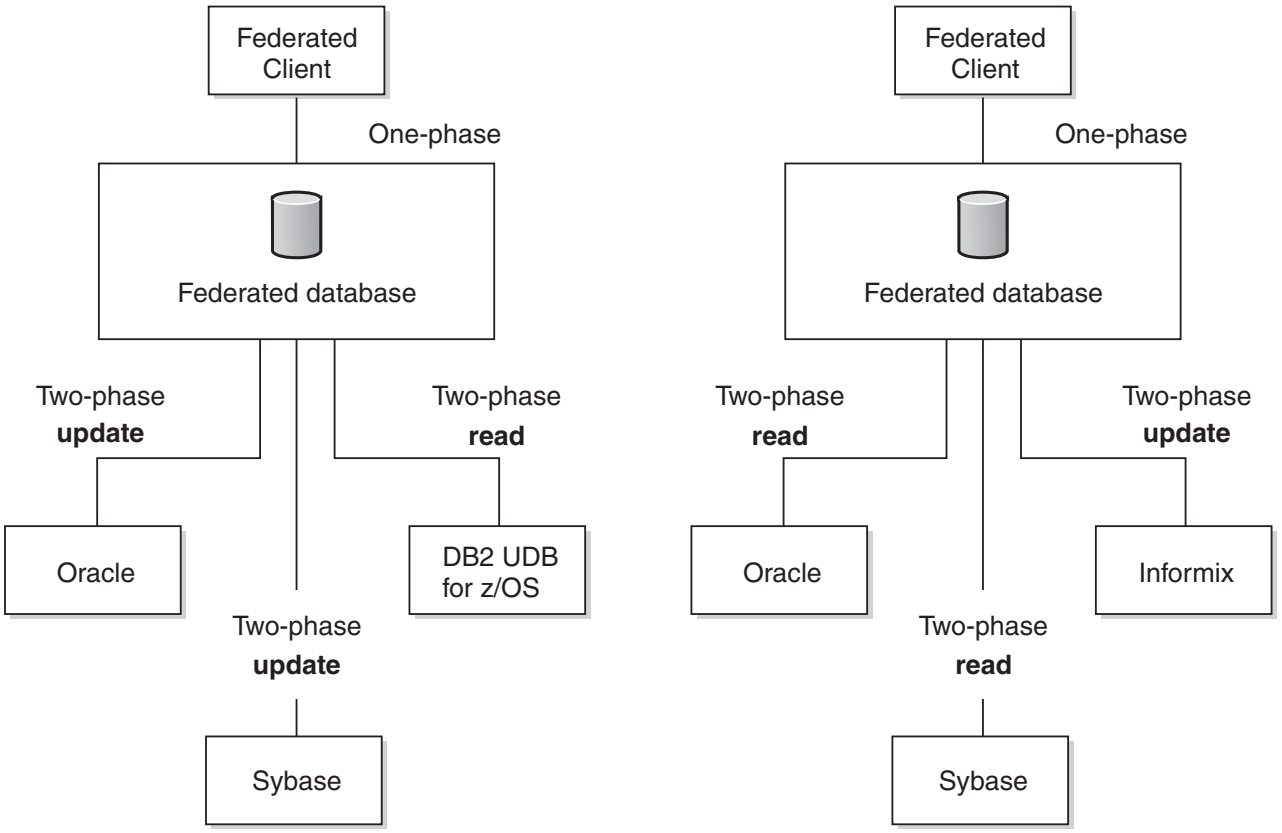


Figure 4. DB2 Type 1 connection with a federated server as the transaction manager

The following figure shows a DB2 Type 2 connection with a federated server functioning as the resource manager. In this configuration, all federated data sources must be supported for federated two-phase commit and enabled for federated two-phase commit.

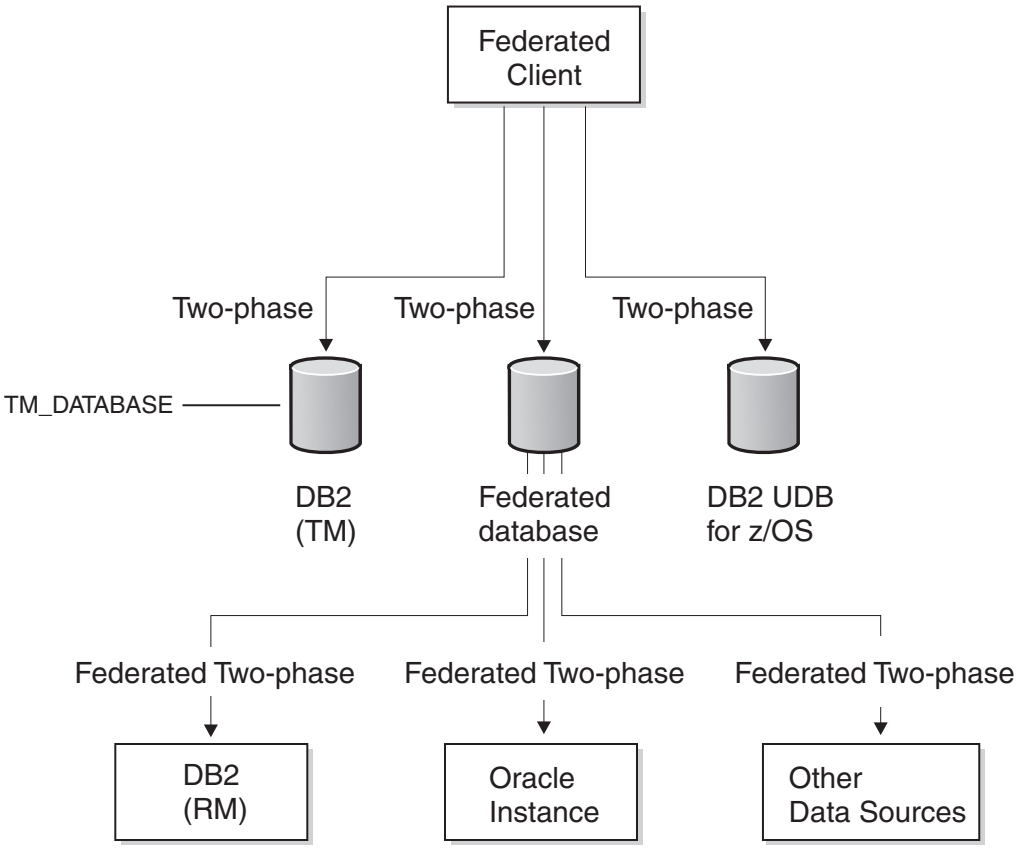


Figure 5. DB2 Type 2 connection with federated server as the resource manager

The following figure shows a DB2 Type 2 connection with a federated server functioning as the transaction manager.

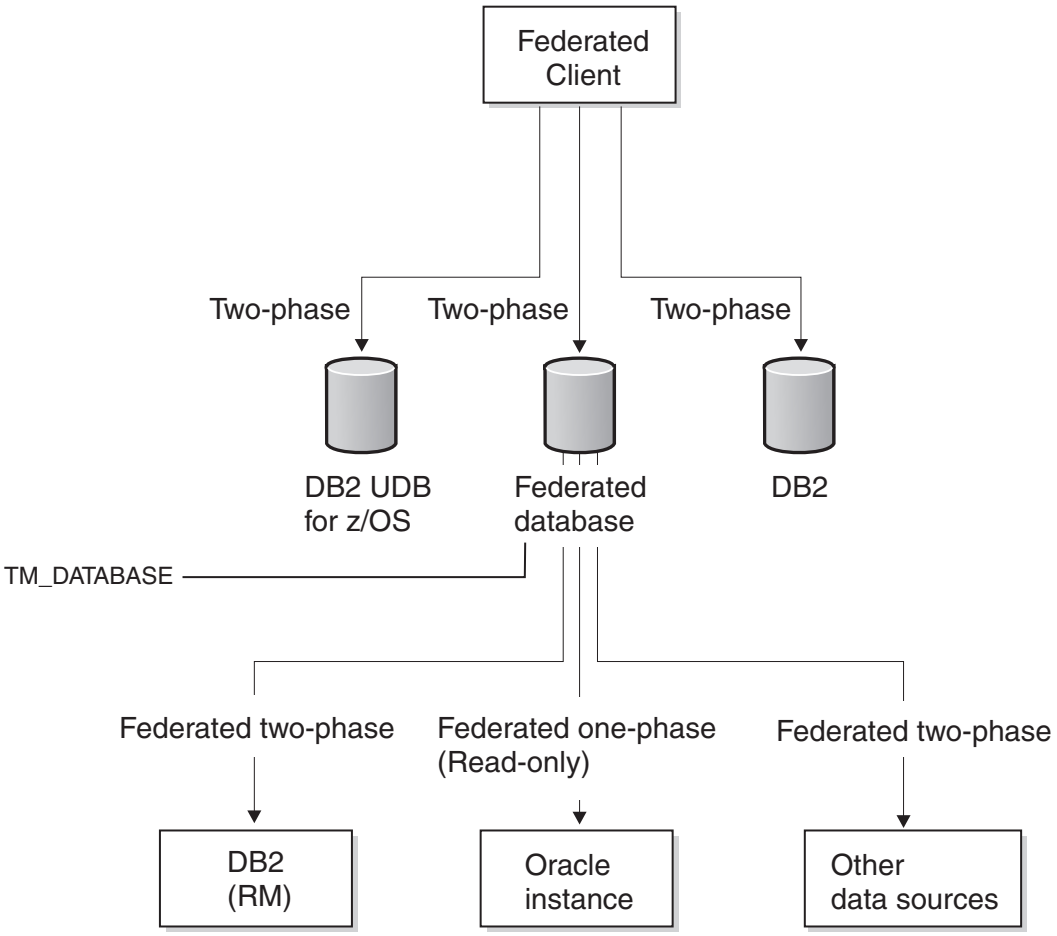


Figure 6. DB2 Type 2 connection with a federated server as the transaction manager

The following figure shows an XA connection with a federated server functioning as the resource manager.

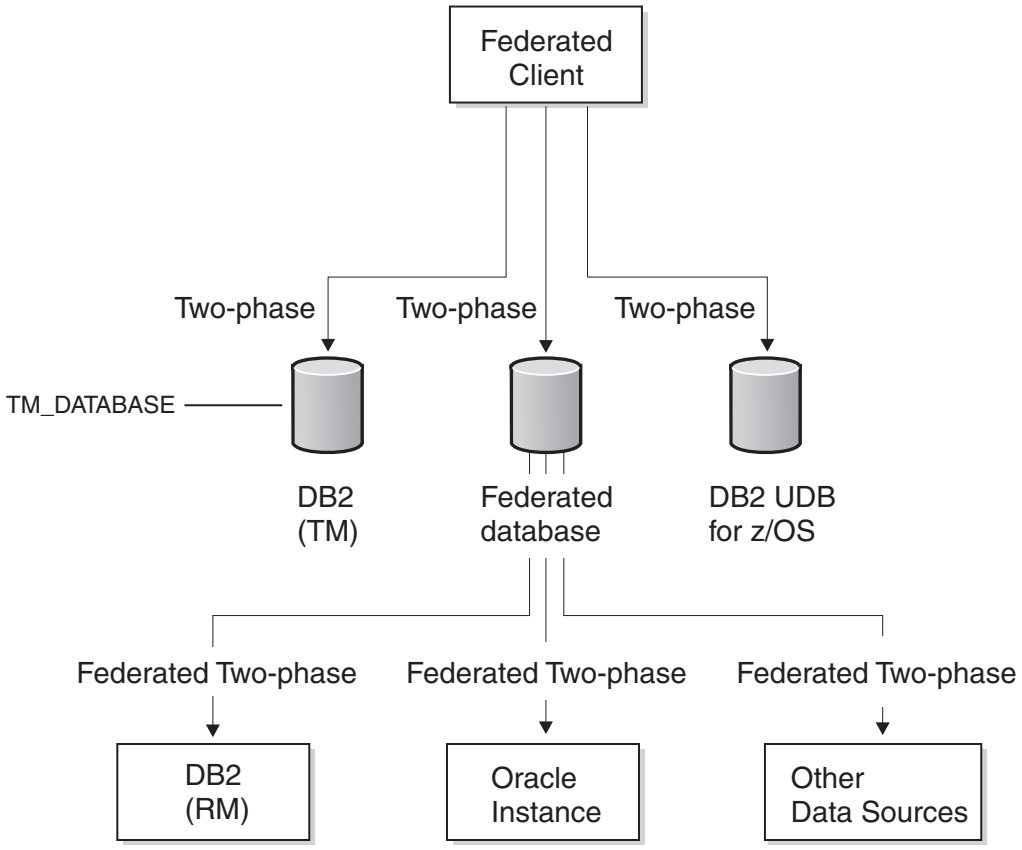


Figure 7. XA connection with federated server as the resource manager

Related reference

CONNECT (Type 1) statement

CONNECT (Type 2) statement

How federated two-phase commit transactions are processed

The federated server maintains data consistency and atomicity of the data sources that it manages. The range of possible transactions depends on the type of connection and whether the federated server is the transaction manager or resource manager for the connection.

Atomicity is a database principal in which sets of operations are defined within indivisible transactions. This principal ensures that the database is consistent at all times because if a single operation within the indivisible transaction fails, the whole transaction fails rather than compromising data integrity due to a partial change.

For example, a transaction to transfer funds from one account to another involves withdrawing funds from the first account and adding funds to the second account. If only the withdrawal succeeds, the funds essentially cease to exist in the first account.

The federated server processes federated update requests under strict rules. A federated update is one of the following actions:

- A federated insert, update, or delete operation where the corresponding data source supports insert, update, or delete operations. For example, some data

sources do not support update operations. Some data sources are read-only, in which case federated insert, update, or delete operations are not allowed.

- A successful pass-through operation inside a pass-through session.
- A transparent DDL operation, which is considered to be both a local update and a federated update because it performs database update both locally and remotely.
- A federated stored procedure with MODIFY SQL ACCESS.

Note: Do not use multiple paths to access the same data source within in a single transaction. Such a transaction can become deadlocked. That is, the transaction can hang. For example, do not use multiple federated servers to refer to the same data sources in the same transaction.

The following table lists what happens in a single distributed unit of work transaction including the type of connection, the type of commit, the federated server role in the transaction, what operations are allowed, and how transparent DDL can be used.

Table 9. What happens in a single distributed unit of work transaction

Type of commit	Type of connection	Federated server role	Operations	Transparent DDL
One-phase	DB2 Type 1 or XA local transaction	Sub-transaction manager. Also acts as the transaction coordinator, determines the transaction outcome, and delivers it to each participating resource manager.	One-phase commit and two-phase commit read operations are permitted. One one-phase data source can be updated as long as it is the only update in the transaction.	Allowed and managed according to one-phase commit data source rules. Each statement that is issued must be the only update in a one-phase commit transaction. Cannot coexist with other federated two-phase commit data source updates in the same transaction. It is strongly recommended that COMMIT or ROLLBACK statements be issued before and after transparent DDL transactions occur.
Two-phase	DB2 Type 1 or XA local transaction	Transaction manager. Also acts as the transaction coordinator, determines the transaction outcome, and delivers it to each participating resource manager.	One-phase commit and two-phase commit read operations are permitted. Multiple two-phase data sources can be updated.	Allowed and managed according to two-phase commit data source rules. Can coexist with other federated two-phase or one-phase commit data source updates in the same transaction.

Table 9. What happens in a single distributed unit of work transaction (continued)

Type of commit	Type of connection	Federated server role	Operations	Transparent DDL
One-phase	DB2 Type 2 or XA global transaction	Can be the transaction manager. If not the transaction manager, only relays the outcome from the external transaction coordinator to each participating resource manager.	One-phase commit and two-phase commit read operations are permitted. One-phase updates are not allowed except for DB2-coordinated transactions which can perform federated one-phase updates over a Distributed Relational Database Architecture™ (DRDA) two-phase inbound connection.	Allowed and managed according to one-phase commit data source rules. Each statement that is issued must be the only update in a one-phase commit transaction. Cannot coexist with other federated two-phase commit data source updates in the same transaction. It is strongly recommended that COMMIT or ROLLBACK statements be issued before and after transparent DDL transactions occur.
Two-phase	DB2 Type 2 or XA global transaction	Can be the transaction manager. If not the transaction manager, only relays the outcome from the external transaction coordinator to each participating resource manager.	One-phase commit and two-phase commit read operations are permitted. Multiple two-phase data sources can be updated.	Allowed and managed according to two-phase commit data source rules. Can coexist with other federated two-phase or one-phase commit data source updates in the same transaction.

How data consistency and atomicity are maintained

Federated servers attempt to ensure data consistency and maintain transaction atomicity of data sources.

Any conflict between an application synchronization point setting and the update capability of a target data source results in an error (SQL30090, reason code 18).

Local updates that include DDL made to the federated database cannot be mixed within the same transaction as an update to a federated one-phase data source. Transparent DDL

Using DDL and transparent DDL

Local updates that include DDL made to the federated database cannot be mixed within the same transaction as an update to a federated one-phase data source. Transparent DDL is an exception. For transparent DDL, both local updates and data source updates are allowed regardless of the type of connection and whether the data source is configured for one-phase or two-phase commit.

Transparent DDL creates a table on a remote data source and a nickname in the local federated database for the remote table. A federated server treats transparent DDL transactions as updates.

Transparent DDL provides the ability to create and modify remote tables through the DB2 database system, without the need to use pass-through sessions. The SQL statements for transparent DDL are CREATE TABLE, ALTER TABLE, and DROP TABLE. For example, a transparent DDL CREATE TABLE statement creates a remote table at the data source and a nickname for that table at the federated server. The statement contains a local update operation and a remote update operation.

Some data sources, such as Oracle, do not permit transparent DDL on a federated two-phase commit connection.

Related tasks

Dropping a table

Related reference

CONNECT (Type 1) statement

CONNECT (Type 2) statement

CREATE TABLE statement

ALTER TABLE statement

Enabling two-phase commit for federated transactions

To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

Before you begin

- When you enable federated two-phase commit for a data source, you increase the number of records that are written to both the federated server database log and the data source database log. Consider the impact this has on the administration and maintenance of these log files in order to ensure that they comply with your local policies.
- The data source that you want to use must be a supported federated two-phase commit data source.
- For DB2 Universal Database for iSeries,version 5.3, and earlier and DB2 Universal Database for z/OS data sources, ensure that the configuration parameter SPM_NAME is set to the default value, the server host name. SPM_NAME defaults to a variant of the first seven characters of the TCP/IP host name. DB2 Universal Database for iSeries, version 5.4, and later does not require that you set SPM_NAME.

About this task

The DB2_TWO_PHASE_COMMIT server option enables two-phase commit for data sources. You register a data source server definition by using the CREATE SERVER statement. The value you set for DB2_TWO_PHASE_COMMIT persists for all connections that are established under that server definition. You can change the value at any time by using the ALTER SERVER statement. After the CREATE SERVER or ALTER SERVER statement is successfully committed, the new setting is available for use on subsequent outbound connection requests.

Clients and application programs can use the SET SERVER OPTION to temporarily override the current value of the DB2_TWO_PHASE_COMMIT server option. The SET SERVER OPTION statement must be run immediately after the connection to the federated server database and before any connections are established to the remote data sources. The command is in effect only for the duration of the connection to the federated database. You cannot change the DB2_TWO_PHASE_COMMIT server option once the federated server has established a connection to the remote data source.

When you include the XA_OPEN_STRING_OPTIONS option in a CREATE SERVER statement, you can embed specialized information in the default XA_OPEN string. This embedded information can be any of the following kinds of information:

- Unique IDs for transactions in addition to what IBM WebSphere Federated Server provides
- User-defined parameters about how transactions are handled
- A user-defined string to append to the XA_OPEN request

When an XA_OPEN call is made, usually at the beginning of the first transaction to a remote data source that uses two-phase commit, the wrapper appends the value of the user-defined string onto the default XA_OPEN string for the XA_OPEN call.

You can include both DB2_TWO_PHASE_COMMIT and XA_OPEN_STRING_OPTIONS in a CREATE SERVER, SET SERVER, or ALTER SERVER statement.

Procedure

In general, to enable two-phase commit for a federated data source:

1. Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the DB2_TWO_PHASE_COMMIT option set to Y.
2. **Optional:** Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the XA_OPEN_STRING_OPTIONS option.

Server option examples

This example shows how to set two-phase commit by using the CREATE SERVER statement:

```
CREATE  SERVER  Net8_Server  TYPE  ORACLE VERSION  8.1.7 WRAPPER  NET8
OPTIONS (DB2_TWO_PHASE_COMMIT  'Y');
```

This example shows how to disable two-phase commit by using the ALTER SERVER statement:

```
ALTER  SERVER  Net8_Server OPTIONS  (SET  DB2_TWO_PHASE_COMMIT  'N');
```

This example shows how to set an XA trace file to D:\Temp\sybase_xa.log for the Sybase wrapper using the ALTER SERVER statement and the XA_OPEN_STRING_OPTIONS server option:

```
 ALTER SERVER Ctlib_Server OPTIONS (ADD XA_OPEN_STRING_OPTIONS
'-LD:\Temp\sybase_xa.log');
```

This example shows how to temporarily disable two-phase commit by using the SET SERVER OPTION statement:

```
SET  SERVER  OPTION  DB2_TWO_PHASE_COMMIT TO 'N'  FOR SERVER  Net8_Server;
```

Related concepts

"Two-phase commit for federated transactions" on page 104
A federated system can use two-phase commit for transactions that access one or more data sources. Two-phase commit uses the industry standard X/Open XA protocol to coordinate the processing of distributed unit of work transactions.

Related tasks

"Configuring DRDA data sources" on page 117
The federated server provides connectivity to DB2 data sources by using the open DRDA protocol. This support is equivalent to that provided by the DB2 Connect™ server.

"Configuring Oracle data sources" on page 118
There are several requirements and restrictions for using Oracle data sources for federated two-phase commit.

"Configuring Informix data sources" on page 119
There are several requirements and restrictions for using Informix data sources for federated two-phase commit.

"Configuring Microsoft SQL Server data sources" on page 120
There are several requirements and restrictions for using Microsoft SQL Server data sources for federated two-phase commit.

"Configuring Sybase data sources" on page 121
There are several requirements and restrictions for using Sybase data sources for federated two-phase commit.

Related reference

CREATE SERVER statement

ALTER SERVER statement

SET SERVER OPTION statement

RESET DATABASE MANAGER CONFIGURATION command

UPDATE DATABASE MANAGER CONFIGURATION command

resync_interval - Transaction resync interval configuration parameter

tm_database - Transaction manager database name configuration parameter

spm_name - Sync point manager name configuration parameter

svcename - TCP/IP service name configuration parameter

Data source requirements and configuration for federated two-phase commit transactions

Before enabling federated two-phase commit for a data source, you must ensure it is a supported data source.

Federated systems support two-phase commit operations with the following data sources:

- DB2 family data sources through the Distributed Relational Database Architecture (DRDA) protocol:
  - DB2 Universal Database for Linux, UNIX, and Windows, version 8.1 or later
  - DB2 Universal Database for z/OS, version 7.1 or later
  - DB2 Universal Database for iSeries, version 5.3 or later
- Informix IDS, version 7.31 or later, version 9.40 or later, version 10.0 or later
- Informix XPS, version 8.40 or later

- Microsoft SQL Server 2000 and Microsoft SQL Server 2005 for a federated server only on Windows
- Oracle, version 8.1.7 or later, with the XA library
- Sybase Adaptive Server Enterprise, version 12 or later, with the XA library for a federated server only on Windows

If you attempt to enable federated two-phase commit for an unsupported data source, you receive an SQL1881N error.

Related concepts

"Two-phase commit for federated transactions" on page 104
A federated system can use two-phase commit for transactions that access one or more data sources. Two-phase commit uses the industry standard X/Open XA protocol to coordinate the processing of distributed unit of work transactions.

Configuring DRDA data sources

The federated server provides connectivity to DB2 data sources by using the open DRDA protocol. This support is equivalent to that provided by the DB2 Connect server.

In addition, for two-phase commit, the federated server interacts with each data source using the industry-standard XA model.

Before you begin

Restrictions:

- Not all DB2 data sources support XA over DRDA natively. For those that do not, such as DB2 Universal Databasefor z/OS and DB2 Universal Database for iSeries, the federated server uses the syncpoint manager (SPM). The syncpoint manager performs a mapping between the XA and non-XA two-phase-commit flows that all DB2 servers support. When federated two-phase commit access is provided by using the syncpoint manager, not all XA semantics are supported because of incompatibilities between federated support and the syncpoint manager. For example, transactions cannot be nested. All transactions must be committed or rolled back before starting a new transaction.
- Federated two-phase commit supports Multiple Virtual Storage (MVS™), but MVS does not allow a SAVEPOINT statement to be issued in a federated two-phase commit transaction. You can control the rollback policy when a two-phase commit error occurs by using the IUD_APP_SVPT_ENFORCE server option.
- In a DB2-coordinated transaction, a DB2 Universal Database for z/OS client can perform a federated one-phase update over a Distributed Relational Database Architecture (DRDA) two-phase inbound connection. However, such an update cannot be completed over an XA DRDA two-phase inbound connection. Nor can a mix of one-phase and two-phase updates over a DRDA two-phase inbound be completed.
- The XA_OPEN_STRING_OPTIONS server option is not supported for DRDA data sources. If you use the option, an SQL1881 error is returned.

Requirements:

- For those DB2 data sources that support XA by using SPM rather than by supporting XA natively, ensure that the SPM_NAME and SVCENAME parameters in the database manager configuration are properly set to their defaults.

Procedure

To configure a DRDA data source:

Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the DB2_TWO_PHASE_COMMIT option set to Y.

The DRDA wrapper automatically generates the following XA OPEN string for DRDA data sources:

```
DB=dbname,UID=uid,PWD=password,TPM=FDB2,HOLD_CURSOR=T
```

Related concepts

Wrapper and wrapper modules

Related tasks

"Enabling two-phase commit for federated transactions" on page 114
To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

Related reference

spm_name - Sync point manager name configuration parameter

svcename - TCP/IP service name configuration parameter

SAVEPOINT statement

CREATE SERVER statement

ALTER SERVER statement

SET SERVER OPTION statement

Configuring Oracle data sources

There are several requirements and restrictions for using Oracle data sources for federated two-phase commit.

Before you begin

Restrictions:

- Pass-through DDL and transparent DDL directed to Oracle both fail with SQL30090 reason code 21 (ORA-2089) Regular SQL submitted in pass-through sessions works.

Requirements:

- The Oracle client used on the federated server should be a complete installation to ensure that all libraries relevant to XA are present. Ensure that djxlinkOracle was successfully run so that all DB2 and Oracle libraries are available and linked properly. Note that the djxlink scripts are run automatically if you install the Oracle client before you install the federated server.
- You must give the following privileges to all users that run two-phase commit transactions from the federated server:
  - grant select on dba_pending_transactions to USERID;
  - grant select on dba_2pc_pending to USERID;

  - grant force transaction to USERID;
- Optionally, you can also give the following privilege to users that run two-phase commit transactions from the federated server:
  - grant force any transaction to USERID;
- If you intend to run more than 10 two-phase commit transactions simultaneously, consider increasing the distributed_transactions parameter of your Oracle server found in the init.ora file.

Procedure

To configure an Oracle data source:

1. Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the DB2_TWO_PHASE_COMMIT option set to Y.

   The Oracle wrapper automatically creates the following XA OPEN string for Oracle data sources:

   ```
   Oracle_XA=Acc=P<i>uid/password</i>+SesTm=0+DB=<i>dbname</i>+SqlNet=<i>dblink</i>+Threads=true
   ```

   For example:

   ```
   XA_OPEN_STRING_OPTIONS '+LogDir=/home/user/directory+DbgFl=0x7'
   ```

2. Specify additional XA options by using the XA_OPEN_STRING_OPTIONS server option.

Related concepts

"Troubleshooting federated two-phase commit issues" on page 126
Troubleshooting federated two-phase commit issues is often specific to the data source that is causing the issue.

Related tasks

"Enabling two-phase commit for federated transactions" on page 114
To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

Configuring access to Oracle data sources

Related reference

CREATE SERVER statement

ALTER SERVER statement

SET SERVER OPTION statement

Configuring Informix data sources

There are several requirements and restrictions for using Informix data sources for federated two-phase commit.

Before you begin

Restrictions:

- You cannot access Informix nicknames with a mix of a two-phase commit server and a one-phase commit server in a single connection to a federated server.
- The WITH HOLD cursor option is not supported.
- The XA_OPEN_STRING_OPTIONS server option is not supported for Informix data sources.

Requirements:

- The Informix database must have logging enabled.

- The Informix XA library only allows one connection per thread. As a result, the federated server cannot access Informix data sources with multiple servers that are enabled for federated two-phase commit in a single connection. If an application needs to use multiple servers that are enabled for federated two-phase commit, complete the optional steps in the following procedure.

Procedure

To configure an Informix data source:

1. Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the DB2_TWO_PHASE_COMMIT option set to Y.

   The Informix wrapper generates the following XA OPEN string automatically for Informix data sources:

   ```
   DB=dbname;RM=rmname;CON=con;USER=user;PASSWD=password
   ```

2. If an application needs to use multiple servers that are enabled for federated two-phase commit, complete the following steps:
   a. Copy the Informix wrapper libraries: libdb2informix.a, libdb2informixF.a, and libdb2informixU.a.
   b. Define multiple instances of the Informix wrapper by specifying a different copy of the Informix wrapper libraries in the LIBRARY clause within the CREATE SERVER statement.
   c. Define each federated two-phase commit server for the different wrapper instances.

   For example:

   ```
   CREATE WRAPPER wrapper1 library 'libdb2informix.a'
   CREATE SERVER server1 type informix version 9.4 wrapper wrapper1 options
     (node 'inf1', dbname 'firstdb', db2_two_phase_commit 'Y');
   CREATE WRAPPER wrapper2 library 'libdb2informix2.a'
   CREATE SERVER server2 type informix version 9.4 wrapper wrapper2 options
     (node 'inf2', dbname 'seconddb', db2_two_phase_commit 'Y');
   ```

Related tasks

"Enabling two-phase commit for federated transactions" on page 114
To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

Configuring access to Informix data sources

Related reference

CREATE SERVER statement

ALTER SERVER statement

SET SERVER OPTION statement

DECLARE CURSOR statement

Configuring Microsoft SQL Server data sources

There are several requirements and restrictions for using Microsoft SQL Server data sources for federated two-phase commit.

Before you begin

Restrictions:

- For Microsoft SQL Server data sources, federated two-phase commit is only supported by IBM WebSphere Federated Server installed on Windows.
- The DB2 isolation level is not propagated to the Microsoft SQL server.

Requirements:

- For federated two-phase commit to work with Microsoft SQL, an extra server option, XA_OPEN_STRING_OPTIONS, must be added to the server:

  ```
  alter server S1 options(add xa_open_string_options
  'RMRecoveryGuid=c200e360-38c5-11ce-ae62-08002b2b79ef');
  ```

  where `RMRecoveryGuid` = resource manager ID.

  The resource manager ID is available in the following location of the Microsoft SQL Server Registry:

  ```
  [HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\MSSQLServer]
   "ResourceMgrID" = "{resource manager ID}"
  ```

Procedure

To configure a Microsoft SQL Server data source:

1. Run the CREATE SERVER, ALTER SERVER, or SET SERVER statement with the DB2_TWO_PHASE_COMMIT option set to Y.

   The Microsoft SQL Server wrapper automatically generates the following XA OPEN string for Microsoft SQL Server data sources:

   `TM=`*`tmname`*
2. Specify additional XA options in addition to the required RMRRecovery Guid value by using the XA_OPEN_STRING_OPTIONS server option.

   Related tasks

   "Enabling two-phase commit for federated transactions" on page 114
   To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

   Configuring access to Microsoft SQL Server data sources

   Related reference

   CREATE SERVER statement

   ALTER SERVER statement

   SET SERVER OPTION statement

Configuring Sybase data sources

There are several requirements and restrictions for using Sybase data sources for federated two-phase commit.

Before you begin

Restrictions:

- For Sybase data sources, federated two-phase commit is only supported by IBM WebSphere Federated Server installed on Windows.
- Pass-through DDL and transparent DDL directed to Sybase both fail with an SQL910N error. Regular SQL submitted in pass-through sessions works.

Requirements:

- The Sybase database administrator must have a license for distributed transaction management of Sybase Adaptive Server Enterprise (ASE) and must enable the feature with the following command in the isql tool:

  ```
  sp_configure 'enable dtm', 1
  ```

  Sybase ASE must be restarted for this parameter to take effect.

- The user name that is specified in an open string must have the dtm_tm_role in the corresponding Sybase ASE. The administrative user can assign this role with the following command in the isql tool:

  ```
  sp_role "grant", dtm_tm_role, user_name
  ```

- For a data source for Sybase Adaptive Server Enterprise (ASE) to act as a resource manager to a federated database, a logical resource manager (LRM) entry must exist in the xa_config file in the $SYBASE/$SYBASE_OCS/config directory (version 12 or later) that maps the resource manager name to the Sybase ASE name. Consult the Sybase ASE XA documentation for more information.

  The LRM name is used by the federated server in the XA OPEN string. The federated server uses the Sybase ASE node name for the LRM name.

- Ensure that the server name specified in the XA configuration file xa_config is present in the initialization file sql.ini in the $SYBASE/ini directory.

Procedure

To configure a Sybase data source:

1. Install the Sybase XA library file libxadtm.dll on the federated server.
2. Before you use federated two-phase commit functions, create the following LRM entries in the $SYBASE/$SYBASE_OCS/config/xa_config file. If you do not have write permissions for the xa_config file, create an xa_config file in another directory and set its absolute path in the XACONFIGFILE environment variable in the db2dj.ini file:

   ```
   ;one comment line is required
   lrm=lrm_name
   server=server_name
   ```

   where *server_name* is an entry name in the $SYBASE/ini/sql.ini file.
3. The Sybase wrapper automatically creates the following default XA_OPEN string for Sybase data sources:

   ```
   -Nrmname -Uuserid -Ppassword
   ```

   If you need to specify other options for the XA_OPEN string, use the XA_OPEN_STRING_OPTIONS server option.

   Related concepts

   "Troubleshooting federated two-phase commit issues" on page 126
   Troubleshooting federated two-phase commit issues is often specific to the data source that is causing the issue.

   Related tasks

   "Enabling two-phase commit for federated transactions" on page 114
   To use federated two-phase commit for specific data sources, you must enable the associated federated servers. The enablement process involves preparing the federated server and modifying the data source server definition.

   Configuring access to Sybase data sources

   Related reference

   CREATE SERVER statement

   ALTER SERVER statement

   SET SERVER OPTION statement

Recovering from federated two-phase commit problems

A federated system can recover from problems during two-phase commit with automatic resynchronization or manual recovery of indoubt transactions.

Related concepts

"Two-phase commit for federated transactions" on page 104
A federated system can use two-phase commit for transactions that access one or more data sources. Two-phase commit uses the industry standard X/Open XA protocol to coordinate the processing of distributed unit of work transactions.

Resynchronization for federated systems

Federated two-phase commit includes an automatic process that attempts to handle errors during commit transactions.

Besides errors on the federated server, a federated environment increases the potential for errors that result from network, communications, or data source failures.

To ensure data integrity, the federated server handles these errors during the federated two-phase commit process:

First phase error
If a database communicates that it has failed to prepare to commit a unit of work, the federated server rolls back the unit of work during the second phase of the commit process. During the second phase, the federated server sends a rollback message to all participating data sources that are waiting for the transaction outcome.

Second phase error
Error handling at this phase depends upon whether the second phase commits or rolls back the transaction. The second phase only rolls back the transaction if the first phase encountered an error.

If one of the participating data sources fails to commit or rollback the unit of work, possibly due to a communications failure, the federated server retries the commit or rollback through a process called resynchronization. Resynchronization is initiated and managed by the federated server automatically. The calling application is informed that the commit was successful through the SQLCA (SQL communications area) if the application connects to the federated server through a two-phase commit connection. The calling application is disconnected from the federated server if the application connects to the federated server through a one-phase commit connection.

Most data sources are not capable of initiating resynchronization if a failure occurs in a federated system. The federated server initiates the resynchronization process.

In certain circumstances, a federated server might fail during transaction processing, for example, due to a power failure. Resynchronization usually resolves any distributed unit of work transactions without intervention.

Resynchronization attempts to complete all indoubt transactions. As part of normal resynchronization, the resynchronization agent connects to the resource manager database for a transaction and issues a commit or rollback decision. The federated transaction manager then propagates that decision to the data sources that participated in the distributed unit of work transaction.

Related concepts

Error recovery during two-phase commit

Manually recovering indoubt transactions

If you cannot wait for resynchronization to automatically resolve indoubt transactions, you can resolve the indoubt transactions manually. This process is sometimes referred to as heuristic processing.

For example, the communications link between the external transaction manager and a federated transaction manager fails in the middle of a transaction. If you have enough information about the transaction, you can free resources on the federated server and on remote data sources by rolling back the transaction from the federated server.

Use heuristic processing only when you know the reason for the transaction failure, and you must free locked resources immediately. In most situations, let the automated resynchronization recover transactions. There are multiple layers of transaction management in a federated system. Recovering transactions heuristically is a complex and potentially risky process.

There are three basic ways to perform heuristic processing:

- The LIST INDOUBT TRANSACTIONS command

  You can use this command line command to perform heuristic processing.
- The Indoubt Transaction Manager window

  You can use this graphical user interface tool to perform heuristic processing.
- Heuristic APIs

  You can use these APIs within your applications to perform heuristic processing.

The specific operations and tasks that you use to perform heuristic processing varies, depending on the circumstances of the error.

In federated systems, when a heuristic processing request is sent to a federated transaction manager, the resulting decision to commit or roll back must be compatible with the actual status of the indoubt transaction on the federated server. Otherwise, an error message is returned.

The status of federated two-phase commit in doubt transactions is slightly different than basic DB2 two-phase commit in doubt transactions:

- A status of (d) means that the transaction is missing commit acknowledgement from one or more federated data sources.
- A status of (b) means that the transaction is missing rollback acknowledgement from one or more federated data sources.

If you cannot successfully commit or rollback a transaction with status (d) or (b), you can specify that the transactions be forgotten by using option (f). However, when you use the (f) option, all records of the transaction are erased from the federated server and you must manually clean up any remaining synchronization problems on the involved data sources. Only use the (f) option when it is absolutely necessary, such as when a remote server crashes or connections to remote servers are dropped and there is an urgent requirement to free up resources and use it with caution.

Note: Because the (d) and (b) statuses are new for WebSphere Federation Server v9.1, down-level federated clients cannot support them. If you use a down-level client to manually recover indoubt transactions, the (d) and (b) statuses are mapped to status (m), instead, which is not an accurate value. To prevent inaccurate information when you are manually recovering indoubt transactions, be sure to use a v9.1 federated client. By default, the computer that runs WebSphere Federated Server v9.1 includes the v9.1 federated client.

Related concepts

Indoubt transaction management APIs

Related tasks

Resolving indoubt transactions manually

Related reference

LIST INDOUBT TRANSACTIONS command

Tracing distributed unit of work transaction states across data sources

If you decide to resolve indoubt transactions manually instead of letting synchronization resolve them automatically, tracing transactions in the federated system is essential. When you are tracing an indoubt distributed unit of work transaction, the only way to determine the data source or data sources that failed is to capture the XID for the failed transaction.

You must look for that XID in the data source database managers for all of the data sources that a federated server might have accessed as a part of the distributed unit of work transaction.

To determine the identifier and state of each transaction that is involved in the distributed unit of work that you want to trace, issue the LIST INDOUBT TRANSACTIONS command in your application database, the federated database, and any data sources in the distributed unit of work transaction.

Note: Each data source that supports the two-phase commit protocol might use a different command. Search for the string XID in the command documentation for your specific data source.

The federated transaction manager generates an XID in hexadecimal format during transaction processing. This XID begins with the format identifier of F2PC, which is the number 46325243 in hexadecimal format. The federated transaction manager sends the XID to the data sources. Before a federated server sends an XID to a data source, however, the federated server changes the XID so that it conforms to the XID format of that data source. These modifications include updating the branch qualifier length section of the XID and adding the branch qualifier section of the XID.

You might need to compare XIDs across several data sources, so you must know which part of the XID that you can accurately compare across your environment when you trace an XID.

For example, the transaction manager is a DB2 database. When you issue the LIST INDOUBT TRANSACTIONS command at the database, a string in hexadecimal representation is returned for the transaction XID that is similar to the following one:

```
463250430000019 00000000473931453346313 5 2E47453934000000 000000000000E80000
```

This string is actually composed of several distinct parts:

Format ID	Transaction identifier length	Branch qualifier length	Transaction identifier
46325043	00000019	00000000	4739314533463135 2E47453934000000 000000000000E800 00

The values listed in the string are hexadecimal. For example, the transaction identifier length (hexadecimal 19) represents the decimal value 25.

If that same XID is passed from a federated server that acts as a federated transaction manager to a data source that acts as a resource manager, the XID string is appended. For example, the XID that is passed to a resource manager for a DB2 database system for Windows data source changes to the following format:

```
463252430000019 00000001473931453346313 5 2E47453934000000  000000000000E8000001
```

Here is that same changed XID string broken down into the standard parts:

Format ID	TID length	Branch qualifier length	Transaction identifier	Branch qualifier
46325043	00000019	00000001	4739314533463135 2E47453934000000 000000000000E800 00	01

The branch qualifier length now has a 1, and the branch qualifier section is added. However, the transaction identifier field has not changed. You can still trace the XID across different data sources by limiting your search string input to the transaction identifier section.

Related reference

LIST INDOUBT TRANSACTIONS command

Troubleshooting federated two-phase commit issues

Troubleshooting federated two-phase commit issues is often specific to the data source that is causing the issue.

Applications must handle error codes that indicate that transactions have timed out, specifically -913 and -918 error codes in addition to checking for -911 error codes.

Oracle data source troubleshooting

Try the following methods to troubleshoot issues with Oracle data sources.

- To log information:

  ```
  db2 "alter server ora1 options (add XA_OPEN_STRING_OPTIONS
  '+LogDir=C:\temp+DbgFl=0x7')
  ```

  Where *C:\temp* is the full path to the location where you want the log file created. Logging information can significantly slow performance, so log information only when you are troubleshooting issues.

- To display trace information, add the following lines to the Oracle client sqlnet.ora file:

  ```
  TRACE_LEVEL_CLIENT=16
  TRACE_DIRECTORY_CLIENT=C:\temp
  ```

  You can also set TRACE_LEVEL_CLIENT to a lower number, such as 4 or 8. Displaying trace information can significantly slow performance, so enable trace level information only when you are troubleshooting issues.
- To list pending transactions exist on the Oracle data source:

  ```
  select * from dba_pending_transactions where formatid=1177702467;
  select * from dba_2pc_pending;
  ```

  To list the state of a transaction and the ID of a transaction:

  ```
  select A.STATE, A.LOCAL_TRAN_ID, A.FAIL_TIME, A.GLOBAL_TRAN_ID,
  B.FORMATID || '.' || B.GLOBALID || '.' || b.BRANCHID as fmt_xid
  from dba_2pc_pending A, dba_pending_transactions B
  where A.GLOBAL_TRAN_ID = B.FORMATID || '.' ||
  B.GLOBALID and STATE='prepared' and B.FORMATID=1177702467;
  ```
- To resolve transactions manually:

  ```
  rollback force '4.31.157818';
  commit force '10.24.154537'
  ```

  where '4.31.157818' is the Oracle field A.LOCAL_TRAN_ID that corresponds to the XID in GLOBAL_TRAN_ID.

Sybase data source troubleshooting

Try the following methods to troubleshoot issues with Sybase data sources.

- Check the Sybase XA log file syb_xa_log located in the $SYBASE directory.
- Use the db2diag tool to check the db2diag.log file.
- To check a transaction on the Sybase server:

  ```
  $ isql -Uuser_name -Ppassword -Sserver_name
  1> sp_transactions
  2> go
  ```

  If you find an invalid or unnecessary transaction, ask the Sybase administrator to delete the transaction.

  Related concepts

  Analyzing db2diag.log files using db2diag

  Related tasks

  "Configuring Oracle data sources" on page 118
  There are several requirements and restrictions for using Oracle data sources for federated two-phase commit.

  "Configuring Sybase data sources" on page 121
  There are several requirements and restrictions for using Sybase data sources for federated two-phase commit.

Federated two-phase commit performance

Data sources that are configured for two-phase commit transactions incur a performance penalty when you compare them to data sources that are configured for one-phase commit transactions.

When a data source uses two-phase commit, the federated server acts as a coordinator, ensuring that all participants are correctly synchronized. This coordination is achieved by additional logging on the federated server and additional communication between data sources. As such, a federated transaction accessing a data source for two-phase commit requires more processing than a transaction that accesses a data source for one-phase commit. Consequently, a data source must be enabled for two-phase commit only when the federated transactions require two-phase commit.

A federated transaction that only requires one-phase commit, such as an update to a single site, but which is run on a data source for two-phase commit is likely to incur a performance penalty when you compare it to the same transaction that is run on a data source without two-phase commit.

A transaction under two-phase commit control incurs the following additional processing regardless of whether or not the transaction requires two-phase commit:

1. All transactions require additional database log writes to the federated server log file.

   The additional log writes enable the federated server to track the data sources in the transaction in order to coordinate subsequent commit and rollback operations.
2. All transactions require additional database communication between the federated server and the data source.
3. IUD (Insert, Update, and Delete) operations require one or more additional database log writes to the remote data source.

Most of the additional processing occurs at the transaction boundaries and is not influenced by the contents within the transaction. As a result, the percentage increase in elapsed time for a short transaction is greater than that for a long transaction.

Concurrent transactions cause the federated server to write multiple log records from the log buffer to the log file in a single write operation. Consequently, applications running federated two-phase commit transactions concurrently incur less overhead than applications running the same transactions serially.

Related concepts

"Two-phase commit for federated transactions" on page 104
A federated system can use two-phase commit for transactions that access one or more data sources. Two-phase commit uses the industry standard X/Open XA protocol to coordinate the processing of distributed unit of work transactions.

Elements of performance

Improving federated two-phase commit performance

You can take steps to improve the performance of transactions in a federated two-phase commit configuration.

Consider the following possible configurations:

- In order to reduce the time spent writing the additional log records to the federated server, place the federated database log files on a device that is capable of fast write transactions, preferably a device that has a write cache. Generally, the federated server log write transactions cause most of the additional processing time. Correct placement of the log files is likely to improve

the two-phase commit performance. This improvement is particularly true for read-only transactions. The placement of the federated server log files is controlled by the NEWLOGPATH database configuration parameter.

- For IUD (Insert, Update, and Delete) transactions, place the remote data source log files on media that are capable of fast writes.
- To reduce the overhead introduced by additional XA messages that are sent between the federated server and data sources:
  - Place the federated server on the same computer as one of the two-phase commit data sources.
  - If you cannot place the federated server on the same computer as a data source, increase the network speed and reduce latency between the federated server and the data sources to help improve performance.

For applications, only enable two-phase commit for a data source when at least one transaction in the application requires two-phase commit. For each server, set the default value for DB2_TWO_PHASE_COMMIT to N and use the SET SERVER OPTION statement to enable two-phase commit within the applications that specifically require two-phase commit.

Related concepts

Quick-start tips for performance tuning

Related reference

newlogpath - Change the database log path configuration parameter

Chapter 9. Insert, update, and delete operations

Authorization privileges for INSERT, UPDATE, and DELETE statements

The privileges required to issue INSERT, UPDATE, and DELETE statements on nicknames are similar to the privileges required to issue these same statements on tables. In addition, you must hold adequate privileges on the data source to perform select, insert, update, and delete operations on the underlying object.

You can grant or revoke SELECT, INSERT, UPDATE, and DELETE privileges on a nickname.

However, granting or revoking privileges on a nickname does not grant or revoke privileges at the data source. At the data source, the privileges must be granted or revoked for the REMOTE_AUTHID specified in the user mapping at the federated server.

The privileges held by the authorization ID of the statement must include the necessary privileges on the nickname (for the federated database to accept the request). The user ID at the data source that is mapped to the authorization ID (through a user mapping) must have the necessary privileges on the underlying table object (for the data source to accept the request).

When a query is submitted to the federated database, the authorization privileges on the nickname in the query are checked. The authorization requirements of the data source object referenced by the nickname are only applied when the query is actually processed. If you do not have SELECT privilege on the nickname, then you can not select from the data source object that the nickname refers to.

Likewise, just because you have UPDATE privilege on the nickname does not mean you will automatically be authorized to update the data source object that the nickname represents. Passing the privileges checking at the federated server does not imply that you will pass the privilege checking at the remote data source. Through user mappings, a federated server authorization ID is mapped to the data source user ID. The privilege checking is enforced at the data source.

Federated system INSERT, UPDATE, and DELETE restrictions

Some restrictions apply to using INSERT, UPDATE, or DELETE statements in a federated system.

The following restrictions apply to updates on nicknames:

- A data source object that is read-only, such as a JOIN view, cannot be updated
- You cannot perform insert, update, and delete operations on federated views created with UNION ALL statements. Federated views created with UNION ALL statements are read-only views.

Unsupported data sources

The federated system does not support insert, update and delete operations against nicknames on nonrelational data sources.

The data sources that federation does not support include:

- BLAST
- BioRS
- Entrez
- Excel
- HMMER
- Table-structured files
- Web services
- WebSphere Business Integration
- XML

Referential integrity in a federated system

In a federated system, the federated database does not enforce referential integrity among data sources.

However, referential integrity constraints at a data source can affect nickname updates. Suppose that you need to insert data that is on the federated server into a nickname. When the federated server sends the insert to the data source, it violates a referential integrity constraint at that data source. The federated server maps the resulting error to a federated error.

Applications are responsible for referential integrity between data sources.

INSERT, UPDATE, and DELETE statements and large objects (LOBs)

You can perform read operations on remote LOBs on federated systems. Write operations on LOBs are supported for some data sources.

With federation you can perform read operations on LOBs that are located in any relational data source. You can perform write operations on LOBs that are located in the following data sources:

- Oracle (Version 8 or higher), using the NET8 wrapper
- DB2 for z/OS (Version 7 or higher), DB2 for iSeries (Version 5), and DB2 Database for Linux, UNIX, and Windows (Version 7 or higher), using the DRDA wrapper

Under certain conditions, you can perform write operations on LOBs in other data sources by altering the nickname column type to a VARCHAR.

Preserving statement atomicity in a federated system

During update operations, federated systems always attempt to keep data in an atomic state at the completion of a DML statement. When data is in an atomic state, it is guaranteed to either be successfully processed or to remain unchanged.

When a client or application issues an INSERT, UPDATE, or DELETE statement on a nickname, a federated server internally processes that statement either as a single DML statement, or as a series of multiple DML statements. If a federated server must send multiple DML statements to a target data source for processing, it is possible that data atomicity might be compromised. To prevent compromising data atomicity, federated systems use data source savepoint APIs to monitor the series of multiple DML statements.

Some data sources or versions of data sources do not externalize savepoint APIs that the federated system can use. Under these circumstances, federated INSERT, UPDATE, or DELETE statements are run without the protection of savepoint APIs.

When an error occurs during federated insert, update, delete transactions, partial update results can occur at the data sources. To correct inconsistency problems, a federated system automatically performs an internal transaction rollback before it returns an SQLCODE error to the applications.

The following data sources do not externalize savepoint APIs that the federated server can use:

- DB2 for iSeries
- DB2 for VM and VSE
- Informix
- Microsoft SQL Server
- ODBC
- Teradata

When an entire insert, update, or delete transaction is pushed down to the data source for processing, the federated server assumes that the data source will preserve the statement atomicity if an error occurs. When only part of the insert, update, or delete transaction is pushed down to the data source for processing, the entire transaction is rolled back if an error occurs.

The IUD_APP_SVPT_ENFORCE server option controls this behavior. By default, this server option is set to 'Y', which rolls back transactions that encounter an error. You can use the CREATE SERVER, ALTER SERVER, or SET SERVER OPTION statement to change the default behavior by altering the server option setting. This change applies to all data source objects that are accessed through the server you specify.

Scenarios for IUD_APP_SVPT_ENFORCE server option behavior

Suppose that you create the nickname INFMX_UT for an Informix table named UT. The UT table has four integer columns, i1, i2, i3, and i4. The i1 column is a unique index column. Assume that AUTOCOMMIT is not activated.

The UT table is empty. You issue an INSERT statement on nickname INFX_UT to insert the values 1, 22, 34, and 40 into Row 1 of the table. The statement is successful.

Then you issue a multiple row INSERT statement on the nickname INFX_UT to insert three rows of data:

- Row 2: 2, 37, 34, 55
- Row 3: 3 ,42, 59, 40
- Row 4: 1, 55, 62, 75

Default behavior of the IUD_APP_SVPT_ENFORCE server option

By default, the IUD_APP_SVPT_ENFORCE server option is set to 'Y'. When set to 'Y', this server option performs an internal rollback of the entire transaction. Although the first two rows of data from the second INSERT are successfully inserted, the data in the last row to be inserted

violates the unique index requirement on column i1. Consequently, all rows for both INSERT statements are rolled back because the entire transaction is rolled back.

Alternate behavior of the IUD_APP_SVPT_ENFORCE server option

When the IUD_APP_SVPT_ENFORCE server option is set to 'N' the transaction is not rolled back. Table 10 lists the rows of the UT table after the inserts. Because the data in the last row to be inserted violates the unique index requirement on column i1, the Informix server returns an error message to the federated system. The second and third rows of data remain in the table. The federated system returns the SQL error SQL0803N to your application. The SQL0803N error message describes the violation of unique indexes. Your application must handle the error recovery.

Table 10. Example of the Informix UT table

	Column (unique index)	Column	Column	Column
Row	**i1**	**i2**	**i3**	**i4**
Row 1	1	22	34	40
Row 2	2	37	34	55
Row 3	3	42	59	40

Modifying data in a federated system

Inserting data into data source objects

To insert data into data sources, use the nicknames for the data source objects in the INSERT statement.

To insert data using a nickname, all of the following privileges must be true:

- The privileges held by the authorization ID of the statement must include the INSERT privilege on the nickname (for the federated database to accept the request).
- The user ID at the data source must have the INSERT privilege on the underlying table object (for the data source to accept the request).
- The user ID at the data source must be mapped to the authorization ID at the federated server through a user mapping.

Restrictions

Federation does not support INSERT operations with nonrelational data sources.

Procedure

To insert data into data source objects, issue the INSERT statement.

Example: An Informix table consists of two columns. The first column contains INTEGER data and the second column contains VARCHAR data (up to 20 characters). The nickname *infx_table_nn* is registered with the federated server for the Informix table.

You can issue INSERT, UPDATE, and DELETE statements on the Informix table using the *infx_table_nn* nickname. The following statement inserts a new row of information into the Informix table:

```
INSERT INTO db2user1.infx_table_nn VALUES(1,'Walter')
```

Updating data in data source objects

To update data into data sources, use the nicknames for the data source objects in the UPDATE statement.

Before you begin

To update data using a nickname, all of the following privileges must be true:

- The privileges held by the authorization ID of the statement must include the UPDATE privilege on the nickname (for the federated database to accept the request)
- The user ID at the data source must have the UPDATE privilege on the underlying table object (for the data source to accept the request)
- The user ID at the data source must be mapped to the authorization ID at the federated server through a user mapping.

Restrictions

Federation does not support UPDATE operations with some data sources, see "Federated system INSERT, UPDATE, and DELETE restrictions" on page 131.

To update data into data source objects, issue the UPDATE statement.

Example: An Informix table consists of two columns. The first column contains INTEGER data and the second column contains VARCHAR data (up to 20 characters). The nickname *infx_table_nn* is registered with the federated server for the Informix table.

You can issue INSERT, UPDATE, and DELETE statements on the Informix table using the *infx_table_nn* nickname. The following statement updates a row of information in the Informix table:

```
UPDATE db2user1.infx_table_nn SET c2='Bill' WHERE c1=2
```

Deleting data from data source objects

To delete data from data sources, use the nicknames for the data source objects in the DELETE statement.

Before you begin

To delete data using a nickname, all of the following privileges must be true:

- The privileges held by the authorization ID of the statement must include the DELETE privilege on the nickname (for the federated database to accept the request)
- The user ID at the data source must have the DELETE privilege on the underlying table object (for the data source to accept the request)
- The user ID at the data source must be mapped to the authorization ID at the federated server through a user mapping.

Restrictions

Federation does not support DELETE operations with some data sources, see "Federated system INSERT, UPDATE, and DELETE restrictions" on page 131.

Procedure

To delete data from data source objects, issue the DELETE statement.

Example: An Informix table consists of two columns. The first column contains INTEGER data and the second column contains VARCHAR data (up to 20 characters). The nickname *infx_table_nn* is registered with the federated server for the Informix table.

You can issue INSERT, UPDATE, and DELETE statements on the Informix table using the *infx_table_nn* nickname. The following statement deletes a row of information in the Informix table:

```
DELETE FROM infx_table_nn WHERE c1=3
```

Assignment semantics in a federated system

When you assign data to a nickname column, the data type might change based on the assignment rules that WebSphere Federation Server uses. You should understand the assignment rules so you get the results that you expect.

The rules for determining the target data type of an assignment to a nickname column are:

- Determine the local source type: The local source type is determined by the local column type or the local result type of the expressions. If the source is constant, the local source type is the same as the type of the constant.
- Determine the target type:
  - If the assignment source has no type, such as parameter markers and NULLs, then the target type is MIN(local_target_type, remote_target_type), where local_target_type is the updated column local data type and remote_target_type is the updated column data source data type. The remote_target_type refers to the default forward type mapping type of the remote target column's data type.
  - If the assignment source is not NULL or parameter markers, then the target type is MIN(local_target_type, remote_target_type, local_source_type).

The definition of MIN(type1, type2)

- Type1 and type2 are not exactly the same.
- MIN(type1,type2) = MIN(type2, type1)
- MIN(type1, type2) = remote_target_type(local_target_type), when MIN(type1, type2) = DECIMAL(0,0)
- BLOB is only compatible with BLOB, so MIN(BLOB(x), BLOB(y))=BLOB(z) where z=min(x,y)
- TIME and DATE data types are not compatible.
- Datetime types and character strings are compatible.
- In Unicode databases, character strings and graphic strings are compatible.

The following tables list the minimum of two data types for numeric, character string, graphic string, and date and time data types.

Table 11. Numeric data types

type1	type2	MIN(type1, type2)
SMALLINT	SMALLINT or INTEGER or BIGINT or REAL or DOUBLE	SMALLINT
INTEGER	BIGINT or REAL or DOUBLE	INTEGER
BIGINT	REAL or DOUBLE	BIGINT
REAL	DOUBLE	REAL
DECIMAL(w,x)	SMALLINT	DECIMAL(p,0) where p=w-x, if p<5; SMALLINT, otherwise
DECIMAL(w,x)	INTEGER	DECIMAL(p,0) where p=w-x, if p<11; INTEGER, otherwise
DECIMAL(w,x)	BIGINT	DECIMAL(p,0) where p=w-x, if p<19; BIGINT, otherwise
DECIMAL(w,x)	DECIMAL(y,z)	DECIMAL(p,s) where p=min(w,y)+min(w-x,y-z), s=min(x,z)
DECIMAL(w,x)	DOUBLE or REAL	DECIMAL(w,x)

The following table lists the minimum of two data types for character string data types.

Table 12. Character string data types

type1	type2	MIN(type1, type2)
CHAR(x)	CHAR(y) or VARCHAR(y) or LONG VARCHAR or CLOB(y)	CHAR(z) where z=min(x,y)
VARCHAR(x)	VARCHAR(y) or LONG VARCHAR or CLOB(y)	VARCHAR(z) where z=min(x,y)
LONG VARCHAR	CLOB(y)	LONG VARCHAR where x>32700, CLOB(x) where x<=32700
CLOB(x)	CLOB(y)	CLOB(z) where z=min(x,y)

The following table lists the minimum of two data types for graphic string data types.

Table 13. Graphic string data types

type1	type2	MIN(type1, type2)
GRAPHIC(x)	GRAPHIC(y) or VARGRAPHIC(y) or LONG VARGRAPHIC or DBCLOB(y)	GRAPHIC(z) where z=min(x,y)
VARGRAPHIC(x)	VARGRAPHIC(y) or LONG VARGRAPHIC or DBCLOB(y)	VARGRAPHIC(z) where z=min(x,y)
LONG VARGRAPHIC	DBCLOB(y)	LONG VARGRAPHIC where x>32700, DBCLOB(x) where x<=32700
DBCLOB(x)	DBCLOB(y)	DBCLOB(z) where z=min(x,y)

The following table lists the minimum of two data types for data and time data types.

Table 14. Date and time data types

type1	type2	MIN(type1, type2)
DATE	TIMESTAMP	DATE
TIME	TIMESTAMP	TIME

If the data of data type CHAR you are inserting is shorter than the target length, the data source pads the rest of the column.

If you are inserting data of DATE or TIME data type into a remote column of TIMESTAMP data type, the data source pads the rest of the column.

Assignment semantics in a federated system - examples

The following table shows several examples of the application of federated assignment semantics in queries given a local type and remote type.

Table 15. Examples of assignment semantics

Local type	Remote type	Your query	Generated remote query
FLOAT	INTEGER	set c1=123.23	set c1=INTEGER(123.23)
INTEGER	FLOAT	set c1=123.23	set c1=INTEGER(123.23)
FLOAT	INTEGER	set c1=123	set c1=123
CHAR(10)	CHAR(20)	set c1='123'	set c1='123' ('123' has a type VARCHAR(3) and it's the shortest of all)
CHAR(10)	CHAR(20)	set c1=char23col	set c1=CHAR(char23col, 10)
CHAR(10)	CHAR(20)	set c1=expr1	• set c1=expr1 -- if expr1 returns char(n) n<= 10 • set c1=CHAR(expr1, 10) if expr1 returns char(n) n>10
TIMESTAMP	DATE	set c1= current timestamp	set c1=DATE(current timestamp)

Chapter 10. Working with nicknames

Nicknames in a federated system

When you want to select or modify data source data, you query the nicknames by using the SELECT, INSERT, UPDATE, and DELETE statements. You submit queries in DB2 SQL to the federated database.

You can join data from local tables and remote data sources with a single SQL statement, as if all the data is local. For example, you can join data that is in the following objects:

- A local DB2 table in the federated database, an Oracle table, and a Sybase view
- A DB2 UDB for z/OS table on one server, a DB2 UDB for z/OS table on another server, and an Excel spreadsheet

By processing SQL statements as if the data sources were ordinary relational tables or views within the federated database, the federated system can join relational data with data in nonrelational formats.

Tables and views in the federated database are *local objects*. Do not create nicknames for these objects. Instead use the actual object name in your SQL statements.

Remote objects are objects not located in the federated database. You need to create nicknames for these objects. For example:

- Tables and views in another database or instance on the federated system
- Tables and views in another database or instance on another system
- Tables and views in data sources such as Oracle, Sybase, and ODBC

WITH HOLD syntax

You can use the WITH HOLD syntax on a cursor that is defined on a nickname.

If you use the syntax and the data source does not support the WITH HOLD syntax, the attribute is ignored.

Triggers

A nickname cannot be an update target in a trigger.

You can include SELECT statements on nicknames in the trigger body. You cannot include INSERT, UPDATE, or DELETE statements on nicknames in the trigger body.

Accessing data with nicknames

With a federated system, you can easily access data, regardless of where it is stored. To access data, you create nicknames for your data source objects, such as tables and views.

For example, if the nickname DEPT represents the remote table EUROPE.PERSON.DEPT, you can use the statement SELECT * FROM DEPT to

query information in the remote table. When you query a nickname, you do not have to remember the connection details about the data source. Thus, you do not need to be concerned with these issues:

- The name of the object at the data source
- The server on which the data source object resides
- The data source type on which the object resides, such as Informix and Oracle
- The query language or SQL dialect that the data source uses
- The data type mappings between the data source and federated server
- The function mappings between the data source and federated server

The underlying metadata in the federated database catalog provides the federated server with the information that it needs to process your queries. This metadata is gathered from the data sources when the federated server and database are set up and configured to access the data sources.

After the federated system is set up, you use the nicknames to query the data sources or to further enhance the federated system configuration.

The SQL statements you can use with nicknames

SQL statements support the use of nicknames.

Table 16. Common SQL statements for use with nicknames

SQL statement	Description	Authorization required
ALTER NICKNAME	Modifies an existing nickname by changing the local column name, the local data type, the federated column options, or the informational constraints. The table or view at the data source is not affected.	• SYSADM or DBADM • ALTER or CONTROL privilege on the nickname • ALTERIN privilege on the schema, if the schema name of the nickname exists • Definer of the nickname, as recorded in the DEFINER column of the catalog view for the nickname
ALTER TABLE	Changes a remote table that was created through the federated database by using transparent DDL. You cannot alter tables that were created natively on the data source. Can use informational constraints to add a referential integrity constraint to a nickname.	• SYSADM or DBADM • ALTER or CONTROL privilege on the nickname • ALTERIN privilege on the schema, if the schema name of the nickname exists
COMMENT ON	Adds or replaces comments in the catalog descriptions of various objects, including functions, function mappings, indexes, nicknames, servers, server options, type mappings, and wrappers.	• SYSADM or DBADM • ALTER or CONTROL privilege on the object • ALTERIN privilege on the schema • Definer of the object, as recorded in the DEFINER column of the catalog view for the object

Table 16. Common SQL statements for use with nicknames (continued)

SQL statement	Description	Authorization required
CREATE ALIAS	Defines an alias for a nickname.	• SYSADM or DBADM • IMPLICIT_SCHEMA authority on the database, if the implicit or explicit schema name of the alias does not exist • CREATEIN privilege on the schema, if the schema name of the alias refers to an existing schema
CREATE INDEX with SPECIFICATION ONLY clause	Creates an index specification (metadata) that indicates to the query optimizer that a data source object has an index. No actual index is created, only the specification is created.	• SYSADM or DBADM • CONTROL or INDEX privilege on the underlying data source object — *and either* IMPLICIT_SCHEMA authority on the database, or CREATEIN privilege on the schema
CREATE TABLE with the OPTIONS clause	Creates a remote table through the federated database using transparent DDL.	• SYSADM or DBADM • CREATETAB privilege on the database and USE privilege on the table space —*and either* IMPLICIT_SCHEMA authority on the database, or CREATEIN privilege on the schema
CREATE TABLE with the AS fullselect and DATA INITIALLY DEFERRED REFRESH clauses	Creates a materialized query table using a fullselect that references a nickname.	• SYSADM or DBADM • CREATETAB privilege on the database and USE privilege on the table space —*and either* IMPLICIT_SCHEMA authority on the database, or CREATEIN privilege on the schema • CONTROL privilege on the table or view • SELECT privilege on the table or view and ALTER privilege if REFRESH DEFERRED is specified
CREATE VIEW	Creates a view that references one or more nicknames.	• SYSADM or DBADM • CONTROL or SELECT privilege on the nickname—*and either* IMPLICIT_SCHEMA authority on the database, or CREATEIN privilege on the schema
DELETE	Deletes rows from the data source object, such as a table or view that has a nickname.	• SYSADM or DBADM • DELETE privilege on the nickname and DELETE privilege on the underlying data source object • CONTROL privilege on the underlying data source object

Table 16. Common SQL statements for use with nicknames (continued)

SQL statement	Description	Authorization required
DROP	Deletes an object, such as a nickname, federated view, or index specification. The table, view, or index at the data source is not affected. When the tables that are created by using the transparent DDL are dropped, the corresponding nickname for that table is also dropped.	• SYSADM or DBADM • DROPIN privilege on the schema for the object • CONTROL privilege on the object
GRANT	Grants privileges on nicknames and federated views, such as ALTER, DELETE, INDEX, INSERT, SELECT, or UPDATE. Privileges at the data source must be granted separately.	• SYSADM or DBADM • WITH GRANT OPTION for each identified privilege • CONTROL privilege on the object
INSERT	Inserts rows into the data source object, such as a table or view that has a nickname.	• SYSADM or DBADM • INSERT privilege on the nickname and INSERT privilege on the underlying data source object • CONTROL privilege on the underlying data source object
LOCK TABLE	Causes the remote object at the data source to be locked. Prevents concurrent application processes from changing a data source table that has a nickname.	• SYSADM or DBADM • SELECT privilege on the underlying table • CONTROL privilege on the underlying table.
REVOKE	Revokes privileges on nicknames and federated views, such as ALTER, DELETE, INDEX, INSERT, SELECT, or UPDATE. Privileges at the data source must be revoked separately.	• SYSADM or DBADM • CONTROL privilege on the object
SELECT	Selects rows from the data source object, such as a table or view that has a nickname.	• SYSADM or DBADM • SELECT privilege on the nickname and SELECT privilege on the underlying data source object • CONTROL privilege on the underlying data source object
UPDATE	Updates the values in specified columns in rows in the data source object, such as a table or view that has a nickname.	• SYSADM or DBADM • UPDATE privilege on the nickname and UPDATE privilege on the underlying data source object • CONTROL privilege on the underlying data source object

When you submit a query to the federated database, the authorization privileges for the nickname in the query are checked. The authorization requirements of the data source object that the nickname refers to are only applied when the query is processed at the data source.

To select, insert, update, or delete data with a nickname, the authorization ID of the statement must include these privileges:

- The appropriate privilege on the nickname for the federated database to accept the request
- The appropriate privilege on the underlying table object for the data source to accept the request

For example to update a data source using a nickname, you need UPDATE privilege on the nickname and UPDATE privilege on the underlying data source object.

Accessing new data source objects

To access new data source objects, you must create nicknames for them. Use the CREATE NICKNAME statement for data sources that do not have nicknames.

Before you begin

The federated system must be configured to access the data source.

A server definition for the data source server on which the object resides must exist in the federated database. You create a server definition with the CREATE SERVER statement.

To insert, update, or delete data with a nickname, all of the following privileges must be true:

- The privileges held by the authorization ID of the statement must include the necessary SELECT, INSERT, UPDATE, and DELETE privileges on the nickname for the federated database to accept the request.
- The user ID at the data source must have the necessary SELECT, INSERT, UPDATE, and DELETE privileges on the underlying table object for the data source to accept the request.
- The user ID at the data source must be mapped to the authorization ID at the federated server through a user mapping.

You must have the required authorizations for the CREATE NICKNAME statement.

- SYSADM or DBADM
- IMPLICIT_SCHEMA authority on the federated database, if the implicit or explicit schema name of the nickname does not exist
- CREATEIN privilege on the schema, if the schema name of the nickname exists

About this task

Periodically, you need to access data source objects that lack nicknames. These might be new objects added to a data source, such as a newly created view. These might be existing objects that were not registered with the federated server when it was initially setup. In either case, you must create a nickname for the object.

Procedure

To access new data source objects, issue the CREATE NICKNAME statement.

Creating nicknames for relational and nonrelational data sources

The CREATE NICKNAME statement differs slightly for relational and nonrelational data sources.

For relational data sources the CREATE NICKNAME statement syntax is:

```
CREATE NICKNAME nickname_name FOR server_name."remote_schema"."object_name"
```

nickname_name
A unique nickname for the data source object. Nicknames can be up to 128 characters in length.

The nickname is a two-part name—the schema and the nickname. If you omit the schema when creating the nickname, the schema of the nickname will be the authentication ID of the user creating the nickname. The default values for the schema name are chosen based on the following hierarchy:

1. CURRENT SCHEMA special register
2. SESSION_USER special register
3. SYSTEM USER special register
4. Authorization ID connected to the database

FOR *server_name."remote_schema"."object_name"*
A three-part identifier for the remote data source object. If your data source does not support schemas, omit the schema from the CREATE NICKNAME statement.

- *server_name* is the name assigned to the data source server in the CREATE SERVER statement.
- *remote_schema* is the name of the remote schema to which the object belongs.
- *object_name* is the name of the remote object that you want to access.

OPTIONS (*options_list***)**
Information about the nickname that enables the SQL Query Compiler and the wrapper to efficiently execute queries.

For some nonrelational data sources the CREATE NICKNAME statement syntax is:

```
CREATE NICKNAME nickname_name column_definition_list
     FOR SERVER server_name
     OPTIONS (options_list)
```

nickname_name
A unique nickname for the data source object, as described above for relational data sources.

column_definition_list
A list of nickname columns and data types.

FOR SERVER *server_name*
The local name that you created for the remote server in the server definition information CREATE SERVER statement.

OPTIONS (*options_list***)**
Information about the nickname that enables the SQL Query Compiler and the wrapper to efficiently execute queries.

Accessing data sources using pass-through sessions

You can submit SQL statements directly to data sources by using a special mode that is called pass-through. The pass-through session allows you to submit SQL statements in the SQL dialect for that data source.

Use a pass-through session when you want to perform an operation that is impossible with the SQL. For example, use a pass-through session to create a procedure, create an index, or perform queries in the native dialect of the data source.

The data sources that support pass-through only accept SQL statements in a pass-through session.

Similarly, you can use a pass-through session to perform actions that are not supported by SQL, such as certain administrative tasks. The administrative tasks that you can perform depend on the data source. For example, for DB2 Database for Linux, UNIX, and Windows, you can run the statistics utility for the data source, but you cannot start or stop the remote database.

You can query only one data source at a time in a pass-through session. Use the SET PASSTHRU command to open a session. When you use the SET PASSTHRU RESET command it closes the pass-through session. If you use the SET PASSTHRU command instead of the SET PASSTHRU RESET command, the current pass-through session is closed and a new pass-through session is opened.

You can use the WITH HOLD syntax on a cursor in a pass-through session. If you use the syntax and the data source does not support the WITH HOLD syntax, the attribute is ignored.

You cannot use pass-through sessions with nonrelational data sources.

Accessing heterogeneous data through federated views

A *federated view* is a view in the federated database whose base tables are located at remote data sources. The federated view references base tables with nicknames, instead of the data source table names.

Before you begin

You must have one of the following authorizations to issue the CREATE VIEW statement:

- SYSADM or DBADM
- For each nickname in any fullselect, both:
  - CONTROL or SELECT privilege on the underlying table or view
  - One of the following authorities or privileges:
    - IMPLICIT_SCHEMA authority on the federated database, if the implicit or explicit schema name of the view does not exist
    - CREATEIN privilege on the schema, if the schema name of the view refers to an existing schema

Privileges for the underlying objects are not considered when defining a view on a federated database nickname.

Restrictions

Federated views with UNION ALL statements are read-only views.

Federated views that include more than one nickname in the FROM clause are read-only views.

Federated views that include only one nickname in the FROM clause might be read-only views.

- If the nickname in the FROM clause is to a nonrelational data source, the federated view is read-only.
- If you include other nicknames as predicates or as subqueries when you create the view, the federated view can be updated.

About this task

When you query from a federated view, data is retrieved from the remote data source. The action of creating a federated database view of data source data is sometimes called "creating a view on a nickname." This is because you reference the nicknames instead of the data sources when you create the view.

These views offer a high degree of data independence for a globally integrated database, just as views defined on multiple local tables do for centralized relational database managers.

Procedure

To create a federated view, issue the CREATE VIEW statement.

Authorization requirements of the data source for the table or view referenced by the nickname are applied when the query is processed. The authorization ID of the statement can be mapped to a different remote authorization ID by a user mapping.

Creating federated views - examples

These examples show how to create federated views to access data from several data sources. The examples show the syntax for the CREATE VIEW statement for federation.

Example: Creating a federated view that merges similar data from several data source objects

You are working with customer data on separate servers: one in Europe, one in Asia, and one in South America. The European customer data is in an Oracle table. The nickname for that table is ORA_EU_CUST. The Asian customer data is in a Sybase table. The nickname for that table is SYB_AS_CUST. The South American customer data is in an Informix table. The nickname for that table is INFMX_SA_CUST. Each table has columns containing the customer number (CUST_NO), the customer name (CUST_NAME), the product number (PROD_NO), and the quantity ordered (QUANTITY). To create a view from these nicknames that merge this customer data, issue the following statement:

```
CREATE VIEW FV1
   AS SELECT * FROM ORA_EU_CUST
   UNION
   SELECT * FROM SYB_AS_CUST
   UNION
   SELECT * FROM INFMX_SA_CUST
```

Example: Joining data to create a federated view

You are working with customer data on one server and sales data on another server. The customer data is in an Oracle table. The nickname for that table is ORA_EU_CUST. The sales data is in a Sybase table. The nickname for that table is SYB_SALES. You want to match up the customer information with the purchases that are made by those customers. Each table has a column containing the customer number (CUST_NO). To create a federated view from these nicknames that joins this data, issue the following statement:

```
CREATE VIEW FV4
   AS SELECT A.CUST_NO, A.CUST_NAME, B.PROD_NO, B.QUANTITY
   FROM ORA_EU_CUST A, SYB_SALES B
   WHERE A.CUST_NO=B.CUST_NO
```

Creating a nickname on a nickname

You can create a nickname on a nickname.

Procedure

To create a nickname on a nickname, follow the steps in this example.

Example: Access a Microsoft Excel spreadsheet from an AIX® federated server

You have a federated server on AIX and a federated server on Windows. You want to access an Excel spreadsheet from both federated servers. However, the Excel wrapper is only supported on federated servers on Windows. To access the Excel spreadsheet from the AIX federated server:

1. On the Windows federated server, install WebSphere Federation Server.
2. Configure the Windows federated server to access Excel data sources.
3. On the Windows federated server, create a nickname for the Excel spreadsheet.
4. On the AIX federated server, install WebSphere Federation Server.
5. Configure the AIX federated server to access DB2 family data sources.
6. On the AIX federated server, create a nickname for the Excel nickname on the Windows federated server.

Selecting data in a federated system

How you select data in a federated systems depends on the type of data source that you select from.

Some of the types of distributed requests used with a federated system are requests that query:

- A single remote data source
- A local data source and a remote data source
- Multiple remote data sources
- A combination of remote and local data sources

To select data from the data sources, use the nicknames for the data source objects in the SELECT statement.

The federated database is a local data source. Tables and views in the federated database are local objects. You do not create nicknames for these objects, you use the actual object name in the SELECT statements.

Remote data sources include: another DB2 Database for Linux, UNIX, and Windows database instance on the federated server, another DB2 Database for Linux, UNIX, and Windows database instance on another server, and data sources other than DB2 Database for Linux, UNIX, and Windows. Objects that reside on remote data sources are remote objects.

Selecting data in a federated system - examples

This topic illustrates a scenario in which a federated server access multiple data sources and provides examples of SELECT statements.

Example: A federated server is configured to access a DB2 for OS/390® data source, a DB2 for iSeries data source, and an Oracle data source. Stored in each data source is a table that contains sales information. This configuration is depicted in the following figure.

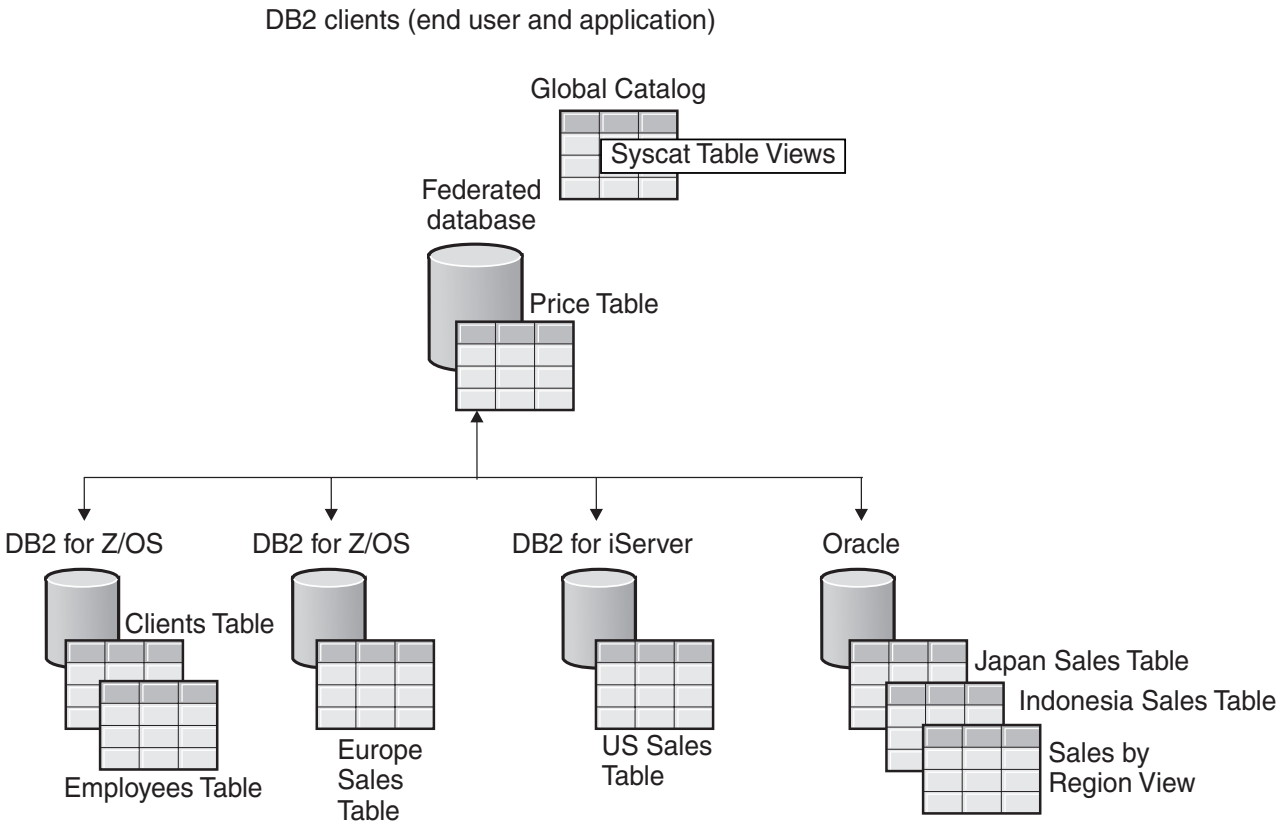


Figure 8. Sample federated system with DB2 and Oracle data sources

The sales tables include columns that record the customer number (CUST_NO), the quantity ordered (QUANTITY), and the product number ordered (PROD_NO). Also, a local table in the federated database contains price information. The price table includes columns that record the product number (PROD_NO) and the current price (PRICE).

The nicknames for the remote data source objects are stored in the SYSCAT.TABLES tables, as shown in the following figure. The TYPE column indicates the type of object, such as nickname (N), local table (T), or view (V).

Data source information

Data source object name	Type of object	Location
PRICES	Local table	Federated database
EUROPE_SALES	Remote table	DB2 for z/OS database
US_SALES	Remote table	DB2 for iSeries database
JAPAN_SALES	Remote table	Oracle database
SALES_BY_REGION	Remote view	Oracle database

SYSCAT Tables

TABNAME	TYPE
PRICES	T
FED_PRICES	N
Z_EU_SALES	N
iS_US_SALES	N
ORA_JAPANSALES	N
ORA_REGIONSALES	N
.....	

Figure 9. Tables and nicknames for sample queries

To select data using a nickname, all of the following privileges must be true:

- The privileges held by the authorization ID of the statement must include the SELECT privilege on the nickname (for the federated database to accept the request).
- The user ID at the data source must have the SELECT privilege on the underlying table object (for the data source to accept the request).
- The user ID at the data source must be mapped to the authorization ID at the federated server through a user mapping.

The following SELECT statement examples use the sample federated system described above.

Example: Querying a single data source:

Z_EU_SALES contains the products ordered by your European customers. It also includes the quantity ordered at each sale. This query uses a SELECT statement with an ORDER BY clause to list the sales in Europe and sorts the list by customer number:

```
SELECT CUST_NO, PROD_NO, QUANTITY
     FROM Z_EU_SALES
     ORDER BY CUST_NO
```

Example: Joining a local data source and a remote data source:

PRICES is a table that resides in the federated database. It contains the price list for the products that you sell. You want to select the prices from this local table that correspond to the products listed in Z_EU_SALES. You also want to sort the result set by the customer number.

```
SELECT sales.CUST_NO, sales.PROD_NO, sales.QUANTITY
     FROM Z_EU_SALES sales, PRICES
     WHERE sales.PROD_NO=PRICES.PROD_NO
     ORDER BY sales.CUST_NO
```

Example: Querying multiple remote data sources:

You want to gather all the sales information from each region and order the result set by product number.

```
WITH GLOBAL_SALES (Customer, Product, Quantity) AS
     (SELECT CUST_NO, PROD_NO, QUANTITY FROM Z_EU_SALES
       UNION ALL
     SELECT CUST.NO,PROD.NO, QUANTITY FROM iS_US_SALES
       UNION ALL
     SELECT CUST.NO,PROD.NO, QUANTITY FROM ORA_JAPANSALES)
     SELECT Customer, Product, Quantity
     FROM GLOBAL_SALES
     ORDER BY Product
```

A view at the Oracle data source lists the sales for Japan and Indonesia. The nickname for this view is ORA_SALESREGION. You want to combine this information with the sales from the United States and display the product prices next to each sale.

```
SELECT us_jpn_ind.CUST_NO, us_jpn_ind.PROD_NO,
        us_jpn_ind.QUANTITY, us_jpn_ind.QUANTITY*PRICES.PRICE
     AS SALEPRICE FROM
     (SELECT CUST_NO, PROD_NO, QUANTITY
     FROM ORA_SALESREGION
     UNION ALL
     SELECT CUST_NO, PROD_NO, QUANTITY
     FROM iS_US_SALES us ) us_jpn_ind,PRICES
     WHERE us_jpn_ind.PROD_NO = PRICES.PROD_NO
     ORDER BY SALEPRICE DESC
```

Informational constraints on nicknames

You can use informational constraints on nicknames to improve the performance of queries. Informational constraints are rules that the optimizer uses to improve performance but that the database manager does not enforce.

You can specify the following for nicknames:

- Referential constraints
- Check constraints
- Functional dependency constraints
- Primary key constraints
- Unique constraints

Specifying informational constraints on nicknames (DB2 Control Center)

To improve the performance of queries on remote data sources, you can specify informational constraints to nicknames. However, the federated server does not enforce or check the constraints. You can specify informational constraints on a nickname from the DB2 Control Center or the DB2 command line.

Restrictions

After you define informational constraints on a nickname, you can only alter the column names for that nickname after you remove the informational constraints.

About this task

For relational data sources, you can specify informational constraints when you alter a nickname.

For nonrelational data sources, you can specify informational constraints when you create a nickname or alter a nickname.

Procedure

To specify informational constraints on a nickname from the DB2 Control Center:

1. Select the **Nicknames** folder:
   - If you are creating a nickname, in the object details pane of the DB2 Control Center, click **Create Nicknames** from the list of actions. The Nicknames wizard opens.
   - From the Create Nicknames window, open the Add Nickname notebook or the Properties notebook for a nickname:
   a. If you plan to create a single nickname, click **Add**. The Add Nickname notebook opens.
   b. If the discover feature generated a list of nicknames, select the nickname that you want to add informational constraints to. Then click **Properties**. The Properties notebook opens.
   - If you are altering a nickname, click on the nickname that you want to alter. In the object details pane of the DB2 Control Center, click **Alter** from the list of actions. The Alter Nickname notebook opens.
2. On the Keys page, set the referential integrity constraints for the nickname. You can set a primary key, unique key, or foreign key constraint.
3. On the Check Constraints page, set the check constraints or functional dependency constraints for the nickname.
4. Click **OK** to set the informational constraints and close the notebook.

Specifying informational constraints on nicknames (DB2 command line)

To improve the performance of queries on remote data sources, you can add informational constraints to nicknames. However, the federated server does not enforce or check the constraints. You can specify informational constraints on a nickname from the DB2 Control Center or the DB2 command line.

Restrictions

After you define informational constraints on a nickname, you can only alter the column names for that nickname after you remove the informational constraints.

About this task

For relational data sources, you can specify informational constraints when you alter a nickname.

For nonrelational data sources, you can specify informational constraints when you create a nickname or alter a nickname.

Procedure

To specify informational constraints on a nickname from the DB2 command line, issue the CREATE NICKNAME statement or the ALTER NICKNAME statement with the appropriate constraint attributes.

Specifying informational constraints on nicknames - examples

These examples illustrate the use of informational constraints on nicknames. You use the CREATE or ALTER NICKNAME statements for check constraints, referential constraints, and other data structures.

Example: Informational check constraint

In the following remote table, the data in the salary column is always greater than 10000.

```
CREATE TABLE account.salary (
    empno INTEGER NOT NULL PRIMARY KEY,
    salary  INTEGER NOT NULL
);
```

Create a nickname for this table:

```
CREATE NICKNAME account.salary FOR myserv.account.salary;
```

Then add informational check constraints for the nickname by issuing the following statement:

```
ALTER NICKNAME account.salary ADD CONSTRAINT cons1 CHECK( salary > 10000 )
NOT ENFORCED
ENABLE QUERY OPTIMIZATION;
```

Example: Informational referential constraint: nickname to nickname

In this example, there are two nicknames N1 and N2. Column F1 of nickname N2 contains the key value in column P1 of nickname N1. You can define the referential constraint on nickname N2 by issuing the following statement:

```
ALTER NICKNAME SCHEMA1.N2 ADD CONSTRAINT ref1
                        FOREIGN KEY (F1) REFERENCES SCHEMA1.N1 (P1)
                        NOT ENFORCED;
```

Example: Informational referential constraint: nickname to table

In this example, nickname N3 with column F1 contains the key value in column P1 of table T1. You can define the referential constraint on nickname N3 by issuing the following statement:

```
ALTER NICKNAME SCHEMA1.N3 ADD CONSTRAINT ref1
                        FOREIGN KEY (F1) REFERENCES SCHEMA1.T1 (P1)
                        NOT ENFORCED;
```

Example: Informational referential constraint: table to nickname

In this example, table T2 with column F1 contains the key value in column P1 of nickname N4. You can define the referential constraint on table T2 by issuing the following statement:

```
ALTER TABLE SCHEMA1.T2 ADD CONSTRAINT ref1
                        FOREIGN KEY (F1) REFERENCES SCHEMA1.N4 (P1)
                        NOT ENFORCED;
```

Example: Functional dependency

In this example, the column pair C1 and C2 uniquely determine the value in the column P1. You can define the functional dependency by issuing the following statement:

```
ALTER NICKNAME SCHEMA1.NICK1 ADD CONSTRAINT FD1 CHECK( P1 DETERMINED BY (C1,C2) )
                        NOT ENFORCED ENABLE QUERY OPTIMIZATION;
```

Example: Table-structured file

This statement defines a primary key for a table-structured file:

```
CREATE NICKNAME MY_FILE (
       X INTEGER NOT NULL,
       Y INTEGER,
       PRIMARY KEY (X) NOT ENFORCED
)  FOR SERVER MY_SERVER OPTIONS(FILE_PATH '/usr/pat/DRUGDATA1.TXT');
```

Star schema

The statement creates four dimension tables and one fact table:

```
CREATE TABLE SCHEMA.FACT (
       LOCATION_CODE   INTEGER NOT NULL,
       PRODUCT_CODE    INTEGER NOT NULL,
       CUSTOMER_CODE   INTEGER NOT NULL,
       SDATE           DATE NOT NULL,
       SALES           INTEGER NOT NULL
);

CREATE TABLE SCHEMA.LOCATION (
       LOCATION_CODE   INTEGER NOT NULL PRIMARY KEY,
       STATE           CHAR(2) NOT NULL,
       SHOP_ID         INTEGER NOT NULL,
       ...
);

CREATE TABLE SCHEMA.PRODUCT (
       PRODUCT_CODE    INTEGER NOT NULL PRIMARY KEY,
       PRODUCT_CAT     INTEGER NOT NULL,
       PRODUCT_NAME    VARCHAR(20) NOT NULL,
       ...
);

CREATE TABLE SCHEMA.CUSTOMER (
       CUSTOMER_CODE   INTEGER NOT NULL PRIMARY KEY,
       NAME            VARCHAR(20) NOT NULL,
       TEL             VARCHAR(10) NOT NULL,
       ...
);

CREATE TABLE SCHEMA.TIMEDIM (
       SDATE           DATE NOT NULL UNIQUE,
       YEAR            INTEGER NOT NULL,
       QUARTER         INTEGER NOT NULL,
       ...
);
```

The federated server creates the following nicknames for the fact table and the dimension tables:

```
CREATE NICKNAME SCHEMA.FACT FOR SERVER.SCHEMA.FACT;
CREATE NICKNAME SCHEMA.LOCATION FOR SERVER.SCHEMA.LOCATION;
CREATE NICKNAME SCHEMA.PRODUCT FOR SERVER.SCHEMA.PRODUCT;
CREATE NICKNAME SCHEMA.CUSTOMER FOR SERVER.SCHEMA.CUSTOMER;
CREATE NICKNAME SCHEMA.TIMEDIM FOR SERVER.SCHEMA.TIMEDIM;
```

You can define the following foreign key relationship by issuing the following statements:

```
ALTER NICKNAME SCHEMA.FACT ADD CONSTRAINT L1 FOREIGN KEY (LOCATION_CODE)
                               REFERENCES SCHEMA.LOCATION(LOCATION_CODE)
                               NOT ENFORCED ENABLE QUERY OPTIMIZATION;

ALTER NICKNAME SCHEMA.FACT ADD CONSTRAINT P1 FOREIGN KEY (PRODUCT_CODE)
                               REFERENCES SCHEMA.PRODUCT(PRODUCT_CODE)
                               NOT ENFORCED ENABLE QUERY OPTIMIZATION;

ALTER NICKNAME SCHEMA.FACT ADD CONSTRAINT C1 FOREIGN KEY (CUSTOMER_CODE)
                               REFERENCES SCHEMA.CUSTOMER(CUSTOMER_CODE)
                               NOT ENFORCED ENABLE QUERY OPTIMIZATION;

ALTER NICKNAME SCHEMA.FACT ADD CONSTRAINT S1 FOREIGN KEY (SDATE)
                               REFERENCES SCHEMA.TIMEDIM(SDATE)
                               NOT ENFORCED ENABLE QUERY OPTIMIZATION;
```

When the value of the TEL column in the CUSTOMER nickname is unique, you can add the following informational unique constraint with this statement:

```
ALTER NICKNAME SCHEMA.CUSTOMER ADD CONSTRAINT U1 UNIQUE( TEL )
                                   NOT ENFORCED ENABLE QUERY OPTIMIZATION;
```

When the value of the SHOP_ID column in the LOCATION nickname uniquely determines the value of the LOCATION_ID column, you can define the following functional dependency with this statement:

```
ALTER NICKNAME SCHEMA.LOCATION
ADD CONSTRAINT F1 CHECK( LOCATION_ID DETERMINED BY SHOP_ID )
                         NOT ENFORCED ENABLE QUERY OPTIMIZATION;
```

Because the value of the QUARTER column in the TIMEDIM nickname is between 1 and 4, you can define the following informational check constraint with this statement:

```
ALTER NICKNAME SCHEMA.TIMEDIM
ADD CONSTRAINT Q1 CHECK( QUARTER BETWEEN 1 AND 4 )
                         NOT ENFORCED ENABLE QUERY OPTIMIZATION;
```

The statements in this example create nicknames for remote tables. Nicknames have primary keys when remote tables have primary keys. When you create nickname on views, nicknames lack primary keys. In this case, you can change the nickname to add an informational primary key constraint. For example:

```
CREATE NICKNAME SCHEMA.LOCATION FOR SERVER.SH.V_LOCATION;
ALTER NICKNAME SCHEMA.LOCATION
      ADD CONSTRAINT P1 PRIMARY KEY ( LOCATION_CODE ) NOT ENFORCED;
```

Chapter 11. Nickname statistics

Nickname statistics update facility - overview

You retrieve the statistics for a nickname to provide the query optimizer with accurate and current information about the nickname. The optimizer uses this information to determine an optimal access plan for a query.

When you register a nickname for a data source object, the federated server adds information about that data source object to the system catalog on the federated database. The query optimizer uses this information to plan how to retrieve data from the data source object. The federated database does not automatically detect changes to the data source objects, so the information in the system catalog can become outdated.

You can retrieve the currently available statistics on database nicknames, columns, and indexes from the remote data source. You can retrieve nickname statistics by using the DB2 Control Center or the DB2 Command Line Processor.

You can retrieve nickname statistics for the following relational data sources:

- DB2 family (DRDA)
- Informix
- Microsoft SQL Server
- ODBC
- Oracle (NET8)
- Sybase (CTLIB)
- Teradata

You can retrieve nickname statistics for the following nonrelational data sources:

- BioRS
- Excel
- Table-structured files
- XML on the root nickname

You can update the following statistics using the catalog-based method for statistics collection:

- CARD
- FPAGES
- NPAGES
- OVERFLOW
- COLCARD
- HIGH2KEY
- LOW2KEY
- NLEAF
- NLEVELS
- CLUSTERFACTOR
- CLUSTERRATIO

- FULLKEYCARD
- FIRSTKEYCARD

You can update the following statistics using the data-based method for statistics collection:

- CARD
- COLCARD
- HIGH2KEY
- LOW2KEY
- FULLKEYCARD
- FIRSTKEYCARD

You can retrieve the statistics of a single nickname or all nicknames in a schema on a specific server definition. You can also retrieve statistics for nicknames on data source objects with Oracle Label Security. If the database is restricted, only users with the appropriate access can view the HIGH2KEY and LOW2KEY statistics. For unrestricted databases, the nickname statistics on objects with Oracle Label Security are exposed and might pose a security risk. In those instances, you can restrict access to the system catalogs that contain sensitive information. If any part of the retrieval fails, the database rolls back the changes.

You can retrieve nickname statistics on any operating system that supports the federated server.

Methods of retrieving nickname statistics

You can choose the method to use for statistics collection, and you can limit your choices to specific columns and indexes. The catalog-based method has better performance. By contrast, the data-based method provides more up-to-date statistics, but takes longer to run.

You can choose one of the following methods for statistics collection:

- Catalog-based method

  The catalog-based method copies statistics from data source catalog tables to the federated catalog table. Only those statistics that can be semantically mapped to federated statistics are copied. However, the nickname statistics are only as accurate and up-to-date as the information currently in the catalog at the remote source. If statistics information is out-of-date, the nickname statistics collected are also out-of-date. When you use the catalog-based method, ensure that statistics on the remote source are current.

  Because statistics are copied from the remote source catalog to the catalog on the federated server, the catalog-based method of statistics collection is generally very fast.

- Data-based method

  The data-based method does not depend on the statistics at the remote source. This method generates its own statistics empirically through the results of the queries that it issues against the nicknames. With this method, the statistics that are collected accurately reflect the remote data.

  The data-based method can be slow if the row size of the nicknames involved is large. The queries typically involve large sorts and aggregates. For this reason, choose data-based statistics collection only if satisfactory statistics cannot be obtained by the catalog-based method.

If you want to increase the performance of the data-based method at the expense of the quality of the statistics gathered, limit statistics collection to the types of columns and indexes for which the benefit is greatest. Those types of columns include columns that are involved in predicates, join keys, grouping operations, or columns that are part of one or more indexes.

With the catalog-based method, you generally do not need to limit statistics collection to specific columns or indexes, because the overhead of this collection method is very low.

Retrieving nickname statistics

You can retrieve the statistics for a nickname to ensure that the query optimizer uses the information about the nickname that is available at the data source. You can update nickname statistics for a single nickname, multiple nicknames, or all nicknames from the DB2 Control Center or the DB2 command line.

About this task

The query optimizer also gathers statistics for data source objects with Oracle Label Security, including the HIGH2KEY and LOW2KEY nickname statistics. For databases that are restricted, only users with the appropriate access level can see the statistics. For unrestricted databases, you can restrict access or set the HIGH2KEY and LOW2KEY nickname statistics to an empty or meaningless value.

Before you begin

The following prerequisites apply when you use the command line prompt to update statistics:

- The federated server creates the log file on the server. The directories that you list in the path must exist.
- The privileges for the fenced user ID of the federated instance must include the privilege to create the log file in the specified location.

Restrictions

The user ID that you use to connect to the federated database must be mapped to the remote data source table.

If the local column name or type does not represent the default type mapping from the remote column name or type, the nickname statistics update utility will not retrieve column statistics.

About this task

By default, statistics collection is attempted for all columns and all indexes of a nickname. You can also limit statistics to specific columns and indexes and specify a log file.

Procedure

To update nickname statistics from the DB2 command line or within your application, call the stored procedure SYSPROC.NNSTAT.

To update nickname statistics from the DB2 Control Center:

1. Select the nicknames for which you want to retrieve current statistics.
2. If a DB2 tools catalog does not exist, a window appears from which you can create the DB2 tools catalog. Create the tools catalog when you want to schedule the update to the nickname statistics.
3. Specify the settings for the update.

Retrieving statistics for multiple nicknames (DB2 Control Center)

You can update statistics for multiple nicknames from the DB2 Control Center or the DB2 command line.

Procedure

To update nickname statistics for multiple nicknames from the DB2 Control Center:

1. Select the required nicknames:
   - Nicknames with server definition:
   a. Expand the **Federated Database Objects** folder. Select the wrapper folder that you want to work with.
   b. Expand the **Server definitions** folder. Select the server folder that contains the nicknames that you want to work with.
   c. Double-click the **Nicknames** folder.
   d. Right-click the names of the nicknames and select **Statistics.**
   e. Select **Update**. The Statistics Update notebook for multiple nicknames opens.
   - To update the statistics for multiple nicknames that are associated with a specific database definition:
   a. Expand the **Databases** folder. Select the **Database** folder that contains the nickname definitions that you want to work with.
   b. Double-click the **Nicknames** folder.
   c. Right-click the names of the nicknames that you want to update and select **Statistics.**
   d. Select **Update**. The Statistics Update notebook for multiple nicknames opens.
2. Specify the settings for the update:
   - Statistics Update page:
   a. View the nicknames in the Statistics Update window and ensure that these are the nicknames that you want to update statistics for. You can uncheck nicknames that you do not want to update.
   b. Select Details <uicontrol> to choose column-level and index-level statistics to update. The Details notebook opens.
   - Details page:
   a. Select all columns of the nickname for update or select specific columns for update.
   b. Select all indexes of the nickname for update or select specific indexes for update.
   c. Select an existing log file or type the fully qualified path for a new log file.
   - Method page, choose one of the following methods for statistics collection:
   a. Catalog-based method, valid for relational nicknames
   b. Data-based method, valid for relational and nonrelational nicknames

c. Both methods, the default

- **Optional:** On the Schedule page, specify when you want the update for the nickname statistics to run.

Retrieving statistics for a single nickname (DB2 Control Center)

You can retrieve statistics for a single nickname associated with a server definition. You can update nickname statistics from the DB2 Control Center or the DB2 command line.

Procedure

To update nickname statistics for a single nickname from the DB2 Control Center:

1. Select the required nicknames.
   - Nicknames with server definitions:
   a. Expand the **Federated Database Objects** folder. Select the wrapper folder that you want to work with.
   b. Expand the **Server definitions** folder. Select the server folder that contains the nickname that you want to work with.
   c. Double-click the **Nicknames** folder.
   d. Right-click the name of the nickname and select **Statistics.**
   e. Select **Update**. The Statistics Update notebook for a single nickname opens.
   - Nicknames with database definition:
   a. Expand the **Databases** folder. Select the **Database** folder that contains the nickname definitions that you want to work with.
   b. Double-click the **Nicknames** folder.
   c. Right-click the name of the nickname that you want to update and select **Statistics.**
   d. Select **Update**. The Statistics Update notebook for a single nickname opens.
2. Specify the settings for the update:
   - Statistics Update page:
   a. Select all columns of the nickname for update or select specific columns for update.
   b. Select all indexes of the nickname for update or select specific indexes for update.
   c. Select an existing log file or type the fully qualified path for a new log file.
   - Method page, choose one of the following methods for statistics collection:
   a. Catalog-based method, valid for relational nicknames.
   b. Data-based method, valid for relational and nonrelational nicknames.
   c. Both methods, the default.
   - **Optional:** On the Schedule page, specify when you want the update for the nickname statistics to run.

Retrieving nickname statistics from the command line - examples

These examples show how to retrieve nickname statistics from the command line by using the SYSPROC.NNSTAT stored procedure.

Example: Retrieving all statistics

The federated server retrieves the statistics for the nicknames on the DB2SERV server and does not create a log.

```
CALL SYSPROC.NNSTAT('DB2SERV',NULL,NULL,NULL,NULL,0,NULL,?)
```

Retrieving all statistics for a schema and returning a log

Example: Retrieving statics with the catalog-based method

The federated server retrieves the statistics for the nickname STAFF in the ADMIN schema. Statistics are gathered for columns 1 through 5 and indexes 1 and 2. The catalog-based method is used to collect statistics. The federated server writes the log to the /home/iiuser/reportlogs/log1.txt file.

```
CALL SYSPROC.NNSTAT(
 NULL, 'ADMIN','STAFF','COL1, COL2, COL3, COL4, COL5','IND1, IND2',1,
  '/home/iiuser/reportlogs/log1.txt',?)
```

In this example, the federated server retrieves the statistics for all the nicknames on the DB2Serv server in the admin schema. The federated server writes the log to the /home/iiuser/stats/recent.log file.

```
CALL SYSPROC.NNSTAT(
 'DB2Serv', 'admin', NULL, NULL, NULL, 0, '/home/iiuser/stats/recent.log', ?)
```

Creating a DB2 tools catalog

When you update the statistics for a nickname, you can use a DB2 tools catalog to schedule subsequent updates to nickname statistics. If you lack a DB2 tools catalog, you are prompted to create a catalog. You can create a DB2 tools catalog from DB2 Control Center or the command line prompt, but you only can schedule the update with the DB2 Control Center.

Before you begin

The DB2 administration server must be installed.

Procedure

To create a DB2 tools catalog from the DB2 Control Center:

1. The Statistics Update window opens when you update nickname statistics.
2. Select the system that you want to create a database on for the DB2 tools catalog.

The database must be on a cataloged system that has no metadata storage. If the system you want is not cataloged, you must catalog the system before you create the database for the DB2 tools catalog.

Viewing the status of the updates to nickname statistics (DB2 Control Center)

After you request an update to the statistics for a nickname, you can view the status of the update. You can view the status of updates to nickname statistics from the DB2 Control Center or the DB2 command line.

Procedure

To view the status of updates to nickname statistics from the DB2 Control Center:

1. Select **Update Statistics**.
2. Select **View Results** and review the status information.

Viewing the status of the updates to nickname statistics (DB2 command line)

After you request an update to the statistics for a nickname, you can view the status of the update. You can view the status of updates to nickname statistics from the DB2 Control Center or the DB2 command line.

Procedure

To view the status of updates to nickname statistics from the DB2 command line, look in the SYSPROC.FED_STATS table.

The following example shows a describe of the table: SYSPROC.FED_STATS (The actual length of the columns is reduced to simplify the example.):

```
db2 describe table sysproc.fed_stats

Column Type Type
name                          schema    name                Length   Scale Nulls
------------------------------ --------- ------------------ -------- ----- ------
SERVER                        SYSIBM    VARCHAR              128     0     Yes
SCHEMA                        SYSIBM    VARCHAR              128     0     Yes
NICKNAME                      SYSIBM    VARCHAR              128     0     Yes
STATS_UPDATE_TIME             SYSIBM    TIMESTAMP             10     0     No
LOG_FILE_PATH                 SYSIBM    VARCHAR             1000     0     Yes
SQLCODE                       SYSIBM    INTEGER                4     0     Yes
SQLSTATE                      SYSIBM    CHARACTER              5     0     Yes
STATUS                        SYSIBM    VARCHAR             1000     0     Yes
8 record(s) selected.


db2 "select * from sysproc.fed_stats"

SERVER  SCHEMA  NICKNAME STATS_UPDATE_TIME           LOG_FILE_PATH  SQLCODE
------  ------- -------- -------------------------  -------------  ----------
ORA8    HAROLDL NICK1    2006-05-02-12.03.24.927112  -             1791 42704

SQLSTATE  STATUS
--------  ------------------------------------------------------------------
SQL1791N  The specified server definition, schema, or nickname does not exist.

1 record(s) selected.

.
```

SYSPROC.NNSTAT stored procedure

Retrieve currently available statistics on one or more nicknames. The statistics are saved in the system catalog on the federated database.

Authorization

SYSPROC.NNSTAT is a fenced procedure. The privileges for the fenced user ID of the federated instance must include the privilege to create the log file in the specified location.

Syntax

```
CALL SYSPROC.NNSTAT(
                    SERVER          VARCHAR(128)
                    SCHEMA          VARCHAR(128)
                    NICKNAME        VARCHAR(128)
                    COLNAMES        CLOB(2M)
                    INDEXNAMES      CLOB(2M)
                    METHOD          SMALLINT
                    LOG_FILE_PATH   VARCHAR(1000)
                    OUT_SQLCODE     INTEGER
                    OUT_TRACE       VARCHAR(2000)
                    )
```

Parameters

Server The server on which the federated server gathers the nickname statistics. This server is what the user registers to define a data source in the federated database. If you specify one nickname, you can specify NULL for this parameter.

Schema

If NULL is specified, the federated server retrieves all the nicknames under the given server. If the Server parameter is NULL, the federated server retrieves the statistics of the nickname under the given schema. If the Schema parameter and Nickname parameter are NULL and you specify a server, the federated server retrieves statistics on the given server.

Nickname

The name of the nickname. If you specify a nickname, you must also specify a schema.

Colnames

The names of the columns that are specified as column-name identifiers.

You can specify this parameter for a single nickname only. If you specify column names, you must also specify a schema and a nickname.

If NULL is specified, statistics are collected for all columns. NULL is the default.

Only the specified columns are processed. If an empty string (″) is specified, columns are not processed.

Indexnames

The names of the indexes that are specified as index-name identifiers.

You can specify this parameter for a single nickname only. If you specify index names, you must also specify a schema and a nickname. Only the specified indexes are processed.

If NULL is specified , statistics are collected for all indexes. NULL is the default.

If an empty string (″) is specified, indexes are not processed.

Method

The method for collecting statistics information from the data source.

0 or NULL

The catalog-based method is used first. If this method fails, then the data-based method is used. This is the default.

1 Catalog-based statistics collection. The catalog-based method maps information from remote catalogs to local statistics for the nickname. This method is valid for relational nicknames only.

2 Data-based statistics collection. The data-based method queries data from the remote table to calculate the values for local statistics. This method is valid for both relational and nonrelational nicknames.

This method is the default for relational nicknames if the catalog-based method fails for a given nickname. Typically, the reason that statistics cannot be collected is because nicknames are defined for remote views. In this case, statistics are not available at the remote source.

Log_File_Path
The path name and file name for the log file. The federated server creates the log file on the server. The directories that you list in the path must exist. On Windows, use two backslashes to specify the log path. For example: c:\\temp\\nnstat.log. If you specify NULL, the federated server does not create a log.

Output parameters

out_SQLCode
The SQL error as a result of the statistics.

out_Trace
The trace.

Example 1: In this example, the federated server retrieves the statistics for the nickname STAFF in the ADMIN schema. Statistics are gathered for columns 1, 3, 4, 6, 7, and 10 and indexes 1 through 3. The data-based method is used to collect statistics. The federated server writes the log to the /home/iiuser/reportlogs/log1.txt file.

```
CALL SYSPROC.NNSTAT(
    NULL, 'ADMIN','STAFF','COL1, COL3, COL4, COL6, COL7, COL10',
    'IND1, IND2, IND3',2,'/home/iiuser/reportlogs/log1.txt',?)
```

Example 2: In this example, the federated server retrieves the statistics for all the nicknames on the DB2SERV server in the ADMIN schema. The federated server writes the log to the c:\\reports\\log1.txt file.

```
CALL SYSPROC.NNSTAT(
    'DB2SERV','ADMIN',NULL,NULL,NULL,0,'c:\\reports\\log1.txt',?)
```

Example 3: In this example, the federated server retrieves the statistics for all the nicknames on the DB2SERV server and does not create a log.

```
CALL SYSPROC.NNSTAT(
    'DB2SERV',NULL,NULL,NULL,NULL,0,NULL,?)
```

Chapter 12. Importing and exporting data for nicknames

You can use the IMPORT command to import data into a nickname and the EXPORT command to export data from a query that references a nickname.

You can only use the IMPORT commands with the following data sources:

- DB2 family
- Informix
- Microsoft SQL Server
- Oracle
- Sybase
- Teradata

Restrictions for importing data into nicknames

There are restrictions on using the IMPORT command to import data into a nickname.

The following restrictions apply when you use the IMPORT command to import data into a nickname:

- The remote object on which the nickname is defined must be a table. You cannot import data into a nickname that is defined on a view or synonym.
- The supported file types are IXF, ASC and DEL.
- The ALLOW WRITE ACCESS clause must be specified. This clause invokes the online import mode. The ALLOW WRITE ACCESS clause allows concurrent applications both read and write access to the import target table.
- You cannot use the COMMITCOUNT AUTOMATIC mode with nicknames.
- The COMMITCOUNT *n* must be specified with *n* being a valid, nonzero number.
- Only the INSERT and INSERT_UPDATE operations are supported with nicknames.
- The column types that are not supported with nicknames are LOBs and generated columns. To import LOB data to a remote table, the corresponding nickname column must be a VARCHAR data type.
- The following `filetype` modifiers are not supported with nicknames:

  ```
  dldelfiletype
  generatedignore
  generatedmissing
  identityignore
  identitymissing
  indexixf
  indexschema
  lobsinfile
  nodefaults
  no_type_idfiletype
  usedefaults
  ```

- Hierarchy (typed table) is not supported with nicknames.

If you submit an IMPORT command that does not adhere to these restrictions, the SQL error code -27999N is returned. For example:

```
SQL27999N The requested IMPORT operation into a remote target (nickname) cannot be
 performed.  Reason code = "reason_code"
```

IMPORT command with nicknames - examples

The examples show you how to import data into nicknames from different file types.

DEL file type

This example uses the INSERT_UPDATE option to import data from a DEL file type:

```
IMPORT FROM import_file_1.del OF DEL
   ALLOW WRITE ACCESS
   COMMITCOUNT 50
   INSERT_UPDATE INTO NICKNAME_1;
```

IXF file type

This example uses the INSERT option to import data from an IXF file type:

```
IMPORT FROM import_file_1.ixf OF IXF
   ALLOW WRITE ACCESS
   COMMITCOUNT 20
   INSERT INTO NICKNAME_1;
```

ASC file type

This example uses the INSERT option to import data from an ASC file type. The example includes the STRIPTBLANKS file modifier to truncate any trailing blank spaces in the data. The METHOD L parameter specifies the start and end of the column numbers.

```
IMPORT FROM import_file_1.asc OF ASC MODIFIED BY STRIPTBLANKS
   METHOD L(1 6, 8 32, 34 44, 46 48)
   ALLOW WRITE ACCESS
   COMMITCOUNT 20
   INSERT INTO NICKNAME_1;
```

Restrictions for exporting data using nicknames

There are restrictions on using the EXPORT command to export data from a query that references a nickname.

The following restrictions apply when you use the EXPORT command to export data using a nickname:

- The description of the target table that is necessary to perform the import CREATE operation is not saved in the IXF file format. Use the db2look utility to collect the information that you need to recreate the table.
- You can export data into the IXF and DEL file types. The ASC file type is not supported for exporting data from nicknames.

Chapter 13. Error tolerance in nested table expressions

Error tolerance is a mechanism that allows query execution to continue while tolerating certain SQL errors in nested table expressions. With error tolerance, instead of receiving an error for a part of a query and terminating the entire query, you can obtain maximum query results from available data.

When the federated server encounters an allowable error, the server allows the error and continues processing the remainder of the query rather than returning an error for the entire query. The result set that the federated server returns can be a partial or an empty result.

When the federated server tolerates errors, it returns query results even when the data sources that the query accesses are not available. This mechanism is useful when you need to return as much information as is available, despite incomplete query results. For example, consider a doctor who needs data about a particular type of medical condition. A query is run to gather information from remote data sources at several different hospitals. If one or more hospital databases are not available, the results from only the available databases are still very valuable to the doctor.

Queries that contain UNION ALL branches can benefit from error tolerance. Without this mechanism, if processing of one branch of the query encounters an error, the federated server stops processing the query. With this mechanism, when you specify the error to tolerate on that same branch of the query, the federated server tolerates the error and continues to navigate to the rest of the available branches. The UNION ALL operation returns the results from any available data sources.

Example: The following query selects data from three nicknames on three different servers:

```
SELECT c1 from nickname1_on_server1
UNION ALL
SELECT c1 from nickname2_on_server2
UNION ALL
SELECT c1 from nickname3_on_server3
```

When nickname2_on_server2 is not available, or when the remote server server2 is not available during query processing, you can obtain the result set from nickname1_on_server1 and from nickname3_on_server3 by tolerating the errors with nickname2 and server2. A result set from two of the three query branches is equivalent to running the following query:

```
SELECT c1 from nickname1_on_server1
UNION ALL
SELECT c1 from nickname3_on_server3
```

You can specify the SQL errors that you want to allow in a nested table expression during query processing. The types of errors that the federated server tolerates are errors with remote connections, authorization, and authentication.

Specifying nested table expressions for error tolerance

You specify the errors to tolerate in a nested table expression with the RETURN DATA UNTIL clause.

About this task

When you use the RETURN DATA UNTIL clause, you must specify the error codes that you want to tolerate. The following table lists the errors that are allowed in the *specific-condition-value* clause. You must specify an SQLSTATE, or an SQLSTATE and SQLCODE, that matches a permissible error code. The SQLCODEs listed in the table are required.

Table 17. Errors allowed in the specific-condition-value clause

SQLSTATE	Error code	SQLCODE
08001	SQL30080N	-30080
08001	SQL30081N	-30081
08001	SQL30082N	-30082
08001	SQL1336N	-1336
08004	Any	Any
28000	Any	Any
42501	Any	Any
42512	Any	Any
42704	SQL0204N	-204
42720	Any	Any

Procedure

To specify nested table expressions for error tolerance, create an SQL statement that contains the RETURN DATA UNTIL clause.

```
RETURN DATA UNTIL specific-condition-value
```

RETURN DATA UNTIL
Any rows that are returned from the fullselect, before the specified condition is encountered, are returned in the result set from the fullselect.

specific-condition-value
Specifies the condition and values for error tolerance.

FEDERATED
Required keyword. The specific condition that you specify must only include errors that occur at a federated data source.

SQLSTATE VALUE *string-constant*
You can specify a specific condition as an SQLSTATE value. The length of the string constant must be 5 when VALUE is specified. An SQLSTATE value can be narrowed down to a particular SQLCODE value. You can specify multiple SQLCODE values that share the same SQLSTATE in one *specific-condition-value*.

Nested table expressions for error tolerance - example

The following examples illustrate how to use the RETURN DATA UNTIL clause to return query results when one or more federated data sources are not available.

Example: The following SQL statement selects data from three different servers: SQLServer, Oracle, and Sybase.

```
SELECT c1 FROM
TABLE RETURN DATA UNTIL
FEDERATED SQLSTATE '08001' SQLCODE -30080, -30082
WITHIN(SELECT c1 FROM n1_from_SQLServer) etq1
UNION ALL
SELECT c1 FROM
TABLE RETURN DATA UNTIL
FEDERATED SQLSTATE '08001' SQLCODE -30080, -30082
WITHIN (SELECT c1 FROM n2_from_Oracle) etq2
UNION ALL
SELECT c1 FROM
TABLE RETURN DATA UNTIL
FEDERATED SQLSTATE '08001' SQLCODE -30080, -30082
WITHIN(SELECT c1 FROM n3_from_Sybase) etq3;
```

Scenario 1: One server is not available.

In this scenario, the Oracle server is not available. The SQLServer server and Sybase server are available. In this situation, the query in the second branch of the UNION ALL operation returns an empty result set with the 30080 error, which is specified to be tolerated. The query returns the results from n1_from_SQLServer and from n3_from_Sybase. Warning sqlwarn5='E' is issued.

The result set is equivalent to running the following query:

```
SELECT c1 FROM n1_from_SQLServer
UNION ALL
SELECT c1 FROM n3_from_Sybase;
```

Scenario 2: All servers are not available.

In this scenario, all servers (SQLServer, Oracle, and Sybase) are unavailable. In this situation, the UNION ALL operation returns an empty result set. Warning sqlwarn5='E' is issued.

Scenario 3: All servers are available.

If all of the servers are available, the result set of the query is equivalent to running the same query without specifying the RETURN DATA UNTIL clause.

Data source support for nested-table-expressions for error tolerance

Error tolerance is supported for several relational data sources and for nonrelational nicknames.

Error tolerance in nested-table-expressions is supported for the following relational data sources:

- DB2 family (DRDA)
- Informix
- Microsoft SQL Server
- ODBC
- Oracle (NET8)
- Sybase (CTLIB)
- Teradata

You can use nonrelational nicknames within nested-table-expressions for error tolerance. The federated server can tolerate the allowable connection, authentication, or authorization errors when the nonrelational wrappers return a permissible error code.

Restrictions on nested-table-expressions for error tolerance

Some restrictions apply when you define error tolerant nested-table-expressions.

A query or view is read-only when you define the query or view with an expression that contains the RETURN DATA UNTIL clause. Cursors that are declared in expressions with the RETURN DATA UNTIL clause are read-only. Errors are returned for each of those situations.

Chapter 14. Monitoring a federated system

Health indicators for federated nicknames and servers

You can use health indicators in the DB2 Health Center to monitor the status of your federated nicknames and servers.

The health indicator for nicknames is db.fed_nicknames_op_status. The health indicator for server definitions is db.fed_servers_op_status. The federated health indicators are installed when the health monitor is installed.

By default, the Health Center does not activate the federated health indicators. You must activate the indicators.

When the state of a nickname or server is not normal, the health indicators issue an alert. You can view the results of monitoring by using the Health Center or the command line.

Federated servers that use AIX, HP-UX, Linux, Microsoft Windows, and Solaris operating systems support the health indicators.

Table 18 describes the health indicators for federated nicknames and servers.

Table 18. Nickname and server health indicators

Health indicator	Description
db.fed_nicknames_op_status	Indicates the aggregate health of all the relational nicknames defined in a database on a federated server. Alerts you if a nickname is invalid. Provides details about the invalid nicknames and recommends actions that you can take to repair them.
db.fed_servers_op_status	Indicates the aggregate health of all the federated servers defined in a database on a federated server. Alerts you if a server is unavailable. Provides details about the unavailable servers and recommends actions that you can take to make them available.

The health indicators can evaluate the following data sources:

- DB2 family (DRDA)
- Excel
- Informix
- Microsoft SQL Server
- ODBC
- Oracle (NET8)
- Sybase (CTLIB)
- Table-structured files
- Teradata
- XML (root nicknames only)

Activating the federated health indicators

To monitor the health of nicknames and servers, you must activate the federated health indicators. The health indicator for nicknames is db.fed_nicknames_op_status. The health indicator for server definitions is db.fed_servers_op_status.

Procedure

To activate the federated health indicators, you can open the DB2 Health Center and configure the health indicators or use the command line processor.

Monitoring the health of federated nicknames and servers

Monitoring nickname and server status can help you determine and resolve problems in your federated system. You can monitor the status of federated nicknames and servers by using health indicators in the Health Center.

Before you begin

- Ensure that SELECT privileges on the nicknames are defined on the federated server.
- Set the FEDERATED database manager configuration parameter to YES.
- If the data source requires authentication, the data source must have user mappings from the Health Monitor's ID. The Health Monitor uses this mapping to connect to the data source.

Restrictions

"Health indicators for federated nicknames and servers" on page 171 lists the data sources that the health indicators can evaluate.

About this task

You can view the results of monitoring by using the Health Center or the command line. Use the DB2 Control Center or DB2 command line processor to resolve the problems that are identified by the health indicators.

Procedure

To do this task from the command line, issue the GET HEALTH SNAPSHOT command.

To do this task from the DB2 Control Center:

1. Open the Health Center.
2. Open the Recommendation Advisor to view recommendations for how to resolve the invalid nicknames or unavailable servers.
3. To do this task from the command line, issue the GET HEALTH SNAPSHOT command.

Monitoring the health of federated nicknames and servers - example

This topic provides an example of the health snapshot of a database.

The federated health indicator names are db.fed_nicknames_op_status and db.fed_servers_op_status. You must enable these health indicators by using either the Health Center or by using the following commands in the CLP:

```
db2 update alert cfg for databases using db.fed_nicknames_op_status set
 THRESHOLDSCHECKED YES
db2 update alert cfg for databases using db.fed_servers_op_status set
 THRESHOLDSCHECKED YES
```

The following command retrieves a database health snapshot including the federated health indicators, if they have been enabled:

```
db2 get health snapshot for database on <database_name>
```

In this example, the database name is fedhi. The output of this command indicates that both health indicators are in normal states. Normal means that the nicknames and the servers are valid.

```
             Database Health Snapshot

Snapshot timestamp                          = 02/10/2006 12:10:55.063004

Database name                               = FEDHI
Database path                               = C:\DB2\NODE0000\SQL00006\
Input database alias                        = FEDHI
Operating system running at database server= NT
Location of the database                    = Local
Database highest severity alert state       = Attention

Health Indicators:

    Indicator Name                             = db.fed_servers_op_status
       Value                                   = 0
       Evaluation timestamp                    = 02/10/2006 12:09:10.961000
       Alert state                             = Normal

    Indicator Name                             = db.fed_nicknames_op_status
       Value                                   = 0
       Evaluation timestamp                    = 02/10/2006 12:09:10.961000
       Alert state                             = Normal

    Indicator Name                             = db.db_op_status
       Value                                   = 0
       Evaluation timestamp                    = 02/10/2006 12:08:10.774000
       Alert state                             = Normal

    Indicator Name                             = db.sort_shrmem_util
       Value                                   = 0
       Unit                                    = %
       Evaluation timestamp                    = 02/10/2006 12:08:10.774000
       Alert state                             = Normal


    Indicator Name                             = db.spilled_sorts
       Value                                   = 0
       Unit                                    = %
       Evaluation timestamp                    = 02/10/2006 12:09:10.961000
       Alert state                             = Normal
```

Snapshot monitoring of federated systems - Overview

You can use the snapshot monitor to capture information about federated data sources and any connected applications at a specific time.

Snapshots are useful for determining the status of a federated system. Taken at regular intervals, snapshots are also useful for observing trends and foreseeing potential problems.

The output of the snapshot monitor is available in the following formats:

- In textual form, through the snapshot monitor command-line processor interface.
- As the output of table functions. This output is useful for writing queries that limit output.

The snapshots that are particularly useful for federated workloads include:

Dynamic SQL statement snapshot
Provides a snapshot of all dynamic SQL statements currently in the statement cache that includes federated and non-federated statements.

Application snapshot
Provides information about a specific application, including the text of the currently running SQL statement.

Remote databases snapshot
Provides information about a specific federated system database.

All remote databases snapshot
Provides information about each federated system database that is active.

Remote applications snapshot
Provides application-level information for each federated system application that is active.

Monitoring federated queries

By monitoring queries, you can determine how your federated system is performing. To help you understand how your federated system is processing a query, you can get a snapshot of the remote query.

About this task

The snapshot monitor tracks two aspects of each query processed by the federated server:

- The entire federated query as submitted by the application, which references nicknames, local tables, or both.
- For queries involving nicknames, one or more *remote fragments*. Remote fragments are the statements that are automatically generated and submitted to remote data sources in their native dialects on behalf of the federated query.

Monitoring federated queries requires that you consider both the work done locally at the federated server and work done at remote servers in response to remote query fragments. The dynamic SQL statement snapshot and the SNAPSHOT_DYN_SQL table function contain information about individual federated queries as they are submitted to the federated server, and about remote query fragments that the federated server, sends to other data sources.

Before you begin

You must set the STATEMENT monitor switch ON for the federated database to collect snapshot information for remote queries.

Procedure

To monitor queries on the federated server, while connected to the federated database, use one of the following methods:

- Textual output:

  ```
  GET SNAPSHOT FOR DYNAMIC SQL on dbname
  ```

  where *dbname* is the name of the federated server database.

- Table function:

  ```
  CREATE TABLE table snap AS (SELECT * FROM TABLE(SNAPSHOT_DYN_SQL ('dbname', -1))
   as snaptab) definition only;
  INSERT INTO snap (SELECT * FROM TABLE(SNAPSHOT_DYN_SQL ('dbname', -1))
   as snaptab);
  ```

  You can then write a query against the snap table that contains one row per query (federated or non-federated) and one row per query fragment in the statement cache of the server.

  The name of remote query fragments is the server to which they were sent prepended to the remote query text, in square brackets, in the stmt_text field of the table function. For example, you can use the following query to look for long-running remote fragments:

  ```
  SELECT total_exec_time, rows_read, total_usr_cpu_time, num_executions,
   substr(stmt_text,1,30)
  FROM TABLE(SNAPSHOT_DYN_SQL ('dbname', -1))AS snaptab
  -- remote fragments only
  WHERE stmt_text LIKE '[%]%'
  ORDER BY total_exec_time;
  ```

By comparing the execution time of an entire federated statement with the execution times of remote fragments sent to other data sources on behalf of the statement, you can understand where most of the processing time is spent.

To determine which query fragments are sent to remote sources on behalf of a federated query, consult an EXPLAIN execution plan for the query.

Snapshot monitoring of federated queries - example

This topic provides an example of output for the text-based dynamic SQL snapshot of a federated query that involves a remote Oracle data source.

The following statement retrieves a snapshot of all statements currently in the statement cache, including federated statements and remote fragments sent to other data sources:

```
GET SNAPSHOT FOR DYNAMIC SQL ON <database_name>
```

The database name is the name of the local federated database.

The output in the following example is the result of the statement:

```
GET SNAPSHOT FOR DYNAMIC SQL ON FEDDB
```

The example shows a federated statement and one remote fragment that is pushed down by that federated statement. You can identify remote fragments by finding the remote server name prepended to the remote statement text in square brackets. In this example, the remote Oracle server is named ORA9. The first entry shows the federated SQL statement that references nicknames, including its overall elapsed time. The second entry shows the remote statement sent to the source [ORA9] that references the remote Oracle table names.

```
  Dynamic SQL Snapshot Result

Database name                       = FEDDB

Number of executions                = 1
Number of compilations              = 1
Worst preparation time (ms)         = 475
Best preparation time (ms)          = 475
Internal rows deleted               = 0
Internal rows inserted              = 0
Rows read                           = 5
Internal rows updated               = 0
Rows written                        = 0
Statement sorts                     = 0
Statement sort overflows            = 0
Total sort time                     = 0
Buffer pool data logical reads      = Not Collected
Buffer pool data physical reads     = Not Collected
Buffer pool temporary data logical reads   = Not Collected
Buffer pool temporary data physical reads  = Not Collected
Buffer pool index logical reads     = Not Collected
Buffer pool index physical reads    = Not Collected
Buffer pool temporary index logical reads  = Not Collected
Buffer pool temporary index physical reads = Not Collected
Buffer pool xda logical reads       = Not Collected
Buffer pool xda physical reads      = Not Collected
Buffer pool temporary xda logical reads    = Not Collected
Buffer pool temporary xda physical reads   = Not Collected
Total execution time (sec.ms)       = 1.816884
Total user cpu time (sec.ms)        = 0.000000
Total system cpu time (sec.ms)      = 0.020000
Statement text                      = select count(*) from orat.supplier,
         orat.nation where s_nationkey =
         n_nationkey and n_name <> 'FRANCE'


Number of executions                = 1
Number of compilations              = 1
Worst preparation time (ms)         = 0
Best preparation time (ms)          = 0
Internal rows deleted               = 0
Internal rows inserted              = 0
Rows read                           = 1
Internal rows updated               = 0
Rows written                        = 0
Statement sorts                     = 0
Statement sort overflows            = 0
Total sort time                     = 0
Buffer pool data logical reads      = Not Collected
Buffer pool data physical reads     = Not Collected
Buffer pool temporary data logical reads   = Not Collected
Buffer pool temporary data physical reads  = Not Collected
Buffer pool index logical reads     = Not Collected
Buffer pool index physical reads    = Not Collected
Buffer pool temporary index logical reads  = Not Collected
Buffer pool temporary index physical reads = Not Collected
Buffer pool xda logical reads       = Not Collected
Buffer pool xda physical reads      = Not Collected
Buffer pool temporary xda logical reads    = Not Collected
Buffer pool temporary xda physical reads   = Not Collected
Total execution time (sec.ms)       = 1.337672
Total user cpu time (sec.ms)        = 0.000000
Total system cpu time (sec.ms)      = 0.000000
Statement text                      = [ORA9] SELECT COUNT(*) FROM "TPCH"."NATION"
         A0, "TPCH"."SUPPLIER" A1 WHERE
         (A0."N_NAME" <> 'FRANCE                   ') AND
         (A1."S_NATIONKEY" = A0."N_NATIONKEY")
```

The snapshot did not collect any buffer pool information, because buffer pool information is not applicable to remote queries.

Federated database systems monitor elements

This topic describes the monitor elements that provide information about federated systems.

A federated system access to diverse data sources that can reside on different platforms, both IBM and other vendors, relational and nonrelational. It integrates access to distributed data and presents a single database image of a heterogeneous environment to its users.

The following elements list information about the total access to a data source by applications running in a federated system and information about access to a data source by a given application running in a federated server instance. They include:

- datasource_name - Data Source Name monitor element
- disconnects - Disconnects monitor element
- insert_sql_stmts - Inserts monitor element
- update_sql_stmts - Updates monitor element
- delete_sql_stmts - Deletes monitor element
- dynamic_sql_stmts - Dynamic SQL Statements Attempted monitor element
- create_nickname - Create Nicknames monitor element
- passthrus - Pass-Through monitor element
- stored_procs - Stored Procedures monitor element
- remote_locks - Remote Locks monitor element
- sp_rows_selected - Rows Returned by Stored Procedures monitor element
- select_time - Query Response Time monitor element
- insert_time - Insert Response Time monitor element
- update_time - Update Response Time monitor element
- delete_time - Delete Response Time monitor element
- create_nickname_time - Create Nickname Response Time monitor element
- passthru_time - Pass-Through Time monitor element
- stored_proc_time - Stored Procedure Time monitor element
- remote_lock_time - Remote Lock Time monitor element

The following example shows the dynamic_sql_statement snapshot:

```
Statement text = [ORACLE817]SELECT A0.C1,A0.C2 FROM ORA_T A0 WHERE A0.C3 = :H0
```

For all remote statements, the Statement text starts with the remote data source name, inside square brackets, followed by the actual text sent to the remote data source.

Chapter 15. Unicode support for federated data sources

Unicode support for federated systems

Relational and nonrelational wrappers and user-defined functions can run on a database in the Unicode code page (UTF-8).

The database in the Unicode code page provides federated server environments that are platform independent. The database can manipulate data that is stored in various code pages on different data sources.

The wrappers and user-defined functions that support Unicode are:

- Relational wrappers
  - DRDA
  - Informix
  - MS SQL Server
  - ODBC
  - OLE DB
  - Oracle
  - Sybase
  - Teradata
- Nonrelational wrappers and user-defined functions
  - BioRS wrapper
  - BLAST wrapper
  - Entrez wrapper
  - Excel wrapper
  - HMMER wrapper
  - KEGG user-defined functions
  - MQ user-defined functions
  - Table-structured file wrapper
  - Web services user-defined functions
  - Web services wrapper
  - WebSphere® Business Integration wrapper
  - XML wrapper

In Figure 10 on page 180 a company has branch offices in different countries. Each branch office stores customer data with its own databases in their own code page. The Microsoft SQL Server database stores data in code page A. The Oracle database stores data in code page B. Code page A and code page B are in different territories. To integrate the data from the different territories, the company can set the federated database's code page to Unicode. The company can then join the tables to see the total number of purchase orders, regardless of territory.

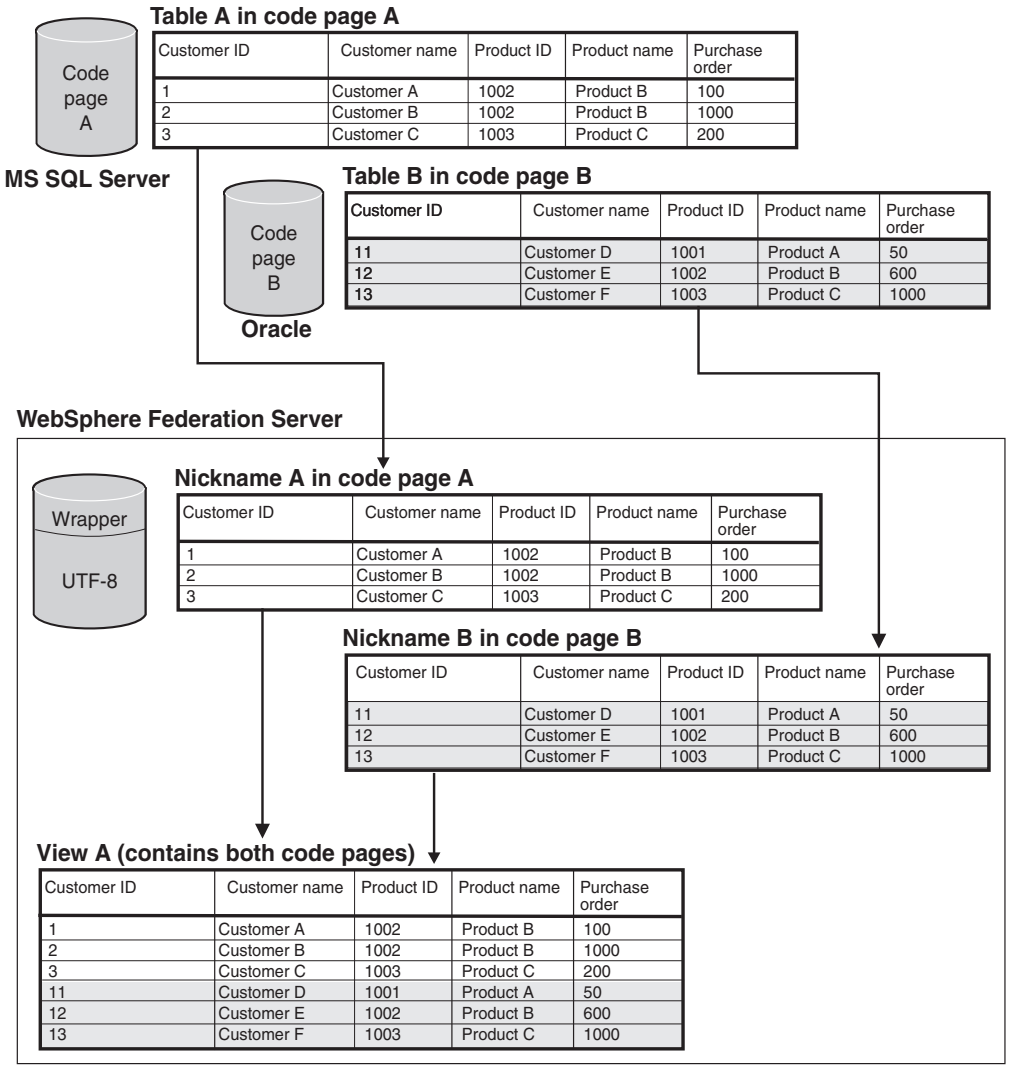


Figure 10. Unicode example

Specifying the client code page for Unicode support of Microsoft SQL Server and ODBC data sources

To ensure correct code page conversion for Microsoft SQL Server and ODBC data sources, you must specify the client code page if the code page differs from the federated database code page.

Procedure

To specify the client code page, issue a CREATE SERVER statement with the CODEPAGE option set to the value of the client code page.

Example: If the data source is Microsoft SQL Server and the federated server is on Windows and the default system locale of the operating system is set to Japanese (Shift-JIS), the CODEPAGE server option must be set to either 943 (Shift-JIS) or 1202 (UTF-16LE). To specify the 1202 code page for the Microsoft SQL server data source named FEDSERVERW, issue the following statement:

```
CREATE SERVER FEDSERVERW TYPE MSSQLSERVER VERSION 2000 WRAPPER MSSQLODBC3
       OPTIONS(NODE 'SAMPLE', DBNAME 'TESTDB', CODEPAGE '1202');
```

Example: If the data source is Microsoft SQL Server and the federated server is running on UNIX and the AppCodePage or IANAAppCodePage setting of the DataDirect Connect client is 6 (Shift-JIS), the CODEPAGE server option must be set to either 943 (Shift-JIS) or 1208 (UTF-8). To specify the 1208 code page for the Microsoft SQL server data source named FEDSERVERU, issue the following statement:

```
CREATE SERVER FEDSERVERU TYPE MSSQLSERVER VERSION 2000 WRAPPER MSSQLODBC3
       OPTIONS(NODE 'SAMPLE', DBNAME 'TESTDB', CODEPAGE '1208');
```

Supported Unicode code pages for the MSSQL and ODBC wrapper CODEPAGE option

Valid code page values are those that DB2 Database for Linux, UNIX, and Windows supports plus those shown in the following table.

Table 19. Supported Unicode code pages for the MSSQL and ODBC wrapper CODEPAGE option

CODEPAGE option value	Description
1200	Codepage1200 - UCS-2 (big-endian)
1202	Codepage1202 - UCS-2 (little-endian)
1208	Codepage1208 - UTF-8
1232	Codepage1232 - UTF-32 (big-endian)
1234	Codepage1234 - UTF-32 (little-endian)

Specifying the file code page for Unicode support of table-structured file data sources

To ensure correct code page conversion for table-structured file data sources data sources, you must specify the file code page if the code page differs from the federated database code page.

Restrictions

You can use the CODEPAGE option only in a Unicode federated database.

About this task

Valid values are those that DB2 Database for Linux, UNIX, and Windows supports. The default value is the code page of the federated database.

Procedure

To specify the code page of a table-structured file, issue the CREATE NICKNAME statement with the CODEPAGE option set to the code page of the data in the table-structured file.

Example: The code page of the data in a file named DRUGDATA1.TXT is 943. To specify the code page of a table-structured file as 943, issue the following CREATE NICKNAME statement:

```
CREATE NICKNAME DRUGDATA1(Dcode Integer NOT NULL, Drug CHAR(20),
         Manufacutuer CHAR(20))
       FOR SERVER biochem_lab
       OPTIONS(FILE_PATH '/usr/pat/DRUGDATA1.TXT',CODEPAGE '943',
      COLUMN_DELIMITER '.',
       SORTED 'Y', KEY_COLUMN 'DCODE', VALIDATE_DATA_FILE 'Y');
```

Specifying the file code page for Unicode support of table-structured file data sources - example

The following example shows you how to specify the code page of a table-structured file.

The code page of the data in a file named DRUGDATA1.TXT is 943. To specify the code page of a table-structured file as 943, issue the following CREATE NICKNAME statement:

```
CREATE NICKNAME DRUGDATA1(Dcode Integer NOT NULL, Drug CHAR(20),
         Manufacutuer CHAR(20))
       FOR SERVER biochem_lab
       OPTIONS(FILE_PATH '/usr/pat/DRUGDATA1.TXT',CODEPAGE '943',
      COLUMN_DELIMITER '.',
       SORTED 'Y', KEY_COLUMN 'DCODE', VALIDATE_DATA_FILE 'Y');
```

Errors when remote and federated code point sizes are different

When the code point size differs between the federated database and the remote data source, you can get truncated data returned or insertion or update failures.

When you select data from the remote data source, that data is truncated if the character string conversion results in a larger number of bytes than the size of the nickname column. If the truncated data ends in a dangling character, blanks fill the remaining bytes. Also, you can insert or update data that is larger than the nickname column size if the converted data size is smaller than, or equal to, the remote column size.

If the federated database has a smaller code point size than the remote data source, insertion or update of data can fail. Insertions or updates fail if the character string conversion results in a larger number of bytes than the size of the remote data source column.

Adjust nickname column size or remote table column size to prevent data truncation or a truncation error.

Chapter 16. Tuning the performance of a federated system

Publications about federated performance

The following IBM documents contain detailed information about performance tuning:

- Parallelism in WebSphere Information Integrator V8.2, at http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0502harris/
- Data Federation with IBM DB2 Information Integrator V8.1, at http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg247052.html?Open
- "Using the federated database technology of IBM DB2 Information Integrator", at ftp://ftp.software.ibm.com/software/data/pubs/papers/iifed.pdf
- "DB2 Information Integrator XML Wrapper Performance", at ftp://ftp.software.ibm.com/software/data/pubs/papers/db2iixmlwrapper.pdf

Query analysis

An important part of query processing is the analysis that determines how to tune the query for optimal performance.

To obtain data from data sources, clients (users and applications) submit queries in SQL to the federated database. The SQL compiler then consults information in the global catalog and the data source wrapper to help it process the query. This includes information about connecting to the data source, server attributes, mappings, index information, and nickname statistics.

As part of the SQL compiler process, the *query optimizer* analyzes a query. The compiler develops alternative strategies, called *access plans*, for processing the query. The access plans might call for the query to be:

- Processed by the data sources
- Processed by the federated server
- Processed partly by the data sources and partly by the federated server

The federated database evaluates the access plans primarily on the basis of information about the data source capabilities and attributes of the data. The wrapper and the global catalog contain this information. The federated database decomposes the query into segments that are called *query fragments*. Typically it is more efficient to pushdown a query fragment to a data source, if the data source can process the fragment. However, the query optimizer takes into account other factors such as:

- The amount of data that needs to be processed.
- The processing speed of the data source.
- The amount of data that the fragment will return.
- The communication bandwidth.

Pushdown analysis is only performed on relational data sources. Nonrelational data sources use the request-reply-compensate protocol.

The following figure illustrates the steps performed by the SQL compiler when it processes a query.

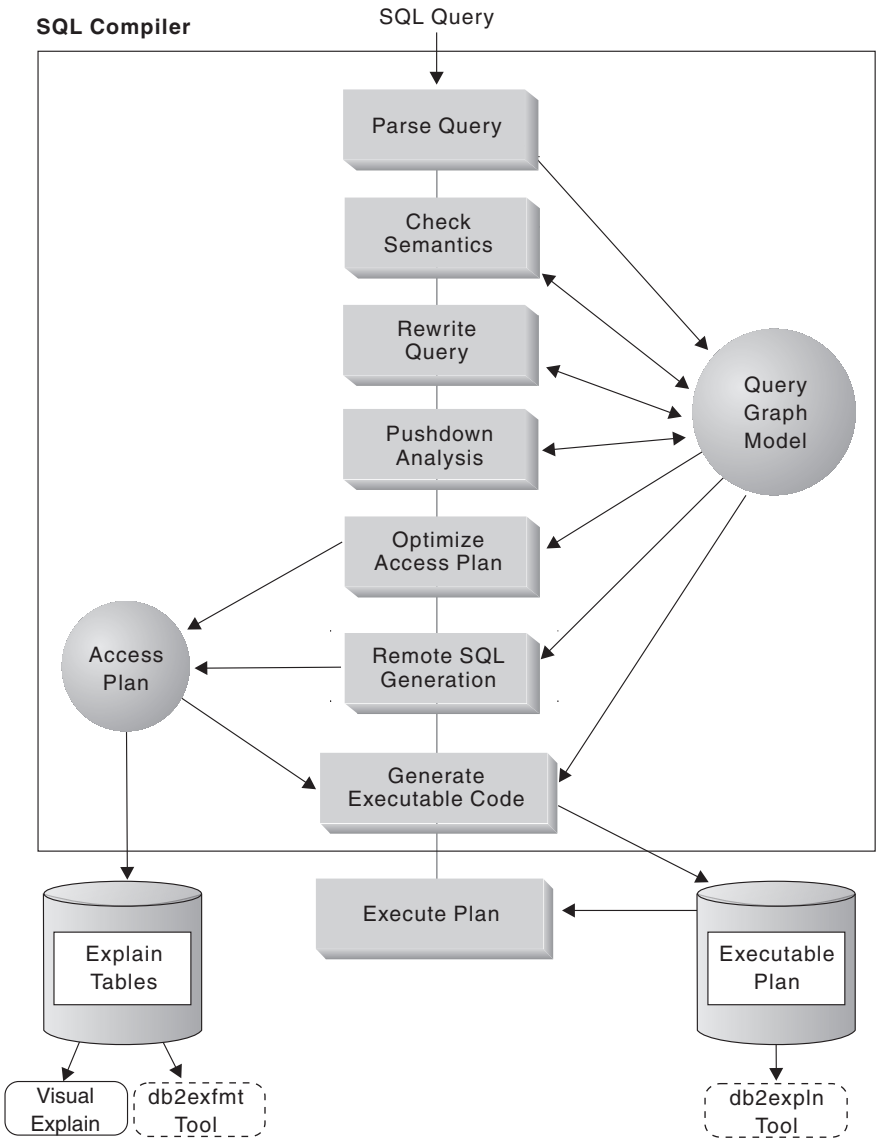


Figure 11. SQL compiler query analysis flowchart

The query optimizer generates local and remote access plans for processing a query fragment, based on resource cost. The federated database then chooses the plan it believes will process the query with the least resource cost.

If any of the fragments are to be processed by data sources, the federated server submits these fragments to the data sources. After the data sources process the fragments, the results are retrieved and returned to the federated server. If the federated database performed any part of the processing, it combines its results with the results retrieved from the data source. The federated server then returns the results to the client.

The primary task of pushdown analysis is to determine which operations can be evaluated remotely. Pushdown analysis does this based on the SQL statement it receives and its knowledge of the capabilities and semantics of the remote data

source. Based on this analysis, the query optimizer evaluates the alternatives and chooses the access plan based on cost. The optimizer might choose to not perform an operation directly on a remote data source because it is less cost-effective. A secondary task is to attempt to rewrite the query to compensate for the difference in semantics and SQL operations between the federated server and the data source so that the query is better optimized.

The final access plan selected by the optimizer can include operations evaluated at the remote data sources. For those operations that are performed remotely, the SQL compiler creates efficient SQL phrased in the SQL dialect of the remote data source during the generation phase. The process of producing an optimal query plan that takes all sources into account is called *global optimization*.

For nonrelational sources, the wrappers use the request-reply-compensate protocol.

Pushdown analysis

Pushdown analysis tells the query optimizer if a remote data source can perform an operation. An operation can be a function, such as relational operator, system or user functions, or an SQL operator (GROUP BY, ORDER BY, and so on). The optimizer then makes a cost-based decision about whether or not to push down the operator. Even if pushdown analysis determines that a particular operation can be performed at the remote source, the optimizer might decide to execute it locally at the federated server, if doing so appears to consume fewer resources.

Pushdown analysis is performed on relational data sources. Nonrelational sources use the request-reply-compensate protocol.

Functions that cannot be pushed-down, can significantly impact query performance. Consider the effect of forcing a selective predicate to be evaluated locally instead of at the remote data source. This approach could require the federated server to retrieve the entire table from the remote data source, and then filter the table locally using the predicate. If your network is constrained—and the table is large—query performance could suffer.

Operators that are not pushed-down can also significantly impact query performance. For example, having a GROUP BY operator aggregate remote data locally could, once again, require the federated server to retrieve the entire table from the remote data source.

For example, suppose that the nickname EMP references the table EMPLOYEE. This table has 10,000 rows. One column contains the zip codes, and one column contains the salary for each employee. The following query is sent to the federated server to count the number of employees per city who earn greater than 50,000 that live in a particular ZIP code range:

```
SELECT CITY, COUNT(*) FROM EMP
 WHERE ZIP BETWEEN 'CA1' AND 'CA5' AND SALARY > 50000
 GROUP BY CITY;
```

When the SQL compiler receives this statement, it considers several possibilities:

- The collating sequences of the data source and the federated server are the same. It is likely that both predicates will be pushed down, because they are likely to reduce the size of the intermediate result set sent from the data source to the federated server. It is usually more efficient to filter and group results at the data source instead of copying the entire table to the federated server and performing

the operations locally. Pushdown analysis determines if operations can be performed at the data source. Since the collating sequences are the same, the predicates and the GROUP BY operation can take place at the data source.

- The collating sequences are the same, and the query optimizer knows that the federated server is very fast. It is possible that the query optimizer will decide that performing the GROUP BY operation locally is the best (least cost) approach. The predicates will be pushed-down to the data source for evaluation. This is an example of pushdown analysis combined with global optimization.
- The collating sequences are not the same. Most likely, the SALARY predicate will be pushed down to the data source, because numeric columns are sorted the same, regardless of collating sequence. However, the predicate on ZIP will not be pushed down because it is order-dependant on a character column. The GROUP BY will not be pushed down unless the predicates on both ZIP and SALARY are also pushed down.

The SQL compiler will consider the available access plans, and then choose the plan that is the most efficient.

In general, the goal is to ensure that the query optimizer considers pushing down the functions and operators to the data sources for evaluation. Many factors can affect whether a function or an SQL operator is evaluated at a remote data source. The key factors which influence the query optimizer are: server characteristics, nickname characteristics, and query characteristics.

Server characteristics affecting pushdown opportunities

Server characteristics that affect pushdown include SQL support, collating sequence, federated server options, and type and function mappings.

The factors that affect pushdown opportunities for nonrelational data sources are different than the factors that affect pushdown opportunities for relational data sources. The SQL dialect is not a factor for most nonrelational data sources, since they do not use SQL.

The following topics describe the data source-specific factors that can affect pushdown opportunities.

SQL differences

SQL characteristics that affect pushdown include SQL capabilities, restrictions, limitations, and SQL specific to a server.

- SQL capabilities. Each data source supports a variation of the SQL dialect and different levels of functionality. For example, consider the GROUP BY list. Most data sources support the GROUP BY operator. However some data sources have restrictions on the number of items on the GROUP BY list, or restrictions on whether an expression is allowed on the GROUP BY list. If there is a restriction at the remote data source, the federated server might perform the GROUP BY operation locally.
- SQL restrictions. Each data source can have different SQL restrictions. For example, some data sources require parameter markers to bind in values to remote SQL statements. Therefore, parameter marker restrictions must be checked to ensure that each data source can support such a bind mechanism. If the federated server cannot determine a good method to bind in a value for a function, this function must be evaluated locally.

- SQL limitations. The federated server might allow the use of larger integers than the remote data sources. However, limit-exceeding values cannot be embedded in statements that are sent to the data sources. Therefore, the function or operator that operates on this constant must be evaluated locally.
- Server specifics. Several factors fall into this category. One example is sorting NULL values (highest, or lowest, depending on the ordering). For example, if the NULL value is sorted at a data source differently from the federated server, ORDER BY operations on a nullable expression cannot be remotely evaluated.

Collating sequence

If you set the COLLATING_SEQUENCE server option to 'Y', you are telling the federated database that the data source collating sequence matches the collating sequence of the federated server. This setting allows the optimizer to consider pushing down order-dependent processing to a data source, which can improve performance.

If the data source collating sequence is not the same as the federated database collating sequence and you set the COLLATING_SEQUENCE server option to 'Y', you can receive incorrect results. For example, if your plan uses merge joins, the optimizer might push down ordering operations to the data sources. If the data source collating sequence is not the same, the join results might not have a correct result set. Set the COLLATING_SEQUENCE server option to 'N', if you are not sure that the collating sequence at the data source is identical to the federated database collating sequence.

Alternatively, you can configure a federated database to use the same collating sequence that a data source uses. You then set the COLLATING_SEQUENCE server option to 'Y'. This allows the optimizer to consider pushing down order-dependent operations on character columns.

To determine if a data source and the federated database have the same collating sequence, consider the following factors:

- National language support

  The collating sequence is related to the language supported on a server. Compare the federated database NLS information for your operating system to the data source NLS information.
- Data source characteristics

  Some data sources are created using case-insensitive collating sequences, which can yield different results from the federated database in order-dependent operations.
- Customization

  Some data sources provide multiple options for collating sequences or allow the collating sequence to be customized.

When a query fragment from a federated server requires sorting, the place where the sorting is processed depends on several factors. If the federated database's collating sequence is the same as the data source collating sequence, the sort can take place at the data source or at the federated server. The query optimizer can determine if a local sort or a remote sort is the most efficient way to complete the query.

Numeric comparisons, in general, can be performed at either location even if the collating sequence is different. You can get incorrect results, however, if the weighting of null characters is different between the federated database and the data source.

Likewise, for comparison statements, be careful if you are submitting statements to a case-insensitive data source. The weights assigned to the characters ″I″ and ″i″ in a case-insensitive data source are the same. For example, in a case-insensitive data source with an English code page, **STEWART**, **SteWArT**, and **stewart** would all be considered equal. The federated database, by default, is case-sensitive and would assign different weights to the characters.

If the collating sequences of the federated database and the data source differ, the federated server retrieves the data to the federated database, so that it can do the sorting locally. The reason is that users expect to see the query results ordered according to the collating sequence defined for the federated server; by ordering the data locally, the federated server ensures that this expectation is fulfilled.

If your query contains an equality predicate on the character column, it is possible to push down that portion of the query even if the collating sequences are different (set to 'N'). For example, the predicate C1 = 'A' would retrieve the same values regardless of collating sequence and therefore could be pushed down to a data source that has a different collating sequence than the federated server. However, such predicates cannot be pushed down when the collating sequence at the data source is case-insensitive (COLLATING_SEQUENCE='I'). When a data source is case-insensitive, the results from C1= 'A' and C1 = 'a' are the same, which is not consistent with a case-sensitive environment such as DB2 Database for Linux, UNIX, and Windows.

Administrators can create federated databases with a particular collating sequence that matches the data source collating sequence. This approach can speed performance if all data sources use the same collating sequence or if most or all column functions are directed to data sources that use the same collating sequence. For example, in DB2 for z/OS, sorts defined by ORDER BY clauses are implemented by a collating sequence based on an EBCDIC code page. If you want to use the federated server to retrieve DB2 for z/OS data sorted in accordance with ORDER BY clauses, it is advisable to configure the federated database so that is uses a predefined collating sequence based on the EBCDIC code page.

If the collating sequences at the federated database and the data source differ, and you need to see the data ordered in the data source's sequence, you can submit your query in a pass-through session, or define the query in a data source view.

Federated server options

The server options that you set refine the knowledge that the federated server has about the remote data source.

The previously listed factors that affect pushdown opportunities are characteristics of the database servers, and you can not change them. By carefully considering the following server options, it is possible to improve query performance:

- COLLATING_SEQUENCE. If a data source has a collating sequence that differs from the federated database collating sequence, any order-dependent operations on character values cannot be remotely evaluated at the data source. An example is executing MAX column functions on a nickname character column at a data source with a different collating sequence. Because results might differ if the

MAX function is evaluated at the remote data source, the federated database will perform the aggregate operation and the MAX function locally.

- VARCHAR_NO_TRAILING_BLANKS. This option is for varying-length character strings that contain no trailing blanks. Some data sources, such as Oracle, do not apply blank-padded comparison semantics like the federated database does. This padding difference can cause unexpected results.

  For example:

  ```
  'HELLO' = 'HELLO         ' in DB2
  'HELLO' <> 'HELLO         ' in Oracle!
  ```

  If trailing blanks are present on VARCHAR columns on an Oracle data source, you should set this option to 'N' (the default for Oracle). This option impacts performance, because the federated server must compensate for the difference in semantics, but guarantees a consistent result set. Setting this value to 'Y' when an Oracle data source column contains trailing blanks can cause inconsistent results.

  If you are certain that all VARCHAR and VARCHAR2 columns at a data source contain no trailing blanks, consider setting this server option for a data source. Ensure that you consider all objects that can potentially have nicknames, including views.

  Recommendation: Set this option on a column by column basis using the VARCHAR_NO_TRAILING_BLANKS column option.

- DB2_MAXIMAL_PUSHDOWN. This option specifies the primary criteria that the query optimizer uses when choosing an access plan. The query optimizer can choose access plans based on cost or based on the user requirement that as much query processing as possible be performed by the remote data sources. With DB2_MAXIMAL_PUSHDOWN set to 'Y', reducing network traffic becomes the overriding criteria for the query optimizer. The query optimizer uses the access plan that performs the fewest number of "sends" to the data sources. Setting this server option to 'Y' forces the federated server to use an access plan that might not be the lowest cost plan. Using an access plan other than the lowest cost plan can decrease performance. When the DB2_MAXIMAL_PUSHDOWN server option is set to 'Y' a query that will result in a Cartesian product will not be pushed down the remote data sources. Queries that will result in a Cartesian product will be processed by the federated database. The DB2_MAXIMAL_PUSHDOWN server option does not need to be set to 'Y' for the federated server to pushdown query processing to the remote data sources. When this server option is set to 'N' (the default), the query optimizer will pushdown query processing to the data sources. However, the primary criteria the optimizer uses when the option is set to 'N' is cost instead of network traffic.

"Server characteristics affecting global optimization" on page 197 describes the COMM_RATE, CPU_RATIO, and IO_RATIO server options that also can affect query performance.

Type and function mapping factors

Both the default data type mappings and default function mappings are built into the data source wrappers. Data type mappings describe the relationship between the data source data type and the federated server data type. You can customize default data type mappings. Function mappings describe the relationship between a data source function and a semantically equivalent function at the federated server. In certain cases, the federated database will compensate for functions mappings that are not supported by a data source.

The default data type mappings are designed so that sufficient buffer space is given to each data source data type to avoid runtime buffer overflow and truncation. You can customize the type mapping for a specific data source or for a particular nickname to suit specific applications and in some cases improve performance. For example, Oracle DATE types can contain both a date and a timestamp portion and are therefore mapped to DB2 TIMESTAMPs by default. If you are accessing an Oracle date column, and you know that it contains only date portions (no timestamps), you can use the ALTER NICKNAME statement to change the local data type of the nickname from TIMESTAMP to DATE. When evaluating predicates based purely on a date, such as SalesDate=DATE('1996-01-04'), this change bypasses the use of a SCALAR function that is used to extract the date from the date and timestamp information held in the column, which can improve performance.

The federated database compensates for functions that are not supported by a data source. Functional compensation usually involves retrieving the necessary data from the data source and applying the function locally, which often has a performance impact. There are several cases where function compensation occurs:

- A function simply does not exist at the data source. Some of the SYSFUN functions, for example, do not exist on DB2 for z/OS and OS/390 data sources, and thus require local compensation.
- A function exists at the data source; however, the characteristics of the operand violate function restrictions. An example is the IS NULL relational operator. Most data sources support it, but some have restrictions such as only allowing a column name on the left hand side of the IS NULL operator.
- A function, if evaluated remotely, can return a different result. An example is the '>' (greater than) operator. For those data sources with different collating sequences, the greater than operator can return different results than if it is evaluated locally by the federated database.

Nickname characteristics affecting pushdown opportunities

Nickname characteristics that affect pushdown include the local data type of a nickname column, federated column options, and materialized query tables.

There are several nickname-specific factors that can affect pushdown opportunities. The local data type of a nickname column can affect the number of possibilities in a joining sequence evaluated by the optimizer. Nicknames can be flagged with a column option to indicate the columns contain no trailing blanks. This gives the SQL compiler the opportunity to generate a more efficient form of a predicate for the SQL statement sent to the data sources.

Local data type of a nickname column

Ensure that the local data type of a column does not prevent a predicate from being evaluated at the data source.

The default data type mappings are provided to avoid any possible overflow. However, a joining predicate between two columns of different data types or lengths will prevent the hash join technique from being considered by the optimizer. For the optimizer to consider the hash join, both the data types and lengths of the joining columns must match exactly. For example, Oracle data source columns that are designed to hold just integer values are often created as NUMBER within the Oracle database, which defaults to NUMBER (38). A nickname column for this Oracle data type is given the local data type FLOAT because the range of a DB2 integer is only roughly equal to NUMBER (9). In this

case, joins between a DB2 integer column and an Oracle column that is defined as NUMBER (but only holding integer values) cannot use the hash join technique because the Oracle column is mapped as a FLOAT type. However, if the domain of this Oracle NUMBER column can be accommodated by the DB2 INTEGER data type, you can change its local data type with the ALTER NICKNAME statement. Then the optimizer can consider the hash join technique, which might improve performance.

Federated column options

You can define federated column options that the query optimizer uses for developing access plans.

The column options tell the wrapper to handle the data in a column differently than it normally would handle it. The SQL complier and query optimizer use the metadata to develop better plans for accessing the data. The federated database treats the object that a nickname references as if it is a table. As a result, you can set column options for any data source object that you create a nickname for.

The ALTER NICKNAME statement can be used to add or change column options for nicknames. There are two column options:

- NUMERIC_STRING. This column option applies to character type columns (CHAR and VARCHAR). Suppose that a data source has a collating sequence that differs from the federated database collating sequence. The federated server would not sort any columns that contain character data at the data source. It would return the data to the federated database and perform the sort locally. However, suppose that the column is a character data type and contains only numeric characters ('0','1',...,'9'). You can indicate this by assigning a value of 'Y' to the NUMERIC_STRING column option. This gives the query optimizer the option of performing the sort at the data source because numerics, even when represented as character strings, always sort the same, regardless of collating sequence. If the sort is performed remotely, you can avoid the overhead of porting the data to the federated server and performing the sort locally.
- VARCHAR_NO_TRAILING_BLANKS. Unlike the server option with the same name, this column option can be used to identify specific Oracle columns that contain no trailing blanks. The SQL compiler pushdown analysis step will then take this information into account when checking all operations performed on columns which have this setting. Based on the VARCHAR_NO_TRAILING_BLANKS setting, the SQL compiler can generate a different but semantically equivalent form of a predicate that is used in the remote SQL statement sent to the data source. A value of 'Y' is likely to enable the use of remote indexes (if available) which can improve query performance.

Query characteristics affecting pushdown opportunities

An SQL operator that references multiple data sources affects pushdown.

A query can reference a SQL operator that involves nicknames from multiple data sources. When the federated server combines the results from two referenced data sources by using one operator, the operation must take place at the federated server. An example of this is a set operator, like UNION. The operator cannot be evaluated at a remote data source directly.

Analyzing where a query is evaluated

Detailed query optimizer information is kept in Explain tables separate from the actual access plan itself. This information allows for in-depth analysis of an access plan. By examining the SHIP operator of a federated access plan, you can determine what SQL operations were pushed down to a data source and which operations were executed at the federated server.

Explain tables are accessible on all supported operating systems, and contain information for both static and dynamic SQL statements. The following tools are typically used to obtain access plan information from the explain tables:

- Explain table format tool. Use the db2exfmt tool to present the information from the explain tables in a predefined format.
- db2expln and dynexpln tools. You can use these tools to understand the access plan chosen for a particular SQL statement. You can also use the integrated Explain Facility in the DB2 Control Center in conjunction with Visual Explain to understand the access plan chosen for a particular SQL statement. Both dynamic and static SQL statements can be explained using the Explain Facility. One difference from the Explain tools is that with Visual Explain the Explain information is presented in a graphical format. Otherwise the level of detail provided in the two methods is equivalent. To fully use the output of db2expln, and dynexpln you must understand:
  - The different SQL statements supported and the terminology related to those statements (such as predicates in a SELECT statement)
  - The purpose of a package (access plan)
  - The purpose and contents of the system catalog tables
  - General application tuning concepts

You can also access explain tables using SQL statements. This allows for easy manipulation of the output, for comparison among different queries, or for comparisons of the same query over time.

Analyzing where a query is evaluated with the DB2_MAXIMAL_PUSHDOWN server option

You can use the DB2_MAXIMAL_PUSHDOWN server option in conjunction with the Explain utilities to determine whether a particular operator was not pushed down to execute at a data source because of a cost-based optimizer decision or because pushdown analysis determined it was not possible.

Procedure

To run the Explain tools on a query with the DB2_MAXIMAL_PUSHDOWN server option:

1. Set the DB2_MAXIMAL_PUSHDOWN server option to 'N'. This is the default setting for this option. Pushdown analysis determines which parts of the SQL can be pushed down. Then the query optimizer generates all the alternative plans that do not violate the criteria set by pushdown analysis. The query optimizer estimates the cost of each plan, and will select the plan with the lowest estimated cost. You can analyze the operators that were pushed down to the data source by viewing the details of the appropriate SHIP operator. If an operator you expect to be pushed down was not pushed down, proceed to step 2.

2. Set the DB2_MAXIMAL_PUSHDOWN server option to 'Y'. Use the Explain tools to analyze the SQL statement again. The plan displayed in the Explain output shows all of the SQL operations that can be pushed down to the data source.
   - If the operator is pushed down after resetting the option to 'Y', the optimizer determined that it was more cost-efficient to execute the operator locally, rather than remotely. If the operator is not pushed down after resetting the option to 'Y', it is likely that pushdown analysis did not allow the operator to be executed remotely.
   - If the optimizer made a cost-based decision not to push down the operator, consider checking the nickname statistics to ensure that they are accurate. If pushdown analysis made the decision not to push down the operator, consider checking server options, data type mappings, and function mappings.

Understanding access plan evaluation decisions

The topics in this section list typical access plan analysis questions, and areas that you can investigate to increase pushdown opportunities.

Why isn't this predicate being evaluated remotely?

This question arises when a predicate is very selective and thus could be used to filter rows and reduce network traffic. Remote predicate evaluation also affects whether a join between two tables of the same data source can be evaluated remotely.

Areas to examine include:

- Server options. How do the settings for the server options COLLATING_SEQUENCE and VARCHAR_NO_TRAILING_BLANKS affect where the predicate is evaluated?
- Subquery predicates. Does this predicate contain a subquery that pertains to another data source? Does this predicate contain a subquery involving an SQL operator that is not supported by this data source? Not all data sources support set operators in a predicate.
- Predicate functions. Does this predicate contain a function that cannot be evaluated by this remote data source? Relational operators are classified as functions.
- Predicate bind requirements. Does this predicate, if remotely evaluated, require bind-in of some value? If so, would it violate SQL restrictions at this data source?
- Global optimization. The optimizer decided that local processing is more cost-effective.

Why isn't the GROUP BY operator evaluated remotely?

There are several areas that you can check to determine why a GROUP BY operator is not evaluated remotely.

The areas that you can check include:

- Is the input to the GROUP BY operator evaluated remotely? If the answer is no, examine the input.
- Does the data source have any restrictions on this operator? Examples include:
  - Limited number of GROUP BY items

  - Limited byte counts of combined GROUP BY items
  - Column specification only on the GROUP BY list
- Does the data source support this SQL operator?
- Global optimization. The optimizer decided that local processing is more cost-effective.

Why isn't the SET operator evaluated remotely?

You can check the operands and check data source restrictions to determine why the SET operator is not evaluated remotely.

Considerations:

- Are both of its operands completely evaluated at the same remote data source? If the answer is no and it should be yes, examine each operand.
- Does the data source have any restrictions on this SET operator? For example, are large objects or long fields valid input for this specific SET operator?

Why isn't the ORDER BY operation evaluated remotely?

You can check the input to the operation, what the clause contains, and check data source restrictions to determine why the ORDER BY operator is not evaluated remotely.

Considerations:

- Server options. How do the settings for the server options COLLATING_SEQUENCE and VARCHAR_NO_TRAILING_BLANKS affect where the predicate is evaluated?
- Is the input to the ORDER BY operation evaluated remotely? If the answer is no, examine the input.
- Does the ORDER BY clause contain a character expression? If yes, does the remote data source have a different collating sequence than the federated server collating sequence?
- Does the data source have any restrictions on this operator? For example, is there a limited number of ORDER BY items? Does the data source restrict column specification to the ORDER BY list?

Why is a remote INSERT with a fullselect statement not completely evaluated remotely?

You can check several elements of the subselect to determine why a remote INSERT with a fullselect statement is not completely evaluated remotely.

Considerations:

- Could the subselect be completely evaluated on the remote data source? If no, examine the subselect.
- Does the subselect contain a set operator? If yes, does this data source support set operators as input to an INSERT?
- Does the subselect reference the target table? If yes, does this data source allow this syntax?

Why is a remote INSERT with VALUES clause statement not completely evaluated remotely?

You can check the VALUES clause and the expression to determine why a remote INSERT with a VALUES clause statement is not completely evaluated remotely.

Considerations:

- Can the VALUES clause be completely evaluated at the remote data source? In other words, does an expression contain a function not supported by the remote data source?
- Does the expression involve a scalar subquery? Is that syntax supported?
- Does the expression reference the target table? Is that syntax supported?

Why is a remote, searched UPDATE statement not completely evaluated remotely?

You can check elements of the SET clause and search condition to determine why a remote, searched UPDATE statement is not completely evaluated remotely.

Considerations:

- Can the SET clause be completely evaluated at the remote data source? In other words, does an update expression contain a function not supported by the remote data source?
- Does the SET clause involve a scalar subquery? Does the data source allow this syntax?
- Can the search condition be completely evaluated at the remote data source? If the answer is no, examine the search condition instead.
- Does the search condition or SET clause reference the target table? Does the data source allow this syntax?
- Does the search condition or SET clause reference the target table with correlation? Does the data source allow this syntax?

Why is a positioned UPDATE statement not completely evaluated remotely?

This happens when the federated database chooses to evaluate the update expression locally before sending the UPDATE statement to the data source. This approach should not significantly affect performance.

Considerations:

- Can the SET clause be completely evaluated at the remote data source? In other words, does an update expression contain a function not supported by the remote data source?
- Does the SET clause involve a scalar subquery? Does the data source allow this syntax?

Why is a remote, searched DELETE statement not completely evaluated remotely?

You can check elements of the search condition to determine why a remote, searched DELETE statement is not completely evaluated remotely.

Considerations:

- Can the search condition be completely evaluated at the remote data source? If the answer is no, examine the search condition instead.
- Does the search condition reference the target table? Does the data source allow this syntax?
- Does the search condition reference the target table with correlation? Does the data source allow this syntax?

Data source upgrades and customization

When data sources are upgraded or customized, you need to update global catalog information.

The SQL compiler relies on information that is stored in the global catalog to provide it with the SQL capabilities of the data sources. This information periodically needs to be updated. The SQL capabilities of the data sources might change in new versions of the data sources. When data sources are upgraded or customized, update the global catalog information so that the SQL compiler is using the most current information.

Use SQL DDL statements, such as CREATE FUNCTION MAPPING and ALTER SERVER, to update the catalog.

Pushdown of predicates with function templates

In a federated system, each remote data source has its own functions. Most of these functions have semantically equivalent DB2 functions and have associated function mappings by default. However, some remote source functions might not have equivalent functions on the federated server. Consequently, only the remote data source can execute these functions. To write queries that use these functions, you must create a function template on the federated server.

A function template acts as a local description of the remote function. You create a function template with the CREATE FUNCTION statement by using the AS TEMPLATE clause. There is no executable code associated with the function template at the federated server. When the template is defined, you can use it to create a function mapping, which maps the function template to its remote counterpart. Then it is possible to refer to the function template in the SQL statements that are issued to the federated server and for the function to be evaluated at the data source.

The query optimizer makes cost-based decisions to determine where a predicate can be evaluated. When possible, the optimizer generates a plan to evaluate a predicate with a function template at the corresponding remote server. In some cases, it might not be possible for the optimizer to generate a plan that evaluates the function template at the data source. When this occurs you might receive an SQL0142N error with the following error message:

```
The SQL statement is not supported.
```

To avoid this error, the query can be rewritten to enable pushdown while maintaining the semantics of the original query.

For a function template to be pushed down, it must be defined with the DETERMINISTIC and NO EXTERNAL ACTION clauses.

Global optimization

The SQL compiler works in three phases, which help to produce an optimal access strategy for evaluating a query referencing a remote data source. These phases are pushdown analysis, global optimization, and remote SQL generation.

For a query submitted to the federated database, the access strategy might involve breaking down the original query into a set of query fragments and then combining the results of these query fragments.

Using the output of the pushdown analysis phase as a recommendation, the query optimizer decides where each operation is evaluated. An operation might be evaluated locally at the federated server or remotely at the data source. The decision is based on the output of the sophisticated cost model that the optimizer uses. This model determines:

- The cost to evaluate the operation
- The cost to transmit the data or messages between the federated server and the data sources

The goal of global optimization is to produce an access plan that optimizes the query operations on all data sources globally, across the federated system. An access plan that is globally optimal has the least overall cost of execution in a federated system. The remote SQL generation phase reverse translates the globally optimal plan into query fragments that are executed by individual data sources.

The SQL compiler has a knowledge base that contains characteristics of supported data sources and metadata about the data at those data sources. The optimizer does not generate SQL, query fragments, or plan hints that the remote data source cannot understand or accept.

Many factors can affect the output from global optimization and thus affect query performance. The key factors are server characteristics and nickname characteristics.

Relational and nonrelational wrappers differ in the details of how an access plan is produced, but the concept and final effect are the same.

Server characteristics affecting global optimization

When you create or alter a server definition, some of the options that you choose can affect query performance.

You provide the query optimizer with information about the data source server characteristics through the server option settings. The server option settings are part of the data source server definition. You can set server options in the CREATE SERVER statement, when you initially establish the server definition. Use the ALTER SERVER statement to add server options to an existing server definition. The server option settings are stored in the federated database global catalog.

These options are classified as location options (such as the data source computer name), security options (such as authentication information), and performance options (such as the CPU ratio).

The performance options help the optimizer determine if the evaluation of an operation can be done at a data source and if the evaluation of an operation on the data source makes execution faster. The server options affecting performance that might require your tuning are:

- CPU_RATIO
- IO_RATIO
- COMM_RATE
- COLLATING_SEQUENCE
- PLAN_HINTS
- VARCHAR_NO_TRAILING_BLANKS

Use caution when tuning the CPU_RATIO, IO_RATIO, or COMM_RATE server options as you can get unexpected errors if the cost calculation for a query causes overflows or underflows. In most cases, the default values for these options are sufficient. Typically, ensuring that the statistics about the objects referenced in your queries are correct is more important than tuning the values of these server options.

Relative ratio of CPU speed

This value indicates the ratio between the CPU speed of the federated server and the CPU speed of the server on which the remote data source is located.

This value is defined as the CPU speed of the federated server divided by the CPU speed of the server for the remote data source. For example, if the CPU speed for the federated server is twice as fast as the CPU speed for the remote server, then CPU_RATIO should be set to 2. If the CPU speed for the federated server is only one third as fast as the CPU speed for the remote server, then CPU_RATIO should be set to 0.33.

When you do not set the CPU ratio server option explicitly, the federated optimizer uses a default value of 1, which indicates that the federated CPU speed and the data source CPU speed are equal.

A low ratio indicates that the data source server CPU is faster than the federated server CPU. For low ratios, the optimizer will consider pushing-down operations that are CPU-intensive to the data source. A low ratio is a value that is less than 1.

Relative ratio of I/O speed

This value indicates the ratio between the I/O rate of the federated server and the I/O rate of the server on which the remote data source is located.

This value is defined as the I/O rate for the federated server divided by the I/O rate for the remote server. For example, if the I/O rate for the federated server is twice the I/O rate for the remote server, then IO_RATIO should be set to 2. But if the I/O rate for the federated server is half that of the remote server, then IO_RATIO should be set to 0.5.

When you do not set the I/O ratio server option explicitly, the federated optimizer uses a default value of 1, which indicates that the I/O rates of both the federated and remote servers are equal.

A low IO_RATIO of less than one indicates that the remote server has a higher I/O rate than the federated server. In this case, the optimizer will tend to favor pushing down I/O-intensive operations to the remote data source. A low ratio is a value that is less than 1.

Communication rate between the federated server and the data source

A low communication rate indicates slow network communication between the federated server and the data source.

The setting of the COMM_RATE server option determines the communication rate. The COMM_RATE represents the speed of the network connection between the data source server and the federated server. The rate is measured in megabytes per second. The default is 2 MBPS.

Lower communication rates encourage the query optimizer to reduce the number of messages sent to or from this data source. If the COMM_RATE server option is set to a very small number, the optimizer produces a query requiring minimal network traffic.

Data source collating sequence

The collating sequence that you choose might affect performance of the federated database. You can use the COLLATING_SEQUENCE server option to indicate if a data source collating sequence matches the local federated database collating sequence.

The federated server can push down order-dependent processing that involves character data to the data source, if the COLLATING_SEQUENCE server option indicates that the collating sequence of the data source and the federated database match. If a data source collating sequence does not match the federated database collating sequence, the optimizer considers data that is retrieved from this data source unordered. The federated database will retrieve the relevant data and perform all order-dependent processing on character data locally, which can slow down the query and affect performance.

Remote plan hints

Plan hints are statement fragments that provide extra information to data source optimizers.

Use the PLAN_HINTS server option to generate remote plan hints. This information can, for certain query types, improve query performance. The plan hints can help the data source optimizer decide whether to use an index, which index to use, or which table join sequence to use.

You should run some tests to determine if this server option will improve the performance of your queries.

You cannot code your own plan hints in a query.

If plan hints are enabled, the query sent to the data source contains additional information. For example, a statement sent to an Oracle optimizer with plan hints could look like this:

```
SELECT /*+ INDEX (table1, tlindex)*/
 col1
 FROM table1
```

The plan hint is the string `/*+ INDEX (table1, tlindex)*/`

Nickname characteristics affecting global optimization

There are several nickname-specific factors that can affect global optimization, including the index information and the global catalog statistics.

It is important that the index information and global catalog statistical data available to the SQL compiler is current.

Index specifications

The SQL compiler uses index information to optimize queries.

The index information for a data source table is only acquired when the nickname is created for that table. After the nickname is created, any changes to the index on that data source table are not updated on the federated server. When the remote index information changes, you can update the index information stored on the federated server by dropping the nickname for the table and creating the nickname again. Alternatively, if a new index is added for the data source table, you can define an index specification for the nickname on the federated server.

Index information is not gathered for nicknames on objects that do not have indexes such as views, synonyms, or nonrelational data source objects.

If an object that has a nickname defined for it does not have an index, you can create an index specification for it. Index specifications build an index definition in the global catalog. The index specification is not an actual index. Use the CREATE INDEX statement with the SPECIFICATION ONLY clause to create an index specification. The syntax for creating an index specification on a nickname is similar to the syntax for creating an index on a local table.

Consider creating index specifications when:

- A table acquires a new index.
- You create a nickname for a data source object that does not contain indexes such as a view or a synonym.

When you create an index specification (SPECIFICATION ONLY) on a nickname and specify that the index is unique, the federated database does not verify that the column values in the remote table are unique. If the remote column values are not unique, then queries against the nickname that include that index column might return incorrect data or result in errors.

Consider your needs before issuing CREATE INDEX...SPECIFICATION ONLY statements on a nickname for a data source view:

- If the remote view is a simple SELECT statement on a data source table with an index, creating an index specification on the nickname that matches the index on the data source table can significantly improve query performance.
- If an index specification is created for a remote view that is not a simple SELECT statement (for example, a view created by joining two tables), query performance might suffer.

For example, consider an index specification that is created for a remote view that is a join of two tables. The optimizer can choose that view as the inner element in a nested loop join. The query might have poor performance because the join will be evaluated several times. An alternative is to create nicknames for each of the tables referenced in the data source view and create a federated view that references both nicknames.

Global catalog statistics

The global catalog on the federated server contains statistical information that is used to optimize queries.

The federated server relies on statistics for data source objects for which nicknames have been defined to optimize queries that involve those nicknames. These statistics are retrieved from the data source when you create a nickname for a data source object using the CREATE NICKNAME statement. The federated database verifies the presence of the object at the data source, and then attempts to gather existing data source statistical data. Information useful to the optimizer is read from the data source catalogs and put into the global catalog on the federated server. Because some or all of the data source catalog information might be used by the optimizer, it is advisable to update statistics (using the data source command equivalent to RUNSTATS) at the data source before you create a nickname.

Catalog statistics describe the overall size of tables and views, and the range of values in associated columns. The information retrieved includes, but is not limited to:

- The number of rows in a nickname object
- The number of pages that a nickname occupies
- The number of distinct values in each column of a table
- The number of distinct values in columns of an index
- The highest/lowest values of a column

While the federated database can retrieve the statistical data held at a data source, it cannot automatically detect updates to existing statistical data at data sources. Furthermore, the federated database has no mechanism for handling object definition or structural changes to objects at the data sources (such as when a column is added to a table).

If the statistical data or structural characteristics for a remote object on which a nickname is defined change, you have three choices for updating the statistics:

- Run the equivalent of RUNSTATS at the data source. Then drop the current nickname and create the nickname again. This is the recommended method for updating statistics.

  An advantage of this method is that in addition to updated statistics, any information about new indexes or structural changes to the remote object will be reflected in the new nickname. A disadvantage of this method is that any views or packages based on the old nickname will be invalidated.
- Use the nickname statistics update facility in the DB2 Control Center. Alternatively, use the underlying stored procedure SYSPROC.NNSTAT(), available from the command line processor.

  The nickname statistics update facility (or SYSPROC.NNSTAT()) only updates the nickname statistics; it does not alter the nickname to match any structural changes to the remote object. For example, if the remote object has a new column, then updating nickname statistics does not add a column to the nickname.
- Manually update the statistics in the SYSSTAT.TABLES catalog view. Use this method only when you know that the statistical information on the remote data source is incorrect or incomplete.

Updating row changes

If a large number of rows are added to or deleted from an object at the data source, the federated database is not aware of these changes because the catalog statistics for the nickname continue to indicate the old number of rows.

However you might notice degradation in performance because the optimizer continues to make decisions based on nickname statistics information that is no longer accurate. After updating statistics for the remote object at the data source, you can update the statistics for the nickname to ensure that the optimizer can use accurate statistics when it generates and chooses access plans for processing queries on the data source.

Updating statistics when columns change

When there are structural changes to a data source object, for example, when a column is added to a table, you must complete several steps to update the statistics for that object in the federated database catalog.

About this task

If columns at the data source are added, deleted, or altered, you might notice incorrect results or receive an error message. For example, assume that the nickname *EUROSALES* refers to the *europe* table in a Sybase database. If a new column called *CZECH* is added to the table, the federated database will not be aware of the *CZECH* column. Queries that reference that column will result in an error message.

Procedure

To update the statistics for that object when column changes occur:

1. Run the utility on the data source that is equivalent to DB2 RUNSTATS. This will update the statistics stored in the data source catalog.
2. Drop the current nickname for the data source object using the DROP NICKNAME statement.
3. Re-create the nickname using the CREATE NICKNAME statement.

Analyzing global optimization

Detailed information about access plans, including some of the information that the global optimizer uses to choose the optimal plan, is kept in explain tables separate from the actual access plan itself.

This information allows for in-depth analysis of an access plan. The explain tables are accessible on all supported operating systems, and contain information for both static and dynamic SQL statements. You can access the explain tables using SQL statements. This allows for easy manipulation of the output, for comparison among different queries, or for comparisons of the same query over time.

There are multiple ways to get global access plan information from the Explain tables:

- You can use the Explain table format tool, db2exfmt, to present the information from the explain tables in a predefined format.
- You can also use the integrated Explain Facility in the DB2 Control Center in conjunction with Visual Explain to understand the access plan that is chosen for a particular SQL statement.

  Both dynamic and static SQL statements are explained by using the Explain Facility. One difference between Visual Explain and db2exfmt is that Visual Explain presents the information in a graphical format while db2exfmt presents the information in text format. The two methods provide the same level of detail.

- You can use the db2expln and dynexpln tools to understand the access plan that is chosen for a particular SQL statement.

  To fully understand the output of db2exfmt, Visual Explain, db2expln, or dynexpln you must understand:

  – The different SQL statements supported and the terminology related to those statements (such as predicates in a SELECT statement)
  – The purpose of a package (access plan)
  – The purpose and contents of the system catalog tables
  – Basic query processing operators such as joins, group-by, aggregation, and sorts

Understanding access plan optimization decisions

This section lists typical optimization questions, and areas you can investigate to improve performance.

Why isn't a join between two nicknames of the same data source being evaluated remotely?

You can check elements of the join operation, join predicates, and the number of rows in the result to evaluate why a join between two nicknames of the same data source is not evaluated remotely

Areas to examine include:

- Join operations. Can the data source support a join?
- Join predicates. Can the join predicate be evaluated at the remote data source?
- Number of rows in the join result. You can determine the number of rows with Visual Explain. Does the join produce a much larger set of rows than the two nicknames combined? Do the numbers match reality? If the answer is no, consider updating the nickname statistics with the SYSPROC.NNSTAT() stored procedure.

Why isn't the GROUP BY operator being evaluated remotely?

Areas to examine include:

- Operator syntax. Verify that the operator can be evaluated at the remote data source.
- Number of rows. Check the estimated number of rows in the GROUP BY operator input and output using Visual Explain. Are these two numbers very close? If the answer is yes, the optimizer considers it more efficient to evaluate this GROUP BY locally. Also, do these two numbers reflect reality? If the answer is no, consider updating the nickname statistics using the SYSPROC.NNSTAT() stored procedure.

Why is the statement not being completely evaluated remotely?

The federated server seeks to ensure that query semantics and results obtained for federated queries are exactly the same as if they had been completely evaluated by DB2 Database for Linux, UNIX, and Windows. The pushdown analysis phase of the query compiler decides whether pushing down processing to remote sources will maintain DB2 semantics. Thus, federated query operations can be safely pushed down only if corresponding operations at the remote source have the same meaning and result. The most common reason for failure to completely push down processing of a query to a single remote source is the existence of subtle differences in functionality between the federated server and the remote source for one or more operations in the query.

The optimizer performs cost-based optimization. Even if pushdown analysis indicates that every operator can be evaluated at the remote data source, the optimizer still relies on its cost estimate to generate a globally optimal plan. There are many factors that contribute to the decision to choose that plan. Suppose that the remote data source can process every operation in the original query. However, its CPU speed is much slower than the CPU speed of the federated server. It might turn out to be more beneficial to perform the operations at the federated server instead. If the desired performance is not achieved, verify the server statistics in the SYSSTAT.SERVEROPTIONS catalog table.

Why does a plan generated by the optimizer and completely evaluated remotely, have much worse performance than the original query executed directly at the remote data source?

Areas to examine include:

- The remote SQL statement generated by the query optimizer. In addition to the replacement of nicknames by corresponding remote table names, the generated remote SQL statement typically differs from the original federated statement in the following ways:
  - The ordering of predicates in the query might have changed.
  - Predicates found in the original query might have been removed, replaced by equivalent ones, or augmented by additional predicates.
  - Subqueries might have been rewritten as joins.
  - Additional functions that do conversion or string truncation might have been added to maintain DB2 semantics

  With the exception of the last item listed above, these changes usually have a favorable impact on performance. However, in a few cases, the changes might cause the remote query optimizer to generate a different (and slower) plan than it would have for the original query

  A good query optimizer should not be sensitive to the predicate ordering of a query. Unfortunately, not all DBMS optimizers are identical. It is likely that the optimizer at the remote data source will generate a different plan based on the input predicate ordering. If this is true, this is a problem inherent in the remote optimizer. Consider either modifying the predicate ordering or contacting the service organization of the remote data source for assistance.

  Also, check for predicate replacements. A good query optimizer should not be sensitive to equivalent predicate replacements. It is possible that the optimizer at the remote data source will generate a different plan based on the input predicate. For example, some optimizers cannot generate transitive closure statements for predicates.
- The number of returned rows. You can get this number from Visual Explain. If the query returns a large number of rows, network traffic is a potential bottleneck.
- Additional functions. Does the remote SQL statement contain more functions than the original query? Some of the extra functions might be generated to convert data types. Ensure that they are necessary.

System monitor elements affecting performance

The federated database system monitor gathers statistical information regarding the current state of the database manager, and activity information such as counters and other measurements of database processing.

In a federated system, you can use the database system monitor to gather information about database activity, system performance, and application performance.

The Timestamp monitor switch is used to track the response times of interactions that the federated database has with a data source. The federated data elements tracked by the timestamp switch are:

- Create nickname response time
- Delete response time
- Insert response time
- Pass-through time
- Query response time
- Remote lock time
- Stored procedure time
- Update response time

The default setting for the Timestamp monitor switch is ON.

Recommendation: You can increase performance by changing the setting for the Timestamp monitor switch to OFF for all applications. If one application has the Timestamp switch set to ON, the system will continue to collect the response times. Therefore, you will not increase performance by turning off the timestamp switch for only some of your applications.

Turning off the switch does have other implications.

- Turning off the Timestamp monitor switch for all applications requires that you stop and restart the DB2 instance to implement the change.
- Turning off the Timestamp monitor switch disables the gathering of timestamp information for both federated and non-federated applications. The local database will not receive timestamp information either.

If you need timestamp information for local non-federated applications, then you should not turn off the Timestamp monitor switch.

You can set the timestamp switch to OFF for all applications by using this command:

```
update dbm cfg using dft_mon_timestamp off
```

Then issue:

```
db2stop
db2start
```

Stopping and starting the federated server will ensure that the switch is off for all applications.

Specific information about each of the elements tracked by the timestamp switch is discussed in a separate topic.

Chapter 17. Parallelism with queries that reference nicknames

Queries that contain nicknames can participate in three types of intra-query parallelism.

The three types of intra-query parallelism are:

- Intrapartition query parallelism on single partition, multiprocessor configurations
- Interpartition query parallelism on multiple partition configurations
- Mixed query parallelism that consists of both intrapartition and interpartition parallelism where each partition runs on an SMP computer

Intrapartition parallelism with queries that reference nicknames

Intrapartition parallelism refers to the process of dividing a query into multiple concurrent parts that run in parallel by multiple processes on a single database partition.

In federated queries, the part of a query that involves local data can run in parallel while the part that involves nicknames runs serially, using a single agent process.

When multiple processors can work on the local portions of the query, the performance of queries that reference local tables and nicknames can improve.

The DFT_DEGREE database configuration parameter and the CURRENT DEGREE special register control the degree of intrapartition parallelism.

Enabling intrapartition parallelism with queries that reference nicknames

For queries that reference local tables and nicknames in a multiprocessor environment, you can enable intrapartition parallelism. The federated server can then process the local tables in parallel.

Restrictions

The federated system can process only the portion of a query that references local tables in parallel. The coordinator partition processes all operations on the remote portion of a query in serial.

Procedure

To enable intrapartition parallelism:

1. Set the INTRA_PARALLEL database configuration parameter to YES.
2. Set the MAX_QUERYDEGREE database configuration parameter to a value greater than 1.
3. Set the DFT_DEGREE database configuration parameter to a value greater than 1, or set the special register CURRENT DEGREE. If you set the DFT_DEGREE parameter to ANY, the default level of intrapartition parallelism equals the number of processors on the computer.

Intrapartition parallelism with queries that reference nicknames - examples of access plans

You can use the DB2 Explain facility to view the access plan that the optimizer uses during query processing. The following examples show how the optimizer accesses nickname data in an intrapartition parallelism environment.

Example 1: Without parallelism support

In this example, the federated server processes the join of the local table, ORDERS, and nickname, ITEMS, serially. No intrapartition parallelism is used.

```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```

```
            RETURN
           (   1)
             |
           HSJOIN
           (   2)
         /----+---\
     TBSCAN         SHIP
     (   3)        (   4)
       |             |
   TABLE: NEWTON  NICKNM: NEWTON
     ORDERS         ITEMS
```

Example 2: With parallelism support

In this example of a join, the query can run faster by having the local table read in parallel before the serial join with the nickname. The LTQ operator indicates where parallelism is introduced into the plan.

```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```

```
            RETURN
            (   1)
              |
            HSJOIN
            (   2)
          /----+---\
      LTQ           SHIP
     (   3)        (   5)
       |             |
     TBSCAN   NICKNM: NEWTON
     (   4)        ITEMS
       |
   TABLE: NEWTON
    ORDERS
```

Example 3: Intrapartition parallelism with aggregation

In this example, the database aggregates the local table data in parallel in the partition, improving the execution of the aggregation. The join of the local table and the nickname occurs serially.

```
SELECT *
FROM ITEMS A
WHERE ID =
   (SELECT MAX(ID)
```

```
FROM ORDERS
WHERE NUMBER = 10)


         RETURN
         (   1)
            |
         NLJOIN
         (   2)
        /----+---\
    GRPBY         SHIP
   (   3)        (   7)
     |             |
    LTQ      NICKNM: NEWTON
   (   4)         ITEMS
     |
     1
   GRPBY
   (   5)
     |
    TBSCAN
   (   6)
     |
TABLE: NEWTON
   ORDERS
```

Interpartition parallelism with queries that reference nicknames

Interpartition parallelism refers to the process of dividing a single query into multiple parts that run in parallel on different partitions of a partitioned database.

In queries that reference local and remote data, the federated server can distribute the remote data to each of the local partitions. Figure 12 on page 210 and Figure 13 on page 210 show the concept of interpartition parallelism that involves local and remote data sources.

Figure 12 on page 210 shows how this type of query is processed without interpartition parallelism. The remote nickname data and the local partitioned data are processed serially at the coordinator partition. This technique does not exploit the parallel power of the database partitions because most processing is performed on a single partition. If data volumes are very large, this technique is likely to result in long-running queries.

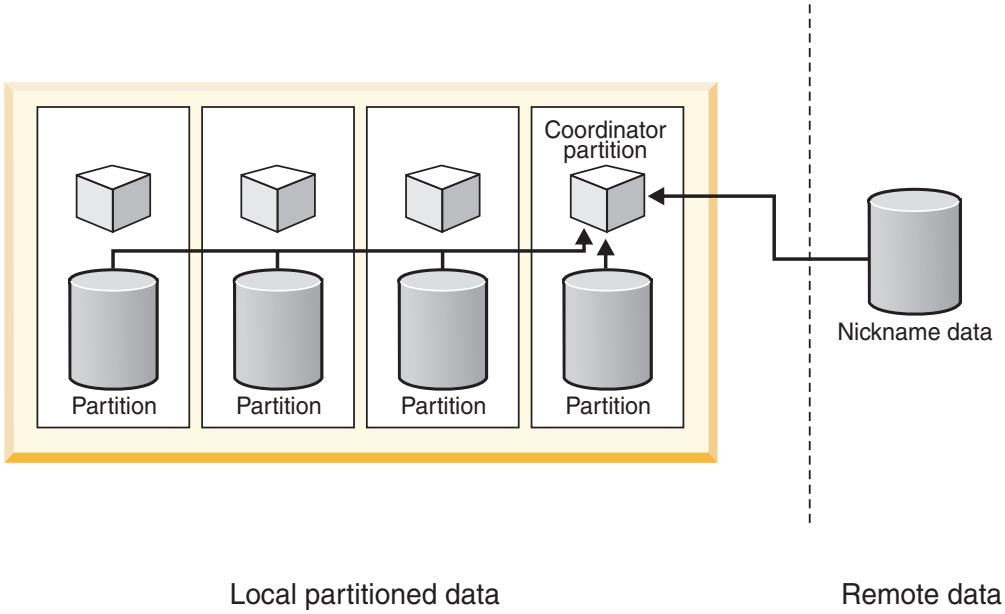


Figure 12. Query without interpartition parallelism of local and remote data sources

Figure 13 shows how processing occurs when the optimizer distributes the nickname data to the partitions. The coordinator partition fetches the nickname data and distributes the data to the database partitions for parallel processing. When parallel processing is complete, the results are sent back to the coordinator partition for final processing before they are returned to the application.

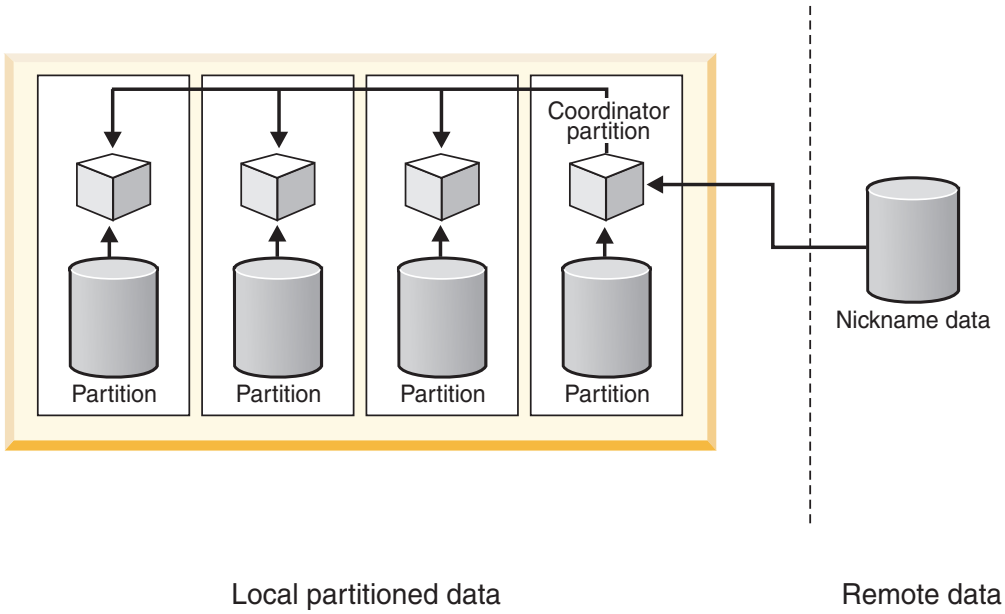


Figure 13. Query with interpartition parallelism of local and remote data sources

Figure 14 on page 211 and Figure 15 on page 211 show the concept of interpartition parallelism that involves only remote data sources.

Figure 14 shows serial processing of the remote nickname data at the coordinator partition. The coordinator partition, which also acts as the federated server, retrieves the nickname data and process it serially.

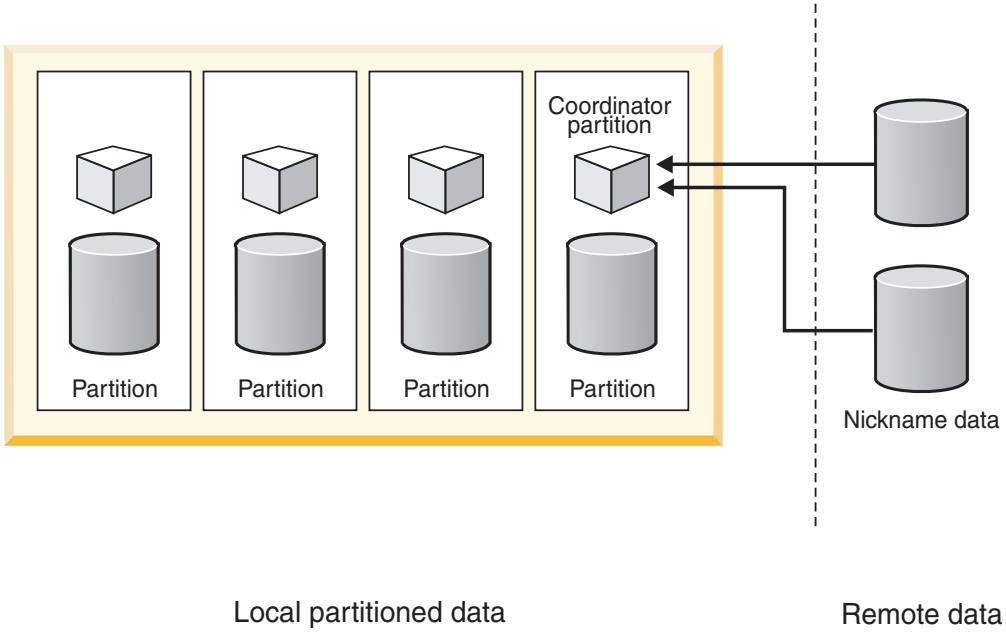


Figure 14. Query without interpartition parallelism that references remote data sources only.

Figure 15 shows the coordinator partition distributing the data across a computational partition group. Computational partition groups allow the optimizer to generate access plans that distribute nickname data across the partitions of a partitioned database server for processing in parallel.

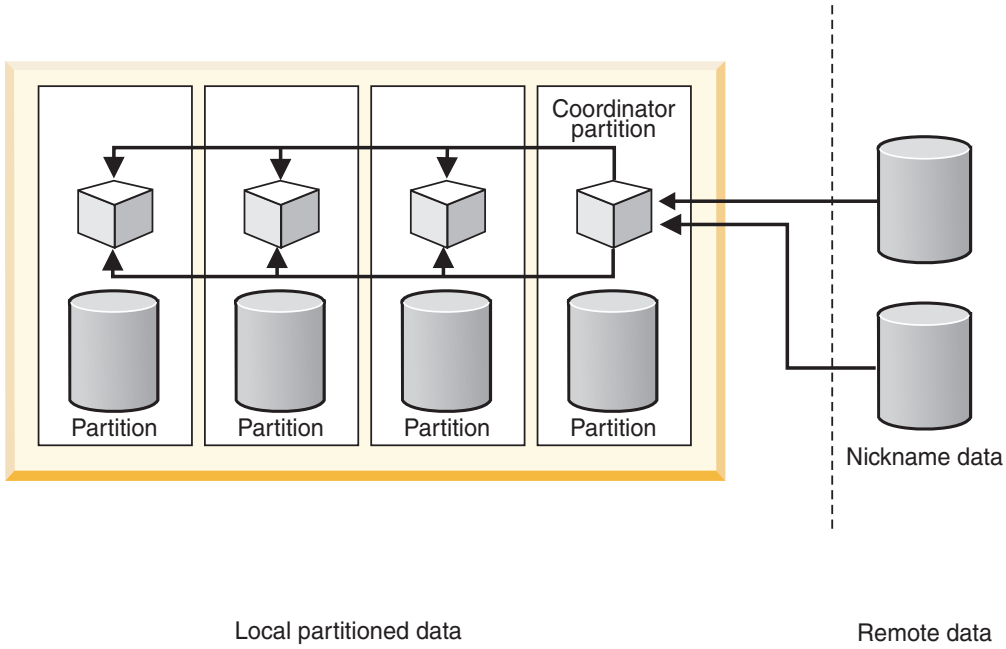


Figure 15. Query with interpartition parallelism that references remote data sources only.

Regardless of the plan that the optimizer chooses, access to the nickname data always occurs serially from the coordinator partition.

Enabling interpartition parallelism with queries that reference nicknames

You can configure a partitioned federated server so that queries involving nicknames can potentially run in parallel on multiple partitions. Parallel execution might significantly reduce the elapsed time of federated queries in a partitioned environment.

Restrictions

Only parts of a query that reference nicknames that use fenced wrappers can run in parallel. Any parts of a query that reference nicknames that use trusted wrappers cannot run in parallel.

About this task

In a partitioned database environment, federated queries can run in parallel under the following conditions:

- They involve a combination of nicknames that are defined by using a fenced wrapper and local partitioned tables
- They involve nicknames defined using a fenced wrapper, and a computational partition group is defined.

You do not need to set any database parameters or database configuration parameters in a partitioned environment to make interpartition parallelism available for federated queries.

Procedure

To enable interpartition parallelism:

1. Issue the CREATE WRAPPER or ALTER WRAPPER statement with the DB2_FENCED option set to Y.
2. **Optional:** Set up a computational partition group, if you are interested in enabling parallelism for portions of queries that involve only nicknames. If queries involve a combination of nicknames and local partitioned tables, you do not need to set up a computational partition group.

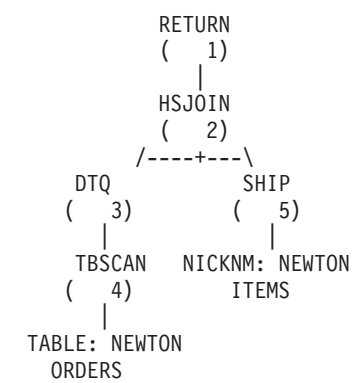
Interpartition parallelism with queries that reference nicknames - examples of access plans

You can use the DB2 Explain facility to view the access plan that the optimizer uses during query processing. The following examples shows how the optimizer accesses nickname data in an interpartition parallelism environment.

Example 1: Trusted mode

In this example, the nickname uses a trusted wrapper. The database serially performs the join between the local table and the nickname at the coordinator partition. The database brings the local data, which is distributed over two partitions, to the coordinator partition. The federated server then joins the local data with the nickname data. The database serially joins nicknames that are defined by using a trusted wrapper at the coordinator partition. The database cannot distribute the data across multiple partitions to create a parallel join.

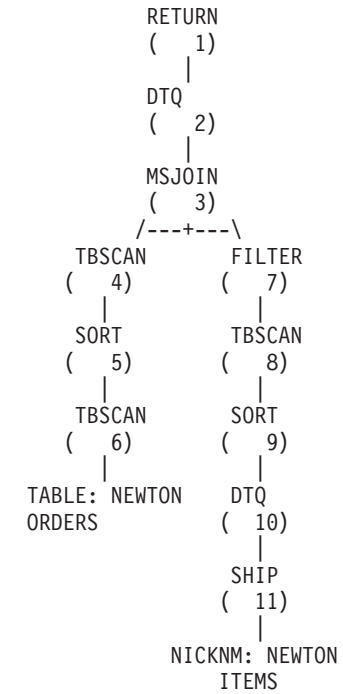
```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```



Example 2: Fenced mode

In this example, the nickname uses a fenced wrapper. The federated server distributes the nickname data to the other partitions and performs the join with the local data in parallel. The DTQ (Distributed Table Queue) operator above the SHIP indicates that the nickname data is distributed to the local partitions using hash partitioning to achieve a co-located parallel join. In a co-located parallel join, nickname data is distributed to the local partitions in such a way that matching nickname and local data for the join will always be located on the same partition.

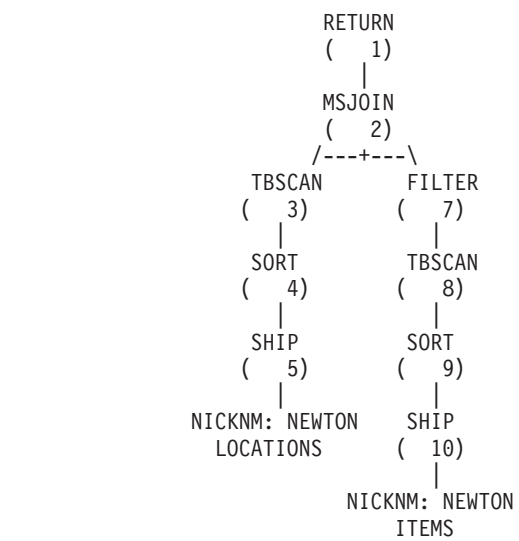
```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```



Example 3: Fenced mode without a computational partition group

In this example, the two nicknames use a fenced wrapper, and a computational partition group is not defined. The federated server performs the join at the coordinator partition. The federated server does not distribute the data to the other partitions for processing. The lack of TQ operators above any of the SHIP operators indicates that the nickname data is not distributed across the partitions.

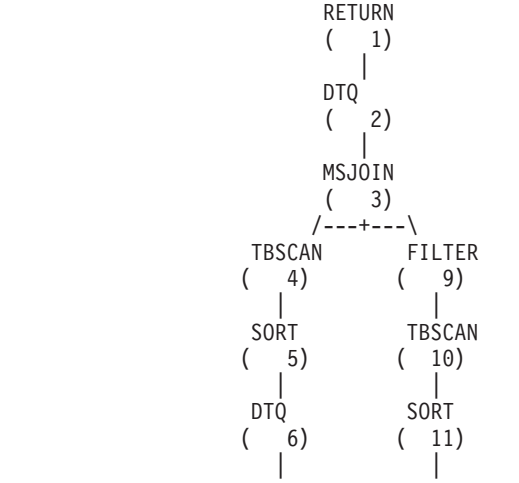
```
SELECT *
FROM ITEMS A, LOCATIONS B
WHERE A.ID1 = B.ID1
```

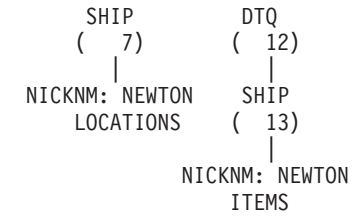


Example 4: Fenced mode with a computational partition group

In this example, the nicknames use fenced wrappers, and a computational partition group is defined. In this case, the optimizer selects a plan that distributes the data from the coordinator partition to the other partitions in the computational partition group. The DTQ operators above both nicknames hash-partition the incoming remote data so that matching join keys are located on the same partition of the computational partition group. The join takes place in parallel on each partition, and the results are then collected at the coordinator partition.

```
SELECT *
FROM ITEMS A, LOCATIONS B
WHERE A.ID = B.ID
```





Computational partition groups

A computational partition group defines a set of database partitions that the optimizer can use to dynamically redistribute nickname data. A computational partition group is for the portions of queries that involve only nicknames.

The coordinator partition fetches nickname data serially and then redistributes the data across the partitions in the computational partition group, at which point parallel processing occurs. The use of computational partition groups by the optimizer often results in performance improvements, particularly when large amounts of nickname data are involved or the queries are complex.

A computational partition group is a database partition group, other than IBMCATNODEGROUP, that is specified in the system catalog, SYSCAT.DBPARTITIONGROUPS.

You use the DB2_COMPPARTITIONGROUP registry variable to specify the computational partition group.

Defining a computational partition group

Defining a computational partition group enables the optimizer to use a plan that distributes nickname data to the partitions of the computational partition group. You define a computational partition group to enable interpartition query parallelism for queries or parts of queries that reference only nicknames.

Before you begin

All partition groups used to represent the computational partition group on all the databases in the instance must have the same name. You can define these partition groups differently in each database, but they must have the same name. For example, three databases called DB1, DB2, and DB3 define a computational partition group that contains different nodes:

- DB1: CPG contains nodes 1, 2, 3, and 4
- DB2: CPG contains nodes 49, 50, and 53
- DB3: CPG contains nodes 78 and 96

You can set the db2set variable to the name CPG. The name CPG is common to all databases, but the contents of the CPG are different for each database.

Restrictions

The optimizer uses computational partition groups for only the parts of a query that reference nicknames without referencing local data.

Procedure

To define a computational partition group, issue the following command at the DB2 command line.

```
db2set DB2_COMPPARTITIONGROUP=partitiongroup_name
```

partitiongroup_name is the name of the partition group that you want to define as the computational partition group. The partition group must already be defined.

The following example shows how to define the computational partition group, FINANCE3, using the DB2_COMPPARTITIONGROUP registry variable.

```
db2set DB2_COMPPARTITIONGROUP=FINANCE3
```

Interpartition parallelism with queries that reference nicknames - performance expectations

For queries that reference a combination of local partitioned tables and nicknames, the optimizer can choose an execution plan that redistributes nickname data across appropriate partitions.

Redistribution plans can make queries run faster if the amount of nickname data in the join is smaller than the amount of local partitioned data. If the amount of nickname data in the join is considerably larger than the local data, then a parallel plan with redistribution of the nickname data is unlikely to be used. If the optimizer does not choose a parallel plan, the federated server performs the joins serially between nicknames and local tables at the coordinator partition.

For joins between two nicknames, an execution plan that distributes the data among all partitions of a computational partition group can be beneficial if it involves a large amount of data. The advantage of processing the large join in parallel offsets the additional cost of redistributing the data across multiple partitions. If the amount of nickname data is relatively small, the join is not expensive enough to merit the extra cost of redistributing the data across partitions. In general, the optimizer chooses computational partition group plans if the nicknames involved are large; otherwise, the federated server joins the nicknames serially at the coordinator partition.

Mixed parallelism with queries that reference nicknames

For queries that contain local tables and nicknames in a partitioned environment, the optimizer can use both intrapartition and interpartition parallelism. Interpartition parallelism is an option for the optimizer in a partitioned environment. Intrapartition parallelism is an option, if it has been enabled in the database configuration or database manager configuration.

For interpartition parallelism, the federated server can distribute remote data among partitions and process data in parallel within each partition.

For intrapartition parallelism, multiple subagent processes within a partition are used to process local data in parallel.

Enabling mixed parallelism with queries that reference nicknames

You can improve the performance of queries that reference local and remote data by the use of intrapartition and interpartition parallelism.

Procedure

To enable interpartition parallelism on a partitioned federated server:

1. Issue the CREATE WRAPPER or ALTER WRAPPER statement with the DB2_FENCED option set to Y.
2. **Optional:** Set up a computational partition group to enable parallelism for nickname-only joins.

To enable intrapartition parallelism on a federated server:

1. Set the MAX_QUERYDEGREE database configuration parameter to a value greater than 1.
2. Set the DFT_DEGREE database configuration parameter to a value greater than 1, or you must set the special register CURRENT DEGREE. If you set the DFT_DEGREE parameter to ANY, the default level of intrapartition parallelism equals the number of SMP processors on the computer.

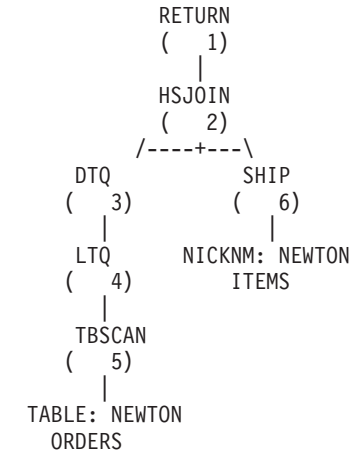
Mixed parallelism with queries that reference nicknames - examples of access plans

You can use the DB2 Explain facility to view the access plan that the optimizer uses during query processing. The following examples shows how the optimizer accesses nickname data in an environment that uses both intrapartition parallelism and interpartition parallelism.

Example 1: Trusted mode

The following example shows a join between a local table and a nickname in trusted mode. The federated server processes the local data in parallel in each partition before it joins the local data with the nickname data at the coordinator partition. The federated server does not process the nickname data in parallel across the partitions or the processors on any given partition.

```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```

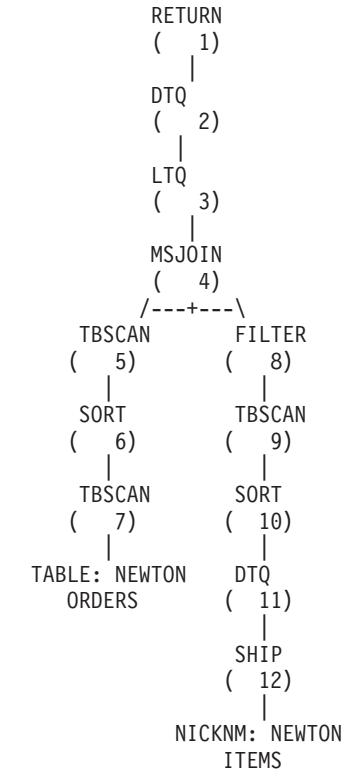


Example 2: Fenced mode

The following example shows a join between a local partitioned table and a nickname in fenced mode. The federated server serially fetches the nickname data onto the coordinator partition and then distributes this data to the other partitions in the database. The data is then joined in parallel with the local data across all appropriate database partitions. Within each partition, multiple subagents are reading the local table and joining to the nickname data. This operation is

intrapartition parallelism, identified in the plan by the LTQ operator. The result is returned to the coordinator partition for final processing and returned to the application.

```
SELECT *
FROM ORDERS A, ITEMS B
WHERE A.ID1 = B.ID1 AND B.ITEM = 3
```



Chapter 18. Asynchronous processing of federated queries

Asynchrony is a method of improving query performance by running multiple parts of an access plan concurrently to reduce the elapsed time for a given query.

In a federated system, data is distributed across systems at multiple data sources, and each system has its own resources. Asynchrony overlaps the operations that use those resources so that multiple systems remain active at the same time. Overlapping operations enables multiple parts of an access plan to run concurrently rather than serially.

Complex queries that involve time-consuming operations on remote data sources can benefit from asynchrony. Asynchrony enables the following types of operations to take place concurrently:

- Two or more operations on remote data sources
- Operations on the federated server and at least one remote data source

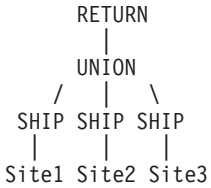
As query operations consume resources, asynchrony can benefit systems in which resources are idle or when only one of the data sources, or the federated system, is doing work at any point in time.

Asynchronous processing of federated queries - examples

Examples of queries with a union and with a merge join illustrate the difference between query operations with and without asynchronous processing.

Example: Query with a union operation

A simple query performs a union operation on data from three different data sources. The computation required to generate the data on each data source is time-consuming. The access plan looks like this:

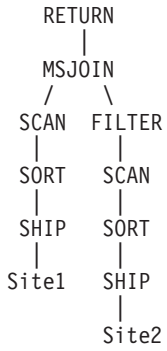


Without asynchrony, the union operation reads data from the query branches one branch at a time, from left to right. When the data from the Site1 server is read, Site2 server and Site3 server are idle. For this example, each branch of the union takes about two hours to return the result rows from one site. The total execution time of the three branches is approximately the sum of the time it takes to process each branch, in this example, about six hours.

With asynchrony, each branch of the union starts to process at the same time, and the three remote servers are active concurrently. The run time of the query is roughly equivalent to the run time of the slowest branch of the union. In this example, the run time is reduced to approximately two hours (about 66% faster) in comparison to six hours without asynchrony.

Example: Query with a merge join operation

A query that joins data from two different data sources uses a merge join operation (MSJOIN). The optimizer access plan looks like this:



Without asynchrony, the merge join operator first processes the outer (left) branch and does not process the inner (right) branch until the left branch starts to return rows. For this example, each branch executes a complex query and therefore takes a long time to execute. The approximate total time to execute the merge join is the sum of the time it takes to execute each branch.

With asynchrony, both branches of the merge join start at the same time, thus reducing the overall execution time of the query.

Asynchrony optimization

The query optimizer makes decisions about the asynchronous processing of remote operations in a query execution plan. *Asynchrony optimization* is the process by which the optimizer analyzes an existing query execution plan and looks for opportunities to allow remote operations to execute concurrently.

Access plans without asynchrony

In an execution plan, the SHIP or RPD operator defines a portion of the plan that is executed at a remote data source.

Without asynchrony, the SHIP or RPD operator becomes active and initiates remote processing only when its data is required by other operators located above the SHIP or RPD operator in the execution plan.

Access plans optimized for asynchrony

The optimizer can make the remote operations that the SHIP or RPD operator defines execute asynchronously.

In an asynchronous operation, a table queue (TQ) operator is inserted directly above the SHIP or RPD operator in the execution plan. The TQ operator defines a portion of the plan, called a *subplan*. A separate process or thread, with its own memory, runs the subplan. A subplan initiates immediately when the query starts.

You can think of the TQ operator as a pipe between the SHIP or RPD operator (producer of data) and the operator above it (consumer of data) in the plan. This pipe decouples the execution of the SHIP in the subplan below it from the main plan, and allows the asynchronous exchange of data between the two plan sections.

A TQ operator that appears directly above a SHIP or RPD operator in the plan enables the remote operations that the SHIP or RPD operator define to initiate at

the beginning of the query and to deliver results to the federated server asynchronously. When asynchrony is beneficial for a given remote operation, the optimizer places a TQ operator directly above the corresponding SHIP or RPD operator in the plan.

TQ operators occur in execution plans for different purposes. A TQ operator usually denotes parallel operations in partitioned databases or in databases enabled for intrapartition parallelism. Another type of TQ operator, that enables asynchronous execution of a subplan, is called an asynchrony TQ (ATQ).

The optimizer makes a given SHIP or RPD operator asynchronous when:

- Query performance will improve
- The number of ATQs is below the per-server and per-query limits
- The operator is not already asynchronous due to the use of another optimization technique
- Restrictions on asynchrony are not violated.
- The semantics of the query do not change

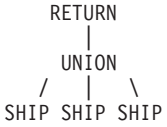
Access plans - examples

Examples of access plans illustrate the difference between plan execution with and without asynchrony optimization.

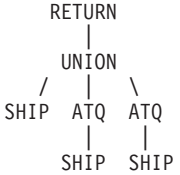
The first two examples show how the union and merge join plans in "Asynchronous processing of federated queries - examples" on page 219 look when asynchrony is enabled.

For simplicity, the examples show plans with SHIP operators only. Asynchrony optimization transforms the plan the same way for RPD operators as for SHIP operators. SHIP and RPD operators are interchangeable, unless otherwise noted.

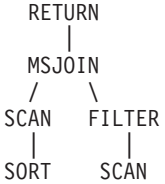
Example 1a: Plan without asynchrony

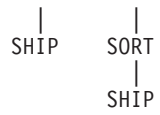


Example 1b: Plan with asynchrony

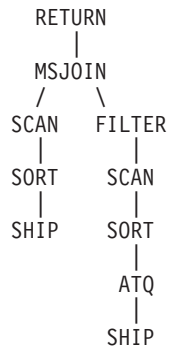


Example 2a: Plan without asynchrony

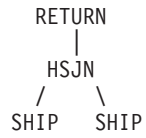




Example 2b: Plan with asynchrony

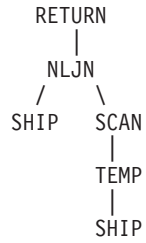


Example 3a: Plan without asynchrony

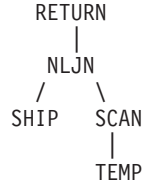


Example 3b: Plan with asynchrony

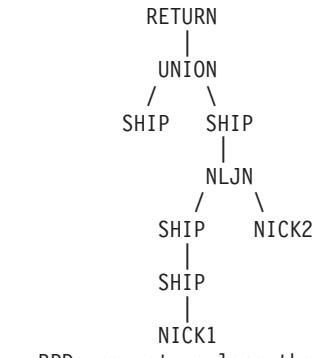
Example 4a: Plan without asynchrony



Example 4b: Plan with asynchrony

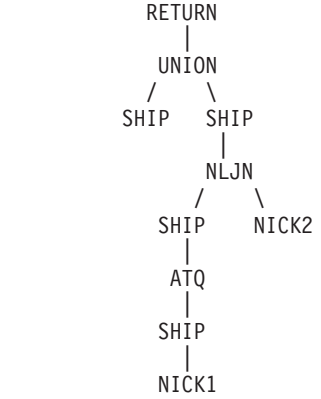


Example 5a: Plan without asynchrony

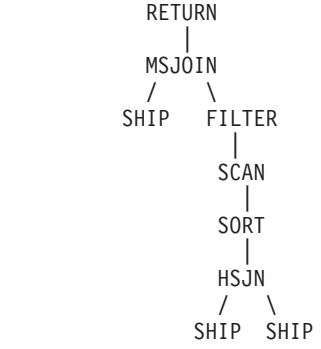


RPDs cannot replace the SHIP-SHIP pair in this plan.

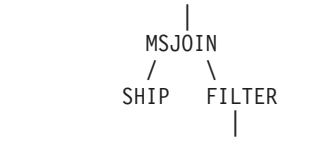
Example 5b: Plan with asynchrony

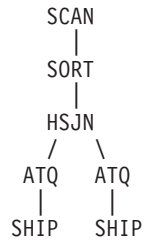


Example 6a: Plan without asynchrony

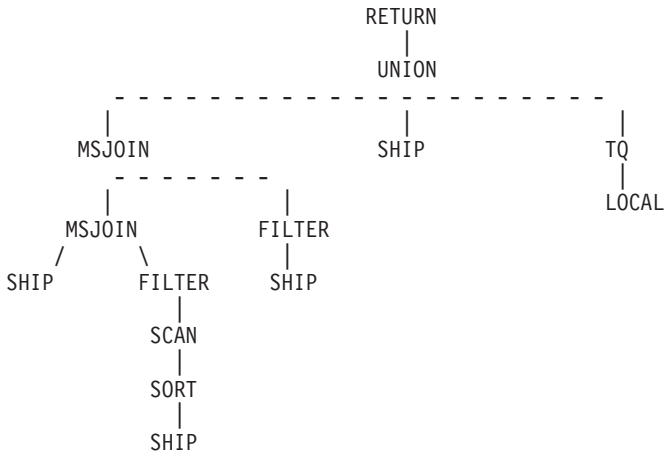


Example 6b: Plan with asynchrony

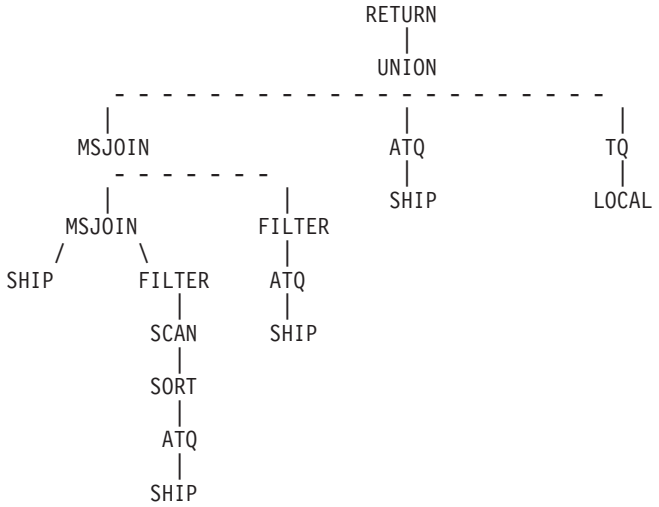




Example 7a: Plan without asynchrony



Example 7b: Plan with asynchrony



Controlling resource consumption

In addition to enabling asynchronous execution of a remote query, the ATQ operator affects the federated server and remote data sources.

Because each ATQ operator creates a new process, and consumes some memory (for buffering), inserting numerous ATQs into an execution plan might use too many system resources on the federated server. In addition, if several SHIP or RPD operators in a query execute on a particular remote data source, making several of the operators asynchronous opens multiple concurrent cursors on that data source and can produce an unacceptable load.

To control resource consumption on the federated system, or on the data sources, you can set configuration parameters. The parameters define limits on the total number of ATQs allowed within a query and on the number of ATQs allowed for each server within a query.

Enabling asynchrony optimization

To enable asynchrony optimization, you specify the number of asynchrony TQ operators for a given query and set a server option for the data source.

Restrictions

Asynchrony optimization requires:

- A federated system with the database partitioning feature (DPF) that includes more than one logical database partition.
- Access to data sources through fenced wrappers.

This optimization does not support these objects:

- Nicknames on a data source that are accessed through a trusted wrapper.
- Queries with insert, update, or delete operations.

Procedure

To enable asynchronous processing:

1. Set one or more of the following parameters:
   - Set the FEDERATED_ASYNC database manager configuration parameter to a value between 0 and MAXAGENTS/4 (MAXAGENTS divided by 4) or to ANY. MAXAGENTS is a database configuration parameter that specifies the maximum number of ATQs allowed per query. The value ANY allows the optimizer to determine the number of ATQs for a given access plan. The default is 0.
   - Optionally, set the FEDERATED_ASYNCHRONY bind option for the static statements in the package to override the configuration parameter setting for the query. The default is 0.
   - Optionally, set the CURRENT FEDERATED ASYNCHRONY special register to dynamically override the database manager configuration parameter setting and the bind option for the query.

   These parameters form the following hierarchy:

   a. Special register
   b. Bind option
   c. Database manager configuration parameter

   The special register value, if specified, takes precedence over the bind option, which in turn takes precedence over the database manager configuration parameter.
2. Set the DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option to a numeric value. This server option specifies the maximum number of asynchronous requests that the server allows for the query. The default is 1. Therefore, one SHIP operator or one RPD operator that belongs to a server is considered for asynchrony in a query.

   The range for the server option is -1 to (maxagents / 4).

- -1 = ANY. The optimizer determines the number of ATQs to use for the SHIPs or RPDs for a given data source.
- 0 = Asynchrony is disabled for the SHIPs or RPDs for this server
- 1 is the default.

Setting the server option to a numeric value other than -1 indicates the maximum ATQs to use for the SHIPs or RPDs for the server provided that other criteria are satisfied.

The default for the ODBC data source only is 0. Asynchrony requires multiple cursors per connection. This value should be 0 for the ODBC wrapper unless the data source supports multiple cursors per connection.

Database manager configuration parameter: FEDERATED_ASYNC

This parameter determines the maximum number of asynchrony table queues (ATQs) in the access plan that the federated server supports.

Type: Numeric

You can assign a number to this parameter or use the keyword 'ANY'.

Default: 0

Asynchrony is disabled by default.

Range: 0 to (MAXAGENTS / 4) inclusive, or ANY

MAXAGENTS is the total number of agents (coordinator agents and subagents) in the system. You specify the total number of agents in the MAXAGENTS database manager configuration parameter.

When ANY is specified, the optimizer determines the number of ATQs for the access plan. The optimizer assigns an ATQ to all eligible SHIP or remote pushdown operators in the plan. The value that is specified for DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option limits the number of asynchronous requests.

Usage notes

The FEDERATED_ASYNC configuration parameter supplies the default or starting value for the special register and the bind option. You can override the value of FEDERATED_ASYNC configuration parameter by setting the value of the special register, bind option, or prep option to a higher or a lower number.

If the special register or the bind option do not override the FEDERATED_ASYNC configuration parameter, the value of the parameter determines the maximum number of ATQs in the access plan that the federated server allows. If the special register or the bind option overrides this parameter, the value of special register or bind option determines the maximum number of ATQs in the plan.

Any changes to the FEDERATED_ASYNC configuration parameter affect dynamic statements as soon as the current unit of work commits. Subsequent dynamic statements recognize the new value automatically. A restart of the federated

database is not needed. Embedded SQL packages are not invalidated nor implicitly rebound when the value of the FEDERATED_ASYNC configuration parameter changes.

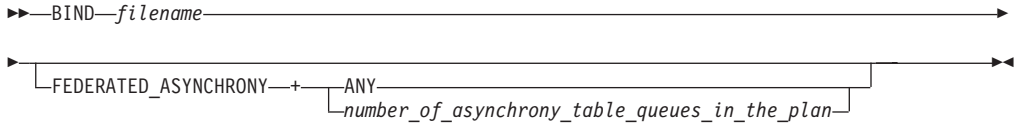
If you want the new value of the FEDERATED_ASYNC configuration parameter to affect static SQL statements, you need to rebind the package.

Bind and precompile options: FEDERATED_ASYNCHRONY

The FEDERATED_ASYNCHRONY option is used with the BIND and PRECOMPILE commands to specify the maximum number of asynchrony table queues (ATQs).

Bind option: FEDERATED_ASYNCHRONY

The FEDERATED_ASYNCHRONY bind option specifies the maximum number of ATQs that the federated server supports in the access plan for programs that use embedded SQL. You use the FEDERATED_ASYNCHRONY option in the BIND command.



Type: Numeric

You can assign a number to this parameter or use the keyword 'ANY'.

Range: 0 to (MAXAGENTS / 4) inclusive, or ANY

MAXAGENTS is the total number of agents (coordinator agents and subagents) in the system. You specify the total number of agents in the MAXAGENTS database manager configuration parameter.

When ANY is specified, the optimizer determines the number of ATQs for the access plan. The optimizer assigns an ATQ to all eligible SHIP or remote pushdown operators in the plan. The value that is specified for DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option limits the number of asynchronous requests.

Usage notes

For an embedded SQL program, if the bind option is not explicitly specified the static statements in the package are bound using the FEDERATED_ASYNC configuration parameter. If the FEDERATED_ASYNCHRONY bind option is specified explicitly, that value is used for binding the packages and is also the initial value of the special register. Otherwise, the value of the database manager configuration parameter is used as the initial value of the special register. The FEDERATED_ASYNCHRONY bind option influences dynamic SQL only when it is explicitly set.

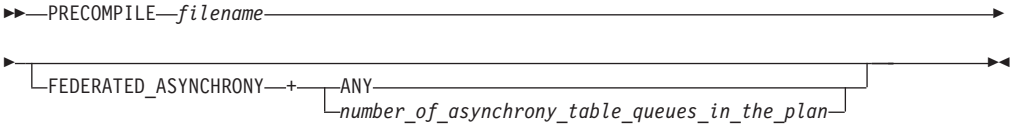
The value of the FEDERATED_ASYNCHRONY bind option is recorded in the FEDERATED_ASYNCHRONY column in the SYSCAT.PACKAGES catalog table.

When the bind option is not explicitly specified, the value of FEDERATED_ASYNC configuration parameter is used and the catalog shows a value of '-2' for the FEDERATED_ASYNCHRONY column.

If the FEDERATED_ASYNCHRONY bind option is not explicitly specified when a package is bound, and if this package is implicitly or explicitly rebound, the package is rebound using the current value of the FEDERATED_ASYNC configuration parameter.

Precompile option: FEDERATED_ASYNCHRONY

The FEDERATED_ASYNCHRONY precompile option specifies the maximum number of ATQs that the federated server supports in the access plan for programs that use embedded SQL. You use the FEDERATED_ASYNCHRONY option in the PRECOMPILE command.



Type: Numeric

You can assign a number to this parameter or use the keyword 'ANY'.

Range: 0 to (MAXAGENTS / 4) inclusive, or ANY

MAXAGENTS is the total number of agents (coordinator agents and subagents) in the system. You specify the total number of agents in the MAXAGENTS database manager configuration parameter.

When ANY is specified, the optimizer determines the number of ATQs for the access plan. The optimizer assigns an ATQ to all eligible SHIP or remote pushdown operators in the plan. The value that is specified for DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option limits the number of asynchronous requests.

Usage notes

For an embedded SQL program, if the FEDERATED_ASYNCHRONY precompile option is not explicitly specified the static statements in the package are bound using the FEDERATED_ASYNC configuration parameter. If the FEDERATED_ASYNCHRONY option is specified explicitly, that value is used for binding the packages and is also the initial value of the special register. Otherwise, the value of the database manager configuration parameter is used as the initial value of the special register. The FEDERATED_ASYNCHRONY precompile option influences dynamic SQL only when it is explicitly set.

Server option: DB2_MAX_ASYNC_REQUESTS_PER_QUERY

This server option applies to all relational and nonrelational data sources. You can specify this option in the CREATE SERVER or ALTER SERVER statements.

Type: Numeric

You can assign a number to this parameter or use the keyword 'ANY'.

Default: 1, except for the ODBC wrapper

All of the wrappers, except for ODBC, support multiple cursors per connection. Because you can use the ODBC wrapper to access a data source that might not support multiple cursors, the default value for the DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option for the ODBC wrapper is 0.

Range: -1 to (maxagents / 4)

- -1 = ANY. The optimizer determines the number of ATQs to use for the SHIPs or RPDs for a given data source.
- 0 = Asynchrony is disabled for the SHIPs or RPDs for this server
- 1 is the default.

Setting the server option to a numeric value other than -1 indicates the maximum ATQs to use for the SHIPs or RPDs for the server provided that other criteria are satisfied.

Usage notes

The DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option specifies the maximum number of asynchronous requests that the data source server wants added by the asynchrony optimization (per connection, per query) in addition to other optimizations. The value of this option translates into how many SHIP or remote pushdown operators that access the data source server should receive an ATQ in a query.

This server option specifies an upper limit for the optimizer to use for the number of ATQs for SHIP or remote pushdown operators accessing the data source server. The DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option does not account for requests originating for reasons other than asynchrony optimization that might result in concurrent requests.

If you change the value for the DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option, packages can be rebound to the database. If this occurs, performance degradation might occur during the rebinding.

The DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option applies only when the FEDERATED_ASYNC database manager configuration parameter is set to a non-zero value to allow asynchrony.

Tuning considerations for asynchrony optimization

When asynchrony is enabled, you need to consider several factors that affect performance.

If your system has available process, memory, and CPU resources, enabling asynchrony can improve the performance of your federated queries. Enabling asynchrony can also increase the use of remote-source systems by the federated server, because a remote source can potentially process more than one request at a time on behalf of a federated query. If your system has resource constraints, asynchrony might degrade performance.

You can tune your system for asynchrony by changing configuration parameters to achieve the degree of asynchrony that is best for your system.

Each asynchronous TQ that asynchrony optimization introduces into a plan requires an additional subagent. If enough subagents are available in the system, consider tuning the MAXAGENTS database manager configuration parameter.

Restrictions on asynchrony optimization

When the optimizer applies asynchrony optimization to a given query, some restrictions apply to the number of ATQs that can be used in the query execution plan.

The number of SHIPs or RPDs that are eligible to be coupled with an ATQ in a plan and benefit from asynchrony might be greater than either the maximum that is set by the FEDERATED_ASYNC parameter or the per server limit of the DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option. In this case, the optimizer chooses SHIPs or RPDs to couple with an ATQ in such a way that:

- The total number of ATQs in the plan is less than or equal to the value set for the following parameters, in the order listed:
  1. FEDERATED ASYNCHRONY special register, if specified
  2. FEDERATED_ASYNCHRONY bind or precompile option, if specified
  3. FEDERATED_ASYNC parameter
- The total number of ATQs for a given server is less than or equal to the DB2_MAX_ASYNC_REQUESTS_PER_QUERY server option for that server.
- The eligible SHIPs or RPDs can benefit from being coupled with an ATQ and improve query performance.

Determining if asynchrony optimization is applied to a query

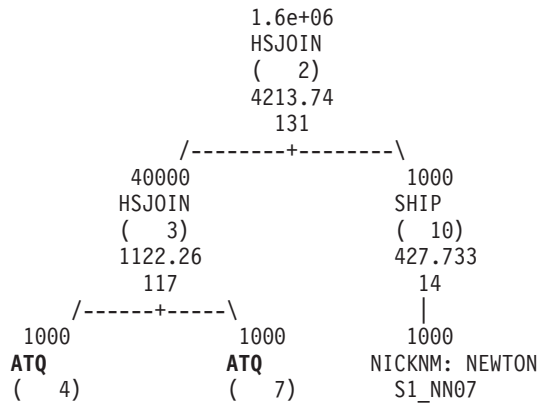
To determine if asynchrony optimization is applied to a given query, you can use one of several methods to check the access plan for the query and check for a specific operator.

You can check any of the following outputs for the query in question:

- db2exfmt output
- Visual explain output
- dynexpln output

Each output shows asynchronous query requests in the access plan as ATQ. The 'origin' property in the description of the ATQ shows 'Asynchrony'.

The following plan fragment shows how the ATQ operator is used and shows its detailed properties.



```
      511.795           532.669
        16                101
        |                 |
       1000              1000
      SHIP              SHIP
      (   5)            (   8)
      478.773           499.887
        16                101
        |                 |
       1000              1000
 NICKNM: NEWTON    NICKNM: NEWTON
      S2_NN02           S3_NN15

- show the origin of the ATQ as ASYNCHRONY:
 4) TQ    : (Table Queue)
  Cumulative Total Cost:   511.795
  Cumulative CPU Cost:   2.79486e+06
  Cumulative I/O Cost:   16
  Cumulative Re-Total Cost:  68.9489
  Cumulative Re-CPU Cost:  1.72372e+06
  Cumulative Re-I/O Cost:  0
  Cumulative First Row Cost:  30.8308
  Cumulative Comm Cost:  18.2176
  Cumulative First Comm Cost: 0
  Estimated Bufferpool Buffers:  16
  Remote communication cost: 538.297

  Arguments:
  ---------
  JN INPUT: (Join input leg)
   OUTER
  LISTENER: (Listener Table Queue type)
   FALSE
  TQMERGE : (Merging Table Queue flag)
   FALSE
  TQORIGIN: (Table Queue Origin type)
   ASYNCHRONY
  TQREAD  : (Table Queue Read type)
   READ AHEAD
  TQSEND  : (Table Queue Write type)
   DIRECTED
  UNIQUE  : (Uniqueness required flag)
   FALSE
```

If you do not find any ATQ operators in the access plan for a query, verify that the following conditions are met:

- The system is a federated system that is enabled for DPF.
- The query accesses nicknames on a data source that is accessed through a fenced wrapper.
- Asynchrony is enabled using the database manager configuration parameter FEDERATED_ASYNC, the special register CURRENT FEDERATED ASYNCHRONY, or the bind option FEDERATED_ASYNCHRONY.
- The server option DB2_MAX_ASYNC_REQUESTS_PER_CONNECTION is set to a non-zero value for the nicknames at each server.

Chapter 19. Materialized query tables and federated systems

Materialized query tables and federated systems – overview

A materialized query table is a table that caches the results of a query. When you submit the query again, the database engine can return the data from the materialized query table instead of repeating the query computation.

You can use materialized query tables with nicknames to improve the performance of a query and to encapsulate a part of logic. Materialized query tables are used when you create cache tables.

The SQL optimizer determines if a query will run more efficiently with a materialized query table than the base tables or nicknames. The optimizer uses the following factors to select a materialized query table:

- The materialized query table must match part or all of the query.
- The refresh age criterion must be met.
- The access plan that uses a materialized query table must be cheaper than the access plan that uses the base tables or nicknames.

Materialized query tables that involve nicknames for objects from the following data sources are supported:

- Relational data sources
  - DRDA
  - Informix
  - ODBC
  - Oracle
  - Sybase
  - MS SQL Server
  - Teradata
- Nonrelational data sources
  - BioRS
  - BLAST
  - Entrez
  - Excel
  - HMMER
  - Table-structured files
  - Web Services
  - WebSphere® Business Integration
  - XML

Creating a federated materialized query table

You use materialized query tables to cache data locally and to improve the performance of your queries. You can use nicknames from relational and nonrelational data sources to create materialized query tables.

Restrictions

- "Data source specific restrictions for materialized query tables" on page 234
- If a query has a function template in a predicate or a select list, the function template must be part of the materialized query table.

- "Restrictions on using materialized query tables with nicknames" on page 236

Procedure

To create a materialized query table, issue a CREATE TABLE statement that references the nicknames that represent the remote data source objects that you want to include.

You can populate a user-maintained materialized query table by using an INSERT statement in a subselect statement. For example:

```
insert into my_mqt (select ..from n1, n2 where ..)
```

where the select portion of the query matches the materialized query table definition. The optimizer might use my_mqt to replace the select portion of the query. In that case, the statement becomes:

```
insert into my_mqt (select .. from my_mqt);
```

In this case, the materialized query table becomes the source of the insert operation. To prevent this from happening, you can by issue one of the following commands to temporarily disable the materialized query table:

```
SET CURRENT REFRESH AGE 0
SET CURRENT MAINTAINED TABLE TYPE FOR OPTIMIZATION SYSTEM
```

Data source specific restrictions for materialized query tables

When you create materialized query tables, you need to be aware of restrictions for specific data sources.

This topic describes the restrictions on creating materialized query tables for the following data sources:
- Bio-RS
- BLAST
- Entrez
- HMMER
- Table-structured files
- Web services
- XML

Bio-RS and Entrez Search restrictions

These wrappers require at least one predicate in the WHERE clause. You must create a materialized query table that satisfies the predicate requirements of the wrappers. If you do not specify a predicate, a refresh of the materialized query table fails.

BLAST and HMMER restrictions

If you use the BLAST and HMMER wrappers, the data source requires predicates on some columns in a query. The wrapper provides a default value for some columns, and predicates on these columns can be omitted from the query. For other columns, you must specify a predicate. If you do not specify a predicate, a refresh of the materialized query table fails.

The wrapper does not access the data source when a materialized query table is created. In the following example, the materialized query table is successfully created.

```
CREATE TABLE MY_MQT AS (SELECT * FROM BLAST_NICK)
   DATA INITIALLY DEFERRED REFRESH DEFERRED ENABLE QUERY OPTIMIZATION;
```

However, when you refresh the table, the BLAST data source is contacted to retrieve data. The BLAST data source issues an error because it needs a predicate on the column blast_seq, and such a predicate is not available in the query.

For optional predicates, if you issue a query with non-default values, you must provide those values when you create the materialized query table. If you do not specify a predicate, you might get incorrect output.

In the following example, when you create the materialized query table, assume that the wrapper provides the default values for the optional parameters and that these default values are provided as IN_ARG1 = 2 and IN_ARG2 < 30.

```
CREATE TABLE MY_MQT AS (SELECT * FROM BLAST_NICK WHERE BLAST_SEQ = '12345')
   DATA INITIALLY DEFERRED REFRESH DEFERRED ENABLE QUERY OPTIMIZATION;
```

When you issue the following query, the optimizer matches the query to the materialized query table. If the optimizer chooses the materialized query table plan, this query yields incorrect results. The query returns incorrect results because the materialized query table contains data with IN_ARG1 = 2, but the query requested data with IN_ARG1 = 3.

```
SELECT * FROM BLAST_NICK WHERE BLAST_SEQ = '12345' AND IN_ARG1 = 3;
```

When you create a materialized query table, you must explicitly specify the values for some of the predicates even if the wrapper supplies default values. Otherwise, the optimizer might not route the query to a matching materialized query table. This can happen because you did not explicitly specify these predicate values when you created the materialized query table.

The following example shows how the optimizer would fail to select the materialized query table when the query matches the materialized query table. If the materialized query table explicitly contains a fixed input column with a default value and the query does not have this predicate specified, the query will not be routed to the materialized query table.

```
CREATE TABLE K55ADMIN.BLAST_NICK1_M1 AS (
   SELECT SCORE, E_VALUE, QUERYSTRANDS
   FROM K55ADMIN.BLAST_NICK1
   WHERE BLASTSEQ='ATGATCGGATCGAATTCGAT'
   AND E_VALUE < 10) DATA INITIALLY DEFERRED REFRESH DEFERRED;
```

If you issue the following query, the query is not routed to the materialized query table. The query is not routed because the optimizer determines that the materialized query table did not specify E_VALUE < 10.

```
SELECT SCORE, E_VALUE, QUERYSTRANDS
FROM K55ADMIN.BLAST_NICK1 WHERE BLASTSEQ='ATGATCGGATCGAATTCGAT';
```

Table-structured file restrictions

If you define a nickname for a table-structured file with the DOCUMENT option, the materialized query table must have a predicate that specifies the file path. If you do not specify a predicate, a refresh of the materialized query table fails.

Web services restrictions

You can only create a materialized query table over a flattened view of a hierarchy of nicknames. You cannot create a materialized query table for each nickname of a hierarchy.

XML restrictions

You cannot create a materialized table on a child table.

If you define a nickname for an XML table with the DOCUMENT option, the materialized query table requires a predicate that specifies the file path. If you do not specify a predicate, a refresh of the materialized query table fails.

Restrictions on using materialized query tables with nicknames

Restrictions on materialized query tables that reference nicknames should be considered when you are tuning your federated system.

Oracle Label Security for data source objects

Nicknames for data source objects with Oracle Label Security cannot be cached, and materialized query tables cannot be created on them.

System-maintained materialized query tables

WebSphere Federation Server does not support system-maintained materialized query tables that reference nicknames in a partitioned database environment.

To work around this restriction, you can use user-maintained materialized query tables.

For example, for a nonrelational nickname named DEPART, you can issue the following commands to simulate a system-maintained materialized query table.

```
SET CURRENT MAINTAINED TABLE TYPES FOR OPTIMIZATION ALL;

CREATE TABLE AST1(C1, C2)
 AS (SELECT EMPNO, FIRSTNME FROM DEPART WHERE EMPNO>'000000')
 DATA INITIALLY DEFERRED REFRESH DEFERRED
 ENABLE QUERY OPTIMIZATION MAINTAINED BY USER;

SET INTEGRITY FOR AST1 ALL IMMEDIATE UNCHECKED;

INSERT INTO AST1 (SELECT EMPNO, FIRSTNME FROM DEPART WHERE EMPNO>'000000');

SET CURRENT REFRESH AGE ANY;
```

The following SELECT statement can be answered by the materialized query table defined above:

```
SELECT EMPNO, FIRSTNME FROM DEPART
 WHERE EMPNO > '000000' AND FIRSTNME LIKE 'AN%';
```

Chapter 20. Cache tables

You use cache tables to store data that you access frequently but that does not change often.

A cache table can improve query performance by storing the data locally instead of accessing the data directly from the data source.

You can cache data from these data sources:

- DB2 family
- Informix
- Microsoft SQL Server
- Oracle
- Sybase

A cache table consists of these components:

- A nickname on your federated database system. The nickname has the same column definitions and same data access as the data source table.
- One or more materialized query tables that you define on the nickname. The type of materialized query tables is FEDERATED_TOOL maintained materialized query table. The materialized query table usually contains a subset of high-use data from the data source table.
- A replication schedule for each materialized query table. The replication schedule keeps the local materialized query tables current with your data source tables. You define the replication schedule.

The following figure illustrates a cache table.

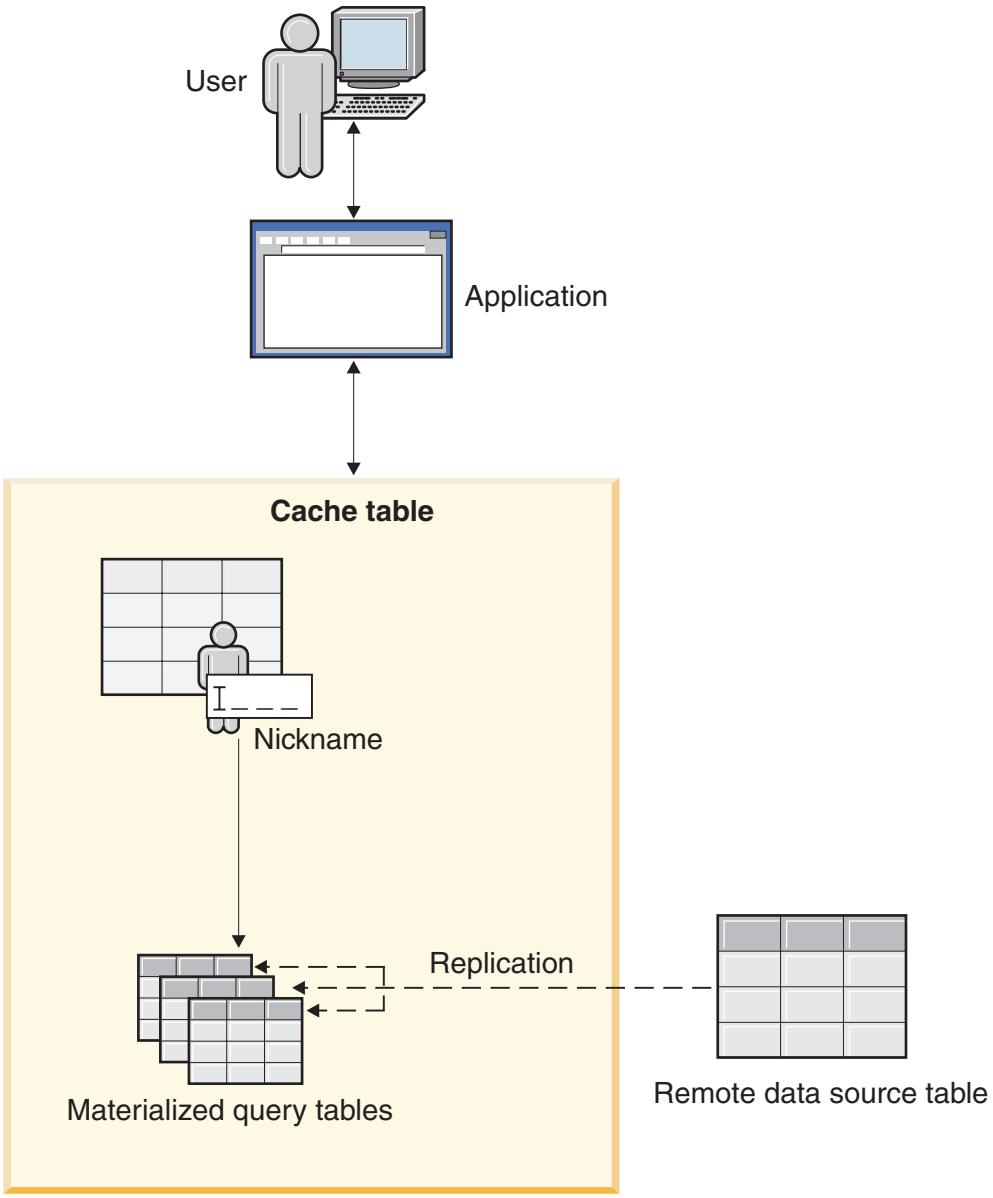


Figure 16. Cache table.

The cache table has the same name as the nickname. You can associate a cache table with only one data source table.

When a cache table is enabled, the query optimizer directs queries to the cache table if the data that the query requests can be found in the materialized query table.

Creating cache tables

You use a wizard in the Control Center to create a cache table. The wizard creates the nickname, the materialized query table, and the replication schedule that are required for the cache table.

Before you begin

- Set the FEDERATED parameter to YES on the federated server. The FEDERATED parameter is a database manager configuration parameter.
- To access Informix data sources, install and configure the Informix Client software development kit (SDK) in the federated server.
- To cache data from DB2 Database for Linux, UNIX, and Windows tables, configure the DB2 database for archive logging.
- The federated database or the source database must be on the computer from which you are creating the cache tables. If the federated database or the source database are not local, you must catalog the databases on the local computer. The alias name that you use when you catalog the database must be the same name as the database name.
- The user ID in the user mapping between the databases must have the authority to create tables in the source database.

Procedure

To create a cache table:

1. In the Control Center, expand the **Cache Objects** folder.
2. Right-click the **Cache Tables** folder and click **Create**.
3. Complete the steps in the Cache Table wizard to create the cache table. You can create the cache table quickly by only specifying values for the required fields and using the default settings for the remaining fields. To change the default settings:
   a. On the Materialized Query Table page, click **Advanced MQT Settings** to change the default settings or to select a subset of columns for the materialized query table.
   b. On the Replication page, click **Advanced Settings** to change the default settings for replicating data from the data source table to the materialized query table.

   In some circumstances, caching is not be enabled when you complete the wizard. You must enable caching to start the replication Capture and Apply programs.

The Cache Table wizard creates one materialized query table when you create the cache table. You can create additional materialized query tables to store other data from the same data source.

Modifying the settings for materialized query tables

You cannot modify the settings for materialized query tables directly. You must use alternative methods to change the replication and materialized query table settings.

Procedure

To modify the settings for a materialized query table:

1. Use the Materialized Query Table Details window to view the settings for the materialized query table and the replication schedule.
   a. In the Control Center, expand the **Cache Objects** folder.
   b. Right-click the cache table and click **Properties**.
   c. Select the materialized query table and click **Details** to view the current settings.

2. If you need to make changes to the replication settings, use the Replication Center. You cannot change the replication settings for the materialized query table from the Materialized Query Table Details window.
3. If you need to change the settings for a materialized query table, drop the materialized query table and create another materialized query table. For example, if you need to add another column to the materialized query table, drop the materialized query table and create a materialized query table with the new settings.

Adding materialized query tables to a cache table

You can create additional materialized query tables for the same cache table. Use the additional materialized query tables to store other data from the same data source table.

About this task

When you create a cache table, the federated server stores data locally from the data source in a materialized query table. The criteria that you specify in the Cache Table wizard determines which data to store in the materialized query table.

For example, you have an initial materialized query table that contains information about customers in Asia. You can create another materialized query table that contains information about customers in South America.

Procedure

To add a materialized query table to an existing cache table:

1. In the Control Center, expand the **Cache Objects** folder and the **Cache Tables** folder.
2. Right-click the appropriate cache table and click **Properties**.
3. Click **Add**.
4. Complete the steps in the Cache Table wizard to create the additional materialized query table.

You can also access the Cache Table Properties window from a nickname object. Right-click the nickname and click **Caching**.

Routing queries to cache tables

You can route queries to the materialized query table for the cache table or to data source by routing the queries to the nickname.

Before you begin

- A cache table must exist for the nickname that you query.
- Enable caching to the materialized query table.
- Set the following database configuration parameters:
  - Materialized table type for optimization (DFT_MTTB_TYPES)
  - Query optimization class (DFT_QUERYOPT)

  Use the **Configure Database** notebook in the Control Center or the command line to set these parameters.

Procedure

To change the routing for queries:

1. In the Control Center, expand the **Cache Objects** folder and the **Cache Tables** folder.
2. Right-click the appropriate cache table and click **Properties**.
3. Select the materialized query table and click **Check Status**.
4. Select where you want the queries to route the queries:
   - Select **Route to the materialized query table**. All queries for the data source are sent to the materialized query table. If the data in the materialized query table cannot satisfy the query, the query is then routed to the data source using the nickname.
   - Select **Route to the nickname**. All queries are sent to the data source by using the nickname. The materialized query table is changed to a regular database table. Replication to the materialized query table continues unless you disable caching.
5. To change the routing immediately, select the **Remove all cached dynamic SQL statements** check box. The dynamic SQL statements that are currently in the package cache are deleted.
6. Click **OK**.

You can also access the Cache Table Properties window from a nickname object. Right-click the nickname and click **Caching**.

Enabling and disabling the replication cache settings

You start and stop data replication for the materialized query table by changing the cache settings.

Before you begin

For DB2 Database for Linux, UNIX, and Windows data sources, set the type of database logging to archive logging.

About this task

When you enable caching, the Capture and Apply programs are started if the programs are not already running. At the same time, the subscription-set member is enabled. Enabling the subscription-set member tells the Apply program to keep the data in the materialized query table synchronized with the data in the data source table.

When you disable caching, the subscription-set member is disabled. The data from the data source is not replicated to the materialized query table.

Important: If you do not change the routing setting, your queries are routed to the materialized query table even though the data is not replicated to the materialized query table.

Procedure

To enable or disable the replication cache settings for a materialized query table:

1. In the Control Center, expand the **Cache Objects** folder and the **Cache Tables** folder.
2. Right-click on the appropriate cache table and click **Properties**.

3. Select the materialized query table and click **Check Status**.
4. Select either **Enable caching** or **Disable caching**. To view the enable or disable commands, click **Show Statements**.
5. Click **OK** .

You can also access the **Cache Table Properties** window from a nickname object. Right-click the nickname and click **Caching**.

Dropping materialized query tables from a cache table

When you no longer want to store data locally in a materialized query table, you can drop the materialized query table from the cache table.

About this task

If a cache table has only one materialized query table, dropping the materialized query table also drops the cache table.

To ensure that the materialized query table is completely removed from your system, use the Control Center to drop the materialized query table from a cache table.

Procedure

To drop a materialized query table from a cache table:

1. In the Control Center, expand the **Cache Objects** folder and the **Cache Tables** folder in the object tree.
2. Right-click the appropriate cache table and click **Properties**.
3. Select the appropriate materialized query table and click **Remove**.

Dropping cache tables

When you no longer want to store data locally in a cache table, you can drop the cache table.

About this task

When you drop a cache table, the federated server performs the following actions:

- Drops the materialized query tables that are for the cache table.
- Removes the replication schedule for the data sources and the materialized query tables.
- Drops the nickname for the data source, if the nickname was created when you created the cache table. If you used an existing nickname when you created the cache table, the federated server does not drop the nickname.

Procedure

To drop a cache table:

1. In the Control Center, expand the **Cache Objects** folder and the **Cache Tables** folder in the object tree.
2. Right-click the appropriate cache table and click **Drop**.

Chapter 21. How client applications interact with data sources

To client applications, the data sources in a federated system appear as a single collective database. To obtain data from data sources, applications submit queries in DB2 SQL to the federated database. The federated database then distributes the queries to the appropriate data sources, and either returns this data to the applications or performs the requested action.

The federated database can join data from local tables and remote data sources in the same SQL statement. For example, you can join data that is located in a local DB2 table, an Informix table, and a Sybase view in a single SQL statement. By processing SQL statements as if the data sources were ordinary relational tables or views within the federated database, the federated system can join relational data and nonrelational data.

In a federated system, you can access data sources through nicknames. A *nickname* is a federated database object that an application uses to reference a data source object, such as a table or view. To write to a data source—for example, to update a data source table—an application can use DB2 SQL (with nicknames). Alternatively, applications can use the SQL dialect of the data source (without nicknames) in a special session called *pass-through* to access the data sources directly.

Applications that use DB2 SQL and nicknames can access any data types that the federated database recognizes.

The federated database catalog contains information about the objects in the federated database and information about objects at the data sources. Because the catalog contains information about the entire federated database, it is called a *global catalog*.

Chapter 22. Nicknames in your applications

Reference data source objects by nicknames in SQL statements

With a federated system, you use the nicknames defined for data source objects to represent the objects in your SQL statements. The federated system does not recognize fully-qualified data source, schema, and object names in SQL statements.

Data source objects must have nicknames registered in the federated database before you can include them in your queries. In general, you can specify nicknames in an SQL statement where you can specify local tables in a SQL statement.

Example: Using nicknames in SELECT, INSERT, UPDATE, and DELETE statements

You define the nickname NFXDEPT to represent a table in an Informix table called PERSON.DEPT, where:

- PERSON is the data source schema
- DEPT is the data source table name

The statement SELECT * FROM NFXDEPT is allowed from the federated server. However, the statement SELECT * FROM PERSON.DEPT is not allowed (except in a pass-through session). The federated server does not have PERSON.DEPT registered as a nickname.

Example: Using nicknames in the CREATE TABLE statement

You want to create a local table based on a remote table for which you have defined a nickname. An example of the CREATE TABLE statement is:

```
CREATE TABLE table_name LIKE nickname
```

Nicknames in DDL statements

Data source objects must have nicknames registered in the federated database before you can include them in your DDL statements. This topic provides some examples of DDL statements that you use with federated systems.

Using nicknames in the COMMENT ON statement

The COMMENT ON statement adds or replaces comments in the federated database global catalog. The COMMENT ON statement is valid with a nickname and columns that are defined on a nickname. This statement does not update data source catalogs.

Using nicknames in the GRANT and REVOKE statements

The GRANT and REVOKE statements are valid with a nickname for certain privileges and for all users and groups. However, DB2 UDB does not issue a corresponding GRANT or REVOKE statement on the object on the data source that the nickname references.

For example, suppose that user JON creates a nickname for an Oracle table that had no index. The nickname is ORAREM1. Later, the Oracle DBA defines an index for this table. User EILEEN now wants the DB2 federated database to know that this index exists, so that the query optimizer can devise strategies to access the table more efficiently. EILEEN can inform the federated database that a new index exists, by creating an index specification for ORAREM1.

The information about the index is stored in the SYSSTAT.INDEXES catalog view. Use the GRANT statement to give EILEEN the index privilege on this nickname, so that she can create the index specification.

```
GRANT INDEX ON NICKNAME ORAREM1 TO USER EILEEN
```

To revoke user EILEEN's privileges to create an index specification on nickname ORAREM1, use the REVOKE statement:

```
REVOKE INDEX ON ORAREM1 FROM USER EILEEN
```

Data source statistics impact applications

When a nickname is created for a data source object, the federated database global catalog is updated with information about that object. The query optimizer uses this information to plan how to retrieve data from the object.

It is important to make sure that the data source information is current. The federated database does not automatically detect changes to data source objects.

Database object statistics stored in the global catalog

The information stored in the global catalog about a data source object, depends on the type of object. For database tables and views, the name of the object, the column names and attributes, are stored in the global catalog.

In the case of a table or nickname, the information also includes:

- Statistics. For example, the number of rows and the number of pages on which the rows exist. To ensure that the federated database obtains the latest statistics, run the data source equivalent of the RUNSTATS command on the table before you create the nickname.
- Index descriptions. If the table has no indexes, you can supply the catalog with metadata that an index definition typically contains. For example, assume that a nickname is created for a remote table, and that an index is subsequently created on the table at the data source. You can create an index specification at the federated server that represents this remote index. You create an index specification by issuing the CREATE INDEX statement and referencing the nickname for the table. You use the SPECIFICATION ONLY clause with the CREATE INDEX statement to produce only an index specification. The index specification informs the federated optimizer that a remote index exists. However, only metadata is generated. No index is actually created on the federated server. In addition, no statistical information is supplied to the global catalog. If you give the index specification exactly the same signature as the remote index (that is, the same name, and the same columns in the same order), you can use SYSPROC.NNSTAT to update statistics on the nickname and index specification.

To determine what data source information is stored in the global catalog, query the SYSCAT.TABLES and SYSCAT.COLUMNS catalog views. To determine what

data source index information is stored in the catalog, or what a particular index specification contains, query the SYSCAT.INDEXES catalog view.

Updating statistics using the SYSSTAT view instead of the SYSCAT view

SYSCAT views are read-only catalog views in the SYSCAT schema. SYSSTAT views are updatable catalog views that contain statistical information that the optimizer uses. SYSSTAT views are in the SYSSTAT schema.

If you issue an UPDATE or INSERT operation on a view in the SYSCAT schema, it will fail. Use the updatable catalog views in the SYSSTAT schema to manually modify statistics on nicknames.

Defining column options on nicknames

Column options are parameters in the CREATE NICKNAME and ALTER NICKNAME statements. You can specify column options when you initially create a nickname or by modifying an existing nickname.

The information that you provide through the column options is stored in the global catalog.

Nonrelational data sources

Column options are unique for each nonrelational wrapper. These options are typically set when you issue the CREATE NICKNAME statement.

Relational data sources

There are two column options you can use for relational data sources: NUMERIC_STRING and VARCHAR_NO_TRAILING_BLANKS.

Setting the NUMERIC_STRING column option

If a data source string column contains only numeric digits, and no other characters including blanks, set the NUMERIC_STRING column option to 'Y'.

Setting the NUMERIC_STRING column option to 'Y' allows queries that use this column to be optimized for sorting operations and comparison operations. For example:

```
ALTER NICKNAME nickname
   ALTER COLUMN local_column_name
   OPTIONS (SET NUMERIC_STRING 'Y')
```

Setting the VARCHAR_NO_TRAILING_BLANKS column option

If the data source string column does not contain trailing blanks, set the VARCHAR_NO_TRAILING_BLANKS column option to 'Y'.

Some data sources, such as Oracle, do not use the same blank-padded string comparison logic that the federated database uses. This applies to data types such as VARCHAR and VARCHAR2. As a result, predicates that involve these data types must be rewritten by the query optimizer to ensure consistent query results.

Rewriting query statements can impact performance. Setting this option for a specific column provides the query optimizer with information about these columns so that it can generate more efficient SQL statements.

For example:

```
ALTER NICKNAME nickname
   ALTER COLUMN local_column_name
   OPTIONS (SET VARCHAR_NO_TRAILING_BLANKS 'Y')
```

Chapter 23. Creating and using federated views

A view in that includes a reference to a nickname in the fullselect is a *federated view*. The base tables are referenced in the federated view using nicknames, instead of using the data source table names.

Restrictions

Federated views that are created from multiple data source objects are read-only views and cannot be updated.

Federated views that are created from only one data source object might or might not be read-only views.

- A federated view created from a single nonrelational data source is read-only.
- A federated view created from a single relational data source might allow updates, depending on what is included in the CREATE VIEW statement.

About this task

The advantages of using federated views are similar to the advantages of using views defined on local tables in a centralized relational database manager:

- Views provide an integrated representation of the data
- You can exclude table columns that contain confidential or sensitive data from a view

Procedure

You create a federated view from data source objects that have nicknames. The action of creating a federated database view of data source data is sometimes called "creating a view on a nickname". This phrase reflects the fact that for the federated view to be created, the CREATE VIEW statement fullselect must reference the nickname of each data source table and view that the federated view is to contain.

Creating federated views - examples

This topic provides examples of creating federated views.

Example: Creating a federated view that merges similar data from several data source objects

You are working with customer data on three separate servers, one in Europe, one in Asia, and one in South America. The Europe customer data is in an Oracle table. The nickname for that table is ORA_EU_CUST. The Asia customer data is in a Sybase table. The nickname for that table is SYB_AS_CUST. The South America customer data resides in an Informix table. The nickname for that table is INFMX_SA_CUST. Each table has columns containing the customer number (CUST_NO), the customer name (CUST_NAME), the product number (PROD_NO), and the quantity ordered (QUANTITY). The syntax to create a view from these three nicknames that merges this customer data is:

```
CREATE VIEW FV1
   AS SELECT * FROM ORA_EU_CUST
   UNION
   SELECT * FROM SYB_AS_CUST
   UNION
   SELECT * FROM INFMX_SA_CUST
```

Example: Joining data to create a federated view

You are working with customer data on one server and sales data on another server. The customer data is in an Oracle table. The nickname for that table is ORA_EU_CUST. The sales data is in a Sybase table. The nickname for that table is SYB_SALES. You want to match up the customer information with the purchases made by those customers. Each table has a column containing the customer number (CUST_NO). The syntax to create a federated view from these two nicknames that joins this data is:

```
CREATE VIEW FV4
   AS SELECT A.CUST_NO, A.CUST_NAME, B.PROD_NO, B.QUANTITY
   FROM ORA_EU_CUST A, SYB_SALES B
   WHERE A.CUST_NO=B.CUST_NO
```

Chapter 24. Maintain data integrity with isolation levels

The isolation level defines the degree of isolation for an application process from other application processes that are running concurrently.

You can maintain data integrity for a data source table by requesting that the table rows be locked at a specific isolation level.

Locking occurs at the base table row at the data source. The database manager, however, can replace multiple row locks with a single table lock. This action is called *lock escalation*. An application process is guaranteed at least the minimum requested lock level.

The isolation levels for the federated database are as follows:

RR Repeatable read

RS Read stability

CS Cursor stability (default)

UR Uncommitted read

The types of isolation are the statement level isolation and connection level isolation.

You can set the isolation when you perform the following actions:

- Precompile or bind an application. You can specify isolation levels when you prepare or bind an application. The isolation level specified in the BIND and PREP command is the default isolation level when the federated server connects to the remote data source.
- Use the WITH clause in an SQL statement. This action is called statement level isolation. You can use the WITH clause in the SELECT, UPDATE, INSERT, and DELETE statements.

If the federated server does not find an isolation level for a statement, the federated server uses the isolation level that was established when the federated server connected to the data source.

The following table lists the data sources that use connection level isolation, the isolation levels that they use, and the equivalent isolation levels on the federated server.

Table 20. Data sources and isolation levels

Data sources	Most restrictive isolation level	More restrictive isolation level	Less restrictive isolation level	Least restrictive isolation level
Federated database	Repeatable read	Read stability	Cursor stability	Uncommitted read
DB2 family of products	Repeatable read	Read stability*	Cursor stability	Uncommitted read
Microsoft SQL Server	Serializable	Repeatable read	Read committed	Read Uncommitted
Informix	Repeatable read	Repeatable read	Cursor stability	Dirty read

Table 20. Data sources and isolation levels (continued)

Data sources	Most restrictive isolation level	More restrictive isolation level	Less restrictive isolation level	Least restrictive isolation level
ODBC	Serializable	Repeatable read	Read committed	Read Uncommitted
Oracle	Serializable	Serializable	Read committed	Read committed
Sybase	Level 3	Level 3	Level 1	Level 0
*For DB2 UDB for VM and VSE Server data sources, the isolation level is repeatable read.				

The CURRENT ISOLATION special register is not used by the federated server when it connects to a data source.

The nonrelational data sources do not have a concept like isolation levels. The OLE DB and Teradata do have the concept of isolation levels but are not supported by the federated server. There is no isolation-level mapping between the federated database isolation levels and the OLE DB, Teradata, and nonrelational data sources.

Statement level isolation in a federated system

For federated data sources, you must use the WITH isolation clause to specify the isolation of a statement.

You must use the WITH isolation clause in your statement if you want to use statement level isolation. If you use attributes with the Call Level Interface (CLI) or other application API for statement level isolation, it does not affect statement isolation.

The data sources that support statement level isolation in a federated system are the DB2 family of products and Microsoft SQL server. The statement isolation is sent to remote data sources for the DB2 family of products and SQL server.

Use the DB2_STATEMENT_ISOLATION server option to turn on or off statement level isolation. You can specify this option in the CREATE SERVER and ALTER SERVER statements. The server option is automatically set to 'Y'.

You can use the WITH isolation clause in these statements:

SELECT
SELECT INTO
Searched DELETE
INSERT
Searched UPDATE
DECLARE CURSOR

Lock request clause

You can use the lock request clause in a SELECT or SELECT INTO statement. The federated data sources that support the lock request clause are the DB2 UDB for Linux, UNIX, and Windows, and the DB2 UDB for z/OS.

Limitations using the WITH clause to set the isolation level

The following conditions apply to isolation levels that are specified for statements:

- The WITH clause cannot be used in subqueries.
- The UR isolation level applies only if the result table of the fullselect or the SELECT INTO statement is read-only. In other cases, the UR isolation level for the statement is changed from UR to CS for the DB2 family of data sources. For the SQL server data source, the UR isolation level is upgraded to Read committed.
- If you specify the lock request clause for the following data sources, the clause is ignored by the federated server.
  - DB2 for iSeries
  - DB2 for VM
  - Microsoft SQL server

Connection level isolation in a federated system

The federated server maps your isolation level to a corresponding one at the data source.

For each connection to the data source, the wrapper determines the isolation level.

When the federated server connects to the data source, the isolation level at the remote data source is set to a level that is equivalent to the level of the federated server. If there is no exact equivalent, then the federated server sets the isolation level to the next restrictive level. After a connection is made to a data source, the isolation level for the duration of the connection cannot be changed.

All wrappers except Teradata keep track of the connection isolation level. When setting up a connection, the wrappers set the connection isolation level to the equivalent of the current DB2 isolation level. The current DB2 isolation level is the isolation level of the current section (first federated statement to a data source). The Teradata wrapper is always in the READ isolation level, the default, because the Teradata wrapper does not have a way to change the connection isolation level.

Chapter 25. Federated LOB support

With a federated database system, you can access and manipulate large objects (LOBs) at remote data sources.

A federated system supports SELECT operations on LOBs at DRDA, Informix, Microsoft SQL Server, Oracle, and Sybase data sources. For example:

```
SELECT empname, picture FROM infmx_emp_table
   WHERE empno = '01192345'
```

Where *picture* represents a LOB column and *infmx_emp_table* represents a nickname referencing an Informix table containing employee data.

A federated system supports SELECT, INSERT, UPDATE, and DELETE operations on LOBs at the following data sources, using the DRDA wrapper:

- DB2 for z/OS (Version 7 or higher)
- DB2 for iSeries (Version 5)
- DB2 UNIX, and Windows (Version 7 or higher)

The read and write operations supported by DB2 Database for Linux, UNIX, and Windows, Version 9 are listed in the following table:

Table 21. Read and write support for LOBs

Data source	Type of operations
DB2 for z/OS, DB2 for iSeries, DB2 Database for Linux, UNIX, and Windows[1]	read and write
BioRS	read only
BLAST	read and bind-in
Entrez	read only
HMMER	read and bind-in
Informix	read only
Microsoft SQL Server	read only
Oracle (NET8 wrapper) [2]	read only
ODBC	read only
Sybase	read only
Teradata	read only
Web services	read only and bind-out for CLOB only
WebSphere Business Integration	read only and bind-out for CLOB only
XML	read only

Note:

1. DB2 for iSeries Version 5 (or later) is required for LOB support. DB2 Information Integrator Version 8 cannot access DB2 UDB for Linux, UNIX, and Windows Version 7 LOB data.
2. To run insert, update, and delete operations on Oracle LONG columns, you need to migrate the remote columns from LONG to LOBs and recreate the nicknames.

Teradata LOBs

Teradata LOBs are slightly different than DB2 LOBs. Teradata does not have any data types as large as the LOBs supported in DB2 UDB. However, there are some Teradata data types that can be up to 64000 bytes long. These data types are CHAR, VARCHAR, BYTE, VARBYTE, GRAPHIC, and VARGRAPHIC. These Teradata data types are mapped to DB2 LOB data types when the length of the Teradata data type exceeds the limits of the corresponding DB2 data type.

LOB lengths

Some data sources, such as Oracle and Informix, do not store the lengths of LOB columns in their system catalogs. When you create a nickname on a table, information from the data source system catalog is retrieved including column length. Since no length exists for the LOB columns, the federated database assumes that the length is the maximum length of a LOB column in DB2 Database for Linux, UNIX, and Windows. The federated database stores the DB2 Database for Linux, UNIX, and Windows maximum length in the federated database catalog as the length of the nickname column.

LOB locators

Applications can request LOB locators for LOBs that are stored in remote data sources. A LOB locator is a 4-byte value stored in a host variable. An application can use the LOB locator to refer to a LOB value (or LOB expression) held in the database system.

Using a LOB locator, an application can manipulate the LOB value as if the LOB value was stored in a regular host variable. When you use LOB locators, there is no need to transport the LOB value from the data source server to the application (and possibly back again).

The federated database can retrieve LOBs from remote data sources, store them at the federated server, and then issue a LOB locator on the stored LOB. LOB locators are released when:

- Applications issue FREE LOCATOR SQL statements
- Applications issue COMMIT statements
- The DB2 federated instance is restarted

Restrictions on LOBs

Federated systems impose some restrictions on LOBs.

The following restrictions apply to LOBs:

- The federated database is unable to bind remote LOBs to a file reference variable
- LOBs are not supported in pass-through sessions

Performance considerations for LOB processing

When developing federated applications that fetch and process LOB data, application designers and database administrators need to understand how LOB processing affects performance.

When an application fetches data from a federated data source, the federated server must fetch the data into its own application buffers before sending the data

to the application. Because LOBs are not processed in a buffer pool, the LOB data must first pass through a temporary table space defined for the federated server. To help improve performance and reduce resource consumption, application designers should only materialize LOB data when necessary.

Similarly, when the federated server updates remote LOB data, the data must pass through a temporary table space assigned to the federated server before it is passed to the data source.

Transient LOBs use the temporary table space assigned to the federated server. Therefore, database administrators might need to increase the size of this temporary table space to ensure that the working area is sufficient for processing the LOBs.

Recommendation: To maximize performance when working with LOBs, define the temporary table space as System Managed (SMS) and ensure that the temporary table space is located on disks with a high I/O bandwidth.

Using the DB2 Call Level Interface to access federated LOBs

The federated server supports two DB2 CLI APIs for selecting LOB data:

- The SQLFetch API fetches the LOB from the federated server or data source into the application buffers in a single operation.
- The SQLGetData API fetches the LOB a chunk at a time and can require repeated calls to the API to fetch the entire LOB into the application buffers.

Recommendation: For optimal performance, use the SQLGetData API when fetching LOBs through a federated server.

The federated server supports the SQLExecute and SQLPutData APIs for updating LOB data. The SQLExecute API updates the LOB data in a single operation, whereas the SQLPutData API can require repeated calls to send all of the LOB data from the application buffers to the server. Each API performs at the same level in a federated environment.

Trusted and fenced wrappers

Nicknames created for wrappers defined as trusted or fenced perform equally when fetching or updating LOBs.

Chapter 26. Distributed requests

Distributed requests for querying data sources

Queries submitted to the federated database can request results from a single data source, but typically are requests that include multiple data sources. Because a typical query is distributed to multiple data sources, it is called a *distributed request*..

In general, a distributed request uses one or more of three SQL conventions to specify where data is to be retrieved from subqueries, set operators, and join subselects..

Distributed requests for querying data sources - examples

In this example, the federated server is configured to access a DB2 for z/OS data source, a DB2 for iSeries data source, and an Oracle data source. Stored in each data source is a table that contains employee information. The federated server references these tables by nicknames that point to where the tables reside.

zOS_EMPLOYEES
Nickname for a table on a DB2 for z/OS data source that contains employee information.

iSERIES_EMPLOYEES
Nickname for a table on a DB2 for iSeries data source that contains employee information.

ORA_EMPLOYEES
Nickname for a table on an Oracle data source that contains employee information.

ORA_REGIONS
Nickname for a table on an Oracle data source that contains information about the regions that the employees live in.

The following examples illustrate the three SQL conventions used with distributed requests, using the nicknames defined for each of the tables.

Example: A distributed request with a subquery

iSERIES_EMPLOYEES contains the phone numbers of employees who live in Asia. It also contains the region codes associated with these phone numbers, but it does not list the regions that the codes represent. ORA_REGIONS lists both codes and regions. The following query uses a subquery to find the region code for China. Then it uses the region code to return a list of those employees in iSERIES_EMPLOYEES who have a phone number in China.

```
SELECT name, telephone FROM db2admin.iSERIES_employees
   WHERE region_code IN
   (SELECT region_code FROM dbadmin.ora_regions
   WHERE region_name = 'CHINA')
```

Example: A distributed request with set operators

The federated server supports three set operators: UNION, EXCEPT, and INTERSECT.

- Use the UNION set operator to combine the rows that satisfy any of two or more SELECT statements.
- Use the EXCEPT set operator to retrieve those rows that satisfy the first SELECT statement but not the second.
- Use the INTERSECT set operator to retrieve those rows that satisfy both SELECT statements.

All three set operators can use the ALL operand to indicate that duplicate rows are not to be removed from the result. This eliminates the need for an extra sort.

The following query retrieves all employee names and region codes that are present in both iSERIES_EMPLOYEES and zOS_EMPLOYEES, even though each table resides in a different data source.

```
SELECT name, region_code
   FROM as400_employees
INTERSECT
SELECT name, region_code
   FROM zOS_employees
```

Example: A distributed request for a join

A relational join produces a result set that contains a combination of columns retrieved from two or more tables. You should specify conditions to limit the size of the rows in the result set.

The query below combines employee names and their corresponding region names by comparing the region codes listed in two tables. Each table resides in a different data source.

```
SELECT t1.name, t2.region_name
   FROM dbadmin.iSERIES_employees t1, dbadmin.ora_regions t2
   WHERE t1.region_code = t2.region_code
```

Optimizing distributed requests with server options

In a federated system, use parameters called *server options* to supply the global catalog with information that applies to a data source as a whole, or to control how the federated database interacts with a data source.

About this task

Server options describe the capabilities of a particular data source and enhance the knowledge that the federated server has about that data source. For example, you can:

- Use the VARCHAR_NO_TRAILING_BLANKS server option to inform the optimizer that every VARCHAR column residing on the data source server is free of trailing blanks. This server option is for Oracle data source only. Use this option only when you are certain that all VARCHAR2 columns for every object that is referenced by a nickname on the server has no trailing blanks. Otherwise, use a column option to specify the columns for individual objects on the server that have no trailing blanks. The column option is also named VARCHAR_NO_TRAILING_BLANKS.
- Set the PLAN_HINTS server option to a value that enables DB2 to provide Oracle data sources with statement fragments, called plan hints. Plan hints can

help a data source optimizer decide which index to use in accessing a table, and which table join sequence to use in retrieving data for a result set.

Typically, the database administrator sets server options for a federated system. However, a programmer can make good use of the server options that help optimize queries. For example, for data sources ORACLE1 and ORACLE2, the PLAN_HINTS server option is set to the default, 'N' (no, do not furnish this data source with plan hints). You write a distributed request that selects data from ORACLE1 and ORACLE2. You expect that plan hints would help the optimizers at these data sources improve their strategies for accessing this data. You could override the default with a setting of 'Y' (yes, furnish the plan hints) while your application is connected to the federated database. When the connection to the data sources is terminated, the setting will automatically revert back to 'N'.

Procedure

To set server options:

1. Use the SET SERVER OPTION statement to set or change server options. To ensure that the setting takes effect, specify the SET SERVER OPTION statement immediately following the CONNECT statement. The server option is set for the duration of a connection to the federated database.
2. **Recommendation:** Prepare the statement dynamically. The SET SERVER OPTION statement affects only dynamic SQL statements.

Chapter 27. Using pass-through sessions within applications

Querying data sources directly with pass-through

This topic describes when and how to use pass-through sessions.

About this task

Pass-through sessions are useful when:

- Applications must create objects at the data source or perform INSERT, UPDATE, or DELETE operations.
- The federated database does not support a unique data source operation.

Procedure

To query data sources directly with pass-through:

- Use the SET PASSTHRU statement to start a pass-through session and access a server directly. This statement can be issued dynamically. An example of this statement is: SET PASSTHRU ORACLE1 This SET PASSTHRU statement opens a pass-through session to the data source using the server name ORACLE1. ORACLE1 is the name you registered for the data source server when you created the server definition.
- When the pass-through session is opened, ensure that you use the true name of the object and not the nickname when you reference objects in a pass-through session. You must use the SQL dialect of the data source, unless the federated database is the data source that is being referenced.
- If a static statement is submitted in a pass-through session, it is sent to the federated server for processing. If you want to submit an SQL statement to a data source for processing, you must prepare it dynamically in the pass-through session and have it executed while the session is still open. To prepare statements dynamically in a pass-through session:
  - To submit a SELECT statement, use the PREPARE statement with it, and then use the OPEN, FETCH, and CLOSE statements to access the results of your query.
  - For a supported statement other than SELECT, you have two options. You can use the PREPARE statement to prepare the supported statement, and then the EXECUTE statement to execute it. Alternatively, you can use the EXECUTE IMMEDIATE statement to prepare and execute the statement.

If you issue the COMMIT or ROLLBACK command during a pass-through session, this command will complete the current unit of work, but does not end the pass-through session.

Federated pass-through considerations and restrictions

This topic explains the considerations and restrictions you need to be aware of when you use a pass-through session.

The following considerations and restrictions apply to all data sources:

- Statements prepared within a pass-through session must be executed within the same pass-through session. Statements prepared within a pass-through session,

but executed outside of the same pass-through session, will fail and result in a SQLSTATE 56098 error. Statements prepared outside of a pass-through session, but executed within a pass-through session, are handled as a SET PASSTHRU statement.

- An application can issue multiple SET PASSTHRU statements, however only the last session is active. When a new SET PASSTHRU statement is invoked, it terminates the previous SET PASSTHRU statement. You cannot pass through to more than one data source in the same pass-through session.
- If multiple pass-through sessions are used in an application, be sure to issue a COMMIT before you open another pass-through session. This will conclude the unit of work for the current session.
- Parameter markers are not supported in pass-through sessions. Use host variables instead of parameter markers.
- You can use the WITH HOLD semantics on a cursor defined in a pass-through session. However, you will receive an error if you try to use the semantics (with a COMMIT) and the data source does not support the WITH HOLD semantics.
- Host variables defined in SQL statements within a pass-through session must take the form :H*n* where H is uppercase and *n* is a unique whole number. The values of *n* must be numbered consecutively beginning with zero.
- Pass-through does not support LOBs.
- Pass-through does not support stored procedure calls.
- Pass-through does not support the SELECT INTO statement.
- Pass-through does not support SQL or external user-defined functions.
- You cannot execute dynamic SQL COMMIT or ROLLBACK statements during a pass-through session.
- When performing update or delete operations during a pass-through session, you cannot use the WHERE CURRENT OF CURSOR condition.
- In pass-through mode, dynamic SQL statements are executed remotely, whereas static SQL statements are sent to the federated server for processing. If you want to submit an SQL statement to a data source for processing, you must prepare it dynamically in the pass-through session and it must be executed while the session is still open.
- Static SET PASSTHRU statements in SQL-bodied stored procedures are blocked when stored procedures are created, and error SQL0104N is issued. To get into and out of pass-through mode, use the EXECUTE IMMEDIATE statement.

  Example:

```
create procedure stp()
dynamic result sets 1
language sql
modifies sql data
begin
declare stmt varchar(100);

DECLARE cur1 CURSOR WITH RETURN TO CALLER
FOR SELECT * FROM t1 ;

set stmt = 'set passthru mvs7';
execute immediate stmt;

set stmt = 'insert into t1 values (20, ''passthru_insert'')';
execute immediate stmt;
commit;

insert into t1 values (20, 'stp_insert');
commit;
```

```
OPEN cur1;

end
DB20000I The SQL command completed successfully.

call stp()


Result set 1
------------------------------
10 local
20 stp_insert

2 record(s) selected.

Return Status = 0


select * from t1

C1 C2
------------------------------
10 remote
20 passthru_insert

2 record(s) selected.


set passthru reset
DB20000I The SQL command completed successfully.

select * from t1

C1 C2
------------------------------
10 local
20 stp_insert

2 record(s) selected.
```

- The return from a stored procedure or compound SQL statement does not terminate pass-through mode automatically. To exit from pass-through mode, the SET PASSTHRU RESET statement must be called explicitly, as shown in the example above.

Pass-through sessions to Oracle data sources

This topics identifies some SQL considerations to be aware of before you submit SQL statements to Oracle data sources in a pass-through session.

- When a remote client issues a SELECT statement from a command line processor (CLP) in pass-through mode and the client code is an SDK before DB2 Universal Database Version 5, the SELECT will elicit an SQLCODE -30090 with reason code 11. To avoid this error, remote clients must use an SDK that is at Version 5 or higher.
- Any DDL statement issued on an Oracle server is performed at parse time and is not subject to transaction semantics. The operation, when complete, is automatically committed by Oracle. If a rollback occurs, the DDL is not rolled back.
- When you issue a SELECT statement from raw data types, use the RAWTOHEX function to receive the hexadecimal values. When you perform an INSERT into raw data types, provide the hexadecimal representation.

Chapter 28. Federated system security

Overview of the user mapping plugin for external repositories

You can develop a plugin that retrieves user mappings from an external repository instead of storing the user mappings on the federated server.

The plugin that you develop must retrieve the user mappings from the repository and pass the authentication information to the federated server. Federation includes a sample plugin that is designed for a Lightweight Directory Access Protocol (LDAP) server repository. To work with other types of external repositories, you can develop your own plugin or extend the sample LDAP plugin. After you create the plugin, you must configure the federated server to use the plugin by specifying an option on the wrapper or server definition.

Advantages of using an external repository to store user mappings

Storing user mappings in an external repository can provide improved security and reduced maintenance. An external repository can be shared by many federated servers.

By default, user mappings are stored in the catalog table on the federated server. You can query the catalog view SYSCAT.USEROPTIONS for details. The remote passwords that are stored in the database catalog use the encryption algorithm that is provided by the federated server. Because the remote passwords expire on a regular basis, and you store the user mappings on each federated server, the routine maintenance of updating passwords can become time consuming.

With an external repository, you can store user mappings in a central repository that many federated servers can use to retrieve user mappings.

An external repository such as a Lightweight Directory Access Protocol (LDAP) server, offers increased security. On an LDAP server, you can encrypt the remote passwords with an encryption algorithm and a secret key. If you choose to enable secure communication using secure sockets layer (SSL), you can protect the information that is passed between the federated server and the external repository. The external repository must support SSL communications to use this functionality.

You must develop your own plugin that matches the security and encryption settings of your external repository.

Relationship between the federated server and the user mapping plugin

The user mapping plugin allows the federated server to retrieve user mappings from an external repository. When DB2_UM_PLUGIN option is set and the federated server tries to establish a connection to the data source, the server uses the plugin to retrieve the user mapping for that data source.

The plugin acts as a gateway between the federated server and the external repository. When the federated server needs to access the repository for user mappings, the federated server provides the following information to the plugin:

- Local instance name
- Local database name
- Remote server name
- Local user ID

The plugin uses this information to locate the appropriate user mapping in the external repository. For example, the plugin finds the appropriate user entry in a Lightweight Directory Access Protocol (LDAP) server and requests that the LDAP server return the remote user ID and remote password.

Other types of external repositories will store and return the user mappings using other methods. You must develop a plugin that contains the code to allow the federated server to interface with the external repository.

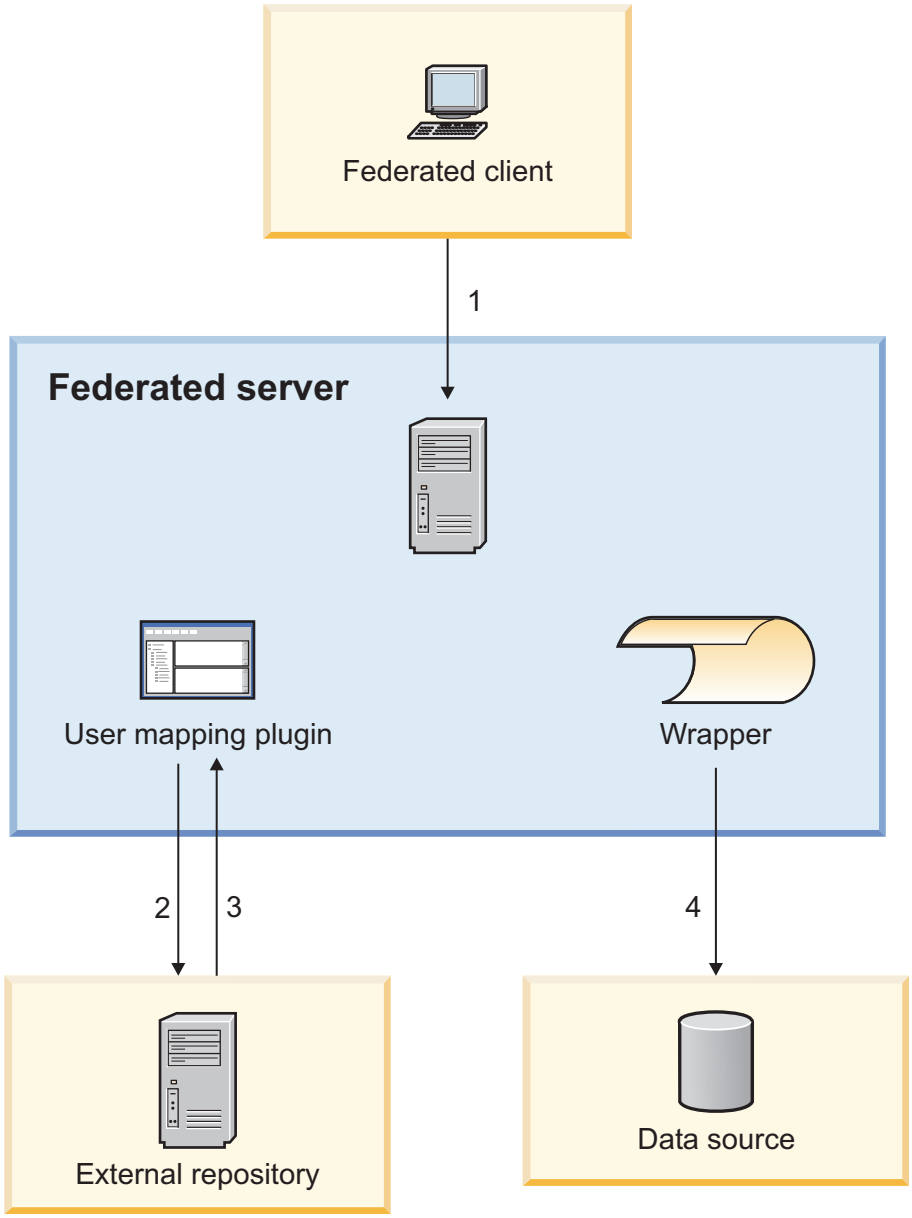


Figure 17. Relationship between the federated server, the user mapping plugin, the external repository and the data source.. When a DB2 client requests a connection through the federated server to data that resides on the remote data source, the local user ID of the client is sent to the federated server (1). If the DB2_UM_PLUGIN option is set for a wrapper, the federated server loads the user mapping plugin and provides the plugin with the local instance name, local database name, remote server name, and local user ID. The plugin uses this information to locate (2) the user mapping in the repository. The plugin retrieves the remote user mapping and performs any necessary decryption (3). The plugin returns the remote user ID and remote password to the federated server. The federated server tries to connect to the data source through the wrapper (4). The remote user ID and remote password that was retrieved from the external repository is used to connect to the data source.

User mapping plugin architecture

You can develop your own plugin for retrieving user mappings from an external repository. The UserMappingRepository, UserMappingCrypto, UserMappingEntry, UserMappingOption, and UserMappingException classes provide the interface and utilities to develop a plugin to return the user mappings from the external repository to the federated server.

The following table lists the interface classes and utility classes that comprise the user mapping plugin architecture. You must extend the interface classes to work with your external repository. The utility classes can be used without modification.

Table 22. Interface classes and utility classes

Interface classes	Utility classes
UserMappingCrypto class	UserMappingEntry class
UserMappingRepository class	UserMappingOption class
	UserMappingException class

Some functions or methods that are in the interface classes act as utility functions. For example, you can use the getChars() function and getBytes() function in the UserMappingCrypto class without modification.

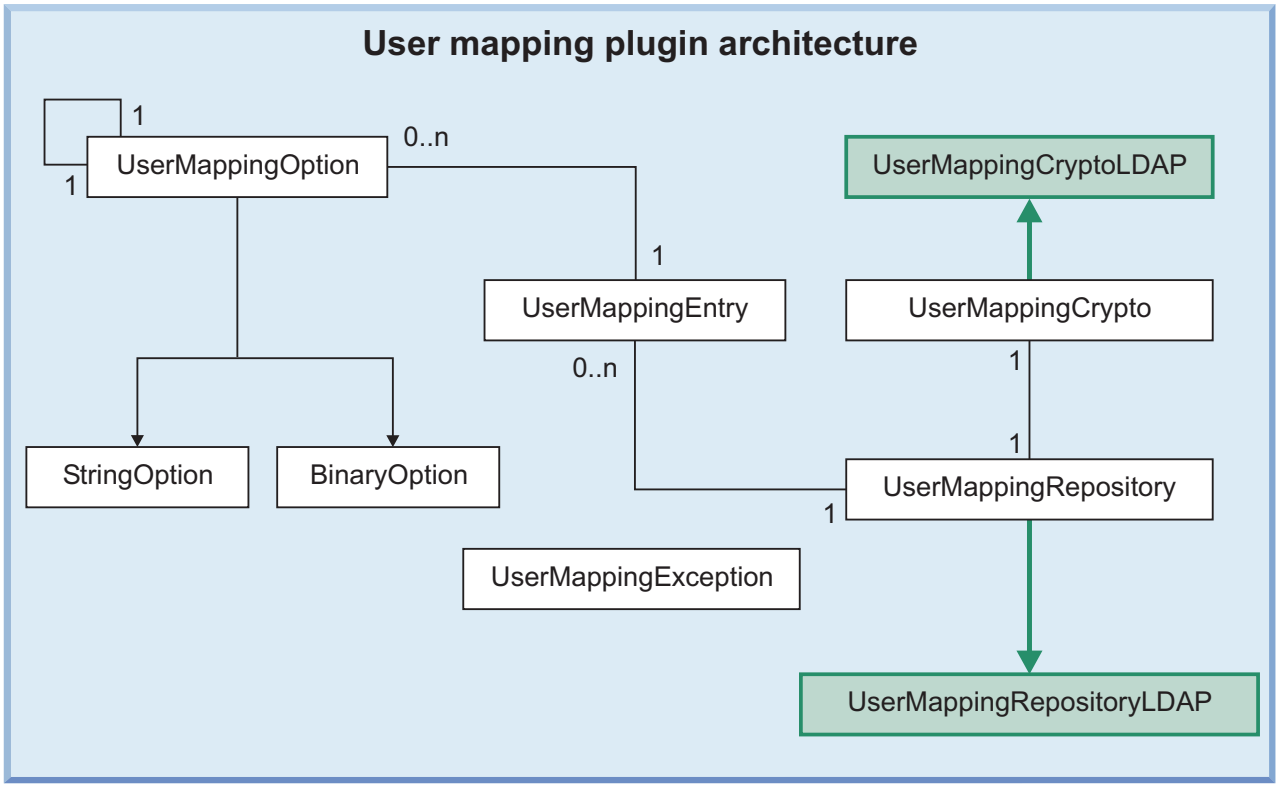


Figure 18. Diagram of the classes for the architecture of the user mapping plugin. The diagram shows the relationship of the classes to each other. The *0..n* term, means there can be multiple zero or more of those objects. For example, a UserMappingEntry object can have multiple UserMappingOption objects but there can be only be a single UserMappingRepository object. The UserMappingCryptoLDAP class and UserMappingRepositoryLDAP class are darker to show that they are extended from their parent classes.

UserMappingRepository class

The UserMappingRepository class is an abstract class that does not have a constructor. You must create a subclass of the UserMappingRepository class or modify the subclass in the LDAP sample plugin.

The UserMappingRepository class contains the following public methods: getVersionNumber(), getCrypto(), connect(), disconnect(), fetchUM(), and lookupUM(). You must create your own subclass of the UserMappingRepository

that contains these functions. In these functions, write the code that allows the plugin to interact with your external repository.

Public methods

int getVersionNumber()
Returns the version number of the plugin development kit that is used by the plugin.

UserMappingCrypto getCrypto()
Returns the UserMappingCrypto object that is associated with this UserMappingRepository object.

abstract void connect()
You must implement your own method for connecting to your repository within this function.

abstract void disconnect()
You must implement your own method for disconnecting from your repository within this function.

abstract void fetchUM(UserMappingEntry **um)**
You must implement your own method for retrieving the user mapping from your repository within this function. The **um** parameter contains the detailed query information that is used to determine which user mapping to retrieve.

UserMappingEntry lookupUM(UserMappingRepository repository, String **iiInstanceName**, String **iiDatabaseName**, String **iiRemoteServerName**, String **iiAuthid)**
This function is used primarily for testing the plugin. The function uses the **iiInstanceName**, **iiDatabaseName**, **iiRemoteServerName**, and **iiAuthid** parameters as input to create and initialize the UserMappingEntry class. The function calls the `connect`, `fetchUM`, and `disconnect` methods.

Sample plugin files

You can view the implementation of these functions in a sample plugin that retrieves user mappings from an LDAP server. The files are in the sqllib/samples/federated/umplugin/ldap/ directory. The functions from this class are used in the UserMappingRepositoryLDAP.java and UserMappingLookupLDAP.java files.

UserMappingCrypto class

If your external repository encrypts or encodes your remote passwords, you must create your own subclass of the UserMappingCrypto class. The constructor of the subclass that you create is used to construct the cryptography object. The methods of the cryptography class are called by other classes when the user mapping passwords need to be encrypted, decrypted, encoded, or decoded.

The UserMappingCrypto class contains the following public methods: encrypt(), decrypt(), encode(), and decode(). In these functions, you must write your code for encrypting, decrypting, encoding, and decoding the remote password. The getBytes() and getChars() functions are utility functions that are inherited and can be used without modification.

Your encryption, decryption, encoding, and decoding methods must match the encryption and encoding methods that are used by your external repository for protecting the stored passwords.

Public methods

abstract byte[] encrypt(byte[] **plainValue)**
Implement the encryption algorithm that matches the encryption algorithm that is used by your external repository.

abstract byte[] decrypt(byte[] **encryptedValue)**
Implement the decryption algorithm that reverses the encryption algorithm that is used by your external repository and returns the password.

abstract string encode(byte[] **bytes)**
Write or implement a function that encodes the **bytes** parameter into a string. This function encodes the encrypted value, which is in bytes, into a string.

abstract byte[] decode(String[] **string)**
Write or implement a function that decodes the **string** parameter into bytes. This function is used to decode the retrieved password, which is a string, into bytes so that the value can be decrypted.

byte[] getBytes(char[] **chars)**
This function is inherited and can be used without modification. The function transforms each character of a string into a byte.

char[] getChars(byte[] **bytes)**
This function is inherited and can be used without modification. The function transforms each byte into a character.

Protected attributes

SecretKey key
The secret key that is used to encrypt and decrypt the remote passwords.

Cipher cipher
The algorithm that is used to encrypt the password by using the secret key.

Sample plugin files

You can view the implementation of these functions in a sample plugin that retrieves user mappings from an LDAP server. The files are located in the sqllib/samples/federated/umplugin/ldap/ directory. The functions from this class are used in the UserMappingRepositoryLDAP.java and UserMappingSetupLDAP.java sample files.

UserMappingEntry class

The UserMappingEntry class is a utility class that creates and holds the user mapping options. The methods in the UserMappingEntry class are called by the fetchUM() and lookupUM() functions from the UserMappingRepository class.

The UserMappingEntry class contains the following public methods: UserMappingEntry(), getRepository(), getIIInstanceName(), getIIDatabaseName(), getIIRemoteServerName(), getIIAuthID(), getFirstOption(), and addOption().

Public methods

UserMappingEntry(UserMappingRepository **repository**, String **iiInstanceName**, String **iiDatabaseName**, String **iiRemoteServerName**, String **iiAuthID**)
This constructor is used to instantiate the UserMappingEntry object with the following input parameters:

- **iiInstance** - Instance name of the federated server

- **iiDatabase** - Database name on the federated server
- **iiRemoteServerName** - Remote server name for the data source
- **iiAuthid** - Local user ID that is associated with the user mapping

UserMappingRepository getRepository()
Returns the UserMappingRepository object that is associated with this UserMappingEntry.

string getIIInstanceName()
Returns the name of the instance.

string getIIDatabaseName()
Returns the name of the database.

string getIIRemoteServerName()
Returns the name of the remote server.

string getIIAuthID()
Returns the name of the local user ID that is associated with the user mapping.

UserMappingOption getFirstOption()
Returns the first UserMappingOption object that belongs to this UserMappingEntry object.

void addOption(UserMappingOption **newOption)**
Adds a new UserMappingOption object to this UserMappingEntry object.

Sample plugin files

You can view the use of these functions in a sample plugin that retrieves user mappings from an LDAP server. The files are in the sqllib/samples/federated/umplugin/ldap/ directory. The functions from this class are used in the UserMappingRepositoryLDAP.java and UserMappingLookupLDAP.java files.

UserMappingOption class

The UserMappingOption class is a utility class that contains the functions for providing the federated server with the two components of the user mapping: the remote user ID and the remote password.

The UserMappingOption class contains the following public methods: getEntry(), getName(), setName(), getNextOption(), setNextOption(), and getValue(). The sqllib/samples/federated/umplugin/ldap/UserMappingRepositoryLDAP.java sample file, contains a sample of using the UserMappingOption object in an LDAP sample plugin.

Public methods

UserMappingEntry getEntry()
Returns the UserMappingEntry object that is associated with this UserMappingOption object.

string getName()
Returns the name of the option.

void setName()
Sets the name of the option.

UserMappingOption getNextOption()
Returns the next option.

void setNextOption(UserMappingOption **nextOption)**
Sets the next option.

abstract object getValue()
Returns the value of the option. This function must be implemented to return either a string value or binary value. See the StringOption and BinaryOption methods that are listed below.

abstract object setValue()
Sets the value of the option. This function must be implemented to set either a string value or binary value. See the StringOption and BinaryOption methods that are listed below.

StringOption class that is extended from the UserMappingOption class

The StringOption class contains the following public methods: getValue() and setValue().

Public methods

object getValue()
Returns the value of the string option when the option is a string.

void setValue(string **value)**
Sets the value of the string option.

BinaryOption class that is extended from the UserMappingOption class

The BinaryOption class contains the following public methods: getValue() and setValue().

Public methods

object getValue()
Returns the value of the binary option.

void setValue(byte[] **value)**
Sets the value of the binary option.

Sample plugin files

You can view the use of these functions in a sample plugin that retrieves user mappings from an LDAP server. The files are in the sqllib/samples/federated/umplugin/ldap/ directory. The functions from this class are used in the UserMappingRepositoryLDAP.java and UserMappingLookupLDAP.java files.

UserMappingException class

The UserMappingException class is a subclass of the java.lang.Exception class and is used by the plugin to report errors.

The UserMappingRepository class contains the following public methods: UserMappingException(), getErrorNumber(), and getErrorMessage().

Public methods

UserMappingException(int **errorNumber)**
The constructor that is used for instantiating the UserMappingException object

that is used for reporting errors. The **errorNumber** parameter that is sent to the UserMappingException object defines the type of error that is reported.

getErrorNumber()
Returns the error number of the exception.

The UserMappingRepositoryLDAP.java file, contains this function for catching and reporting errors for testing an LDAP plugin.

getErrorMessage()
Returns the error message of the exception.

The UserMappingRepositoryLDAP.java file, contains a method for catching and reporting errors for testing an LDAP plugin.

Error messages

Possible error numbers, constant names, and error messages are listed in the following table.

Table 23. Error numbers and messages

Error number	Constant name	Error message
1	INITIALIZE_ERROR	The plugin failed to initialize.
2	CONNECTION_ERROR	Unable to connect to the repository.
3	AUTHENTICATION_ERROR	Unable to authenticate with the repository.
4	LOOKUP_ERROR	Lookup on the repository failed.
5	DECRYPTION_ERROR	Decryption failed.
6	DISCONNECT_ERROR	Unable to disconnect from the repository.
7	INVALID_PARAMETER_ERROR	Invalid parameter.
8	UNAUTHORIZED_CALLER	The caller is not authorized to call the plugin.

Sample plugin files

You can view the use of these functions in a sample plugin that retrieves user mappings from an LDAP server. The files are in the sqllib/samples/federated/umplugin/ldap/ directory. The functions from this class are used in each of the sample Java™ files.

LDAP sample plugin

A sample plugin for retrieving user mappings from a Lightweight Directory Access Protocol (LDAP) server is provided in the sqllib/samples/federated/umplugin/ldap/ directory. Before you can use the sample plugin, you must modify the plugin to match the settings of your LDAP server.

By default, user mappings are stored locally on each federated server for a data source. An LDAP server stores objects (for example, user entries) in a directory tree. These objects can have attributes (for example, passwords). The LDAP server is often used store information about users and their related information such as the e-mail address of a user.

The LDAP plugin is used to retrieve user mapping information from an LDAP server. Many user mappings can be stored on one LDAP server and be accessed by many federated servers. The LDAP system administrators can choose the level of security that they want to protect the user attributes, such as passwords or other sensitive information that should not be transmitted as plain text.

Because the directory structure and security settings can vary on each server, you must modify the sample plugin to match the settings of your LDAP server.

Description of files for the LDAP sample plugin

The sample plugin contains Java source code that allows the plugin to connect to a Lightweight Directory Access Protocol (LDAP) server, retrieve the user mappings, and decrypt the remote password. Because server settings can vary, you must develop your own LDAP plugin to match your server's encryption settings, directory structure, and other settings.

By using the files that are described below, you can develop a plugin for retrieving user mappings from an LDAP server or you can extend the sample to use with any external repository. The sample code is in the sqllib/samples/federated/umplugin/ldap/ directory. Before you begin to develop your plugin, copy the sample files to an empty working directory.

The sample plugin

You can develop your own LDAP plugin by using the sample code as a starting point. You can also find instructions in the sample folder for the LDAP plugin for testing the sample code with sample LDAP directory entries.

The sample plugin consists of four Java source files and two LDAP Lightweight Directory Interchange Format (LDIF) files. The following table contains a description of each file.

Table 24. Description of the files in the sample plugin

File name	Description
README.txt	This file contains a condensed version of the instructions and documentation for using the sample plugin.
UserMappingCryptoLDAP.java	This Java class contains the code that implements the security measures for encrypting, decrypting, encoding, and decoding the user mappings that are retrieved from the LDAP server. This file must be modified to work with your LDAP server.
UserMappingSetupLDAP.java	This Java class creates the configuration file that stores the LDAP connection information and other configuration parameters, including: IP address or host name, SSL or non-SSL, user ID, and password.
UserMappingRepositoryLDAP.java	This Java class contains the code for connecting, disconnecting, and fetching user mappings from the LDAP server. The code for this file uses the schema that is defined in the schema.ldif file. If you want to change the schema, you must also change a section in this file.

Table 24. Description of the files in the sample plugin (continued)

File name	Description
UserMappingLookupLDAP.java	This Java class contains the code to perform a LDAP lookup test. You can test your plugin using this file before testing the plugin on the federated server.
schema.ldif	This file is loaded into your LDAP server to define the schema. The Lightweight Directory Interchange Format (LDIF) file contains objects and attributes that are added to the LDAP server.
entry.ldif	The entry.ldif file adds user entries to the LDAP server. The user entries use the objects and attributes from the schema.ldif file to store the user mappings.

You can develop the plugin to use the settings that are appropriate for your LDAP server. For example, the encryption and encoding settings in the plugin must match the encryption and encoding settings on your LDAP server.

Developing a plugin for retrieving user mappings from an external repository

By using the sample code from the Lightweight Directory Access Protocol (LDAP) plugin as a starting point, you can develop a plugin for retrieving user mappings from an external repository. The sample files are located in the /sqllib/samples/federated/umplugin/ldap/ directory.

You should restrict access to the code for the encryption and decryption algorithms to protect the security of your repository and your remote passwords.

Before you begin

To develop a plugin you must have the following installed:

- The Java Development Kit (JDK) version 1.4 or later
- The db2umplugin.jar file. This Java Archive (JAR) file is installed with the DB2 server installation or the DB2 Application Development client installation.
- The LDAP sample files, which are installed in the /sqllib/samples/federated/umplugin/ldap/ directory. The sample files are installed with the DB2 Application Development Client installation.

About this task

The plugin that you develop must be able to connect to your repository, retrieve the user mappings, and decrypt the remote passwords. The repository that you are using might change the functions that you include in your plugin, for example, if you repository stores passwords as plain text then you would not need a decryption function. A sample plugin that you can reference is in the /sqllib/samples/federated/umplugin/ldap/ directory. The following procedure is a general guide for developing your plugin.

Procedure

To develop your plugin, the following steps are required:

Extending the sample LDAP plugin files to other external repositories

A sample plugin is provided for retrieving user mappings from a Lightweight Directory Access Protocol (LDAP) server. You can modify the plugin to retrieve user mappings from other repositories. The sample plugin provides a starting point to help you reduce your development time.

Each of the sample files accomplishes a different task in the process of retrieving user mappings. Many of the functions and classes in the sample code will need to be modified to work with a different repository. Refer to the user mapping plugin architecture when you develop your plugin. The important functions are listed in the following table.

Table 25. Functions and classes that must be modified in the sample plugin to extend the plugin to other repositories.

File name	Function or class name	Class to reference

UserMappingCryptoLDAP.java	`UserMappingCryptoLDAP()`
	`encrypt()`
	`decrypt()`
	`getKey()`
	`decode()`
	`encode()`
		UserMappingCrypto class
		UserMappingException class

UserMappingRepositoryLDAP.java	`class UserMappingRepositoryLDAP(String configFilePath)`
	`connect()`
	`disconnect()`
	`fetchUM()`
		UserMappingRepository class
		UserMappingEntry class
		UserMappingOption class
		UserMappingException class

UserMappingSetupLDAP.java	You will need to modify the file to create a configuration file that stores the values that your plugin requires for connecting to and retrieving user mappings from the external repository. If you choose to manually create a configuration file, ensure that any passwords that you store in the file are encrypted. The configuration file name should match the name of the repository class (for example, UserMappingRepositoryXXXX class and UserMappingRepositoryXXXX.cfg file).	

Security considerations for the user mapping plugin

When developing and using your plugin, you are sending sensitive user IDs and passwords between multiple sources. You can protect your information by restricting access to code, auditing plugin usage, and encrypting communication.

You must develop your plugin to use security settings that match the security settings that are used by your external repository. You should choose an external repository that allows you to encrypt sensitive information that is stored. The user mapping cryptography class must contain the encryption schema and secret key that allows for decrypting and decoding the passwords. You should restrict access to the source code of the plugin so that this information stays secure.

If auditing is turned on, any attempt to access a user mapping through plugin by the federated server has a VALIDATE audit record. You can configure the db2audit tool to capture VALIDATE records: `db2audit configure scope VALIDATE`.

You can also protect the communication between the federated server and the external repository by using a secure socket layer (SSL). When you create your configuration file for the user mapping plugin, you can specify that the plugin uses SSL to protect communications.

Modifying the UserMappingCryptoLDAP sample file

You need to modify the code that is in the UserMappingCryptoLDAP.java file to match the encryption or encoding that is used on your Lightweight Directory Access Protocol (LDAP) server. Because the encryption methods should be secret and unique, this task only directs you to the sections or functions that you must modify.

You should restrict access to the code for the encryption and decryption algorithms to protect the security of your repository and your remote passwords.

Before you begin

To use the user mapping plugin for LDAP servers, you need to know how the passwords are protected in the LDAP server.

About this task

You will need to write the code for implementing your security measures that match your LDAP configuration. The UserMappingCryptoLDAP.java file provides the functions for encrypting, decrypting, encoding, and decoding the remote passwords for the data sources.

Procedure

To modify the security functions in the UserMappingCryptoLDAP file:

1. Open the UserMappingCryptoLDAP.java file with a text editor.
2. Below the IBM copyright and legal disclaimer, import the packages that your code will reference. The sample plugin uses the javax.crypto and javax.crypto.spec Java packages. These packages provide the classes for the cipher (encrypting and decrypting) and the key and algorithm parameters.

   If you choose to use your own packages, replace these packages with your packages.
3. Update the following functions:

   `public UserMappingCryptoLDAP()`
   Replace the code for the cipher with the code for the cipher that matches the password encryption that is used by your LDAP server.

   `public byte[] encrypt(byte[] plainValue)`
   This function provides the code that encrypts the passwords so that they can be stored on the LDAP server. This function also encrypts the LDAP connection password that is stored in the configuration file.

   Replace this function's code with your own code that encrypts the **plainValue** parameter.

   `public byte[] decrypt(byte[] encryptedValue)`
   Replace this function's code with your own code that decrypts the **encryptedValue** parameter.

   `private SecretKey getKey()`

   Replace this function's code with the code to provide the plugin with the key that is used to encrypt and decrypt your passwords.

   `public byte[] decode(String string)`
   The passwords are first encrypted and then encoded. This function provides the code for decoding the passwords before the passwords are

decrypted. The encrypted passwords are encoded to transform the binary output of the encrypted password into ASCII characters.

Replace this function's code with your own code that decodes the **string** parameter.

public String encode(byte[] bytes)
The passwords are first encrypted and then encoded. This function provides the code for encoding the binary output of the encrypted passwords. The encrypted passwords are encoded to transform the binary output of the encrypted password into ASCII characters.

Replace this function's code with your own code that encodes the **bytes** parameter.

Modifying the UserMappingRepositoryLDAP sample file

The UserMappingRepositoryLDAP.java sample file contains the functions for connecting to the Lightweight Directory Access Protocol (LDAP) server, retrieving the user mappings, and disconnecting. You need to modify the sample code to match the object classes that are defined in the directory structure of your LDAP server.

About this task

To retrieve the user mappings from the LDAP server, the plugin must search the directory for the user entries that have the attributes that define the user mapping entry. The following information is stored as attributes of a user:

- Remote server name
- Instance name
- Database name
- Remote user name
- Remote user password

The sample code assumes that the user entry will be identified by the *inetOrgPerson* object class and that the user mapping entry will be identified by the *IIUserMapping* object class. The Lightweight Directory Interchange Format (LDIF) sample files (schema.ldif and entry.ldif) are used to load the sample schema and sample entries into the LDAP server.

The code in the UserMappingRepositoryLDAP.java file assumes that the schema.ldif file contains the LDIF code for objects and attributes with the following names:

- *IIUserMapping* object
  - *IIRemoteServerName* attribute
  - *IIInstanceName* attribute
  - *IIDatabaseName* attribute
  - *IIRemotePassword* attribute
  - *uid* attribute

You must modify the UserMappingRepositoryLDAP.java file to match the schema in your LDAP server. The UserMappingRepositoryLDAP.java file must search for and retrieve the user mapping entries from the LDAP server. The LDIF files are provided as a sample schema and method of storing user mapping entries.

Procedure

To modify the schema that is used by the UserMappingRepositoryLDAP.java file:

1. Locate the code: `private String UserObjectClassName = "inetOrgPerson"` and replace the *inetOrgPerson* value with the name of the object class for user entries that is used by your LDAP server.
2. **Optional:** Change the attribute names that are used by the plugin. Replace the values for the *IIRemoteServerAttrName*, *IIInstanceAttrName*, *IIDatabaseAttrName*, and *IIRemotePasswordAttrName* variables with the attribute names that you choose.
3. If you choose to use the LDIF files, ensure that the schema in the LDIF files matches the structure that is searched by the UserMappingRepositoryLDAP.java file.

Compiling the user mapping plugin files

When you finish modifying your source files for your user mapping plugin, you must compile the Java source files.

About this task

In the commands below, if your full path contains spaces, then the full path must be enclosed in quotation marks, for example, "C:\program files\sqllib\java\db2umplugin.jar". *%DB2PATH%* is the path to your Windows DB2 sqllib folder. *$DB2PATH* is the path to your the sqllib folder on UNIX.

The commands below assume that you are using the file naming convention that is based on the names of the classes. Replace the *XXXX* with the name that you chose.

Procedure

To compile the Java source files:

1. Issue the compile command:

   Windows:

   ```
   javac -classpath %DB2PATH%\sqllib\java\db2umplugin.jar; ^
          "%CLASSPATH%" -d . ^
          .\UserMappingRepositoryXXXX.java ^
          .\UserMappingCryptoXXXX.java ^
          .\UserMappingSetupXXXX.java ^
          .\UserMappingLookupXXXX.java
   ```

   UNIX:

   ```
   javac -classpath $DB2PATH/java/db2umplugin.jar: \
          $CLASSPATH -d . \
          ./UserMappingRepositoryXXXX.java \
          ./UserMappingCryptoXXXX.java \
          ./UserMappingSetupXXXX.java \
          ./UserMappingLookupXXXX.java
   ```

2. Archive the Java class files into a single Java Archive (JAR) file: The period symbol after the output file name, directs the command to find and place the files in the same directory. If you change the directory, use the appropriate file path for your operating system (for example, /home/user/folder or C:\test\folder).

   ```
   jar -cfM0 UserMappingRepositoryXXXX.jar .
   ```

Creating the configuration file for the user mapping plugin

The configuration file for the sample LDAP plugin stores the connection information for your LDAP server. After compiling the Java source files, you can run the configuration program that prompts you to enter your LDAP connection information.

About this task

The application uses your input to create the UserMappingRepositoryLDAP.cfg file to store the configuration information. If a password is required to connect to your LDAP server, the password is stored in the configuration file. The password is encrypted with the algorithm that you provided in the UserMappingCryptoLDAP.java file.

In the commands below, if your full path contains spaces, then the full path must be enclosed in quotation marks, for example, "C:\program files\sqllib\java\ db2umplugin.jar". *%DB2PATH%* is the path to your Windows DB2 sqllib folder. *$DB2PATH* is the path to your the sqllib folder on UNIX.

The application prompts you to enter the following configuration information:

- Host name or IP address of the LDAP server
- Port number (default is 389) of the LDAP server
- Distinguished Name of the LDAP subtree (for example, ou=ii,o=ibm,c=us)
- User ID for connecting to the LDAP server
- Password for connecting to the LDAP server
- SSL configuration

Procedure

To create the configuration file for the sample LDAP plugin:

Windows:

```
java -classpath %DB2PATH%\sqllib\java\db2umplugin.jar; ^
 .\UserMappingRepositoryLDAP.jar;"%CLASSPATH%" UserMappingSetupLDAP
```

UNIX:

```
java -classpath $DB2PATH/sqllib/java/db2umplugin.jar: \
 ./UserMappingRepositoryLDAP.jar:$CLASSPATH UserMappingSetupLDAP
```

Testing the user mapping plugin

You can test the functionality of your user mapping plugin outside of the federated server. To test outside of the federated server, you can develop an application that attempts to connect to and retrieve user mappings from your external repository.

About this task

You can develop a simple program that calls the `lookupUM()` method that your UserMappingRepositoryXXXX class inherited from the UserMappingRepository class to connect and retrieve user mappings from your external repository. You can view the UserMappingLookupLDAP.java file that is located in the sqllib/samples/federated/umplugin/ldap/ directory.

In the commands below, if your full path contains spaces, then the full path must be enclosed in quotation marks, for example, "C:\program files\sqllib\java\db2umplugin.jar". *%DB2PATH%* is the path to your Windows DB2 sqllib folder. *$DB2PATH* is the path to your the sqllib folder on UNIX.

The test program must take the parameters that are required for identifying the user mapping:

- *remoteServerName* - Remote server name for the data source
- *iiAuthid* - Local user ID that is associated with the user mapping
- *iiInstance* - Instance name of the federated server
- *iiDatabase* - Database name on the federated server

Procedure

To test the user mapping plugin:

Windows:

```
java -classpath %DB2PATH%\sqllib\java\db2umplugin.jar; ^
  .\UserMappingRepositoryXXXX.jar;"%CLASSPATH%" ^
  UserMappingLookupXXXX -server remoteServerName ^
   -authid iiAuthid ^
  -instance iiInstance -database iiDatabase
```

UNIX:

```
java -classpath $DB2PATH/sqllib/java/db2umplugin.jar: ^
  ./UserMappingRepositoryXXXX.jar:$CLASSPATH \
  UserMappingLookupXXXX -server remoteServerName \
  -authid iiAuthid \
  -instance iiInstance -database iiDatabase
```

Installing the user mapping plugin files

After compiling and testing the Java files for your user mapping plugin, you must install the files on the federated server.

About this task

In the commands below, *sqllib* is the full path to your DB2 installation.

The commands below assume that you are using the file naming convention that is based on the names of the classes. Replace the *XXXX* with the name that you chose.

Procedure

To install the compiled files:

Copy your UserMappingRepository*XXXX*.jar and UserMappingRepository*XXXX*.cfg files into the *sqllib*/function/directory.

Placing the files in this directory allows the federated server to load and call the classes contained in those files.

You can now configure the wrapper or server option on your federated server to use the user mapping plugin to retrieve the user mappings from your repository.

Configuring access to the user mapping plugin

You must set the DB2_UM_PLUGIN option to configure the federated server to access the user mappings in the external repository.

Before you begin

Before you configure the federated server to access the user mappings in an external repository, you must:

- Develop a user mapping plugin
- Install the plugin on the federated server
- Update your database manager configuration:

```
db2 update dbm cfg using JDK_PATH <your_jdk_path>
db2 terminate
db2stop
db2start
```

About this task

The DB2_UM_PLUGIN option must contain the full path of the class, including the package name. If you develop a package, you must include the package name before the class name for example, `'package.classname'`. If you use the LDAP sample plugin, specify the value `'UserMappingRepositoryLDAP'` in the DB2_UM_PLUGIN option. The sample plugin is not developed as a package.

Where you specify the DB2_UM_PLUGIN option impacts how the federated server uses the user mapping plugin:

- If you specify the DB2_UM_PLUGIN option on a wrapper, all of the servers that use the wrapper will use the value in that DB2_UM_PLUGIN wrapper option.
- If you specify the DB2_UM_PLUGIN option on a server definition, that server will use the value in that DB2_UM_PLUGIN server option.
- If you specify the DB2_UM_PLUGIN option on both a wrapper and server definition, the value that is specified on the server definition takes precedence over the value that is specified on the wrapper. A warning message, SQL20351W, is returned that indicates that the server option is not overwritten by the wrapper option.
- If the DB2_UM_PLUGIN is specified for a wrapper or server option, you can still use DDL to create a user mapping. A warning message, SQL20352W, is returned that indicates that the user mapping is stored in the federated database catalog but will not be used to access the remote server.

Procedure

To configure the federated server to access the user mapping external repository:

Choose how you want to implement the user mapping plugin:

Method	SQL statement

Specify the user mapping plugin when you create a wrapper	`CREATE WRAPPER net8`
	`  OPTIONS (`
	`  DB2_UM_PLUGIN 'UserMappingRepositoryLDAP'`
	`  );`

Method	SQL statement

Alter an existing wrapper to specify the user mapping plugin	`ALTER WRAPPER drda (`
	`  ADD DB2_UM_PLUGIN 'UserMappingRepositoryLDAP'`
	`  );`

Specify the user mapping plugin when you create a server definition	`CREATE SERVER server_name`
	`  TYPE date_source_type`
	`  VERSION version_number`
	`  WRAPPER net8 OPTIONS (`
	`  DB2_UM_PLUGIN 'UserMappingRepositoryLDAP'`
	`  );`

Alter an existing server definition to specify the user mapping plugin	`ALTER SERVER server_name`
	`  OPTIONS (`
	`  ADD DB2_UM_PLUGIN 'UserMappingRepositoryLDAP'`
	`  );`

After you set the DB2_UM_PLUGIN option, the federated server uses the plugin to retrieve user mappings from the external repository server by using the connection information that you specify in the UserMappingRepository*XXXX*.cfg file. *XXXX* is the name of the plugin.

To alter the wrapper to use a different plugin

```
ALTER WRAPPER drda (
  SET DB2_UM_PLUGIN 'com.package_name.um.UserMappingRepositoryXXXX'
  );
```

To alter a server definition to use a different plugin

```
ALTER SERVER server_name
  OPTIONS (
     SET DB2_UM_PLUGIN 'UserMappingRepositoryXXXX'
     );
```

Troubleshoot the user mapping plugin

After you configure the federated server to use the user mapping plugin, there are several ways that you can troubleshoot problems.

Before you test the plugin in the federated server, you should develop a Java application to test if the plugin can connect to and retrieve user mappings from the external repository. If the test is successful, then you can try to integrate the plugin with the federated server.

JAVA_HEAP_SZ parameter

The **JAVA_HEAP_SZ** parameter determines the maximum size of the heap that is used by the Java interpreter to service Java stored procedures and user defined functions.

If you encounter memory problems, try setting the heap size to 1024 or higher.

Error messages

If an error message is returned when using the plugin, the information in the error message can help you to determine the problem. There are reason codes and user responses in the error messages that should help you determine the problem.

SQL20349N

The federated server cannot access the user mappings that are in the external repository. One reason why you might receive this error is that the

user mapping plugin is not in the correct location. If you use the sample LDAP plugin on a UNIX federated server, ensure that UserMappingRepositoryLDAP.jar file and the UserMappingRepositoryLDAP.cfg file are in sqllib/function directory.

SQL20350N
The authentication at the user mapping repository failed. The proper credentials, such as passwords or certificates, to retrieve the user mappings were not communicated to the user mapping repository.

SQL20351W
You specified the DB2_UM_PLUGIN option on a wrapper and the DB2_UM_PLUGIN option is already specified on a server definition that uses the wrapper. The value that is set on the server definition takes precedence over the value that you set on the wrapper.

SQL20352W
You specified DB2_UM_PLUGIN option for the server. User mappings for this server are read from the external user mapping repository. The CREATE USER MAPPING, ALTER USER MAPPING, and DROP USER MAPPING statements only affect user mappings in the federated catalog table.

db2diag.log file

The diagnostic log is a file that contains text information that is logged by the federated database. This information is used for problem determination and is intended for IBM customer support.

You can view the file with a text editor. You can also use the db2diag tool can help you to analyze the information that is contained in the db2diag.log file to troubleshoot problems with your user mapping plugin.

Example 1
In the db2diag.log file, search for the term `JNI_Env::check_exception`. This term marks the first error record for the user mapping plugin. The record indicates the line in the plugin source code where an error has occurred. In the following example, the line in the code that is in error is line 119 of the UserMappingRepositoryLDAP.java file.

```
2006-02-18-19.05.43.966176-480 I6780A253          LEVEL: Severe
PID     : 3932214              TID  : 2828        PROC : db2fmp (Java)
INSTANCE: einstein               NODE : 000
FUNCTION: DB2 UDB, Query Gateway, Sqlqg_JNI_Env::check_exception, probe:150
2006-02-18-19.05.43.966953-480 I7034A388          LEVEL: Error
PID     : 3932214              TID  : 2828        PROC : db2fmp (Java)
INSTANCE: einstein               NODE : 000
FUNCTION: DB2 UDB, Query Gateway, Sqlqg_JNI_Env::check_exception, probe:190
DATA #1 : String, 108 bytes
com.ibm.ii.um.UserMappingException
    at UserMappingRepositoryLDAP.<init>(UserMappingRepositoryLDAP.java:119)
```

Example 2
In the db2diag.log file, search for the term `Error fetching user mappings`. The record for this error includes the plugin name, the instance name, the database name, the remote server name, and the local authentication ID. These values are passed from the federated server to the plugin as input. Ensure that the values , including the case sensitivity, are correct.

```
2006-02-18-19.05.43.967308-480 I7423A371          LEVEL: Severe
PID     : 3932214              TID  : 2828        PROC : db2fmp (Java)
INSTANCE: einstein               NODE : 000
```

```
MESSAGE : Error fetching user mappings through plugin:
DATA #1 : String, 110 bytes
   plugin=UserMappingRepositoryLDAP,instance=EINSTEIN,database=SAMPLE,
   remote_server=DRDASERVER1,local_authid=NEWTON
```

db2trc - trace command

If you are unable to troubleshoot the problem. you can use db2trc command to generate a trace. You will need to send the trace record to IBM for analysis.

Oracle Label Security and federated systems

When you create nicknames against data source objects with Oracle Label Security, the nicknames are not cached and materialized query tables cannot be created on them. The secure data remains on the Oracle system and only users with the appropriate authority can see it.

Administrators can use Oracle Label Security to apply security policies to each row in a table. The security policies determine a user's level of access to the data in the table that is based on the authorities that are granted to their user ID or session ID. See the documentation about Oracle Label Security from the Oracle Corporation for detailed information about configuring and managing security policies and labels.

Federated access to tables with Oracle Label Security

When you create a nickname on an Oracle data source object, the federated server automatically detects whether the data source uses Oracle Label Security. If Oracle Label Security is being used, the nickname is not cached. You can use the ALTER NICKNAME statement to allow or disallow caching. For example, if you created a nickname on a data source object with Oracle Label Security before federated support for the feature was available, you can alter the nickname to disallow caching. If you created a nickname on a data source object with Oracle Label Security and Oracle Label Security is removed, you can alter the nickname to allow caching.

A database administrator can choose to hide a label to prevent some users from knowing that it exists. In this case, the Label column is hidden in the table. Nicknames with hidden label columns are not cached.

Chapter 29. Federated system and data source configuration parameters

Views in the global catalog table containing federated information

Most of the catalog views in a federated database are the same as the catalog views in any other DB2 Version 9.1 for Linux, UNIX, and Windows database.

There are several unique views that contain information pertinent to a federated system, such as the SYSCAT.WRAPPERS view.

The DB2 Version 8 SYSCAT views are read-only. If you issue an UPDATE or INSERT operation on a view in the SYSCAT schema, it will fail. Using the SYSSTAT views is the recommended way to update the system catalog. Change applications that reference the SYSCAT view to reference the updatable SYSSTAT view instead.

The following table lists the SYSCAT views which contain federated information. These are read-only views.

Table 26. Catalog views typically used with a federated system

Catalog views	Description
SYSCAT.CHECKS	Contains check constraint information that you defined.
SYSCAT.COLCHECKS	Contains columns referenced by a check constraint.
SYSCAT.COLUMNS	Contains column information about the data source objects (tables and views) that you created nicknames for.
SYSCAT.COLOPTIONS	Contains information about column option values that you set for a nickname.
SYSCAT.CONSTDEP	Contains the dependency of an informational constraint that you defined.
SYSCAT.DATATYPES	Contains data type information about local built-in and user-defined DB2 data types.
SYSCAT.DBAUTH	Contains the database authorities held by individual users and groups.
SYSCAT.FUNCMAPOPTIONS	Contains information about option values that you have set for a function mapping.
SYSCAT.FUNCMAPPINGS	Contains the function mappings between the federated database and the data source objects.
SYSCAT.INDEXCOLUSE	Contains columns that participate in an index.
SYSCAT.INDEXES	Contains index specifications for data source objects.
SYSCAT.INDEXOPTIONS	Contains information about index options.

Table 26. Catalog views typically used with a federated system (continued)

Catalog views	Description
SYSCAT.KEYCOLUSE	Contains columns that participate in a key defined by a unique key, primary key, or foreign key constraint.
SYSCAT.NICKNAMES	Contains information about nicknames that you created.
SYSCAT.REFERENCES	Contains information about referential constraints that you defined.
SYSCAT.ROUTINES	Contains local DB2 user-defined functions, or function templates. Function templates are used to map to a data source function.
SYSCAT.REVTYPEMAPPINGS	This view is not used. All data type mappings are recorded in the SYSCAT.TYPEMAPPINGS view.
SYSCAT.ROUTINEOPTIONS	Contains information about federated routine option values.
SYSCAT.ROUTINEPARMOPTIONS	Contains information about federated routine parameter option values.
SYSCAT.ROUTINEPARMS	Contains a parameter or the result of a routine defined in SYSCAT.ROUTINES.
SYSCAT.ROUTINESFEDERATED	Contains information about federated routines that you defined.
SYSCAT.SERVERS	Contains server definitions that you create for data source servers.
SYSCAT.TABCONST	Each row represents a table and nickname constraints of type CHECK, UNIQUE, PRIMARY KEY, or FOREIGN KEY.
SYSCAT.TABLES	Contains information about each local DB2 table, federated view, and nickname that you create.
SYSCAT.TYPEMAPPINGS	Contains forward data type mappings and reverse data type mappings. The mapping is to local DB2 data types from data source data types. These mappings are used when you create a nickname on a data source object.
SYSCAT.USEROPTIONS	Contains user authorization information that you set when you create user mappings between the federated database and the data source servers.
SYSCAT.VIEWS	Contains information about local federated views that you create.
SYSCAT.WRAPOPTIONS	Contains information about option values that you have set for a wrapper.
SYSCAT.WRAPPERS	Contains the name of the wrapper and library file for each data source that you create a wrapper for.

The following table lists the SYSSTAT views which contain federated information. These are read-write views that contain statistics you can update.

Table 27. Federated updatable global catalog views

Catalog views	Description
SYSSTAT.COLUMNS	Contains statistical information about each column in the data source objects (tables and views) that you have created nicknames for. Statistics are not recorded for inherited columns of typed tables.
SYSSTAT.INDEXES	Contains statistical information about each index specification for data source objects.
SYSSTAT.ROUTINES	Contains statistical information about each user-defined function. Does not include built-in functions. Statistics are not recorded for inherited columns of typed tables.
SYSSTAT.TABLES	Contains information about each base table. View, synonym, and alias information is not included in this view. For typed tables, only the root table of a table hierarchy is included in the view. Statistics are not recorded for inherited columns of typed tables.

Nickname column options for federated systems

You can specify column information in the CREATE NICKNAME or ALTER NICKNAME statements using parameters called nickname column options.

The following table lists the nickname column options for data source. Table two contains a complete listing of nickname column options.

Table 28. Nickname column options for relational data sources

Data source	DOCUMENT	ESCAPE_ INPUT	FOREIGN_ KEY	NUMERIC_ STRING	PRIMARY_ KEY	TEMPLATE	XPATH
DB2 UDB for iSeries				X			
DB2 UDB for z/OS and OS/390				X			
DB2 for VM and VSE				X			
DB2 Version 9.1 for Linux, UNIX, and Windows				X			
Informix				X			
Microsoft SQL Server				X			
ODBC				X			
OLE DB							
Oracle				X			
Sybase				X			
Teradata				X			

Table 29. Nickname column options for nonrelational data sources

Options	BLAST	Script	Table-structured files	WebSphere Business Integration	Web services	XML
DELIMITER	X					
DOCUMENT			X			X

Table 29. Nickname column options for nonrelational data sources (continued)

Options	BLAST	Script	Table-structured files	WebSphere Business Integration	Web services	XML
ESCAPE_INPUT				X	X	
FINAL_XDROPOFF	X					
FOREIGN_KEY				X	X	X
INDEX	X					
INPUT_MODE		X				
MASK_LOWER_CASE	X					
POSITION		X				
PRIMARY_KEY				X	X	X
QUERY_GENETIC_CODE	X					
SWITCH		X				
SWITCH_ONLY		X				
TEMPLATE				X	X	
VALID_VALUES		X				
XDROPOFF_GAPPED	X					
XDROPOFF_UNGAPPED	X					
XPATH				X	X	X

Table 30. Column options and their settings

Option	Description and valid settings	Default setting
DEFAULT	Specifies a new default value for the following input fixed columns: • E_value • QueryStrands • GapAlign • NMisMatchPenalty • NMatchReward • Matrix • FilterSequence • NumberOfAlignments • GapCost • ExtendedGapCost • WordSize • ThresholdEx This new value overrides the preset default values. The new default value must be of the same type as the indicated value for a given column.	

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
DELIMITER	The delimiter characters to be used to determine the end point of the definition line information for the column on which this option appears. If more than one character appears in this option's value, then the first occurrence of any one of the characters signals the end of this field's information. The default is end of line. This option is required, unless you want the last specified column to contain the remainder of the definition line.	The default delimiter is end of line.

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
DOCUMENT	For table-structured files: Specifies the kind of table-structured file. This wrapper supports only the value FILE for this option. Only one column can be specified with the DOCUMENT option per nickname. The column that is associated with the DOCUMENT option must be a data type of VARCHAR or CHAR. Using the DOCUMENT nickname column option instead of the FILE_PATH nickname option implies that the file that corresponds to this nickname will be supplied when the query runs. If the DOCUMENT option has the FILE value, the value that is supplied when the query runs is the full path of the file whose schema matches the nickname definition for this nickname. For XML: Specifies that this column is a DOCUMENT column. The value of the DOCUMENT column indicates the type of XML source data that is supplied to the nickname when the query runs. This option is accepted only for columns of the root nickname (the nickname that identifies the elements at the top level of the XML document). Only one column can be specified with the DOCUMENT option per nickname. The column that is associated with the DOCUMENT option must be a VARCHAR data type. If you use a DOCUMENT column option instead of the FILE_PATH or DIRECTORY_PATH nickname option the document that corresponds to this nickname is supplied when the query runs. The valid values for the DOCUMENT option are: **FILE** Specifies that the value of the nickname column is bound to the path name of a file. The data from this file is supplied when the query runs. **DIRECTORY** Specifies that the value of the nickname column is bound to the path name of a directory that contains multiple XML data files. The XML data from multiple files is supplied when the query runs. The data is in XML files in the specified directory path. The XML wrapper uses only the files with an .xml extension that are located in the directory that you specify. The XML wrapper ignores all other files in this directory. **URI** Specifies that the value of the nickname column is bound to the path name of a remote XML file to which a URI refers. The URI address indicates the remote location of this XML file on the Web. The URI can contain a colon-separated IPv6 address if it is enclosed in square brackets (for example, http://[1080:0:0:0:8:800:200C:417A]). **COLUMN** Specifies that the XML document is stored in a relational column.	

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
ELEMENT_NAME	Specifies the BioRS element name. The case sensitivity of this name depends on the case sensitivity of the BioRS server and on the value of the CASE_SENSITIVE server option. You must specify the BioRS element name only if it is different from the column name.	
ESCAPE_INPUT	Specifies whether XML special characters are replaced in XML input values or not. Use this option to include XML fragments as input, such as XML fragments with repeating elements. The TEMPLATE column option must be defined on columns that use the ESCAPE_INPUT column option. The column data type must be VARCHAR or CHAR. Valid values are: **Y** If the XML input contains special characters these are replaced with their counterpart characters that XML uses to represent the input characters. **N** Input characters are preserved exactly as they appear.	Y
FINAL_XDROPOFF	The X dropoff value for the final gapped alignment, measured in bits. The value 0 invokes the default behavior.	50 bits for blastn and megablast queries. 0 bits for tblastx queries. 25 bits (INTEGER data types) for all other query types.
FOREIGN_KEY	Indicates that this nickname is a child nickname and specifies the name of the corresponding parent nickname. A nickname can have at most one FOREIGN_KEY column option. The value for this option is case sensitive. The table that is designated with this option holds a key that is generated by the wrapper. The XPATH option must not be specified for this column. The column can be used only to join parent nicknames and child nicknames. A CREATE NICKNAME statement with a FOREIGN_KEY option will fail if the parent nickname has a different schema name. Unless the nickname that is referred to in a FOREIGN_KEY clause was explicitly defined as lowercase or mixed case by enclosing it in quotation marks in the corresponding CREATE NICKNAME statement, then when you refer to this nickname in the FOREIGN_KEY clause, you must specify the nickname in uppercase. When this option is set on a column, no other option can be set on the column.	
INDEX	The ordinal number of the column on which this option appears in the group of definition line columns. This option is required.	
INPUT_MODE	Specifies the input mode for a column. Valid values are CONFIG or FILE_INPUT. The wrapper passes the specified value to the script daemon.	

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
IS_INDEXED	Indicates whether the corresponding column is indexed (whether the column can be referenced in a predicate). The valid values are Y and N. The value Y can be specified only for columns whose corresponding element is indexed by the BioRS server.	When a nickname is created, this option is automatically added with the value Y to any columns that correspond to a BioRS indexed element.
MASK_LOWER_CASE	Use lowercase filtering with a FASTA sequence.	
NUMERIC_STRING	Specifies whether a column contains strings of numeric characters. **Y** This column contains strings of numeric characters '0', '1', '2', ... '9'. It does not contain blanks. If this column contains only numeric strings followed by trailing blanks, do not specify Y. When you set NUMERIC_STRING to Y for a column, you are informing the optimizer that this column contains no blanks that could interfere with sorting of the column's data. Use this option when the collating sequence of a data source is different from the collating sequence that the federated server uses. Columns that use this option are not excluded from remote evaluation because of a different collating sequence. **N** This column is either not a numeric string column or is a numeric string column that contains blanks.	N
POSITION	An integer value for positional parameters. This option applies only to input columns. If the positional value is set to an integer, then this input must be in this position in the command line. If this option is set, the switch is inserted into the appropriate location when the query is run. If POSITION is set to -1, the option is added as the last command line option. POSITION integer values cannot be duplicated in a nickname. This option is not required.	
PRIMARY_KEY	Indicates that this nickname is a parent nickname. The column data type must be VARCHAR(16). A nickname can have at most one PRIMARY_KEY column option. YES is the only valid value. The column that is designated with this option holds a key that is generated by the wrapper. The XPATH option must not be specified for this column. The column can be used only to join parent nicknames and child nicknames. When this option is set on a column, no other option can be set on the column.	
QUERY_GENETIC_CODE	Query genetic code uses default = 1.	
REFERENCED_OBJECT	This option is valid only for columns whose BioRS data type is Reference. This option specifies the name of the BioRS databank that is referenced by the current column. The case sensitivity of this name depends on the case sensitivity of the BioRS server and on the value of the CASE_SENSITIVE server option.	

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
SOAPACTIONCOLUMN	A column to dynamically specify the URI SOAPACTION attribute from the Web Service Description Language (WSDL) format. This option is specified on only the root nickname. When this option is set on a column, no other option can be set on the column. The URL can contain a colon-separated IPv6 address if it is enclosed in square brackets (for example, http://[1080:0:0:0:8:800:200C:417A]).	
SWITCH	A character string to specify a parameter for the script on the command line. This option applies only to input columns.	
SWITCH_ONLY	Enables the use of switches without a command line argument. If the SWITCH_ONLY option is specified with a value of Y, then valid input values are Y or N. For an input value of Y, only the switch is added to the command line. For an input value of N, no value is added to the command line.	
TEMPLATE	The column template fragment to use to construct the XML input document. The fragment must conform to the specified template syntax.	
URLCOLUMN	A column to dynamically specify the URL for the Web service endpoint when you run a query. This option is specified on only the root nickname. When this option is set on a column, no other option can be set on the column. The URL can contain a colon-separated IPv6 address if it is enclosed in square brackets (for example, http://[1080:0:0:0:8:800:200C:417A]).	
VALID_VALUES	A semicolon-separated set of valid values for a column.	

Table 30. Column options and their settings (continued)

Option	Description and valid settings	Default setting
VARCHAR_NO_ TRAILING_BLANKS	This option applies to data sources that have variable character data types that do not pad the length with trailing blanks during comparison. Some data sources, such as Oracle, do not have blank-padded character comparison semantics that return the same results as the DB2 for Linux, UNIX, and Windows comparison semantics. Set this option when you want it to apply only to a specific VARCHAR or VARCHAR2 column in a data source object. **Y** Trailing blanks are absent from these VARCHAR columns, or the data source has blank-padded character comparison semantics that are similar to the semantics on the federated server. The federated server sends character comparison operations to the data source for processing. **N** Trailing blanks are present in these VARCHAR columns, and the data source has blank-padded character comparison semantics that are different than the federated server. The federated server processes character comparison operations if it is not possible to compensate for equivalent semantics. For example, rewriting the predicate.	N for affected data sources
XDROPOFF_GAPPED	The X dropoff value for gapped alignment, measured in bits. The value 0 invokes the default behavior.	30 bits for blastn queries. 20 bits for megablast queries. 15 bits INTEGER data types) for all other query types.
XDROPOFF_UNGAPPED	The X dropoff value for ungapped extension measured in bits. The value 0.0 invokes the default behavior. For blastn queries, the default is 20 bits. For megablast queries, the default is 10 bits. For all other query types, the default is 7 bits (REAL data types).	20 bits for blastn queries. 10 bits for megablast queries. 7 bits (REAL data types) for all other query types.
XPATH	Specifies the XPath expression in the XML document that contains the data that corresponds to this column. The wrapper evaluates the XPath expression after the CREATE NICKNAME statement applies this XPath expression from this XPATH nickname option.	

Function mapping options for federated systems

The primary purpose of function mapping options, is to provide information about the potential cost of executing a data source function at the data source.

WebSphere Federation Server supplies default mappings between existing built-in data source functions and built-in DB2 functions. For most data sources, the default function mappings are in the wrappers. To use a data source function that the federated server does not recognize, you must create a function mapping between a data source function and a counterpart function at the federated database.

Pushdown analysis determines if a function at the data source is able to execute a function in a query. The query optimizer decides if pushing down the function processing to the data source is the least cost alternative.

The statistical information provided in the function mapping definition helps the query optimizer compare the estimated cost of executing the data source function with the estimated cost of executing the DB2 function.

Table 31. Function mapping options and their settings

Option	Valid settings	Default setting
DISABLE	Disable a default function mapping. Valid values are 'Y' and 'N'.	'N'
INITIAL_INSTS	Estimated number of instructions processed the first and last time that the data source function is invoked.	'0'
INITIAL_IOS	Estimated number of I/Os performed the first and last time that the data source function is invoked.	'0'
IOS_PER_ARGBYTE	Estimated number of I/Os expended for each byte of the argument set that's passed to the data source function.	'0'
IOS_PER_INVOC	Estimated number of I/Os per invocation of a data source function.	'0'
INSTS_PER_ARGBYTE	Estimated number of instructions processed for each byte of the argument set that's passed to the data source function.	'0'
INSTS_PER_INVOC	Estimated number of instructions processed per invocation of the data source function.	'450'
PERCENT_ARGBYTES	Estimated average percent of input argument bytes that the data source function will actually read.	'100'
REMOTE_NAME	Name of the data source function.	local name

Nickname options for federated systems

Some nickname options are required and cannot be dropped. Other nickname options cannot be added if specific nickname options are already set.

Table 32 on page 300 list the nickname options for each data source. Table 33 on page 300 describes each nickname option and lists the valid and default values.

Table 32. Available nickname options

Data source	BUSOBJ_NAME	COLUMN_DELIMITER	DATASOURCE	FILE_PATH	NAMESPACES	REMOTE_OBJECT	REQUIRE_PREDICATE	RESULTSET_SIZE	STREAMING	TEMPLATE	TIMEOUT	VALIDATE	XPATH
BioRS						X							
BLAST			X								X		
Entrez						X							
Excel				X									
HMMER			X								X		
Script			X		X				X		X	X	X
Table- structured files		X		X									
Web services					X				X	X			X
WebSphere Business Integration	X				X		X	X		X			X
XML				X	X				X				X

Table 33 describes each nickname option and lists the valid and default settings.

Table 33. Nickname options and their settings

Option	Description and valid settings	Default setting
BUSOBJ_NAME	The name of the XML schema definition file (.xsd) that represents the business object. For example sap_bapi_customer_get_detail2 . This option must be specified in a parent nickname.	
COLUMN_ DELIMITER	The delimiter that is used to separate columns of a table-structured file, enclosed in single quotation marks. The column delimiter can be more than one character in length. If no column delimiter is defined, the default delimiter is a comma. A single quotation mark cannot be used as a delimiter. The column delimiter must be consistent throughout the file. A null value is represented by two delimiters next to each other or a delimiter followed by a line terminator, if the NULL field is the last one on the line. The column delimiter cannot exist as valid data for a column.	The default delimiter is a comma.

Table 33. Nickname options and their settings (continued)

Option	Description and valid settings	Default setting
DATASOURCE	For BLAST: The name of the data source on which the BLAST search will run. The same string that is used here must be present in the configuration file of the BLAST daemon. This option is required. For HMMER (type PFAM): The name of the HMM Profile database that is to be searched by HMMPFAM. The same string that is used here must be present in the configuration file of the HMMER daeamon. This option is required. For HMMER (type SEARCH): The name of the sequence file that is to be searched by HMMSEARCH. The same string that is used here must be present in the configuration file of the HMMER daeamon. This option is required.	
DIRECTORY_PATH	Specifies the path name of a directory that contains one or more XML files. Use this option to create a single nickname over multiple XML source files. The XML wrapper uses only the files with an .xml extension that are located in the directory that you specify. The XML wrapper ignores all other files in this directory. If you specify this nickname option, do not specify a DOCUMENT column. This option is accepted only for the root nickname (the nickname that identifies the elements at the top level of the XML document).	
FILE_PATH	For Microsoft Excel: Specifies the fully qualified directory path and file name of the Excel spreadsheet that you want to access. This option is required. For table-structured files: The fully qualified path to the table-structured file to be accessed, enclosed in single quotation marks. The data file must be a standard file or a symbolic link, rather then a pipe or another non-standard file type. Either the FILE_PATH or the DOCUMENT nickname column option must be specified. If the FILE_PATH nickname option is specified, then no DOCUMENT nickname column option can be specified. For XML: Specifies the file path of the XML document. If you specify this nickname option, do not specify a DOCUMENT column. This option is accepted only for the root nickname (the nickname that identifies the elements at the top level of the XML document).	
HMMTYPE	Optional: The alphabet that is used in both models and gene sequences. The value can be either NUCLEIC or PROTEIN and is not case sensitive.	PROTEIN
INSTANCE_ PARSE_TIME	Specifies the time (in milliseconds) to parse the data in one row of the XML source document. You can modify the INSTANCE_PARSE_TIME, XPATH_EVAL_TIME, and NEXT_TIME options to optimize queries of large or complex XML source structures. This option is accepted only for columns of the root nickname (the nickname that identifies the elements at the top level of the XML document). The number that you specify can be an integer or a decimal value.	7
KEY_COLUMN	The name of the column in the file that forms the key on which the file is sorted, enclosed in single quotation marks. Use this option for sorted files only. A column that is designated with the DOCUMENT nickname column option must not be specified as the key column. Only single-column keys are supported. Multi-column keys are not allowed. The value must be the name of a column that is defined in the CREATE NICKNAME statement. The column must be sorted in ascending order. The key column must be designated not nullable by adding the NOT NULL option to its definition in the nickname statement. This option is case-sensitive. However, DB2 changes column names to uppercase unless the column is defined with double quotation marks.	If the value is not specified for a sorted nickname, the value is the name of the first column in the nicknamed file.

Table 33. Nickname options and their settings (continued)

Option	Description and valid settings	Default setting

NAMESPACES	The namespaces that are associated with the namespace prefixes that is used in the XPATH and TEMPLATE options for each column. The syntax is:
	`NAMESPACES 'prefix1= "actual_namespace1", prefix2="actual_namespace2" '`
	Separate each namespace with a comma. For example:
	`NAMESPACES ' c="http://www.myweb.com/cust", i="http://www.myweb.com/cust/id", n="http://www.myweb.com/cust/name"'`	

NEXT_TIME	Specifies the time (in milliseconds) that is required to locate subsequent source elements from the XPath expression. You can modify the NEXT_TIME, XPATH_EVAL_TIME, and INSTANCE_PARSE_TIME options to optimize queries of large or complex XML source structures. This option is accepted for root nicknames and non-root nicknames.	1
PARENT	Specified only for a child nickname whose parent was renamed through the REMOTE_OBJECT option. The PARENT option associates a child with a parent when multiple nickname families are defined within a DB2 schema. This name is case-sensitive.	
PROCESSORS	Specifies the number of processors to be used when a BLAST query is evaluated. This option corresponds to the blastall -a option.	1
RANGE	Specifies a range of cells to be used in the data source.	
REMOTE_OBJECT	For BioRS: Specifies the name of the BioRS databank that is associated with the nickname. This name determines the schema and the BioRS databank for the nickname. This name also specifies the relationship of the nickname to other nicknames. The case sensitivity of this name depends on the case sensitivity of the BioRS server and on the value of the CASE_SENSITIVE server option. You cannot use the ALTER NICKNAME statement to change or delete this name. If the name of the BioRS databank that is used in this option changes, you must delete and then create the entire nickname again. For Entrez and OMIM: Specifies the name of the object type that is associated with the nickname. This name determines the schema and NCBI database for the nickname and its relationship to other nicknames. This name is case insensitive.	
REQUIRE_ PREDICATE	Specify Y to require an equality predicate on at least one input column in all queries on the nickname hierarchy, which can limit the size of the result set. If you know that the size of the result set that returns on a query with no predicate will not exceed the JVM memory limit, you can set the value of the REQUIRE_PREDICATE nickname option to N.	
RESULTSET_SIZE	Specifies how many business objects the adapter should return to the wrapper. Specify any other value to have the adapter return the specified number of business objects. You must enable result sets in the wrapper (RESULTSET_ENABLED set to Yes) for the RESULTSET_SIZE option to work. If you specify a nonzero value for RESULTSET_SIZE, an incomplete result might be returned. Any rows that exceed the specified number are discarded, and the wrapper issues a warning message that indicates that an incomplete result set was returned to the application.	0, which specifies that all business objects that match the query should be returned.
SOAPACTION	The URI SOAPACTION attribute from the Web Service Description Language (WSDL) format. This option is required for the root nickname. This option is not allowed with nonroot nicknames. The URL can contain a colon-separated IPv6 address if it is enclosed in square brackets (for example, http://[1080:0:0:0:8:800:200C:417A])	

Table 33. Nickname options and their settings (continued)

Option	Description and valid settings	Default setting
SORTED	Specifies whether the data source file is sorted or unsorted. This option accepts either Y, y, n, or N. Sorted data sources must be sorted in ascending order according to the collation sequence for the current locale, as defined by the settings in the LC_COLLATE National Language Support category. If you specify that the data source is sorted, set the VALIDATE_DATA_FILE option to Y.	N
STREAMING	Specifies whether the XML source document should be separated into logical fragments for processing. The fragments correspond to the node that matches the XPath expression of the nickname. The wrapper then parses and processes the XML source data fragment by fragment. This type of parsing minimizes memory usage. This option is specified on only the root nickname. You can specify streaming for any XML source document (FILE, DIRECTORY, URI, or COLUMN). This option is accepted only for columns of the root nickname (the nickname that identifies the elements at the top level of the XML document). Valid values are: **Y** Streaming mode is enabled. **N** Streaming mode is disabled. Do not set the STREAMING parameter to YES if you set the VALIDATE parameter to YES. If you set both parameters to YES, you will receive an error message.	N
TEMPLATE	For WebSphere Business Integration: The nickname template fragment to use to construct an XML input document. The fragment must conform to the specified template syntax. For Web Services: The nickname template fragment to use to construct a SOAP request. The fragment must conform to the specified template syntax.	
TIMEOUT	For BLAST and HMMER: The maximum time, in minutes, that the wrapper waits for results from the daemon.	For BLAST and HMMER: 60.
URL	The URL for the Web service endpoint. This option is required for the root nickname. This option is not allowed with nonroot nicknames. Supported protocols are HTTP and HTTPS. The URL can contain a colon-separated IPv6 address if it is enclosed in square brackets (for example, http://[1080:0:0:0:8:800:200C:417A]).	
VALIDATE	Specifies whether the XML source document is validated before the XML data is extracted. If this option is set to YES, the nickname option verifies that the structure of the source document conforms to an XML schema or to a document type definition (DTD). This option is accepted only for columns of the root nickname (the nickname that identifies the elements at the top level of the XML document). The XML source document is not validated if the XML wrapper cannot locate the XML schema file or DTD file (.xsd or .dtd). DB2 does not issue an error message if the validation does not occur. Therefore, ensure that the XML schema file or DTD file exists in the location that is specified in the XML source document. Do not set the VALIDATE parameter to YES if you set the STREAMING parameter to YES. If you set both parameters to YES, you will receive an error message.	NO

Table 33. Nickname options and their settings (continued)

Option	Description and valid settings	Default setting
VALIDATE_ DATA_FILE	For sorted files, this option specifies whether the wrapper verifies that the key column is sorted in ascending order and checks for null keys. The only valid values for this option are Y or N. The check is done once at registration time. This option is not allowed if the DOCUMENT nickname column option is used for the file path.	N
XPATH	Specifies the XPATH expression that identifies the elements that represent the individual tuples. The XPATH nickname option for a child nickname is evaluated in the context of the path that is specified by the XPATH nickname option of its parent. This XPATH expression is used as a context for evaluating column values that are identified by the XPATH nickname column options.	
XPATH_EVAL_ TIME	Specifies the time (in milliseconds) to evaluate the XPath expression of the nickname and to locate the first element. You can modify the XPATH_EVAL_TIME, INSTANCE_PARSE_TIME, and NEXT_TIME options to optimize queries of large or complex XML source structures. This option is accepted for root nicknames and nonroot nicknames. The number that you specify can be an integer or a decimal value.	1

Server options for federated systems

Server options are used to describe a data source server.

Server options specify data integrity, location, security, and performance information. Some server options are available for all data sources, and other server options are data source specific.

The common federated server options for relational data sources are:

- Compatibility options. COLLATING_SEQUENCE, IGNORE_UDT
- Data integrity options. IUD_APP_SVPT_ENFORCE
- Data and time options. DATEFORMAT, TIMEFORMAT, TIMESTAMPFORMAT
- Location options. CONNECTSTRING, DBNAME, IFILE
- Security options. FOLD_ID, FOLD_PW, INFORMIX_LOCK_MODE
- Performance options. COMM_RATE, CPU_RATIO, DB2_MAXIMAL_PUSHDOWN, IO_RATIO, LOGIN_TIMEOUT, PACKET_SIZE, PLAN_HINTS, PUSHDOWN, TIMEOUT, VARCHAR_NO_TRAILING_BLANKS

The following table lists the server definition server options applicable for each relational data source.

Table 34. Server options for relational data sources

Data Source	CODEPAGE	COLLATING_SEQUENCE	COMM_RATE	CONNECTSTRING	CONV_EMPTY_STRING	CPU_RATIO	DATEFORMAT	DB2_MAXIMAL_PUSHDOWN	DB2_PRESERVE_CUR_ON_CONNECTION	DB2_UM_PLUGIN	DB2_TWO_PHASE_COMMIT	DBNAME	FOLD_ID	FOLD_PW	IFILE	INFORMIX_CLIENT_LOCALE	INFORMIX_DB_MODE	INFORMIX_LOCK_MODE	IO_RATIO	IUD_APP_SVPT_ENFORCE	LOGIN_TIMEOUT	NODE	PACKET_SIZE	PASSWORD	PLAN_HINTS	PUSHDOWN	TIMEOUT	TIMEFORMAT	TIMESTAMPFORMAT	VARCHAR_NO_TRAILING_BLANKS
DB2 UDB for iSeries		X	X			X		X		X	X	X	X	X					X	X				X		X				X
DB2 UDB for z/OS and OS/390		X	X			X		X		X	X	X	X	X					X	X				X		X				X
DB2 for VM and VSE		X	X			X		X		X		X	X	X					X	X				X		X				X
DB2 for Linux, UNIX, and Windows		X	X			X		X		X	X	X	X	X					X	X				X		X				X
Informix		X	X			X		X		X	X	X	X	X	X	X		X	X	X		X		X		X				
Microsoft SQL Server	X	X	X			X		X	X	X	X	X	X	X					X	X		X		X		X				
ODBC	X	X	X			X	X	X		X		X	X	X					X	X		X		X		X		X	X	X
OLE DB		X		X						X																				
Oracle		X	X			X		X		X	X		X	X					X			X		X	X	X				X
Sybase		X	X		X	X		X		X	X	X	X	X	X				X		X	X	X	X	X	X	X			
Teradata		X	X			X		X		X									X	X		X				X				

The following table lists the server definition server options applicable for each nonrelational data source, except WebSphere Business Integration. The server definition server options for WebSphere Business Integration are listed in Table 36 on page 307.

Table 35. Server options for nonrelational data sources.

Data Source	DAEMON_PORT	DB2_UM_PLUGIN	MAX_ROWS	NODE	PROXY_AUTHID	PROXY_PASSWORD	PROXY_SERVER_NAME	PROXY_SERVER_PORT	PROXY_TYPE	SSL_CLIET_CERTIFICATE_LABEL	SSL_KEYSTORE_FILE	SSL_KEYSTORE_PASSWORD	SSL_VERIFY_SERVER_CERTIFICATE	SOCKET_TIMEOUT	TIMEOUT	USE_CLOB_SEQUENCE
BioRS		X		X					X						X	
BLAST	X			X	X	X	X	X	X	X	X	X	X			X
Entrez			X		X	X	X	X	X					X		
Excel																
HMMER	X			X	X	X	X	X	X	X	X	X	X			X
SCRIPT	X			X	X	X	X	X	X	X	X	X	X			
Table-structured files																
Web services		X								X	X	X	X		X	
XML					X	X	X	X	X	X	X	X	X	X		

The following table lists the server definition server options applicable for WebSphere Business Integration data sources.

Table 36. Server options for WebSphere Business Integration data sources.

Data Source	APP_TYPE	FAULT_QUEUE	MQ_CONN_NAME	MQ_MANAGER	MQ_RESPONSE_TIMEOUT	MQ_SVRCONN_CHANNELNAME	REQUEST_QUEUE	RESPONSE_QUEUE
WebSphere Business Integration	X	X	X	X	X	X	X	X

The following table describes each server option and lists the valid and default settings.

Table 37. Server options and their settings

Option	Description and valid settings	Default setting
APP_TYPE	The type of remote application. Valid values are 'PSOFT', 'SAP', and 'SIEBEL'. This option is required.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
CASE_SENSITIVE	Specifies whether the BioRS server treats names in a case sensitive manner. Valid values are Y or N. **'Y'** The BioRS server treats names in a case sensitive manner. **'N'** The BioRS server does not treat names in a case sensitive manner In the BioRS product, a configuration parameter controls the case sensitivity of the data that is stored on the BioRS server. The CASE_SENSITIVE option is the federated server counterpart to that BioRS system configuration parameter. You must synchronize the BioRS server case sensitivity configuration settings in your BioRS system and in the federated server. If you do not keep the case sensitivity configuration settings synchronized between BioRS and the federated server, errors will occur when you attempt to access BioRS data through federated server. You cannot change or delete the CASE_SENSITIVE option after you create a new BioRS server in federated server. If you need to change the CASE_SENSITIVE option, you must drop and then create the entire server again. If you drop the BioRS server, you must also create all of the corresponding BioRS nicknames again. Federated server automatically drops all nicknames that correspond to a dropped server.	Y
CODEPAGE	Specifies the DB2 code page identifier corresponding to the coded character set of the data source client configuration. You must specify the client's code page if the client's code page and the federated database code page do not match. For data sources that support Unicode, the CODEPAGE option can be set to the DB2 code page identifier corresponding to the supported Unicode encoding of the data source client.	On UNIX or Windows systems with a non-Unicode federated database: The federated database code page. On UNIX systems with a Unicode federated database: 1208 On Windows systems with a Unicode federated database: 1202

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
COLLATING_ SEQUENCE	Specifies whether the data source uses the same default collating sequence as the federated database, based on the NLS code set and the country/region information. **'Y'** The data source has the same collating sequence as the DB2 federated database. **'N'** The data source has a different collating sequence than the DB2 federated database collating sequence. **'I'** The data source has a different collating sequence than the DB2 federated database collating sequence, and the data source collating sequence is insensitive to case (for example, 'STEWART' and 'StewART' are considered equal).	'N'
COMM_RATE	Specifies the communication rate between the federated server and the data source server. Expressed in megabytes per second. Valid values are greater than 0 and less than $1x10^{23}$. Values can be expressed in any valid REAL notation.	'2'
CONNECTSTRING	Specifies initialization properties needed to connect to an OLE DB provider.	None.
CONNECTSTRING	Specifies initialization properties needed to connect to an OLE DB provider.	None.
CONV_EMPTY_STRING	Use for Sybase wrapper that works with replication tasks. When you set the CONV_EMPTY_STRING option into Y, the Sybase wrapper converts an empty string into a space. Set this option to Y when a source data server has a non-nullable character column that stores an empty string and the target data server is Sybase.	N
CPU_RATIO	Indicates how much faster or slower a data source CPU runs than the federated server CPU. Valid values are greater than 0 and less than $1x10^{23}$. Values can be expressed in any valid REAL notation. A setting of 1 indicates that the DB2 federated CPU speed and the data source CPU speed have the same CPU speed, a 1:1 ratio. A setting of 0.5 indicates that the DB2 federated CPU speed is 50% slower than the data source CPU speed. A setting of 2 indicates that the DB2 federated CPU speed is twice as fast as the data source CPU speed.	'1.0'
DATEFORMAT	The date format used by the data source. Enter the format using 'DD', 'MM', and 'YY' or 'YYYY' to represent the numeric form of the date. You should also specify the delimiter such as a space or comma. For example, to represent the date format for '2003-01-01', use 'YYYY-MM-DD'. This field is nullable.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
DAEMON_PORT	Specifies the port number on which the daemon will listen for BLAST or HMMER job requests. The port number must be the same number specified in the DAEMON_PORT option of the daemon configuration file.	BLAST: 4007 HMMER: 4098
DB2_MAXIMAL_ PUSHDOWN	Specifies the primary criteria that the query optimizer uses when choosing an access plan. The query optimizer can choose access plans based on cost or based on the user requirement that as much query processing as possible be performed by the remote data sources. **'Y'** The query optimizer chooses an access plan that pushes down more query operations to the data source than other plans. When several access plans provide the same amount of pushdown, the query optimizer then chooses the plan with the lowest cost. If a materialized query table (MQT) on the federated server can process part or all of the query, then an access plan that includes the materialized query table might be used. The federated database does not push down queries that result in a Cartesian product. **'N'** The query optimizer chooses an access plan based on cost.	'N'
DB2_PRESERVE_CUR_ON_ CONNECTION	Specifies the behavior of cursors for committed or rolled back transactions. If value is set to Y, cursors can remain open on Microsoft SQL Server even if COMMIT or ROLLBACK is sent. If this option is not set or the value is set to N, cursors are closed if COMMIT or ROLLBACK is sent. This option is optional.	None
DB2_TWO_PHASE_COMMIT	Specify to allow or disallow federated two-phase commit for each data source using the CREATE SERVER or ALTER SERVER statements. Set to Y to configure and activate two-phase commit at the database level. Set to N to configure but not activate federated two-phase commit at the database level. Applications can activate or deactivate this option with the SET SERVER OPTION statement.	
DB2_UM_PLUGIN	If you use a plugin for retrieving user mappings from an external repository, specify the class name including package (for example, 'com.ibm.ii.um.ldap.UserMappingRepository'). If you are using the Lightweight Directory Access Protocol (LDAP) sample plugin, you can use just the class name (for example, UserMappingRepository).	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
DBNAME	Name of the data source database that you want the federated server to access. For DB2 database, this value corresponds to a specific database for the initial remote DB2 database connection. This specific database is the database alias for the remote DB2 database that is cataloged at the federated server using the CATALOG DATABASE command or the DB2 Configuration Assistant. Does not apply to Oracle data sources because Oracle instances contain only one database. Does not apply to Teradata.	None.
FAULT_QUEUE	The name of the fault queue that delivers error messages from the adapter to the wrapper. The name must conform to the specifications for queue names for WebSphere MQ. This is a required option.	None.
FOLD_ID (See notes 1 and 4 at the end of this table.)	Applies to user IDs that the federated server sends to the data source server for authentication. Valid values are: **'U'** The federated server folds the user ID to uppercase before sending it to the data source. This is a logical choice for DB2 family and Oracle data sources (See note 2 at end of this table.) **'N'** The federated server does nothing to the user ID before sending it to the data source. (See note 2 at end of this table.) **'L'** The federated server folds the user ID to lowercase before sending it to the data source. If this option is not specified, the federated server tries to send the user ID to the data source in uppercase (unchanged) and in lowercase.	None.
FOLD_PW (See notes 1, 3 and 4 at the end of this table.)	Applies to passwords that the federated server sends to data sources for authentication. Valid values are: **'U'** The federated server folds the password to uppercase before sending it to the data source. This is a logical choice for DB2 family and Oracle data sources. **'N'** The federated server does nothing to the password before sending it to the data source. **'L'** The federated server folds the password to lowercase before sending it to the data source. If this option is not specified, the federated server tries to send the user ID to the data source in uppercase (unchanged) and in lowercase.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
HMMPFAM_OPTIONS	Specifies hmmpfam options such as --null2, --pvm, and --xnu that have no corresponding column name in a reference table that maps options to column names. For example: `HMMPFAM_OPTIONS '--xnu --pvm'` In this example, the daemon runs the HMMPFAM program with options from the WHERE clause of the query, plus the additional options --xnu --pvm.	
HMMSEARCH_ OPTIONS	Allows the user to provide additional command line options to the hmmsearch command. Only valid with type SEARCH. See the HMMER User's Guide for more information.	None.
IFILE	Use to specify the path and name of the Sybase Open Client interfaces file if you do not want to use the default interface file. On Windows NT federated servers, the default is %SYBASE%\ini\sql.ini. On UNIX federated servers, the default is $SYBASE%/interfaces.	None.
INFORMIX_CLIENT_LOCALE	Specifies the CLIENT_LOCALE to use for the connection between the federated server and the data source server. If the INFORMIX_CLIENT_LOCALE option is not specified, the Informix CLIENT_LOCALE environment variable is set to the value specified in the db2dj.ini file (if any). If db2dj.ini does not specify CLIENT_LOCALE, the Informix CLIENT_LOCALE environment variable is set to the Informix locale that most closely matches the code page and territory of the federated database. Any valid Informix locale is a valid value. This option is optional.	None.
INFORMIX_DB_LOCALE	Specifies the DB_LOCALE to use for the connection between the federated server and the data source server. If the INFORMIX_DB_LOCALE option is not specified, the Informix DB_LOCALE environment variable is set to the value specified in the db2dj.ini file (if any). If db2dj.ini does not specify DB_LOCALE, the Informix DB_LOCALE environment variable is not set. Any valid Informix locale is a valid value. This option is optional.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
INFORMIX_LOCK_ MODE	Specifies the lock mode to be set for an Informix data source. The Informix wrapper issues the 'SET LOCK MODE' command immediately after establishing the connection to an Informix data source. Valid values are: **'W'** Sets the Informix lock mode to WAIT. If the wrapper tries to access a locked table or row, Informix waits until the lock is released. **'N'** Sets the Informix lock mode to NOWAIT. If the wrapper tries to access a locked table or row, Informix returns an error. **'n'** Sets the Informix lock mode to WAIT *n* seconds. If the wrapper tries to access a locked table or row and the lock is not released within the specified number of seconds, Informix returns an error. If a deadlock or timeout error occurs when a federated server attempts to connect to an Informix data source, changing the lock mode setting on the federated server can often resolve the error. Use the ALTER SERVER statement to change the lock mode setting on the federated server. For example: `ALTER SERVER TYPE informix VERSION 9 WRAPPER informix OPTIONS (ADD informix_lock_mode '60')`	'W'
IO_RATIO	Denotes how much faster or slower a data source I/O system runs than the federated server I/O system. Valid values are greater than 0 and less than $1x10^{23}$. Values can be expressed in any valid REAL notation. A setting of 1 indicates that the DB2 federated I/O speed and the data source I/O speed have the same I/O speed, a 1:1 ratio. A setting of .5 indicates that the DB2 federated I/O speed is 50% slower than the data source I/O speed. A setting of 2 indicates that the DB2 federated I/O speed is twice as fast as the data source I/O speed.	'1.0'

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
IUD_APP_SVPT_ ENFORCE	Specifies whether the DB2 federated system should enforce detecting or building of application savepoint statements. When set using the SET SERVER OPTION statement, this server option will have no effect with static SQL statements. **'Y'** The federated server rolls back insert, update, or delete transactions if an error occurs in an insert, update, or delete operation and the data source does not enforce application savepoint statements. SQL error code SQL1476N is returned. **'N'** The federated server will not roll back transactions when an error is encountered. Your application must handle the error recovery.	'Y'
LOGIN_TIMEOUT	Specifies the number of seconds for the DB2 federated server to wait for a response from Sybase Open Client to the login request. If you specify 0, the federated server will wait indefinitely for a response.	'0'
MAX_ROWS	For Entrez: Specifies the number of rows that the federated server returns for a query. For OMIM: Limits to the number of records for the root nickname that a query can return. For example, if the MAX_ROWS server option is set to 25, a maximum of 25 records for the root nickname and all of the records for the child-related nicknames are returned. You can specify only positive numbers and zero. When you set the option to be zero, you enable queries to retrieve an unlimited number of rows from the NCBI Web site. However, setting the MAX_ROWS server option to zero or to a very high number can impact your query performance. The MAX_ROWS server option is not required.	Microsoft Windows operating systems: 2000 rows. UNIX-based operating systems: 5000 rows.
MQ_CONN_NAME	The hostname or network address of the computer where the Websphere MQ server is running. An example of a connection name is: 9.30.76.151(1420) where 1420 is the port number. If the port number is excluded a default value of 1414 will be used. This option is optional. If it is omitted, the MQSERVER environment variable (if sepcified in db2dj.ini file) is used to select the channel definition. If MQSERVER is not set, the client channel table is used.	The wrapper uses the MQSERVER environment variable, if specified in the db2dj.ini file, to select the channel definition. If the MQSERVER environment variable is not set, the wrapper uses the client channel table.
MQ_MANAGER	The name of the WebSphere MQ manager. Any valid WebSphere MQ manager name. This option is required.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
MQ_RESPONSE_ TIMEOUT	The amount of time that the wrapper should wait for a response message from the response queue. The value is in milliseconds. You can specify a special value of -1 to indicate that there is no timeout period. This option is optional.	10000
MQ_SVRCONN_ CHANNELNAME	The name of the server-connection channel on the Webspehere MQ Manager that the wrapper should try to connect to. This parameter can be specified only if the MQ_CONN_NAME server option is specified. The default server-connection channel, SYSTEM.DEF.SVRCONN, is used if this option is omitted.	SYSTEM.DEF.SVRCONN
NODE	Relational data sources: Name by which a data source is defined as an instance to its RDBMS. BLAST: Specifies the host name or IP address of the system on which the BLAST daemon process is running. The IP address can be an IPv4 address (for example, 192.168.1.1) or it can be a colon-separated IPv6 address (for example, 1080:0:0:0:8:800:200C:417A). This option is required. HMMER: Specifies the host name or IP address of the server on which the HMMER daemon process runs. The IP address can be an IPv4 address or it can be a colon-separated IPv6 address. This option is required. BioRS: Specifies the host name of the system on which the BioRS query tool is available. The IP address can be an IPv4 address or it can be a colon-separated IPv6 address. This option is optional.	BioRS: localhost
PACKET_SIZE	Specifies the byte size of the packet that the Client-Library uses when sending special packets. If the Sybase wrapper needs to send or receive large amounts of text or image data, a larger packet size might improve efficiency.	
PASSWORD	Specifies whether passwords are sent to a data source. **'Y'** Passwords are sent to the data source and validated. **'N'** Passwords are not sent to the data source and not validated.	'Y'

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
PLAN_HINTS	Specifies whether plan hints are to be enabled. Plan hints are statement fragments that provide extra information for data source optimizers. This information can, for certain query types, improve query performance. The plan hints can help the data source optimizer decide whether to use an index, which index to use, or which table join sequence to use. **'Y'** Plan hints are to be enabled at the data source if the data source supports plan hints. **'N'** Plan hints are not to be enabled at the data source. This option is only available for Oracle and Sybase data sources.	'N'
PORT	Specifies the number of the port the wrapper uses to connect to the BioRS server. This option is optional.	'5014'
PROCESSORS	Specifies the number of processors that the HMMER program uses. This option is equivalent to the --cpu option of the `hmmpfam` command.	None.
PROXY_AUTHID	Specifies the user name to use when the proxy server requires authentication. Contact your network administrator for the user name.	None.
PROXY_PASSWORD	Specifies the password to use when the proxy server requires authentication. Contact your network administrator for the password.	None.
PROXY_SERVER_ NAME	Specifies the proxy server name or the IP address. This field is required if the value of PROXY_TYPE is 'HTTP' or 'SOCKS'. Contact your network administrator for the server name or IP address of the proxy server. The IP address can be an IPv4 address (for example, 192.168.1.1) or it can be a colon-separated IPv6 address (for example, 1080:0:0:0:8:800:200C:417A).	None.
PROXY_SERVER_ PORT	Specifies the proxy server port number. This field is required if the value of PROXY_TYPE is 'HTTP' or 'SOCKS'. Contact your network administrator for the port number of the proxy server.	None.
PROXY_TYPE	Specifies the type of proxy type that is used to access the Internet when behind a firewall. The valid values are 'NONE', 'HTTP', or 'SOCKS'. The default value is 'NONE'. Contact your network administrator for the type of proxy that is used. BLAST, HMMER, and SCRIPT do not support HTTP proxies.	'NONE'

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
PUSHDOWN	**'Y'** DB2 will consider letting the data source evaluate operations. **'N'** DB2 will send the data source SQL statements that include only SELECT with column names. Predicates (such as WHERE=), column and scalar functions (such as MAX and MIN), sorts (such as ORDER BY or GROUP BY), and joins will not be included in any SQL sent to the data source. Default for the ODBC wrapper.	'Y'
RESPONSE_QUEUE	The name of the response queue that delivers query results from the adapter to the wrapper. The name must conform to the specifications for queue names for WebSphere MQ. This option is required.	None.
REQUEST_QUEUE	The name of the request queue that delivers query requests from the wrapper to the adapter. The name must conform to the specifications for queue names for WebSphere MQ. This option is required.	None.
SOCKET_TIMEOUT	Specifies the maximum time in minutes that the DB2 federated server will wait for results from the proxy server. A valid value is any number that is greater than or equal to zero. The default is zero '0'. A value of zero denotes an unlimited amount of time to wait.	0
SSL_CLIENT_CERTIFICATE_LABEL	Specifies the client certificate that is sent during SSL authentication. If the value is not specified, the current DB2 authorization ID will be sent.	None.
SSL_KEYSTORE_FILE	Specifies the name of the certificate storage file to use for SSL/TLS communications. The value that you specify must be a full path that is accessible by the DB2 agent or FMP process. Optional: You can specify the value 'GSK_MS_CERTIFICATE_STORE' to use the native Microsoft certificate storage.	None.
SSL_KEYSTORE_PASWORD	Specifies the password that is used to access the SSL certificate storage. This password is encrypted when it is stored in the DB2 catalog.	None.
SSL_VERIFY_SERVER_CERTIFICATE	Specifies if the server certificate should be verified during SSL authenticate. The values are not case sensitive. To authenticate, use one of the following values: 'Y', 'YES', 'T', or 'TRUE'. To disable, use 'F', 'FALSE', 'N', or 'NO'.	'NO'
TIMEFORMAT	The time format used by the data source. Enter the format using 'hh12', 'hh24', 'mm', 'ss', 'AM', or 'A.M'. For example, to represent the time format of '16:00:00', use 'hh24:mm:ss'. To represent the time format of '8:00:00 AM', use 'hh12:mm:ss AM'. This field is nullable.	None.

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
TIMESTAMPFORMAT	The timestamp format used by the data source. The format follows that for date and time, plus 'n' for tenth of a second, 'nn' for hundredth of a second, 'nnn' for milliseconds, and so on, up to 'nnnnnn' for microseconds. For example, to represent the timestamp format of '2003-01-01-24:00:00.000000', use 'YYYY-MM-DD-hh24:mm:ss.nnnnnn'. This field is nullable.	None.
TIMEOUT	The timeout value that you specify depends on which wrapper that you are using. If you specify 0, DB2 will wait indefinitely for a response. Sybase: Specifies the number of seconds that the DB2 federated server will wait for a response from the Sybase server for any SQL statement. The value of *seconds* is a positive whole number. BioRS: Specifies the time, in minutes, that the BioRS wrapper should wait for a response from the BioRS server. The default value is 10. This option is optional. Web services: Specifies the time, in minutes, that DB2 should wait for a network transfer and the computation of a result.	Sybase: 0 BioRS: 10
USE_CLOB_ SEQUENCE	This option specifies the data type the federated server uses for the BlastSeq or HmmQSeq column. The values can be 'Y' or 'N'. You can use the CREATE NICKNAME or ALTER NICKNAME statement. to override the default data type for the BlastSeq or HmmQSeq column. • If you specify a value of N, the data type is VARCHAR(32000). • If you specify a value of Y, the data type is CLOB(5M). The default value is N, not Y.	'Y'

Table 37. Server options and their settings (continued)

Option	Description and valid settings	Default setting
VARCHAR_NO_ TRAILING_BLANKS	This option applies to data sources which have variable character data types that do not pad the length with trailing blanks during comparison. Some data sources, such as Oracle, do not have blank-padded character comparison semantics that return the same results as the DB2 for Linux, UNIX, and Windows comparison semantics. Set this option when you want it to apply to all the VARCHAR and VARCHAR2 columns in the data source objects that will be accessed from the designated server. This includes views. **Y** Trailing blanks are absent from these VARCHAR columns, or the data source has blank-padded character comparison semantics that are similar to the semantics on the federated server. The federated server pushes down character comparison operations to the data source for processing. **N** Trailing blanks are present in these VARCHAR columns and the data source has blank-padded character comparison semantics that are different than the federated server. The federated server processes character comparison operations if it is not possible to compensate for equivalent semantics. For example, rewriting the predicate.	N for affected data sources.

Notes on this table:

1. This field is applied regardless of the value specified for authentication.
2. Because DB2 stores user IDs in uppercase, the values 'N' and 'U' are logically equivalent to each other.
3. The setting for FOLD_PW has no effect when the setting for password is 'N'. Because no password is sent, case cannot be a factor.
4. Avoid null settings for either of these options. A null setting can seem attractive because DB2 will make multiple attempts to resolve user IDs and passwords; however, performance might suffer (it is possible that DB2 will send a user ID and password up to nine times before successfully passing data source authentication).

Valid server types in SQL statements

Server types indicate the kind of data source that the server definition represents.

Server types vary by vendor, purpose, and operating system. Supported values depend on the wrapper being used.

For most data sources, you must specify a valid server type in the CREATE SERVER statement.

BioRS wrapper

A server type specification is optional for BioRS data sources.

Server Type	Data Source
Not required in the CREATE SERVER statement.	BioRS

BLAST wrapper

A server type specification is required for each type of BLAST search that you want to run for BLAST data sources supported by the BLAST daemon.

Server Type	Data Source
BLASTN	BLAST searches in which a nucleotide sequence is compared with the contents of a nucleotide sequence database to find sequences with regions homologous to regions of the original sequence.
BLASTP	BLAST searches in which an amino acid sequence is compared with the contents of an amino acid sequence database to find sequences with regions homologous to regions of the original sequence.
BLASTX	BLAST searches in which a nucleotide sequence is compared with the contents of an amino acid sequence database to find sequences with regions homologous to regions of the original sequence.
TBLASTN	BLAST searches in which an amino acid sequence is compared with the contents of a nucleotide sequence database to find sequences with regions homologous to regions of the original sequence.
TBLASTX	BLAST searches in which a nucleotide sequence is compared with the contents of a nucleotide sequence database to find sequences with regions homologous to regions of the original sequence.

CTLIB wrapper

The CTLIB wrapper supports Sysbase data sources. A server type specification is required for Sybase data sources supported by the CTLIB client software.

Server Type	Data Source
SYBASE	Sybase

DRDA wrapper

The DRDA wrapper is used for DB2 family data sources. A server type specification is required for the DB2 family data sources.

Table 38. DB2 family data sources

Server Type	Data Source
DB2/UDB	IBM DB2 Version 9.1 for Linux, UNIX, and Windows
DB2/ISERIES	IBM DB2 UDB for iSeries and AS/400
DB2/ZOS	IBM DB2 UDB for z/OS
DB2/VM	IBM DB2 for VM

Entrez wrapper

A server type specification is required for Entrez data sources.

Server Type	Data Source
NUCLEOTIDE	Entrez
OMIM	Entrez
PUBMED	Entrez

Excel wrapper

A server type specification is not required for Excel data sources.

Server Type	Data Source
Not required in the CREATE SERVER statement.	Microsoft Excel

HMMER wrapper

A server type specification is required for each server that you want to run a HMMER search on for HMMER data sources supported by the HMMER daemon.

Server Type	Data Source
PFAM	HMMER
SEARCH	HMMER

Informix wrapper

A server type specification is required for Informix data sources supported by Informix Client SDK software.

Server Type	Data Source
INFORMIX	Informix

MSSQLODBC3 wrapper

A server type specification is required for Microsoft SQL Server data sources supported by the DataDirect Connect ODBC 4.2 (or later) driver or the Microsoft SQL Server ODBC 3.0 (or later) driver.

Server Type	Data Source
MSSQLSERVER	Microsoft SQL Server

NET8 wrapper

A server type specification is required for Oracle data sources supported by Oracle NET8 client software.

Server Type	Data Source
ORACLE	Oracle Version 8.0. or later

ODBC wrapper

A server type specification is required for ODBC data sources supported by the ODBC 3.x driver.

Server Type	Data Source
ODBC	ODBC

OLE DB wrapper

A server type definition is not required for OLE DB providers compliant with Microsoft OLE DB 2.0 or later.

Server Type	Data Source
Not required in the CREATE SERVER statement.	Any OLE DB provider

Table-structured files wrapper

A server type definition is not required for table-structured file data sources.

Server Type	Data Source
Not required in the CREATE SERVER statement.	Table-structured files

Teradata wrapper

A server type definition is required for Teradata data sources supported by the Teradata client software.

Server Type	Data Source
TERADATA	Teradata

Web services wrapper

A server type definition is not required for Web services data sources.

Server Type	Data Source
Not required in the CREATE SERVER statement.	Any Web services data source.

WebSphere Business Integration wrapper

A server type definition is required for business application data sources supported by the WebSphere Business Integration wrapper.

Server Type	Data Source
WBI	WebSphere Business Integration 2.2 or 2.3

XML wrapper

A server type definition is not required for XML data sources.

Server Type	Data Source
Not required in the CREATE SERVER statement.	XML

User mapping options for federated systems

These options are used with the CREATE USER MAPPING and ALTER USER MAPPING statements.

Table 39. User mapping options and their settings

Option	Valid settings	Default setting
ACCOUNTING	DRDA: Used to specify a DRDA accounting string. Valid settings include any string of length 255 or less. This option is required only if accounting information needs to be passed. See the DB2 Connect Users Guide for more information.	None
GUEST	Specifies if the wrapper is to use the guest access mode to the BioRS server. **Y** The wrapper uses the guest access mode to the BioRS server. **N** The wrapper does not use the guest access mode to the BioRS server. When set to a value of Y, this option is mutually exclusive with the REMOTE_AUTHID option and the REMOTE_PASSWORD option. Valid for the BioRS data source.	N
REMOTE_AUTHID	Indicates the authorization ID used at the data source. Valid settings include any string of length 255 or less. Valid for the BioRS and Web services data sources.	The authorization ID you use to connect to DB2.
PROXY_AUTHID	Specifies the password to use when the proxy server requires authentication. Contact your network administrator for the password. Valid for BioRS, Blast, Entrez, HMMER, Script, and Web services data sources.	
PROXY_PASSWORD	Specifies the user name to use when the proxy server requires authentication. Contact your network administrator for the user name. Valid for BioRS, Blast, Entrez, HMMER, Script, and Web services data sources.	

Table 39. User mapping options and their settings (continued)

Option	Valid settings	Default setting
REMOTE_PASSWORD	Indicates the authorization password used at the data source. Valid settings include any string of length 32 or less. If your server requires a password and you do not set this option, you must ensure that the following conditions are met or the connection will fail: • The database manager configuration parameter AUTHENTICATON is set to SERVER. • The server option PASSWORD is omitted or set to Y (the default). • When you connected to the DB2 database, you specified an authorization ID and password. The password that you specified must be the same as the password of your remote server. Valid for the BioRS and Web services data sources.	The password you use to connect to the DB2 if both conditions listed in the valid settings column are met.
SSL_CLIENT_ CERTIFICATE_LABEL	Specifies the client certificate that is sent during SSL authentication. If the value is not specified, the current DB2 authorization ID will be sent. Valid for the Web services data source.	None.

Wrapper options for federated systems

Wrapper options are used to configure the wrapper or to define how the federated server uses the wrapper. Wrapper options can be set when you create or alter the wrapper.

All relational and nonrelational data sources use the DB2_FENCED wrapper option. The ODBC and Teradata wrappers support the DB2_SOURCE_CLIENT_MODE wrapper option. The Entrez data source uses the EMAIL wrapper option. The ODBC data source uses the MODULE wrapper option. BioRS, BLAST, Entrez, HMMER, web services, and XML data sources can use the wrapper options for proxies. The SSL options are supported by the BLAST, HMMER, SCRIPT, web services, and XML wrappers.

Table 40. Wrapper options and their settings

Option	Valid settings		Default setting
DB2_FENCED	Specifies whether the wrapper runs in fenced or trusted mode.		Relational wrappers: N. Nonrelational wrappers from IBM: N. Nonrelational wrappers from third parties: Y.
	Y	The wrapper runs in fenced mode.	
	N	The wrapper runs in trusted mode.	

Table 40. Wrapper options and their settings (continued)

Option	Valid settings	Default setting
DB2_SOURCE_ CLIENT_MODE	Specifies that the data source client is 32-bit and that the database instance on the federated server is 64-bit. When you specify this option, you must also set the DB2_FENCED wrapper option to Y. This option applies only to ODBC and Teradata data sources, and is currently supported only on AIX or Solaris operating systems. The only valid value is 32BIT. The value is not case sensitive. **32BIT** The data source client that is installed on the federated server is 32-bit.	None.
DB2_UM_PLUGIN	If you use a plugin for retrieving user mappings from an external repository, specify the class name including package (for example, 'com.ibm.ii.um.ldap.UserMappingRepository'). If you are using the Lightweight Directory Access Protocol (LDAP) sample plugin, you can use the class name (for example, UserMappingRepository).	None.
EMAIL	Specifies an e-mail address when you register the Entrez wrapper. This e-mail address is included with all queries and allows NCBI to contact you if there are problems, such as too many queries overloading the NCBI servers. This option is required.	
MODULE	Specifies the full path of the library that contains the ODBC Driver Manager implementation or the SQL/CLI implementation. Required for the ODBC wrapper on UNIX federated servers.	On Windows, the default value is odbc32.dll
PROXY_SERVER_ NAME	Specifies the proxy server name or the IP address. This field is required if the value of PROXY_TYPE is 'HTTP' or 'SOCKS'. The IP address can be an IPv4 address (for example, 192.168.1.1) or it can be a colon-separated IPv6 address (for example, 1080:0:0:0:8:800:200C:417A). Contact your network administrator for the server name or IP address of the proxy server.	None.
PROXY_SERVER_ PORT	Specifies the proxy server port number. This field is required if the value of PROXY_TYPE is 'HTTP' or 'SOCKS'. Contact your network administrator for the port number of the proxy server. Specifying the PROXY_SERVER_PORT option in a CREATE SERVER statement overrides the PROXY_SERVER_PORT option in a CREATE WRAPPER statement overrides the the	None.

Table 40. Wrapper options and their settings (continued)

Option	Valid settings	Default setting
PROXY_TYPE	Specifies the type of proxy type that is used to access the Internet when behind a firewall. The valid values are 'NONE', 'HTTP', or 'SOCKS'. The default value is 'NONE'. Contact your network administrator for the type of proxy that is used. The BLAST, HMMER, and SCRIPT wrappers do not support HTTP proxies.	'NONE'
SSL_KEYSTORE_ FILE	Specifies the name of the certificate storage file to use for SSL/TLS communications. The value that you specify must be a full path that is accessible by the DB2 agent or FMP process. Optional: You can specify the value 'GSK_MS_CERTIFICATE_STORE' to use the native Microsoft certificate storage.	None.
SSL_KEYSTORE_ PASWORD	Specifies the password that is used to access the SSL certificate storage. This password is encrypted when it is stored in the DB2 catalog.	None.
SSL_VERIFY_ SERVER_ CERTIFICATE	Specifies if the server certificate should be verified during SSL authenticate. The values are not case sensitive. To authenticate, use one of the following values: 'Y', 'YES', 'T', or 'TRUE'. To disable, use 'F', 'FALSE', 'N', or 'NO'.	'NO'

Chapter 30. Federated system and data source mappings

Default forward data type mappings

The two kinds of mappings between data source data types and federated database data types are forward type mappings and reverse type mappings. In a forward type mapping, the mapping is from a remote type to a comparable local type.

You can override a default type mapping, or create a new type mapping with the CREATE TYPE MAPPING statement.

These mappings are valid with all the supported versions, unless otherwise noted.

For all default forward data types mapping from a data source to the federated database, the federated schema is SYSIBM.

The following tables show the default forward mappings between federated database data types and data source data types.

DB2 Database for Linux, UNIX, and Windows data sources

The following table lists the forward default data type mappings for DB2 Database for Linux, UNIX, and Windows data sources.

Table 41. DB2 Database for Linux, UNIX, and Windows forward default data type mappings (Not all columns shown)

REMOTE_ TYPE NAME	REMOTE_ LOWER _LEN	REMOTE_ UPPER_ LEN	REMOTE_ LOWER_ SCALE	REMOTE_ UPPER_ SCALE	REMOTE_ BIT_ DATA	REMOTE_ DATA_ OPERATORS	FEDERATED_ TYPE NAME	FEDERATED_ LENGTH	FEDERATED_ SCALE	FEDERATED_ BIT_ DATA
BIGINT	-	-	-	-	-	-	BIGINT	-	0	-
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHAR	-	-	-	-	-	-	CHAR	-	0	N
CHAR	-	-	-	-	Y	-	CHAR	-	0	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	-	-	-	-	-	DATE	-	0	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
DOUBLE	-	-	-	-	-	-	DOUBLE	-	-	-
FLOAT	-	-	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	0	N
INTEGER	-	-	-	-	-	-	INTEGER	-	0	-
LONGVAR	-	-	-	-	N	-	CLOB	-	-	-
LONGVAR	-	-	-	-	Y	-	BLOB	-	-	-
LONGVARG	-	-	-	-	-	-	DBCLOB	-	-	-
REAL	-	-	-	-	-	-	REAL	-	-	-
SMALLINT	-	-	-	-	-	-	SMALLINT	-	0	-
TIME	-	-	-	-	-	-	TIME	-	0	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
TIMESTMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	0	N

Table 41. DB2 Database for Linux, UNIX, and Windows forward default data type mappings (Not all columns shown) (continued)

REMOTE_ TYPE NAME	REMOTE_ LOWER _LEN	REMOTE_ UPPER_ LEN	REMOTE_ LOWER_ SCALE	REMOTE_ UPPER_ SCALE	REMOTE_ BIT_ DATA	REMOTE_ DATA_ OPERATORS	FEDERATED_ TYPE NAME	FEDERATED LENGTH	FEDERATED SCALE	FEDERATED_ BIT_ DATA
VARCHAR	-	-	-	-	Y	-	VARCHAR	-	0	Y
VARGRAPH	-	-	-	-	-	-	VARGRAPHIC	-	0	N
VARGRAPHIC	-	-	-	-	-	-	VARGRAPHIC	-	0	N

DB2 for iSeries data sources

The following table lists the forward default data type mappings for DB2 for iSeries data sources.

Table 42. DB2 for iSeries forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHAR	1	254	-	-	-	-	CHAR	-	0	N
CHAR	255	32672	-	-	-	-	VARCHAR	-	0	N
CHAR	1	254	-	-	Y	-	CHAR	-	0	Y
CHAR	255	32672	-	-	Y	-	VARCHAR	-	0	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	-	-	-	-	-	DATE	-	0	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
FLOAT	4	-	-	-	-	-	REAL	-	-	-
FLOAT	8	-	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	1	127	-	-	-	-	GRAPHIC	-	0	N
GRAPHIC	128	16336	-	-	-	-	VARGRAPHIC	-	0	N
INTEGER	-	-	-	-	-	-	INTEGER	-	0	-
NUMERIC	-	-	-	-	-	-	DECIMAL	-	-	-
SMALLINT	-	-	-	-	-	-	SMALLINT	-	0	-
TIME	-	-	-	-	-	-	TIME	-	0	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
TIMESTMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
VARCHAR	1	32672	-	-	-	-	VARCHAR	-	0	N
VARCHAR	1	32672	-	-	Y	-	VARCHAR	-	0	Y
VARG	1	16336	-	-	-	-	VARGRAPHIC	-	0	N
VARGRAPHIC	1	16336	-	-	-	-	VARGRAPHIC	-	0	N

DB2 for VM and VSE data sources

The following table lists the forward default data type mappings for DB2 for VM and VSE data sources.

Table 43. DB2 Server for VM and VSE forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHAR	1	254	-	-	-	-	CHAR	-	0	N
CHAR	1	254	-	-	Y	-	CHAR	-	0	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	-	-	-	-	-	DATE	-	0	-
DBAHW	-	-	-	-	-	-	SMALLINT	-	0	-
DBAINT	-	-	-	-	-	-	INTEGER	-	0	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
FLOAT	4	-	-	-	-	-	REAL	-	-	-
FLOAT	8	-	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	1	127	-	-	-	-	GRAPHIC	-	0	N
INTEGER	-	-	-	-	-	-	INTEGER	-	-	-
SMALLINT	-	-	-	-	-	-	SMALLINT	-	-	-
TIME	-	-	-	-	-	-	TIME	-	0	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
TIMESTMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
VARCHAR	1	32672	-	-	-	-	VARCHAR	-	0	N
VARCHAR	1	32672	-	-	Y	-	VARCHAR	-	0	Y
VARGRAPHIC	1	16336	-	-	-	-	VARGRAPHIC	-	0	N
VARGRAPH	1	16336	-	-	-	-	VARGRAPHIC	-	0	N

DB2 for z/OS data sources

The following table lists the forward default data type mappings for DB2 for z/OS data sources.

Table 44. DB2 for z/OS forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHAR	1	254	-	-	-	-	CHAR	-	0	N
CHAR	255	32672	-	-	-	-	VARCHAR	-	0	N
CHAR	1	254	-	-	Y	-	CHAR	-	0	Y
CHAR	255	32672	-	-	Y	-	VARCHAR	-	0	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	-	-	-	-	-	DATE	-	0	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
FLOAT	4	-	-	-	-	-	REAL	-	-	-
FLOAT	8	-	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	1	127	-	-	-	-	GRAPHIC	-	0	N
INTEGER	-	-	-	-	-	-	INTEGER	-	0	-
ROWID	-	-	-	-	Y	-	VARCHAR	40	-	Y

Table 44. DB2 for z/OS forward default data type mappings (Not all columns shown) (continued)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
SMALLINT	-	-	-	-	-	-	SMALLINT	-	0	-
TIME	-	-	-	-	-	-	TIME	-	0	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
TIMESTMP	-	-	-	-	-	-	TIMESTAMP	-	0	-
VARCHAR	1	32672	-	-	-	-	VARCHAR	-	0	N
VARCHAR	1	32672	-	-	Y	-	VARCHAR	-	0	Y
VARG	1	16336	-	-	-	-	VARGRAPHIC	-	0	N
VARGRAPHIC	1	16336	-	-	-	-	VARGRAPHIC	-	0	N

Informix data sources

The following table lists the forward default data type mappings for Informix data sources.

Table 45. Informix forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BLOB	-	-	-	-	-	-	BLOB	2147483647	-	-
BOOLEAN	-	-	-	-	-	-	CHARACTER	1	-	-
BYTE	-	-	-	-	-	-	BLOB	2147483647	-	-
CHAR	1	254	-	-	-	-	CHARACTER	-	-	-
CHAR	255	32672	-	-	-	-	VARCHAR	-	-	-
CLOB	-	-	-	-	-	-	CLOB	2147483647	-	-
DATE	-	-	-	-	-	-	DATE	4	-	-
DATETIME	0	4	0	4	-	-	DATE	4	-	-
DATETIME	6	10	6	10	-	-	TIME	3	-	-
DATETIME	0	4	6	15	-	-	TIMESTAMP	10	-	-
DATETIME	6	10	11	15	-	-	TIMESTAMP	10	-	-
DECIMAL	1	31	0	31	-	-	DECIMAL	-	-	-
DECIMAL	32	130	-	-	-	-	DOUBLE	8	-	-
FLOAT	-	-	-	-	-	-	DOUBLE	8	-	-
INTEGER	-	-	-	-	-	-	INTEGER	4	-	-
INTERVAL	-	-	-	-	-	-	VARCHAR	25	-	-
INT8	-	-	-	-	-	-	BIGINT	19	0	-
LVARCHAR	1	32672	-	-	-	-	VARCHAR	-	-	-
MONEY	1	31	0	31	-	-	DECIMAL	-	-	-
MONEY	32	32	-	-	-	-	DOUBLE	8	-	-
NCHAR	1	254	-	-	-	-	CHARACTER	-	-	-
NCHAR	255	32672	-	-	-	-	VARCHAR	-	-	-
NVARCHAR	1	32672	-	-	-	-	VARCHAR	-	-	-
REAL	-	-	-	-	-	-	REAL	4	-	-
SERIAL	-	-	-	-	-	-	INTEGER	4	-	-
SERIAL8	-	-	-	-	-	-	BIGINT	-	-	-
SMALLFLOAT	-	-	-	-	-	-	REAL	4	-	-

Table 45. Informix forward default data type mappings (Not all columns shown) (continued)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
SMALLINT	-	-	-	-	-	-	SMALLINT	2	-	-
TEXT	-	-	-	-	-	-	CLOB	2147483647	-	-
VARCHAR	1	32672	-	-	-	-	VARCHAR	-	-	-

Notes:

- For the Informix DATETIME data type, the DB2 UNIX and Windows federated server uses the Informix high-level qualifier as the REMOTE_LENGTH and the Informix low-level qualifier as the REMOTE_SCALE.

  The Informix qualifiers are the "TU_" constants defined in the Informix Client SDK `datatime.h` file. The constants are:

0 = YEAR	8 = MINUTE	13 = FRACTION(3)
2 = MONTH	10 = SECOND	14 = FRACTION(4)
4 = DAY	11 = FRACTION(1)	15 = FRACTION(5)
6 = HOUR	12 = FRACTION(2)	

Microsoft SQL Server data sources

The following table lists the forward default data type mappings for Microsoft SQL Server data sources.

Table 46. Microsoft SQL Server forward default data type mappings

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
bigint [2]	-	-	-	-	-	-	BIGINT	-	-	-
binary	1	254	-	-	-	-	CHARACTER	-	-	Y
binary	255	8000	-	-	-	-	VARCHAR	-	-	Y
bit	-	-	-	-	-	-	SMALLINT	2	-	-
char	1	254	-	-	-	-	CHAR	-	-	N
char	255	8000	-	-	-	-	VARCHAR	-	-	N
datetime	-	-	-	-	-	-	TIMESTAMP	10	-	-
datetimen	-	-	-	-	-	-	TIMESTAMP	10	-	-
decimal	1	31	0	31	-	-	DECIMAL	-	-	-
decimal	32	38	0	38	-	-	DOUBLE	-	-	-
decimaln	1	31	0	31	-	-	DECIMAL	-	-	-
decimaln	32	38	0	38	-	-	DOUBLE	-	-	-
DUMMY2000 [1]	1	38	-84	127	-	-	DOUBLE	-	-	-
float	-	8	-	-	-	-	DOUBLE	8	-	-
floatn	-	8	-	-	-	-	DOUBLE	8	-	-
float	-	4	-	-	-	-	REAL	4	-	-
floatn	-	4	-	-	-	-	REAL	4	-	-
image	-	-	-	-	-	-	BLOB	2147483647	-	Y
int	-	-	-	-	-	-	INTEGER	4	-	-
intn	-	-	-	-	-	-	INTEGER	4	-	-
money	-	-	-	-	-	-	DECIMAL	19	4	-
moneyn	-	-	-	-	-	-	DECIMAL	19	4	-
nchar	1	127	-	-	-	-	CHAR	-	-	N
nchar	128	4000	-	-	-	-	VARCHAR	-	-	N

Table 46. Microsoft SQL Server forward default data type mappings (continued)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
numeric	1	31	0	31	-	-	DECIMAL	-	-	-
numeric	32	38	0	38	-	-	DOUBLE	8	-	-
numericn	32	38	0	38	-	-	DOUBLE	-	-	-
numericn	1	31	0	31	-	-	DECIMAL	-	-	-
ntext	-	-	-	-	-	-	CLOB	2147483647	-	Y
nvarchar	1	4000	-	-	-	-	VARCHAR	-	-	N
real	-	-	-	-	-	-	REAL	4	-	-
smallint	-	-	-	-	-	-	SMALLINT	2	-	-
smalldatetime	-	-	-	-	-	-	TIMESTAMP	10	-	-
smallmoney	-	-	-	-	-	-	DECIMAL	10	4	-
smallmoneyn	-	-	-	-	-	-	DECIMAL	10	4	-
SQL_BIGINT	-	-	-	-	-	-	DECIMAL	-	-	-
SQL_BIGINT [2]	-	-	-	-	-	-	BIGINT	-	-	-
SQL_BINARY	1	254	-	-	-	-	CHARACTER	-	-	Y
SQL_BINARY	255	8000	-	-	-	-	VARCHAR	-	-	Y
SQL_BIT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_CHAR	1	254	-	-	-	-	CHAR	-	-	N
SQL_CHAR	255	8000	-	-	-	-	VARCHAR	-	-	N
SQL_DATE	-	-	-	-	-	-	DATE	4	-	-
SQL_DECIMAL	1	31	0	31	-	-	DECIMAL	-	-	-
SQL_DECIMAL	32	38	0	38	-	-	DOUBLE	8	-	-
SQL_DECIMAL	32	32	0	31	-	-	DOUBLE	8	-	-
SQL_DOUBLE	-	-	-	-	-	-	DOUBLE	8	-	-
SQL_FLOAT	-	-	-	-	-	-	DOUBLE	8	-	-
SQL_GUID	1	4000	-	-	Y	-	VARCHAR	16	-	Y
SQL_INTEGER	-	-	-	-	-	-	INTEGER	4	-	-
SQL_ LONGVARCHAR	-	-	-	-	-	-	CLOB	2147483647	-	N
SQL_ LONGVARBINARY	-	-	-	-	-	-	BLOB	-	-	Y
SQL_NUMERIC	1	31	0	31	-	-	DECIMAL	-	-	-
SQL_REAL	-	-	-	-	-	-	DOUBLE	8	-	-
SQL_SMALLINT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_TIME	-	-	-	-	-	-	TIME	3	-	-
SQL_TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	10	-	-
SQL_TINYINT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_VARBINARY	1	8000	-	-	-	-	VARCHAR	-	-	Y
SQL_VARCHAR	1	8000	-	-	-	-	VARCHAR	-	-	N
text	-	-	-	-	-	-	CLOB	-	-	N
timestamp	-	-	-	-	-	-	VARCHAR	8		Y
tinyint	-	-	-	-	-	-	SMALLINT	2	-	-
uniqueidentifier	1	4000	-	-	Y	-	VARCHAR	16	-	Y
varbinary	1	8000	-	-	-	-	VARCHAR	-	-	Y
varchar	1	8000	-	-	-	-	VARCHAR	-	-	N

Table 46. Microsoft SQL Server forward default data type mappings (continued)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data

Note:

1. This type mapping is valid only with Windows 2000 operating systems.
2. This type mapping is valid only with Microsoft SQL Server Version 2000.

ODBC data sources

The following table lists the forward default data type mappings for ODBC data sources.

Table 47. ODBC forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
SQL_BIGINT	-	-	-	-	-	-	BIGINT	8	-	-
SQL_BINARY	1	254	-	-	-	-	CHARACTER	-	-	Y
SQL_BINARY	255	32672	-	-	-	-	VARCHAR	-	-	Y
SQL_BIT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_CHAR	1	254	-	-	-	-	CHAR	-	-	N
SQL_CHAR	255	32672	-	-	-	-	VARCHAR	-	-	N
SQL_DECIMAL	1	31	0	31	-	-	DECIMAL	-	-	-
SQL_DECIMAL	32	38	0	38	-	-	DOUBLE	8	-	-
SQL_DOUBLE	-	-	-	-	-	-	DOUBLE	8	-	-
SQL_FLOAT	-	-	-	-	-	-	DOUBLE	8	-	-
SQL_INTEGER	-	-	-	-	-	-	INTEGER	4	-	-
SQL_ LONGVARCHAR	-	-	-	-	-	-	CLOB	2147483647	-	N
SQL_ LONGVARBINARY	-	-	-	-	-	-	BLOB	-	-	Y
SQL_NUMERIC	1	31	0	31	-	-	DECIMAL	-	-	-
SQL_NUMERIC	32	32	0	31	-	-	DOUBLE	8	-	-
SQL_REAL	-	-	-	-	-	-	REAL	4	-	-
SQL_SMALLINT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_TYPE_DATE	-	-	-	-	-	-	DATE	4	-	-
SQL_TYPE_TIME	-	-	-	-	-	-	TIME	3	-	-
SQL_TYPE_ TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	10	-	-
SQL_TINYINT	-	-	-	-	-	-	SMALLINT	2	-	-
SQL_VARBINARY	1	32672	-	-	-	-	VARCHAR	-	-	Y
SQL_VARCHAR	1	32672	-	-	-	-	VARCHAR	-	-	N
SQL_WCHAR	1	127	-	-	-	-	CHAR	-	-	N
SQL_WCHAR	128	16336	-	-	-	-	VARCHAR	-	-	N
SQL_WVARCHAR	1	16336	-	-	-	-	VARCHAR	-	-	N
SQL_ WLONGVARCHAR	-	1073741823	-	-	-	-	CLOB	2147483647	-	N

Oracle NET8 data sources

The following table lists the forward default data type mappings for Oracle NET8 data sources.

Table 48. Oracle NET8 forward default data type mappings

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BLOB	0	0	0	0	-	\0	BLOB	2147483647	0	Y
CHAR	1	254	0	0	-	\0	CHAR	0	0	N
CHAR	255	2000	0	0	-	\0	VARCHAR	0	0	N
CLOB	0	0	0	0	-	\0	CLOB	2147483647	0	N
DATE	0	0	0	0	-	\0	TIMESTAMP	0	0	N
FLOAT	1	126	0	0	-	\0	DOUBLE	0	0	N
LONG	0	0	0	0	-	\0	CLOB	2147483647	0	N
LONG RAW	0	0	0	0	-	\0	BLOB	2147483647	0	Y
NUMBER	10	18	0	0	-	\0	BIGINT	0	0	N
NUMBER	1	38	-84	127	-	\0	DOUBLE	0	0	N
NUMBER	1	31	0	31	-	>=	DECIMAL	0	0	N
NUMBER	1	4	0	0	-	\0	SMALLINT	0	0	N
NUMBER	5	9	0	0	-	\0	INTEGER	0	0	N
NUMBER	-	10	0	0	-	\0	DECIMAL	0	0	N
RAW	1	2000	0	0	-	\0	VARCHAR	0	0	Y
ROWID	0	0	0	NULL	-	\0	CHAR	18	0	N
TIMESTAMP [1]	-	-	-	-	-	-	TIMESTAMP	10	-	-
VARCHAR2	1	4000	0	0	-	\0	VARCHAR	0	0	N

Note:

1. This type mapping is valid only for Oracle 9i (or later) client and server configurations.

Sybase data sources

The following table lists the forward default data type mappings for Sybase data sources.

Table 49. Sybase CTLIB forward default data type mappings

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
binary	1	254	-	-	-	-	CHAR	-	-	Y
binary	255	16384	-	-	-	-	VARCHAR	-	-	Y
bit	-	-	-	-	-	-	SMALLINT	-	-	-
char	1	254	-	-	-	-	CHAR	-	-	N
char	255	16384	-	-	-	-	VARCHAR	-	-	N
char null (see varchar)										
datetime	-	-	-	-	-	-	TIMESTAMP	-	-	-
datetimn	-	-	-	-	-	-	TIMESTAMP	-	-	-
decimal	1	31	0	31	-	-	DECIMAL	-	-	-
decimal	32	38	0	38	-	-	DOUBLE	-	-	-
decimaln	1	31	0	31	-	-	DECIMAL	-	-	-

Table 49. Sybase CTLIB forward default data type mappings (continued)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
decimaln	32	38	0	38	-	-	DOUBLE	-	-	-
float	-	4	-	-	-	-	REAL	-	-	-
float	-	8	-	-	-	-	DOUBLE	-	-	-
floatn	-	4	-	-	-	-	REAL	-	-	-
floatn	-	8	-	-	-	-	DOUBLE	-	-	-
image	-	-	-	-	-	-	BLOB	-	-	-
int	-	-	-	-	-	-	INTEGER	-	-	-
intn	-	-	-	-	-	-	INTEGER	-	-	-
money	-	-	-	-	-	-	DECIMAL	19	4	-
moneyn	-	-	-	-	-	-	DECIMAL	19	4	-
nchar	1	254	-	-	-	-	CHAR	-	-	N
nchar	255	16384	-	-	-	-	VARCHAR	-	-	N
nchar null (see nvarchar)										
numeric	1	31	0	31	-	-	DECIMAL	-	-	-
numeric	32	38	0	38	-	-	DOUBLE	-	-	-
numericn	1	31	0	31	-	-	DECIMAL	-	-	-
numericn	32	38	0	38	-	-	DOUBLE	-	-	-
nvarchar	1	16384	-	-	-	-	VARCHAR	-	-	N
real	-	-	-	-	-	-	REAL	-	-	-
smalldatetime	-	-	-	-	-	-	TIMESTAMP	-	-	-
smallint	-	-	-	-	-	-	SMALLINT	-	-	-
smallmoney	-	-	-	-	-	-	DECIMAL	10	4	-
sysname	1	254	-	-	-	-	CHAR	-	-	N
text	-	-	-	-	-	-	CLOB	-	-	-
timestamp	-	-	-	-	-	-	VARCHAR	8	-	Y
tinyint	-	-	-	-	-	-	SMALLINT	-	-	-
unichar[1]	1	254	-	-	-	-	CHAR	-	-	N
unichar[1]	255	16384	-	-	-	-	VARCHAR	-	-	N
unichar null (see univarchar)										
univarchar[1]	1	16384	-	-	-	-	VARCHAR	-	-	N
varbinary	1	16384	-	-	-	-	VARCHAR	-	-	Y
varchar	1	16384	-	-	-	-	VARCHAR	-	-	N

Note:

1. Valid for non-Unicode federated databases.

Teradata data sources

The following table lists the forward default data type mappings for Teradata data sources.

Table 50. Teradata forward default data type mappings (Not all columns shown)

Remote Typename	Remote Lower Len	Remote Upper Len	Remote Lower Scale	Remote Upper Scale	Remote Bit Data	Remote Data Operators	Federated Typename	Federated Length	Federated Scale	Federated Bit Data
BYTE	1	254	-	-	-	-	CHAR	-	-	Y
BYTE	255	32672	-	-	-	-	VARCHAR	-	-	Y
BYTE	32673	64000	-	-	-	-	BLOB	-	-	-
BYTEINT	-	-	-	-	-	-	SMALLINT	-	-	-
CHAR	1	254	-	-	-	-	CHARACTER	-	-	-
CHAR	255	32672	-	-	-	-	VARCHAR	-	-	-
CHAR	32673	64000	-	-	-	-	CLOB	-	-	-
DATE	-	-	-	-	-	-	DATE	-	-	-
DECIMAL	1	18	0	18	-	-	DECIMAL	-	-	-
DOUBLE PRECISION	-	-	-	-	-	-	DOUBLE	-	-	-
FLOAT	-	-	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	1	127	-	-	-	-	GRAPHIC	-	-	-
GRAPHIC	128	16336	-	-	-	-	VARGRAPHIC	-	-	-
GRAPHIC	16337	32000	-	-	-	-	DBCLOB	-	-	-
INTEGER	-	-	-	-	-	-	INTEGER	-	-	-
INTERVAL	-	-	-	-	-	-	CHAR	-	-	-
NUMERIC	1	18	0	18	-	-	DECIMAL	-	-	-
REAL	-	-	-	-	-	-	DOUBLE	-	-	-
SMALLINT	-	-	-	-	-	-	SMALLINT	-	-	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	-	-
VARBYTE	1	32762	-	-	-	-	VARCHAR	-	-	Y
VARBYTE	32763	64000	-	-	-	-	BLOB	-	-	-
VARCHAR	1	32672	-	-	-	-	VARCHAR	-	-	-
VARCHAR	32673	64000	-	-	-	-	CLOB	-	-	-
VARGRAPHIC	1	16336	-	-	-	-	VARGRAPHIC	-	-	-
VARGRAPHIC	16337	32000	-	-	-	-	DBCLOB	-	-	-

Default reverse data type mappings

For most data sources, the default type mappings are in the wrappers.

The two kinds of mappings between data source data types and federated database data types are forward type mappings and reverse type mappings. In a forward type mapping, the mapping is from a remote type to a comparable local type. The other type of mapping is a reverse type mapping, which is used with transparent DDL to create or modify remote tables.

The default type mappings for DB2 family data sources are in the DRDA wrapper. The default type mappings for Informix are in the INFORMIX wrapper, and so forth.

When you define a remote table or view to the federated database, the definition includes a reverse type mapping. The mapping is from a local federated database

data type for each column, and the corresponding remote data type. For example, there is a default reverse type mapping in which the local type REAL points to the Informix type SMALLFLOAT.

Federated databases do not support mappings for LONG VARCHAR, LONG VARGRAPHIC, DATALINK, and user-defined types.

When you use the CREATE TABLE statement to create a remote table, you specify the local data types you want to include in the remote table. These default reverse type mappings will assign corresponding remote types to these columns. For example, suppose that you use the CREATE TABLE statement to define an Informix table with a column C2. You specify BIGINT as the data type for C2 in the statement. The default reverse type mapping of BIGINT depends on which version of Informix you are creating the table on. The mapping for C2 in the Informix table will be to DECIMAL in Informix Version 8 and to INT8 in Informix Version 9.

You can override a default reverse type mapping, or create a new reverse type mapping with the CREATE TYPE MAPPING statement.

The following tables show the default reverse mappings between federated database local data types and remote data source data types.

These mappings are valid with all the supported versions, unless otherwise noted.

DB2 Database for Linux, UNIX, and Windows data sources

The following table lists the reverse default data type mappings for DB2 Database for Linux, UNIX, and Windows data sources.

Table 51. DB2 Database for Linux, UNIX, and Windows reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Federated Bit Data
BIGINT	-	8	-	-	-	-	BIGINT	-	-	-
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHARACTER	-	-	-	-	-	-	CHAR	-	-	N
CHARACTER	-	-	-	-	Y	-	CHAR	-	-	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	4	-	-	-	-	DATE	-	-	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
DOUBLE	-	8	-	-	-	-	DOUBLE	-	-	-
FLOAT	-	8	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	-	N
INTEGER	-	4	-	-	-	-	INTEGER	-	-	-
REAL	-	-	-	-	-	-	REAL	-	-	-
SMALLINT	-	2	-	-	-	-	SMALLINT	-	-	-
TIME	-	3	-	-	-	-	TIME	-	-	-
TIMESTAMP	-	10	-	-	-	-	TIMESTAMP	-	-	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	-	N
VARCHAR	-	-	-	-	Y	-	VARCHAR	-	-	Y
VARGRAPH	-	-	-	-	-	-	VARGRAPHIC	-	-	N

Table 51. DB2 Database for Linux, UNIX, and Windows reverse default data type mappings (Not all columns shown) (continued)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Federated Bit Data
VARGRAPHIC	-	-	-	-	-	-	VARGRAPHIC	-	-	-

DB2 for iSeries data sources

The following table lists the reverse default data type mappings for DB2 for iSeries data sources.

Table 52. DB2 for iSeries reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operations	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHARACTER	-	-	-	-	-	-	CHARACTER	-	-	N
CHARACTER	-	-	-	-	Y	-	CHARACTER	-	-	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	4	-	-	-	-	DATE	-	-	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	NUMERIC	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
DOUBLE	-	8	-	-	-	-	FLOAT	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	-	N
INTEGER	-	4	-	-	-	-	INTEGER	-	-	-
REAL	-	4	-	-	-	-	FLOAT	-	-	-
SMALLINT	-	2	-	-	-	-	SMALLINT	-	-	-
TIME	-	3	-	-	-	-	TIME	-	-	-
TIMESTAMP	-	10	-	-	-	-	TIMESTAMP	-	-	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	-	N
VARCHAR	-	-	-	-	Y	-	VARCHAR	-	-	Y
VARGRAPHIC	-	-	-	-	-	-	VARG	-	-	N

DB2 for VM and VSE data sources

The following table lists the reverse default data type mappings for DB2 for VM and VSE data sources.

Table 53. DB2 for VM and VSE reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHARACTER	-	-	-	-	-	-	CHAR	-	-	-
CHARACTER	-	-	-	-	Y	-	CHAR	-	-	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	4	-	-	-	-	DATE	-	-	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-

Table 53. DB2 for VM and VSE reverse default data type mappings (Not all columns shown) (continued)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
DOUBLE	-	8	-	-	-	-	FLOAT	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	-	N
INTEGER	-	4	-	-	-	-	INTEGER	-	-	-
REAL	-	4	-	-	-	-	REAL	-	-	-
SMALLINT	-	2	-	-	-	-	SMALLINT	-	-	-
TIME	-	3	-	-	-	-	TIME	-	-	-
TIMESTAMP	-	10	-	-	-	-	TIMESTAMP	-	-	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	-	-
VARCHAR	-	-	-	-	Y	-	VARCHAR	-	-	Y
VARGRAPH	-	-	-	-	-	-	VARGRAPH	-	-	N

DB2 for z/OS data sources

The following table lists the reverse default data type mappings for DB2 for z/OS data sources.

Table 54. DB2 for z/OS reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BLOB	-	-	-	-	-	-	BLOB	-	-	-
CHARACTER	-	-	-	-	-	-	CHAR	-	-	N
CHARACTER	-	-	-	-	Y	-	CHAR	-	-	Y
CLOB	-	-	-	-	-	-	CLOB	-	-	-
DATE	-	4	-	-	-	-	DATE	-	-	-
DBCLOB	-	-	-	-	-	-	DBCLOB	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
DOUBLE	-	8	-	-	-	-	DOUBLE	-	-	–
FLOAT	-	8	-	-	-	-	DOUBLE	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	-	N
INTEGER	-	4	-	-	-	-	INTEGER	-	-	-
REAL	-	4	-	-	-	-	REAL	-	-	-
SMALLINT	-	2	-	-	-	-	SMALLINT	-	-	-
TIME	-	3	-	-	-	-	TIME	-	-	-
TIMESTAMP	-	10	-	-	-	-	TIMESTAMP	-	-	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	-	N
VARCHAR	-	-	-	-	Y	-	VARCHAR	-	-	Y
VARGRAPHIC	-	-	-	-	-	-	VARGRAPHIC	-	-	N

Informix data sources

The following table lists the reverse default data type mappings for Informix data sources.

Table 55. Informix reverse default data type mappings

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BIGINT [1]	-	-	-	-	-	-	DECIMAL	19	-	-
BIGINT [2]	-	-	-	-	-	-	INT8	-	-	-
BLOB	1	2147483647	-	-	-	-	BYTE	-	-	-
CHARACTER	-	-	-	-	N	-	CHAR	-	-	-
CHARACTER	-	-	-	-	Y	-	BYTE	-	-	-
CLOB	1	2147483647	-	-	-	-	TEXT	-	-	-
DATE	-	4	-	-	-	-	DATE	-	-	-
DECIMAL	-	-	-	-	-	-	DECIMAL	-	-	-
DOUBLE	-	8	-	-	-	-	FLOAT	-	-	-
INTEGER	-	4	-	-	-	-	INTEGER	-	-	-
REAL	-	4	-	-	-	-	SMALLFLOAT	-	-	-
SMALLINT	-	2	-	-	-	-	SMALLINT	-	-	-
TIME	-	3	-	-	-	-	DATETIME	6	10	-
TIMESTAMP	-	10	-	-	-	-	DATETIME	0	15	-
VARCHAR	1	254	-	-	N	-	VARCHAR	-	-	-
VARCHAR [1]	255	32672	-	-	N	-	TEXT	-	-	-
VARCHAR	-	-	-	-	Y	-	BYTE	-	-	-
VARCHAR [2]	255	2048	-	-	N	-	LVARCHAR	-	-	-
VARCHAR [2]	2049	32672	-	-	N	-	TEXT	-	-	-

Note:

1. This type mapping is valid only with Informix server Version 8 (or lower).
2. This type mapping is valid only with Informix server Version 9 (or higher).

For the Informix DATETIME data type, the federated server uses the Informix high-level qualifier as the REMOTE_LENGTH and the Informix low-level qualifier as the REMOTE_SCALE.

The Informix qualifiers are the ″TU_″ constants defined in the Informix Client SDK `datatime.h` file. The constants are:

0 = YEAR	8 = MINUTE	13 = FRACTION(3)
2 = MONTH	10 = SECOND	14 = FRACTION(4)
4 = DAY	11 = FRACTION(1)	15 = FRACTION(5)
6 = HOUR	12 = FRACTION(2)	

Microsoft SQL Server data sources

The following table lists the reverse default data type mappings for Microsoft SQL Server data sources.

Table 56. Microsoft SQL Server reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BIGINT [1]	-	-	-	-	-	-	bigint	-	-	-
BLOB	-	-	-	-	-	-	image	-	-	-
CHARACTER	-	-	-	-	Y	-	binary	-	-	-
CHARACTER	-	-	-	-	N	-	char	-	-	-
CLOB	-	-	-	-	-	-	text	-	-	-
DATE	-	4	-	-	-	-	datetime	-	-	-

Table 56. Microsoft SQL Server reverse default data type mappings (Not all columns shown) (continued)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
DECIMAL	-	-	-	-	-	-	decimal	-	-	-
DOUBLE	-	8	-	-	-	-	float	-	-	-
INTEGER	-	-	-	-	-	-	int	-	-	-
SMALLINT	-	-	-	-	-	-	smallint	-	-	-
REAL	-	4	-	-	-	-	real	-	-	-
TIME	-	3	-	-	-	-	datetime	-	-	-
TIMESTAMP	-	10	-	-	-	-	datetime	-	-	-
VARCHAR	1	8000	-	-	N	-	varchar	-	-	-
VARCHAR	8001	32672	-	-	N	-	text	-	-	-
VARCHAR	1	8000	-	-	Y	-	varbinary	-	-	-
VARCHAR	8001	32672	-	-	Y	-	image	-	-	-

Note:

1. This type mapping is valid only with Microsoft SQL Server Version 2000.

Oracle NET8 data sources

The following table lists the reverse default data type mappings for Oracle NET8 data sources.

Table 57. Oracle NET8 reverse default data type mappings

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BIGINT	0	8	0	0	N	\0	NUMBER	19	0	N
BLOB	0	2147483647	0	0	Y	\0	BLOB	0	0	Y
CHARACTER	1	254	0	0	N	\0	CHAR	0	0	N
CHARACTER	1	254	0	0	Y	\0	RAW	0	0	Y
CLOB	0	2147483647	0	0	N	\0	CLOB	0	0	N
DATE	0	4	0	0	N	\0	DATE	0	0	N
DECIMAL	0	0	0	0	N	\0	NUMBER	0	0	N
DOUBLE	0	8	0	0	N	\0	FLOAT	126	0	N
FLOAT	0	8	0	0	N	\0	FLOAT	126	0	N
INTEGER	0	4	0	0	N	\0	NUMBER	10	0	N
REAL	0	4	0	0	N	\0	FLOAT	63	0	N
SMALLINT	0	2	0	0	N	\0	NUMBER	5	0	N
TIME	0	3	0	0	N	\0	DATE	0	0	N
TIMESTAMP [1]	0	10	0	0	N	\0	DATE	0	0	N
TIMESTAMP [2]	0	10	0	0	N	\0	TIMESTAMP	6	0	N
VARCHAR	1	4000	0	0	N	\0	VARCHAR2	0	0	N
VARCHAR	1	2000	0	0	Y	\0	RAW	0	0	Y

Note:

1. This type mapping is valid only with Oracle Version 8.
2. This type mapping is valid only with Oracle Version 9 and Version 10.

Sybase data sources

The following table lists the reverse default data type mappings for Sybase data sources.

Table 58. Sybase CTLIB default reverse data type mappings

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BIGINT	-	-	-	-	-	-	decimal	19	0	-
BLOB	-	-	-	-	-	-	image	-	-	-
CHARACTER	-	-	-	-	N	-	char	-	-	-
CHARACTER	-	-	-	-	Y	-	binary	-	-	-
CLOB	-	-	-	-	-	-	text	-	-	-
DATE	-	-	-	-	-	-	datetime	-	-	-
DECIMAL	-	-	-	-	-	-	decimal	-	-	-
DOUBLE	-	-	-	-	-	-	float	-	-	-
INTEGER	-	-	-	-	-	-	integer	-	-	-
REAL	-	-	-	-	-	-	real	-	-	-
SMALLINT	-	-	-	-	-	-	smallint	-	-	-
TIME	-	-	-	-	-	-	datetime	-	-	-
TIMESTAMP	-	-	-	-	-	-	datetime	-	-	-
VARCHAR[1]	1	255	-	-	N	-	varchar	-	-	-
VARCHAR[1]	256	32672	-	-	N	-	text	-	-	-
VARCHAR [2]	1	16384	-	-	N	-	varchar	-	-	-
VARCHAR [2]	16385	32672	-	-	N	-	text	-	-	-
VARCHAR[1]	1	255	-	-	Y	-	varbinary	-	-	-
VARCHAR[1]	256	32672	-	-	Y	-	image	-	-	-
VARCHAR [2]	1	16384	-	-	Y	-	varbinary	-	-	-
VARCHAR [2]	16385	32672	-	-	Y	-	image	-	-	-

Note:

1. This type mapping is valid only for CTLIB with Sybase server version 12.0 (or earlier).
2. This type mapping is valid only for CTLIB with Sybase server version 12.5 (or later).

Teradata data sources

The following table lists the reverse default data type mappings for Teradata data sources.

Table 59. Teradata reverse default data type mappings (Not all columns shown)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
BLOB [1]	1	64000	-	-	-	-	VARBYTE	-	-	-
CHARACTER	-	-	-	-	-	-	CHARACTER	-	-	-
CHARACTER	-	-	-	-	Y	-	BYTE	-	-	-
CLOB [2]	1	64000	-	-		-	VARCHAR	-	-	-
DATE	-	-	-	-	-	-	DATE	-	-	-
DBCLOB [3]	1	32000	-	-	-	-	VARGRAPHIC	-	-	-
DECIMAL	1	18	0	18	-	-	DECIMAL	-	-	-
DECIMAL	19	31	0	31	-	-	FLOAT	-	-	-

Table 59. Teradata reverse default data type mappings (Not all columns shown) (continued)

Federated Typename	Federated Lower Len	Federated Upper Len	Federated Lower Scale	Federated Upper Scale	Federated Bit Data	Federated Data Operators	Remote Typename	Remote Length	Remote Scale	Remote Bit Data
DOUBLE	-	-	-	-	-	-	FLOAT	-	-	-
GRAPHIC	-	-	-	-	-	-	GRAPHIC	-	-	-
INTEGER	-	-	-	-	-	-	INTEGER	-	-	-
REAL	-	-	-	-	-	-	FLOAT	-	-	-
SMALLINT	-	-	-	-	-	-	SMALLINT	-	-	-
TIME	-	-	-	-	-	-	TIME	-	-	-
TIMESTAMP	-	-	-	-	-	-	TIMESTAMP	-	-	-
VARCHAR	-	-	-	-	-	-	VARCHAR	-	-	-
VARCHAR	-	-	-	-	Y	-	VARBYTE	-	-	-
VARGRAPHIC	-	-	-	-	-	-	VARGRAPHIC	-	-	-

Note:

1. The Teradata VARBYTE data type can contain only the specified length (1 to 64000) of a BLOB data type.
2. The Teradata VARCHAR data type can contain only the specified length (1 to 64000) of a CLOB data type.
3. The Teradata VARGRAPHIC data type can contain only the specified length (1 to 32000) of a DBCLOB data type.

Unicode default data type mappings

Unicode default forward data type mappings - Microsoft SQL Server wrapper

The following table lists the default forward data type mapping for the Microsoft SQL Server wrapper when the federated database is a Unicode database.

Table 60. Unicode default forward data type mappings for the Microsoft SQL Server wrapper

UTF-8	Microsoft SQL Server	
Data type	**Data type**	**Length**
CHAR	CHAR	1 to 254 bytes
VARCHAR	CHAR	255 to 8000 bytes
	VARCHAR	1 to 8000 bytes
CLOB	TEXT	-
GRAPHIC	NCHAR	1 to 127 characters
VARGRAPHIC	NCHAR	128 to 16336 characters
	NVARCHAR	1 to 16336 characters
DBCLOB	NTEXT	-

Unicode default reverse data type mappings - Microsoft SQL Server wrapper

The following table lists the default reverse data type mapping for the Microsoft SQL Server wrapper when the federated database is a Unicode database.

Table 61. Unicode default reverse data type mappings for the Microsoft SQL Server wrapper

UTF-8		Microsoft SQL Server
Data type	**Length**	**Data type**
CHAR	1 to 254 bytes	CHAR
VARCHAR	1 to 32672 bytes	VARCHAR
CLOB	1 to 2 147 483 647 bytes	TEXT
GRAPHIC	1 to 127 characters	NCHAR
VARGRAPHIC	1 to 16336 characters	NVARCHAR
DBCLOB	1 to 1 073 741 823 characters	NTEXT

Unicode default forward data type mappings - NET8 wrapper

The following table lists the default forward data type mapping for the NET8 wrapper when the federated database is a Unicode database.

Table 62. Unicode default forward data type mappings for the NET8 wrapper

UTF-8	Oracle	
Data type	**Data type**	**Length**
CHAR	CHAR	1 to 254 bytes
VARCHAR	CHAR	255 to 2000 bytes
	VARCHAR2	1 to 4000 bytes
DBCLOB	NCLOB	
GRAPHIC	NCHAR	1 to 127 characters
VARGRAPHIC	NCHAR	128 to 1000 characters
	NVARCHAR2	1 to 2000 characters

Unicode default reverse data type mappings - NET8 wrapper

The following table lists the default reverse data type mapping for the NET8 wrapper when the federated database is a Unicode database.

Table 63. Unicode default reverse data type mappings for the NET8 wrapper

UTF-8		Oracle
Data type	**Length**	**Data type**
CHAR	1 to 254 bytes	CHAR
VARCHAR	1 to 4000 bytes	VARCHAR2
CLOB	1 to 2 147 483 647 bytes	CLOB
GRAPHIC	1 to 127 characters	NCHAR
VARGRAPHIC	1 to 2000 characters	NVARCHAR2
DBCLOB	1 to 1 073 741 823 characters	NCLOB

Unicode default forward data type mappings - ODBC wrapper

The following table lists the default forward data type mapping for the ODBC wrapper when the federated database is a Unicode database.

Table 64. Unicode default forward data type mappings for the ODBC wrapper

UTF-8	ODBC	
Data type	**Data type**	**Length**
CHAR	SQL_CHAR	1 to 254 bytes
VARCHAR	SQL_CHAR	255 to 32672 bytes
	SQL_VARCHAR	1 to 32672 bytes
CLOB	SQL_LONGVARCHAR	-
GRAPHIC	SQL_WCHAR	1 to 127 characters
VARGRAPHIC	SQL_WVARCHAR	128 to 16336 characters
	SQL_WVARCHAR	1 to 16336 characters
DBCLOB	SQL_WLONGVARCHAR	-

Unicode default reverse data type mappings - ODBC wrapper

The following table lists the default reverse data type mapping for the ODBC wrapper when the federated database is a Unicode database.

Table 65. Unicode default reverse data type mappings for the ODBC wrapper

UTF-8		ODBC
Data type	**Length**	**Data type**
CHAR	1 to 254 bytes	SQL_CHAR
VARCHAR	1 to 32672 bytes	SQL_VARCHAR
CLOB	1 to 2 147 483 647 bytes	SQL_LONGVARCHAR
GRAPHIC	1 to 127 characters	SQL_WCHAR
VARGRAPHIC	1 to 16336 characters	SQL_WVARCHAR
DBCLOB	1 to 1 073 741 823 characters	SQL_WLONGVARCHAR

Unicode default forward data type mappings - Sybase wrapper

The following table lists the default forward data type mapping for the CTLIB wrapper when the federated database is a Unicode database.

Table 66. Unicode default forward data type mappings for the Sybase CTLIB wrapper

UTF-8	Sybase	
Data type	**Data type**	**Length**
CHAR	char	1 to 254 bytes
	nchar	1 to 127 characters
VARCHAR	char	255 to 32672 bytes
	varchar	1 to 32672 bytes
	nchar	128 to 16336 characters
	nvarchar	1 to 16336 characters
CLOB	text	
GRAPHIC	unichar	1 to 127 characters
VARGRAPHIC	unichar	128 to 16336 characters
	univarchar	1 to 16336 characters

Unicode default reverse data type mappings - Sybase wrapper

The following table lists the default reverse data type mapping for the CTLIB wrapper when the federated database is a Unicode database.

Table 67. Unicode default reverse data type mappings for the Sybase CTLIB wrapper

UTF-8		Sybase
Data type	**Length**	**Data type**
CHAR	1 to 254 bytes	char
VARCHAR	1 to 32672 bytes	varchar
CLOB	1 to 2 147 483 647 bytes	text
GRAPHIC	1 to 127 characters	unichar
VARGRAPHIC	1 to 16336 characters	univarchar

Data types supported for nonrelational data sources

For most of the nonrelational data sources, you must specify the column information, including data type, when you create the nicknames to access the data source.

Some of the nonrelational wrappers create all of the columns required to access a data source. These are called *fixed columns*. Other wrappers let you specify some or all of the data types for the columns in the CREATE NICKNAME statement.

The following sections list the wrappers that you can specify the data types for, and the data types that the wrapper supports.

Data types supported by the BioRS wrapper

The following table lists the DB2 data types that the BioRS wrapper supports.

Table 68. BioRS data types that map to DB2 data types

BioRS data types	DB2 data type
	CHARACTER
	CLOB
	VARCHAR

Data types supported by the BLAST wrapper

Some of the data types are automatically set for the fixed columns that the BLAST wrapper creates.

For the definition line fields, you can assign when you create a nickname. If the data in the definition line column is not compatible with the local column data type, you will get an error. For example, if you define a definition line column of type INTEGER and there are values in the column that are not numeric, an error is returned.

The following table lists the DB2 data types that the BLAST wrapper supports.

Table 69. BLAST data types that map to DB2 data types

BLAST data types	DB2 data type
definition line	CLOB (maximum length is 5 megabytes)
definition line	DOUBLE
definition line	FLOAT
definition line	INTEGER
definition line	VARCHAR

Data types supported by the Entrez wrapper

The following table lists the DB2 data types that the Entrez wrapper supports.

Table 70. Entrez data types that map to DB2 data types

Entrez data types	DB2 data type
	CHARACTER
	CLOB (maximum length is 5 megabytes)
	DATE
	DECIMAL
	DOUBLE
	INTEGER
	REAL
	SMALLINT
	TIMESTAMP
	VARCHAR

Data types supported by the Excel wrapper

The following table lists the DB2 data types that the Excel wrapper supports.

Table 71. Excel data types that map to DB2 data types

Excel data types	DB2 data type
	DATE
	FLOAT
	INTEGER
	VARCHAR

Data types supported by the HMMER wrapper

The following table lists the DB2 data types that the HMMER wrapper supports.

Table 72. HMMER data types that map to DB2 data types

HMMER data types	DB2 data type
	CLOB (maximum length is 5 megabytes)
	DOUBLE

Table 72. HMMER data types that map to DB2 data types (continued)

HMMER data types	DB2 data type
	FLOAT
	INTEGER
	VARCHAR

Data types supported by the Script wrapper

The following table lists the DB2 data types that the Script wrapper supports.

Table 73. Script data types that map to DB2 data types

Script data types	DB2 data type
	CLOB (maximum length is 5 megabytes)
	DATE
	DOUBLE
	INTEGER
	VARCHAR

Data types supported by the table-structured file wrapper

The following table lists the DB2 data types that the table-structured file wrapper supports.

Table 74. Table-structured file data types that map to DB2 data types

Table-structured file data types	DB2 data type
	CHARACTER
	CLOB (maximum length is 5 megabytes)
	DECIMAL
	DOUBLE
	FLOAT
	INTEGER
	REAL
	SMALLINT
	VARCHAR

Data types supported by the Web services wrapper

The following table lists the DB2 data types that the Web services wrapper supports. The Web services wrapper uses XML data types.

Table 75. XML data types that map to DB2 data types for the Web services wrapper

XML data types	DB2 data type
	BIGINT
	CHARACTER
	CHARACTER FOR BIT DATA
	CLOB (maximum length is 5 megabytes)

Table 75. XML data types that map to DB2 data types for the Web services wrapper (continued)

XML data types	DB2 data type
	DATE
	DECIMAL
	DOUBLE
	FLOAT
	INTEGER
	REAL
	TIME
	TIMESTAMP
	SMALLINT
	VARCHAR
	VARCHAR FOR BIT DATA

Data types supported by the WebSphere Business Integration wrapper

The following table lists the DB2 data types that the WebSphere Business Integration wrapper supports. The WebSphere Business Integration wrapper uses XML data types.

Table 76. XML data types that map to DB2 data types for the WebSphere Business Integration wrapper

XML data types	DB2 data type
	BIGINT
	CHARACTER
	CHARACTER FOR BIT DATA
	CLOB (limited to 5 megabytes)
	DATE
	DECIMAL
	DOUBLE
	FLOAT
	INTEGER
	REAL
	TIME
	TIMESTAMP
	SMALLINT
	VARCHAR
	VARCHAR FOR BIT DATA

Data types supported by the XML wrapper

The following table lists the DB2 data types that the XML wrapper supports.

Table 77. XML data types that map to DB2 data types for the XML wrapper

XML data types	DB2 data type
	CHARACTER
	CHARACTER FOR BIT DATA
	CLOB (maximum length is 5 megabytes)
	DATE
	DECIMAL
	DOUBLE
	FLOAT
	INTEGER
	REAL
	SMALLINT
	VARCHAR
	VARCHAR FOR BIT DATA

Accessing information about IBM

IBM has several methods for you to learn about products and services.

You can find the latest information on the Web at www.ibm.com/software/data/integration/db2ii/support.html:

- Product documentation in PDF and online information centers
- Product downloads and fix packs
- Release notes and other support documentation
- Web resources, such as white papers and IBM Redbooks™
- Newsgroups and user groups
- Book orders

To access product documentation, go to this site:

publib.boulder.ibm.com/infocenter/db2help/topic/

You can order IBM publications online or through your local IBM representative.

- To order publications online, go to the IBM Publications Center at www.ibm.com/shop/publications/order.
- To order publications by telephone in the United States, call 1-800-879-2755.

To find your local IBM representative, go to the IBM Directory of Worldwide Contacts at www.ibm.com/planetwide.

Contacting IBM

You can contact IBM by telephone for customer support, software services, and general information.

Customer support

To contact IBM customer service in the United States or Canada, call 1-800-IBM-SERV (1-800-426-7378).

Software services

To learn about available service options, call one of the following numbers:

- In the United States: 1-888-426-4343
- In Canada: 1-800-465-9600

General information

To find general information in the United States, call 1-800-IBM-CALL (1-800-426-2255).

Go to www.ibm.com for a list of numbers outside of the United States.

Accessible documentation

Documentation is provided in XHTML format, which is viewable in most Web browsers.

XHTML allows you to view documentation according to the display preferences that you set in your browser. It also allows you to use screen readers and other assistive technologies.

Syntax diagrams are provided in dotted decimal format. This format is available only if you are accessing the online documentation using a screen reader.

Providing comments on the documentation

Please send any comments that you have about this information or other documentation.

Your feedback helps IBM to provide quality information. You can use any of the following methods to provide comments:

- Send your comments using the online readers' comment form at www.ibm.com/software/awdtools/rcf/.
- Send your comments by e-mail to comments@us.ibm.com. Include the name of the product, the version number of the product, and the name and part number of the information (if applicable). If you are commenting on specific text, please include the location of the text (for example, a title, a table number, or a page number).

Notices and trademarks

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785 U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing 2-31 Roppongi 3-chome, Minato-ku
Tokyo 106-0032, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003 U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not

been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. _enter the year or years_. All rights reserved.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM trademarks and certain non-IBM trademarks are marked at their first occurrence in this document.

See http://www.ibm.com/legal/copytrade.shtml for information about IBM trademarks.

The following terms are trademarks or registered trademarks of other companies:

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel®, Intel Inside® (logos), MMX and Pentium® are trademarks of Intel Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product or service names might be trademarks or service marks of others.

Index

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Printed in USA

SC19-1020-00