

IBM® DB2® 通用数据库



数据仓库中心管理指南

版本 7

IBM® DB2® 通用数据库



数据仓库中心管理指南

版本 7

在使用本资料及其支持的产品之前，请阅读第403页的『声明』中的一般信息。

本文档包含 IBM 的专利信息。它在许可证协议下提供，并受版权法保护。本出版物包含的信息不包括任何产品保证，且本手册提供的任何声明不应作如此解释。

通过您当地的 IBM 代表或 IBM 分部可订购出版物，或者，通过致电 1-800-879-2755（在美国）或 1-800-IBM-4YOU（在加拿大）来订购出版物。

当您发送信息给 IBM 后，即授予 IBM 非专有权，IBM 对于您所提供的任何信息，有权利以任何它认为适当的方式使用或分发，而不必对您负任何责任。

© Copyright International Business Machines Corporation 1996, 2001. All rights reserved.

目录

关于本书	ix	定义对 DB2 DRDA 源数据库的特权	47
本书读者	ix	设置 DB2 Connect 网关节点	48
先决条件出版物	ix	建立与 DB2 DRDA 源数据库的联网	48
第1章 关于数据入库	1	从 AS/400 代理访问远程数据库的联网需求	49
什么是数据入库?	1	从 AS/400 代理建立与本地和远程数据库的联网	49
DB2 通用数据库中的数据入库	1	从 OS/390 代理的远程数据库的联网需求	50
主题区	2	使用 DataJoiner	51
仓库源	2	定义 DB2 仓库源	51
仓库目标	2	访问远程文件	54
仓库代理和代理站点	2	通过 Windows NT 或 Windows 2000 文件服务器访问文件	54
过程和步骤	2	用 NFS 访问文件	55
入库任务	4	用 FTP 访问文件	55
第2章 设置仓库	7	使用“使用 FTP 复制文件”访问数据文件	56
启动“数据仓库中心”	7	定义文件源	57
启动仓库服务器和日志记录器	7	访问非 DB2 数据库仓库源	60
启动仓库代理守护进程	8	在 Windows NT 上设置非 DB2 数据库仓库源	60
停止仓库代理守护进程	10	在 AIX 或 Solaris 操作环境中设置非 DB2 数据库仓库源	85
启动“数据仓库中心”管理界面	11	在 OS/2 上设置非 DB2 数据库仓库源	91
定义代理站点	13	在“数据仓库中心”中定义非 DB2 数据库仓库源	95
代理站点配置	13	打开“仓库源”笔记本	95
仓库服务器和仓库代理的联网需求	16	添加有关仓库源的信息	96
在 OS/390 上配置 TCP/IP	16	指定代理站点	96
在 OS/390 上更新环境变量	17	指定数据库信息	96
对“数据仓库中心”定义代理站点	17	导入源表和视图	97
“数据仓库中心”安全性	19	为仓库组授予访问权	98
定义仓库用户	21	指定重试计数	98
定义仓库组	23	定义供 DataJoiner 使用的仓库源	98
第3章 设置仓库源	25	第4章 设置对仓库的访问权	101
选择“数据仓库中心”数据源	25	设置 DB2 通用数据库仓库	101
Windows NT 或 Windows 2000	27	定义对 DB2 通用数据库仓库的特权	101
AIX	33	建立与 DB2 通用数据库仓库的联网	101
Solaris 操作环境	38	设置 DB2 AS/400 版仓库	102
OS/2	42	使用 DB2 Connect	102
设置对“DB2 通用数据库”源数据库的访问权	45	使用 CA/400	103
定义对“DB2 通用数据库”源数据库的特权	46	设置 DB2 OS/390 版仓库	105
建立与“DB2 通用数据库”源数据库的联网	46		
设置对 DB2 DRDA 源数据库的访问权	46		
先决条件产品	47		

定义对 DB2 OS/390 版仓库的特权	105	定义“提交 OS/390 JCL 作业流”	
建立与 DB2 OS/390 版仓库数据库的联网	105	(VWPMVS) 程序的值	176
创建“数据仓库中心”状态表	106	在“数据仓库中心”中设置复制	178
对“数据仓库中心”定义仓库	106	复制表	179
设置 DB2 EEE 仓库	106	在“数据仓库中心”中定义复制源	180
定义对 DB2 EEE 数据库的特权	106	定义用户复制、时间点或基本聚集复制步骤	181
建立与 DB2 EEE 数据库的联网	107	定义更改聚集复制步骤	183
对“数据仓库中心”定义 DB2 EEE 数据		定义分级表复制步骤	186
库	107	从 SAP R/3 系统中抽取数据	188
定义供 DataJoiner 使用的仓库	107	定义 SAP 源	188
用 DataJoiner 创建目标表	109	定义每个源业务对象的特性	189
移动目标表	110	定义 SAP 步骤	189
在远程数据库中创建或更新表	110	从 WebSphere Site Analyzer 数据库中抽取数	
定义仓库目标	111	据	190
定义仓库目标特性	111	定义 WebSphere Site Analyzer 源	190
定义主键	116	定义 Web 流量轮询步骤	191
定义外键	116		
在仓库数据库中创建“数据仓库中心”状态表	117		
第5章 定义和运行过程	119	第7章 转换数据	193
定义主题区	120	连接源列	193
定义过程	120	转换代码	195
打开过程	121	给连接添加空值	196
添加源和目标到过程	121	生成星型连接	198
添加步骤到过程中	122	过滤数据	199
步骤子类型	124	添加计算的列	200
将步骤与源和目标链接	136	转换目标表	202
定义步骤子类型的基本值	138	清理数据	202
运行仓库步骤	142	生成键列	207
开发仓库步骤	142	生成周期数据	209
测试仓库步骤	142	反转数据	212
调度仓库过程	146	旋转数据	214
从“数据仓库中心”外部启动步骤	150	更改日期字段的格式	217
第6章 移动数据	155	第8章 计算统计信息	219
选择和插入数据	155	ANOVA 变换器	219
定义步骤	155	“计算统计信息”变换器	221
增量落实	157	“计算小计”变换器	223
装入和导出数据	158	“X 平方”变换器	226
导出数据	159	“相关”变换器	228
装入数据	161	“滑动平均值”变换器	231
操纵文件	173	“回归”变换器	234
定义“使用 FTP 复制文件”(VWPRCPY)		第9章 清理名称和地址数据	237
程序的值	173	要求	237
定义“运行 FTP 命令文件”程序		Trillium Software System 组件	238
(VWPFFTP) 的值	175	将 Trillium Batch System 和“数据仓库中	
		心”一起使用	238

导入 Trillium 元数据	239	定义 DB2 UDB RUNSTATS 程序的值	276
映射元数据	241	定义 DB2 OS/390 RUNSTATS 程序的值	277
限制	241	在“数据仓库中心”中使用 MQSeries	278
编写 Trillium Batch System JCL 文件	242	创建 MQSeries 消息的视图	278
定义 Trillium Batch System 步骤	242	导入 MQSeries 消息和 XML 元数据	279
使用 Trillium Batch System 用户定义的程序	243		
错误处理	244		
错误返回码	245		
日志文件	246		
第10章 数据建模	247		
什么是 IBM ERwin Metadata Extract			
Program?	247		
软件需求	247		
程序文件	248		
创建标记语言文件	248		
将标记语言文件导入“数据仓库中心”	250		
将标记语言文件导入“信息目录管理器”	251		
故障诊断	252		
ERwin 到 DB2“数据仓库中心”的映射	253		
ERwin 到“信息目录管理器”的映射	254		
第11章 刷新“OLAP 服务器”数据库	257		
定义“OLAP 服务器: 缺省			
calc”(ESSCALC1) 仓库程序的值	258		
定义“OLAP 服务器: 使用计算规则的 Calc			
”(ESSCALC2) 仓库程序的值	259		
定义“OLAP 服务器: 自由文本数据装入”			
(ESSDATA1) 仓库程序的值	260		
定义“OLAP 服务器: 使用装入规则从文件装			
入数据”(ESSDATA2) 仓库程序的值	261		
定义“OLAP 服务器: 使用装入规则从 SQL			
表装入数据”(ESSDATA2) 仓库程序的值	262		
定义“OLAP 服务器: 不使用装入规则从文件			
装入数据”(ESSDATA4) 仓库程序的值	263		
定义“OLAP 服务器: 从文件更新轮廓”			
(ESSOTL1) 仓库程序的值	264		
定义“OLAP 服务器: 从 SQL 表更新轮廓”			
(ESSOTL2) 仓库程序的值	266		
第12章 维护仓库数据库	269		
重组数据	269		
定义 DB2 通用数据库 REORG 程序的值	269		
定义 DB2 OS/390 版 REORG 程序的值	270		
定义“DB2 OS/390 实用程序”程序的值	272		
更新系统编目统计信息	276		
		定义 DB2 UDB RUNSTATS 程序的值	276
		定义 DB2 OS/390 RUNSTATS 程序的值	277
		在“数据仓库中心”中使用 MQSeries	278
		创建 MQSeries 消息的视图	278
		导入 MQSeries 消息和 XML 元数据	279
		第13章 用用户定义程序扩展“数据仓库中	283
		心”	283
		定义用户定义程序	283
		定义程序组	283
		指定要使用的程序	284
		指定要使用的代理站点	285
		指定参数	286
		使用用户定义程序步骤	287
		测试使用用户定义程序的步骤	288
		编写自己的程序以配合“数据仓库中心”使用	289
		传送参数	289
		返回状态信息	290
		第14章 维护“数据仓库中心”	295
		备份“数据仓库中心”	295
		扩充仓库	295
		导出和导入“数据仓库中心”元数据	296
		将元数据导出到另一个“数据仓库中心”	296
		导入元数据	298
		发布“数据仓库中心”元数据	301
		计划发布元数据	301
		选择要发布的元数据	302
		更新发布的元数据	304
		调度发布的元数据的定期更新	305
		查看发布更新的日志文件	305
		将删除的对象传播到信息目录	306
		更改活动的仓库控制数据库	306
		在安装期间初始化仓库控制数据库	307
		配置“数据仓库中心”	307
		更改“数据仓库中心”特性	308
		第15章 在“数据仓库中心”中创建星型模式	311
		在“数据仓库中心”中设计仓库模式	312
		定义仓库模式	312
		将表和视图添加至仓库模式	313
		自动连接表	314
		在非键列之间添加连接关系	314
		将仓库模式导出到 DB2 OLAP Integration	
		Server	315

在 DB2 OLAP Integration Server 中使用仓库模式	317
在 DB2 OLAP Integration Server 中创建轮廓和为多维立方装入数据	318
将元数据导出到“数据仓库中心”	320
从“数据仓库中心”中装入多维立方	321
创建填充多维立方的调度表	322
发布关于您的仓库模式的元数据	323

附录A. “数据仓库中心”记录和跟踪数据	325
基本记录功能	325
组件跟踪数据	326
仓库程序和变换器	327
启动错误跟踪文件	328

附录B. 元数据映射	329
信息目录管理器和“数据仓库中心”之间的元数据映射	329
信息目录管理器和 OLAP 服务器之间的元数据映射	338
DB2 OLAP Integration Server 和“数据仓库中心”之间的元数据映射	339

附录C. “数据仓库中心”如何迁移版本 5.2	
定义	341
Visual Warehouse 对象和“数据仓库中心”对象之间的映射	341
主题	341
仓库	341
业务视图	341
Visual Warehouse 程序	342
安全性更改	342

附录D. 定义 Visual Warehouse 5.2 兼容程序的值	343
定义“Visual Warehouse 5.2 DB2 UDB 数据导出”(VWPEXPT1) 程序的值	343
定义“Visual Warehouse DB2 UDB 数据插入装入”(VWPLOADI) 程序的值	344
定义“Visual Warehouse DB2 UDB 数据置换装入”(VWPLOADR) 程序的值	346
定义 Visual Warehouse 5.2 DB2 UDB REORG (VWPREORG) 程序的值	348
定义 Visual Warehouse 5.2 DB2 UDB RUNSTATS (VWPSTATS) 程序的值	349

定义“Visual Warehouse 5.2 将文本文件装入 DB2 UDB EEE”(VWPLDPR) 程序的值(仅 AIX)	350
---	-----

附录E. 数据入库样本	353
安装样本	353
创建样本数据库	353
创建仓库数据库	354
使用样本内容	355
查看样本数据	355
查看和修改样本元数据	356
提升步骤	359
运行步骤	360
查看样本仓库数据	361

附录F. 在“数据仓库中心”中使用 Classic Connect	363
什么是 Classic Connect?	363
它起什么作用?	363
它可访问哪些数据源?	363
如何使用它?	364
它有哪些组件?	364
建立环境	372
硬件和软件需求	372
安装和配置的先决条件产品	372
配置 OS/390 和 Windows NT 之间的通信协议	374
通信选项	374
配置 TCP/IP 通信协议	375
配置 LU 6.2 通信协议	378
配置 Windows NT 客户机	382
安装“Classic Connect 驱动程序”组件	382
安装 CROSS ACCESS ODBC 驱动程序	382
配置数据源	382
从“Visual Warehouse 主机适配器”迁移到 Classic Connect	393

附录G. “数据仓库中心”环境结构	395
“数据仓库中心”环境变量	395
“数据仓库中心”注册表更新	395
“数据仓库中心”配置文件	395

附录H. Microsoft OLE DB 和“数据事务服务”支持	397
软件需求	397
限制	397
为 OLE DB 表函数创建视图	398

为 DTS 数据包创建视图	398	商标	405
附录I. 创建 i2 步骤	399	索引	407
导入 i2 源的元数据	399	与 IBM 联系	417
创建 i2 步骤	400	产品信息	417
声明	403		

关于本书

本书描述使用“IBM® 数据仓库中心”构建和维护仓库所需的步骤。仓库是一个数据库，它包含从您的操作数据源中抽取和转换的信息数据。

本书读者

本书面向使用“数据仓库中心”构建仓库的管理员和仓库设计人员。读者应该对关系数据库管理系统有一个基本的了解、有一些 SQL 知识且能联系上可对各种数据源建立联网的专家。

要使用统计变换器，应该熟悉统计技术。

先决条件出版物

在阅读本书之前，请阅读针对您的操作系统的《DB2 通用数据库快速入门》以安装“数据仓库中心”。若已有“DB2 仓库管理器”，请阅读《DB2 仓库管理器安装指南》以安装代理和变换器。

除本书外，还需要下列各项的文档：

- Windows NT® 操作系统
- 计划用作源和目标的数据库
- 计划使用的联网软件

第1章 关于数据入库

DB2 通用数据库® 提供了“数据仓库中心”，一个使数据仓库处理自动化的组件。可使用“数据仓库中心”定义移动和转换仓库数据的过程。然后，可使用“数据仓库中心”来调度、维护和监控这些过程。

本章概述了数据入库和数据入库任务。有关入库的更多信息详情和有关使用“数据仓库中心”的信息，参见“数据仓库中心”联机帮助。

什么是数据入库？

包含操作数据（运行公司日常事务的数据）的系统中有对商务分析员有用的信息。例如，分析员可以使用在年度的哪段时间在哪些地区销售了哪些产品之类的信息，来查找异常情况或制定将来的销售计划。

但是，当分析员直接访问操作数据时，可能会产生几个问题：

- 他们可能没有查询操作数据库的专门知识。例如，查询 IMS™ 数据库需要使用专用类型的数据操作语言的应用程序。具有查询操作数据库的专门知识的程序员通常全职从事维护数据库及其应用程序的工作。
- 性能对于许多操作数据库，如银行的数据库是至关重要的。系统不能处理用户进行特定查询的情况。
- 操作数据通常不是商务分析员所使用的最佳格式。例如，对于分析员来说，按产品、地区和季节汇总的销售数据比原始数据要有用得多。

数据入库解决了这些问题。在数据入库中，可创建信息数据的仓库。信息数据是从操作数据中抽取并进行转换以用于决策的数据。例如，数据入库工具可以从操作数据库复制所有销售数据、清理该数据、执行计算以汇总该数据、然后将汇总的数据写入独立于操作数据的数据库的目标中。用户可以查询该独立数据库（仓库），而不影响操作数据库。

DB2 通用数据库中的数据入库

以下各节描述将用于创建和维护数据仓库的对象。

主题区

主题区标识与业务的逻辑区相关的过程，并将过程分组。例如，若正构建市场营销和销售数据的仓库，则定义“销售”主题区和“市场营销”主题区。然后在“销售”主题区下添加与销售相关的过程。类似地，在“市场营销”主题区下添加与市场营销数据相关的定义。

仓库源

仓库源标识将为仓库提供数据的表和文件。“数据仓库中心”使用仓库源中的规范来访问数据。这些源几乎可以是与您的网络联网的任何关系或非关系源（表、视图或文件）。

仓库目标

仓库目标是包含已转换的数据的数据库表或文件。类似于仓库源，用户可以使用仓库目标向其他仓库目标提供数据。中央仓库可以向部门的服务器提供数据，或仓库中的主事实表可以向总结表提供数据。

仓库代理和代理站点

仓库代理管理数据源和目标仓库之间的数据传送。AIX[®]、AS/400[®]、OS/2[®]、OS/390[®] 和 Windows NT 操作系统及 Solaris 操作环境上都提供仓库代理。代理使用“开放式数据库联网” (ODBC) 驱动程序或 DB2 CLI 来与不同的数据库通信。

有几种代理可以处理源与目标仓库之间的数据传送。使用的代理计数取决于现存的联网配置和计划移动至仓库的数据量。若需要同一代理的多个进程同时运行，则会生成该代理的附加实例。

代理可以是本地或远程的。本地仓库代理是安装在与仓库服务器相同的工作站上的代理。远程仓库代理是安装在与该仓库服务器联网的另一工作站上的代理。

代理站点是安装代理软件的工作站的逻辑名。代理站点名与 TCP/IP 主机名不同。单个工作站只能有一个 TCP/IP 主机名。然而，可以在单个工作站上定义多个代理站点。逻辑名标识每个代理站点。

缺省代理站点，名为“缺省 DWC 代理站点”，是在仓库控制数据库初始化期间，“数据仓库中心”定义的位于 Windows NT 上的本地代理。

过程和步骤

过程包含为特定仓库用途执行数据转换和移动的一系列步骤。通常，过程将源数据移动至仓库。然后，聚集并汇总该数据库以供仓库使用。过程可以产生单个文本表和一组总结表。过程还可以执行某种特定类型的数据转换。

步骤是仓库内单个操作的定义。通过使用 SQL 语句或调用程序，步骤定义如何移动和转换数据。当运行步骤时，可能在仓库源和仓库目标之间进行数据传送，或对该数据进行任何转换。

步骤是“数据仓库中心”中的逻辑实体，它定义：

- 与其源数据的链接。
- 输出表或文件的定义和与输出表或文件的链接。
- 填充输出表或文件的机制（SQL 语句或程序）和定义。
- 填充输出表或文件所依据的处理选项和调度表。

假设想要“数据仓库中心”执行下列任务：

1. 从不同的数据库抽取数据。
2. 将数据转换为单一格式。
3. 将数据写入数据仓库中的表中。

应创建包含多个步骤的过程。每个步骤执行单独的任务，如从数据库抽取数据或将其转换为正确的格式。可能需要创建多个步骤，来彻底转换和格式化数据并将它放入最终的表中。

当步骤或过程运行时，它可能以下列方式影响目标：

- 用新数据置换仓库目标中的所有数据
- 追加新数据到现存数据
- 追加数据的一个单独版本

有关详情，参见第141页的『定义处理选项』。

可以按需求运行步骤，或者可以调度步骤在设置的时间运行。可以调度步骤只运行一次，或者调度它反复运行，如在每个星期五运行。还可以调度步骤按顺序运行，以便步骤紧接着上一个步骤运行。可以调度步骤根据另一个步骤的完成情况（成功或失败）来运行。若调度过程，该过程中的第一个步骤会在调度的时间运行。

以下各节描述将在“数据仓库中心”中发现的各种步骤。有关步骤的详情，参见第119页的『第5章 定义和运行过程』和“数据仓库中心”的联机帮助。

SQL 步骤

SQL 步骤使用 SQL SELECT 语句从仓库源中抽取数据，并生成 INSERT 语句以将数据插入到仓库目标表中。

程序步骤

程序步骤有多种类型：DB2 AS/400 版程序、DB2 OS/390 版程序、DB2 UDB 版程序、Visual Warehouse 5.2 DB2 程序、“OLAP 服务器”程序、“文件”程序和“复制”。这些步骤运行预定义的程序和实用程序。

变换器步骤

变换器步骤是存储过程和用户定义函数，它们指定可用来转换数据的统计变换器或仓库变换器。可使用变换器清除、反转和旋转数据；生成主键和周期表；以及计算各种统计信息。

在变换器步骤中，指定统计变换器或仓库变换器中的一个。当运行该过程时，变换器步骤将数据写入一个或多个仓库目标。

用户定义程序步骤

用户定义程序步骤是“数据仓库中心”内的逻辑实体，它表示想要“数据仓库中心”启动的公司特定转换。因为每个公司都有独特的数据转换需求，所以公司可以选择编写他们自己的程序步骤，或使用一些工具，象 ETI 或 Vality 提供的那些工具。

例如，可以编写执行下列功能的用户定义程序：

1. 从表中导出数据。
2. 操纵该数据。
3. 将该数据写入临时输出资源或仓库目标。

入库任务

创建数据仓库涉及下列任务：

- 标识源数据（或操作数据）并定义它用作仓库源。
- 创建数据库以用作仓库并定义仓库目标。
- 为将在仓库中定义的过程组定义主题区。
- 通过在过程中定义步骤，指定如何移动源数据和将它转换为用于仓库数据库的格式。
- 测试定义的步骤并调度它们自动运行。
- 通过定义安全性和监控数据库的使用情况管理仓库。
- 若拥有“DB2 仓库管理器”程序包，则创建仓库中的数据的信息目录。信息目录是包含公司元数据的数据库。公司元数据帮助用户在该组织中标识和查找可供使用的数据和信息。仓库的用户可以搜索该目录以确定仓库中提供什么数据。
- 为仓库中的数据定义星型模式模型。星型模式是一个专门的设计，由多个维表和一个事实表组成，维表描述业务的各个方面，事实表包含该业务的事实或计

量。例如，对于制造业公司，维表可以是产品、市场和时间。事实表包含有关每个地区按季节订购的产品的事务信息。

通过使用《商业智能教程》、查看《DB2 通用数据库快速导览》以及阅读“数据仓库中心”联机帮助或阅读本书，可了解有关这些任务和其他任务的更多详情。

第2章 设置仓库

要设置仓库，需要登录到“数据仓库中心”，并指定“数据仓库中心”访问源数据库和目标数据库时将使用的代理站点。还需要为将为仓库定义的对象设置安全性。最后，需要设置主题区，它们将包含有关填充仓库所需过程的信息。

本章描述设置仓库所需的步骤，包括：

- 『启动“数据仓库中心”』
- 第13页的『定义代理站点』
- 第19页的『“数据仓库中心”安全性』

构建仓库的过程是一个迭代的过程。可以从一组特定的假设开始，构建一个反映那些假设的仓库。最后可能想要更改这些假设以及它们反映的仓库。因为“数据仓库中心”易于安装和使用，所以可进行更改而不必重建仓库。可以迭代评估您的假设，并可对仓库进行任意多次更改。“数据仓库中心”会随着您进行调整。

启动“数据仓库中心”

要启动“数据仓库中心”，需要启动“数据仓库中心”的组件。有些组件会自动启动，但有些则需要人工启动。

安装仓库服务器和日志记录器之后，在启动 Windows NT 时它们会自动启动。仓库代理可自动启动或人工启动。可从“DB2 控制中心”人工打开“数据仓库中心”管理界面。

启动仓库服务器和日志记录器

仓库服务器和仓库日志记录器作为“Windows NT 服务”运行。要启动它们，必须在初始化仓库控制数据库之后重新启动系统。以后每次启动 Windows NT 时，仓库服务器和日志记录器将自动启动，除非将它们更改成人工操作或停止 DB2 服务器。如果停止 DB2 服务器，将丢失与本地和远程数据库的连接。在停止然后重新启动 DB2 服务器以复原连接后，必须停止并重新启动仓库服务器和日志记录器。

要人工启动仓库服务器和日志记录器，可使用下列选项之一：

- 在 Windows NT 桌面上：
 1. 单击**开始** → **设置** → **控制面板** → **服务**。
 2. 向下滚动列表，直到找到**仓库服务器**。单击**启动**，并单击**确定**。

设置仓库

- 在 DOS 命令提示符处输入如下命令:

```
NET START VWKERNEL
```

启动仓库代理守护进程

仓库服务器可用作仓库代理，称作本地代理。本地代理定义为所有“数据仓库中心”活动的缺省仓库代理。当仓库服务器启动时，本地代理自动启动。

若安装了 AIX 或“Solaris 操作环境”仓库代理，仓库代理守护进程会自动启动。

若正使用远程 Windows NT 仓库代理或 OS/2 仓库代理，可在安装期间选择是自动还是人工启动仓库代理守护进程。

若正使用 AS/400 或 OS/390 仓库代理，必须人工启动它。

有关安装仓库代理的更多信息，参见《DB2 仓库管理器安装指南》。

人工启动 Windows NT 仓库代理守护进程

若安装了不是本地代理的 Windows NT 仓库代理，可象启动任何 Windows NT 服务那样人工启动该仓库代理守护进程:

1. 单击**开始** → **设置** → **控制面板** → **服务**。
2. 向下滚动列表，直到找到**仓库代理守护进程**。进行任何想要的更改，并单击**确定**。

人工启动 OS/2 仓库代理守护进程

要人工启动 OS/2 仓库代理，可在 OS/2 命令提示符处输入如下命令:

```
START  
/B "IBM Visual Warehouse-NT OS/2 Daemon" d:\directory\VWDAEMON.EXE
```

其中 d:\directory 是安装仓库代理和仓库代理守护进程的驱动器和目录。

人工启动 AS/400 仓库代理守护进程

安装 AS/400 仓库代理之后，需要启动仓库代理守护进程。

启动代理守护进程的用户简要表应具有 *JOBCTL 权限。

要启动代理守护进程，在 AS/400 命令提示符处输入 STRVWD。STRVWD 命令在 QIWH 子系统中启动 QIWH/IWHVWD (仓库代理守护进程)。这会导致该仓库代理守护进程启动的所有仓库代理过程在 QIWH 子系统中启动。

要验证仓库代理守护进程是否已启动:

1. 在 AS/400 命令提示符处输入 WRKACTJOB。

2. 查找作业 VWD 和 IWH4MSGQ。若显示了这些作业，则仓库代理守护进程已启动。

仓库代理守护进程以后台作业方式运行。

有时可能想要验证您启动的 AS/400 仓库代理守护进程是否仍在运行。例如，您要在工作周使用 AS/400 仓库代理，然后回家过周末。下周一回来时，在启动新的代理进程之前，应验证该仓库代理守护进程是否是活动的。

要验证 AS/400 仓库代理守护进程是否活动：

1. 在 AS/400 命令提示符处，输入 WRKACTJOB。显示活动的作业。
2. 查找与启动仓库代理守护进程时所用的用户标识相关的函数 PGM-IWHVWD。若未显示该函数，表示仓库代理不活动。

启动 OS/390 仓库代理守护进程

为 OS/390 仓库代理配置完系统之后，需要启动仓库代理守护进程。OS/390 代理和 OS/390 代理守护进程都在“UNIX[®] 系统服务” (USS) 平台上运行。

要在前台启动仓库代理守护进程：

1. 通过 OS/390 主机名和 USS 端口 telnet 至 OS/390 上的 USS。
2. 在命令行上输入 vwd 以在前台启动代理守护进程。

要在后台启动仓库代理守护进程，可在“UNIX 系统服务” shell 中的命令行上输入 vwd>[/usr/lpp/DWC/logs/vwd.log] 2>&1 &。其中 [/usr/lpp/DWC/logs/vwd.log] 是您要发送守护进程消息的目录和路径名。

要从 UNIX shell 验证仓库代理守护进程是否正在运行，可在 UNIX shell 命令行上输入 ps -e | grep vwd。

若仓库代理守护进程正在运行，且您有权查看该任务，将返回类似于以下消息的消息：

```
$ ps -ef | grep vwd
MVSUSR2      198   16777537  - 13:13:22 tty0013  0:00 grep vwd
MVSUSR2     16777446  16777538  - 09:57:21 tty0002  0:00 vwd
```

若仓库代理守护进程没有运行，或您无权查看该任务，将返回类似于以下消息的消息：

```
$ ps -ef | grep vwd
MVSUSR2      198   16777537  - 13:13:22 tty0013  0:00 grep vwd
```

要从 OS/390 控制台验证仓库代理守护进程是否正在运行，可在 OS/390 命令提示符处输入 D OMVS,A=ALL。

设置仓库

若仓库代理守护进程正在运行，在返回的消息中将显示带 `vwd` 字符串的任务。显示类似于以下示例的消息：

```
D OMVS,A=ALL
BPX0040I 13.16.15 DISPLAY OMVS 156
OMVS      000E ACTIVE          OMVS=(00)
USER      JOBNAME ASID        PID        PPID STATE   START      CT_SECS
MVSUSR2  MVSUSR24 00C5      16777446   16777538 HRI    09.57.20    .769
  LATCHWAITPID=          0 CMD=vwd
```

验证仓库服务器和仓库代理之间的通信

“数据仓库中心”使用 **TCP/IP** 与远程代理站点通信。为使这种通信能够进行，仓库服务器必须能识别代理站点的全限定主机名。代理站点也必须能识别仓库服务器的全限定主机名。

要验证一个站点是否能识别另一个站点的全限定主机名，可从命令提示符处使用 **ping** 命令。

例如，有一个仓库代理站点的全限定主机名为 `abc.xyz.commerce.com`。要验证仓库服务器是否能识别该代理站点的全限定主机名，可从 **DOS** 命令提示符处输入：

```
ping abc.xyz.commerce.com
```

确保既验证了从代理站点到仓库服务器工作站的通信，又验证了其反向通信。

停止仓库代理守护进程

有时可能需要停止仓库代理守护进程，例如，需要在 **Windows NT** 上更改仓库代理和仓库代理守护进程的环境变量时。

停止 **Windows NT** 仓库代理守护进程

要停止 **Windows NT** 仓库代理守护进程：

1. 单击**开始** → **设置** → **控制面板** → **服务**。
2. 选择**仓库代理守护进程**。
3. 单击**停止**。
仓库代理守护进程停止。
4. 单击**确定**。

停止 **AIX** 或“**Solaris** 操作环境”仓库代理守护进程

需要停止 **AIX** 或“**Solaris** 操作环境”仓库代理守护进程的唯一情况，是想要更改仓库代理及其相应仓库代理守护进程的环境变量。

要成功更改仓库代理中的一个及其相应仓库代理守护进程的环境变量：

1. 通过编辑 `IWH.ENVIRONMENT` 文件，更改仓库代理和仓库代理守护进程的环境变量。
2. 显示仓库代理守护进程的进程标识。在 AIX 或 “Solaris 操作环境” 命令行上，输入：

```
ps -ef|grep vwd
```

以显示进程标识。

3. 停止仓库代理守护进程。在命令行上输入：

```
kill processid
```

其中 `processid` 是步骤 2 中显示的进程标识。

缺省情况下，当仓库代理守护进程停止时，它立即重新启动。当仓库代理守护进程重新启动时，它使用新的环境变量。

停止 AS/400 仓库代理守护进程

有时可能需要停止 AS/400 仓库代理守护进程。

要停止仓库代理守护进程，可在 AS/400 命令提示符下输入 `ENDVWD`。

当输入此命令时，仓库代理守护进程停止或显示一个作业列表。若显示作业列表，则结束状态为 `ACTIVE` 的作业。

停止 OS/390 仓库代理守护进程

若需要停止代理守护进程：

1. 确定守护进程的进程标识。

通过在命令行上输入 `ps -ef | grep vwd` 或 `D OMVS,A=ALL`，可确定守护进程的进程标识。

2. 停止守护进程。

通过在命令行上输入 `kill [-9] pid`，可停止守护进程，其中 `pid` 是守护进程的进程标识。

启动“数据仓库中心”管理界面

要启动“数据仓库中心”管理界面：

1. 启动“DB2 控制中心”：

- 在 Windows NT 上，单击开始 → 程序 → IBM DB2 → 控制中心。
- 在 AIX 或 “Solaris 操作环境” 中，输入如下命令：

```
db2jstrt 6790
db2cc 6790b
```

设置仓库

- 在“DB2 控制中心”窗口中单击**工具** → **数据仓库中心**。“数据仓库中心登录”窗口打开。
- 若是第一次登录，单击**高级**。
“高级”窗口打开。
- 在**控制数据库**字段中，为仓库控制数据库的名称输入 ODBC 系统数据集名（系统 DSN）。
- 在**服务器主机名**字段中，输入安装仓库服务器的工作站的 TCP/IP 主机名。



- 单击**确定**。
“高级”窗口关闭。
下次登录时，“数据仓库中心”将使用您在“高级”窗口中指定的设置。
- 在“登录”窗口的**用户标识**字段中，输入有效的“数据仓库中心”用户标识。
若是第一次登录，输入缺省仓库用户标识。
- 在**密码**字段中，输入该用户标识的密码。



- 单击**确定**。
“数据仓库中心登录”窗口关闭。

定义代理站点

“数据仓库中心”使用本地代理作为所有“数据仓库中心”活动的缺省代理。然而，您可能想在与包含仓库服务器的工作站不同的站点上使用仓库代理。需要对“数据仓库中心”定义代理站点，它是安装代理的工作站。“数据仓库中心”使用此定义标识要在其上启动该代理的工作站。

要使用仓库代理，必须完成下列任务：

1. 确定在何处放置仓库代理。
2. 建立与本地和远程数据库的联网。
3. 对于 OS/390 仓库代理，为其配置 TCP/IP 和设置环境变量并为仓库代理守护进程设置权限。
4. 启动代理守护进程。（参见第8页的『启动仓库代理守护进程』。）
5. 验证仓库服务器和仓库代理之间的通信。
6. 对“数据仓库中心”定义该代理站点。

代理站点配置

仓库代理接收来自仓库服务器的 SQL 命令，并将这些命令传送给源数据库或目标数据库。

可用下列三种配置中的一种设置仓库代理：

仓库代理、源和目标位于同一工作站。

仓库代理安装在包含同一数据库中的仓库源表和仓库目标表的系统上，如第14页的图1中所示。

设置仓库

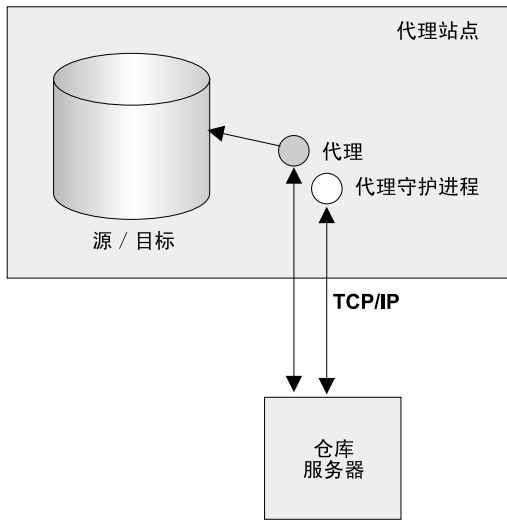


图 1. 仓库代理、仓库源和仓库目标位于同一系统。

仓库服务器也可位于与仓库代理、仓库源和仓库目标相同的系统上。

在此配置中，仓库代理传送从源表抽取数据的 SQL 语句。若有必要，仓库代理转换数据，然后将它写入目标表。

仓库代理和仓库目标位于同一系统。

仓库代理安装在包含仓库目标表的系统上，如第15页的图2中所示。

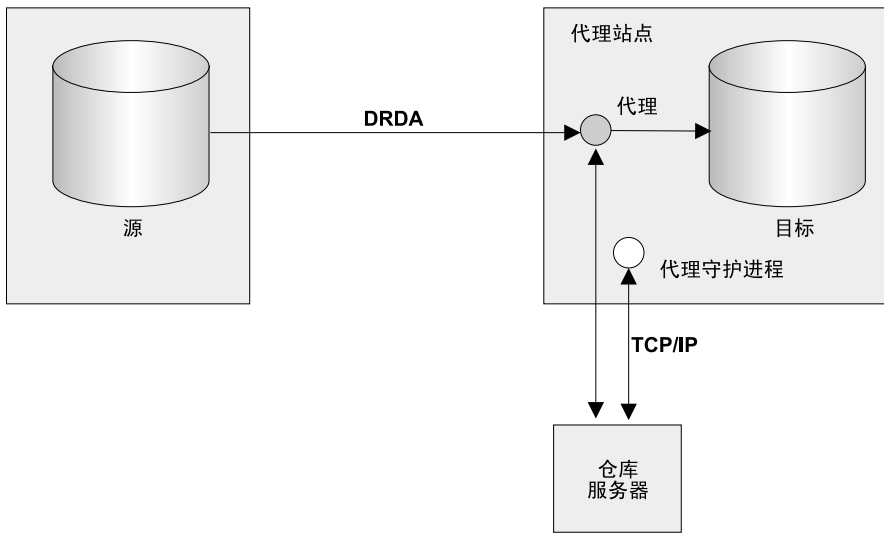


图 2. 仓库代理和仓库目标位于同一系统。

在此配置中，仓库代理传送从远程源抽取数据的 SQL 语句。若有必要，仓库代理转换数据，然后将它写入本地数据库上的目标表。

当源和目标位于不同的系统时，此配置提供最佳性能。

仓库代理和仓库源位于同一系统。

仓库代理安装在包含仓库源的系统中，如第16页的图3中所示。

设置仓库

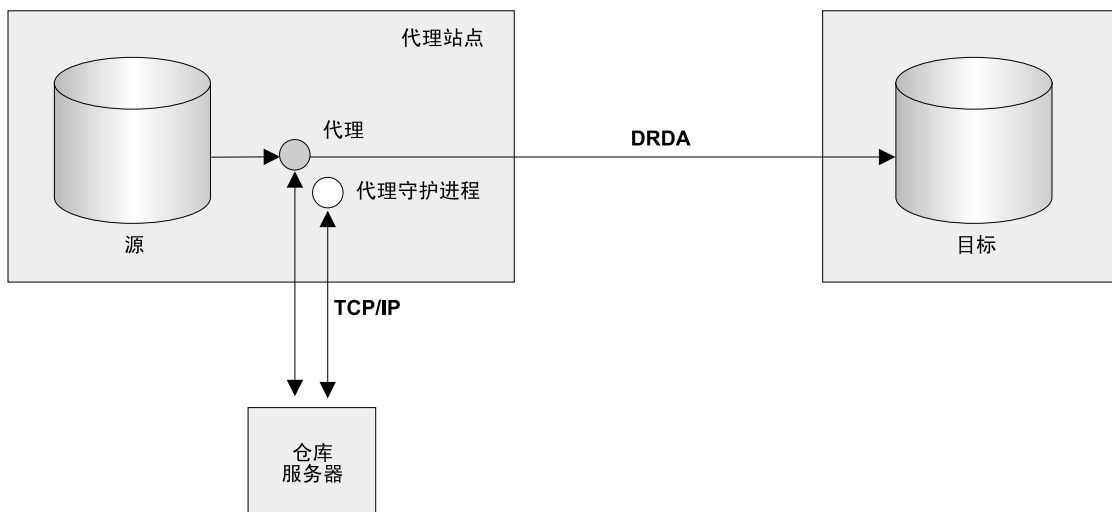


图 3. 仓库代理和仓库源位于同一系统。

在此配置中，仓库代理传送从本地数据库上的仓库源抽取数据的 SQL 语句。若有必要，仓库代理转换数据，然后将它写入远程数据库上的目标表。

对数据设置访问权和确定仓库代理的位置之后，必须为仓库定义安全性。有关定义仓库安全性的信息，参见第19页的『“数据仓库中心”安全性』。

仓库服务器和仓库代理的联网需求

仓库服务器使用 TCP/IP 与仓库代理和仓库代理守护进程通信。为使这种通信能够进行，仓库服务器必须能识别仓库代理的全限定主机名。仓库代理也必须能识别仓库服务器的全限定主机名。

缺省情况下，仓库服务器用端口 11001 给仓库代理守护进程发送消息，用端口 11000 接收响应。仓库日志记录器使用端口 11002。若另一应用程序使用缺省“数据仓库中心”端口号中的一个，可更改“数据仓库中心”使用的端口号。若需要更改端口号，参见 *DB2 Universal Database Troubleshooting Guide*。

在 OS/390 上配置 TCP/IP

要在 OS/390 上配置 TCP/IP:

1. 在 OS/390 环境中转至 /etc/services 文件或 'TCPIP.ETC.SERVICES' 文件。
2. 给文件添加下列服务:

端口名	端口号
-----	-----

vwkernel	11000/tcp
vwd	11001/tcp
vwlogger	11002/tcp

在 OS/390 上更新环境变量

要更新环境变量，将以下变量添加到启动代理守护进程的用户主目录的 .profile 文件中：

变量	注释
export VWS_LOGGING=/u/mydir/logs/	/u/mydir/logs/ 是想要在其中放置代理日志的目录。
export DSNAOINI='DBA1.INSTALIB(DSNAOINI)'	有关 INI 文件的详情，参见 <i>DB2 UDB for OS/390 ODBC Guide and Reference</i> 。
export STEPLIB='DSN710.SDSNLOAD'	DSN710.SDSNLOAD 是 DB2 装入库。

将以下变量添加到 INI 文件：

表 1.

变量	注释
export DSNAOTRC='DBA1.DSNAOTRC'	仅当想要接收 CLI 跟踪时才使用。

对“数据仓库中心”定义代理站点

使用“代理站点”笔记本对“数据仓库中心”定义代理站点。

若正使用与仓库服务器安装在一起的本地代理，不必定义本地代理。它自动定义为缺省代理站点。

若想要代理访问用户对象，如 LAN 驱动器，应指定“数据仓库中心”作为用户过程运行而不是作为系统过程运行。

对“数据仓库中心”定义代理站点不需要任何权限级别。

要对“数据仓库中心”定义代理站点：

1. 打开“代理站点”笔记本：
 - a. 在“数据仓库中心”目录树上，展开**管理**文件夹。
 - b. 右键单击**代理站点**文件夹，并单击**定义**。“代理站点”笔记本打开。

The screenshot shows the 'Define Agent Site' dialog box with the following fields and values:

- Agent Site: Warehouse Sources and Targets | Programs
- Name: San Jose
- Administrator: Adrienne
- Description: This agent site has access to Marketing data in San Jose.
- Notes: (empty)
- Hostname: 9.999.99.99
- Operating system: Windows NT
- Agent module name: (empty)
- Operating system logon:
 - User ID: adrienne
 - Password: (masked with asterisks)
 - Verify password: (masked with asterisks)

Buttons: OK, Cancel, Help

2. 在**名称**字段中，为对“数据仓库中心”定义的代理站点输入名称。此名称最多可为 80 字符长。
3. 可选：在**管理员**字段中，输入负责此代理站点定义的仓库管理员的姓名。
4. 可选：在**描述**字段中，为代理定义输入业务描述。描述最多可以是 254 个字符。
5. 可选：在**注释**字段中，输入可能对访问“代理站点”笔记本的用户有帮助的详细信息。
6. 在**主机名**字段中，输入安装代理的系统或工作站的 IP 地址，或输入机器主机名。以 n.nnn.nn.nnn 形式输入 IP 地址。在**操作系统**字段中，选择运行仓库代理所使用的操作系统。
7. 可选：在**代理模块名**字段中，可按“IBM 软件支持”的指示信息更改仓库代理程序的名称。缺省值为 IWH2AGNT。
8. 在**用户标识**字段中，为代理站点输入有效的用户标识（若需要一个的话）。若不指定用户标识，代理将其假设为仓库服务器（仅适用于缺省代理）或代理守护进程的特征。例如，若仓库代理守护进程作为系统过程运行，则仓库代理也作为系统过程运行。若指定用户标识，仓库代理继承由该用户标识指定的用户特征。

9. 若提供用户标识，在**密码**字段中输入与该用户标识相关的密码。在**验证密码**字段中再次输入该密码。
10. 若已定义仓库源和目标，单击**仓库源和目标**选项卡。在**可用的仓库源和目标**列表中，选择想要代理访问的仓库源，并单击 **>**。要选择多个仓库源，按住 **Ctrl** 键，并单击想要添加的源。然后，单击 **>**。若想要添加列表中的所有项目，单击 **>>**。
11. 可选：若仓库代理将运行仓库程序、变换器或用户定义程序，单击**程序**选项卡。在**可用的程序**列表中，单击仓库代理将运行的程序或变换器，并单击 **>**。要选择多个程序，按住 **Ctrl** 键并单击想要添加的程序。然后，单击 **>**。若想要添加**可用的程序**列表中的所有项目，单击 **>>**。
12. 单击**确定**以对“数据仓库中心”定义代理站点。

若正在代理站点上安装仓库程序，必须对“数据仓库中心”定义该程序，然后才可使用它。有关详情，参见第283页的『定义用户定义程序』。

“数据仓库中心”安全性

因为“数据仓库中心”存储各种数据库和系统的用户标识和密码，所以具有独立于数据库和操作系统安全性的“数据仓库中心”安全性结构。此结构由仓库组和仓库用户组成。用户通过属于某个仓库组来获取对“数据仓库中心”对象的特权和访问权。仓库组是仓库用户和特权的已命名的分组，它是用户执行功能的授权。仓库用户和仓库组不需要与为仓库控制数据库定义的 **DB** 用户和 **DB** 组匹配。

在初始化期间，指定仓库控制数据库的 **ODBC** 名称、有效的 **DB2** 用户标识和密码。“数据仓库中心”授权此用户标识和密码更新仓库控制数据库。在“数据仓库中心”中，此用户标识定义为缺省仓库用户。

提示：对于仓库控制数据库支持的每种操作系统，缺省仓库用户需要不同类型的数据库和操作系统授权。有关详情，参见《*DB2 仓库管理器安装指南*》。

当登录到“数据仓库中心”时，“数据仓库中心”通过比较您的用户标识和定义的仓库用户，来验证您是否有权打开“数据仓库中心”管理界面。

若不想定义安全性，可登录为缺省仓库用户，并访问所有“数据仓库中心”对象和执行所有“数据仓库中心”功能。缺省仓库用户是缺省仓库组的一部分。此仓库组对“数据仓库中心”中定义的所有对象具有访问权，除非从仓库组除去这些对象。

然而，您可能想要不同的用户组对“数据仓库中心”内的对象具有不同的访问权。例如，仓库源和仓库目标包含其相应数据库的用户标识和密码。可能想要限制对包含敏感数据（如人员数据）的仓库源和仓库目标的访问权。

设置仓库

通过对仓库组指定特权，限制用户可执行的操作。在“数据仓库中心”中，可对组指定两种特权：管理特权和操作特权。

管理特权

该仓库组中的用户可定义和更改仓库用户和仓库组、更改“数据仓库中心”特性、导入元数据和在创建对象时定义哪些仓库组对它们具有访问权。

操作特权

该仓库组中的用户可监控调度处理的状态。

（由具有管理特权的用户）将特权指定给组。为了使仓库用户拥有某种特权，他们必须属于拥有该特权的仓库组。

除了特权之外，仓库组还包含组中用户对其具有访问权的对象的列表。可指定对源、目标和过程的访问权。

例如，可定义对应使用“数据仓库中心”的某人的仓库用户。然后可定义有权访问某些仓库源的仓库组，并添加新用户到该新仓库组。该新用户有权访问该组中包括的仓库源。

可给予用户各种授权。可在仓库组中包括任何不同种类的授权。也可将一个仓库用户包括在多个仓库组中。用户所属仓库组的组合是该用户的所有授权。

当用户给“数据仓库中心”定义新对象且没有管理特权时，该用户所属的所有组在缺省情况下将对该新对象具有访问权。他们可指定访问权的组的列表限制为他们所属的组。对象笔记本的“安全性”页将不可用于该用户。

用户可从源访问的表或视图的列表也将由他们的组成员资格限制，以便他们能在他们具有访问权的表和视图中进行选择。另外，用户可通过“数据仓库中心”使用的操作集将由该用户具有的安全性级别限制。例如，若用户不属于对某个对象具有访问权的组，该用户将不能访问该对象的特性。

要获取“数据仓库中心”对象和定义或编辑对象所需的组成员资格之间的关系摘要，参见联机帮助中的『数据仓库中心安全性』。

“数据仓库中心”通过将数据库的用户标识和密码包括为仓库源和仓库目标特性的一部分，来使用数据库管理器的安全性。

图4显示仓库用户、仓库组与仓库数据库的用户标识和密码之间的关系:

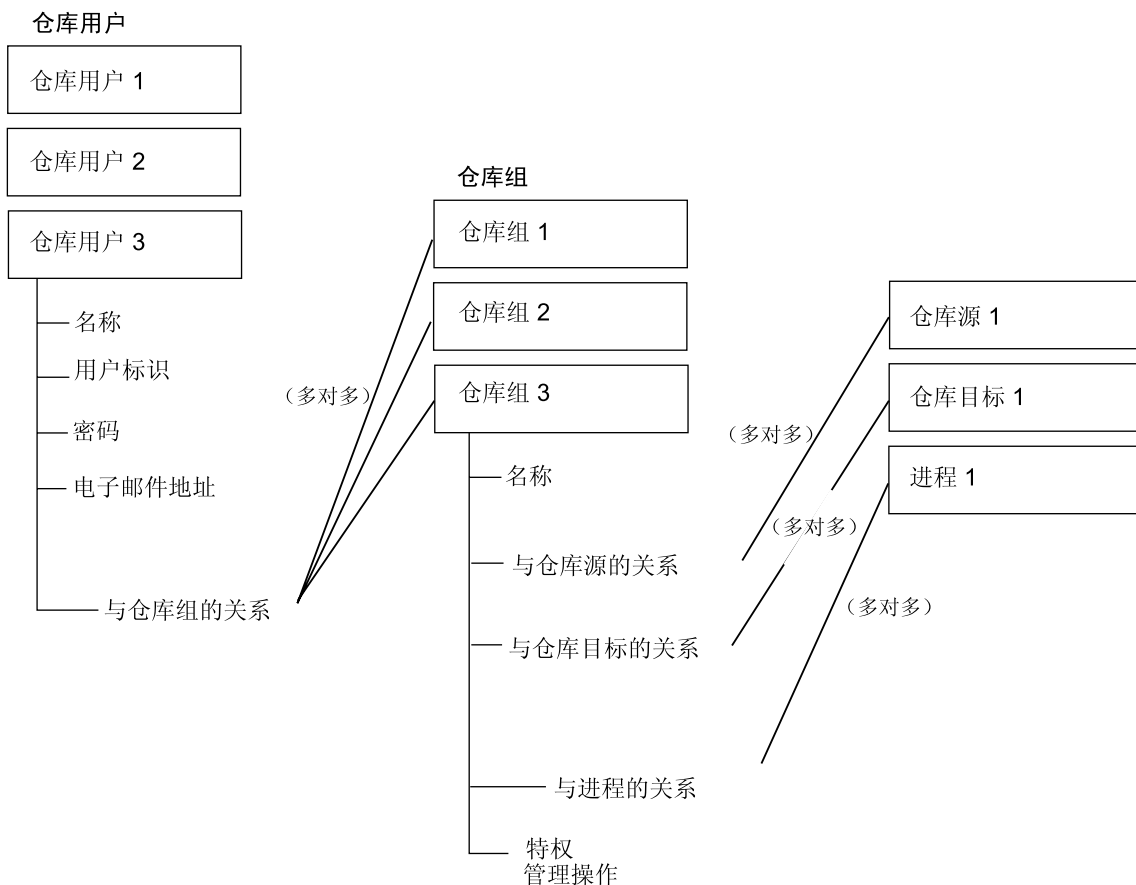


图4. 仓库用户、仓库组与仓库数据库的用户标识和密码之间的关系

定义仓库用户

“数据仓库中心”通过用户标识控制访问权。当用户登录时，将用户标识与“数据仓库中心”中定义的仓库用户进行比较，以确定该用户是否有权访问“数据仓库中心”。通过定义新的仓库用户，可授权另外的用户访问“数据仓库中心”。

新用户的用户标识不需要对操作系统或仓库控制数据库的授权。该用户标识仅存在于“数据仓库中心”内。

要定义仓库用户:

1. 在“数据仓库中心”主窗口的左边，单击**管理**文件夹。

设置仓库

2. 展开仓库用户和组目录树。
3. 右键单击仓库用户文件夹，并单击定义。
“定义仓库用户”笔记本打开。

特性 - Tutorial User

教程用户

仓库用户 | 安全性

名称: 教程用户

管理员: Adrienne

描述: 这是为教程创建的用户。

注释:

“数据仓库中心” 登录

用户标识: tutuser

口令: *****

验证口令: *****

电子邮件通知

电子邮件地址:

活动用户(A)

确定(O) 取消 帮助

4. 在名称字段中，输入用户的业务名。
该名称在“数据仓库中心”内标识用户标识。此名称最多可为 80 字符长，包括空格在内。
5. 在管理员字段中，输入此用户的联系人。
6. 在描述字段中，输入该用户的简短描述。

提示: 可使用描述和注释字段提供关于仓库定义的元数据。然后可在仓库的信息目录中发布此元数据。仓库用户可搜索该元数据来查找包含要查询的信息的仓库。

7. 在用户标识字段中，输入新用户标识。
用户标识不得长于 60 字符，且不能包含空格、短划线或特殊字符（如 @、#、\$、%、>、+、=）。它可包括下划线字符。
8. 在密码字段中，输入密码。然后，在验证密码字段中再次输入该密码。

密码必须至少有六个字符，且不能包含空格、短划线或特殊字符。

提示：可在“定义仓库用户”笔记本的此页上更改密码。

9. 验证已选择**活动用户**复选框。

提示：可清除此复选框以临时取消用户对“数据仓库中心”的访问权，而不必删除用户定义。

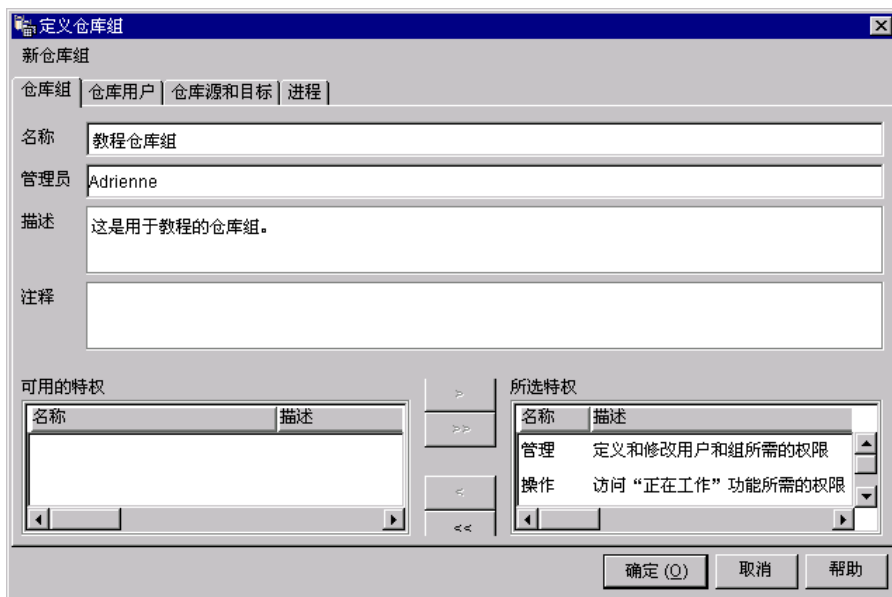
10. 单击**确定**保存仓库用户并关闭笔记本。

定义仓库组

在“数据仓库中心”中，仓库组包括在“数据仓库中心”中执行某些任务和访问对象的授权。要授权一个或多个用户执行任务，必须定义一个仓库组，然后将用户添加到该组。

要定义仓库组：

1. 在“数据仓库中心”主窗口中，右键单击**仓库组**文件夹，并单击**定义**。“仓库组”笔记本打开。



2. 在**名称**字段中，输入新仓库组的名称。
3. 可选：在**管理员**字段中，输入此新仓库组的联系人。
4. 可选：在**描述**字段中，输入新仓库组的简短描述。

设置仓库

5. 可选：在**注释**字段中，输入任何附加信息，管理员可能需要此信息来了解这个仓库组。

6. 在**可用的特权**列表中，选择想要指定给该仓库组的特权，并单击 **>**。选择的特权移动至**选择的特权**列表。

若想要指定**可用的特权**列表中的所有特权，单击 **>>**。

可从下列特权中选择：

管理 该仓库组中的用户可定义和更改仓库用户和仓库组、更改“数据仓库中心”特性、导入元数据和在创建对象时定义哪些仓库组对它们具有访问权。

操作 该仓库组中的用户可监控调度处理的状态。

7. 在“仓库用户”页上，从**可用的用户**列表中选择想要包括在仓库组中的仓库用户，并单击 **>**。选择的仓库用户会移动至**选择的用户**列表中。

若想要包括**可用的用户**列表中现有的所有仓库用户，单击 **>>**。

8. 在“仓库源和目标”页上，从**可用的仓库源和目标**列表中，选择仓库组将具有访问权的仓库源和仓库目标，并单击 **>**。选择的仓库源和目标移动至**选择的仓库源和目标**列表中。

若想要仓库组对**可用的仓库源和目标**列表的所有仓库源和目标具有访问权，单击 **>>**。

提示：可从“定义仓库组”笔记本、“定义仓库源”或“定义仓库目标”笔记本中，授权仓库组访问仓库源和仓库目标。

9. 在“过程”页上，从**可用的过程**列表中选择仓库组将具有访问权的过程，并单击 **>**。选择的过程会移动至**选择的过程**列表中。

若想要仓库组对**可用的过程**列表中的所有过程具有访问权，单击 **>>**。

提示：可从“定义仓库组”笔记本或“定义过程”笔记本中，授权仓库组访问过程。

10. 单击**确定**保存仓库用户组并关闭笔记本。

第3章 设置仓库源

在可创建访问数据源的步骤之前，必须执行下列任务：

- 确定将在仓库中使用哪些数据源。
- 设置已决定在源和仓库代理之间使用的联网。
- 对“数据仓库中心”定义作为仓库源的源。

本章描述“数据仓库中心”处理的数据源的类型，并告诉您如何设置对它们的访问权。

选择“数据仓库中心”数据源

“数据仓库中心”支持很多种关系和非关系数据源。参见表2和表3以获取受支持的源的列表。

表 2. 受支持的 *IBM* 仓库源的版本和发行版级别

源数据库	版本 / 发行版
IMS	5.1
DB2 通用数据库 Windows NT 版	5.2 - 7.1
DB2 通用数据库扩充企业版	5.2 - 7.1
DB2 通用数据库 OS/2 版	5.2 - 7.1
DB2 通用数据库 AS/400 版	3.7 - 5.1
DB2 通用数据库 AIX 版	5.2 - 7.1
DB2 通用数据库 Solaris 操作环境版	5.2 - 7.1
DB2 通用数据库 OS/390 版	5.1 - 7.1
IBM DataJoiner [®]	2.1.1
DB2 VM 版	5.3.4 或更高版本
DB2 VSE 版	7.1

表 3. 受支持的非 *IBM* 仓库源的版本和发行版级别

源数据库	Windows NT	AIX
Informix	7.2.2 - 8.2.1	7.2.4 - 9.2.0
Oracle	7.3.2 - 8.1.5	8.1.5
Microsoft [®] SQL Server	7.0	

设置仓库源

表 3. 受支持的非 IBM 仓库源的版本和发行版级别 (续)

源数据库	Windows NT	AIX
Microsoft Excel	97	
Microsoft Access	97	
Sybase	11.5	11.9.2

可来自下列数据库和文件的数据填充“数据仓库中心”仓库:

- 任何 DB2 系列的数据库
- Oracle
- Sybase
- Informix
- Microsoft SQL Server
- IBM DataJoiner

有关与“数据仓库中心”配合使用 DataJoiner 的更多信息, 参见第98页的『定义供 DataJoiner 使用的仓库源』。

- “多虚拟存储器”(OS/390)、“虚拟机”(VM)和局域网(LAN)文件
- IMS 和“虚拟存储器访问方法”(VSAM)(带有 Data Joiner Classic Connect)

有关这些产品的版本和发行版及产品先决条件的信息, 请参阅《DB2 仓库管理器安装指南》。

也可使用“数据仓库中心”类属 ODBC 驱动程序来访问其他数据库和文件, 如基于工作站的电子表格和 Lotus Notes 数据库。

“数据仓库中心”还包括了几个可用来访问非 IBM 数据的 Merant ODBC 驱动程序。有关安装这些驱动程序的信息, 请参阅针对您的操作系统的《DB2 通用数据库快速入门》。

特定的仓库代理支持特定的源。表4概述了哪些仓库代理支持哪些源。

表 4. 仓库代理对源的支持

数据源	Windows NT 或 Windows 2000 代理	AIX 代理	Solaris 操作环境代理	IBM Operating System/2 (OS/2) 代理	AS/400	OS/390
DB2 系列	✓	✓	✓	✓	✓	✓
DataJoiner	✓	✓				

表 4. 仓库代理对源的支持 (续)

数据源	Windows NT 或 Windows 2000 代理	AIX 代理	Solaris 操作环境代理	IBM Operating System/2 (OS/2) 代理	AS/400	OS/390
Oracle	✓	✓	✓	✓		
Sybase	✓	✓	✓	✓		
Microsoft SQL Server	✓	✓	✓			
Informix	✓	✓	✓			
类属 ODBC 驱动程序	✓	✓	✓	✓		
本地文件	✓	✓	✓	✓	✓	✓ ²
远程文件	✓	✓	✓	✓	✓	
IMS	✓	✓	✓ ¹	✓ ¹		✓ ³
VSAM	✓	✓	✓ ¹	✓ ¹		✓ ³

1. 当 DataJoiner 在远程工作站上时，要访问 IMS 或 VSAM 数据，需在代理站点上编目 DataJoiner 所驻留的节点，并编目 DataJoiner 数据库。
2. 平面文件不能作为 OS/390 的 ODBC 源来访问，但它们可用作仓库实用程序的源。

Windows NT 或 Windows 2000

表5列示 Windows NT 或 Windows 2000 上受支持的数据源，并说明需要做些什么才能与那些源连接。

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
“DB2 通用数据库”数据库	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: “DB2 通用数据库版本 7”服务器或 DB2 客户机</p>	<ol style="list-style-type: none"> 1. 在代理站点上安装 DB2 服务器或 DB2 客户机。 2. 编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。 4. 将数据库实用程序和 ODBC (CLI) 与数据库绑定。

设置仓库源

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
DB2 DRDA 数据库	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: DB2 Connect[®]</p>	<ol style="list-style-type: none"> 1. 在网关站点上安装 DB2 Connect。 2. 在代理站点上编目网关站点的节点。 3. 在代理站点上编目 DB2 Connect 数据库。 4. 标识对源数据库具有访问权的用户标识和密码。
IMS (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 必须使用以下程序之一:</p> <ul style="list-style-type: none"> • CROSS ACCESS ODBC 驱动程序和 DataJoiner Classic Connect • DataJoiner 和 DataJoiner Classic Connect <p>如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: 若正使用 CROSS ACCESS ODBC 驱动程序, 为 ODBC 若正使用 DataJoiner, 为 TCP/IP 或 APPC</p> <p>客户机使能器程序: 无</p>	<p>若正使用 CROSS ACCESS ODBC 驱动程序</p> <ol style="list-style-type: none"> 1. 建立从代理站点到主机的链接。 2. 在主机上安装和配置数据服务器。 3. 在代理站点上安装和配置 CROSS ACCESS ODBC 驱动程序。 4. 标识对源数据库具有访问权的用户标识和密码。 <p>若正使用 DataJoiner, 从 DataJoiner 工作站执行下列过程:</p> <ol style="list-style-type: none"> 1. 建立从工作站到主机的链接。 2. 在主机上安装和配置适配器。 3. 标识对源数据库具有访问权的用户标识和密码。 <p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
VSAM (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则必须以下程序之一:</p> <ul style="list-style-type: none"> • CROSS ACCESS ODBC 驱动程序和 DataJoiner Classic Connect • DataJoiner 和 DataJoiner Classic Connect <p>如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: 若正使用 CROSS ACCESS ODBC 驱动程序, 为 ODBC 若正使用 DataJoiner, 为 TCP/IP 或 APPC</p> <p>客户机使能器程序: 无</p>	<p>若正使用 CROSS ACCESS ODBC 驱动程序</p> <ol style="list-style-type: none"> 1. 建立从代理站点到主机的链接。 2. 在主机上安装和配置数据服务器。 3. 在代理站点上安装和配置 CROSS ACCESS ODBC 驱动程序。 4. 标识对源数据库具有访问权的用户标识和密码。 <p>若正使用 DataJoiner, 从 DataJoiner 工作站执行下列过程:</p> <ol style="list-style-type: none"> 1. 建立从工作站到主机的链接。 2. 在主机上安装和配置适配器。 3. 标识对源数据库具有访问权的用户标识和密码。 <p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。
OS/390 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。</p>
VM 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。</p>

设置仓库源

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
Sybase	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: Sybase Open Client-Library 10.0.4 或更高版本以及相应的 Sybase Net-Library。</p> <p>对于 Intel® 系统, 为 Sybase Open Client Library 版本 10.03 (或更新版本) 和适当的 Sybase Net-Library</p> <p>对于 Alpha 系统, 为 Sybase Open Client Library 版本 11.01 (或更新版本) 和适当的 Sybase Net-Library</p>	<ol style="list-style-type: none">1. 在代理站点上安装 Open Client。2. 根据客户机使能器的指示信息来编目远程数据库。3. 标识对源数据库具有访问权的用户标识和密码。

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
Oracle	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: Oracle SQL*Net V2</p>	<ol style="list-style-type: none"> 1. 在代理站点上安装 SQL*NET。 2. 根据客户机使能器指示信息编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。
	<p>数据库客户机需求</p> <p>对于 Oracle 版本 7: SQL*Net 的当前版本的相应 DLL, 加上 OCIW32.DLL。例如, SQL*Net 2.3 需要 ORA73.DLL、CORE35.DLL、NLSRTL32.DLL、CORE350.DLL 和 OCIW32.DLL。</p> <p>对于 Oracle 版本 8: 要访问版本级别为 8.0.3 或更高版本的远程 Oracle8 数据库服务器, 应安装 Oracle Net8 客户机版本 7.3.4.x、8.0.4 或更高版本。</p> <p>在 Intel 系统上, 安装 Oracle Net8 客户机的相应 DLL (例如 Ora804.DLL、PLS804.DLL 和 OCI.DLL) 到您的路径中。</p>	

设置仓库源

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
Informix	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: 对于 Informix 5、6 和 7.x, 为 i-connect 7.2 对于 Informix 7.x 和 9.x, 为 i-connect 9.x</p> <p>数据库客户机需求 Informix-Connect for Windows Platforms 2.x 或 Informix-Client Software Developer's Kit for Windows Platforms 2.x</p>	<ol style="list-style-type: none"> 1. 在代理站点上安装 i-connect。 2. 根据客户机使能器指示信息编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。
Microsoft SQL Server	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: 对于访问版本 6.0 的 DBMS, 为 Microsoft SQL Server DB-Library and Net-Library 版本 6.0 对于访问版本 7.0 的 DBMS, 为 Microsoft SQL Server DB-Library and Net-Library 版本 7.0</p>	<ol style="list-style-type: none"> 1. 在代理站点上安装 Microsoft SQL Server DB-Library 和 Net-Library。 2. 根据客户机使能器指示信息编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。
本地文件	<p>数据库访问程序: 无</p> <p>源 / 代理连接: TCP/IP</p> <p>客户机使能器程序: 无</p>	<p>若有需要, 提供访问前命令和 / 或访问后命令。</p>

表 5. Windows NT 或 Windows 2000 上受支持的数据源的联网需求 (续)

源数据库或文件:	Windows NT 或 Windows 2000 的先决条件产品:	如何连接:
远程文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	1. 获取 FTP 服务器或安装字符串。 2. 标识对源文件具有访问权的用户标识和密码。 有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。 有关使用 FTP 访问文件的更多信息, 参见第55页的『用 FTP 访问文件』。
Microsoft Access	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: 无	1. 使用类属 ODBC 连接字符串。 2. 参见 Microsoft Access 帮助主题以获取 Microsoft Access 支持的 ANSI SQL 数据类型的映射。 有关创建使用 Microsoft Access 作为源数据库的仓库的详细指示信息, 参见第78页的『Microsoft Access』。
Microsoft Excel	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: 无	1. 使用类属 ODBC 连接字符串。 2. 参见 Microsoft Excel 帮助主题以获取 Microsoft Excel 支持的 ANSI SQL 数据类型的映射。 有关创建使用 Microsoft Excel 作为源数据库的仓库的详细指示信息, 参见第82页的『Microsoft Excel』。

注意事项:

¹ 如果是在 Windows 95 或 Windows 98 操作系统上使用 Sybase 客户机, 则必须包括以下语句作为 autoexec.bat 文件中的最后一条语句:

```
CALL C:\SYBASE\BIN\WSYBSET.BAT
```

² 参阅《仓库管理器安装指南》获取关于 OS/390 代理的更多信息。

AIX

第34页的表6列示 AIX 上受支持的数据源, 并说明需要做些什么才能与那些源连接。

设置仓库源

AIX 仓库代理有两个版本：一个用于 ODBC 访问，一个用于 CLI 访问。下表中列示的非 DB2 数据库要求安装仓库代理的 ODBC 版本。

要验证已安装仓库代理的哪个版本，输入如下命令：

```
ls -l /usr/bin/IWH2AGNT
```

若命令返回 db2cli，表示正在使用 DB2 CLI 版本。若命令返回 ivodbc，表示正在使用 ODBC 版本。

如果正在使用的“数据仓库中心”AIX 代理已链接且可访问 Merant ODBC 源，并且将访问 DB2 数据库，应更改 .odbc.ini 文件的 DB2 源部分中的 Driver= 属性的值，如下所示：

```
Driver=/usr/lpp/db2_07_01/lib/db2_36.o
```

以下示例是一个可用于 AIX 的样本 ODBC 源条目：

```
[SAMPLE] Driver=/usr/lpp/db2_07_01/lib/db2_36.o  
Description=DB2 ODBC Database  
Database=SAMPLE
```

有关安装和配置仓库代理的详情，参见《DB2 仓库管理器安装指南》。

表 6. AIX 上受支持的数据源的联网需求

源数据库或文件:	AIX 的先决条件产品:	如何连接:
“DB2 通用数据库” 数据库	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: “DB2 UDB 版本 7” 服务器或 DB2 客户机	1. 在代理站点上安装 DB2 服务器或 DB2 客户机。 2. 编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。 4. 将数据库实用程序和 ODBC (CLI) 与数据库绑定。
DB2 DRDA 数据库	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: DB2 Connect	1. 在网关站点上安装 DB2 Connect。 2. 在代理站点上编目网关站点的节点。 3. 在代理站点上编目 DB2 Connect 数据库。 4. 标识对源数据库具有访问权的用户标识和密码。

表 6. AIX 上受支持的数据源的联网需求 (续)

源数据库或文件:	AIX 的先决条件产品:	如何连接:
IMS (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则需要 DataJoiner 和 DataJoiner Classic Connect 如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: TCP/IP 或 APPC</p> <p>客户机使能器程序: 无</p>	<p>从 DataJoiner 工作站:</p> <ol style="list-style-type: none"> 1. 建立从工作站到主机的链接。 2. 在主机上安装和配置适配器。 3. 标识对源数据库具有访问权的用户标识和密码。 <p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。
VSAM (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则需要 DataJoiner 和 DataJoiner Classic Connect 如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: TCP/IP 或 APPC</p> <p>客户机使能器程序: 无</p>	<p>从 DataJoiner 工作站:</p> <ol style="list-style-type: none"> 1. 建立从工作站到主机的链接。 2. 在主机上安装和配置适配器。 3. 标识对源数据库具有访问权的用户标识和密码。 <p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。
OS/390 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第5页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第5页的『用 FTP 访问文件』。</p>
VM 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第5页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第5页的『用 FTP 访问文件』。</p>

设置仓库源

表 6. AIX 上受支持的数据源的联网需求 (续)

源数据库或文件:	AIX 的先决条件产品:	如何连接:
Sybase	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: “Intersolv 版本 3.6 驱动程序管理器”和 Sybase 驱动程序</p> <p>数据库客户机需求 在非 DCE 环境中: (ibsyb15 ODBC 驱动程序): libct 库 在 DCE 环境中 (ibsyb1115 ODBC 驱动程序): Sybase 11.1 client library libct_r</p>	<ol style="list-style-type: none">1. 安装 AIX 仓库代理使用 ODBC 进行访问的版本。2. 标识对源数据库具有访问权的用户标识和密码。3. 安装 Open Client.
Oracle	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: “Intersolv 版本 3.6 驱动程序管理器”和 Oracle 驱动程序</p> <p>数据库客户机需求 对于 Oracle 版本 7: Oracle7 SQL*Net 和 Oracle7 SQL*Net 共享库 (通过 genclntsh 脚本构建) 对于 Oracle 版本 8: Oracle8 Net8 和 Oracle8 SQL*Net 共享库 (通过 genclntsh8 脚本构建)</p>	<ol style="list-style-type: none">1. 安装 AIX 仓库代理使用 ODBC 进行访问的版本。2. 标识对源数据库具有访问权的用户标识和密码。

表 6. AIX 上受支持的数据源的联网需求 (续)

源数据库或文件:	AIX 的先决条件产品:	如何连接:
Informix	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: “Intersolv 版本 3.6 驱动程序管理器” 和 Informix 驱动程序 数据库客户机需求 Informix-Connect 和 ESQL/C 版本 9.1.4 或更高版本	<ol style="list-style-type: none"> 1. 安装 AIX 仓库代理使用 ODBC 进行访问的版本。 2. 标识对源数据库具有访问权的用户标识和密码。
Microsoft SQL Server	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: “数据仓库中心” ODBC 驱动程序管理 ¹	<ol style="list-style-type: none"> 1. 安装 AIX 仓库代理使用 ODBC 进行访问的版本。 2. 标识对源数据库具有访问权的用户标识和密码。
本地文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	若有需要, 提供访问前命令或 / 和访问后命令。
远程文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	<ol style="list-style-type: none"> 1. 获取 FTP 服务器或安装字符串。 2. 标识对源文件具有访问权的用户标识和密码。 <p>有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。</p>

设置仓库源

表 6. AIX 上受支持的数据源的联网需求 (续)

源数据库或文件:	AIX 的先决条件产品:	如何连接:
注释:		
1 “数据仓库中心” ODBC 驱动程序和驱动程序管理器是 Merant 提供的 DataDirect Connect ODBC 驱动程序和驱动程序管理器。		
2 Open Client 是连接到 Windows NT 或 Windows 2000 操作系统上的 Sybase 源所需的。		
3 参阅《仓库管理器安装指南》以获取关于 OS/390 代理的更多信息。		

Solaris 操作环境

第39页的表7列示 Solaris 操作环境中受支持的数据源，并说明需要做些什么才能与那些源连接。

Solaris 操作环境仓库代理有两个版本：一个用于 ODBC 访问，一个用于 CLI 访问。下表中列示的非 DB2 数据库要求安装仓库代理的 ODBC 版本。

要验证已安装仓库代理的哪个版本，输入如下命令：

```
ls -l /usr/bin/IWH2AGNT
```

若命令返回 db2cli，表示正在使用 DB2 CLI 版本。若命令返回 ivodbc，表示正在使用 ODBC 版本。

如果正在使用的“数据仓库中心代理 Solaris 操作环境版”已链接且可访问 Merant ODBC 源，并且将访问 DB2 数据库，应更改 .odbc.ini 文件的 DB2 源部分中的 Driver= 属性的值，如下所示：

```
Driver=/opt/IBMDB2/V7.1/lib/libdb2_36.so
```

以下示例是一个可用于 Solaris 操作环境的样本 ODBC 源条目：

```
[SAMPLE] Driver=/opt/IBMDB2/V7.1/lib/libdb2_36.so
Description=DB2 ODBC Database
Database=SAMPLE
```

有关安装和配置仓库代理的详情，参见《DB2 仓库管理器安装指南》。

表 7. Solaris 操作环境中受支持的数据源的联网需求

源数据库或文件:	Solaris 操作环境的先决条件产品:	如何连接:
DB2 Universal Database	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: “DB2 UDB 版本 7” 服务器或 DB2 客户机</p>	<ol style="list-style-type: none"> 1. 在代理站点上安装 DB2 服务器或 DB2 客户机。 2. 编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。 4. 将数据库实用程序和 ODBC(CLI) 与数据库绑定。
DB2 DRDA 数据库	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: DB2 Connect</p>	<ol style="list-style-type: none"> 1. 在网关站点上安装 DB2 Connect。 2. 在代理站点上编目网关站点的节点。 3. 在代理站点上编目 DB2 Connect 数据库。 4. 标识对源数据库具有访问权的用户标识和密码。
IMS (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则需要 DataJoiner 和 DataJoiner Classic Connect 如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: ODBC (到 DataJoiner 数据库)</p> <p>客户机使能器程序: 无</p>	<p>从 DataJoiner 工作站:</p> <ol style="list-style-type: none"> 1. 建立从工作站到主机的链接。 2. 在主机上安装和配置适配器。 3. 标识对源数据库具有访问权的用户标识和密码。 <p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。
VSAM (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则需要 DataJoiner 和 DataJoiner Classic Connect 如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: ODBC (到 DataJoiner 数据库)</p> <p>客户机使能器程序: 无</p>	<p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。

设置仓库源

表 7. Solaris 操作环境中受支持的数据源的联网需求 (续)

源数据库或文件:	Solaris 操作环境的先决条件产品:	如何连接:
OS/390 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第 55 页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第 55 页的『用 FTP 访问文件』。</p>
VM 文件	<p>数据库访问程序: FTP 或 NFS</p> <p>源 / 代理连接: TCP/IP (FTP 或 NFS)</p> <p>客户机使能器程序: 无</p>	<p>建立从代理站点到主机的链接。</p> <p>有关使用 NFS 访问文件的详情, 参见第 55 页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第 55 页的『用 FTP 访问文件』。</p>
Sybase	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: 对于 Sybase 11, 为 “Intersolv 版本 3.6 驱动程序管理器” 和 Sybase 驱动程序。</p> <p>数据库客户机需求 在 DCE 环境中 (ibsyb1115 ODBC 驱动程序): Sybase 11.1 客户机库 libct_r。 在非 DCE 环境中 (ibsyb15 ODBC 驱动程序): libct 库</p>	<ol style="list-style-type: none">1. 安装 Solaris 操作环境仓库代理使用 ODBC 进行访问的版本。2. 标识对源数据库具有访问权的用户标识和密码。3. 安装 Open Client。

表 7. Solaris 操作环境上受支持的数据源的联网需求 (续)

源数据库或文件:	Solaris 操作环境的先决条件产品:	如何连接:
Oracle	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: 对于 Oracle 版本 7.3.2, 为 “Intersolv 版本 3.6 驱动程序管理器” 和 Oracle 驱动程序。</p> <p>数据库客户机需求 对于 Oracle 版本 7: Oracle7 SQL*Net 和 Oracle7 SQL*Net 共享库 (通过 genclntsh 脚本构建)。 对于 Oracle 版本 8: Oracle8 Net8 和 Oracle8 SQL*Net 共享库 (通过 genclntsh8 脚本构建)。</p>	<ol style="list-style-type: none"> 1. 安装 Solaris 操作环境仓库代理使用 ODBC 进行访问的版本。 2. 标识对源数据库具有访问权的用户标识和密码。
Informix	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: “Intersolv 版本 3.6 驱动程序管理器” 和 Informix 驱动程序。</p> <p>数据库客户机需求 Informix-Connect 和 ESQL/C 版本 9.1.3 或更高版本。</p>	<ol style="list-style-type: none"> 1. 安装 Solaris 操作环境仓库代理使用 ODBC 进行访问的版本。 2. 标识对源数据库具有访问权的用户标识和密码。
Microsoft SQL Server	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: 数据仓库中心 ODBC 驱动程序管理器</p>	<ol style="list-style-type: none"> 1. 安装 AIX 仓库代理使用 ODBC 进行访问的版本。 2. 标识对源数据库具有访问权的用户标识和密码。

设置仓库源

表 7. Solaris 操作环境中受支持的数据源的联网需求 (续)

源数据库或文件:	Solaris 操作环境的先决条件产品:	如何连接:
本地文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	若有需要, 提供访问前命令或 / 和访问后命令。
远程文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	1. 获取 FTP 服务器或安装字符串。 2. 标识对源文件具有访问权的用户标识和密码。 有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。 有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。

注释:

¹ “数据仓库中心” ODBC 驱动程序和驱动程序管理器是 Merant 提供的 DataDirect Connect ODBC 驱动程序和驱动程序管理器。

² Open Client 是连接到 Windows NT 或 Windows 2000 操作系统上的 Sybase 源所需的。

³ 参阅《仓库管理器安装指南》以获取关于 OS/390 代理的更多信息。

OS/2

表8列示了在 OS/2 上受支持的数据源, 并说明需要做些什么才能与那些源相连接。

表 8. OS/2 上受支持的数据源的联网需求

源数据库或文件:	OS/2 的先决条件产品:	如何连接:
“DB2 通用数据库” 数据库	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: “DB2 UDB 版本 7” 服务器或 DB2 客户机	1. 在代理站点上安装 DB2 服务器或 DB2 客户机。 2. 编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。 4. 将数据库实用程序和 ODBC(CLI) 与数据库绑定。

表 8. OS/2 上受支持的数据源的联网需求 (续)

源数据库或文件:	OS/2 的先决条件产品:	如何连接:
DB2 DRDA 数据库	<p>数据库访问程序: 无</p> <p>源 / 代理连接: ODBC</p> <p>客户机使能器程序: DB2 Connect</p>	<ol style="list-style-type: none"> 1. 在网关站点上安装 DB2 Connect。 2. 在代理站点上编目网关站点的节点。 3. 在代理站点上编目 DB2 Connect 数据库。 4. 标识对源数据库具有访问权的用户标识和密码。
LAN BLOB 文件	<p>数据库访问程序: 无</p> <p>源 / 代理连接: TCP/IP</p> <p>客户机使能器程序: 无</p>	<ol style="list-style-type: none"> 1. 获取 FTP 服务器或安装字符串。 2. 标识对源文件具有访问权的用户标识和密码。 <p>有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。</p>
IMS (OS/390)	<p>数据库访问程序: 如果没有使用 OS/390 代理, 则需要 DataJoiner 和 DataJoiner Classic Connect 如果使用的是 OS/390 代理, 则需要 Classic Connect</p> <p>源 / 代理连接: ODBC (到 DataJoiner 数据库)</p> <p>客户机使能器程序: 无</p>	<p>若正使用 DataJoiner, 从代理站点执行如下过程:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。
VSAM (OS/390)	<p>数据库访问程序: DataJoiner 和 DataJoiner Classic Connect</p> <p>源 / 代理连接: ODBC (到 DataJoiner 数据库)</p> <p>客户机使能器程序: 无</p>	<p>从代理站点:</p> <ol style="list-style-type: none"> 1. 编目 DataJoiner 所驻留的节点。 2. 编目 DataJoiner 数据库。

设置仓库源

表 8. OS/2 上受支持的数据源的联网需求 (续)

源数据库或文件:	OS/2 的先决条件产品:	如何连接:
OS/390 文件	数据库访问程序: FTP 或 NFS 源 / 代理连接: TCP/IP (FTP 或 NFS) 客户机使能器程序: 无	建立从代理站点到主机的链接。 有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。 有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。
VM 文件	数据库访问程序: FTP 或 NFS 源 / 代理连接: TCP/IP (FTP 或 NFS) 客户机使能器程序: 无	建立从代理站点到主机的链接。 有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。 有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。
Sybase	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: Sybase Open Client Library 版本 10 OS/2 版, 为 “Intersolv 版本 3.6 驱动程序管理器” 和 Sybase 驱动程序	<ol style="list-style-type: none">1. 在代理站点上安装 Open Client。2. 根据客户机使能器指示信息编目远程数据库。3. 标识对源数据库具有访问权的用户标识和密码。
Oracle	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: Oracle SQL*Net V2.1.4 OS/2 版, 为 “Intersolv 版本 3.6 驱动程序管理器” 和 Oracle 驱动程序	<ol style="list-style-type: none">1. 在代理站点上安装 SQL*NET。2. 根据客户机使能器指示信息编目远程数据库。3. 标识对源数据库具有访问权的用户标识和密码。

表 8. OS/2 上受支持的数据源的联网需求 (续)

源数据库或文件:	OS/2 的先决条件产品:	如何连接:
Informix	数据库访问程序: 无 源 / 代理连接: ODBC 客户机使能器程序: Informix-Net OS/2 版, 为 “Intersolv 版本 3.6 驱动程序管理器” 和 Informix 驱动程序	<ol style="list-style-type: none"> 1. 在代理站点上安装 Informix-Net。 2. 根据客户机使能器指示信息编目远程数据库。 3. 标识对源数据库具有访问权的用户标识和密码。
本地文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	若有需要, 提供访问前命令或 / 和访问后命令。
远程文件	数据库访问程序: 无 源 / 代理连接: TCP/IP 客户机使能器程序: 无	<ol style="list-style-type: none"> 1. 获取 FTP 服务器或安装字符串。 2. 标识对源文件具有访问权的用户标识和密码。 <p>有关使用 NFS 访问文件的详情, 参见第55页的『用 NFS 访问文件』。</p> <p>有关使用 FTP 访问文件的详情, 参见第55页的『用 FTP 访问文件』。</p>

注释:

¹ 参阅《仓库管理器安装指南》以获取关于 OS/390 代理的更多信息。

设置对“DB2 通用数据库”源数据库的访问权

可使用任何“DB2 通用数据库”数据库作为仓库的源数据库。要获取关于使用 DB2 通用数据库的更多信息, 请参阅《DB2 通用数据库安装和配置补遗》。

要设置对“DB2 通用数据库”源数据库的访问权:

1. 定义对源数据库的特权。
2. 建立与源数据库的联网。

设置仓库源

定义对“DB2 通用数据库”源数据库的特权

源系统的系统管理员必须设置具有下列特权的用户标识:

- BINDADD (数据库级特权)
- CONNECT (数据库级特权)

另外, 下列 SYSIBM 系统表需要显式 SELECT 特权:

- SYSIBM.SYSTABLES
- SYSIBM.SYSCOLUMNS
- SYSIBM.SYSDBAUTH
- SYSIBM.SYSTABAUTH
- SYSIBM.SYSINDEXES
- SYSIBM.SYSRELS
- SYSIBM.SYSTABCONST

想要访问的任何表也需要显式 SELECT 特权。

建立与“DB2 通用数据库”源数据库的联网

在系统管理员定义必需的特权之后, 在代理站点建立与源数据库的联网:

1. 若数据库是远程的, 则设置与数据库的通信。
2. 若数据库是远程的, 则编目该节点。
3. 编目该数据库。
4. 若正使用 Windows NT、Windows 2000 或 OS/2 仓库代理, 或正使用 AIX 或 Solaris 操作环境使用 ODBC 的版本, 则将该数据库登录为 ODBC 系统 DSN。若正在使用的 AIX、OS/390 或 Solaris 操作环境仓库代理使用 CLI 界面, 则使用 DB2 编目实用程序编目该数据库。
5. 将数据库实用程序和 ODBC(CLI) 与数据库绑定。每种类型的客户机只需要绑定一次。

设置对 DB2 DRDA 源数据库的访问权

要使用网关访问 DB2 Distributed Relational Database Architecture™ (DRDA®) 源数据库:

1. 验证网关服务器安装有先决条件产品
2. 定义对源数据库的特权。
3. 设置 DB2 Connect™ 网关站点。
4. 建立与源数据库的联网。

要通过直接连接访问 DB2 Distributed Relational Database Architecture™ (DRDA®) 源数据库，建立与源数据库的联网。参见第48页的『建立与 DB2 DRDA 源数据库的联网』以获取循序渐进的指示信息。

先决条件产品

若要访问下列其中一个源数据库中的数据，需要网关站点。为 DRDA 配置该站点：

- DB2 通用数据库 AS/400 版®
- DB2 通用数据库 OS/390 版
- DB2 VM 版
- DB2 VSE 版

有关 DRDA 的详情，参见 *Distributed Relational Database Architecture Connectivity Guide*。

有关 DB2 Connect 的信息，参见下列书籍：

- 您的产品版本的《DB2 Connect 快速入门》
- 《DB2 Connect 用户指南》
- 《安装和配置补遗》

定义对 DB2 DRDA 源数据库的特权

源系统的系统管理员必须在为 DRDA 配置的服务器上设置具有下列特权的用户标识：

- 对于所有 DRDA 服务器，该用户标识必须有权与数据库“连接”。
 - 另外，下列系统表以及想要访问的任何表，需要显式 SELECT 特权：
 - SYSIBM.SYSTABLES
 - SYSIBM.SYSCOLUMNS
 - SYSIBM.SYSDBAUTH
 - SYSIBM.SYSTABAUTH
 - SYSIBM.SYSINDEXES
 - SYSIBM.SYSRELS
 - SYSIBM.SYSTABCONST
- 对于 DB2 OS/390 版，该用户标识必须具有下列其中一种授权：
 - SYSADM
 - SYSCTRL
 - BINDADD 和 CREATE IN COLLECTION NULLID 授权

设置仓库源

- 对于 DB2 VSE 版或 DB2 VM 版，该用户标识必须具有 DBA 权限。
要使用 BIND 命令上的 GRANT 选项，NULLID 这个用户标识必须有权给其他用户授予对下列表的权限：
 - SYSTEM.SYSCATALOG
 - SYSTEM.SYSCOLUMNS
 - SYSTEM.SYSINDEXES
 - SYSTEM.SYSTABAUTH
 - SYSTEM.SYSKEYCOLS
 - SYSTEM.SYSSYNONYMS
 - SYSTEM.SYSKEYS
 - SYSTEM.SYSCOLAUTH
- 对于“DB2 通用数据库 AS/400 版”，用户标识必须在 NULLID 集合上具有 CHANGE 或更高权限

设置 DB2 Connect 网关节点

在用必需的特权设置您的用户标识之后，在网关节点上执行下列任务：

1. 对于 Windows NT 或 Windows 2000，安装 Microsoft SNA Server。
2. 安装 DB2 Connect。
3. 配置 DB2 Connect 系统与源数据库通信。
4. 更新 DB2 节点目录、系统数据库目录和 DCS 目录。

建立与 DB2 DRDA 源数据库的联网

需要在仓库代理站点建立与源数据库的联网：

1. 设置与 DB2 Connect 站点的通信。
2. 编目 DB2 Connect 站点的节点。
3. 编目该数据库。
4. 使用“ODBC 管理员”将该数据库登录为系统 DSN。
5. 如《DB2 Connect 用户指南》中所述的那样，将 DB2 Connect 实用程序与 DRDA 服务器绑定。每种类型的客户机只需要绑定一次。

提示：在 Windows 32 位系统上，可使用“DB2 UDB 客户机配置辅助”完成此任务。参见《IBM DB2 通用数据库：快速入门 Windows 版》以获取有关用“DB2 UDB 客户机配置辅助”登录数据库的信息。

从 AS/400 代理访问远程数据库的联网需求

可通过使用“IBM 分布式关系数据库体系结构”(DRDA)或基于 TCP/IP 的 DRDA 的“系统网络体系结构”(SNA)联网,通过 AS/400 代理访问远程数据库。

必须使用 DRDA 联网 访问下列远程数据库:

- DB2 通用数据库 AS/400 版
- DB2 通用数据库 OS/390 版

当满足下列条件时,可从 AS/400 代理连接至远程数据库:

- 与远程数据库的 SNA 连接是正确的。
- 远程数据库已编目在“AS/400 关系数据库目录”中。

提示: 若满足下列条件,应能从“数据仓库中心”连接至远程数据库并查询它:

- 可从 AS/400 代理连接至远程数据库。
- 可从 AS/400 交互式 SQL 设施 (STRSQL) 查询远程数据库。

有关使用 DRDA 通过 SNA 与不同类型的数据库连接的详情,参见红皮书 *Distributed Relational Database Cross Platform Connectivity and Application*。也可联机查看此书,网址为 <http://www.redbooks.ibm.com>。

要获取关于使用基于 TCP/IP 的 DRDA 的更多信息,参阅《仓库管理器安装指南》

从 AS/400 代理建立与本地和远程数据库的联网

必须在代理站点上的“AS/400 关系数据库”目录中,编目计划用作仓库源或目标的本地和远程数据库的名称。必须也在代理访问的远程工作站上编目这些数据库名称。

在代理将访问的远程工作站上,必须将在代理站点上编目的本地数据库名编目为远程数据库名。同理,在代理将访问的远程工作站上,必须将在代理站点上编目的远程数据库名编目为本地数据库名。

例如,Fred 正在创建数据仓库。他想要编目名为“销售”的数据库和名为“费用”的数据库的数据库名。名为“销售”的数据库与 AS/400 代理位于同一工作站。名为“费用”的数据库位于该代理将访问的远程工作站。第50页的表9描述 Fred 应如何在每个工作站上编目每个数据库。

设置仓库源

表 9. 如何编日本地和远程数据库名

数据库名	位置	在代理站点上编目为本地或远程	在远程机器上编目为本地或远程
销售	代理站点	本地	远程
费用	代理访问的机器	远程	本地

若源数据库和目标数据库位于同一工作站，必须将一个编目为本地，另一个编目为远程。

要将数据库名称项添加至“AS/400 关系数据库”目录，在 AS/400 命令提示符处输入如下命令：

```
ADDRDBDIRE databasename locationname
```

其中 `databasename` 是 AS/400 数据库的名称，`locationname` 是 AS/400 工作站的位置名。必须指定数据库是本地的还是远程的。

确保提供数据库名和位置名，即使它们是同一名称。

对于本地数据库，位置名为 `*LOCAL` 关键字。对于每个远程数据库，位置字段必须包含 `SNA LU` 名称。

注意：若更改“远程数据库目录”中的数据库的名称，必须更新参考它的每个仓库源。未那样做会导致仓库源数据库连接错误。

也可使用 **WRKRDBDIRE** 命令查看、添加、更改和除去“关系数据库”目录项。要使用此命令，在 AS/400 命令提示符处输入该命令。显示当前定义的远程数据库名的列表。选项集显示在窗口的顶部。

有关详情，参见这些命令中的每一个的 AS/400 联机帮助。

从 OS/390 代理的远程数据库的联网需求

通过使用基于 TCP/IP 的“IBM 分布式关系数据库体系结构” (DRDA)，可从 OS/390 代理访问远程数据库。

要访问下列远程数据库，需要 DRDA 联网：

- DB2 通用数据库 OS/390 版
- DB2 通用数据库 AS/400 版
- Oracle、Sybase、Informix 和其他类似的数据库产品（当使用 DataJoiner[®] 时）

有关使用 DRDA 与不同类型的数据库连接的详情，参见红皮书 *IBM Distributed Relational Database Architecture Reference*。

OS/390 代理允许通过 Classic Connect ODBC 驱动程序访问 IMS 和 VSAM 远程数据库。不能使用 Classic Connect ODBC 驱动程序直接访问“DB2 通用数据库”数据库。有关对 IMS 和 VSAM 数据库使用 Classic Connect 的详情，参见第 363 页的『附录 F. 在“数据仓库中心”中使用 Classic Connect』。

使用 DataJoiner

当使用 OS/390 代理访问“DataJoiner 版本 2”和其他 DataJoiner 数据源时，需要将“系统网络体系结构”(SNA) LU 6.2 用作通信协议。虽然 DataJoiner 在版本 2.1 和版本 2.1.1 中支持 TCP/IP 作为应用程序请求器，但它没有应用程序服务器。另外，不能将 DataJoiner 用作 OS/390 的目标，因为 DataJoiner 不支持 DRDA 的两阶段落实，而“DB2 通用数据库 OS/390 版”需要它。

有关使用 DRDA 通过 SNA 与不同类型的数据库连接的详情，参见 IBM 红皮书 *Distributed Relational Database Cross Platform Connectivity*。可联机查看此书，网址为 <http://www.redbooks.ibm.com>。

定义 DB2 仓库源

若正使用对仓库代理来说是远程的源数据库，必须在包含仓库代理的工作站上注册该数据库。

“数据仓库中心”支持使用普通 SQL 标识符的源表。普通标识符：

- 必须以字母开始
- 可以包括大写字母、数字和下划线
- 不能是保留字

若表的普通标识符的一部分是小写字母，则“数据仓库中心”将小写字母存储为大写字母。

“数据仓库中心”不支持使用定界标识符的源表。定界标识符：

- 包括在引号内
- 可以包括大写和小写字母、数字、下划线和空格
- 可以包含双引号（用两个连续的引号表示）

为了节省时间，可以将元数据从特定类型的表、文件和视图中导入到“数据仓库中心”。导入元数据可节省人工定义这些源所需的时间。有关详情，参见联机帮助。

设置仓库源

为了帮助您确定想使用数据源中的哪些表，可以使用“样本内容”功能来查看源表中的数据。每次可查看一个表中的数据。“数据仓库中心”显示该表的所有列名，而不管列中是否存在数据。最多可显示 200 行的数据。

可以在导入表的定义之前或之后查看数据。

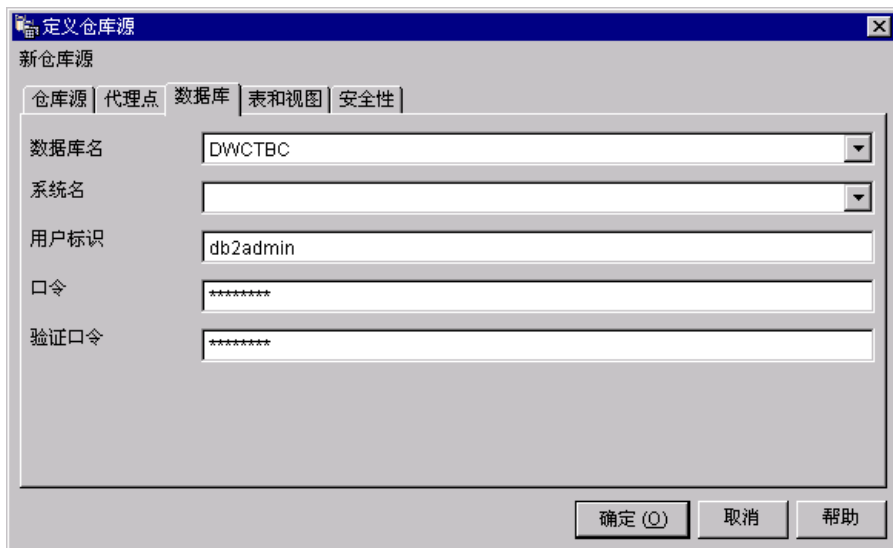
任何仓库用户都可定义仓库源，但只有属于对仓库源具有访问权的仓库组的仓库用户才有权更改它。

对 DB2 VM 版的数据库定义仓库源（它是通过 DRDA 网关访问的）时，使用 CLOB 和 BLOB 数据类型有以下限制：

- 不能使用“样本内容”功能来查看数据类型为 CLOB 和 BLOB 的数据。
- 不能在 SQL 步骤中使用数据类型为 CLOB 和 BLOB 的列。这是 DB2 VM 版的版本 5.2 服务器中的一个已知限制，在该版本中 LOB 对象不能使用 DRDA 传送到 DB2 版本 7 客户机中。

要定义“DB2 通用数据库”仓库源：

1. 右键单击**仓库源**文件夹。
2. 单击**定义** → **DB2 系列**，然后从菜单中单击适用于您的操作系统的 DB2 通用数据库的版本（例如 **DB2 UDB Windows NT 版**）。
“定义仓库源”笔记本打开。
3. 在**名称**字段中，输入仓库源的业务名。
将在“数据仓库中心”中始终使用此名称引用该仓库源。
4. 在**管理员**字段中，输入该仓库源的联系人。
5. 在**描述**字段中，输入该数据的简短描述。
6. 单击**代理站点**页。可选择的站点显示在**可用的代理站点**列表中。
7. 从**可用的代理站点**列表中，选择想要对其授予仓库源访问权的站点名，并单击 **>**。该代理站点添加至**选择的代理站点**列表。该站点现在可访问仓库源。
要包括**可用的代理站点**列表中的所有项目，单击 **>>**。
8. 单击**数据库**选项卡。



9. 在**数据库名**字段中，输入物理数据库的名称。如果正在 OS/2 上定义仓库源，应以大写字母输入数据源的名称。如果正在定义 OS/390 数据源，这是数据库的位置名。
10. 在**用户标识**字段中，输入对该数据库具有访问权的用户标识。
11. 在**密码**字段中，输入将访问该数据库的用户标识的密码。
12. 在**验证密码**字段中，再次输入该密码。
13. 单击**代理站点**页。可选择的站点显示在**可用的代理站点**列表中。
14. 从**可用的代理站点**列表中，选择想要对其授予仓库源访问权的站点名，并单击 **>**。该代理站点添加至**选择的代理站点**列表。该站点现在可访问仓库源。要包括**可用的代理站点**列表中的所有项目，单击 **>>**。
15. 单击**表和视图**选项卡。
因为表在 DB2 数据库中，所以可从“DB2通用数据库”导入表定义，而无需人工定义它们。
在“OS/400[®] 版本 4 发行版 2”和“OS/400 版本 4 发行版 3”上，必须选择**视图**文件夹来导入系统表。
16. 展开**表**文件夹。
“过滤器”窗口打开。
17. 单击**确定**。
“数据仓库中心”显示进度窗口。导入可能要用几分钟。

设置仓库源

若仓库源选择了多个代理站点，仓库服务器将名称排在第一（取决于用户的语言环境）的代理站点用于导入。

例如，仓库源选择了如下三个代理站点：“缺省代理”、“AIX 代理”和“OS/390™ 代理”。仓库服务器将名为“AIX 代理”的代理站点用于导入。

导入完成之后，“数据仓库中心”在**可用的表和视图**列表中列示导入的对象。

18. 从**可用的表和视图**列表中，选择想要包括在仓库源中的表和视图，然后单击 >。

选择的表和视图移动至**选择的表和视图**列表。

19. 单击**安全性**选项卡。
20. 选择仓库组给组中的用户授予创建使用此仓库源的步骤的能力。
21. 单击 >。

该仓库组移动至**选择的安全性组**列表。

接受笔记本中的其余值。有关这些值的详情，参见联机帮助中的『仓库源』。

22. 单击**确定**保存更改并关闭“定义仓库源”笔记本。
23. 单击**仓库** → **特性**以打开“数据仓库中心”的“特性”笔记本。
24. 单击**处理选项**选项卡。
25. 在**计数**字段中，输入想要重试抽取的次数值，或使用该字段右边的按钮上的向上和向下箭头，在您可选择的值列表中滚动。
26. 单击**确定**保存所作的更改并关闭“数据仓库中心”的“特性”笔记本。

访问远程文件

步骤可以使用几种方法来访问远程工作站上的文件。本节列示了可为步骤设置的一些访问方法。

通过 Windows NT 或 Windows 2000 文件服务器访问文件

可将数据文件用作步骤的源文件。若文件不在代理站点上，而是通过 Windows NT 或 Windows 2000 文件服务器访问，要明白下列要求。访问 LAN 服务器上的远程文件的要求与这些要求类似。

代理站点必须具有有权访问该文件的用户标识和密码。代理站点必须包含一个执行 NET USE 命令的 .bat 文件。该文件必须至少包含下列行：

```
NET USE drive: /DELETE  
NET USE drive: //hostname/sharedDrive password /USER:userid
```

其中:

- *drive* 是代表代理站点上的共享驱动器的驱动器盘符。
- *hostname* 是远程工作站的 TCP/IP 主机名
- *sharedDrive* 是远程工作站上包含该文件的驱动器
- *password* 是访问共享驱动器所需的密码
- *userid* 是访问共享驱动器所需的用户标识

若驱动器盘符正在使用中, 则该文件的第一行释放它。该文件的第二行建立连接。

当您定义代理站点时, 指定用于访问该文件的用户标识和密码。

当定义的仓库源时, 在从“仓库源”笔记本的“文件”笔记本页打开的“高级”窗口中的**访问前命令**字段中, 指定该 .bat 文件。

还可定义类似的 .bat 文件, 以便在“数据仓库中心”处理完文件之后删除与远程驱动器的链接。若这样做, 则在“高级”窗口中的**访问后命令**字段中指定该 .bat 文件。

要使用源数据文件, 您还必须向 ODBC 将该文件注册为名为 IWH_TEXT 的系统 DSN。使用适当的驱动程序, 如 VISWHSE 3.6 32 位文本文件 (*.*)。

用 NFS 访问文件

从代理站点访问文件的另一种方法是使用 TCP/IP 的“网络文件系统”(NFS) 协议。使用 NFS 时, 在 NFS 命令中必须提供用户标识(若从 Hummingbird 使用 Maestro, 则该命令 NFS LINK)。必须在从“仓库源”笔记本的“文件”笔记本页打开的“高级”窗口中的**访问前命令**字段中, 指定这些访问命令。

若代理站点未安装 NFS, 则使用 NET USE 命令来访问 NFS, 就象第54页的『通过 Windows NT 或 Windows 2000 文件服务器访问文件』中描述的那样。

要使用源数据文件, 您还必须向 ODBC 将该文件注册为名为 IWH_TEXT 的系统 DSN。使用适当的驱动程序, 如 VISWHSE 3.6 32 位文本文件 (*.*)。

用 FTP 访问文件

可以使用 FTP 来访问远程工作站上的数据文件。作为一条准则, 若文件不超过 20 MB, 则使用 FTP。当将使用远程文件的步骤提升为测试方式时, 将象将该步骤提升为生产方式时那样传送这些文件。若文件比较大, 则提升要花很长时间, 并且代理站点的空间可能会用尽。

设置仓库源

避免此问题的一种方法是，在测试期间，在远程工作站上放置一个伪文件。另一方法是使用“使用 FTP 复制文件”而不是使用 FTP（参见『使用“使用 FTP 复制文件”访问数据文件』）。

如果计划使用 FTP 命令 GET 或 PUT，将文件从 OS/390 主机传送到另一远程主机，需要为您的 OS/390 系统更改 JCL 模板中的帐户信息。要获取更多关于更改“数据仓库中心”模板，以在 OS/390 系统上支持 FTP 的信息，请参阅《仓库管理器安装指南》。

要使用 FTP 访问数据文件：

1. 将文件存储在 FTP 站点上。
2. 用**远程文件的仓库源类型**声明该文件。
3. 在从“仓库源”笔记本的“文件”笔记本页打开的“高级”窗口中，指定**系统名、用户标识和密码**。

当将使用此源的步骤提升为测试方式时，“数据仓库中心”将该文件传送到代理站点上的临时文件。

如果访问安全 UNIX[®] 或“UNIX 系统服务”系统上的远程文件时有问题，请验证您是否对用户代理站点定义中指定的 \$HOME 目录有写访问权。

例如，代理站点的主机名为 glacier.stl.ibm.com。您想要使用用户标识vwinst2，使用 FTP 将文件从远程站点 kingkong.stl.ibm.com 传送到代理站点。vwinst2/.netrc 文件必须包含如下项目：

```
machine glacier.stl.ibm.com login vwinst2
```

使用“使用 FTP 复制文件”访问数据文件

可使用“使用 FTP 复制文件”来访问远程工作站上的数据文件。若文件大于 20 MB，请使用“使用 FTP 复制文件”。当将步骤提升为测试状态时，“数据仓库中心”不运行仓库程序，因此将不传送文件。还可为“使用 FTP 复制文件”指定目标文件的位置。

要使用“使用 FTP 复制文件”来访问文件：

1. 用**本地文件的仓库源类型**声明该文件。
2. 定义两个步骤来访问这么大的文件：
 - a. 定义第一个步骤使用“使用 FTP 复制文件”仓库程序。
使用此步骤将文件复制至代理站点。
 - b. 定义第二个步骤使用为该文件创建的仓库源。

该步骤将该文件作为本地文件访问。此文件是第一个步骤的输出文件。

定义文件源

可在仓库源中定义一个或多个文件。

在对“数据仓库中心”定义该文件之前，不能查看“本地文件”或“远程文件”仓库源中的数据。

要定义文件源：

1. 右键单击**仓库源**文件夹。
2. 单击**定义** → **平面文件**，然后单击**本地文件**或**远程文件**。
“仓库源”笔记本打开。
3. 在**名称**字段中，输入仓库源的业务名。
4. 在**管理员**字段中，输入仓库源的联系人的姓名。
5. 在**描述**字段中，输入该数据的简短描述。
6. 单击**代理站点**选项卡。可选择的站点显示在**可用的代理站点**列表中。
7. 从**可用的代理站点**列表中，选择想要对其授予仓库源访问权的站点名，并单击 **>**。该代理站点添加至**选择的代理站点**列表。现在可使用该站点访问仓库源。
8. 单击**文件**选项卡。
9. 右键单击**文件**列表的空白区域，并单击**定义**。
“定义仓库源文件”笔记本打开。
10. 在**文件名**字段中，输入全限定路径和文件。
文件的名称不得包含空格。在 **UNIX** 系统上，文件名是区分大小写的。
11. 可选：在**描述**字段中，输入文件的简短描述。
12. 可选：在**注意事项**字段中，输入关于该文件的注意事项。
13. 可选：在**业务名**字段中，输入文件的业务名。

- 单击**参数**选项卡。



- 从**文件类型**列表中，选择文件类型。
- 在**字段定界符**字符字段中，指定将用作定界符的字符。仅当从**文件类型**列表选择了**字符**时，此字段才是可用的。
- 若文件的第一行包含列名，则选择**首行包含列名**复选框。

“数据仓库中心”将忽略列名并从包含数据的文件的第二行开始抽取数据。

若正定义远程文件仓库源，指定下列 FTP 传送格式之一。

 - 选择 **ASCII** 以用 ASCII 格式传送文件。
 - 选择**二进制**以用二进制格式传送文件。
- 单击**字段**选项卡。

“数据仓库中心”读取在“仓库源文件”页上指定的文件。它基于文件中的字段定义列，并在**字段**列表中显示列定义。它在**文件预览**区域中显示样本数据。最多可显示 10 行样本数据。可滚动以查看所有样本数据。
- 若想要更改“数据仓库中心”生成的列名，双击该列名。
- 输入该列的新名称。
- 按 **Enter** 键。

新名称显示在**列名字段**中。在下图中，COL001 重命名为 STATE。



22. 单击**确定**。
“文件”笔记本关闭。
23. 如果需要指定一个预访问或后访问命令，可单击“仓库源”笔记本的“文件”页上的**高级**按钮。
高级窗口打开。
24. 在**访问前命令**字段中，输入用于访问逻辑文件的命令。
25. 在**访问后命令**字段中，输入访问文件之后使用的命令。
26. 单击**确定**保存更改并关闭窗口。
27. 单击**安全性**选项卡。
28. 选择仓库组给组中的用户授予创建使用此仓库源的步骤的能力。
29. 单击 **>**。该仓库组移动至**选择的安全性组**列表。
30. 单击**确定**保存更改并关闭“定义仓库源”笔记本。
31. 单击**仓库** **→ 特性**以打开“数据仓库中心”的“特性”笔记本。
32. 单击**处理选项**选项卡。
33. 在**计数**字段中，输入想要重试抽取的次数值，或使用该字段右边的按钮上的向上和向下箭头，在您可选择的值列表中滚动。

设置仓库源

34. 单击**确定**保存所作的更改并关闭“数据仓库中心”的“特性”笔记本。

访问非 DB2 数据库仓库源

可访问非 DB2 源中的数据。在创建任何将访问这些源的步骤之前，必须定义这些源。

访问某些源可能有多种方法。例如，可使用“数据仓库中心”ODBC 驱动程序访问许多非 DB2 数据库。但是也可使用 DataJoiner 访问这些数据库。有关与“数据仓库中心”配合使用 DataJoiner 的详情，参见第98页的『定义供 DataJoiner 使用的仓库源』。

在 Windows NT 上设置非 DB2 数据库仓库源

以下各节描述如何在 Windows NT 上设置 Informix、Sybase、Oracle、Microsoft SQL Server、Microsoft Access、Microsoft Excel、IMS 和 VSAM 源。

Informix

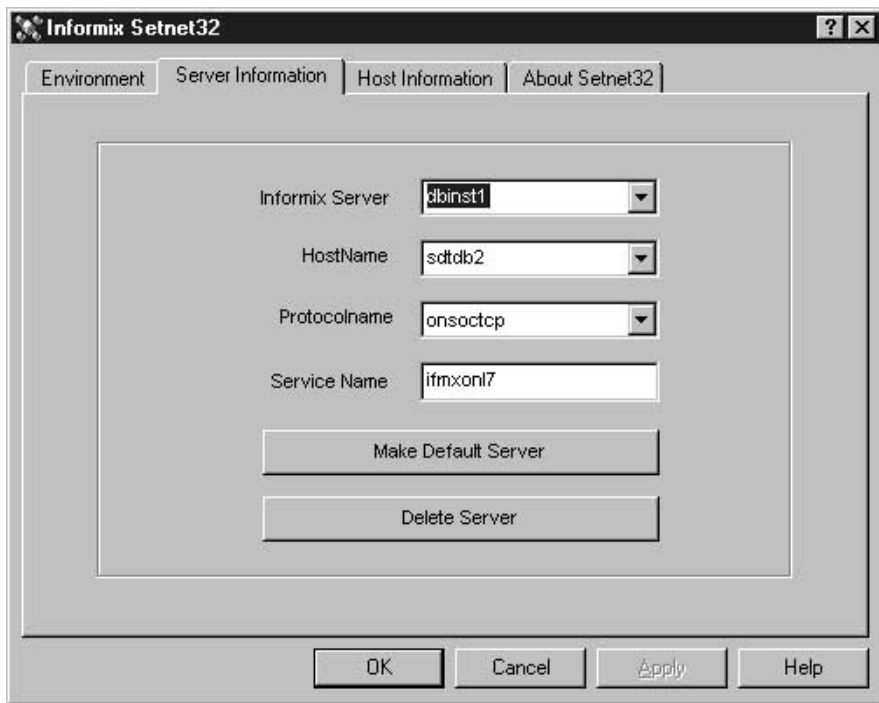
本节包含执行下列任务的步骤：

- 配置 Informix 客户机
- 安装和配置 ODBC 驱动程序

配置 Informix 客户机： 要设置对 Informix 客户机的访问权，必须使用 Informix-Setnet 32 位实用程序配置 Informix 服务器和主机信息。

要配置 Informix 服务器信息：

1. 单击**开始** → **程序** → **Informix 客户机** → **Setnet32**。
Setnet32 窗口打开。
2. 单击**服务器信息 (Server Information)**选项卡。

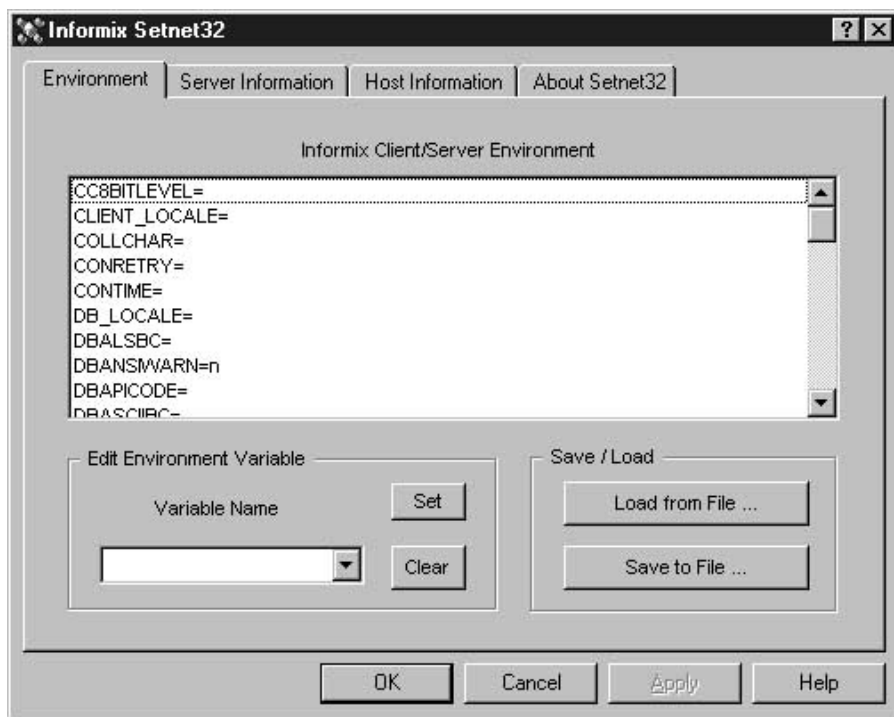


3. 从 **Informix 服务器 (Informix Server)** 列表选择现存 Informix 数据库服务器，或在 **Informix 服务器 (Informix Server)** 字段中输入新数据库服务器的名称。
4. 从主机名 (**HostName**) 下拉列表中选择数据库服务器的主机，或在主机名 (**HostName**) 字段中输入新主机的名称。
5. 从协议名 (**Protocolname**) 列表中选择必需的网络协议。
6. 在服务名 (**Service Name**) 字段中，输入主机上指定给数据库服务器的服务名或端口号。

必须在客户机工作站上的 Windows NT 安装目录中的 服务文件中定义该服务名。

设置仓库源

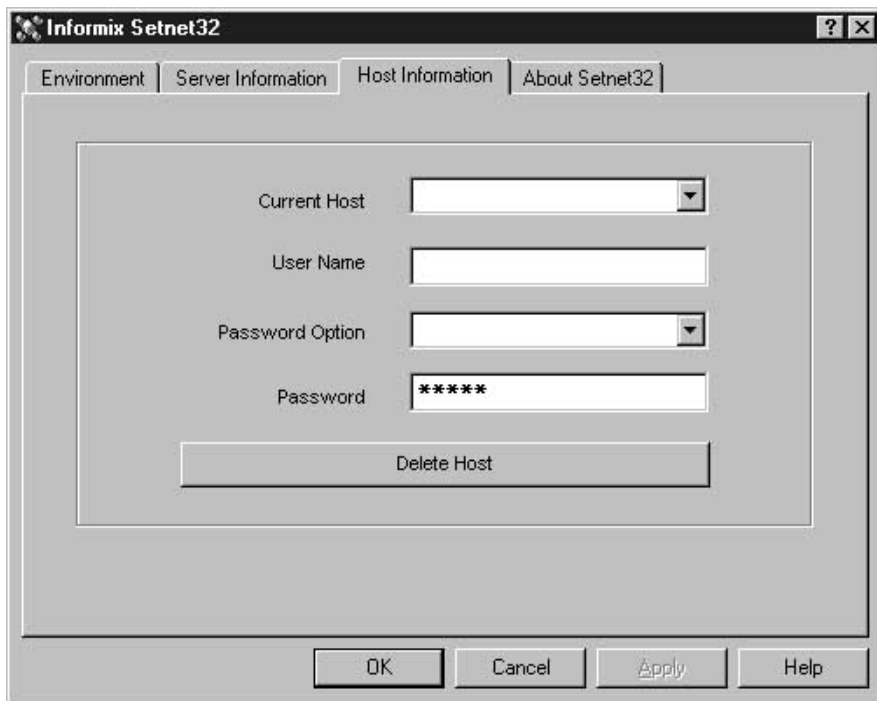
- 单击**环境 (Environment)** 选项卡。



- 在**变量名 (Variable Name)** 字段中输入 INFORMIXSERVER。
- 单击**设置 (Set)**。
- 单击**确定 (OK)**。

要配置 Informix 主机信息:

- 单击 Setnet32 窗口中的**主机信息 (Host Information)** 选项卡。



2. 从当前主机 (**Current Host**) 列表中选择想要建立网络连接的主机名，或在当前主机 (**Current Host**) 字段中输入一个名称以定义新主机名。
3. 确保当前主机 (**Current Host**) 字段中的用户名是所选主机上的帐户的用户名。
4. 在密码 (**Password**) 字段中输入密码。
5. 单击确定 (**OK**)。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Informix 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后，必须通过将 Informix 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 Windows NT 中为 ODBC 驱动程序注册 DSN:

1. 单击开始 → 设置 → 控制面板。
2. 双击 **ODBC 数据源**。

设置仓库源

3. 单击**系统 DSN** 选项卡。
4. 单击**添加**。
5. 选择想注册的 ODBC 驱动程序。
6. 单击**完成**。
驱动程序设置窗口打开。
7. 单击**一般事项**选项卡。
8. 在**数据源名字段**中输入数据库别名。
9. 在**数据库描述**字段中输入数据库的描述。
10. 在**数据库名字段**中输入数据库的名称。
11. 单击**连接**选项卡。
12. 在**缺省用户名**字段中输入用户标识。
13. 在**主机名字段**中输入服务器的名称。
14. 在**服务名字段**中输入服务名。
15. 从**协议类型**列表中选择 **onsoctcp**。
16. 单击**确定**。
17. 从**系统数据源**窗口中选择希望的数据库别名。
18. 单击**确定**。
19. 关闭 ODBC 窗口。

Sybase

本节包含执行下列任务的步骤:

- 配置 Sybase 客户机
- 安装和配置 ODBC 驱动程序

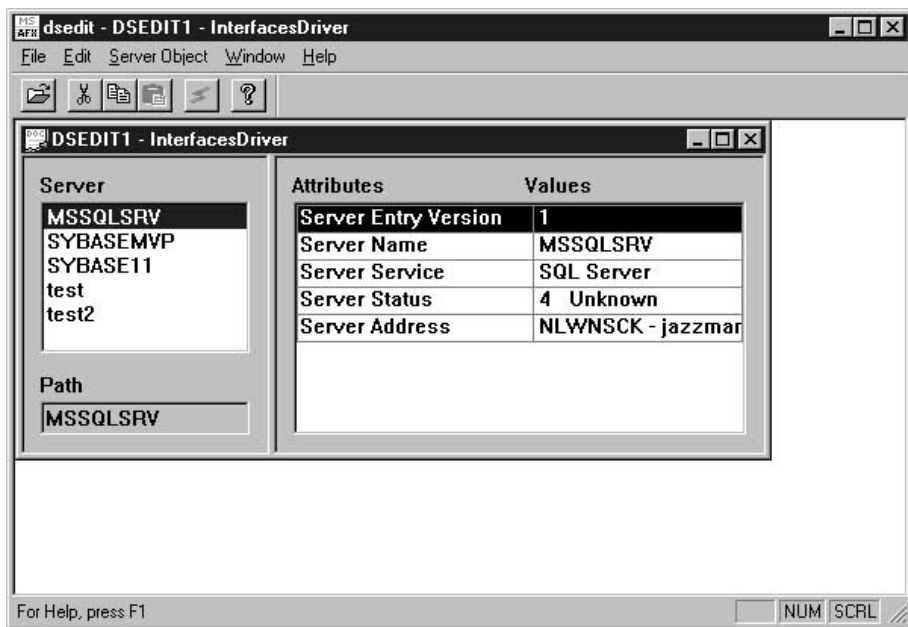
配置 Sybase 客户机: 要配置 Sybase 客户机:

1. 单击**开始** → **程序** → **Sybase Windows NT 版** → **DSEdit**。
DSEdit 窗口打开。



2. 从 **DS 名 (DS Name)** 列表中选择 DS 名。
3. 单击**确定 (OK)**。

InterfacesDriver 窗口打开。



4. 从“服务器” (Server) 窗口的列表中选择想要配置的服务器。
5. 若未列示想要使用的服务器，则给列表添加新的服务器。

要将新的服务器添加到服务器列表中:

- a. 在“服务器”(Server)窗口中右键单击。
 - b. 单击**添加 (Add)**。
 - c. 在**名称 (Name)**字段中输入服务器的名称。
在“服务器”(Server)窗口的右屏面中,将显示下列服务器属性:“项目版本”(Entry version)、“服务器名”(Server name)、“服务器服务”(Server service)、“服务器状态”(Server status)和“服务器地址”(Server address)。应保留“项目版本”(Entry version)、“服务器服务”(Server service)和“服务器状态”(Server status)的缺省值。
 - d. 选择**服务器地址 (Server address)**属性。
 - e. 从**服务器对象 (Server Object)**菜单中选择**修改属性 (Modify Attribute)**。
任选地,可双击该属性,或右键单击该属性并选择**修改属性 (Modify Attribute)**。
 - f. 从协议列表中选择 **NLWNSCK**。
若想使用的协议不在列表中,可:
 - 1) 单击**添加 (Add)**。
 - 2) 从**协议 (Protocol)**列表中选择协议类型。
 - 3) 在**网络地址 (Network Address)**字段中输入服务器地址和端口号。
 - 4) 单击**确定 (OK)**。
 - g. 验证添加的服务器已突出显示。
 - h. 单击**确定 (OK)**。
 - i. 验证想配置的服务器已突出显示。
6. 关闭 **DSEDIT** 程序。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

要启用对 Sybase 的定界标识符支持,可选择“ODBC 驱动程序设置”笔记本的“高级”页中的**启用引用的标识符**复选框。

若没有访问 Sybase 数据库所需的 ODBC 驱动程序,可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息,参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后,必须通过将 Sybase 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 Windows NT 中为 ODBC 驱动程序注册 DSN:

1. 单击开始 → 设置 → 控制面板。
2. 双击**ODBC 数据源**。
3. 单击系统 **DSN** 选项卡。
4. 单击**添加**。
5. 选择想注册的 ODBC 驱动程序。
6. 单击**完成**。
驱动程序设置窗口打开。
7. 单击**一般事项**选项卡。
8. 在**数据源名**字段中输入数据库别名。
9. 在**数据库描述**字段中输入数据库的描述。
10. 在**服务器名**字段中输入服务器的名称。
11. 在**数据库名**字段中输入数据库的名称。
12. 单击**连接**选项卡。
13. 在**缺省登录标识**字段中输入登录标识。
14. 在**工作站标识**字段中输入工作站的标识。
15. 单击**确定**。
16. 从“系统数据源”窗口中选择希望的数据库别名。
17. 单击**确定**。
18. 关闭 ODBC 窗口。

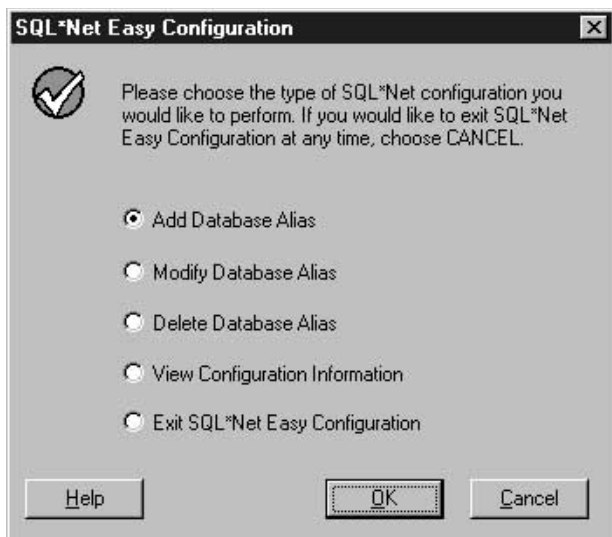
Oracle

本节包含执行下列任务的步骤:

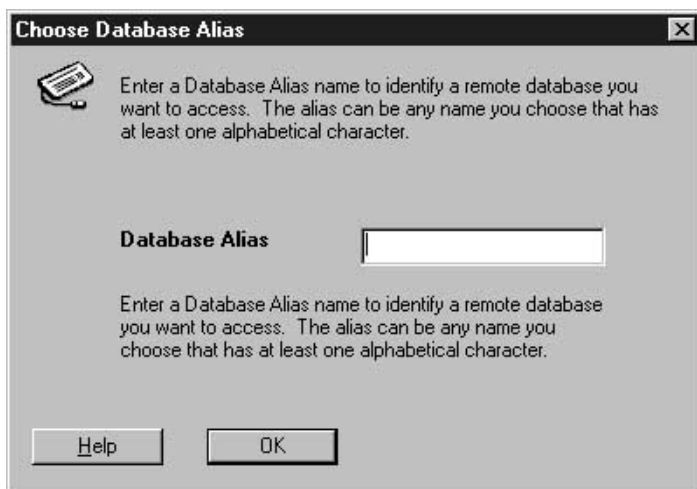
- 配置 Oracle 客户机
- 安装和配置 ODBC 驱动程序

配置 Oracle 7 客户机: 要配置 Oracle 7 客户机:

1. 单击开始 → 程序 → **Oracle Windows NT 版** → **SQL Net 轻松配置**。
“SQL Net 轻松配置”(SQL Net Easy Configuration) 窗口打开并显示使用客户机配置的几个选项。

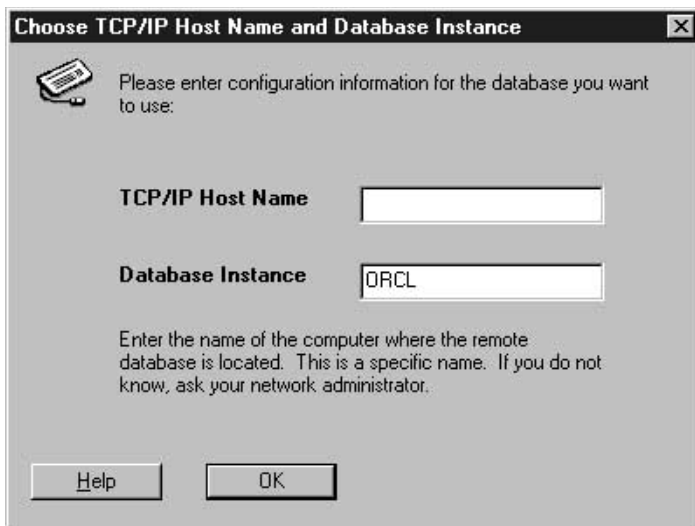


2. 单击想要的客户机配置选项旁边的单选按钮。
可添加新的客户机配置，或更改或查看现存配置。
3. 若单击**添加数据库别名 (Add database alias)**:
 - a.



单击**确定 (OK)**。

- b. 在**数据库别名 (Database alias)** 字段中输入数据库别名。
4. 单击**确定 (OK)**。



5. 在 **TCP/IP 主机名 (TCP/IP host name)** 字段中输入 TCP/IP 主机名。
6. 在**数据库实例 (Database instance)** 字段中输入数据库实例，或使用缺省值。
7. 单击**确定**。

确认窗口打开。

8. 复查确认窗口中的信息。
9. 若信息正确，单击**是 (Yes)**。

若确认窗口中的信息不正确：

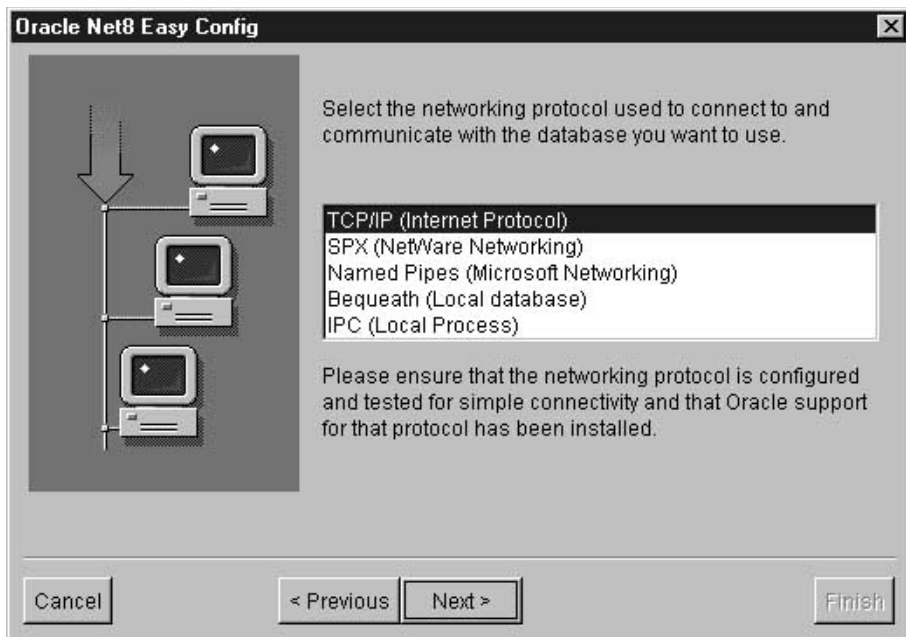
- a. 单击**上一步 (Back)**，直到到达具有错误消息的窗口。
- b. 校正信息。
- c. 重复步骤 7 至 9。

配置 Oracle 8 客户机： 要配置 Oracle 8 客户机：

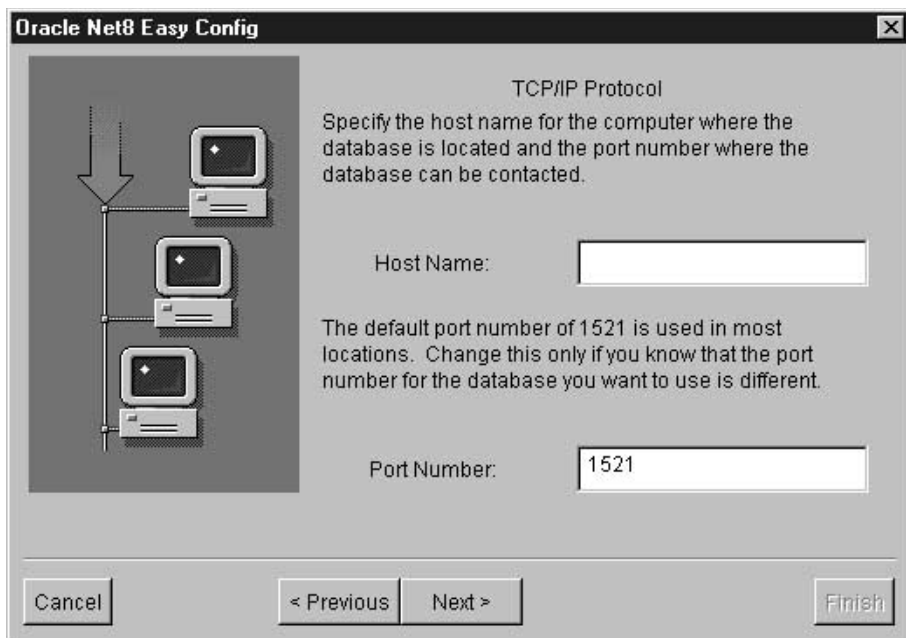
1. 单击**开始** → **程序** → **Oracle Windows NT 版** → **SQL Net8 轻松配置**。
“Oracle Net8 轻松配置” (Oracle Net8 Easy Configuration) 窗口打开并显示使用客户机配置的几个选项。



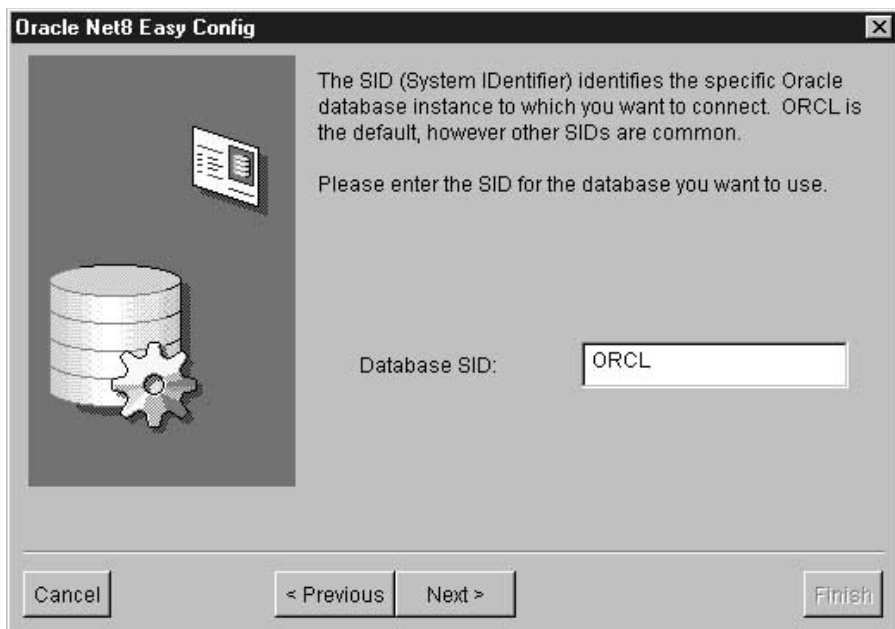
2. 单击想要的客户机配置选项旁边的单选按钮。
可添加新的客户机配置，或更改或查看现存配置。
3. 若单击添加 (**Add**)，必须还在**新服务名 (New Service Name)** 字段中输入数据库别名。
4. 单击下一步 (**Next**)。
5. 从“协议” (Protocol) 窗口的列表中选择联网协议的类型。



- 单击下一步 (**Next**)。
- 在“TCP/IP 协议” (TCP/IP Protocol) 窗口的**主机名 (Host Name)** 字段中输入 TCP/IP 主机名。



8. 在“TCP/IP 协议” (TCP/IP Protocol) 窗口的端口号 (**Port Number**) 字段中输入 TCP/IP 端口号。
9. 单击下一步 (**Next**)。
10. 在数据库 **SID (Database SID)** 字段中输入数据库 SID。



11. 单击下一步 **(Next)**。
“测试” (Test) 窗口打开。
12. 单击**测试 (Test)** 以测试配置。
若不想测试该配置，单击下一步 **(Next)**。
若单击**测试 (Test)**:
 - a. 在**用户标识 (User ID)** 字段中输入用户标识。
 - b. 在**密码 (Password)** 字段中输入密码。
 - c. 单击**测试 (Test)**。
13. 单击下一步 **(Next)**。
14. 单击**完成 (Done)**。
确认窗口打开。
15. 复查确认窗口中的信息。
16. 若确认窗口中的信息正确，单击**完成 (Finish)**。
若确认窗口中的信息不正确:
 - a. 单击**前一屏 (Previous)**，直到到达具有错误消息的窗口。
 - b. 校正信息。
 - c. 重复步骤 9 至 16。

设置仓库源

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Oracle 数据库所需的 ODBC 驱动程序, 可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息, 参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后, 必须通过将 Oracle 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 Windows NT 中为 ODBC 驱动程序注册 DSN:

1. 单击开始 → 设置 → 控制面板。
2. 双击**ODBC 数据源**。
3. 单击**系统 DSN** 选项卡。
4. 单击**添加**。
5. 选择想注册的 ODBC 驱动程序。
6. 单击**完成**。
驱动程序设置窗口打开。
7. 选择**一般事项**选项卡。
8. 在**数据源名字段**中输入数据库别名。
9. 在**数据库名字段**中输入数据库的名称。
10. 在**服务器名字段**中输入服务器的名称。
11. 单击**确定**。
12. 从“系统数据源”窗口中选择希望的数据源别名。
13. 单击**确定**。
14. 关闭 ODBC 窗口。

Microsoft SQL Server

本节包含执行下列任务的步骤:

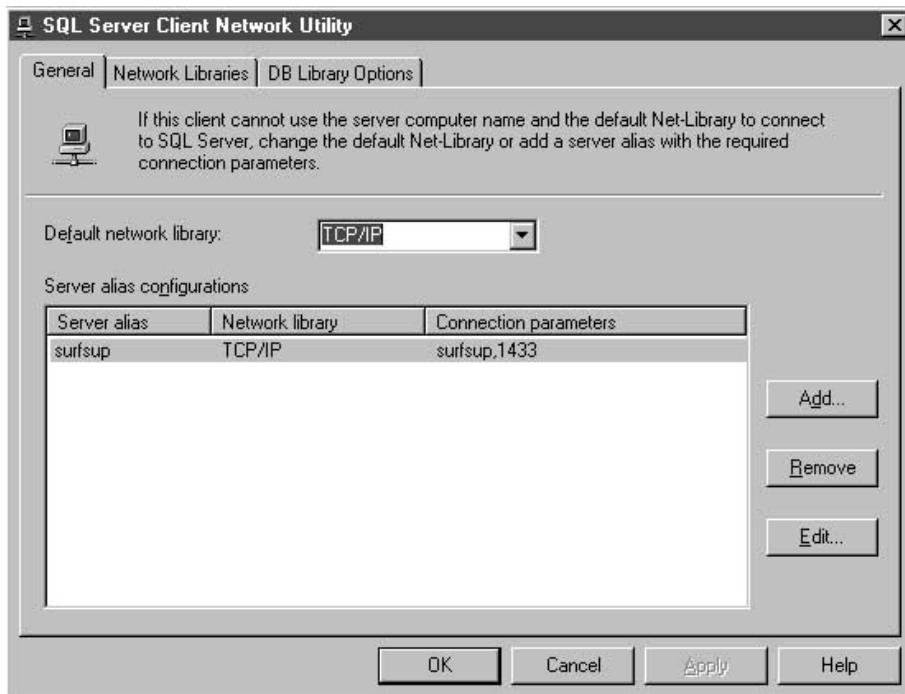
- 配置 Microsoft SQL Server 客户机
- 安装和配置 ODBC 驱动程序

配置 Microsoft SQL Server 客户机: 要设置对 Microsoft SQL Server 客户机的访问权, 必须使用“Microsoft SQL Server 客户机网络实用程序”配置 Microsoft SQL Server 客户机软件。

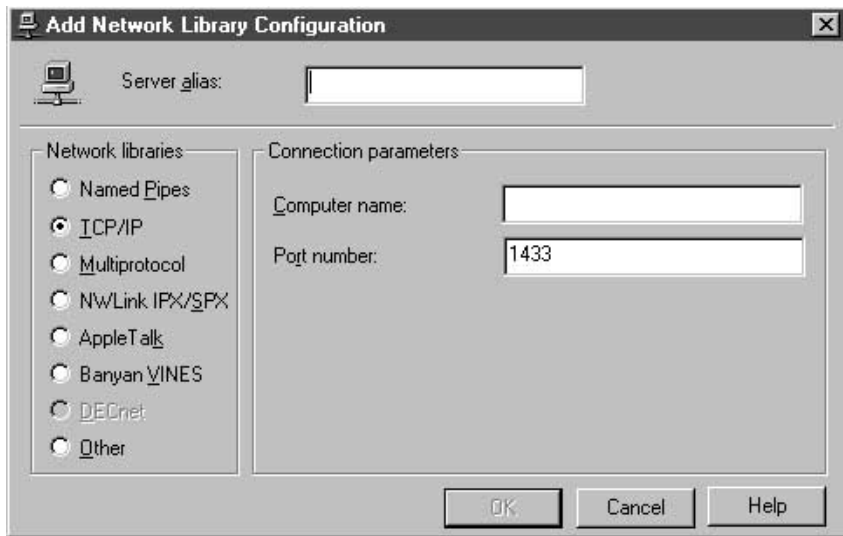
要配置 Microsoft SQL Server 客户机:

1. 单击开始 → 程序 → **Microsoft SQL Server Windows NT 版** → 客户机网络实用程序。

“SQL Server 客户机网络实用程序” (SQL Server Client Network Utility) 窗口打开。



2. 在缺省网络库 (**Default network library**) 列表中, 单击 **TCP/IP**。
3. 从服务器别名配置 (**Server alias configurations**) 列表中选择服务器名。任选地, 单击添加 (**Add**) 以将新服务器添加到列表中。“添加网络库配置” (Add Network Library Configuration) 窗口打开。



- a. 在服务器别名 (**Server alias**) 字段中输入服务器的名称。
 - b. 在计算机名 (**Computer name**) 字段中输入服务器的名称。
此名称应是在服务器别名 (**Server alias**) 字段中输入的同一名称。
 - c. 在端口号 (**Port number**) 字段中输入端口号。
 - d. 单击**确定 (OK)** 关闭“添加网络库配置”(Add Network Library Configuration) 窗口。
 - e. 从服务器别名配置 (**Server alias configurations**) 列表中选择刚才创建的服务器名。
4. 单击**确定 (OK)**。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

如果使用 Merant ODBC 驱动程序来访问 Microsoft SQL Server，应验证系统路径是否包含 `sqllib\odbc32` 目录。

要启用对 Microsoft SQL Server 的定界标识符支持，可选择“ODBC 驱动程序设置”笔记本的“高级”页中的启用引用的标识符复选框。

若没有访问 Microsoft SQL Server 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后，必须通过将 Microsoft SQL Server 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

设置仓库源

要在 Windows NT 中为 ODBC 驱动程序注册 DSN:

1. 单击**开始** → **设置** → **控制面板**。
2. 双击**ODBC 数据源**。
3. 单击**系统 DSN** 选项卡。
4. 单击**添加**。
5. 选择想注册的 ODBC 驱动程序。
6. 单击**完成**。
“ODBC SQL 服务器驱动程序设置”窗口打开。
7. 单击**一般事项**选项卡。
8. 在**数据源名**字段中输入数据库别名。
9. 在**描述**字段中输入数据库的描述。
10. 在**服务器名**字段中输入想要使用的服务器的名称。
11. 在**数据库名**字段中输入想要使用的数据库的名称。
12. 单击**高级**选项卡。
13. 在**缺省登录**字段中输入用户名。
14. 在**工作站标识**字段中输入工作站标识。
15. 单击**确定**。
16. 从“系统数据源”窗口中选择希望的数据库别名。
17. 单击**确定**。
18. 关闭 ODBC 窗口。

Microsoft Access

本节包含执行下列任务的步骤:

- 创建和编目 Microsoft Access 数据库
- 创建和编目目标仓库数据库
- 定义使用该数据库的仓库

创建和编目 Microsoft Access 数据库: 要创建 Microsoft Access 数据库:

1. 打开 Microsoft Access。
2. 单击**数据库向导**。
3. 单击**确定**。
“新建”窗口打开，并显示数据库模板。
4. 选择一个模板。
5. 单击**确定**。

6. 在文件名字段中输入数据库的名称。
7. 单击**创建**。
“数据库向导”打开。
8. 按照提示，并单击**完成**以创建数据库。
记下该数据库的路径和文件名，因为以后您将使用它。
9. 创建表，并向表中输入数据。

要将该数据库中编目在 ODBC 中：

1. 单击**开始** → **设置** → **控制面板**。
2. 双击 **ODBC**。
3. 单击**系统 DSN**。
若没有**系统 DSN**按钮，参见 *DB2 Universal Database Troubleshooting Guide*。
4. 单击**添加**。
5. 从已安装的 **ODBC 驱动程序**列表中选择 **Microsoft Access 驱动程序**。
6. 单击**确定**。
7. 在**数据源名**字段中输入数据库别名。
8. 可选：在**描述**字段中输入数据库的描述。
9. 单击**选择**。
10. 从列表框中选择该数据库的路径和文件名。
11. 单击**确定**。
12. 单击**高级**。
13. 在**登录名**字段中输入用户标识。
14. 在**密码**字段中输入密码。
15. 在**选项**列表中，验证 **FIL** 的值是否为 **MS Access**。
16. 在“设置高级选项”窗口中单击**确定**。
17. 在“ODBC Microsoft Access 设置”窗口中单击**确定**。
18. 单击**关闭**。

创建和编目目标仓库数据库： 要在 DB2 中创建目标仓库数据库：

1. 通过单击**开始** → **程序** → **IBM DB2** → **控制中心**，启动“DB2 控制中心”。
2. 右键单击**数据库**文件夹，并单击**使用向导创建数据库**。“创建数据库向导”打开。
3. 在**数据库名**字段中输入数据库的名称。

设置仓库源

4. 从**缺省驱动器**列表中，选择数据库的驱动器。
5. 可选：在**注释**字段中，输入数据库的描述。
6. 单击**完成**。此向导中的所有其他字段和页是可选的。创建数据库并将它列示在“DB2 控制中心”中。

要将目标仓库数据库编目在 ODBC 中：

1. 单击**开始** → **设置** → **控制面板**。
2. 双击 **ODBC**。
3. 单击**系统 DSN**。

若没有**系统 DSN** 按钮，参见 *DB2 Universal Database Troubleshooting Guide*。

4. 单击**添加**。
5. 在已安装的 **ODBC** 驱动程序列表中，单击 **IBM DB2 ODBC** 驱动程序。
6. 单击**确定**。
7. 在**数据源名**字段中输入数据库别名。
8. 可选：在**描述**字段中输入数据库的描述。
9. 单击**选择**。
10. 从列表框中选择该数据库的路径和文件名。
11. 单击**确定**。
12. 单击**关闭**。

定义使用数据库的仓库： 要为创建的数据库创建“数据仓库中心”定义：

1. 按照第95页的『在“数据仓库中心”中定义非 DB2 数据库仓库源』中的指示信息，为 Microsoft Access 数据库创建仓库源。为列示的参数指定下列值：
 - 在“数据库”页上：
 - 选择**定制 ODBC 连接字符串**复选框。
 - 在 **ODBC 连接字符串**字段中输入下列字符串：

```
DSN=database-alias;UID=userID;PWD=password;
```

database-alias
在 ODBC 中注册 Microsoft Access 数据库所用的名称。

userID 在 ODBC 中注册 Microsoft Access 数据库时指定的用户标识。

password
在 ODBC 中注册 Microsoft Access 数据库时指定的密码。
 - 在“代理站点”页上，指定您注册 Microsoft Access 源数据库和 DB2 仓库数据库所在的代理站点。

- 在“表”页上:
 - a. 展开表文件夹以便从 Microsoft Access 数据库导入表定义。
 - b. 在“过滤器”窗口上单击**确定**以列示数据库中的所有表。
验证在数据库中创建的表是否在**可用的表**列表中。
 - c. 从**可用的表**列表中选择要使用的表。
 - d. 单击 **>**。
该表移动至**选择的选表**列表。

注:

Microsoft Access 数据库的 Memo 字段在“数据仓库中心”源中表示为数据类型 LONG VARCHAR, 其列大小超过了 1 GB。要支持实际的系统配置, “数据仓库中心”会截断超过 128 KB 的值。要避免截断仓库源中 Memo 字段的值, 应在使用步骤中的表前, 将接收 Memo 字段数据的列的数据类型从 LONG VARCHAR 更改为 CLOB。如果不更改列的数据类型, 任何大于 128 KB 的值都将被截断。

对 CLOB 数据类型的 DRDA 支持是 OS/390 和 OS/400 所必需的。对 CLOB 数据类型的 OS/390 支持是从 DB2 版本 6 开始的。对 CLOB 数据类型的 OS/400 支持是从带有 DB 修订包 4 或更高版本 (PTF SF99104) 的修订包的版本 4, 发行版 4 开始的。对于 OS/400, 日期为 1999 年 2 月的安装磁盘版本 4, 发行版 4 也包含了对 CLOB 数据类型的支持。

2. 按照第101页的『第4章 设置对仓库的访问权』中的指示信息, 为 DB2 数据库创建仓库。
3. 按照第122页的『添加步骤到过程中』中的指示信息创建步骤。给予该步骤下列属性:
 - 它使用来自 Microsoft Access 数据库的仓库源的一个或多个源表。
 - 它在 DB2 仓库数据库中创建目标表。
4. 将步骤提升为测试方式。
5. 通过右键单击步骤并单击**测试**来运行该步骤。
6. 验证在 Microsoft Access 数据库中创建的数据是否在仓库数据库中。在“DB2 命令行处理器”窗口中输入如下命令:


```
select * from prefix.database-name
```

prefix 仓库数据库的前缀 (如 IWH)。

设置仓库源

database-name

仓库数据库的名称。

您应该看到您在 Microsoft Access 数据库中输入的数据。

Microsoft Excel

本节包含执行下列任务的步骤:

- 创建和编目 Microsoft Excel 数据库
- 创建和编目目标仓库数据库
- 定义使用该数据库的仓库

创建和编目 Microsoft Excel 数据库: 要创建 Microsoft Excel 数据库:

1. 打开 Microsoft Excel。
2. 单击**文件** → **新建**。
“新建”窗口打开,并显示电子表格模板。
3. 选择一个模板。
4. 单击**确定**。
5. 单击**文件** → **保存**以保存该电子表格。
6. 在**文件名**字段中输入电子表格的名称。
7. 单击**保存**。

记下该电子表格的路径和文件名,因为以后您将使用它。

8. 将数据输入该电子表格中。

要将该数据库中编目在 ODBC 中:

1. 单击**开始** → **设置** → **控制面板**。
2. 双击 **ODBC**。
3. 单击**系统 DSN**。
若没有**系统 DSN**按钮,参见 *DB2 Universal Database Troubleshooting Guide*。
4. 单击**添加**。
5. 从已安装的 **ODBC** 驱动程序列表中选择 **Microsoft Excel** 驱动程序。
6. 单击**确定**。
7. 在**数据源名**字段中输入数据库别名。
8. 可选: 在**描述**字段中输入数据库的描述。
9. 从**版本**列表中选择 **Excel 97**。
10. 单击**选择工作簿**。

11. 从列表框中选择该数据库的路径和文件名。
12. 单击**确定**。
13. 在“ODBC Microsoft Excel 设置”窗口中单击**确定**。
14. 单击**关闭**。

若正使用 Microsoft Excel 95/97 ODBC 驱动程序访问 Excel 电子表格，需要为电子表格内的每个工作表创建一个命名的表。要为每个工作表创建命名的表：

1. 选择想要包括在表中的列和行。
2. 单击**插入** → **名称** → **定义**。
3. 确保“定义名称”窗口的**参考**中包含了已在“步骤 1”中选择的单元。如果不包含，则单击**参考**字段最右边的图标，以包括所选择的所有单元。
4. 为标记的数据输入名称（或使用缺省名称）。
5. 单击**确定**。

现在可在定义仓库源时导入表，而无需标记**包括系统表**复选框。

创建和编目目标仓库数据库： 要在 DB2 中创建目标仓库数据库：

1. 通过单击**开始** → **程序** → **IBM DB2** → **控制中心**，启动“DB2 控制中心”。
2. 右键单击**数据库**文件夹，并单击**使用向导创建数据库**。“创建数据库向导”打开。
3. 在**数据库名字**段中输入数据库的名称。
4. 从**缺省驱动器**列表中，选择数据库的驱动器。
5. 在**注释**字段中，输入数据库的描述。
6. 单击**完成**。此向导中的所有其他字段和页是可选的。创建数据库并将它列示在“DB2 控制中心”中。

要将目标仓库数据库编目在 ODBC 中：

1. 单击**开始** → **设置** → **控制面板**。
2. 双击 **ODBC**。
3. 单击**系统 DSN**。

若没有**系统 DSN**按钮，参见 *DB2 Universal Database Troubleshooting Guide*。

4. 单击**添加**。
5. 从已安装的 **ODBC 驱动程序**列表中选择 **IBM DB2 ODBC 驱动程序**。
6. 单击**确定**。
7. 在**数据源名字**段中输入数据库别名。

设置仓库源

8. 可选: 在**描述**字段中输入数据库的描述。
9. 单击**选择**。
10. 从列表框中选择该数据库的路径和文件名。
11. 单击**确定**。
12. 单击**关闭**。

定义使用数据库的仓库: 要为创建的数据库创建“数据仓库中心”定义:

1. 按照第95页的『在“数据仓库中心”中定义非 DB2 数据库仓库源』中的指示信息, 为 Microsoft Excel 电子表格创建仓库源。为列示的参数指定下列值:
 - 在“数据库”页上:
 - 选择定制 **ODBC 连接字符串** 复选框。
 - 在 **ODBC 连接字符串** 字段中输入下列字符串:
`DSN=database-alias;UID=userID;PWD=password;`
database-alias
在 ODBC 中注册 Microsoft Excel 电子表格所用的名称。
userID 在 ODBC 中注册 Microsoft Excel 电子表格时指定的用户标识。
password
在 ODBC 中注册 Microsoft Excel 电子表格时指定的密码。
 - 在“代理站点”页上, 指定注册 Microsoft Excel 源电子表格和 DB2 仓库数据库所在的代理站点。
 - 在“表”页上:
 - a. 展开**表**文件夹以便从 Microsoft Excel 电子表格导入表定义。
 - b. 在“过滤器”窗口上单击**确定**以列示电子表格中的所有表。
验证在数据库中创建的表是否在**可用的表**列表中。
 - c. 从**可用的表**列表中选择要使用的表。
 - d. 单击 **>**。
该表移动至**选择的表**列表。
2. 按照第101页的『第4章 设置对仓库的访问权』中的指示信息, 为 DB2 数据库创建仓库目标。
3. 按照第122页的『添加步骤到过程中』中的指示信息创建步骤。给予该步骤下列属性:
 - 它使用来自 Microsoft Excel 电子表格的仓库源的一个或多个源表。
 - 它在 DB2 仓库数据库中创建目标表。
4. 将步骤提升为测试方式。

5. 通过右键单击步骤并单击**测试**来运行该步骤。
6. 验证在 Microsoft Access 数据库中创建的数据是否在仓库数据库中。在“DB2 命令行处理器”窗口中输入如下命令：

```
select * from prefix.database-name
```

prefix 仓库数据库的前缀（如 IWH）。

database-name

仓库数据库的名称。

您应该看到您在 Microsoft Access 数据库中输入的数据。

IMS 和 VSAM

若仓库源使用 IMS 或 VSAM 数据库中的操作数据，则在“数据仓库中心”中使用 Classic Connect。使用 Classic Connect 将非关系数据映射为伪关系数据。然后使用 CROSS ACCESS ODBC 驱动程序访问伪关系数据。然后可在“数据仓库中心”中定义对应伪关系数据的 IMS 或 VSAM 仓库源。

有关在“数据仓库中心”中使用 Classic Connect 的详情，参见第363页的『附录F. 在“数据仓库中心”中使用 Classic Connect』。

在 AIX 或 Solaris 操作环境中设置非 DB2 数据库仓库源

以下各节描述如何在 AIX 或 Solaris 操作环境中设置 Informix、Sybase、Oracle 和 Microsoft SQL Server 源。

Informix

本节包含执行下列任务的步骤：

- 配置 Informix 客户机
- 安装和配置 ODBC 驱动程序

配置 Informix 客户机： 要在 AIX 或 Solaris 操作环境中配置 Informix 客户机：

1. 打开 sqlhosts 文件。
2. 添加新的项目列表到该文件。
为保持相同格式，复制并粘贴示例列表。然后如下列步骤所述更改项目信息。
3. 输入数据库名。
4. 输入协议类型 onsoctcp。
5. 输入主机名。
6. 输入端口名。

端口名必须与客户机上 Windows 安装目录中的 /etc/services 文件中定义的端口名相同。

设置仓库源

图5显示包含新项目列表的 `sqlhosts` 文件的示例。

```
# Informix V5
database1    olsoctcp    test0      ifmxfrst1
database2    olsoctcp    test0      ifmxfrst2
```

图 5. 已完成的 `sqlhosts` 文件项目列表

安装和配置 ODBC 驱动程序： ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Informix 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后，必须设置对 Informix 数据库的访问权。

要设置对 Informix 数据库的访问权：

1. 打开 `.odbc.ini` 文件。
2. 在文件的顶部定义 Informix 数据库别名。
3. 将新项目添加到该文件。
4. 输入 ODBC 驱动程序的目录位置。
5. 输入数据库的描述。
6. 输入数据库名。
7. 输入服务器主机名。
8. 输入登录标识。
9. 输入该登录标识的密码。
10. 输入服务器名。
11. 输入服务名。
12. 保存并关闭该文件。

第87页的图6显示 Informix 数据库已完成的项目的示例。

```
[INF72]
Driver=/home/db2_07_01/3.6/odbc/lib/ivinf12.so
Description=Informix7.23
Database=test7
HostName=xxyy.zzz.ibm.com
LoginID=informix
Password=password
ServerName=ifmx72
Service=ifmxon72
```

图 6. 一个完整的 Informix 数据库的 .odbc.ini 条目

Sybase

本节包含执行下列任务的步骤:

- 配置 Sybase 客户机
- 安装和配置 ODBC 驱动程序

配置 Sybase 客户机: 要在 AIX 或 Solaris 操作环境中配置 Sybase 客户机:

1. 打开 interfaces 文件。
2. 添加新的项目列表到该文件。
为保持相同格式，复制并粘贴示例列表。然后如下列步骤所述更改项目信息。
3. 输入想要用于 CONNECT 语句的数据库别名。
4. 输入查询。
5. 输入要使用的协议。
6. 输入服务器主机名。
7. 输入端口号。

图7显示包含新项目列表的 interfaces 文件的示例。

```
Sybase11
    query tcp ether superman 2000
```

图 7. 已完成的 interfaces 文件项目列表

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Sybase 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

设置仓库源

安装 ODBC 驱动程序之后，必须设置对 Sybase 数据库的访问权。

要设置对 Sybase 数据库的访问权：

1. 打开 `.odbc.ini` 文件。
2. 在文件的顶部定义 Sybase 数据库别名。
3. 将新项目添加到该文件。
4. 输入 ODBC 驱动程序的目录位置。
5. 输入数据库的描述。
6. 将数据库类型定义为 **master**。
7. 输入服务器名。
8. 输入登录标识。
9. 输入该登录标识的密码。
10. 输入 `interfaces` 文件的目录位置。
11. 可选：要启用定界标识符支持，要编辑 Sybase 数据源以包括连接属性 `EQI=1`。
12. 保存并关闭该文件。

图8显示 Sybase 数据库的已完成项目的示例。

```
[Sybase10]
Driver=/home/db2_07_01/3.6/odbc/lib/ivsyb1112.so
Description=Sybase 10 ODBC Database
Database=master
ServerName=Sybase10
LogonID=sybase
Password=password
InterfacesFile=/public/sdt_lab/sybase/AIX/System10/interfaces
```

图 8. Sybase 数据库的已完成的 `.odbc.ini` 项目

Oracle

本节包含执行下列任务的步骤：

- 配置 Oracle 客户机
- 安装和配置 ODBC 驱动程序

配置 Oracle 客户机： 要在 AIX 或 Solaris 操作环境中配置 Oracle 客户机：

1. 打开 `tnsnames.ora` 文件。
2. 添加新的项目列表到该文件。

为保持相同格式，复制并粘贴示例列表。然后如下列步骤所述更改项目信息。

3. 输入要使用的协议。
4. 输入服务器主机名。
5. 输入端口名。
6. 输入 SID。

图9显示包含新项目列表的 `tnsnames.ora` 文件的示例。

```
# Oracle 8.1.5
Oracle8i=
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS =
        (PROTOCOL = TCP)
        (HOST = superman)
        (PORT = 2000)
      )
    )
  (CONNECT_DATA =
    (SID=oracle8i)
  )
)
```

图9. 已完成的 `tnsnames.ora` 文件项目列表

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Oracle 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后，必须设置对 Oracle 数据库的访问权。

要设置对 Oracle 数据库的访问权:

1. 打开 `.odbc.ini` 文件。
2. 在文件的顶部定义 Oracle 数据库别名。
3. 将新项目添加到该文件。
4. 输入 ODBC 驱动程序的目录位置。
5. 输入服务器名。
6. 输入数据库的描述。
7. 保存并关闭该文件。

图10显示 Oracle 数据库的已完成项目的示例。

```
[Oracle_8]
Driver=/home/db2_07_01/3.6/lib/ivor814.so
ServerName=Oracle8
Description=Oracle 8 ODBC Database
```

图 10. Informix 数据库的已完成的 .odbc.ini 项

Microsoft SQL Server

本节包含执行下列任务的步骤:

- 配置 Microsoft SQL Server 客户机
- 安装和配置 ODBC 驱动程序

配置 Microsoft SQL Server 客户机: 要在 AIX 或 Solaris 操作环境中配置 Microsoft SQL Server 客户机, 在 AIX 或 Solaris 操作环境客户机上和安装 Microsoft SQL Server 的 Windows 服务器上都配置 TCP/IP。因为“Microsoft SQL Server 客户机”配置在 INI 文件中与 ODBC 驱动程序配置集成在一起, 所以无需任何其他操作来配置您的客户机。有关详情, 参见您的操作系统的 *Merant DataDirect Connect ODBC Reference*。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Microsoft SQL Server 数据库所需的 ODBC 驱动程序, 可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息, 参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后, 必须设置对 Microsoft SQL Server 数据库的访问权。

要设置对 Microsoft SQL Server 数据库的访问权:

1. 打开 .odbc.ini 文件。
2. 在文件的顶部定义 Microsoft SQL Server 数据库别名。
3. 将新项目添加到该文件。
4. 输入运行 Microsoft SQL Server 的服务器的网络地址。
5. 输入驱动程序应使用的 ANSI 定义的行为的首选项。
6. 输入数据库名。
7. 输入用户标识。
8. 输入该用户标识的密码

9. 输入用于识别 SQL 语句中的引号的首选项。
10. 输入与数据库配合使用的 TDS 的版本。
11. 若正使用 SQL Server 6.5, 输入用于创建 SQL Prepare 的临时存储过程的首选项。
12. 保存并关闭该文件。

有关配置 ODBC 驱动程序和 .odbc.ini 值和属性的详情, 参见您的操作系统的 *Merant DataDirect Connect ODBC Reference*。

图11显示 Microsoft SQL Server 数据库的样本项示例。

```
[MSSQL6.5]
Driver=/home/db2_07_01/3.6/odbc/lib/ivmsss14.so
Address=xyyyy.zzz.ibm.com
AnsiNPW=yes
Database=test7
UID=MSSQL
PWD=password
QuotedID=no
TDS=4.2
UseProcForPrepare=1
```

图 11. Microsoft SQL Server 数据库的样本 .odbc.ini 条目

在 OS/2 上设置非 DB2 数据库仓库源

Informix

本节包含执行下列任务的步骤:

- 配置 Informix 客户机
- 安装和配置 ODBC 驱动程序

配置 Informix 客户机: 要在 OS/2 上配置 Informix 客户机:

1. 双击 **Informix** 图符。
2. 双击 **Setnet** 图符。
 - “Informix-Net OS/2 版实用程序”窗口打开。
3. 在**主机名**字段中输入主机名。
4. 在**用户名**字段中输入用户名。
5. 在**服务名**字段中输入服务名。
6. 从**协议名**下拉列表中选择协议。

设置仓库源

7. 选择**密码**字段中密码选项旁边的单选按钮。
8. 单击**保存**。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Informix 数据库所需的 ODBC 驱动程序, 可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息, 参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后, 必须通过将 Informix 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 OS/2 中为 ODBC 驱动程序注册 DSN:

1. 单击**ODBC 管理员**图符。
若给 OS/2 代理安装了“ODBC 管理员”, 而不是更新现存的“ODBC 管理员”, 则该图符位于“数据仓库中心代理”文件夹中。
2. 选择想注册的 ODBC 驱动程序。
3. 单击**添加**。
4. 选择想注册的 ODBC 驱动程序, 然后单击**确定**。
5. 选择希望的数据库别名, 并单击**确定**。
6. 关闭 ODBC 窗口。

Sybase

本节包含执行下列任务的步骤:

- 配置 Sybase 客户机
- 安装和配置 ODBC 驱动程序

配置 Sybase 客户机: 要在 OS/2 上配置 Sybase 客户机:

1. 双击 **Sybase** 图符。
2. 双击 **SQL 编辑**图符。
“打开”窗口打开。
3. 选择要打开的文件。
4. 单击**打开**。
5. 从窗口中的列表中选择想要的服务器。
“服务器”窗口打开。
若未列示想要的服务器:

- a. 单击**编辑** → **添加服务**。
 - b. 在**名称**字段中输入想要配置的服务器的名称。
 - c. 单击**编辑** → **添加服务**。
必须设置主服务和查询服务。
 - d. 在**服务**列表中单击**主**。
 - e. 从**网络驱动程序**列表中选择想要的网络驱动程序。
 - f. 在**连接信息**字段中输入服务器的名称和端口号。
 - g. 单击**确定**。
 - h. 单击**编辑** → **添加服务**。
 - i. 在**服务**列表中单击**查询**。
 - j. 从**网络驱动程序**下拉列表中选择想要的网络驱动程序。
 - k. 在**连接信息**字段中输入服务器的名称和端口号。
 - l. 单击**确定**。
6. 单击**退出** → **退出编辑服务**。
 7. 关闭窗口并退出程序。

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Sybase 数据库所需的 ODBC 驱动程序，可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息，参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后，必须通过将 Sybase 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 OS/2 中为 ODBC 驱动程序注册 DSN:

1. 单击**ODBC 管理员**图符。
若给 OS/2 代理安装了“ODBC 管理员”，而不是更新现存的“ODBC 管理员”，则该图符位于“数据仓库中心代理”文件夹中。
2. 选择想注册的 ODBC 驱动程序。
3. 单击**添加**。
4. 选择想注册的 ODBC 驱动程序，然后单击**确定**。
5. 选择希望的数据库别名，并单击**确定**。
6. 关闭 ODBC 窗口。

Oracle

本节包含执行下列任务的步骤:

- 配置 Oracle 客户机
- 安装和配置 ODBC 驱动程序

配置 Oracle 客户机: 要在 OS/2 上配置 Oracle, 使用 SQL*Net OS/2 版客户机。

要配置 Oracle SQL*Net OS/2 版:

1. 转至 \oraos2\network\admin 目录。
2. 打开 TNSNAMES.ORA 文件。
3. 为您的数据库添加新的连接描述符到该文件。
连接描述符必须包括下列信息:
 - a. 描述符别名。
 - b. 协议适配器类型。
 - c. 端口号
 - d. 主机名。
 - e. 系统 ID。
4. 在每个连接描述符字段中输入您的客户机的信息。
5. 保存并关闭该文件。

图12显示包含新连接描述符项的 tnsnames.ora 文件的示例。

```
testtcp = (DESCRIPTION=
          (ADDRESS=
            (PROTOCOL=tcp)(PORT=1111)(HOST=oracle))
          (CONNECT_DATA=(SID=ORACLE7)
          ))
```

图 12. 已完成的 tnsnames.ora 连接描述符项

安装和配置 ODBC 驱动程序: ODBC 驱动程序用于注册“数据仓库中心”将访问的源、目标和控制数据库。

若没有访问 Oracle 数据库所需的 ODBC 驱动程序, 可使用“定制”安装选项从“DB2 通用数据库”光盘获取该驱动程序。有关如何选择和安装适当的驱动程序的信息, 参见您的操作系统的 *IBM DB2 通用数据库快速入门*。

安装 ODBC 驱动程序之后, 必须通过将 Oracle 数据库注册为 ODBC 中的系统数据库源名 (DSN) 来设置对该数据库的访问权。

要在 OS/2 中为 ODBC 驱动程序注册 DSN:

1. 单击**ODBC 管理员**图符。
若给 OS/2 代理安装了“ODBC 管理员”，而不是更新现存的“ODBC 管理员”，则该图符位于“数据仓库中心代理”文件夹中。
2. 选择想注册的 ODBC 驱动程序。
3. 单击**添加**。
4. 选择想注册的 ODBC 驱动程序，然后单击**确定**。
5. 选择希望的数据库别名，并单击**确定**。
6. 关闭 ODBC 窗口。

在“数据仓库中心”中定义非 DB2 数据库仓库源

要在“数据仓库中心”中定义非 DB2 仓库源，必须完成下列任务:

1. 在“数据仓库中心”管理客户机中打开“仓库源”笔记本。
2. 添加有关仓库源的信息。
3. 指定代理站点访问仓库源。
4. 添加仓库源的数据库信息。
5. 将源表和视图导入到仓库源中，或手工创建表。要获取关于创建源表的信息，参见联机帮助中的“定义仓库源表”。
6. 授权仓库组访问仓库源。
7. 指定代理站点访问仓库源。
8. 指定每隔多久重试数据抽取。

如果数据源是使用 Relational Connect 和 Create Nickname 语句创建的，该数据源将不可用于导入表到“数据仓库中心”中。要使用数据源作为源表，应在“数据仓库中心”定义一个非 DB2 的仓库源，但不要导入源表。必须手工创建表，并确保仓库源表中的列映射到数据源中的列。

以下过程适用于 Informix、Sybase、Oracle 和 Microsoft SQL Server 仓库源。

打开“仓库源”笔记本

1. 单击**开始** → **程序** → **IBM DB2** → **控制中心**。
“控制中心”打开。
2. 在“控制中心”工具栏中单击“数据仓库中心”图符。
“数据仓库中心管理客户机”打开。
3. 右键单击**仓库源**文件夹。

设置仓库源

- 单击**定义**，然后单击想要创建的“仓库”源的类型。
“仓库源”笔记本打开。

添加有关仓库源的信息

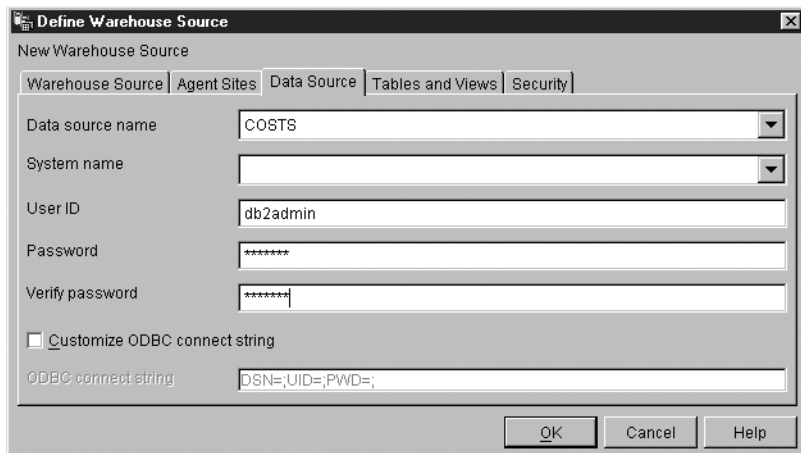
- 在“仓库源”页上，在**名称**字段中为仓库源输入一个业务名。取决于所选的源类型，笔记本中剩余的页可能会不可用。
将使用此名称在“数据仓库中心”中参考您的仓库源。
- 可选：在**管理员**字段中输入仓库源的联系人的姓名。
- 可选：在**描述**字段中输入数据的简短描述。
- 可选：在**注释**字段中输入有关仓库源的任何附加信息。

指定代理站点

- 单击**代理站点**选项卡。
可选择的站点显示在**可用的代理站点**列表中。
- 在**可用的代理站点**列表中，选择想要对其授予仓库源访问权的站点名。
- 单击 **>**。
该代理站点移动至**选择的代理站点**列表。现在可使用该站点访问仓库源。
要包括**可用的代理站点**列表中的所有项目，单击 **>>**。

指定数据库信息

- 单击**数据源**选项卡。



The screenshot shows a dialog box titled "Define Warehouse Source" with a sub-title "New Warehouse Source". It has five tabs: "Warehouse Source", "Agent Sites", "Data Source", "Tables and Views", and "Security". The "Data Source" tab is selected. The fields are as follows:

- Data source name: COSTS (dropdown menu)
- System name: (empty dropdown menu)
- User ID: db2admin
- Password: (masked with asterisks)
- Verify password: (masked with asterisks)
- Customize ODBC connect string
- ODBC connect string: DSN=;UID=;PWD=;

Buttons at the bottom: OK, Cancel, Help.

- 在**数据源名称**字段中指定系统数据源的名称。
对于非 DB2 源，数据源名可不同于数据库名。

3. 在**系统名字段**中，指定正定义为仓库源的数据库或文件所在工作站的主机名。
此步骤是可选的，且仅当使用名称与不同工作站上的另一数据库或文件相同的数据库或文件时，才使用该步骤。
4. 在**用户标识**字段中，输入将从代理站点工作站访问该数据库的用户标识。
5. 在**密码**字段中，输入该用户标识的密码。
6. 在**验证密码**字段中再次输入该密码。
7. 可选：选择**定制 ODBC 连接字符串**复选框以输入特定的连接字符串。
8. 若选择了“ODBC 连接字符串”复选框，在 **ODBC 连接字符串**字段中输入 ODBC 连接字符串。

导入源表和视图

如果数据源是使用 Relational Connect 和 Create Nickname 语句创建的，必须手工定义源表。

1. 单击**表和视图**选项卡。
2. 根据想要导入的对象的类型，展开**表或视图**文件夹。
要导入表和视图，双击**可用的表和视图**列表的空白处，并单击**刷新**。
“过滤器”窗口打开。
3. 通过使用“过滤器”窗口选项，指定想要从数据库导入的对象的类型。
“过滤器”窗口上可用的选项取决于您是想导入表还是想导入视图。
 - 选择**仅检索可复制的表**以仅检索支持复制的表。此选项在导入表时可用。
 - 选择**仅检索可复制的视图**以仅检索支持复制的视图。此选项在导入视图时可用。
 - 选择**包括系统表**以从数据库中检索系统表。此选项在导入表时可用。
 - 在**对象模式**字段中，输入包括通配符的搜索字符串，以定义模式名的特征，这些模式名包含将从数据库返回的对象。
例如，XYZ* 返回其模式以这些字符开始的表和视图。
 - 在**对象名字段**中，输入包括通配符的搜索字符串，以定义将从数据库返回的表和视图名的特征。
例如，XYZ* 返回以这些字符开始的表和视图。
4. 单击**确定**提交输入的过滤标准。
系统返回符合过滤标准的对象，并将其列示在**可用的表和视图**列表中。
5. 从**可用的表和视图**列表中，选择想要包括在仓库源中的表、视图或单个列。
6. 单击 **>**。
选择的表、视图和列移动至**选择的表和视图**列表。

设置仓库源

要包括可用的表和视图列表中的所有项目，单击 >>。

为仓库组授予访问权

1. 单击安全性选项卡。
2. 从可用的仓库组列表中，选择想要授予对仓库源的访问权的仓库组。
3. 单击 >。
选择的组移动至选择的仓库组列表。
要包括可用的仓库组列表中的所有项目，单击 >>。
4. 单击确定保存更改并关闭“定义仓库源”笔记本。

指定重试计数

1. 在主“数据仓库中心”窗口中，单击仓库 → 特性以打开“数据仓库中心”的“特性”笔记本。
2. 单击处理选项选项卡。
3. 在计数字段中，指定数据抽取将重试多少次。
4. 在小时和分钟字段中，指定在“数据仓库中心”重试抽取前所经过的时间量。

定义供 DataJoiner 使用的仓库源

DataJoiner 为访问步骤的数据提供了几项便利条件。通过使用本机数据库协议，可使用 DataJoiner 直接访问非 IBM 数据库，而无须使用对那些数据库的 ODBC 支持。您也可使用 DataJoiner 写入 Oracle 数据库或其他非 IBM 数据库。利用 DataJoiner，您可使用单个 SQL 语句和单个接口来访问和连接不同数据源中的数据。接口隐藏不同的 IBM 和非 IBM 数据库之间的差异。DataJoiner 优化 SQL 语句以改善性能。

可定义利用 DataJoiner 优点的“数据仓库中心”步骤。首先，定义使用 DataJoiner 数据库的仓库。然后定义向那些仓库进行写入的步骤。

DataJoiner 目标数据库不支持“数据仓库中心”变换器。

在阅读本节之前，阅读第7页的『第2章 设置仓库』、第25页的『第3章 设置仓库源』、第101页的『第4章 设置对仓库的访问权』和第119页的『第5章 定义和运行过程』以了解仓库的基本组件。您还应熟悉在 DataJoiner 中创建服务器映射和别名。

有关在 DataJoiner 中创建服务器映射的详情，参见 *DB2 DataJoiner: Planning, Installation, and Configuration Guide*。

有关在 DataJoiner 中创建表和别名的信息，参见 *DB2 DataJoiner: Administration Guide*。

在“数据仓库中心”中，为每个 DataJoiner 源数据库定义单独的仓库源定义。在定义仓库源之前，您必须通过 DataJoiner 的服务器映射将每个源数据库映射为 DataJoiner 数据库。您可能还必须创建用户映射，该映射将 DataJoiner 的用户标识和密码映射为源数据库的用户标识和密码。在“数据仓库中心”中为资源所定义的用户标识和密码就是相应的 DataJoiner 数据库的用户标识和密码。还必须为每个要与“数据仓库中心”一起使用的数据库源表创建别名。

以下示例显示如何创建服务器映射和为表创别名：

```
CREATE SERVER MAPPING FROM oracle1 TO NODE "oranode"
TYPE Oracle VERSION 7.2 PROTOCOL "sqlnet"
```

```
create user mapping from USER to server oracle1 authid "iwhserve" password "VWPW"
```

```
create nickname iwh.oracle_target for ORACLE1.iwhserve.oratar
```

- CREATE SERVER MAPPING 命令（为了便于阅读，它横跨多行）定义一个称作 Oracle 1 的源数据库，其中：

oracle1

标识 DataJoiner 中的远程数据库的名称。

oranode

Oracle TNSNAMES 文件中定义的项目，它标识目的地 Oracle TCP/IP 主机和端口。

Oracle 数据库的类型。

7.2 数据库的版本。

sqlnet 访问该数据库所用的方法，如对 Oracle 使用 SQLNET，或对 Sybase 使用 dblib 或 ctlib。

DATABASE 的值是空值，因为 Oracle 每个节点只允许一个数据库。对于其他的一些数据库源，可指定一个数据库。有关在 DataJoiner 中创建服务器映射的详情，参见 *DB2 DataJoiner: Planning, Installation, and Configuration Guide*。

- create user mapping 命令指定 DataJoiner 将用来连接远程数据库 (Oracle) 的用户标识。关键字 USER 是 DB2 专用的注册表，它指定当前登录的用户。用户将使用指定的用户标识和密码 (*iwhserve* 和 *VWPW*) 来连接远程 Oracle 数据库

设置仓库源

- `create nickname` 命令为远程 Oracle 表 `ORACLE1.iwhserve.oratar` 指定别名 `iwh.oracle_target`。在 `DataJoiner` 中创建别名之后，可用该别名来替代远程表的名称。例如，若您指定 `SELECT * FROM iwh.oracle_target`，则 `DataJoiner` 从 Oracle 表 `iwhserve.oratar` 返回数据。

创建映射和别名之后，可定义仓库源。要定义每个仓库源的源表，导入 `DataJoiner` 别名作为表定义。在前例中，应从 `DataJoiner` 中导入 `iwh.oracle_target`。

第4章 设置对仓库的访问权

每组用户将访问一个仓库。为仓库选择的操作系统取决于生产能力计划、用户工作的操作系统和网络约束。DB2 系列的数据库可支持各种各样的能力需求，因此可选择适合于您的仓库的数据库。

仓库用户必须对仓库数据库具有访问权。他们可使用任何可访问 DB2 数据的应用程序。

设置 DB2 通用数据库仓库

可使用任何“DB2 通用数据库”数据库作为仓库。有关通过“DB2 通用数据库”服务器或任何 DB2 客户机与“DB2 通用数据库”数据库连接的更多信息，请参阅《DB2 通用数据库安装和配置补遗》。

定义对 DB2 通用数据库仓库的特权

在设置仓库之前，需要在目标系统上创建一个数据库。然后设置对数据库具有下列特权的用户标识：

- CREATETAB（数据库级特权）
- BINDADD（数据库级特权）
- CONNECT（数据库级特权）

另外，下列系统表需要显式 SELECT 特权：

- SYSIBM.SYSTABLES
- SYSIBM.SYSCOLUMNS
- SYSIBM.SYSDBAUTH
- SYSIBM.SYSTABAUTH
- SYSIBM.SYSINDEXES
- SYSIBM.SYSRELS
- SYSIBM.SYSTABCONST

建立与 DB2 通用数据库仓库的联网

定义必需的特权之后，在代理站点上执行下列任务：

1. 若数据库是远程的，则设置与目标系统的通信。
2. 若数据库是远程的，则编目目标系统的节点。

3. 编目该数据库。
4. 若正在使用 Windows NT 或 OS/2 仓库代理，或正在使用 AIX、OS/390 或 Solaris 操作环境仓库代理使用 ODBC 的版本，则将该数据库注册为 ODBC 系统 DSN。若正在使用 AIX 或 Solaris 操作环境使用 CLI 接口的仓库代理，则使用 DB2 编目实用程序编目该数据库。
5. 将数据库实用程序和 ODBC (CLI) 与目标数据库绑定。仅在每种类型的客户机上执行一次绑定。

设置 DB2 AS/400 版仓库

可以使用 DB2 Connect 或 Client Access/400 (CA/400) 来访问 DB2 AS/400 版仓库。

使用 DB2 Connect

要用 DB2 Connect 来设置对 DB2 AS/400 版仓库的访问权：

1. 验证网关服务器是否安装了先决条件产品。
2. 定义对仓库的特权。
3. 设置 DB2 Connect 网关站点。
4. 建立与仓库的联网。

先决条件产品

若要访问 DB2 AS/400 版仓库中的数据，需要网关站点。通过安装 DB2 Connect 来配置 DRDA 的站点。

有关 DRDA 的更多信息，参见 *Distributed Relational Database Architecture Connectivity Guide*。

有关 DB2 Connect 的信息，参见下列书籍：

- 《DB2 Connect 个人版快速入门》或《DB2 Connect 企业版快速入门》
- 《DB2 Connect 用户指南》
- 《安装和配置 DB2 客户机》

定义对 DB2 AS/400 版仓库的特权

目标系统的系统管理员必须设置对 NULLID 集合具有 CHANGE 权限或更高权限的用户标识。

另外，下列系统表需要显式 SELECT 特权：

- SYSIBM.SYSTABLES
- SYSIBM.SYSCOLUMNS

- SYSIBM.SYSINDEXES
- SYSIBM.SYSREFCST
- SYSIBM.SYSCST

用户标识还需要 ALLOBJ 特权来创建 AS/400 集合。

设置 DB2 Connect 网关站点

定义了必需的特权之后，在网关站点上执行下列任务：

1. 安装 DB2 Connect。
2. 配置 DB2 Connect 系统，以与目标数据库通信。
3. 更新 DB2 节点目录、系统数据库目录和 DCS 目录。

建立与 DB2 AS/400 版仓库的联网

在代理站点上执行下列任务：

1. 设置与 DB2 Connect 站点的通信。
2. 编目 DB2 Connect 站点的节点。
3. 编目该数据库。
4. 使用“ODBC 管理员”将该数据库注册为系统 DSN。
5. 如《DB2 Connect 用户指南》中描述的那样，将 DB2 Connect 实用程序与 DRDA 服务器绑定。仅在每种类型的客户机上执行一次绑定。

使用 CA/400

要设置 CA/400 以便访问 DB2 AS/400 版数据库：

1. 在 AS/400 上安装最新的 Client Access PTF。

要查看说明 AS/400 需要哪些 PTF 的最新信息 APAR 的完整列表，可访问如下 Web 页面：

<http://www-1.ibm.com/servers/eserver/series/clientaccess/>

2. 选择 **Support Information** 下的 **Service Packs**。

在 Client Access Service Packs Web 页面上，可以找到有关最新工作站服务包和主机 PTF 的信息。

要检查您的操作系统的当前级别：

- a. 输入如下命令：

```
GO LICPGM
```

- b. 选择“选项 10”。
- c. 按 F11 键。

3. 确保已设置 AS/400 QUSER 密码（且未到期）。若需要的话，可将该密码设置为无密码。
4. 在 WINNT\SYSTEM32\DRIVERS\ETC\HOSTS 文件中输入 AS/400 系统和 Windows NT 工作站的 TCP/IP 主机名。例如：

```
9.123.456.7    demo400
9.123.456.8    mypc
```

5. 在 Windows NT 工作站上安装 CA/400。
6. 验证 AS/400 系统上的所有服务器都在运行。在 Windows NT 工作站上，从 DOS 命令提示符处输入如下命令：

```
cwbping
hostname ip
```

您将接收到类似以下的响应：

```
C:\>cwbping demo400 ip
```

```
IBM AS/400 Client Access Windows 95/NT 版
版本 3 发行版 1 级别 2
```

```
(C) Copyright IBM Corporation and Others 1984, 1995.
All rights reserved.
```

```
U.S. Government Users Restricted Rights - Use, duplication or
disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
Licensed Materials - Property of IBM
[9.180.160.17]
```

```
ping 服务器 “端口映射器” 成功
ping 服务器 as-central 成功
ping 服务器 as-database 成功
ping 服务器 as-dtaq 成功
ping 服务器 as-file 成功
ping 服务器 as-netprt 成功
ping 服务器 as-rmtcmd 成功
ping 服务器 as-signon 成功
```

若服务器未启动，则在 AS/400 系统上输入以下命令来启动服务器：

```
STRHOSTSVR SERVER (*ALL)
```

7. 验证 AS/400 是否具有已定义的 DB2 数据库名。可通过在 AS/400 上发出“处理关系数据库目录条目” (WRKRDBDIRE) 命令来检查此条目。

若 AS/400 没有 DB2 数据库名，当您尝试访问数据库时，将接收到 SQL0114 错误。

8. 在 Windows NT 系统变量（不是用户变量）中，将 DB2CODEPAGE 环境变量设置为 850。
9. 从 Web 下载最新的 CA/400 服务包和 CWBCFG 实用程序。您可获得如何从如下 Web 页面下载 SP45545 和 cwbcfg.exe 的全部详细信息：

```
http://www-1.ibm.com/servers/eserver/iserries/clientaccess/service.htm
```

10. 应用该服务包并运行 CWBCFG 实用程序以允许 CA/400 在 Windows NT 上作为系统任务运行。

若不应用服务包 SP45545 和运行 CWBCFG 实用程序，将接收到如下错误消息：

"IBM.[Client Access ODBC 驱动程序 (32 位)][DB2/400 SQL] 通信故障。COMM RC=0x3"

原因：根据 Rochester 实验室，CA/400 连接是在“当前用户”(HKEY_CURRENT_USERS)下面注册的，但是应在 HKEY_USERS (.Default) 下面注册。

解决方案：您需要在正确的注册表中注册该连接，并使用独立的实用程序 cwbcfg.exe，它可用来为 HKEY_USERS.Default 用户配置会话。cwbcfg 的正确语法是：
cwbcfg /host HOSTNAME /s

其中，HOSTNAME 是 HOSTS 文件中的 AS/400 名（参见步骤 3）。

注意："/s" 选项很重要。执行此命令之后，您可通过运行 regedt32 来验证是否已设置了 NT 注册表并浏览 HKEY_USERS (.Default)

11. 使用 CA/400 ODBC 驱动程序将 AS/400 数据库注册为系统 DSN。添加其他必需的库，例如 DORMERK、QGPL、MYLIBRARY
12. 用 Lotus Approach® 或 Microsoft Query 来测试与 AS/400 数据库的连接。
若您接收到 CA/400 错误，则单击帮助以获取该错误的详细描述。

设置 DB2 OS/390 版仓库

在缺省情况下，“数据仓库中心”以 DB2 通用数据库产品支持的格式创建仓库目标表。您也可调整步骤以使用 DB2 OS/390 版中的目标表。

定义对 DB2 OS/390 版仓库的特权

在设置仓库之前，需要在目标系统上创建一个数据库。然后设置对该数据库具有下列特权的用户标识：

- 创建一个数据库和表空间以用作仓库。
- 赋予某个用户标识对该数据库的 DBADM 特权。

建立与 DB2 OS/390 版仓库数据库的联网

要建立与 DB2 OS/390 版数据库的联网，执行下列任务：

1. 设置对 DB2 OS/390 版的访问权。
2. 与仓库数据库连接。

创建“数据仓库中心”状态表

通过使用以下命令，创建“数据仓库中心”状态表：

```
CREATE TABLE IWH.BVBESTATUS (BVNAME VARCHAR(80) NOT NULL, RUN_ID INT NOT NULL,  
UPDATIME CHAR(26) NOT NULL) IN database-name.tablespace-name
```

database-name

仓库数据库的名称

tablespace-name

包含数据库的表空间的名称

对“数据仓库中心”定义仓库

当为 DB2 OS/390 版仓库定义目标表时，必须指定用于创建表的表空间。若不指定表空间，DB2 OS/390 版会在为给定的子系统定义的缺省 DB2 数据库中创建该表。

要指定表空间：

1. 定义一个仓库。
2. 定义或生成一个目标表。
3. 右键单击该目标表，然后单击**特性**以打开该表的“特性”笔记本。
4. 在**表空间**字段中，指定用于创建该表的表空间。
5. 验证是否已清除了**授权为公共**复选框。“数据仓库中心”创建的 GRANT 命令语法不受 DB2 VM 版或 DB2 VSE 版的支持。
6. 单击**确定**保存更改并关闭“表”笔记本。

将该步骤提升为测试方式时，若指定“数据仓库中心”来创建目标表，“数据仓库中心”会在 DB2 OS/390 版数据库中创建目标表。

设置 DB2 EEE 仓库

在缺省情况下，“数据仓库中心”以 DB2 通用数据库产品支持的格式创建仓库目标表。也可以调整步骤以使用“DB2 UDB 企业扩充版”（DB2 EEE）中的目标表。

定义对 DB2 EEE 数据库的特权

在设置仓库之前，需要在目标系统上创建一个数据库。然后设置对数据库具有下列特权的用户标识：

- 创建一个数据库、节点组和表空间以用作仓库。
- 赋予某个用户标识对该数据库的 BINDADD、CONNECT 和 CREATETAB 特权。

建立与 DB2 EEE 数据库的联网

按照第101页的『建立与 DB2 通用数据库仓库的联网』中的指示信息来设置对 DB2 EEE 的访问权，并从显式 SELECT 特权开始。

对“数据仓库中心”定义 DB2 EEE 数据库

设置了对系统的访问权之后，在“数据仓库中心”中执行下列任务来使用 DB2 EEE 仓库：

1. 定义一个仓库目标。
2. 定义或生成一个目标表。
3. 右键单击该目标表，并单击**特性**。
“表”笔记本打开。
4. 单击**显示 SQL**。
“显示 SQL”窗口打开。
5. 单击**编辑**。
确定窗口打开。单击**是**编辑该 SQL。
6. 将如下子句添加至 CREATE 语句的末尾：

```
IN tablespace-name PARTITIONING KEY (column-name) USING HASHING
```

tablespace-name

包含数据库的表空间的名称

column-name

用于分区的列的名称

7. 单击**确定**。
“编辑 SQL”窗口关闭。
8. 单击**确定**。
“表”笔记本关闭。

将该步骤提升为测试方式时，“数据仓库中心”会在 DB2 EEE 数据库中创建目标表。

定义供 DataJoiner 使用的仓库

IBM DataJoiner 为访问步骤的数据提供了几项便利条件。通过本机数据库协议，可使用 DataJoiner 直接访问非 IBM 数据库，而无须使用对那些数据库的 ODBC 支持。也可使用 DataJoiner 写入非 IBM 数据库。利用 DataJoiner，可使用单个 SQL 语句和单个接口来访问和连接不同数据源中的数据。接口隐藏了不同的 IBM 和非 IBM 数据库之间的差异。DataJoiner 优化了 SQL 语句以改善性能。

可以定义利用 DataJoiner 功能的“数据仓库中心”步骤。首先，定义使用 DataJoiner 数据库的仓库。然后定义向那些仓库进行写入的步骤。

在阅读本节之前，阅读第7页的『第2章 设置仓库』、第25页的『第3章 设置仓库源』、第101页的『第4章 设置对仓库的访问权』和第119页的『第5章 定义和运行过程』以了解仓库的基本组件。还应熟悉在 DataJoiner 中创建服务器映射和别名。

在“数据仓库中心”中，为每个 DataJoiner 目标数据库创建单独的逻辑仓库定义。例如，为每个 Oracle 数据库定义一个仓库，为每个 Sybase 数据库定义另一个仓库，等等。

指定同一个 DataJoiner 数据库作为每个逻辑仓库的数据库。当“数据仓库中心”生成 SQL 以便从源数据库抽取数据并将数据写至目标数据库时，因为 DataJoiner 数据库既是源数据库，又是目标数据库，所以“数据仓库中心”将生成 SELECT INSERT 语句。然后，DataJoiner 优化对 DataJoiner 目标数据库（如 Oracle 和 Sybase）的查询。通过利用 DataJoiner 异机种连接优化，可定义具有来自多个数据库的源的步骤。

要用不同的“数据仓库中心”代理访问仓库，更改数据库的名称、用户标识和密码。

仓库数据库中的 BVBESTATUS 表包含步骤版本的时间戳记。必须在 DataJoiner 数据库或每个远程数据库中创建 BVBESTATUS 表。

若在 DataJoiner 数据库中创建 BVBESTATUS 表，必须使用两阶段落实来使该表与远程仓库数据库保持同步。

若在远程数据库中创建 BVBESTATUS 表，对该表的更新将与远程数据库处于同一落实域。因为“数据仓库中心”要求表的名称是 BVBESTATUS，所以对于每个远程数据库，必须要有不同的 DataJoiner 数据库。一个 DataJoiner 别名不能表示不同数据库的多个表。

要创建 BVBESTATUS 表，使用 CREATE TABLE 语句。例如，要在 Oracle 数据库中创建该表，发出如下命令：

```
CREATE TABLE BVBESTATUS (BVNAME, VARCHAR2(80) NOT NULL,  
                           RUN_ID NUMBER(10) NOT NULL,  
                           UPDATIME CHAR(26) NOT NULL)
```

创建该表之后，在 DataJoiner 中为 IWH.BVBESTATUS 表创建别名。

有关在 DataJoiner 中创建表和别名的详情，参见 *DB2 DataJoiner: Administration Guide*。

有关在 DataJoiner 中创建服务器映射的详情，参见 *DB2 DataJoiner: Planning, Installation, and Configuration Guide*。

用 DataJoiner 创建目标表

对于“DataJoiner 版本 2.1.1”或更高版本，“数据仓库中心”可直接将表创建到远程数据库（如 Oracle）中。

要创建目标表：

1. 创建带有 DataJoiner 数据库中的目标表的步骤。
2. 在“表”笔记本的表空间字段中输入服务器映射名。
在第98页的『定义供 DataJoiner 使用的仓库源』中的示例中，服务器映射名是 Oracle 1。
3. 因为“数据仓库中心”的缺省表名限定符是 IWH，所以确定需要做什么：
 - 请求名为 IWH 的用户标识。
 - 将表名限定符更改为与您的用户标识匹配。
 - 请求给您的用户标识授予创建任何表、插入到任何表中、更新任何表和删除任何表的特权。
 - 请求给您的用户标识授予 DBA 权限。

若目标数据库的用户标识具有用与您的用户标识不同的限定符创建表的特权，可继续第 4 步。

4. 将步骤提升为测试方式。
5. 运行该步骤以验证已将正确的数据写入目标表。
6. 将步骤提升为生产方式。

若 DataJoiner 的版本早于“版本 2.1.1”，可：

- 在 DataJoiner 数据库中创建表，并将它移动至远程数据库中。
- 在远程数据库中创建表或 / 和更新远程数据库中的表。

如果将 DataJoiner 与 Windows NT、UNIX 或 OS/2 上的 DB2 版本 7.1（并带有修订包 2 或更高版本）一起使用，可能会出现错误，指示发生了绑定问题。例如，将 DataJoiner 源与“数据仓库中心版本 7”代理一起使用，可能会得到以下错误：

```
DWC07356E 对于步骤 "?" 的修订版 "0"，代理处理类型为  
"importTableNames" 的命令失败。SQL0001N 绑定或预编译未成功完成。SQL0001N 未找到数据包 "I
```

要校正该问题，可将以下行添加到 db2cli.ini 文件中：

[COMMON]
DYNAMIC=1

在 UNIX 系统上, db2cli.ini 文件位于 ../sqllib/cfg 目录中。在 Windows NT 上, db2cli.ini 文件位于 ../sqllib 目录中。

移动目标表

可在 DataJoiner 数据库中创建并测试步骤, 然后将它移动至远程数据库中:

1. 创建带有 DataJoiner 数据库中的目标表的步骤。
2. 将步骤提升为测试方式。
3. 运行该步骤以验证与源数据库的连接正在工作, 且已将正确的数据写入目标表。
4. 人工将该表移动至远程数据库, 如 Oracle。(也可使用建模或数据字典工具。) DataJoiner 表和 Oracle 表的数据类型必须兼容。
 - a. 在 Oracle 数据库中创建该表。
 - b. 删除 DataJoiner 表。

有关 DB2 数据类型如何映射到 Oracle 数据类型的信息, 参见 *DB2 DataJoiner: Planning, Installation, and Configuration Guide*。

5. 在 DataJoiner 中人工为远程表创建别名。该别名必须与“数据仓库中心”中的步骤的目标表名匹配。
6. 再次运行该步骤, 以测试数据通过 DataJoiner 正确地移动至目标。
7. 将步骤提升为生产方式。

在远程数据库中创建或更新表

可使用“数据仓库中心”更新远程数据库中的现存表。当数据已存在时使用另一工具(如建模工具)来创建仓库模式时, 使用此选项。

1. 创建仓库表或使用现存仓库表。
2. 在 DataJoiner 中创建目标表的别名。
3. 通过在“定义仓库目标”笔记本的表列表中选择该表, 将别名定义导入仓库。

在该表的“特性”笔记本中, “列”页显示已导入的表定义。
4. 在那个仓库中创建一个步骤, 选择导入的表作为该步骤的目标表。如果“步骤”笔记本的“处理选项”页上有一个**数据仓库中心创建的表**复选框, 应验证在提升该步骤时该复选框是否已清除。
5. 将步骤提升为测试方式。
6. 运行该步骤测试对目标表的访问权。

7. 将步骤提升为生产方式。

定义仓库目标

在将仓库的源定义为仓库源之后，定义将包含该数据的仓库目标。“数据仓库中心”支持对您的仓库使用DB2 通用数据库和 DB2 AS/400 版数据库。也可更改用于目标表的 CREATE 语句，以便使用 DB2 EEE 和 DB2 OS/390 版。有关详情，参见第106页的『设置 DB2 EEE 仓库』和第105页的『设置 DB2 OS/390 版仓库』。

如果数据源是使用 Relational Connect 和 Create Nickname 语句创建的，它将不可用于导入表到“数据仓库中心”中。要使用数据源作为源表，可在“数据仓库中心”中定义一个仓库目标，但不要导入目标表。必须手工创建表，并确保仓库目标表中的列映射到数据源中的列。

表10列示了数据仓库中心支持的仓库目标的版本和发行版级别。

表 10. 受支持的 IBM 仓库目标的版本和发行版级别

目标	版本 / 发行版
DB2 通用数据库 Windows NT 版	6 - 7
DB2 通用数据库扩充企业版	6 - 7
DB2 通用数据库 OS/2 版	6 - 7
DB2 通用数据库 AS/400 版	3.1-5.1
DB2 通用数据库 AIX 版	6 - 7
DB2 通用数据库 Solaris 操作环境版	6 - 7
DB2 通用数据库 OS/390 版	5 - 7
DB2 DataJoiner	2.1.1
DB2 DataJoiner/Oracle	8
DB2 VM 版	3.4 - 5.3.4
DB2 VSE 版	3.2, 7.1
CA/400	3.1.2

定义仓库目标特性

任何仓库用户都可以定义仓库目标，但仅属于对仓库目标具有访问权的仓库组的用户可以更改仓库目标。有关定义用户和安全性组的信息，参见联机帮助。

要定义仓库目标：

1. 右键单击**仓库目标**文件夹。

2. 单击**定义** → **DB2 系列**，然后选择适用于您的操作系统的 DB2 通用数据库的版本。

“定义仓库目标”笔记本打开。

3. 在**名称**字段中，输入仓库目标的业务名。
4. 在**管理员**字段中，输入该仓库目标的联系人。
5. 在**描述**字段中，输入该数据的简短描述。
6. 单击“代理站点”页。可选择的站点显示在**可用的代理站点**列表中。
7. 从**可用的代理站点**列表中，选择想要对其授予仓库目标访问权的站点名，并单击 **>**。您的代理站点添加至 **选择的代理站点**列表。现在可使用该站点访问仓库目标。

要包括**可用的代理站点**列表中的所有项目，单击 **>>**。

8. 单击**数据库**选项卡。
9. 在**数据库名字**字段中输入数据库的名称。如果正在 OS/2 上定义仓库目标，应以大写字母输入数据库的名称。对于 OS/390，它必须是数据库的位置名。
10. 在**用户标识** 字段中，输入将从代理站点访问该数据库的用户标识。
11. 在**密码**字段中，输入该用户标识的密码。
12. 在**验证密码**字段中，再次输入该密码。
13. 可选：在“目标”数据库区中，指定是否应在目标数据库中创建“数据仓库中心”变换器。如果不想在目标数据库中创建“数据仓库中心”变换器，可单击**删除或不创建变换器**。

- 如果删除此目标的变换器，在仓库目标中定义的使用变换器的过程或步骤将不运行。如果您以后又决定要使用变换器，可再次创建它们。
- 删除此仓库目标中的变换器将影响到与该数据库相关的其他仓库目标。
- 要创建变换器，并将他们作为受防护的变换器来运行，可单击**创建变换器并注册为受防护的**。
- 要创建变换器，并将他们作为不受防护的变换器来运行，可单击**创建变换器并注册为不受防护的**。
- 将变换器从受防护更改为不受防护，或从不受防护更改为受防护的时，将删除存储过程然后再次创建。此项更改会影响到与此数据库相关的其他仓库目标。

注：在目标数据库中创建“数据创建中心”变换器的选项对 OS/390 不可用，因为 OS/390 不支持脚本。关于在 OS/390 上创建变换器的信息，请参阅《*DB2 仓库管理器安装指南*》。

14. 可选：在“仓库目标”区中，指定是否对变换器启用该仓库目标。

- 要指定仓库目标可以使用变换器，单击**对变换器启用目标**。若已为此仓库目标启用变换器，则**对变换器启用**将已被选择。
- 要指定仓库目标不可以使用变换器，单击**不对变换器启用**。

15. 单击**表和视图**选项卡。

16. 如果数据源是使用 **Relational Connect** 和 **Create Nickname** 语句创建的，要跳到步骤16c。若目标表已经存在，可以从该数据库中导入它们的定义：

a. 展开**表或视图**文件夹。

在“OS/400 版本 4 发行版 2”和“OS/400 版本 4 发行版 3”上，必须选择**视图**文件夹来导入系统表。

“过滤器”窗口打开。

- 可选：选择**仅可复制的表**以仅检索支持复制的表。
- 可选：选择**包括系统表**以从数据库中检索系统表。
- 可选：在**对象模式**字段中，输入包括通配符的搜索字符串，以定义模式名的特征，这些模式名包含将从数据库返回的对象。例如，XYZ* 返回其模式以这些字符开始的表。
- 可选：在**对象名字段**中，输入包括通配符的搜索字符串，以定义将从数据库返回的表名的特征。例如，XYZ* 返回以这些字符开始的表。

b. 单击**确定**。

若仓库目标选择了多个代理站点，仓库服务器将名称排在第一（取决于用户的语言环境）的代理站点用于导入。

例如，仓库目标选择了如下三个代理站点：“缺省代理”、“AIX 代理”和“OS/390代理”。仓库服务器将名为“AIX 代理”的代理站点用于导入。

符合您输入的过滤器标准的对象显示在**可用的表**列表中。

c. 若这些目标表不存在，则定义它们：

1) 右键单击**选择的表**列表中的空白处。

2) 单击**定义**。

“定义仓库目标表”笔记本打开。

3) 在**表模式**字段中，输入创建该仓库数据库所用的用户标识。

4) 在**表名字段**中，为您创建的表输入全限定名。该名称最多可包含 128 个字符，可以是普通标识符或是定界标识符。

“数据仓库中心”支持使用普通 SQL 标识符的目标表。普通标识符：

- 必须以字母开始。
- 可以包括大写字母、数字和下划线。
- 不能是保留字。

若表的普通标识符的一部分含有小写字母，“数据仓库中心”将小写字母存储为大写字母。

“数据仓库中心”不支持使用定界标识符的目标资源表。定界标识符：

- 包括在引号内。
- 可以包括大写字母和小写字母、数字、下划线和空格。
- 可以包含双引号，用两个连续的引号表示。

对于某些仓库目标类型，表名可以区分大小写的或具有其他限制。有关命名约定的信息，参见仓库目标类型的文档。

- 5) 可选：在**表空间**字段中，指定用于创建表的表空间。
- 6) 可选：在**索引表空间**字段中，指定用于对表创建任何索引的表空间。
- 7) 在**表名字段**中，输入目标表的名称。
- 8) 在**描述**字段中，输入该表的描述。
- 9) 在**业务名字段**中输入该表的业务名（用户理解的描述性名称）。
- 10) 若您想要“数据仓库中心”在运行填充此表的步骤的时候创建此表，验证是否已选择“数据仓库中心”创建的表复选框。

当您希望“数据仓库中心”创建目标表时（如目标表是一个 SQL 步骤的结果时），使用此选项。若使用的目标表已定义，则清除此复选框。

- 11) 可选：选择**瞬时数据**复选框以指定源数据是临时的，将在接收数据的目标填满了以后立即删除。
- 12) 若对该数据库拥有访问权的任何人要对该表拥有访问权，则验证已选择**授权公共**复选框。
- 13) 可选：在**修订版数**字段中，指定当过程运行时，您希望此表中的数据被替换的次数。如果希望“数据仓库中心”在每次过程运行时都附加数据，可保留该选项为缺省值 0。
- 14) 在**修订版列**列表中，单击想要用作目标表的修订版列的列。修订版数必须大于 0 此字段才能生效。修订版列的类型必须是整数。如果需要“数据仓库中心”在表中创建一个修订版列，可从列表中选择**<系统生成的>**。
 - 系统生成的列处理起来与其他任何用户定义的列类似。当修订版数设置为 0 或在“修订版”列的列表中选择了另一列时，系统生成的列不会自动除去。如果不选择某列作为修订版列，可以从该表的“特性”笔记本中的“列”页中手工除去它。

- 在已指定了修订版列时，会从“修订版”列列表中除去“系统生成的”选项。要将“系统生成的”选项重新设置为列表，可首先将修订版数设置为 0，然后将修订版数设置为大于 0 的值。
 - 如果创建多个系统生成的修订版列，将赋予每列一个唯一的名称。例如，第一列称为 RUN_ID、第二列称为 RUN_ID1 而第三列称为 RUN_ID2，以此类推。
 - 如果已指定了一列作为修订版列，然后又从列表中选择了另一列，则新选择的列现在将作为修订版列。旧的修订版列会保留在表中，但不再指定为修订版列。如果不希望旧的修订版列保留在表中，必须手工除去它。
- 15) 若该表是要导出至 DB2 OLAP Integration Server 的维表或事实表，则验证已选择 **OLAP 模式** 的部分复选框。
 - 16) 单击**列**选项卡。
 - 17) 在**列**页上，右键单击该表的空白处并单击**添加**，或右键单击一行并单击**插入**。如果单击**添加**，会在所有其他行下添加一个空白行。若单击**插入**，则在所选行之上添加一个空白行。要除去一列，可右键单击该列，然后单击**除去**。
 - 18) 在**列名字段**中输入列的名称。

对于某些数据源类型，列名可以是区分大小写的。有关命名约定的信息，参见数据源类型的文档。
 - 19) 在**数据类型**字段中，为列指定数据类型。
 - 20) 若指定对字节数有要求的数据类型，在**长度**字段中指定长度。

例如，必须为 **CHAR** 数据类型指定长度。
 - 21) 若指定 **DECIMAL** 数据类型，则在**精度**字段中指定精度。

精度指数字的总数，范围为 1 至 131。
 - 22) 若指定 **DECIMAL** 数据类型，则在**小数位**字段指定小数位。

小数位指小数点右边的数字的位数，其范围为 0 至该数字的精度。
 - 23) 为确保正确转换代码页，若 **CHAR** 或 **VARCHAR** 字段包含文本，则选择是**文本**复选框。
 - 24) 若该表中允许空值（不存在或未知），则选择**可空**复选框。
 - 25) 可选：若源是一个复制源，则选择**之前复制**以捕捉该列的前映像。
 - 26) 可选：若该列是为捕捉变化定义的，则选择**之后复制**。
 - 27) 可选：在**描述**字段中，输入列的描述。
 - 28) 单击**确定**。

“表”笔记本关闭。目标表添加在**已选表**列表中的**表**文件夹下。

17. 单击**安全性**选项卡。
18. 选择仓库组以授予组中的用户创建使用此仓库目标的步骤的能力。
19. 单击 **>**。
该仓库组移动至**选择的安全性组**列表。
20. 单击**确定**保存更改并关闭“仓库目标”笔记本。

定义主键

要定义仓库主键:

1. 在“仓库主键”页上，从**可用的列**列表中选择想要定义为仓库主键的列，并单击 **>**。选择的列移动至**仓库主键列**列表。因为数据库管理器使用仓库主键对表数据进行高效访问，所以指定列的次序很重要。
要包括**可用的列**列表中的所有列，单击 **>>**。
2. 要从仓库主键定义中除去一列，从**仓库主键列**列表中，选择该列并单击 **<**。
要从仓库主键定义中除去所有列，单击 **<<**。
3. 可选：在**约束名称**字段中，为该表输入仓库主键约束的名称。为了编辑此字段，必须将列移动至**仓库主键列**列表中。
“数据仓库中心”为仓库主键约束生成缺省名称。

如果为 OS/390 上的表指定一个主键，必须在创建该表后为该表定义一个唯一索引。

定义外键

可以为仓库源表、仓库源视图或仓库目标表定义外键。“数据仓库中心”仅在连接过程中使用外键。“数据仓库中心”不会落实对底层数据库定义的外键。

在定义外键之前，必须知道外键对应的父表的名称和模式。

当步骤处于开发或测试方式时，可以定义外键。若步骤处于开发方式，则在通过将该步骤提升为测试方式创建表时，创建该键。若步骤处于测试方式，则在单击**确定**时，“数据仓库中心”改变该表以添加键。

要定义外键:

1. 在“外键”页上，右键单击表的空白处，并单击**定义**。“定义外键”窗口打开。
2. 可选：在**约束名称**字段中，为该表输入外键约束的名称。
若不提供名称，“数据仓库中心”为外键约束生成缺省名称。
3. 在**对象模式**字段中，选择父表的表模式。

4. 在**对象名字段**中，选择父表的名称。
主键列字段显示与指定的表相关的主键列。
5. 从**可用的列列表**中，选择想要定义为外键的列，并单击 **>**。该列移动至**外键列列表**中。
要包括**可用的列列表**中的所有列，单击 **>>**。
要从该表或视图中除去一个外键定义，从**外键列列表**中选择该列，并单击 **<**。
要除去外键定义中的所有列，单击 **<<**。
6. 单击**确定**。“定义外键”窗口关闭，且定义的外键显示在“外键”页上的外键列表中。

若选择了仓库数据库中存在的表（或已经运行了该步骤），则可以查看目标表中的数据。每次可查看一个表中的数据。“数据仓库中心”显示表中的所有列，最多可显示 200 行。

要查看数据，单击**样本数据**。

“样本数据”窗口打开，并显示数据的样本。

在仓库数据库中创建“数据仓库中心”状态表

用户可以使用 **BVBESTATUS** 表，通过匹配它们的时间戳记来连接表，或按日期范围而不是按版本号来查询版本。

例如，版本号 1010 对用户可能没有任何意义，但抽取数据的日期可能有意义。可以在目标表上创建一个简单的视图，以允许用户按抽取数据的日期来查询数据。

必须人工创建状态表。若该表是用“Visual Warehouse 版本 2.1”创建的，则必须删除该表并再次创建它。

要创建状态表：

1. 与目标仓库数据库连接。
2. 发出如下 **CREATE TABLE** 语句：

```
CREATE TABLE IWH.BVBESTATUS ( BVNAME VARCHAR(80) NOT NULL,  
RUN_ID INT NOT NULL, UPDATIME CHAR(26)NOT NULL );
```

可能需要按如下的方法更改该语句：

- 若目标仓库在 **DB2 Windows NT 版**、**DB2 OS/2 版**、**DB2 AS/400 版**、**DB2 AIX 版**或基于 **UNIX** 的平台上的 **DB2** 中，则使用所示语句。
- 若目标仓库在 **DB2 OS/390 版**中，可能需要标识将在其中创建状态表的数据库和表空间的名称。将以下子句添加至 **CREATE TABLE** 语句末尾：

IN database-name.tablespace-name

database-name

想在其中创建 BVBESTATUS 表的数据库的名称

tablespace-name

想在其中创建 BVBESTATUS 表的表空间的名称

- 若您的目标仓库使用“DB2 版本 5 扩充企业版”，可能需要标识要在其中创建状态表的表空间的名称和数据库的分区键。将以下子句添加至 CREATE TABLE 语句末尾:

```
IN tablespace-name partitioning-key (RUN_ID)  
USING HASHING
```

tablespace-name

想在其中创建 BVBESTATUS 表的表空间的名称

partitioning key

用于确定将特定数据行存储在哪个分区中的列名

第5章 定义和运行过程

定义仓库之后，需要用有用的信息填充仓库。为此，需要了解用户需要什么、什么源数据可用以及“数据仓库中心”如何将源数据转换成信息。

要对与业务逻辑区相关的过程进行标识并分组，应先定义主题区。

例如，若正构建销售和市场营销数据仓库，则定义“销售”主题区和“市场营销”主题区。然后在“销售”主题区下添加与销售相关的过程。类似地，在“市场营销”主题区下添加与市场营销数据相关的定义。

要定义如何为数据仓库移动和转换数据，在主题区内定义一个过程，它包含转换和移动过程中的一系列步骤。

在该过程中，定义数据转换步骤，它指定数据将如何从源格式转换为目标格式。通过包括下列规范，每个步骤定义数据从源格式到目标格式的转换：

- “数据仓库中心”将从中抽取数据的一个或多个源表、视图或文件。

在步骤中使用这些源表之前，必须将这些源定义为仓库源的一部分。（参见第25页的『第3章 设置仓库源』。
- “数据仓库中心”将要对其写入数据的目标表。

根据该步骤中的规范，可指定“数据仓库中心”在仓库数据库中创建表，或者指定“数据仓库中心”更新现存表。
- 数据将如何转换：
 - 通过发出一个 SQL 语句，该语句指定要抽取的数据以及如何将该数据转换为目标格式。

例如，SQL 语句可以从多个源表中选择数据，连接这些表，并将连接后的数据写入目标表。
 - 通过运行仓库程序或变换器。

例如，可能想使用 DB2 批量装入和卸装实用程序来将数据传送到您的仓库。或者可能想使用“清除”变换器清除您的数据。还可将一个外部程序定义成“数据仓库中心”的用户定义程序。

本章描述如何执行这些任务。

定义主题区

在创建过程之前，必须定义主题区。任何用户都可定义或编辑主题区。

要定义主题区：

1. 在“数据仓库中心”目录树上，右键单击**主题区**文件夹，并单击**定义**。
“定义主题区”笔记本打开。
2. 在**名称**字段中，输入主题区的业务名。
该名称可以有 80 个字符，包括空格在内。
3. 可选：在**管理员**字段中，输入此项的联系人姓名，例如您的数据库管理员。
该名称最多可为 80 字符长，并区分大小写。名称的第一个字符必须是字母数字。
4. 可选：在**描述**字段中，输入主题区的简短描述。
最多可输入 254 字符。
5. 可选：在**注释**字段中，输入有关主题区的附加信息。
最多可输入 32,000 个字符。
6. 单击**确定**在“数据仓库中心”目录树中创建主题区。

参见『定义过程』了解如何在此主题区下定义过程。

定义过程

定义过程对象以标识一个步骤序列。该序列可以是数据转换序列或 / 和启动步骤的序列。

定义过程对象：

1. 在“数据仓库中心”窗口中，展开**主题区**目录树。
2. 展开包含过程的**主题区**。
3. 右键单击**过程**文件夹，并单击**定义**。
“定义过程”笔记本打开。
4. 在**名称**字段中，输入过程的名称。
该名称最多可为 80 字符长，并区分大小写。名称的第一个字符必须是字母数字。不能使用英文中的 & 作为第一个字符。
5. 在**管理员**字段中，输入该过程对象的联系人姓名。
6. 可选：在**描述**字段中，输入过程的描述。
最多可输入 254 字符。
7. 可选：在**注释**字段中，输入有关过程的附加信息。

最多可输入 32,000 个字符。

8. 单击**安全性**选项卡。
9. 在**可用的仓库组**列表中，单击想要授予对该过程的访问权的仓库组，并单击 **>**。

若想要选择**可用的仓库组**列表中的所有仓库组，单击 **>>**。

选择的仓库组移动至**选择的仓库组**列表中。

10. 单击**确定**。

“定义过程”笔记本关闭。当展开**过程**文件夹时，新过程就会显示出来。

打开过程

打开过程以使用图形方式定义过程中的数据流。

要打开过程:

1. 右键单击该过程。
2. 单击**打开**。

添加源和目标到过程

要定义数据流，需添加步骤转换的每个源和转换产生的目标表。

要添加源或目标到过程:

1. 单击**添加数据**图符:



在画面上单击想要放置表的位置。“添加数据”窗口打开。

2. 在**可用的源和目标表**列表中，展开**仓库源**或**仓库目标**目录树。

显示在仓库中定义的仓库源或仓库目标列表。

3. 展开仓库源或仓库目标的目录树。

4. 展开**表、视图**或**文件**目录树。

5. 选择想要添加的表、视图或文件。

6. 单击 **>** 以添加该表、视图或文件到**选择的源和目标表**列表。

要将仓库源或仓库目标的所有表、视图或文件添加到过程中，选择**表、视图或文件**文件夹，并单击 **>>**。

7. 单击**确定**将选择的表、视图和文件添加到过程中。如果在某个表已添加到过程中后对该表进行了更改，所作的更改会在保存该过程时保存下来。

选择的表、视图和文件会显示在“过程模型”窗口上。

添加步骤到过程中

需要添加定义如何移动源数据和将它们转换为目标数据的步骤。步骤主要有四种类型：

SQL 步骤

SQL 步骤使用 SQL SELECT 语句从仓库源中抽取数据，并生成 INSERT 语句以将数据插入到仓库目标表中。

仓库程序步骤

仓库程序步骤运行预定义的程序和实用程序。仓库程序步骤有如下几种类型：

- 文件程序
- OLAP 程序
- Visual Warehouse 版本 5.2 程序
- 为特定 DB2 通用数据库平台设计的程序

特定操作系统的仓库程序与该操作系统的代理包装在一起。当安装代理代码时，将同时安装仓库程序。

DSNUTILS 是一个 DB2 OS/390 版存储过程，它在 WLM 和 RRS 环境中运行。通过用户定义的存储过程接口，可以使用 DSNUTILS 来运行任何已安装的 DB2 实用程序。要获取关于配置 DSNUTILS 存储过程的信息，请参阅《仓库管理器安装指南》。

在初始化期间添加每个仓库程序的定义。要查看定义列表，在“数据仓库中心”主窗口中展开**程序和变换器**文件夹。

变换器步骤

变换器步骤是存储过程和用户定义函数，它们指定可用来转换数据的统计变换器或仓库变换器。可使用变换器清除、反转和旋转数据；生成主键和周期表；以及计算各种统计信息。

在变换器步骤中，指定统计变换器或仓库变换器中的一个。当运行该过程时，变换器步骤将数据写入一个或多个仓库目标。

变换器步骤有如下几种类型：

- 统计变换器
- 仓库变换器

除了变换器步骤之外，还有一种变换器是用户定义的函数。变换器可与 SQL 步骤一起使用。

特定操作系统的变换器与该操作系统的代理包装在一起。当安装代理代码时，会同时安装仓库程序。要获取关于在 OS/390 和 AS/400 系统上设置仓库变换器的信息，请参阅《仓库管理器安装指南》。

限制: DataJoiner 目标数据库不支持“数据仓库中心”变换器。

在初始化期间添加每个变换器的定义。要查看定义列表，在“数据仓库中心”主窗口中展开**程序和变换器**文件夹。

在使用变换器之前:

- 在目标数据库中安装变换器。
- 在“仓库目标”笔记本的“数据库”页中，创建变换器并注册它们。
- 在“仓库目标”笔记本的“数据库”页中，对变换器启用目标数据库。
- 更改每个变换器的程序定义以指定代理和目标资源。
- 确保源和目标表在同一数据库中。

在通过 iSeries (AS/400) 仓库代理来使用变换器之前:

1. 在 IWH 集合中创建 Java 存储过程。在 iSeries 命令行上输入 WRKLIB IWH，以检查 iSeries 系统上是否存在 IWH 集合。如果 IWH 集合不存在，可使用 CREATE COLLECTION IWH SQL 语句，在“iSeries SQL 交互式”屏幕上创建集合 IWH。
2. 完成以下步骤:
 - a. 设置环境。
 - b. 注册存储过程。
 - c. 启用变换器的目标数据库。

要获取关于完成这些步骤的更多信息，请参阅《DB2 仓库管理器安装指南》。

复制步骤

复制步骤将任何 DB2 关系数据库中的指定更改从一个位置（源）复制到另一个位置（目标），使两个位置中的数据保持同步。源和目标可在同一机器或分布式网络中的不同机器上的逻辑服务器（如 DB2 数据库或 DB2 OS/390 版子系统或数据共享组）中。

可以使用复制步骤使仓库表与操作表保持同步，不必在每次更新操作表时装入整个仓库表。利用复制，可以使用增量更新使您的数据保持最新。

SAP 步骤

使用“SAP 数据抽取”步骤从 SAP R/3 系统中抽取 SAP 业务对象，并将它们导入“数据仓库中心”。有关详情，参见第188页的『从 SAP R/3 系统中抽取数据』。

Web 流量步骤

使用“Web 流量轮询”可查询 WebSphere Site Analyzer 数据导入的状态，以确定当前的 Web 流量数据是否已从数据导入复制到 WebSphere Site Analyzer Webmart 数据库的表中。在该步骤处于“生产”方式时，不能更改它。有关详情，参见第190页的『从 WebSphere Site Analyzer 数据库中抽取数据』。

i2 步骤

使用 i2 步骤来调度对 i2 业务模型、分布式数据集市以及 OLAP 立方的更新。有关详情，参见第399页的『附录I. 创建 i2 步骤』。

若需要上述步骤中没有提供的某种函数，可以编写自己的仓库程序或变换器，并定义使用那些程序或变换器的步骤。有关编写自己的仓库程序的信息，参见第289页的『编写自己的程序以配合“数据仓库中心”使用』。

每个步骤组（除 SQL 组以外）都有几种步骤子类型。在除 SQL 组以外的所有情况中，可选择一个特定的步骤子类型来移动或转换数据。例如，ANOVA 变换器是“统计”变换器组的子类型。在 SQL 组中，只有一种 SQL 步骤类型。可使用这种步骤在仓库源和目标上执行 SQL 选择操作。

可以将关于步骤的信息（例如主题区、源表名和目标表名）打印到文本文件中。要将步骤信息打印到文件中，右键单击“过程模型”窗口中的步骤图符，然后单击打印—> 打印到文件，指定要将信息打印到的文件的名称。

步骤子类型

表第125页的表11通过第133页的表17按程序组列示步骤子类型。程序组是相关程序的逻辑分组。例如，所有提供的操纵文件的仓库程序在“文件”仓库程序组中。提供的仓库程序和变换器的程序组与“过程模型”窗口左边的图符相对应。

第125页的表11列示文件仓库程序。

表 11. 文件仓库程序

名称	描述	代理站点						OS/390参见...
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2	AS/400	OS/390参见...	
使用 FTP 复制文件 (VWPRCPY)	将代理站点上的文件复制到远程主机, 和从远程主机复制文件到代理站点。	✓	✓	✓	✓	✓	✓	第173页的『定义“使用 FTP 复制文件”(VWPRCPY) 程序的值』
运行 FTP 命令文件 (VWPFTP)	运行指定的任何 FTP 命令文件。	✓	✓	✓	✓	✓	✓	第175页的『定义“运行 FTP 命令文件”程序 (VWPFTP) 的值』
使用 ODBC 将数据导出到文件中 (VWPEXPT2)	在 ODBC 中注册的数据库内包含的表中选择数据, 并将该数据写入定界文件中。	✓	✓	✓	✓			第160页的『定义“使用 ODBC 将数据导出到文件中”(VWPEXPT2) 仓库程序的值』
提交 OS/390 JCL 作业流 (VWPMVS)	将 JCL 作业流提交给 OS/390 系统进行处理。	✓	✓	✓	✓		✓	第176页的『定义“提交 OS/390 JCL 作业流”(VWPMVS) 程序的值』

第126页的表12列示 DB2 仓库程序。

定义和运行过程

表 12. DB2 仓库程序

名称	描述	代理站点				AS/400	OS/390参见...
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2		
DB2 UDB 装入	将数据从定界文件装入到 DB2 UDB 数据库中，置换数据库中的现存数据或追加到数据库中的现存数据中。	✓	✓	✓	✓		第161页的『定义“DB2 通用数据库”装入程序的值』
DB2 AS/400 版置换装入 (VWPLOADR)	将数据从定界文件装入到 DB2 AS/400 版数据库中，用新数据置换数据库中的现存数据。					✓	第166页的『定义“DB2 UDB AS/400 版数据置换装入”(VWPLOADR)程序的值』
DB2 AS/400 版插入装入 (VWPLOADI)	将数据从定界文件装入到 DB2 AS/400 版表中，将新数据追加到数据库中的现存数据中。					✓	第162页的『定义“DB2 UDB AS/400 版数据插入装入”(VWPLOADI)程序的值』
DB2 OS/390 版装入	将记录装入表空间中的一个或多个表。						✓ 第170页的『定义“DB2 OS/390 版装入”程序的值』
DB2 数据导出 (VWPEXPT1)	将数据从本地 DB2 数据库导出到定界文件中。	✓	✓	✓	✓		第159页的『定义 DB2 UDB 导出 (VWPEXPT1) 仓库程序的值』

表 12. DB2 仓库程序 (续)

名称	描述	代理站点						OS/390参见...
		Windows NT 2000	或 AIX	Solaris Op. Env.	OS/2	AS/400		
DB2 runstats (VWPSTATS)	对指定的表运行 DB2 RUNSTATS 实用程序。	✓	✓	✓	✓		✓	第276页的『定义 DB2 UDB RUNSTATS 程 序的值』 第277页的『定 义 DB2 OS/390 RUNSTATS 程 序的值』
DB2 reorg (VWPREORG)	对指定的表运行 DB2 REORG 和 RUNSTATS 实用程 序。	✓	✓	✓	✓		✓	第269页的『定 义 DB2 通用数 据库 REORG 程序的值』 第270页的『定 义 DB2 OS/390 版 REORG 程 序的值』

定义和运行过程

表13列示 OLAP Server™ 仓库程序。

表 13. DB2 OLAP Server® 程序

名称	描述	代理站点						OS/390参见...
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2	AS/400		
OLAP 服务器: 自由文本数据装入 (ESSDATA1)	使用自由格式数据装入, 将数据从逗号定界的文本文件装入多维 “DB2 OLAP 服务器” 数据库中。	✓	✓	✓		✓		第260页的『定义 “OLAP 服务器: 自由文本数据装入” (ESSDATA1) 仓库程序的值』
OLAP 服务器: 使用装入规则将数据从文件中装入 (ESSDATA2)	使用装入规则, 将数据从源文本文件装入多维 “DB2 OLAP 服务器” 数据库中。	✓	✓	✓		✓		第261页的『定义 “OLAP 服务器: 使用装入规则从文件装入数据” (ESSDATA2) 仓库程序的值』
OLAP Server: 使用装入规则将数据从 SQL 表中装入 (ESSDATA3)	使用装入规则, 将数据从 SQL 表装入多维 “DB2 OLAP 服务器” 数据库中。	✓	✓	✓		✓		第262页的『定义 “OLAP 服务器: 使用装入规则从 SQL 表装入数据” (ESSDATA2) 仓库程序的值』
OLAP 服务器: 不使用装入规则将数据从文件中装入 (ESSDATA2)	不使用装入规则, 将数据从文本文件装入多维 “DB2 OLAP 服务器” 数据库中。	✓	✓	✓		✓		第263页的『定义 “OLAP 服务器: 不使用装入规则从文件装入数据” (ESSDATA4) 仓库程序的值』
OLAP 服务器: 从文件更新轮廓 (ESSOTL1)	使用装入规则, 从源文件更新 “DB2 OLAP 服务器” 轮廓。	✓	✓	✓		✓		第264页的『定义 “OLAP 服务器: 从文件更新轮廓” (ESSOTL1) 仓库程序的值』

表 13. DB2 OLAP Server[®] 程序 (续)

名称	描述	代理站点						OS/390参见...
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2	AS/400	OS/390参见...	
OLAP 服务器: 从 SQL 表更新轮廓 (ESSOTL2)	使用装入规则, 从 SQL 表更新 “DB2 OLAP 服务器” 轮廓。	✓	✓	✓		✓		第266页的『定义 “OLAP 服务器: 从 SQL 表更新轮廓” (ESSOTL2) 仓库程序的值』
OLAP 服务器: 缺省 (ESSCALC1)	调用与目标数据库相关的缺省 “DB2 OLAP 服务器” 计算脚本。	✓	✓	✓		✓		第258页的『定义 “OLAP 服务器: 缺省 calc” (ESSCALC1) 仓库程序的值』
OLAP 服务器: 用 calc 规则进行计算 (ESSCALC2)	将指定的计算脚本应用于 “DB2 OLAP 服务器” 数据库。	✓	✓	✓		✓		第259页的『定义 “OLAP 服务器: 使用计算规则的 Calc” (ESSCALC2) 仓库程序的值』

表14列示复制程序。

表 14. 复制程序

名称	描述	代理站点						OS/390参见...
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2	AS/400	OS/390参见...	
基本聚集	创建一个目标表, 它包含在指定的时间间隔内为用户表追加的数据聚集。	✓	✓	✓	✓	✓	✓	第181页的『定义用户复制、时间点或基本聚集复制步骤』
更改聚集	创建一个目标表, 它包含基于为源表记录的更改的数据聚集。	✓	✓	✓	✓	✓	✓	第183页的『定义更改聚集复制步骤』

定义和运行过程

表 14. 复制程序 (续)

		代理站点						
名称	描述	Windows		Solaris			OS/390参见...	
		NT 2000	或 AIX	Op. Env.	OS/2	AS/400		
时间点	利用添加的时间戳记列，创建与源表匹配的目标表。	✓	✓	✓	✓	✓	✓	第181页的『定义用户复制、时间点或基本聚集复制步骤』
分级表	创建一致更改数据表，该表可用作将数据更新为多个目标表的源。	✓	✓	✓	✓	✓	✓	第186页的『定义分级表复制步骤』
用户拷贝	在复制时创建一个与源表准确匹配的目标表。	✓	✓	✓	✓	✓	✓	第181页的『定义用户复制、时间点或基本聚集复制步骤』

因为某些 DB2 仓库程序的命令行界面从 Visual Warehouse 版本 5.2 变化而来，所以单独支持来自 Visual Warehouse 版本 5.2 的 DB2 仓库程序。表15列示版本 5.2 仓库程序。

表 15. Visual Warehouse 版本 5.2 仓库程序

		代理站点						
名称	描述	Windows		Solaris			OS/390参见...	
		NT 2000	或 AIX	Op. Env.	OS/2	AS/400		
DB2 置换装入 (VWPLOADR)	将数据从定界文件装入到 DB2 UDB 数据库中，用新数据置换数据库中的现存数据。	✓	✓	✓	✓			第346页的『定义“Visual Warehouse DB2 UDB 数据置换装入”(VWPLOADR)程序的值』

表 15. Visual Warehouse 版本 5.2 仓库程序 (续)

名称	描述	代理站点					
		Windows NT 或 2000	AIX	Solaris Op. Env.	OS/2	AS/400	OS/390参见...
DB2 插入装入 (VWPLOADI)	将数据从定界文件装入到 DB2 表中，将新数据追加到数据库中的现存数据中。	✓	✓	✓	✓		第344页的『定义“Visual Warehouse DB2 UDB 数据插入装入”(VWPLOADI)程序的值』
将文本文件装入到 DB2 UDB EEE 中 (仅 AIX)(VWPLDPR)	将数据从定界文件装入到 DB2 EEE 数据库中，用新数据置换数据库中的现存数据。		✓				第350页的『定义“Visual Warehouse 5.2 将文本文件装入 DB2 UDB EEE”(VWPLDPR)程序的值(仅 AIX)』
DB2 数据导出 (VWPEXPT1)	将数据从本地 DB2 数据库导出到定界文件中。	✓	✓	✓	✓		第343页的『定义“Visual Warehouse 5.2 DB2 UDB 数据导出”(VWPEXPT1)程序的值』
DB2 runstats (VWPSTATS)	对指定的表运行 DB2 RUNSTATS 实用程序。	✓	✓	✓	✓		第349页的『定义 Visual Warehouse 5.2 DB2 UDB RUNSTATS (VWPSTATS) 程序的值』

定义和运行过程

表 15. Visual Warehouse 版本 5.2 仓库程序 (续)

名称	描述	代理站点						OS/390参见...
		Windows NT 2000	或 AIX	Solaris Op. Env.	OS/2	AS/400	AS/400 版本 4 发行版 5 和更 高版本	
DB2 (VWPREORG)	reorg 对指定的表运行 DB2 REORG 和 RUNSTATS 实用程序。	✓	✓	✓	✓			第348页的『定义 Visual Warehouse 5.2 DB2 UDB REORG (VWPREORG) 程序的值』

有关提供的仓库程序的详情，参见联机帮助。

表16列示仓库变换器。

表 16. 仓库, 变换器

名称	描述	代理站点						OS/390参见...
		Windows NT 2000	或 AIX	Solaris Op. Env.	OS/2	AS/400 版本 4 发行版 5 和更 高版本		
清理数据	置换数据值、除去数据行、剪切数值、执行数字离散化并除去空白。	✓	✓	✓	✓	✓	✓	第 202 页的『清理数据』
生成键表	生成或修改现存表中的唯一键值序列。	✓	✓	✓	✓	✓		第 207 页的『生成键列』
生成周期表	创建包含生成日期、时间或时间戳记值的表，它还包含基于指定的参数或基于行的日期或 / 和时间值的可选列。	✓	✓	✓	✓	✓	✓	第 209 页的『生成周期数据』

表 16. 仓库, 变换器 (续)

代理站点								
名称	描述	Windows		Solaris		AS/400		OS/390参见...
		NT 2000	或 AIX	Op. Env.	OS/2	版本 4 发行版 5 和更 高版本		
反转数据	反转表的行和列, 使行变为列, 列变为行。	✓	✓	✓	✓	✓	✓	第 212 页的『反转数据』
旋转数据	将来自源表中所选列的相关数据分组到目标表中的单个列中。将来自源表中的数据指定为输出表中的特定数据组。	✓	✓	✓	✓	✓	✓	第 214 页的『旋转数据』

表17列示统计变换器。

表 17. 统计变换器

代理站点								
名称	描述	Windows		Solaris		AS/400		OS/390参见...
		NT 2000	或 AIX	Op. Env.	OS/2	版本 4 发行版 5 和更 高版本		
ANOVA	计算一向、二向和三向方差分析; 估计组间和组内的变化并计算估计的比率; 并计算 p 值。	✓	✓	✓	✓	✓	✓	第 219 页的『ANOVA 变换器』

定义和运行过程

表 17. 统计变换器 (续)

		代理站点						AS/400	
名称	描述	Windows		Solaris		OS/2	AS/400 版本 4 发行 版 5 和更 高版 本	OS/390参见...	
		NT 2000	或 AIX	Op. Env.					
计算统计信息	计算单个表的数据列的数目、总计、平均值、方差、标准偏差、标准误差、标准误差、最小值、最大值、范围和变差系数。	✓	✓	✓	✓	✓	✓	第 221 页的 『“计算统计信息”变换器』	
计算小计	使用具有主键的表计算按某个时间周期（每周、每半月、每月、每季度或每年）分组的数值的日常小计。	✓	✓	✓	✓	✓	✓	第 223 页的 『“计算小计”变换器』	
X 平方	执行 X 平方和 X 平方拟合优度测试以确定两个变量值之间的关系，以及值的分布是否符合预期。	✓	✓	✓	✓	✓	✓	第 226 页的 『“X 平方”变换器』	
相关	通过计算任意数目的输入列对上的相关系数 r、协方差、T 值和 P 值，计算两个属性的变化之间的关系。	✓	✓	✓	✓	✓	✓	第 228 页的 『“相关”变换器』	

表 17. 统计变换器 (续)

		代理站点						
名称	描述	Windows		Solaris		AS/400 版本 4 发行 版 5 和更 高版 本	OS/390参见...	
		NT 或 2000	AIX	Op. Env.	OS/2			
滑动平均值	计算简单滑动平均值、指数滑动平均值、滚动总和，重新分布事件以除去数据中的干扰、随机情况和峰值或谷值。	✓	✓	✓	✓	✓	✓	第 231 页的『“滑动平均值”变换器』
回归	通过执行反向的完整模型回归，显示两个不同变量间的关系以及这两个变量的相关程度。	✓	✓	✓	✓	✓	✓	第 234 页的『“回归”变换器』

表18列示的变换器是用户定义函数。

表 18. 用户定义函数变换器

		代理站点						
名称	描述	Windows		Solaris		AS/400	OS/390参见...	
		NT 或 2000	AIX	Op. Env.	OS/2			
格式化日期和时间	更改源表日期字段的格式	✓	✓	✓	✓			第217页的『更改日期字段的格式』

本章其余部分提供有关使用步骤的一般信息。后续各章提供有关定义和使用各种步骤子类型的详细信息。

将步骤与源和目标链接

在定义步骤的值之前，可以使用数据链路来连接步骤与适用的仓库源和目标。在某些情况下，“数据仓库中心”可以为您生成目标表。将步骤与源和目标链接，定义一个从源经过步骤转换到目标的数据流。

要设置步骤使用数据源，使用“过程模型”窗口：

1. 在“过程模型”窗口的左边窗格中单击步骤图符。要查看步骤图符的描述，选择**查看** → **图注**。
2. 在左边窗格中单击步骤图符之后，选择步骤子类型（若可用的话）。
3. 将光标移动至“过程模型”窗口中，并单击。步骤图符就显示在“过程模型”窗口中。

可将步骤与它的源和目标链接。

要链接步骤：

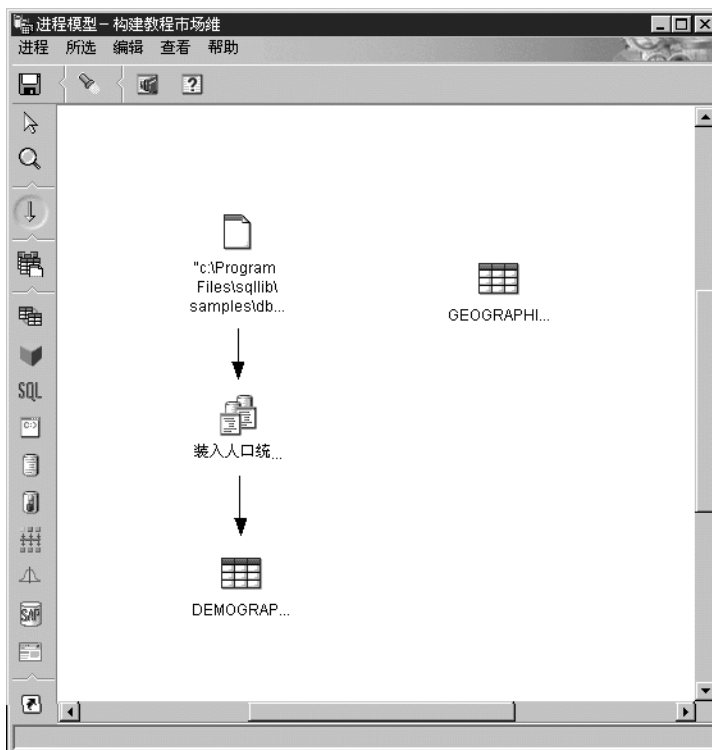
1. 单击**链路**图标：



2. 单击**数据链路**图标：



3. 单击源对象的中部并将它拖动到步骤上。
“数据仓库中心”在源和步骤之间画一条线。
该线指示该源包含步骤的源数据。
4. 单击步骤的中部并将它拖动到目标表。

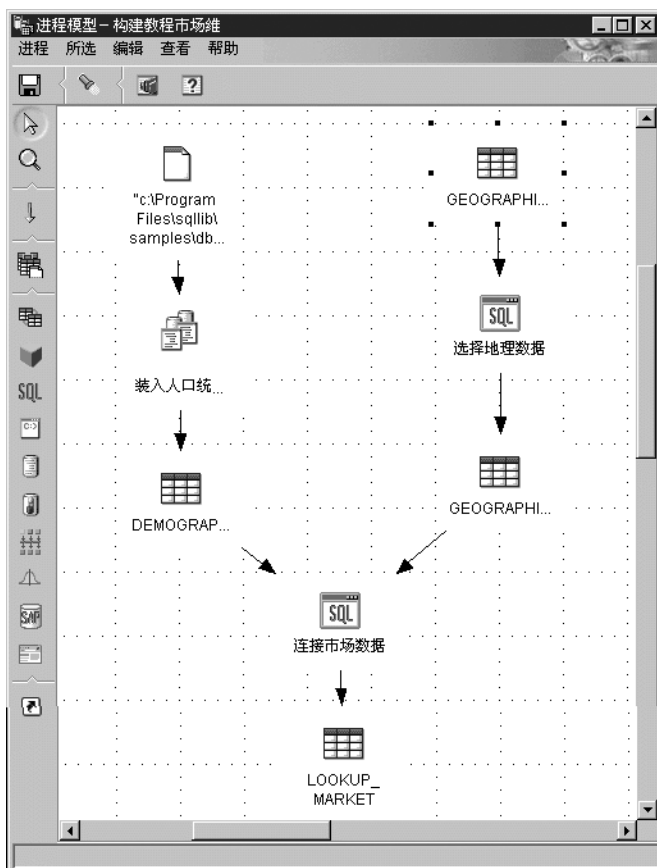


该线指示该目标表将包含步骤的目标数据。

如果更改源或目标的特性，所作的更改会自动保存。如果进行其他任何更改，例如添加一个步骤，则必须显式地保存该更改才能使它永久生效。要保存更改，单击**过程** → **保存**。

在以下示例中，有两个原始源。demographics.txt 文件包含某些城市的人口统计数据。GEOGRAPHIES 表包含有关在哪些地区销售哪些产品的信息。“装入人口统计数据”步骤将人口统计数据装入到 DEMOGRAPHICS_TARGET 目标表中。选择地理数据”步骤选择 GEOGRAPHIES 数据并将它写入到 GEOGRAPHIES_TARGET 目标表中。“连接市场数据”步骤连接两个目标表中的数据，并将数据写入到

LOOKUP_MARKET 目标表中。可使用产生的数据来分析按人口分的销售量。



有关前一示例的详情，包括定义源、步骤和目标的过程，参见《商业智能教程》。

定义步骤子类型的基本值

以下各节告诉您如何定义步骤的值。它们也提供有关步骤的基本信息，包括步骤子类型使用的数据源。例如，某些步骤子类型只使用仓库目标表，而其他子类型只使用仓库源或目标文件。

在定义步骤的值，首先必须打开步骤。右键单击步骤，然后单击**特性**。

每个步骤子类型笔记本由四页组成：

- 第一页由有关步骤的一般信息组成。在除“复制”步骤子类型之外的所有步骤子类型中，此页具有同样的字段和控件。

- “参数”页由定义步骤参数的值组成。此页具有每种步骤子类型的独特字段和控件。
- “列映射”页由“参数”页中选择的列和目标表中的列之间的映射信息组成。在某些情况下，可以从“列映射”页中创建一个目标表。并不是所有步骤都使用“列映射”页。
- “处理选项”页由定义如何运行步骤的值组成。在此页上每组步骤具有不同的字段和控件。

提供关于步骤的一般信息

步骤子类型笔记本的首页按该步骤类型命名。例如，ANOVA 变换器的第一页称作“统计变换器”。步骤子类型笔记本第一页上的所有字段对所有步骤子类型是相同的。

要定义步骤子类型笔记本第一页的值：

1. 在**名称**字段中，为步骤输入一个新名称。或者可保留“数据仓库中心”自动为该步骤提供的名称。
2. 可选：在**管理员**字段中，输入负责维护该步骤的人员的姓名。
3. 可选：在**描述**字段中，输入步骤的业务描述。描述最多可以是 254 个字符。
4. 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。

定义参数值

参见以下的步骤子类型描述，以获取有关定义步骤子类型的参数值的信息。

定义列映射信息

当使用“数据仓库中心”时，容易操纵数据。决定将在您的仓库数据库中使用源数据库中的哪些行和列（或字段）。然后，在步骤中定义那些行和列。

例如，想要创建一些与制造数据相关的步骤。每个制造点维护一个关系数据库，该数据库描述在该点制造的产品。为四个制造点中的每一个，创建一个步骤。第140页的图13显示源表和仓库表之间的初始映射。

定义和运行过程

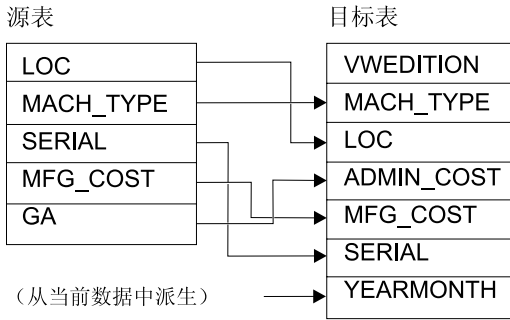


图 13. 源数据到仓库表的映射

只有某些步骤使用列映射。在您步骤定义参数值之后，若列映射页是空白的，这些值将产生在多个列中，且“生成缺省表”按钮不可用，因而您的步骤将不使用列映射。是否提供列映射信息是可选的。

在“列映射”页上，将从“参数”页上定义的转换产生的输出列映射到目标表上的列。在此页上，“参数”页的输出列称作源列。源列列示在该页的左边。与步骤链接的输出表中的目标列列示在该页的右边。使用“列映射”页执行下列任务：

- 要创建映射，可单击源列并将它拖动到目标列。在源列和目标列之间画出一个箭头。
- 要不使用鼠标创建映射，可使用跳格键和箭头键来选择要映射的源和目标列。然后跳格至映射按钮并按 **enter** 键。将在源列和目标列之间画出一个箭头。
- 要通过单击**添加列和映射**按钮来映射所有未映射的列。所有未映射的的源列都是基于列名、数据类型和大小属性映射到目标列的。如果没有匹配的目标列，将添加一个新的目标列。
- 要删除映射，右键单击箭头，并单击**除去**。
- 要重命名目标列，双击列名并输入新名称。也可以通过双击属性来更改目标列的其他任何属性。

创建缺省目标表：

1. 要基于您在步骤中指定的参数来创建缺省目标表：
2. 单击“生成缺省表”按钮。“生成缺省表”窗口打开。
3. 在“仓库目标”列表中，单击将定义目标表的仓库目标。
4. 可选：在“表模式”列表中，输入或单击目标表的表模式的名称。
5. 可选：在“表空间”列表中，输入或单击目标表的表空间的名称。
6. 在“表名”字段中，输入目标表的名称。

7. 单击“确定”。目标列显示在该页的右边。
- 8.

要从目标表中除去通过单击**生成缺省表**或**添加列和映射**按钮定义的列，可右键单击该列，然后单击**除去**。只有在创建了目标表后未保存到步骤笔记本的更改时，才能这样做。

对于某些步骤子类型，可在此页上执行的操作是受限的。对于其他步骤子类型，“参数”页的输出列可能要遵循某些规则。在以下的步骤子类型中的适当地方描述了此信息。

定义处理选项

本节描述需要为所有笔记本公共的“处理选项”字段和控件定义的值。

要为处理选项提供值：

1. 在“处理选项”页上，**填充类型**字段显示步骤子类型填充数据所使用的方法。在某些情况下，可从多个填充类型中选择。表19中显示了这些填充类型：

表 19. 填充类型

填充类型	描述
正常	为一定数目的版本追加数据，然后置换第一个版本的数据。例如，步骤在 12 个月内每月追加一次数据。在第 13 个月，用第 13 个月中运行产生的数据置换第一个月写入的数据。
追加	追加数据。
置换	置换数据。
程序控制	填充由程序管理。
删除	删除、重建和填充表。
复制	填充由复制管理。

2. 在**代理站点**列表中，选择想要运行该步骤的代理站点。此列表中的选项是源表、目标表和变换器或正定义的程序的公共代理站点。
3. 若想要选择随时运行步骤，选择**按需运行**复选框。若不选择此框，则不能从“工作进度”窗口运行该步骤。
4. 若需要重试步骤，在**重试区**中，指定想要该步骤再次运行多少次，以及在步骤下一次运行前想要经过的时间。

当出现下列其中一种问题时，“数据仓库中心”会重试步骤：

- 对一个步骤进行更改，未能运行的步骤依赖它。

定义和运行过程

- “数据仓库中心”无法将消息发送到代理守护进程。当试图运行使用未运行的代理的步骤时，可能出现此错误。
- “数据仓库中心”无法将消息发送到代理。当为代理指定了错误的主机名，或与代理的连接已停止时，可能出现此错误。
- “数据仓库中心”接收来自代理的启动通知时出错。
- “数据仓库中心”接收来自代理的消息时出错。
- 没有足够的页文件空间供该代理使用。
- 代理站点超载。

当返回下列其中一个返回码时，“数据仓库中心”也会重试数据抽取：

- DWC07902
 - DWC07903
 - DWC07904
 - DWC07905
 - DWC07906
 - DWC07907
5. 在**记录表**字段中为变换器步骤指定记录表。
 6. 在**跟踪级别**字段中为变换器步骤指定跟踪级别。

运行仓库步骤

“数据仓库中心”通过将步骤按开发、测试或生产三种方式分类，让您管理步骤的开发。方式确定是否可对步骤进行更改，以及“数据仓库中心”是否会按照步骤的调度表运行它。

开发仓库步骤

第一次创建步骤时，它处于开发方式。在此方式中可以更改任何步骤特性。“数据仓库中心”还没有为步骤在目标仓库中创建一个表。不能运行步骤来测试它，而且“数据仓库中心”不会按照该步骤的自动调度表运行它。

测试仓库步骤

运行步骤用数据填充它们的目标。然后可以验证结果是否符合期望。

在运行步骤之前，必须将它们提升为测试方式。

在步骤特性中，可以指定“数据仓库中心”将为该步骤创建一个目标表。当将步骤提升为测试方式时，“数据仓库中心”就创建该目标表。因此，将步骤提升为

测试方式之后，仅可对目标表进行非破坏性的更改。例如，当目标表的相关步骤处于测试方式时可向目标表添加列，但不能从该目标表除去列。

将这些步骤提升为测试方式之后，可以分别运行每个步骤。“数据仓库中心”不会按照步骤的自动调度表运行它。

将步骤提升为测试方式

要提升步骤：

1. 右键单击步骤。
2. 单击方式 → 测试。

确认窗口会询问您是否想要保存该处理。单击是。

“数据仓库中心”开始创建该目标表，并打开一个进度窗口。

要验证目标表已创建：

1. 若“DB2 控制中心”没有打开，在“数据仓库中心”主窗口中单击工具 → 控制中心来打开“DB2 控制中心”。
2. 展开对象目录树直到看见仓库数据库。
3. 展开数据库。
4. 展开表文件夹。
表的列表会显示在窗口的“内容”窗格中。
5. 验证目标表在该列表中。

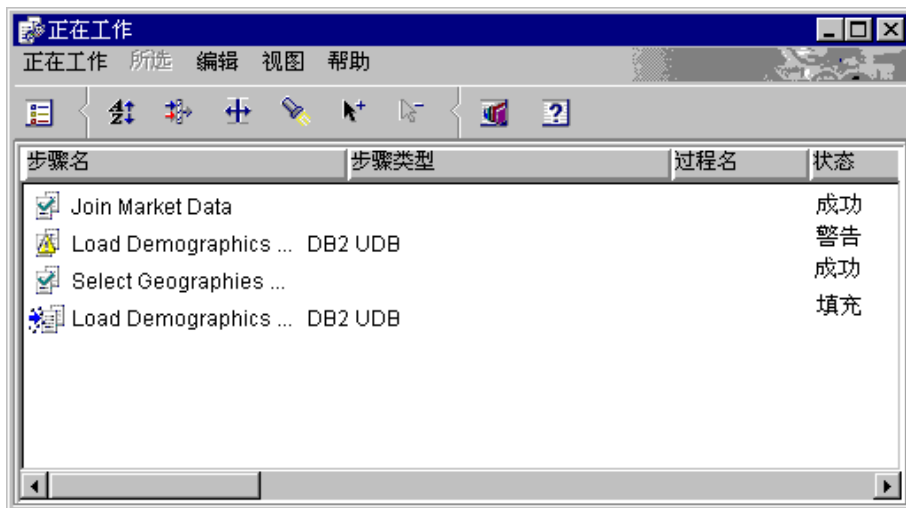
测试步骤

要测试步骤：

1. 在“过程模型”中，右键单击步骤。
2. 单击测试。

该步骤开始运行。“数据仓库中心”发出该步骤的 SQL 语句，或启动仓库程序或变换器。步骤停止运行之后，一个确认窗口打开。

3. 在“数据仓库中心”主窗口中，单击数据仓库中心 → 工作进度。
“工作进度”窗口打开。



使用“工作进度”窗口监视“数据仓库中心”中所有正运行或已调度的步骤的进度。应看见正在运行的步骤项。当该步骤正运行时，它处于**填充**状态。该处理应成功完成。但是，若处理失败：

有关“工作进度”窗口的详情，参见联机帮助中的『“工作进度 — 概述”』。

1. 选择步骤。
2. 单击**记录**。
“日志查看器”窗口打开。
3. 查找消息类型为“运行时错误”的日志记录。
4. 选择一条记录。
5. 单击**详细信息**。

“日志查看器详细信息”窗口打开。

若**错误 RC1** 字段的值为 8410，则程序在处理期间失败。在该程序的联机帮助的“返回码”节中，查找**错误 RC2** 字段的值，它是从该程序返回的值。

变换器错误消息与“数据仓库中心”中的其他信息不同：

- 变换器错误消息以 **DWC14** 开始。
- 变换器错误消息、警告消息和返回的 SQL 代码存储为辅助代码。因此，若消息以 **DWC14** 开始，则变换器（存储过程）导致该错误。若辅助代码包括 **SQLCODE**，则变换器中的 SQL 语句导致该错误。
- 仓库数据库中的输出日志表包含详细的错误消息、警告消息和 SQL 代码。在输出日志表中，消息类型是下列值之一：

E 错误

W 警告

Q SQL 代码

建议: 定期清理输出日志表, 使它们不包含过时的日志数据。

6. 查看程序日志文件以获取关于程序处理的更多信息。这些文件位于 `VWS_LOGGING` 环境变量指定的目录中。`VWS_LOGGING` 在 Windows NT 和 OS/2 上的缺省值是 `xprogram files\sqllib\logging`, 在 UNIX 和 OS/390 上是 `/var/IWH`。

`VWS_LOGGING` 的值是在“配置”笔记本中的跟踪日志目录字段的缺省值。若更改跟踪日志目录字段的值, “数据仓库中心”将日志文件写入指定的新目录, 但 `VWS_LOGGING` 的值不会更改。

首先查看 `trcppp.log` 文件。对于某些错误, 此文档会指示用于问题确定的其他日志文件。

一个常见的问题是将 Windows NT 仓库代理作为系统过程而不是用户过程运行导致的。当仓库代理作为系统过程运行时, 它无权与网络驱动器或产品连接, 因为该过程没有用户标识。此问题的症状包括: 仓库代理找不到仓库程序 (“日志查看器详细信息”窗口中的错误 **RC2** = 128 或错误 **RC2** = 1), 或不能初始化该程序。

若仓库代理作为用户过程运行, 则该仓库代理具有用户的特征, 包括访问用户有权访问的网络驱动器或程序的能力。

为避免这些问题, 执行下列步骤:

- a. 通过执行下列步骤, 将仓库服务器、仓库日志记录器和仓库代理守护进程服务更改为作为用户过程运行:
 - 1) 在 Windows NT 控制面板文件夹中, 双击服务图符。
 - 2) 停止那些服务。
 - 3) 选择服务, 并单击启动。
 - 4) 单击此帐户。
 - 5) 单击此帐户字段后的 ... 按钮选择用户标识。
该用户标识必须在 Windows NT 中拥有管理权限, 并有权访问任何必需的网络驱动器。
 - 6) 输入该用户标识的密码两次。
 - 7) 单击确定。
 - 8) 重新启动服务。
- b. 若正使用提供的 OLAP 服务器程序, 验证 DB2 OLAP 或 Essbase 客户机是否安装在执行该程序的代理的本地驱动器上。

定义和运行过程

- c. 若正使用提供的 OLAP 服务器程序，验证（Essbase 客户机或管理员上设置的）ARBORPATH 变量是否指定了执行该程序的代理的本地驱动器，和是否将其指定为系统变量。

要验证步骤的处理结果：

1. 在“过程模型”窗口中，右键单击目标表。
2. 单击**样本内容**。

“数据仓库中心”显示表中数据的子集。

也可从“DB2 控制中心”查看数据样本。右键单击目标表并单击**样本内容**。

从过程建模器在仓库目标表上运行“样本内容”时，仓库源表、仓库目标表和步骤必须有一个公共的代理站点，否则请求将失败。要避免这种问题，可对要使用的代理站点在“特性”笔记本中的“程序”页上选择步骤。

对目标文件运行“样本内容”时，将使用所选列表上的第一个代理站点。但是，数据库维护操作可能会影响到代理站点列示的次序。如果选择的代理站点与源或目标文件不在同一系统中，样本内容将失败。

调度仓库过程

要调度数据移动和转换，可使用“数据仓库中心”的调度能力，或可与另一个提供调度功能的产品配合使用“数据仓库中心”。

使用“数据仓库中心”的调度功能

在“数据仓库中心”中，启动步骤有两种方法。可指定步骤将在另一个步骤运行之后启动。或者可调度步骤在指定的日期和时间启动。在过程中可组合使用这些运行步骤的方法。可调度第一个步骤在指定的日期和时间运行。然后，可以指定另一个步骤在第一个步骤运行之后启动，并指定第三个步骤在第二个步骤运行之后启动，依次类推。

指定步骤按顺序运行： 若步骤使用另一个步骤转换的数据，可调度该步骤在另一个步骤完成处理之后启动。

要指定步骤按顺序运行：

1. 从“过程模型”窗口中，单击**任务流**图标：



2. 单击下列图符之一：

成功时 指示仅当步骤之前的步骤成功运行时，才会启动该步骤。

完成时 指示当步骤之前的步骤成功运行或失败时，会启动该步骤。

失败时 指示仅当步骤之前的步骤失败时，才会启动该步骤。

有关详情，参见联机帮助中的『调度步骤』。

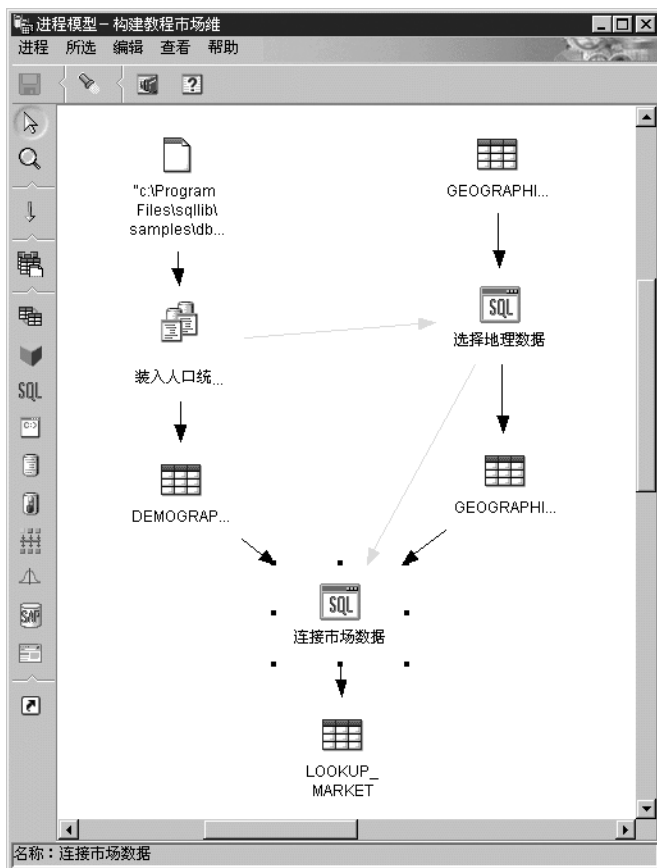
3. 单击将首先运行的步骤。
4. 按住鼠标按钮并将它拖动到将在第一个步骤之后运行的步骤。
5. 释放鼠标按钮。

表示该任务流的箭头会显示在这两个步骤之间的画面上。

6. 对处理序列中的每个步骤，重复步骤 3 到 5。

现在这些步骤将按您指定的次序运行。

例如，在下图中，“装入人口统计数据”步骤首先运行。当它完成运行时，“选择地理数据”步骤开始运行。当“选择地理数据”步骤完成运行时，“连接市场数据”步骤开始运行。



调度步骤在指定的日期和时间启动: 可调度步骤在指定的日期和时间启动。当调度步骤时，可指定步骤将在一个或多个日期和时间运行。也可指定步骤将运行一次或按指定的时间间隔运行，如每周星期六运行。

要调度步骤在指定的日期和时间启动。

1. 右键单击步骤，并单击**调度**。
“调度”笔记本打开。
2. 单击**调度**选项卡。
3. 在**时间间隔**列表中，单击运行该步骤想要按照的时间间隔。
缺省选择是**每周**。
4. 在**频率**列表中，单击频率和天。
缺省选择是**每周星期五**。

5. 在 **启动日期和时间** 字段中，选择步骤将首次运行的日期和时间。在第一次运行之后，该步骤将按您指定的时间间隔和频率运行。

缺省选择是当前日期，晚上 10:00。

6. 在 **结束** 字段中，指定何时结束调度：无限地或在指定的日期。

缺省选择是调度将无限地运行。

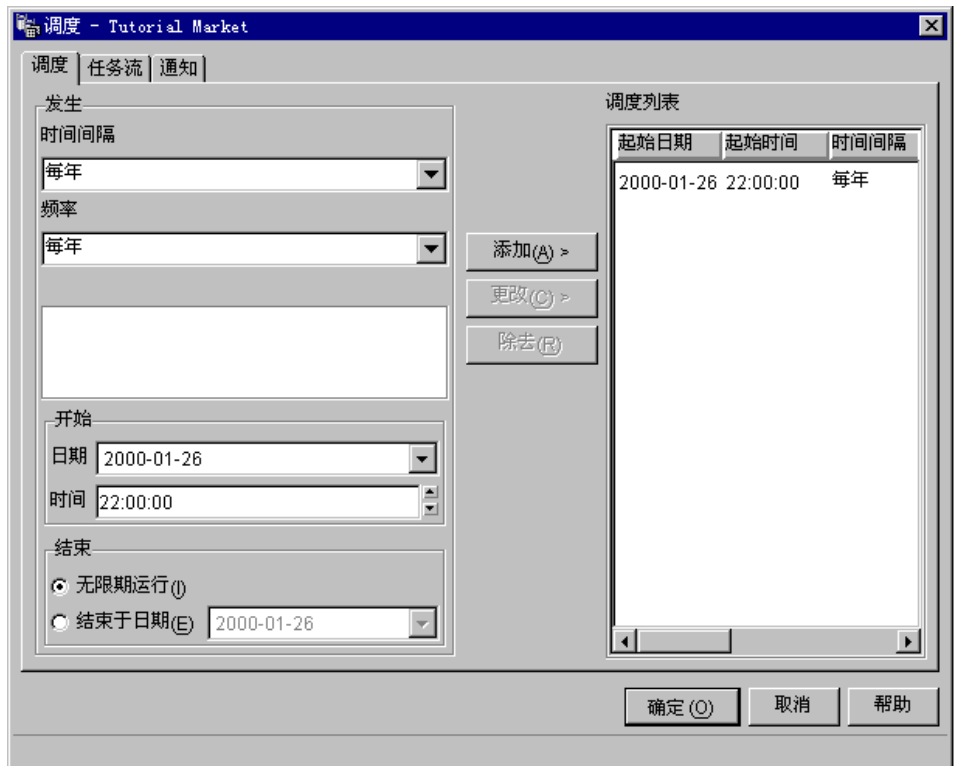
7. 单击 **添加**。

调度添加到 **调度列表**。

8. 单击 **确定**。

创建指定的调度。

在以下示例中，“装入人口统计数据”步骤将每年运行一次，从 2000 年 1 月 26 日开始。调度将不定期地运行。



与其他产品的调度功能配合使用“数据仓库中心”

当定义步骤时，可指定将如何填充它。可以将步骤定义为 **外部填充**，意指“数据仓库中心”将创建目标表，但将由另一个程序来填充它。这个另外的程序已经调度并在“数据仓库中心”之外运行。

定义和运行过程

例如，将您的步骤定义为外部填充。然后可以使用 DPropR 的调度能力填充目标表。

将步骤提升为生产方式

要激活创建的调度和任务流链接，必须将这些步骤提升为生产方式。生产方式指示这些步骤处于它们的最终格式。在生产方式中，仅可更改不会影响步骤产生的数据的那些设置。可以更改调度、处理选项（除填充类型外）或步骤的描述性数据。不能更改步骤的参数。

要将步骤提升为生产方式：

1. 右键单击步骤。
2. 单击方式 > 生产。

“数据仓库中心”打开一个进度窗口。

从“数据仓库中心”外部启动步骤

通过使用外部触发程序，可独立于“数据仓库中心”管理界面启动步骤。外部触发程序是一个调用“数据仓库中心”的仓库程序。

不能从外部触发程序运行过程。

要获取关于从 OS/390 启动外部触发器程序的信息，请参阅《仓库管理器安装指南》。

外部触发程序由 XTServer 和 XTClient 两个组件组成。XTServer 与仓库服务器安装在一起。XTClient 与所有代理类型的仓库代理安装在一起。

要使用外部触发程序来启动具有单字节名称的步骤，必须在仓库服务器工作站和代理站点上安装了 JDK 1.1.7。如果启动的是具有双字节名称的步骤，则必须安装 JDK 1.1.8。也可以使用随“数据仓库中心”和“控制中心”安装的 JDK。

启动外部触发服务器

在给外部触发客户机发出命令之前，必须启动外部触发服务器。

启动外部触发服务器语法如下所示：

XTServer

▶▶—java—XTServer—TriggerServerPort—————▶▶

TriggerServerPort

指定给外部触发服务器的 TCP/IP 端口。

此值可以是 11001、11002、11003 或 11004。

启动外部触发客户机

启动外部触发客户机的语法如下：

XTClient

▶▶—java—XTClient—ServerHostName— ServerPort—DWCUserID—DWCUserPassword—————▶▶

▶—StepName—Command—————▶▶
 └─WaitForStepCompletion─┘ └─RowLimit─┘

ServerHostName

安装仓库服务器的工作站的 TCP/IP 主机名。

指定一个全限定主机名。

ServerPort

给仓库服务器指定的 TCP/IP 端口。外部触发客户机必须使用与外部触发服务器相同的端口。

此值可以是 11001、11002、11003 或 11004。

DWCUserID

拥有“数据仓库中心操作”特权的用户标识。

DWCUserPassword

该用户标识的密码。

StepName

要启动的步骤的名称。

该名称区分大小写。若该名称包括空格，则把它用双引号 (『』) 引起来，如『Corporate Profit』。

Command

下列值之一：

1 填充

定义和运行过程

运行外部触发程序所用的用户标识必须与包含该步骤的过程在同一仓库组中。

2 提升为测试方式

运行外部触发器程序所用的用户标识必须与包含该步骤的过程在同一仓库组中。

3 提升为生产方式

运行外部触发器程序所用的用户标识必须与包含该步骤的过程在同一仓库组中。

4 降级为测试方式

运行外部触发器程序所用的用户标识必须与包含该步骤的过程在同一仓库组中。

5 降级为开发方式

运行外部触发器程序所用的用户标识必须与包含该步骤的过程在同一仓库组中。

WaitForStepCompletion

可选。此参数指示外部触发程序是否将返回步骤处理的结果。选择下列其中一个值：

1 等待步骤完成，若步骤成功返回 0，若步骤失败返回错误。

0 或空白

不等待步骤完成。

RowLimit

可选。此参数指示外部触发程序将抽取源表的所有行还是这些行的子集。可以使用此参数快速测试从大型数据库抽取数据的步骤。选择下列其中一个值：

0 或空白

抽取所有行

n 抽取 *n* 行

仅当步骤处于测试方式时，此参数才有效。

示例

例如，想要使用 db2admin 用户标识和 db2admin 密码来启动 Corporate Profit 步骤。外部触发程序在 dwserver 主机上。发出如下命令：

```
java XTClient dwserver 11004 db2admin db2admin "Corporate Profit" 1
```

状态

当运行外部触发程序时，它向仓库服务器发送消息。若该消息成功发送，外部触发程序返回零值返回码。

若为 *WaitForStepCompletion* 参数指定值 1，则该外部触发程序将一直等待，直到该步骤完成运行，然后返回该运行产生的返回码。

若外部触发程序未能向仓库服务器发送消息，它返回非零返回码。当有通信错误或认证失败时，返回码与“数据仓库中心”函数发出的相应代码匹配。有关“数据仓库中心”操作码的详情，参见《DB2 通用数据库消息与原因码》。

第6章 移动数据

“数据仓库中心”根据您的需要提供几种不同的方法来移动数据:

- 要移动少量数据, 可使用 SQL 步骤选择源数据并将它插入目标表。
- 要移动大量数据, 可使用导出和装入仓库程序将数据从表导出到文件中, 然后将该数据装入另一个表。若 SQL 步骤的性能满足不了您的需要, 尝试使用导出和装入仓库程序。
- 要移动对数据的更改而不是移动整个源, 对复制源使用复制仓库程序。

除了上述方法之外, “DB2 控制中心”具有可用来移动数据的导入和导出实用程序。有关详情, 参见 *DB2 Universal Database Data Movement Utilities Guide and Reference*。

选择和插入数据

可使用 SQL 步骤选择源列, 并将这些列中的数据插入目标表。可指定“数据仓库中心”基于源数据生成目标表, 或使用源数据更新现存表。

可使用仓库源或仓库目标作为 SQL 步骤的源。直到在“过程模型”窗口中将此步骤与源链接, “步骤”笔记本的“参数”页才可用于该步骤。也可在“过程模型”窗口中将此步骤与目标链接。若不将步骤与目标链接, 可指定在该步骤运行时创建表。

提示: 创建修订版化的 SQL 步骤时, 基于用法, 您可能会考虑对修订版列创建非唯一的索引, 以加速删除修订版的性能。应只对大仓库表考虑这样做, 因为在插入少量行时插入性能可能会受影响。

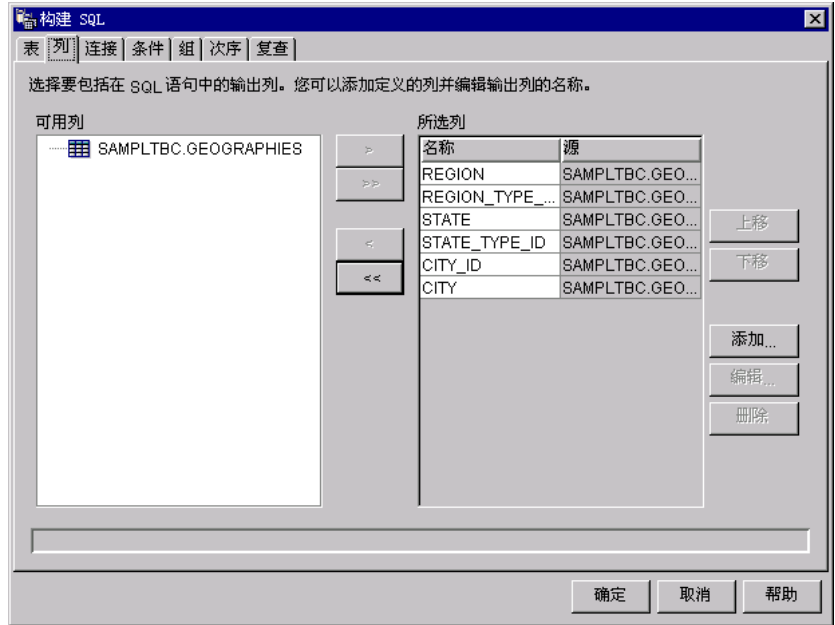
不能更改处于生产方式的 SQL 步骤。

定义步骤

要选择和插入数据:

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“SQL 语句”页上, 通过使用“SQL 辅助”或人工创建 SQL 语句:
 - 要使用“SQL 辅助”生成 SQL, 单击**构建 SQL**。“SQL 辅助”启动。

- a. 单击列选项卡。
- b. 从可用的列列表中，选择想要包括在目标表中的源列，并单击 >。
要将可用的列列表中的所有列包括在目标表中，单击 >>。



- c. 单击完成。
“SQL 辅助” 关闭。
- d. 单击列映射选项卡并查看“数据仓库中心”为您生成的列映射，或创建缺省目标表。参见第140页的『创建缺省目标表』获取关于从“列映射”页中创建缺省目标表的信息。

关闭“SQL 辅助”之后，新生成的 SQL 显示在 **SQL 语句** 字段中。

- 要人工创建 SQL 或编辑“SQL 辅助”生成的 SQL 语句，单击编辑。“SQL 语句” 字段变成可用。“SQL 辅助” 不再可用。

要擦除 SQL 语句字段中的文本，单击清除。若想要创建新的 SQL，必须再次单击编辑。

4. 可选：当完成生成或编辑 SQL 语句时，单击测试以测试 SQL 查询。“数据仓库中心” 返回查询的样本结果。
5. 可选：在“列映射” 页上，查看或修改列映射。有关详情，参见第139页的『定义列映射信息』。

6. 在“处理选项”页上的**填充类型**列表中，指定步骤是在目标中追加数据还是替换目标中的数据。有关填充类型的详情，参见第141页的『定义处理选项』。
7. 在**代理站点**列表中，选择想要运行该步骤的代理站点。此列表中的选择项是源表和目标表公共的代理站点。
8. 若想要选择随时运行步骤，选择**按需运行**复选框。必须使步骤处于测试或生产方式，才能运行它。
9. 可选：若由外部填充步骤，则选择**外部填充**复选框，意指以“数据仓库中心”以外的方式启动它。为使您将模式更改为生产，该步骤不需要以其他任何方式在“数据仓库中心”中运行。
若未选择**外部填充**复选框，为了将方式更改为生产，必须调度该步骤与是另一步骤的输入的瞬时表链接，或由另一程序启动该步骤。
10. 若需要重试步骤，在**重试**区中，指定想要该步骤再次运行多少次，以及在该步骤下一次运行之前要经过的时间。有关“重试”区的详情，参见第141页的『定义处理选项』。
11. 在**返回码**区中，选择在运行步骤时没有返回行或出现 SQL 警告的情况下，想要接收的返回码级别。
12. 在**增量落实**复选框和**每隔 x 行**字段中，指定是否想要指令执行增量落实，若想要，指定想要落实每隔多久执行一次。有关增量落实的详情，参见『增量落实』。
13. 单击**确定**保存更改并关闭步骤笔记本。

增量落实

增量落实是可用于所有 SQL 步骤的选项，允许控制“数据仓库中心”管理的数据的落实范围。当代理移动的数据量太大，以致在指令的全部工作完成之前 DB2 日志文件可能已填满时，或想要保存部分数据时，可使用增量落实。若正移动的数据量超过已分配的最大 DB2 日志文件，SQL 步骤完成时会返回错误。

增量落实选项允许指定在执行落实之前要处理的行数（舍入到 16 的最近因数）。代理以增量落实方式选择和插入数据，直到成功完成数据移动。当数据移动成功完成时，会除去过时的版本（若目标具有这样的版本）。

应考虑下列关于增量落实的情况：

- 若不指定增量落实，且返回了错误，数据将回滚到它的原始状态。
- 在填充类型为“替换”的步骤中，只有插入新数据时才使用增量提交。旧数据将在单个提交作用域期间删除。如果需要删除数据而不产生日志记录，可在运行填充类型为“附加”的 SQL 步骤前运行一个装入空文件的步骤。

移动数据

- 若发出落实命令之后出现错误，填充类型为“不带版本追加”的步骤可包含部分数据。
- 若发出落实命令之后出现错误，填充类型为“带版本追加”的步骤可包含部分数据，但将不删除过时的版本。
- 因为可能进行大量的落实，所以数据库的性能可能降低。
- 若源数据库和目标数据库相同，则不能优化 SQL。若指定增量落实，“数据仓库中心”会使用“选择并插入”逻辑，而不使用优化的“从选择插入”。
- 使用增量落实且具有版本的表可能需要附加的表空间，因为直到成功插入新版本之后，才删除过时的版本。

装入和导出数据

可使用提供的导出程序，如 DB2 数据导出从 DB2 数据库抽取数据，并将它写至文本文件。可使用提供的装入程序，如 DB2 置换装入从文件抽取数据，并将它写至另一 DB2 数据库。

建议：当要装入大量数据时，使用这些程序，而不要使用“数据仓库中心”的 SQL 处理。试验观察哪种方法对您效率更高。

这些大量装入和导出程序对定界数据文件和 DB2 数据库进行操作。数据库服务器不需要驻留在代理站点上，但源或目标文件必须驻留在代理站点上。输入和输出字段的数目和次序必须匹配。

这些程序将日志文件写至 VWS_LOGGING 环境变量指定的目录中。在 Windows NT 和 OS/2 上，VWS_LOGGING 的缺省值是 `x:\program files\sqlib\logging\`，在 UNIX 和 OS/390 上为 `/var/IWH`，其中 `x` 是您安装仓库代理的驱动器。

VWS_LOGGING 环境变量的值是“配置”笔记本中跟踪日志目录字段的缺省值。若更改跟踪日志目录字段的值，“数据仓库中心”将日志文件写至指定的新目录，但 VWS_LOGGING 的值不会更改。

要获取这些程序的参数、限制和返回码的完整列表，参见联机帮助中的『步骤和任务』。

以下各节描述如何定义“DB2 通用数据库”仓库程序的不同步骤子类型。

有关 DB2 UDB 导出和 DB2 UDB 装入仓库程序的各节描述如何定义这些程序的基本值。在联机帮助中描述了有关定义“DB2 通用数据库”插入和“DB2 通用数据库”装入仓库程序的高级特性的信息。

导出数据

可使用提供的仓库程序从 DB2 UDB 数据库或 ODBC 中定义的数据库导出数据。

定义 DB2 UDB 导出 (VWPEXPT1) 仓库程序的值

使用 DB2 UDB 导出的“步骤特性”笔记本创建一个步骤，该步骤可用于将数据从“DB2 通用数据库”表或视图导出到位于代理站点的文件。

数据库服务器不需要在该代理站点上。但目标文件必须在代理站点上。

指定目标文件在代理站点上使用的名称。若目标文件不存在，DB2 UDB 导出将创建它，若目标文件存在，DB2 UDB 导出将替换它。

要求：必须在“过程模型”窗口中将源表或视图与步骤链接。必须将步骤与仓库目标链接。

DB2 UDB 导出步骤不使用“列映射”页。

要定义运行“DB2 通用数据库”导出仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 可选：在“参数”页上，指定导出步骤的信息：
 - 在**定界符**组框中，通过单击向下箭头显示有效定界符的列表并选择一个，来更改定界符类型：
 - 列定界符指定用于对导出到目标文件的每列进行定界的字符。
 - 字符定界符指定用于包括导出到目标文件的字符列的字符。
 - 小数定界符指定在导出到目标文件的数字列中用作小数点的字符。
 - 在 **SELECT 语句** 字段中创建或更新 SQL 语句：
 - 单击**构建 SQL** 打开“SQL 辅助”，它将帮助您构建和生成 SQL 语句。“构建 SQL”窗口打开。
 - 单击**编辑**以编辑显示的 SQL 语句。“SQL 语句”字段变为可用。
 - 单击**测试**以测试显示的 SQL 语句。“样本内容”窗口打开。
 - 单击**清除**以清除显示的 SQL 语句。删除“SELECT 语句”字段中的 SQL 语句。
 - 要指定导出的高级选项，单击**高级**。有关指定这些选项的信息，参见此窗口的联机帮助。
4. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。

5. 单击**确定**保存更改并关闭步骤笔记本。

定义“使用 ODBC 将数据导出到文件中”(VWPEXPT2) 仓库程序的值

使用“使用 ODBC 将数据导出到文件中”(VWPEXPT2) 仓库程序在 ODBC 中注册的数据库中包含的表中选择数据，并将该数据写至定界文件。要在 AIX 或 UNIX 上运行此程序，使用仓库代理的 ODBC 版本。

此步骤子类型使用仓库源或目标文件作为源。在“过程模型”窗口中将源与步骤连接。在代理站点上生成输出文件。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- OS/2

“列映射”页不可用于此步骤。

要定义运行此仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的**列定界符**字段中，单击或输入想要用作列定界符的字符。
4. 使用“SQL 辅助”或人工创建或生成 SQL。
 - 要使用“SQL 辅助”生成 SQL 语句，单击**构建 SQL**。“SQL 辅助”打开。关闭“SQL 辅助”之后，新生成的 SQL 显示在 SELECT 语句字段中。
 - 要人工创建 SQL 或想要编辑“SQL 辅助”生成的 SQL 语句，单击**编辑**。“SELECT 语句”字段变为可用。“SQL 辅助”不再可用。

可选：若想要擦除 SELECT 语句字段中的文本，单击**清除**。若想要创建新 SQL，必须再次单击**编辑**。
5. 可选：当完成生成或编辑 SQL 时，单击**测试**以测试 SQL 查询。“数据仓库中心”返回查询的样本结果。
6. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
7. 单击**确定**保存更改并关闭步骤笔记本。

装入数据

可使用提供的仓库程序，将数据装入“DB2 通用数据库”、“DB2 AS/400 版和 DB2 OS/390 版”数据库。

定义“DB2 通用数据库”装入程序的值

使用“DB2 通用数据库装入步骤特性”笔记本创建一个步骤，该步骤将数据从源或目标文件装入“DB2 通用数据库”表。

可使用仓库源或目标作为此步骤子类型的源。在“过程模型”窗口中将源与步骤子类型链接。然后，将该步骤子类型链接到仓库目标，或指定“数据仓库中心”将创建目标表。

“列映射”页不可用于此步骤。

要定义运行“DB2 通用数据库”装入仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 可选：在“参数”页上，执行下列操作：
 - 从**装入方式**下拉列表中选择装入方式。可能值为：
 - INSERT - 装入表，并将文件中的数据追加到现存表。（这是缺省值。）
 - REPLACE - 装入表，并置换现存表中的所有数据。
 - RESTART - 重新启动异常终止的装入过程。
 - TERMINATE - 终止当前正在执行的装入过程。
 - 在**定界符**组框中，通过单击向下箭头显示有效定界符的列表并选择一个，来更改定界符类型：
 - 列定界符指定用于对要装入的文件中的每列进行定界的字符。
 - 字符定界符指定用于包括要装入的文件中的字符列的字符。
 - 小数定界符指定在要装入的文件中的数字列中用作小数点的字符。
 - 要指定装入的高级选项，单击**高级**。有关指定这些选项的信息，参见此窗口的联机帮助。
 - 要指定步骤是创建缺省目标表的，可选择**基于参数创建仓库目标表**复选框，并提供关于目标表的信息。参见第140页的『创建缺省目标表』以获取关于创建缺省目标表的字段通用信息。
4. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
5. 单击**确定**保存更改并关闭步骤笔记本。

定义 “DB2 UDB AS/400 版数据插入装入” (VWPLOADI) 程序的值

使用 “DB2 UDB AS/400 版数据插入装入” (VWPLOADI) 程序将数据从文本文件装入到 DB2 UDB AS/400 版表。该装入操作将新数据追加到表中现存数据的末尾。

在定义此步骤子类型之前，必须在 “过程建模器” 中将该步骤与仓库源和仓库目标连接。

VWPLOADI 的 AS/400 实现的可接受源文件是 “集成文件系统” (IFS) 根文件系统中的 AS/400 QSYS 源文件成员或流文件。

提示： 通过使用 QSYS 文件成员而不是流文件，可提高性能和存储器利用。CPYFRMIMPF 为 QRESTORE 制作整个流文件的副本，然后将副本装入您的表。有关详情，参见 CPYFRMIMPF 的联机帮助。

目标表仅可驻留在 QSYS 文件系统中。

仅当步骤处于开发方式时，才可对步骤进行更改。

在程序将新数据装入表之前，它将该表导出到备份文件，该备份文件可用于恢复。

“列映射” 页不可用于此步骤。

先决条件： 要使用此程序，必须使下列 PTF 应用于 AS/400 代理站点：

表 20. PTF 需求

OS/400 系统	产品标识	需要的 PTF	用于 VWPLOADI 或 VWPLOADR 的 PTF
版本 4 发行版 2	5769SS1	SF51534	SF46911、SF46976 和 SF49466
版本 4 发行版 3	5769SS1	SF51535	SF49467
版本 4 发行版 4	5769SS1	无	SF56124

这些 PTF 提供运行 VWPLOADI 程序所需的 AS/400 CPYFRMIMPF 和 CPYTOIMPF 命令 (LOAD 和 EXPORT)。这些 PTF 也安装这些命令的联机帮助。

安全性： 此程序和仓库代理运行所用的用户简要表对将装入的表必须至少拥有读 / 写权限。

要求: 下列要求适用于 VWPLOADI 程序。有关 CPYFRMIMPF 命令的限制的消息, 参见 CPYFRMIMPF 命令的联机帮助的限制部分。要查看此命令的联机帮助, 在 AS/400 命令提示符处输入 CPYFRMIMPF, 并按 F1 键。

1. 运行该程序的代理站点的“数据仓库中心”定义必须包括用户标识和密码。数据库服务器不需要在该代理站点上。然而, 源文件必须在数据库服务器上。按在 DB2 服务器系统上的定义指定源文件的全限定名。
2. 若程序在处理期间检测到故障, 则该表被清空。若装入过程出现警告, 则程序返回为成功完成。
3. VWPLOADI 的缺省行为是容忍 LOAD (ERRLVL(*NOMAX)) 期间的所有可恢复数据错误。

要进一步控制此行为, 在 fileMod 字符串参数中包括 ERLVL(n) 关键字, 其中 n = 允许的可恢复错误数。

可在 CPYFRMIMPF 命令的联机帮助中查找有关 ERLVL 关键字的详情。

要定义使用此程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 单击**参数**选项卡。
4. 可选: 要指定装入程序的定界符, 单击**仅指定 MODSTRING**的定界符。否则, 使用**记录、列、字符串和小数点**字段中的缺省项。

指定装入程序的定界符:

- 在**记录**列表中, 指定想要如何指示记录的结束。
 - 在**列**列表中, 指定想要用作列定界符的字符。
 - 在**字符串**字段中, 指定想要用于指示字符串的字符。
 - 在**小数点**字段中, 单击想要用于指示小数点的字符。
5. 可选: 要给程序添加参数, 单击**输入模式字符串参数**。在单选按钮下的字段中输入参数。

此字段用于修改 CPYFRMIMPF 命令期望输入文件具有的文件特性。若省略此参数, 则假设 CPYFRMIMPF 命令期望的所有缺省值都正确。

输入文件的一些缺省特性如下:

- 文件以逗号定界。
- 字符串和日期/时间值包括在引号中。
- 日期和时间值采用 ISO 格式。
- 小数点用句点字符表示。

有关 CPYFRMIMPF 命令的缺省值的详情，参见 CPYFRMIMPF 命令的 AS/400 联机帮助。

FileMod 字符串的格式为:

- a. 该字符串必须包含有效的 CPYFRMIMPF 命令关键字。在 CPYFRMIMPF 命令的联机帮助中描述了该命令的所有有效关键字。
- b. 每个关键字的后面必须紧跟它的值。该值必须包括在圆括号中。
- c. 每个关键字和下一个关键字之间必须用空格分隔。

要求: 某些参数要求将值包括在两个单引号内。例如，FLDDL M 命令必须将值包括在两个单引号内。“数据仓库中心”按如下格式生成 VWPLOADI 的 AS/400 CALL 语句:

```
CALL PGM(QIWH/VWPLOADI) PARM('fromfile' 'totable' 'filemodstring')
```

两个在一起的单引号告知 AS/400 命令提示处理程序，您的参数值包含一个单引号。这防止命令行处理器将单引号与一般的参数结束标记混淆。

6. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
7. 单击**确定**保存更改并关闭步骤笔记本。

跟踪和诊断信息: VWPLOADI 程序提供两种诊断信息:

- 返回码，编档在“数据仓库中心概念”联机帮助中
- VWPLOADI 跟踪

要点: 此程序的成功完成不保证已正确传送数据。要进行更精确的错误处理，使用 ERRLVL 参数。

阅读 VWPLOADI 跟踪文件: VWPLOADI 跟踪文件位于“集成文件系统”中的 /QIBM/UserData/IWH 目录中。

VWPLOADI 跟踪文件具有如下名称格式:

```
VWxxxxxxxx.VWPLOADI
```

其中 xxxxxxxx 是产生该文件的 VWPLOADI 运行的进程 ID。

要从工作站查看跟踪文件:

1. 使用 Client Access/400 将 AS/400 根文件系统映射为逻辑驱动器，或使用 FTP 将文件复制到工作站。
2. 用文本编辑器打开跟踪文件以查看信息。

用 *Client Access/400* 查看 *VWPLOADI* 跟踪： 要使用 *Client Access/400* 将 *AS/400* 系统映射为 NT 工作站上的逻辑驱动器：

1. 设置 *Client Access/400* 与 *AS/400* 系统的基于 TCP/IP 的连接。
2. 打开“Windows NT 文件浏览器”。
3. 在“资源管理器”菜单中，单击工具 → 映射网络驱动器。
4. 输入路径名：
\\hostname\.

其中 hostname 是 *AS/400* 系统的全限定 TCP/IP 主机名。

5. 单击确定。

要求： 若使用 *Client Access/400* 访问跟踪文件，必须对 *Client Access/400* 定义文件扩展名 *.VWPLOADI*。定义此扩展名允许 *Client Access/400* 将具有此扩展名的文件的内容从 EBCDIC 转换为 ASCII。

要对 *Client Access/400* 定义文件扩展名：

1. 在 Windows NT 中，选择开始 → 程序 → **IBM AS400 Client Access** → **Client Access 特性**。
Client Access 笔记本打开。
2. 单击网络驱动器选项卡。
3. 在文件扩展名：字段中，输入 *.VWPLOADR*。
4. 单击添加。
5. 单击应用。
6. 单击确定。

现在应能将该文件装入任何 ASCII 文本编辑器和文字处理器。

AS/400 异常： 若 *VWPLOADI* 发出的任何系统命令有故障，将在 *VWPLOADI* 跟踪文件中记录一个异常码。要获取异常的解释：

1. 在 *AS/400* 命令提示符处，输入 *DSPMSGD RANGE(xxxxxxx)*，其中 xxxxxxx 为异常码。例如，可输入 *DSPMSGD RANGE(CPF2817)*。

显示“显示格式化消息文本”屏面。

2. 选择选项 **30** 以显示所有信息。显示近似如下消息的消息：

```

消息标识. . . : CPF2817
消息文件 . . . : QCPFMSG
库 . . . . . : QSYS
消息 . . . . . : 因为错误, 复制命令结束。
原因 . . . . . : 在复制文件时
  
```

恢复 : 出错。
查看先前列示的消息。
校正错误，然后
再试该请求。

VWPLOADR 跟踪文件的第二行包含发出 WRKJOB 命令所需的信息。

要查看假脱机文件，可剪切并粘贴消息文件的名称到 AS/400 命令提示符处的 WRKJOB 命令之后，并按 Enter 键。查看该作业的假脱机文件以获取有关可能遇到的任何错误的附加信息。

定义“DB2 UDB AS/400 版数据置换装入”(VWPLOADR) 程序的值

使用“DB2 UDB AS/400 版数据置换装入”(VWPLOADR) 程序将数据从文本文件装入到 DB2 UDB AS/400 版表。装入操作完全置换表中的现存数据。

在定义此步骤子类型之前，必须在“过程建模器”中将该步骤与仓库源和仓库目标连接。

VWPLOADR 程序的 AS/400 实现的可接受源文件是“集成文件系统”(IFS) 根文件系统中的 AS/400 QSYS 源文件成员或流文件。

提示：通过使用 QSYS 文件成员而不是流文件，可提高性能和存储器利用。CPYFRMIMPF 为 QRESTORE 复制整个流文件，然后将副本装入您的表。有关详情，参见 CPYFRMIMPF 的联机帮助。

目标表仅可驻留在 QSYS 文件系统中。

仅当步骤处于开发方式时，才可对步骤进行更改。

“列映射”页不可用于此步骤。

先决条件：要使用此程序，必须将第162页的表20中列表的 PTF 应用到 AS/400 代理站点。这些 PTF 提供 AS/400 CPYFRMIMPF 和 CPYTOIMPF 命令 (LOAD 和 EXPORT)。这些是使 VWPLOADR 程序工作的命令。这些 PTF 也安装这些命令的联机帮助。

安全性：此程序和仓库代理运行所用的用户简要表对将装入的表必须至少拥有读 / 写权限。

要求：下列要求适用于 VWPLOADR 程序。有关 CPYFRMIMPF 命令的限制的消息，参见 CPYFRMIMPF 命令的联机帮助的限制部分。要查看此命令的联机帮助，在 AS/400 命令提示符处输入 CPYFRMIMPF，并按 F1 键。

- 运行该程序的代理站点的“数据仓库中心”定义必须包括用户标识和密码。数据库服务器不需要在该代理站点上。然而，源文件必须在数据库服务器上。按在 DB2 服务器上的定义指定源文件的全限定名。

- 若程序在处理期间检测到故障，该表将被清空。若装入出现警告，则程序返回为成功完成。

- VWPLOADR 程序的这种实现不同于其他平台上的 VWPLOADR。尤其是，若装入操作由于某种原因失败，它不删除所有已装入的记录。

正常情况下，此程序每次运行时置换目标表中的所有内容，并自动删除失败运行产生的记录。然而，若装入操作失败，应避免使用目标表中的数据。若目标表中有数据，它将不完整。

- VWPLOADR 的缺省行为是容忍 LOAD (ERRLVL(*NOMAX)) 期间的的所有可恢复数据错误。

要进一步控制此行为，在 fileMod 字符串参数中包括 ERRLVL(n) 关键字，其中 n = 允许的可恢复错误数。

可在 CPYFRMIMPF 命令的联机帮助中查找有关 ERRLVL 关键字的详情。

要定义一个过程，将本地文件装入本地表中。此过程将启动仓库程序“AS/400 通过替换装入”，然后将本地文件装入本地表中：

1. 定义您的源文件的平面文件仓库源。在**文件名字段**中，输入全限定文件名（参阅“iSeries 通过替换装入”文档，以了解命名规则）。
2. 通过仓库提供的“AS/400 通过替换装入”程序创建步骤。要获取关于创建“AS/400 通过替换装入”步骤的更多信息，参见要定义使用此程序的步骤的值。

3. 选择平面文件源，然后将源文件添加到步骤中。

限制：不能对文本文件使用 SQL 语句。SQL iSeries 仓库代理的版本 7 中不受支持。也不能通过 iSeries 仓库代理来使用平面文件的样本内容。数据必须使用“AS/400 通过替换装入”仓库程序来传送。

4. 从仓库目标选择目标表，并与该步骤连接。
5. 将该步骤提升到测试方式然后运行它。目标表现在包含了平面文件中的所有源数据。

要定义使用此程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 单击**参数**选项卡。

4. 可选: 若想要指定装入程序的定界符, 单击**仅指定 MODSTRING** 的定界符。若不单击此单选按钮, 则假设“记录”、“列”、“字符串”和“小数点”字段中的缺省项是正确的。
5. 若单击了**仅指定 MODSTRING** 的定界符, 则指定装入程序的定界符:
 - 在**记录**列表中, 指定想要如何指示记录的结束。
 - 在**列**列表中, 指定想要用作列定界符的字符。
 - 在**字符串**字段中, 指定想要用于指示字符串的字符。
 - 在**小数点**字段中, 单击想要用于指示小数点的字符。
6. 可选: 要给程序添加附加参数, 单击**输入模式字符串参数**。在单选按钮下的字段中输入参数。

此字段用于修改 CPYFRMIMPF 命令期望输入文件具有的文件特性。若省略此参数, 则假设 CPYFRMIMPF 命令期望的所有缺省值都正确。

输入文件的一些缺省特性如下:

- 文件以逗号定界。
- 字符串和日期/时间值包括在引号中。
- 日期和时间值采用 ISO 格式。
- 小数点用句点字符表示。

有关 CPYFRMIMPF 命令的缺省值的详情, 参见 CPYFRMIMPF 命令的 AS/400 联机帮助。

FileMod 字符串的格式为:

- a. 该字符串必须包含有效的 CPYFRMIMPF 命令关键字。在 CPYFRMIMPF 命令的联机帮助中描述了该命令的所有有效关键字。
- b. 每个关键字的后面必须紧跟它的值。该值必须包括在圆括号中。
- c. 每个关键字和下一个关键字之间必须用空格分隔。

注意: 某些参数要求将值包括在两个单引号内。例如, FLDDL M 命令必须将值包括在两个单引号内。因为“数据仓库中心”按如下格式生成 VWPL OADI 的 AS/400 CALL 语句, 所以必须这样做:

```
CALL PGM(QIWH/VWPL OADI) PARM('fromfile' 'totable' 'filemodstring')
```

两个在一起的单引号告知 AS/400 命令提示处理程序, 您的参数值包含一个单引号字符。这防止命令行处理器将您的单引号字符与一般的参数结束标记混淆。

7. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。

8. 单击**确定**保存更改并关闭步骤笔记本。

跟踪和诊断信息: VWPLOADR 程序提供两种诊断信息:

- 返回码, 编档在“数据仓库中心概念”联机帮助中
- VWPLOADR 跟踪

要点: 此程序的成功完成不保证已正确传送数据。要进行更精确的错误处理, 使用 ERRLVL 参数。

阅读 VWPLOADR 跟踪文件: VWPLOADR 跟踪文件位于“集成文件系统”中的 /QIBM/UserData/IWH 目录中。

VWPLOADR 跟踪文件具有下列名称格式:

VWxxxxxxxx.VWPLOADR

其中 xxxxxxxx 是产生该文件的 VWPLOADR 运行的进程 ID。

要从工作站查看跟踪文件:

1. 使用 Client Access/400 将 AS/400 根文件系统映射为逻辑驱动器, 或使用 FTP 将文件复制到工作站。
有关使用 Client Access/400 的信息, 参见『通过 Client Access/400 查看 VWPLOADR 跟踪』。
2. 用文本编辑器打开跟踪文件以查看信息。

通过 Client Access/400 查看 VWPLOADR 跟踪: 要使用 Client Access/400 将 AS/400 系统映射为 NT 工作站上的逻辑驱动器:

1. 设置 Client Access/400 与 AS/400 系统的基于 TCP/IP 的连接。
2. 打开“Windows NT 文件浏览器”。
3. 在“资源管理器”菜单中, 选择工具 → 映射网络驱动器。
4. 输入路径名:
\\hostname\.

其中 hostname 是 AS/400 系统的全限定 TCP/IP 主机名。

5. 单击**确定**。

要求: 若使用 Client Access/400 访问跟踪文件, 必须对 Client Access/400 定义文件扩展名 .VWPLOADR。定义此扩展名允许 Client Access/400 将具有此扩展名的文件的内容从 EBCDIC 转换为 ASCII。

要对 Client Access/400 定义文件扩展名:

1. 在 Windows NT 中，单击开始 → 程序 → **IBM AS400 Client Access** → **Client Access 特性**。
Client Access 笔记本打开。
2. 单击**网络驱动器**选项卡。
3. 在**文件扩展名:** 字段中，输入 **.VWPLOADR**。
4. 单击**添加**。
5. 单击**应用**。
6. 单击**确定**。

现在应能将该文件装入任何 ASCII 文本编辑器和文字处理器。

AS/400 异常: 若 VWPLOADR 发出的任何系统命令有故障，将在 VWPLOADR 跟踪文件中记录一个异常码。要获取异常的解释：

1. 在 AS/400 命令提示符处，输入 **DSPMSGD RANGE(xxxxxxx)**，其中 xxxxxxx 为异常码。例如，可输入 **DSPMSGD RANGE(CPF2817)**。

显示“显示格式化消息文本”屏面。

2. 选择选项 **30** 以显示所有信息。显示近似如下消息的消息：

```
消息标识 . . . : CPF2817
消息文件 . . . : QCPFMSG
库 . . . . . : QSYS
消息 . . . . . : 因为错误，复制命令结束。
原因 . . . . . : 在复制文件时
                  出错。
恢复 . . . . . : 查看先前列示的消息。
                  校正错误，然后
                  再试该请求。
```

VWPLOADR 跟踪文件的第二行包含发出 **WRKJOB** 命令所需的信息。

要查看假脱机文件，可复制并粘贴消息文件的名称到 AS/400 命令提示符处的 **WRKJOB** 命令之后，并按 **Enter** 键。查看该作业的假脱机文件以获取有关可能发生的任何错误的附加信息。

定义“DB2 OS/390 版装入”程序的值

使用“DB2 OS/390 版装入”程序将记录装入表空间中的一个或多个表。

要定义使用此仓库程序的步骤的值：

1. 定义 **DSNUTILS** 存储过程将运行的仓库目标。
2. 创建使用此仓库程序的步骤。
3. 将该步骤连接到目标。
4. 打开步骤笔记本。

5. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
6. 在“参数”页上，若想要预格式化与表相关的表空间和索引空间中的页，则选择**预格式化整个表空间和索引空间**复选框，以在与表相关的表空间和索引空间中预格式化高利用的 RBA（或页）和高分配的 RBA 之间的空闲页。预格式化在装入数据和构建索引之后进行。
7. 要指定是将记录装入空的还是非空的表空间，选择**指定表空间级别的 RESUME 选项**复选框。
 - 单击**否**将记录装入空的表空间。若表空间不是空的，而您未指定 **REPLACE**，则 **LOAD** 过程结束，并产生警告消息。对于包含已删除行或已删除表的行的未分段表空间，使用 **REPLACE** 选项使效率更高。
 - 单击**是**以用尽表空间，这可禁止不同分区的并行处理。若表空间是空的，则发出警告消息，但装入该表空间。装入在表空间中的当前数据结束处开始进行。不重复使用标记为已删除的行或已删除表的行占据的空间。

若表空间不是空的，而您不置换内容（未选择**装入之前将表空间和索引复位成空的**的复选框），则发出DB2 UDB OS/390 版和 z/OS 消息，实用程序作业步骤终止，产生作业步骤状态码 8。对于未分段的表空间，不使用标记为已删除的行或已删除表的行占据的空间。

8. 若想要新装入的行置换表空间中所有表的所有现存行，选择**装入之前将表空间和索引复位成空的**的复选框 (**LOAD REPLACE**)。
9. 指定其中一个**输入数据文件类型**单选按钮，以选择输入数据的代码页类型。若指定 **ASCII**，则不影响数字、日期、时间和时间戳记内部格式。
10. 选择**输入记录的格式**复选框，以标识输入记录的格式。格式必须与 **DB2 卸装格式 (UNLOAD)** 或 **SQL/DS™ 卸装格式**兼容。此操作唯一地确定输入的格式。**INTO TABLE** 选项中不允许字段规范。

DB2 卸装格式指定输入记录格式与 **DB2 卸装格式**兼容。**DB2 卸装格式**是具有 **UNLOAD ONLY** 选项的 **REORG** 的结果。将 **REORG** 实用程序卸装的输入记录装入从中卸装它们的表。不要在 **REORG UNLOAD ONLY** 和 **LOAD FORMAT UNLOAD** 之间添加或更改列规范。**DB2** 将记录重新装入从中卸装它们的相同表。

SQL/DS 卸装格式指定输入记录格式与 **SQL/DS 卸装格式**兼容。要装入的表中列的数据类型必须与 **SQL/DS** 表中相应列的数据类型相同。不能装入比 **DB2** 限制长的 **SQL/DS** 字符串。

11. 选择**不允许字符串中有替换字符**复选框，以禁止将替换字符放在字符串中作为转换的结果。当选择了此复选框且**DB2 UDB OS/390 版**和 **z/OS** 确定将一个替换字符放进了字符串中作为转换的结果时，它执行下列操作之一：

- 若废弃处理是活动的，DB2 发出消息 DSNU310I，并将记录放在废弃文件中。
 - 若废弃处理不活动，DB2 发出消息 DSNU334I，且该实用程序异常终止。
12. 选择输入文件的 **CCSID** 复选框，为输入数据指定最多三个编码字符集标识符 (CCSID)。在一个或所有字段中输入有效的整数值。缺省 CCSID 是安装“DB2 通用数据库 OS/390 版”时选择的 CCSID。若将与 CCSID 相关的任何字段留为空白，该字段将使用安装缺省值。可指定 SBCS、DBCS (双字节字符集) 数据或混合 DBCS (图形和双字节字符集)。
 13. 选择**强制检查**和**参考约束**复选框，让 DB2在有检查约束和参考约束时通知您。当强制约束且装入操作检测到违反时，DB2 删除错误行并发出标识它的消息。若选择强制约束且参考约束存在，则需要排序输入和排序输出数据集。
 14. 在**将在废弃数据集上写入的最大记录数字**段中，指定将在废弃数据集上写入的最大源记录数。值的范围从 0 到 2,147,483,647。若达到该最大数，LOAD 处理异常结束，废弃数据集为空，且不能查看刻废弃了哪些记录。缺省值为 0，意指没有最大值。在这种情况下，可废弃整个输入数据集。
 15. 在**废弃数据集**字段中输入废弃数据集的名称。若指定将废弃的记录数，则这是必需的工作数据集。此数据集保存未装入的记录的副本。它还保存已装入然后除去的记录的副本。它必须是 **BSAM** 服务可读的顺序数据集。**SYSDISC** 是与数据集名相关的数据定义名。
 16. 在**错误数据集**字段中，输入用于错误处理的工作数据集的名称。此数据集存储有关装入处理中出现的错误的信息。若指定了将在装入处理期间废弃的记录数，则此字段是必需的。**SYSERR** 是与数据集名相关的数据定义名。
 17. 单击**高级**打开“DB2 OS/390 版装入表空间”笔记本。必须单击此按钮否则装入操作将失败。在此笔记本中，可指定将数据装入表的附加选项。也可以选择收集表空间和 / 或索引的统计信息。统计信息存储在 DB2 OS/390 版目录中。
 18. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
 19. 单击**确定**保存更改并关闭步骤笔记本。

使用 **LOAD** 实用程序在 **DB2 z/OS** 版和 **OS/390** 版的表之间复制数据

如果想通过将表卸装到平面文件中，然后将该平面文件装入另一表中，从而复制表，通常需要卸装数据、编辑卸装过程的装入控制语句，然后再装入数据。使用 zSeries 仓库代理，可以指定要将数据重新装入另一表中而不在步骤之间停止，并可手工编辑控制语句。

要使用 **LOAD** 实用程序在 **DB2 z/OS** 版和 **OS/390** 版的表之间复制数据:

1. 使用“类属”接口创建一个步骤，该步骤使用 UNLOAD 实用程序或 REORG TABLESPACE 实用程序来卸装文件。这两个实用程序会产生两个输出数据集，一个带有表数据而另一个则带有可添加到 LOAD 实用程序的实用程序控制语句。

以下是一个 DSNUTILS 参数示例，可用于“Reorg 卸装”步骤：

```
UTILITY_ID REORGULX
RESTART NO
UTSTMT REORG TABLESPACE DBVW.USAINENT UNLOAD EXTERNAL
UTILITY_NAME REORG TABLESPACE
RECDSN DBVW.DSNURELD.RECDSN
RECDEVT SYSDA
RECSpace 50
PNCHDSN DBVW.DSNURELD.PNCHDSN
PNCHDEVT SYSDA
PNCHSPACE 3
```

2. 使用“类属”DSNUTILS 接口来创建装入步骤。DSNUTILS 实用程序语句参数指定了一个实用程序控制语句。仓库实用程序接口允许实用程序语句字段中的文件名。可以使用 :FILE: 关键字来指定包含有效控制语句的文件，也可以使用 :TABLE: 关键字来指定要装入的表的名称。
3. 要使用 LOAD 实用程序来处理先前示例中的输出，可在 LOAD 特性中应用以下参数值：

```
UTILITY_ID LOAD
RESTART NO
UTSTMT :FILE:DBVW.DSNURELD.PNCHDSN:TABLE:[DBVW].INVENTORY
UTILITY_NAME LOAD
RECDSN DBVW.DSNURELD.RECDSN
RECDEVT SYSDA
```

4. 在 UTSTMT 字段中，输入从带有 UNLOAD EXTERNAL 选项的 REORG 实用程序产生的装入语句或文件名称。先前的示例将对任何 DB2 z/OS 版和 OS/390 版源表或目标表起作用，不论这些表是在相同的还是不同的 DB2 子系统上。控制语句平面文件可以是 HFS 或本机 OS/390 文件。

要获取关于可用于 OS/390 或 z/OS 的 DSNUTILS 和 DB2 实用程序的更多详细信息，参见 *DB2 for z/OS and OS/390 Utility Guide and Reference*。

操纵文件

以下各节描述如何定义仓库文件程序的不同步骤子类型。

定义“使用 FTP 复制文件” (VWPRCPY) 程序的值

使用仓库程序“使用 FTP 复制文件” (VWPRCPY) 在代理站点和远程主机之间复制文件。

在将文件复制到 OS/390 之前，必须分配它们的数据集。不能传送 VSAM 数据集。当定义使用此仓库程序的步骤时，选择源文件和目标文件。一个文件必须存储在代理站点上，另一个必须存储在 OS/390 系统上。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- AS/400
- OS/390
- OS/2

“列映射”页不可用于此步骤子类型。

要定义运行此仓库程序的步骤：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，单击将文件从远程主机复制到代理站点（GET 方法）或将文件从代理站点复制到远程主机（PUT 方法）。
4. 在“传送类型”区中，单击 ASCII 或二进制。
5. 在远程系统名字段中，输入从中复制文件或将文件复制到的远程系统的名称。
6. 在远程用户标识 字段中，输入用于与远程主机连接的用户标识。
7. 在远程密码字段中，输入该用户标识的密码。
8. 在验证远程密码字段中，再次输入该密码。
9. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。

如果访问安全 UNIX 或“UNIX 系统服务”系统上的远程文件时有问题，请验证您是否对用户代理站点定义中指定的 \$HOME 目录有写访问权。

例如，代理站点的主机名为 glacier.stl.ibm.com。您想要使用用户标识vwinst2，使用 FTP 将文件从远程站点 kingkong.stl.ibm.com 传送至代理站点。vwinst2/.netrc 文件必须包含如下项目：

```
machine glacier.stl.ibm.com login vwinst2
```

定义“运行 FTP 命令文件”程序 (VWPFTP) 的值

使用仓库程序“运行 FTP 命令文件”(VWPFTP)使用 FTP 从远程主机传送文件。当定义使用此仓库程序的步骤时，不要为该步骤指定源表或目标表。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- AS/400
- OS/390
- OS/2

“列映射”页不可用于此步骤。

要定义运行此仓库程序的过程：

1. 在主“数据仓库中心”窗口中，右键单击**仓库源文件夹**，然后单击**定义** → **平面文件** → **远程文件**。在“文件”页上，单击**高级**，然后指定主机名、用户标识和密码。

限制：不能对文本文件使用 SQL 语句。SQL iSeries 仓库代理的版本 7 中不受支持。

2. 在 iSeries 上定义本地平面文件目标。
3. 通过仓库的“使用 FTP 程序复制文件”来创建一个步骤。
4. 将步骤连接到远程源和本地目标。
5. 打开步骤笔记本。
6. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
7. 在“参数”页上的**远程系统名字段**中，输入要访问的远程系统的名称。
8. 在 **FTP 命令文件（全路径名）**字段中输入 FTP 命令文件的路径和文件名。
在 FTP 命令文件中，必须在单独的行上按列示的次序提供下列信息：
 - a. 用户标识
 - b. 密码
 - c. 一个或多个 FTP 命令，每个命令单独一行

在以下示例中，使用用户标识和密码使用 FTP 登录到远程主机，获取远程文件，并将它放在指定的本地目录中：

```
nst1
password
get /etc/services d:/udprcpy.out
quit
```

9. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。
11. 将此步骤与“AS/400 通过替换装入”步骤链接。
12. 将此本地目标文件作为源与“AS/400 通过替换装入”步骤连接。
13. 将本地目标表与该步骤连接。

两个组合步骤可在 iSeries 上的本地文件上装入远程平面文件，然后再将数据装入表中。

如果访问安全 UNIX 或“UNIX 系统服务”系统上的远程文件时有问题，请验证您是否对用户代理站点定义中指定的 \$HOME 目录有写访问权。

例如，代理站点的主机名为 glacier.stl.ibm.com。您想要使用用户标识vwinst2，使用 FTP 将文件从远程站点 kingkong.stl.ibm.com 传送到代理站点。vwinst2/.netrc 文件必须包含如下项目：

```
machine glacier.stl.ibm.com login vwinst2
```

定义“提交 OS/390 JCL 作业流” (VWPMVS) 程序的值

使用“提交 OS/390 JCL 作业流” (VWPMVS) 仓库程序，将驻留在 OS/390 上的 JCL 作业流提交到 OS/390 系统用于执行。

作业必须使 MSGCLASS 和 SYSOUT 按路径发送到保留输出类。

在使用“提交 OS/390 JCL 作业流”仓库程序之前，通过在计划用于该程序的同一用户标识下从 TSO 运行您的 JCL 文件，来测试它。

若 OS/390 主机名、用户标识和密码正确，则此仓库程序运行成功。若想要测试 JCL 产生的结果的有效性，必须编写自己的测试逻辑。

“提交 OS/390 JCL 作业流”仓库程序也在代理站点上接收 JES 日志文件。在提交新作业进行处理之前，它擦除代理站点上来自任何先前作业的 JES 日志文件的副本。在作业完成之后，它还验证 JES 日志文件已下载到代理站点。

“提交 OS/390 JCL 作业流”仓库程序需要在 OS/390 上安装 TCP/IP 3.2 或更高版本。在使用该程序之前，验证已启用 FTP 服务。

当定义使用此仓库程序的步骤时，不要为该步骤指定源表或目标表。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- OS/390
- OS/2

“列映射”页不可用于此步骤。

要定义运行此仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的 **MVS 系统名字** 字段中，输入要运行该作业的 OS/390 主机的名称。
4. 在 **MVS 用户标识** 字段中，输入与 OS/390 主机连接所用的用户标识。
5. 在 **MVS 密码** 字段中，输入该用户标识的密码。
6. 在 **验证 MVS 密码** 字段中，再次输入该密码。
7. 在 **MVS JCL 文件** 字段中，输入要提交的 JCL 文件的名称。名称必须由用户标识加一个字符组成。
8. 在 **本地假脱机文件全路径名** 字段中，输入代理站点上接收 JES 日志文件的文件的路径和文件名。
如果 .netrc 文件不存在，仓库程序会在与 JES 文件相同的目录中创建该文件。
如果 .netrc 文件存在，仓库程序会重命名现有文件并创建一个新文件。在程序处理完成后，它会删除它创建的这个新 .netrc 文件，并重命名原始文件为 .netrc。
9. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。

若访问安全的 UNIX 系统上的远程文件有问题，验证该用户标识的主目录包含 .netrc 文件。 .netrc 文件必须包含一个项目，它包括代理站点的主机名和您想使用的远程用户标识。

例如，代理站点的主机名为 glacier.stl.ibm.com。您想要使用用户标识vwinst2，使用 FTP 将文件从远程站点 kingkong.stl.ibm.com 传送到代理站点。vwinst2/.netrc 文件必须包含如下项目：

```
machine glacier.stl.ibm.com login vwinst2
```

在“数据仓库中心”中设置复制

通过“数据仓库中心”来设置并运行复制需要仓库控制数据库和仓库目标数据库上同时存在复制控制表。

复制控制表必须同时存在于控制数据库和目标数据库中。复制控制表可在 ASN 模式中找到，并且都以 IBMSNAP 开始。如果控制表不存在，它们会当您在“控制中心”中定义复制源时，在数据库中自动创建。

要创建控制表，可编辑在定义复制源时生成的 SQL。除去插入到 IBMSNAP_REGISTER 表中的 SQL 语句以及创建 CD 表的语句。也可使用 DataJoiner Replication Administration (DJRA) 来创建控制表。

要在“数据仓库中心”中设置复制，必须完成以下步骤。

1. 可选：安装 DJRA。

注：复制非 IBM 数据库时 DJRA 是必需的，也建议对 AS/400 系统使用。要获取关于安装 DJRA 的更多信息，请参阅 *Replication Guide and Reference*。

2. 设置 Capture and Apply 程序。要获取关于设置 Capture and Apply 程序的信息，请参阅 *Replication Guide and Reference*。

必须在每个将要使用 Capture and Apply 程序的数据库上绑定这些程序。不需要创建密码文件。“数据仓库中心”将自动为复制预订创建密码文件。

“数据仓库中心”包括了用于复制支持的 JCL 模板。如果计划使用 zSeries 仓库代理来运行 Apply 程序，需要针对 OS/390 或 z/OS 系统更改此模板中的帐户和数据设置信息。有关信息，请参阅《仓库管理器安装指南》。

3. 在“控制中心”或 DJRA 中定义复制源。
4. 在“数据仓库中心”中定义复制步骤。
5. 在与源数据库相同的系统上启动 Capture 程序。
6. 启动 Apply 程序。建议您在与目标数据库相同的系统上启动 Apply 程序，但这不是必需的。

启动 Apply 程序时：

- 使用为步骤定义的不同 Apply 限定符

- 使用目标服务器作为启动 Apply 程序时的控制服务器名称

要获取关于 Apply 程序的参数的更多信息，参见 *Replication Guide and Reference*。

7. 在“数据仓库中心”中将复制步骤提升为测试或生产。

复制表

复制是维护多个位置中的已定义数据集的过程。它涉及将指定的更改从一个位置（源）复制到另一个位置（目标），并使两个位置中的数据保持同步。源和目标可在同一机器或分布式网络中的不同机器上的逻辑服务器（如 DB2 数据库或 DB2 OS/390 版子系统或数据共享组）中。

当想要使仓库表与操作表保持同步，而不必在每次更新操作表时装入整个仓库表时，可使用“数据仓库中心”的复制能力。利用复制，可以使用增量更新使您的数据保持最新。

可使用“数据仓库中心”定义一个复制步骤，它将在任何 DB2 关系数据库之间复制更改。也可使用其他 IBM 产品（如 DB2 DataJoiner 和 IMS DataPropagator™ NonRelational）或非 IBM 产品（如 Microsoft SQL Server 和 Sybase SQL Server），在许多（关系和非关系）数据库产品之间复制数据。所需的复制环境取决于想要何时更新数据和想要如何处理事务。

要用“数据仓库中心”定义复制步骤，您必须属于对将使用该步骤的过程拥有访问权的仓库组。

“数据仓库中心”支持五种复制：

用户复制

产生是复制源的只读副本的目标表，不添加任何复制控制列。这些表看起来象一般的源表，是用于复制的好起点。它们是最常见的目标表类型。

时间点 产生是复制源的只读副本的目标表，添加时间戳记列。时间戳记列最初是空的。当复制更改时，会添加值以指示进行更新的时间。若想要跟踪更改的时间，可使用这种类型的表。

基本聚集

产生概述源表内容的只读表。基本聚集复制表对跟踪一般基础的源表的状态很有用。聚集表使用 SQL 列函数（如 SUM 和 AVG）计算源表的所有内容的汇总，或最近对源表数据进行的更改的汇总。

更改聚集

产生只读表，这些表总结了更改数据表而不是源表的内容。这种复制对跟踪在每个 Apply 程序周期之间进行的更改很有用。

分级表 产生包含已落实事务产生的数据的只读表。也称作一致更改数据表（CCD 表），若这些表是压缩的、非压缩的、完整的或不完整的，它们会包含不同的数据。

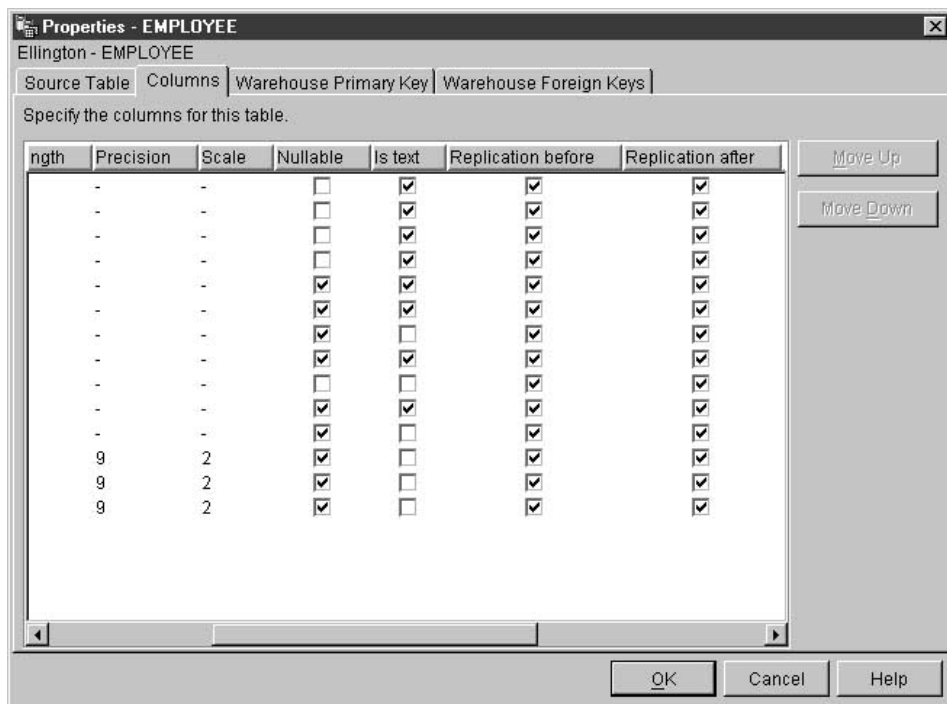
- 压缩的分级表只包含行的最新值。Apply 程序只更新已在压缩表中的行。压缩表对将对远程位置的更改分级和在将热点更新复制到目标之前汇总它们很有用。
- 非压缩的分级表包含行的更改历史。Apply 程序将行追加到非压缩表。非压缩的分级表对审核用途很有用。
- 完整的分级表包含想要从源表复制的所有行。
- 不完整的分级表在创建时空的，在对源表进行更改时追加行。

对于复制步骤，提升为测试方式将创建目标表并生成预订集。第一次运行复制步骤时，会制作全新的副本。将复制步骤提升为生产方式将启用已定义的调度。仅当步骤处于开发方式时，才可对它进行更改。

在“数据仓库中心”中定义复制源

可用定义其他关系源的相同方法，使用“数据仓库中心”定义复制源。除了可从数据库返回的其他数据库对象（表、视图和系统表）之外，可选择返回允许复制的表和视图。在表或视图可用作“数据仓库中心”中的复制源之前，必须使用“DB2 控制中心”或 DJRA 将它定义为用于复制。有关将表或视图定义为复制源的指示信息，参见 *DB2 Replication Guide and Reference*。

当将允许复制的表定义为仓库源表时，前映象列和后映象列由列名之后的前或后标识。



若选择只从源数据库中检索允许复制的表，将仅检索表内允许复制的列。

有关在“数据仓库中心”中定义复制源的指示信息，参见第51页的『定义 DB2 仓库源』。

定义用户复制、时间点或基本聚集复制步骤

用于用户复制或时间点复制步骤的源表必须具有主键。使用“DB2 控制中心”，为想要在用户复制或时间点复制步骤中包括的每个表定义一个主键。

要定义用户复制、时间点或基本聚集复制步骤：

1. 定义一个过程对象。
2. 打开该过程对象。
3. 添加一个或多个仓库源。
4. 添加一个或多个仓库目标。
5. 打开步骤笔记本。

6. 指定步骤的信息:
 - 在**名称**字段中，可为该步骤输入新名称。否则，可保留“数据仓库中心”自动为该步骤提供的名称。
 - 可选：在**管理员**字段中，输入负责维护该步骤的人员的姓名。
 - 可选：在**描述**字段中，输入步骤的业务描述。描述最多可以是 254 个字符。
 - 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。
7. 在“参数”页上，从**可用的列**列表中选择想要复制的列，并单击 **>**。选择的列移动至**选择的列**列表。**可用的列**列表仅显示已对更改捕捉启用的列。
要包括**可用的列**列表中的所有项目，单击 **>>**。
8. 可选：单击**添加计算列**打开一个窗口，可在其中创建派生列。创建的派生列显示在**选择的列**列表中。
9. 可选：要选择将复制哪些行，编写 **WHERE** 语句以对行进行子查询。
10. 在“列映射”页上，将您在“参数”页上定义的输出列映射为目标表中的列。在此页上，“参数”页的输出列称作**源列**。源列列示在该页的左边。与步骤链接的输出表中的目标列列示在该页的右边。参见第139页的『定义列映射信息』获取关于使用“列映射”页可执行的任务的信息。
如果“参数”页不产生输出列，或此步骤不与目标表链接，则不能使用此页来映射列。某些步骤不允许更改列映射。
11. 在“处理选项”页上，从**代理站点**下拉列表中选择想要运行您的步骤的代理站点。此列表中的选择项是源表和目標表的公用代理站点。
12. 复制步骤的“填充”类型仅可具有一个值，即**复制**。
13. 若想要选择随时运行步骤，选择**按需运行**复选框。步骤必须处于测试或生产方式，才能运行它。
14. 可选：若步骤是外部填充的，意指通过不同于“数据仓库中心”的方法启动它，则选择**外部填充**复选框。为将模式更改为生产，步骤不需要以其他任何方式在“数据仓库中心”中运行。
若未选择**外部填充**复选框，为将方式更改为生产，必须调度步骤与是另一步骤的输入的瞬时表链接，或由另一程序启动它。
15. 若需要重试步骤，在**重试**区中，指定想要该步骤再次运行多少次，以及想要在步骤下一次运行前经过的时间。
16. 在**复制控制数据库**字段中，选择包含 **Apply** 程序的复制控制表的控制数据库或子系统。
17. 在**数据库类型**列表中，选择复制控制数据库的数据库类型。
18. 在**用户标识** 字段中，输入访问复制控制数据库的用户标识。
19. 在**密码**字段中，输入将访问该数据库的用户标识的密码。

20. 在**验证密码**字段中，再次输入该密码。
21. 在**预订集名字段**中，输入预订集的名称。此名称最多可包含 18 个字符，且可以是普通或定界限定符。
22. 可选：在 **Apply 限定符** 字段中，输入 Apply 限定符的名称。对于定义的每个复制步骤，Apply 限定符名称必须唯一。若不指定 Apply 限定符，“数据仓库中心”会为您生成一个。
23. 可选：在**事件名字段**中，输入事件名。事件名是 Apply 程序读取的事件表中记录的事件的名称。对于定义的每个复制步骤，事件名必须唯一。若不指定事件名，“数据仓库中心”会为您生成一个。
24. 在**分块系数**字段中，指定预订循环期间可用于复制更改数据的分钟数。
25. 单击**确定**保存更改并关闭笔记本。
26. 将步骤与仓库源链接。
27. 将步骤与仓库目标链接。
28. 启动 Capture 程序。在启动 Apply 程序前必须启动 Capture 程序。要启动 Capture 程序，可在 DOS 窗口中输入以下命令：


```
ASNCCP source-database COLD PRUNE
```

COLD 参数指定 COLD 开始并删除 CD 表中任何现有的数据。PRUNE 参数告诉 capture 程序要维护 IBMSNAP_PRUNCNTL 表。

在 Capture 程序仍在运行时，转至步骤 。
29. 启动 Apply 程序。可在启动了 Capture 程序后的任何时候启动 Apply 程序。要获取关于启动 Apply 程序的信息，参见 *Replication Guide and Reference*。
30. 将步骤提升为测试方式。
31. 运行步骤测试它。

步骤运行时，将更新 IBMSNAP_SUBS_EVENT 表中的事件记录而 IBMSNAP_SUBS_SET 中的预订记录是活动的。如果在更新事件记录后更新原始源表而预订记录是活动的，更改的数据将移动到 CD 表中。如果在更改的数据移动到 CD 表后运行复制步骤，而 Apply 程序再次运行，则更改的数据将从 CD 表移动到目标表中。
32. 调度步骤。
33. 将步骤提升为生产方式。

定义更改聚集复制步骤

更改聚集复制步骤产生使用控制表中的更改数据的表，但不包含源表的内容。

要定义更改聚集复制步骤：

1. 定义一个过程对象。
2. 打开该过程对象。
3. 添加一个或多个仓库源。
4. 添加一个或多个仓库目标。
5. 打开步骤笔记本。
6. 指定步骤的信息：
 - 在**名称**字段中，可为该步骤输入新名称，或可保留“数据仓库中心”自动为该步骤提供的名称。
 - 可选：在**管理员**字段中，输入负责维护该步骤的人员的姓名。
 - 可选：在**描述**字段中，输入步骤的业务描述。描述最多可以是 254 个字符。
 - 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。
7. 在“参数”页上，从**可用的列**列表中选择想要复制的列，并单击 **>**。选择的列移动至**选择的列**列表。**可用的列**列表仅显示已对更改捕捉启用的列。
若想要包括**可用的列**列表中的所有项目，单击 **>>**。
8. 可选：单击**添加计算列**打开一个窗口，可在其中创建派生列。创建的派生列显示在**选择的列**列表中。
9. 可选：要选择将复制哪些行，编写 **WHERE** 语句以对行进行子查询。
10. 可选：要添加计算列，添加 **GROUP BY** 语句。可根据 **GROUP BY** 语句中定义的组对行进行分组。
11. 在“列映射”页上，将您在“参数”页上定义的输出列映射为目标表上的列。在此页上，“参数”页的输出列称作**源列**。源列列示在该页的左边。与步骤链接的输出表的目标列列示在该页的右边。参见第139页的『定义列映射信息』获取关于使用“列映射”页可执行的任务的信息。
如果“参数”页不产生输出列，或此步骤不与目标表链接，则不能使用此页来映射列。某些步骤不允许更改列映射。
12. 在“处理选项”页上，从**代理站点**下拉列表中选择想要运行您的步骤的代理站点。此列表中的选择项是源表和目标的公用代理站点。
13. 复制步骤的“填充”类型仅可具有一个值，即**复制**。
14. 若想要选择随时运行步骤，选择**按需运行**复选框。步骤必须处于测试或生产方式，才能运行它。
15. 可选：若步骤是外部填充的，意指通过不同于“数据仓库中心”的方法启动它，则选择**外部填充**复选框。为将模式更改为生产，步骤不需要以其他任何方式在“数据仓库中心”中运行。
若未选择**外部填充**复选框，为将方式更改为生产，必须调度步骤与是另一步骤的输入的瞬时表链接，或由另一程序启动它。

16. 若需要重试步骤，在**重试**区中，指定想要该步骤再次运行多少次，以及想要在步骤下一次运行前经过的时间。
17. 在**复制控制数据库**字段中，选择包含 Apply 程序的复制控制表的控制数据库或子系统。
18. 在**数据库类型**列表中，选择复制控制数据库的数据库类型。
19. 在**用户标识**字段中，输入访问复制控制数据库的用户标识。
20. 在**密码**字段中，输入将访问该数据库的用户标识的密码。
21. 在**验证密码**字段中，再次输入该密码。
22. 在**预订集名字段**中，输入预订集的名称。此名称最多可包含 18 个字符，且可以是普通或定界限定符。
23. 可选：在 **Apply 限定符**字段中，输入 Apply 限定符的名称。对于定义的每个复制步骤，Apply 限定符名称必须唯一。若不指定 Apply 限定符，“数据仓库中心”会为您生成一个。
24. 可选：在**事件名字段**中，输入事件名。事件名是 Apply 程序读取的事件表中记录的事件的名称。对于定义的每个复制步骤，事件名必须唯一。若不指定事件名，“数据仓库中心”会为您生成一个。
25. 在**分块系数**字段中，指定预订循环期间可用于复制更改数据的分钟数。
26. 单击**确定**保存更改并关闭笔记本。
27. 将步骤与仓库源链接。
28. 将步骤与仓库目标链接。
29. 要启动 Capture 程序，可在 DOS 窗口中输入以下命令：

```
ASNCCP source-database COLD PRUNE
```

COLD 参数指定 COLD 开始并将删除 CD 表中任何现有的数据。PRUNE 参数告诉 capture 程序要维护 IBMSNAP_PRUNCNTL 表。保留 Capture 程序运行。在启动 Apply 程序前必须启动 Capture 程序。
30. 启动 Apply 程序。可在启动了 Capture 程序后的任何时候启动 Apply 程序。要获取关于启动 Apply 程序的信息，参见 *Replication Guide and Reference*。
31. 将步骤提升为测试方式。
32. 运行步骤测试它。

步骤运行时，将更新 IBMSNAP_SUBS_EVENT 表中的事件记录而 IBMSNAP_SUBS_SET 中的预订记录则设置为活动的。如果在更新事件记录后更新原始源表而预订记录是活动的，更改的数据将移动到 CD 表中。如果在更改的数据移动到 CD 表后运行复制步骤，而 Apply 程序再次运行，则更改的数据将从 CD 表移动到目标表中。
33. 调度步骤。

34. 将步骤提升为生产方式。

定义分级表复制步骤

分级表复制步骤产生包含来自已落实事务的数据的只读表。用于分级表复制步骤的源表必须具有主键。使用“DB2 控制中心”，为想要包括在分级表复制步骤中的每个表定义一个主键。

要定义分级表复制步骤：

1. 定义一个过程对象。
2. 打开该过程对象。
3. 添加一个或多个仓库源。
4. 添加一个或多个仓库目标。
5. 打开步骤笔记本。
6. 指定步骤的信息：
 - 在**名称**字段中，可为该步骤输入新名称。否则，可保留“数据仓库中心”自动为该步骤提供的名称。
 - 可选：在**管理员**字段中，输入负责维护该步骤的人员的姓名。
 - 可选：在**描述**字段中，输入步骤的业务描述。描述最多可以是 254 个字符。
 - 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。
7. 在“参数”页上，从**可用的列**列表中选择想要复制的列，并单击 **>**。选择的列移动至**选择的列**列表。**可用的列**列表仅显示已对更改捕捉启用的列。
要包括“可用的列”列表中的所有项目，单击 **>>**。
8. 可选：单击**添加计算列**打开一个窗口，可在其中创建派生列。创建的派生列显示在**选择的列**列表中。
9. 可选：要选择将复制哪些行，编写 **WHERE** 语句以对行进行子查询。
10. 可选：单击**分级表选项**以更改表的任何特性。若选择“分级表”的复制类型，则此选项可用。
11. 在“列映射”页上，将您在“参数”页上定义的输出列映射为目标表上的列。在此页上，“参数”页的输出列称作**源列**。源列列示在该页的左边。与步骤链接的输出表中的目标列列示在该页的右边。参见第139页的『定义列映射信息』获取关于使用“列映射”页可执行的任务的信息。
如果“参数”页不产生输出列，或此步骤不与目标表链接，则不能使用此页来映射列。某些步骤不允许更改列映射。
12. 在“处理选项”页上，从**代理站点**下拉列表中选择想要运行您的步骤的代理站点。此列表中的选择项是源表和目標表的公用代理站点。
13. 复制步骤的“填充”类型仅可具有一个值，即**复制**。

14. 若想要选择随时运行步骤，选择**按需运行**复选框。步骤必须处于测试或生产方式，才能运行它。
15. 可选：若步骤是外部填充的，意指通过不同于“数据仓库中心”的方法启动它，则选择**外部填充**复选框。为将模式更改为生产，步骤不需要以其他任何方式在“数据仓库中心”中运行。
若未选择**外部填充**复选框，为将方式更改为生产，必须调度步骤与是另一步骤的输入的瞬时表链接，或由另一程序启动它。
16. 若需要重试步骤，在**重试区**中，指定想要该步骤再次运行多少次，以及想要在步骤下一次运行前经过的时间。
17. 在**复制控制数据库**字段中，选择包含 Apply 程序的复制控制表的控制数据库或子系统。
18. 在**数据库类型**列表中，选择复制控制数据库的数据库类型。
19. 在**用户标识**字段中，输入访问复制控制数据库的用户标识。
20. 在**密码**字段中，输入将访问该数据库的用户标识的密码。
21. 在**验证密码**字段中，再次输入该密码。
22. 在**预订集名**字段中，输入预订集的名称。此名称最多可包含 18 个字符，且可以是普通或定界限定符。
23. 可选：在 **Apply 限定符**字段中，输入 Apply 限定符的名称。对于定义每个复制步骤，Apply 限定符名称必须唯一。若不指定 Apply 限定符，“数据仓库中心”会为您生成一个。
24. 可选：在**事件名字段**中，输入事件名。事件名是 Apply 程序读取的事件表中记录的事件的名称。对于定义每个复制步骤，事件名必须唯一。若不指定事件名，“数据仓库中心”会为您生成一个。
25. 在**分块系数**字段中，指定预订循环期间可用于复制更改数据的分钟数。
26. 单击**确定**保存更改并关闭笔记本。
27. 将步骤与仓库源链接。
28. 将步骤与仓库目标链接。
29. 启动 Capture 程序。在启动 Apply 程序前必须启动 Capture 程序。要启动 Capture 程序，可在 DOS 窗口中输入以下命令：

```
ASNCCP source-database COLD PRUNE
```


COLD 参数指定 COLD 开始并删除 CD 表中任何现有的数据。PRUNE 参数告诉 capture 程序要维护 IBMSNAP_PRUNCNTL 表。在 Capture 程序仍在运行时，转至步骤 30。
30. 启动 Apply 程序。可在启动了 Capture 程序后的任何时候启动 Apply 程序。要获取关于启动 Apply 程序的信息，参见 *Replication Guide and Reference*。

31. 将步骤提升为测试方式。
32. 运行步骤测试它。

步骤运行时，将更新 IBMSNAP_SUBS_EVENT 表中的事件记录而 IBMSNAP_SUBS_SET 中的预订记录是活动的。如果在更新事件记录后更新原始源表而预订记录是活动的，则更改的数据将移动到 CD 表中。如果在数据移动到 CD 表后运行复制步骤，而 Apply 程序再次运行，则更改的数据将从 CD 表移动到目标表中。
33. 调度步骤。
34. 将步骤提升为生产方式。

从 SAP R/3 系统中抽取数据

通过“SAP 数据抽取”步骤，可将 SAP R/3 系统中定义的 SAP 业务对象的数据导入 DB2 数据仓库中。业务对象与业务组件提供了一个 R/3 业务函数的面向对象的视图。然后可使用 DB2 和“数据仓库中心”的功能来进行数据分析、数据转换或数据挖掘。

“SAP R/3 连接器”在 Microsoft Windows NT 或 Windows 2000（带服务包 1 或更高版本）上运行。SAP R/3 服务器可在任何受支持的操作系统上。

定义 SAP 源

定义 SAP 源时，您会见到关于 SAP 业务对象的所有元数据，包括键字段和参数名。还会见到关于参数名，如数据类型、精度、刻度、长度以及该参数是否是必需参数的信息。还会见到所有与 SAP 业务对象相关的基本和详细参数。

要对“数据仓库中心”定义 SAP 源：

1. 右键单击**仓库源**文件夹，单击**定义**然后单击**SAP**。可在“数据仓库中心”中定义 SAP 业务对象前，它必须存在于 R/3 系统中。
2. 在**名称**字段中，输入 SAP 源的名称。
3. 可选：在**管理员**字段中，输入负责源定义的人员的姓名。
4. 可选：在**描述**字段中，输入源的简短描述。
5. 可选：在**注意事项**字段中，输入关于源的信息，这些信息可能会对可访问源特性的用户有用。
6. 要“定义仓库源”窗口的“代理站点”页上，将一个或多个代理站点移动到**所选的代理站点**列表中。选择安装了 SAP 连接器的代理站点。
7. 在“定义仓库源”窗口的“SAP 源”页上，指定 SAP 源的连接类型是应用程序服务器不是一组服务器（与在 R/3 中定义的一样）。

对于应用程序服务器，提供应用程序服务器名和服务器的系统号（与对 R/3 定义的一样）。也可以选择提供网关主机和服务名以及特定目的的（本地工作站上的 `saprfc.ini` 文件中的一个条目）。

对于一组服务器，提供服务器的系统标识、消息服务器名、组名，并可任选提供它的目的地（本地工作站上的 `saprfc.ini` 文件中的一个条目）。对于应用程序服务器或一组服务器都应提供与 R/3 服务器连接的客户机号、用户标识和密码。

8. 在“定义仓库源”窗口的“业务对象”页上，展开“业务对象”文件夹并从在所选 R/3 系统上定义的业务对象列表中选择要定义为此仓库源的一部分的特定 SAP 业务对象。

“数据仓库中心”从 R/3 系统中检索业务对象列表时可能会花好几分钟，所用的时间取决于 R/3 系统上的网络流量和工作负荷。

定义每个源业务对象的特性

对“数据仓库中心”定义了 SAP 源后，可以为每个源业务对象定义特性：

1. 展开树中 SAP 源对象的列表，然后单击 SAP 源的**业务对象**文件夹。
2. 右键单击导航器的详细信息视图中的业务对象之一，然后单击**特性**。
3. 在“特性”笔记本的“源业务对象”页上。业务对象的描述是可视的，您可编辑它。
4. 在“特性”笔记本的“参数”页上，可以看到为 R/3 系统中的所选业务对象定义的参数，以及数据类型、长度、精度和刻度（如果适用的话）。还可以看到哪个参数是键字段，哪个参数是必需参数。
5. 如果需要业务对象的详细信息，必须映射业务对象的键字段以从“特性”笔记本的“列映射”页上的业务对象的 `GetList` 方法（商务应用程序编程接口的部件）中导出参数。单击右边的键字段，并拖动光标至一个导出参数，以映射这个键字段和这个导出参数。必须映射所有的键字段，但并不是所有的导出参数都需要映射。

定义 SAP 步骤

要定义此步骤的特性：

1. 在“过程模型”窗口中，单击 **SAP** 图标，然后单击画布以将些步骤添加到过程中。
2. 将该步骤链接到 SAP 业务对象。
3. 右键单击画布上的“SAP 步骤”然后单击**特性**。
4. 在**名称**字段中，输入 SAP 步骤的名称。
5. 可选：在**管理员**字段中，输入负责步骤定义的人员的姓名。

6. 可选: 在**描述**字段中, 输入步骤的简短描述。
7. 可选: 在**注意事项**字段中, 输入关于步骤的信息, 这些信息可能会对可访问源特性的用户有用。
8. 在“特性”笔记本的“输入参数”页上, 可以看到一个参数列表, 以及数据类型、长度和精度(如果适用的话)、缺省值以及对该参数是否是必需参数的指示。如果需要的话, 可以编辑值以指定从 SAP 系统中导入数据的过滤标准。
9. 在“输出参数”页上, 从业务对象的可用参数列表中选择要包括在此抽取步骤中的参数。
10. 在“列映射”页上, 可以生成缺省表以与所选的输出参数相匹配, 或将输出参数映射到现有的表。将输出参数视为要映射到仓库目标中的列的源列。有关详情, 参见第139页的『定义列映射信息』。
11. 在“处理选项”页上, 指定填充类型、代理站点、落实计数以及重试计数和间隔。落实计数指定 SAP 步骤将在把指定的记录数插入目标表中之前, 发出 SQL COMMIT 语句。缺省的落实计数为 0, 表示 SAP 步骤将把所有记录插入一个工作单元中, 然后再发出 COMMIT 语句。

从 WebSphere Site Analyzer 数据库中抽取数据

IBM WebSphere Site Analyzer (WSA) 是 IBM WebSphere 系列的 Web 服务器和应用程序服务器的部件。WSA 可帮助您分析流向 Web 站点以及从 Web 站点流出的流量。

“Web 连接器”使您可从 WebSphere Site Analyzer 数据库或数据集中抽取数据然后放到数据仓库中。“Web 连接器”提供了轮询步骤, 它检查 WSA 是否从它的数据导入(日志文件、表和点击流数据)中复制 Web 流量数据到数据集中。

此检查成功后, SQL 步骤就可以把 Web 流量数据从数据集市复制到仓库目标中。然后可使用 DB2 和 DB2 仓库管理器来进行数据分析、数据转换和数据挖掘。也可以将 WebSphere Commerce 数据与 Web 流量数据结合, 以进行更完整的 Web 站点分析。定义 WSA 源后, 可以从“数据仓库中心”的进程建模器中定义 Web 流量轮询步骤。

“Web 连接器”在与仓库代理相同的平台上运行: Windows NT、Windows 2000、AIX 或 Solaris 操作环境。

定义 WebSphere Site Analyzer 源

要对“数据仓库中心”定义新的 WebSphere Site Analyzer 源:

1. 右键单击**仓库源**文件夹，然后单击**定义**，再单击 **WebSphere Site Analyzer**。可在“数据仓库中心”中定义 WSA 源之前，WSA Webmart 数据库必须已存在于 WSA 系统上。一旦成功运行了 WSA 数据收集，WSA Webmart 数据库就会存在于 WSA 系统上。
2. 在“定义仓库源”窗口的“仓库源”页上，输入业务名，任选输入管理员姓名、描述和 WSA 源的注释。
3. 在“定义仓库源”窗口的“代理站点”页上，将一个或多个代理站点移动到所选的代理站点框中。
4. 在“定义仓库源”窗口的“数据库”页上，在与数据库相关的字段中填入 DB2 数据库，如果 WSA webmart 是 Oracle 数据库，则定制 ODBC 连接字符串。对于“数据库”页的“WebSphere Site Analyzer 特性”区，输入以下信息：

站点名称

在定义对其收集 Web 流量数据的“WSA 项目”时，在“项目信息”笔记本中指定的项目名。

主机名 安装了 IBM WebSphere Site Analyzer 的一个或多个虚拟主机的名称。通常是安装了 IBM WebSphere Site Analyzer 的工作站的名称。

端口号 WebSphere Application Server Webserver 所监听的端口号。缺省端口号是 80。如果保留此字段为空白，当 WSA 在非安全环境中运行时，将使用端口号 80。当 WSA 在安全环境中运行时，使用缺省端口号 443。如果某个环境是使用 SSL 进行加密的，即将它视为安全环境。

用户标识

用于登录到 IBM WebSphere Site Analyzer 的用户标识。用户标识的长度必须少于 128 个字符。取决于 IBM WebSphere Site Analyzer 的配置，此字段是可选的。

密码 用于登录到 IBM WebSphere Site Analyzer 的用户标识的密码。密码长度必须小于 64 个字符。取决于 IBM WebSphere Site Analyzer 的配置，此字段是可选的。

5. 在“定义仓库源”笔记本的“Web 源数据”页上，从可用的数据导入、表和视图列表中选择那些想要包括在此新仓库源中的特定项。

定义 Web 流量轮询步骤

要定义 Web 流量轮询步骤：

1. 在“过程模型”窗口中，单击 **Web 流量** 图标，然后单击画布以将这些步骤添加到过程中。
2. 右键单击画布上的 **Web 流量轮询** 步骤，然后单击**特性**。

3. 在“特性”笔记本的“Web 流量程序”页上，输入业务名，并任选输入管理员姓名、描述以及该 Web 流量步骤的注释。
4. 在“特性”笔记本的“参数”页上，单击仓库源名并选择一项可用的 WSA 源。将一个或多个要作为此步骤的一部分来轮询的可用数据导入移动到**选择的数据导入**列表。可用的数据导入是那些在定义 Web 流量源时指定的数据导入。
5. 在“特性”笔记本的“处理选项”页上，指定代理站点、轮询间隔和持续时间以及代理站点重试计数和间隔。轮询间隔和轮询持续时间指定“Web 流量轮询”步骤检查所选数据导入的数据收集状态的频率。轮询持续时间必须大于轮询时间间隔。轮询持续时间的缺省值是 60 分钟，而轮询时间间隔的缺省值是 20 分钟。在轮询持续时间和轮询时间间隔都设置为缺省值时，Web 流量步骤在返回最后的轮询状态前，最多轮询 WSA 服务器 3 次。如果选择的所有数据导入的数据收集状态都是成功，步骤可成功终止。
步骤保持已启用，并在调度该步骤运行时激活。

第7章 转换数据

本章描述“数据仓库中心”提供的转换数据的方法。它描述如何使用 SQL 步骤和仓库变换器来转换数据。

在许多情况下，有多种可以转换数据的方法。例如，若想要清理数据，有下列选择：

表 21. 清理数据的方法

方法	描述	有关详情，参见：
WHERE 子句	构建 SQL WHERE 子句以限制从源表中抽取的行。	第199页的『过滤数据』
公式和表达式	使用公式和表达式消除不需要的数据和创建需要的数据。使用“SQL 辅助”中的“表达式”窗口来指定公式、常量和记号。	第200页的『添加计算的列』
“清理数据”变换器	使用“清理数据”变换器在表上执行基于规则的查找和置换操作。	第202页的『清理数据』
仓库程序	使用仓库程序在源数据上执行先前的方法中不可用的任何功能。	第283页的『第13章 用用户定义程序扩展“数据仓库中心”』

连接源列

若选择多个源，还需指定这些列中的行将如何在表的连接中相关。可以使用“SQL 辅助”定义连接。

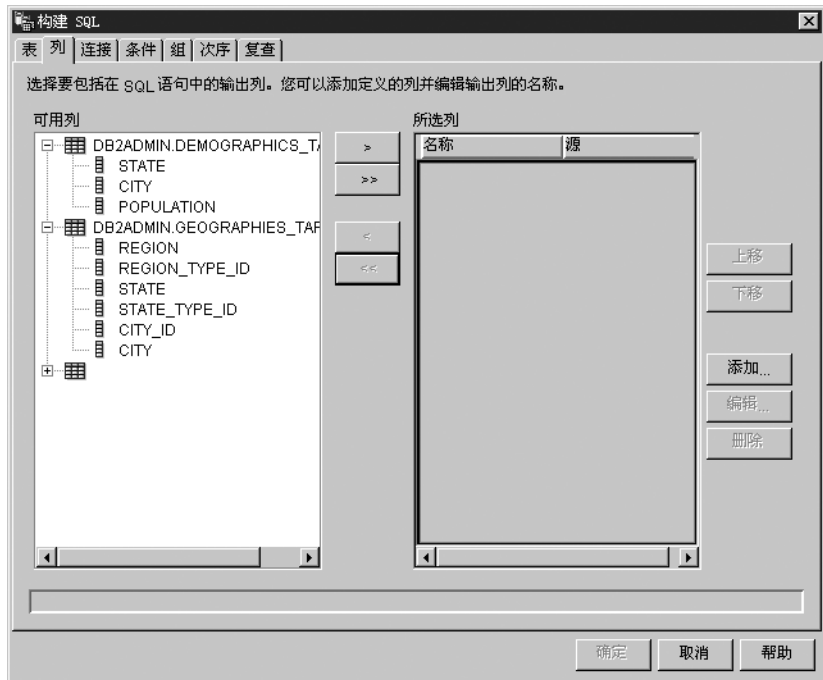
在定义连接之前，必须将源与步骤链接。

要连接源表：

1. 定义 SQL 步骤。（参见第155页的『选择和插入数据』。）
2. 打开步骤。
3. 单击 **SQL 语句** 选项卡。
4. 单击 **构建 SQL** 以自动生成 SQL。（否则，可以输入自己的 SQL。）

“SQL 辅助” 打开。

5. 单击**表**选项卡并验证已列示出源。
6. 单击**列**选项卡。



7. 从**可用的列**列表中，选择一列并单击 **>**。
该列从**可用的列**列表移动至**选择的列**列表中。
单击 **>>**，将**可用的列**列表中的所有列添加到**选择的列**列表中。
8. 单击**连接**选项卡。
9. 在其中一个表中选择一列。表按它们在“表”页上的**选择的表**列表中显示的次序显示。
10. 在另一个表中选择一列。
若这两列的数据类型兼容，一条连接这两列的灰线就显示出来，并且“连接”按钮变为可用。
若这两列的数据类型不兼容，在窗口底部的状态区会显示一条错误消息。
11. 单击**连接类型**按钮创建该连接。
“SQL 辅助” 在选择的列之间画出一条红线，它指示表已通过该列连接起来。



12. 要请求附加的连接，重复前述步骤。

要除去连接：

- a. 选择已连接的列，或通过单击 > 或 < 找到想要除去的连接。红线指示当前选择的连接。蓝线指示其他连接。
- b. 单击**断开连接**。除去该连接线。

13. 单击**复查**选项卡查看刚才构建的 SQL 语句。

14. 单击**确定**。

“SQL 辅助”关闭。

提示：要使用“SQL 语句”页上的**测试**按钮，源表必须存在。若指定“数据仓库中心”创建这些表，必须将把那些表作为目标表链接的步骤提升为测试方式，以创建这些表。

15. 单击**确定**。

“步骤特性”笔记本关闭。

转换代码

在许多生产环境下，源数据包括编码信息。例如，可以使用代码来引用整个数据库中的部件号。还有一个表，它使部件号与每个部件的序列号和描述相关。在这些情况下，您想要仓库中的信息包括每个部件的名称和描述。为此，必须将解码表与包含编码部件号的源数据组合在一起。

转换数据

首先，需要将解码表和编码部件号表定义为仓库源的一部分。然后，选择那些表作为步骤的源表。再单击“SQL 辅助”的“连接”页上的**连接**来连接这些表。

另一种方法是使用 CASE 语句将该数据解码。例如，月份列中的数据是用数字编码的，而您想将该数据转换为包含月份缩写的字符串。发出如下语句：

```
CASE TBC.ORDER_HISTORY.ORDERMONTH WHEN 1 THEN 'Jan'  
WHEN 2 THEN 'Feb' WHEN 3 THEN 'Mar' WHEN 4 THEN 'Apr'  
WHEN 5 THEN 'May' WHEN 6 THEN 'Jun' WHEN 7 THEN 'Jul'  
WHEN 8 THEN 'Aug' WHEN 9 THEN 'Sep' WHEN 10 THEN 'Oct'  
WHEN 11 THEN 'Nov' WHEN 12 THEN 'Dec'  
END
```

给连接添加空值

缺省情况下，连接假设为内连接。还可以通过在“SQL 辅助”的“连接”页上单击**连接类型**，来请求其他类型的连接。下列连接类型可用：

- 内连接
- 左外连接
- 右外连接
- 全外连接

若您的数据库支持 OUTER JOIN 关键字，可扩展内连接以从一个表中添加在另一个表中没有匹配行的那些行。

例如，您想连接两个表，以获得每个部门的经理的姓。第一个表是“部门”表，列示每个部门经理的员工号。第二个表是“员工”表，列示每个员工的员工号和姓。但是，某些部门没有经理；在这些情况下，部门经理的员工号是空值。要包括所有部门（无论它们是否有经理）和经理（若有的话）的姓，生成左外连接。左外连接包括第一个表中与第二个表匹配的行，或为空值的行。产生的 SQL 语句如下：

```
SELECT DEPTNO, DEPTNAME, EMPNO, LASTNAME  
FROM DEPARTMENT LEFT OUTER JOIN EMPLOYEE  
ON MGRNO = EMPNO
```

右外连接与左外连接相同，只是包括第二个表中与第一个表匹配的行或为空值的行。全外连接包括两个表中的匹配行和空行。

例如，有两个表，表 1 和表 2，包含下列数据：

表 1	
列 A	列 B
1	A

表 1	
2	B
3	C

表 2	
列 C	列 D
2	X
4	2

指定“列 A = 列 C”的连接条件。不同类型的连接的结果表如下：

内连接

列 A	列 B	列 C	列 D
2	B	2	X

左外连接

列 A	列 B	列 C	列 D
1	A	空	空
2	B	2	X
3	C	空	空

右外连接

列 A	列 B	列 C	列 D
2	B	2	X
空	空	4	2

全外连接

列 A	列 B	列 C	列 D
1	A	空	空
2	B	2	X
3	C	空	空
空	空	4	2

若指定值 (a,c), 将获得以下结果:

1
2
3
4

生成星型连接

可以生成一个星型连接，它是星型模式中定义的源表的连接。星型模式是一种专门的设计，它由下列类型的表组成:

- 维表，描述一个业务的各个方面
- 事实表，包含有关该业务的事实

例如，若您有一个销售书的邮购业务，一些维表是“客户”、“书籍”、“目录”和“财政年度”。事实表包含有关在该财政年度期间每个用户从每个目录中订购的书籍的信息。

每个维表包含一个主键，它是您选择来标识该表中的行的一列或多列。事实表包含外键，它们与维表中的主键相对应。外键是一个表中的一列，它的允许值必须是另一个表的主键。

当请求星型连接时，“数据仓库中心”将维表的主键与事实表的外键连接。在上例中，“客户”表具有主键“客户号”，每本书具有主键“书号”(ISBN)。每个表中的每个订单包含外键“客户号”和“书号”。星型连接将客户和书的信息与订单组合在一起。

有关在“数据仓库中心”中定义主键和外键的信息，参见第111页的『定义仓库目标』。有关定义星型模式和将它导出到 OLAP Integration Server 的信息，参见第311页的『第15章 在“数据仓库中心”中创建星型模式』。

过滤数据

在大多数情况下，当创建步骤时，只想要源数据的一个子集。可能只想抽取满足某个标准的那些行。可使用“数据仓库中心”构建 SQL WHERE 子句，以限制从源表中抽取的行。

例如，可以定义一个步骤，从源表的最新版本中选择行：

```
WHERE TBC.ORDER_HISTORY.RUN_ID = &cur_edtn.IWHDATA.TBC.ORDER_HISTORY
```

RUN_ID 列包含有关该步骤版本的信息。&cur_edtn 记号表示当前的步骤版本。因此，此 WHERE 子句选择其中的步骤版本等于当前版本的行。

要构建 WHERE 子句，使用“SQL 辅助”的“条件”页。

要排除重复行，选择排除重复行 (SELECT DISTINCT) 复选框。此操作消除结果集中每组重复行中的所有行，只留下一行。

要指定搜索条件：

1. 在“步骤特性”笔记本的“SQL 语句”页中，单击构建 SQL。
“SQL 辅助”打开。
2. 单击条件选项卡。
3. 从可用的列列表中，选择用于搜索的列。
4. 从运算符列表中，选择一个运算符。
5. 在值列表中输入一个或多个值。每行输入一个值。在值列表中，单击清除以除去所有值。

若在运算符列表中选择了“Between”运算符，必须在值列表中的每一行输入两个值。用和号 (&) 分隔这两个值。

可通过单击**查找**来搜索适当的值。

可在**值**列表中指定主机变量。为此，单击**值**列表中的一行，然后单击**添加变量**。**添加变量**窗口打开。输入变量名，然后单击**确定**。“添加变量”窗口关闭，该变量添加到“条件”页上的**值**列表中。

可在**值**列表中指定参数。若指定一个参数。则在搜索条件中使用它的值。以格式 `:parm` 指定参数，其中 `parm` 是参数名。例如，`:empid` 是名为 `empid` 的参数有效规范。

- 单击**添加**将该条件添加到**条件**字段中。
- 要为语句指定附加搜索条件，单击 **And** 或 **Or**。然后，如前述步骤所述构建第二个搜索条件。当单击**添加**时，后续条件追加到已显示在**条件**字段中的那些条件之后。

可编辑**条件**字段中的文本。可通过在**条件**字段中突出显示想要除去的条件部分，然后按键盘上的 **Delete** 键，来除去搜索条件。

要构建复杂条件，单击**高级表达式**。“表达式构建器 - 条件”窗口打开。有关使用**表达式构建器**的帮助，参见联机帮助中的『构建表达式』。

添加计算的列

也可定义根据其他列的值计算的一些列。例如，您只需要订购了商品的月份。可使用 `SQL DATE` 函数将订单日期转换为 `DATE` 数据类型格式。然后，使用 `MONTH` 函数返回该日期的月份部分。用于计算的列的 `SQL` 语句如下：

```
MONTH(DATE(TBC.ORDERS_MONTH.OrderDate))
```

还可以使用计算的列来汇总数据。在许多情况下，源数据包含的信息远比您想复制到仓库中的信息详细。您需要从源数据获得的仅仅是某种类型的汇总。您可能需要源数据库中各元素的平均值、汇总或计数，而不需要所有数据。

“数据仓库中心”允许您容易和准确地定义对源数据进行汇总的步骤。可使用标准 `SQL` 聚集函数 (`AVG`、`COUNT`、`MAX`、`MIN` 和 `SUM`) 和 `SQL GROUP BY` 子句来创建对源数据进行汇总的步骤。

汇总步骤减轻网络上的负载。它们在通过网络复制源数据之前对数据进行聚集。还可创建使用汇总技术汇总其他步骤的复合步骤。汇总可减小创建的目标仓库的大小。

要创建具有此类型汇总的步骤，单击“`SQL` 辅助”的“表达式构建器”窗口中**函数**字段中的 **SUM**函数。

例如，步骤将某月中出售的所有商品汇总，并以千美元为单位表示金额：

SUM(TBC.ITEMS_MONTH.Amount)/1000

要定义计算的列:

1. 在“步骤特性”笔记本的“SQL 语句”页中, 单击**构建 SQL**。
“SQL 辅助”打开。
2. 单击**列选项卡**。
3. 单击**添加**。
“表达式构建器”窗口打开。

可在**表达式**字段中输入您的表达式, 或可使用“表达式构建器”中的字段和控件来构建表达式。要构建表达式:

- a. 使用**列**、**运算符**和**条件**列表选择表达式的组成部分。双击特定的列、运算符或条件关键字将它添加到**表达式**字段中。您双击的每个项目将追加到**表达式**字段中的表达式中, 因此要确保按您想要它们显示的次序来选择这些项目。
- b. 将特定值添加到表达式。在**值**字段中输入一个值, 然后单击选择标记将该值添加到**表达式**字段。
- c. 要将函数添加到表达式:
 - 1) 从**函数**列表选择一个类别。**函数**字段下的列表变为显示指定类别的函数。
 - 2) 在**函数**字段下的列表中双击一个函数。
 - 3) “函数参数”窗口打开。选择该函数的格式并指定参数值。
 - 4) 单击**确定**。“函数参数”窗口关闭。该函数及其参数显示在“表达式构建器”的**表达式**字段中。
- d. 要将常量添加到表达式:
 - 1) 从**常量**列表选择一个类别。**常量**字段下的列表变为显示指定类别的常量。
 - 2) 在**常量**字段下的列表中双击一个常量。该常量添加到**表达式**字段中的表达式中。
- e. 使用下列按钮操作表达式:
 - 根据需要单击 **And**、**Or**、**=**、**<>**、**(和)**, 以将那些运算符添加到表达式中。
 - 单击**清除**除去**表达式**字段中的所有输入。
 - 单击**撤销**以除去在**表达式**字段中所做的最近一次更改。
 - 单击**重做**以返回到在**表达式**字段中所做最近一次更改。

转换数据

- f. 在完成表达式之后，单击**确定**。“表达式构建器”窗口关闭，且该列表表达式添加到“列”页上的**选择的列**列表中。
- g. 单击新建列的**名称**字段，并输入该列的名称。
- h. 按 Enter 键。
- i. 单击**上移**和**下移**，将该列移动至表中适当的位置。

转换目标表

使用仓库变换器执行下列基本数据转换：

- 清理数据
- 生成键列
- 生成周期表
- 反转数据
- 旋转数据

清理数据

使用“清理数据”变换器在表上执行基于规则的查找和置换操作。该变换器在步骤访问的源表的数据列中查找指定的值。然后该变换器用指定的置换值更新步骤写入的表中的相应列。可从输入表中选择多列转移到输出表。“清理数据”变换器不为转移列定义规则或参数。

在装入或导入之后，作为过程的一部分，可使用清理数据变换器清理数据值，并使其标准化。不要使用此变换器作为通用数据列编辑器。

可使用清理数据变换器执行下列任务：

- 置换所选数据列中丢失的、无效的或与适当的替代值不一致的值
- 除去不适合的数据行
- 剪切数值
- 执行数字离散化
- 除去文本中多余的空白
- 将列从源表复制到目标表

仅当源表和目标表在同一数据库中时，才可使用“清理数据”变换器。源表必须是单个仓库表。目标表是缺省目标表。

可以选择在查找字符串时忽略大小写和空白，并可对数字数据指定容差值。

仅当步骤处于开发方式时，才可对步骤进行更改。

指定的每个清理转换使用下列四种清理类型之一：

查找和置换

执行基本查找和置换功能。

离散化 在值范围之内执行查找和置换功能。

剪切 在值范围之内或之外执行查找和置换功能。

转移 在输入表中指定列以将其复制到输出表。

先决条件：在可使用“清理数据”变换器之前，必须为清理类型创建一个规则表。规则表指定“清理数据”变换器在查找并替换过程中将使用的值。规则表必须与源表和目标表位于同一数据库中。

为清理变换器创建规则表

规则表必须至少包含两列。一列包含查找值。另一列包含置换值。每一列中的各行彼此对应。

例如，规则表中的列 1 和 列 2 具有如下值：

列 1	列 2
办公桌	椅子
餐桌	灯

假设列 1 包含查找值，列 2 包含替换值。当运行该步骤时，“清理数据”变换器在源列中搜索值“办公桌”。无论它在何处找到值“办公桌”，它都会将值“椅子”写入目标列的相应字段中。

“清理数据”变换器将没有列示在查找列中的值直接复制到目标表中。在该示例中，值“凳子”没有列示在包含查找值的列中。若所选源列包含值“凳子”，“清除”变换器将把“凳子”写入目标列的相应字段中。

下表描述必须包含在每种清理类型的规则表中的列：

清理类型	规则表中的最少列数	列的用途
查找和置换	2 - 3	第一列包含查找值。 第二列包含替换值。 第三列（若指定的话）包含规则次序信息，但不是必需的。

离散化	3 - 4	<p>第一列包含查找值范围的下界。</p> <p>第二列包含查找值范围的上界。</p> <p>第三列包含置换值。</p> <p>第四列（若指定的话）包含规则次序信息，但不是必需的。</p>
剪切	3 - 5	<p>第一列包含查找值范围的下界。</p> <p>第二列包含查找值范围的上界。</p> <p>第三列包含置换值的下界。</p> <p>第四列包含查找值的上界。</p> <p>第五列（若指定的话）包含规则次序信息，但不是必需的。</p> <p>提示： 当使用“剪切”清理类型时，可在值范围之外运行查找并替换过程。</p>
转移	全部不允许	不使用规则表。

可使用“步骤”笔记本对输出列排序。可在“步骤”笔记本的“列映射”页上更改列名。

定义“清理数据”变换器

要定义“清理数据”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，定义想要在该步骤运行时想执行的清理转换。右键单击“变换器”定义表上的任何地方，并单击**添加**。表中会添加一行。
4. 在刚才添加的行的“输入列”标题下，单击并选择包含将被清理的数据的列名。
5. 在**清理类型**列表中，选择当步骤运行时将在输入列上执行的清理转换类型。

要不作任何更改将输入列复制到目标表中，单击**转移**。“规则”列表和“参数”列表变为不可用。继续步骤 7。

6. 在“规则”列表中，单击...指定步骤将对所选清理类型使用的规则表。一个窗口打开。该窗口上的字段因所做的选择而各异。使用此窗口定义所选清理类型的规则:

- 要定义“查找和置换”的规则:
 - a. 在**规则表**列表选择一个规则表。
 - b. 选择包含查找值的列。在**可用的列**列表中单击一列，然后单击**查找列**字段旁边的 **>**。
 - c. 选择包含置换值的列。在**可用的列**列表中单击一列，然后单击**置换列**字段旁边的 **>**。
 - d. 可选: 选择包含在规则表中读取行的顺序的列。在**可用的列**列表中单击一列，然后单击**按列排序**字段旁边的 **>**。
 - e. 单击**确定**保存更改并关闭窗口。

限制:

- 查找值的规范取决于仓库源中输入列的数据类型。查找列的数据类型必须与输入列的数据类型匹配。
 - 在数字查找列中仅允许出现数字。
 - 模式字符串仅允许用于字符列。模式字符串可包含代表任一单个字符的 **_** 或代表零个或多个字符的 **%**。若模式字符串中包含 **_** 或 **%**，则使用规则转义字符。例如，假设转义字符是 ****。要指定字符 **%**，在规则表的查找列中输入 **\%**。对于数字或字符数据，数据库空值可用作查找值或置换值。
- 要定义“离散化”的规则:
 - a. 在**规则表**列表选择一个规则表。
 - b. 在包含下界值的规则表中选择要搜索的列。在**可用的列**列表选择一列，然后单击**下界列**字段旁边的 **>**。
 - c. 在包含上界值的规则表中选择要搜索的列。在**可用的列**列表选择一列，然后单击**上界列**字段旁边的 **>**。
 - d. 在包含置换值的规则表中选择一列。在**可用的列**列表选择一列，然后单击**替换值列**字段旁边的 **>**。
 - e. 可选: 选择包含在规则表中读取行的顺序的列。在**可用的列**列表中单击一列，然后单击**按列排序**字段旁边的 **>**。选择的列应为“整数”类型。
 - f. 单击**确定**保存更改并关闭窗口。

注释: 可使用字符串或数字数据类型。按所用字符集的语言环境特定次序进行比较。

- 要定义“剪切”的规则:
 - a. 在**规则表**列表表中选择一个规则表。
 - b. 在包含下界值的规则表中选择要搜索的列。在**可用的列**列表中单击一列，然后单击**下界列**字段旁边的 **>**。
 - c. 从包含上界值的规则表中选择要搜索的列。在**可用的列**列表中单击一列，然后单击**上界列**字段旁边的 **>**。
 - d. 从包含下界置换值的规则表中选择一列。在**可用的列**列表中单击一列，然后单击**下置换值列**字段旁边的 **>**。
 - e. 从包含上界置换值的规则表中选择一列。在**可用的列**列表中单击一列，然后单击**上置换值列**字段旁边的 **>**。
 - f. 可选: 选择包含在规则表中读取行的顺序的列。在**可用的列**列表中单击一列，然后单击**按列排序**字段旁边的 **>**。选择的列应为“整数”类型。
 - g. 单击**确定**保存更改并关闭窗口。

注: 仅可对数字数据类型使用“剪切”清理类型。要避免下界或上界剪切，在规则表中指定数据库空值。数据库空值还可用作置换值。

7. 在“参数”列中，单击...。根据在**输入列**字段中指定的列的数据类型，打开“字符串参数”窗口或“数值参数”窗口。

使用“字符串参数”窗口中的控件，定制“清理数据”变换器如何执行查找和置换任务。要有关字符串参数的信息:

- a. 在“指定字符串设置”区中，选择适当的复选框以:
 - 压缩空白。除去前导和结尾空格。将两个以上的连续内部空格压缩为一个。
 - 在匹配期间忽略空白。从输入列、查找、上界和下界列中除去空白。
 - 忽略大小写。当“清理”变换器执行匹配处理时忽略大小写。
 - 若一个值与规则表中的列匹配，则不复制它。若一个值与规则表中的列匹配，则不复制该值。
- b. 在“转义”字符字段中，输入一个转义字符。例如，可以指定 \。不能使用字符 % 或 _。空字段不假定空白字符为转义字符。
- c. 单击**确定**保存更改并关闭窗口。

使用“数字参数”窗口中的控件为您的转换指定一个数字容差。缺省值是 0.0。数值容差是应用于数字查找值的数字。数值容差允许置换查找值或查找

值范围之外不远的值。例如，使用“离散化”清理类型以如下方式置换查找值数字范围：用 1000 置换 23.50 -24.50。若指定 0.10 的容差，输入数据 23.40 和 24.60 也将由 1000 置换。

要指定数字容差：

- a. 在“数字容差”字段中，输入数字容差。
 - b. 若想要“清理数据”变换器不复制与规则表中的列匹配的值，选择**发现规则匹配时不将记录复制到输出复选框**。
 - c. 单击**确定**保存更改并关闭窗口。
8. 定义所想数目的清理类型。完成时，单击**列映射**选项卡。若想要修改列映射或创建缺省目标表，参见第139页的『定义列映射信息』。
 9. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
 10. 单击**确定**保存更改并关闭步骤笔记本。

生成键列

使用“生成键表”变换器将唯一键添加到仓库表中。

可用两种方法更改键列中的值：

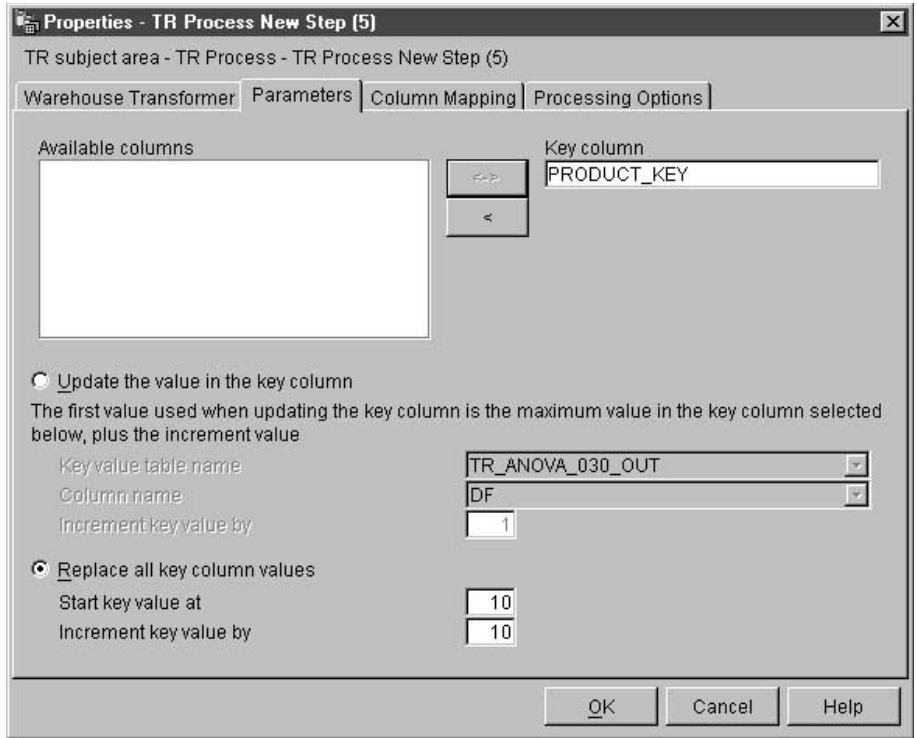
- 可基于仓库目标中的目标表或另一个表中数据类型为 INT 或 SMALLINT 的列更改值。
- 可置换目标表的键列中的所有键值。

“生成键表”变换器使用仓库目标表作为源。该变换器对仓库目标上的表进行写入。在定义此步骤之前，在“过程模型”窗口中用指向该步骤的箭头将仓库目标与该步骤链接。仅当步骤处于开发方式时，才可对步骤进行更改。

列映射不可用于此步骤。

要定义“生成键表”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的**可用的列**列表中，单击给它指定键值的列。然后，单击 **>**。选择的列不必是主键列。



- 4. 指定是要基于所选步骤中的列更新键列中的值，还是要置换现存键列中的所有键列值。

要基于所选步骤中的列更新键列中的值:

- a. 单击**更新键列中的值**。
- b. 在**键值表名**列表中，选择包含该目标表的仓库目标中的表。选择的表必须包含生成的键值所基于的列。
- c. 在**列名**列表中，在刚才指定的表中单击生成的键值所基于的列。
- d. 在**键值增量**字段中，输入一个数字增量值。

选择“更新键列中的值”时，变换器只更新那些没有键值的表中的行。向表中插入附加行时，再次运行变换器之前，键值为空。

要避免此问题，在第一次运行变换器之后，遵循4a中的步骤。

要置换现存键列中的所有键列值:

- a. 单击**置换所有键列值**。
- b. 在**起始键值**字段中输入一个起始键值。

- c. 在**键值增量**字段中输入一个数字增量值。
5. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
6. 单击**确定**保存更改并关闭步骤笔记本。

生成周期数据

使用“生成周期表”变换器来创建一个包含日期信息列的周期表，当评价其他数据，如确定某个时间周期内的销售量时，可使用这些日期信息。

可用下列方法使用周期表：

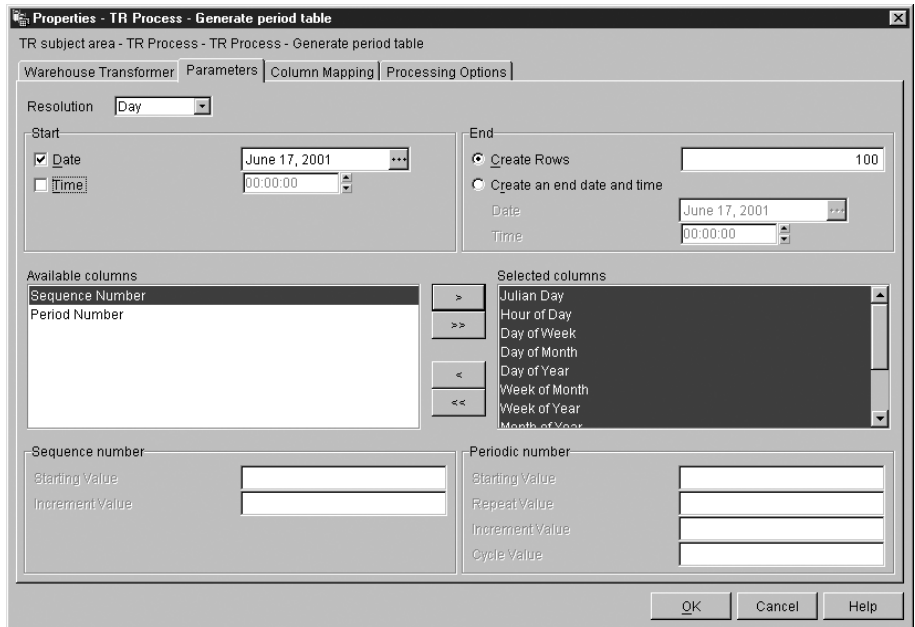
- 通过基于每行的生成日期/时间值创建一个或多个输出列来分析趋势。
- 将周期表与其他数据表连接。可将产生的表用于基于用户需求的更简单的 SQL `SELECT` 语句。
- 作为在变换器将其他列添加至生成的周期表之后创建更复杂的数据表的起点。

“生成周期表”变换器仅在目标表上工作。为成功使用该变换器，必须将变换器与目标连接。

仅当步骤处于开发方式时，才可对步骤定义进行更改。

要定义“生成周期表”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的**分辨率**列表中，单击周期表的日期、时间或时间戳记列的分辨率。



4. 通过选择日期或 / 和时间复选框指定起始周期值。然后，设置起始日期 / 时间：
 - 要创建具有“日期”类型列的周期表，选择日期复选框。然后，单击...。显示出一个日历。在日历上使用 << 和 >> 按钮在月份中滚动。找到起始月份之后，在日历中单击一个日期。
 - 要创建具有“时间”类型列的周期表，选择时间复选框。然后，在时间字段中单击小时、分或秒，并使用箭头按钮指定您的时间。
 - 要创建具有“时间戳记”类型列的周期表，选择这两个复选框，并指定一个起始日期和起始时间。
5. 使用行数或 / 和结束日期和时间指定结束周期值：
 - 要指定将生成的周期行的总数，单击**创建行**，在**创建行**字段中输入周期行的数目。
使用**创建行**字段作为替代项来设置周期表的结束日期或时间。在此字段中，输入变换器将在“周期表”中生成的周期行的数目。此字段是可选的。
若决定不使用此字段，则下列情况为真：
 - 若“生成周期表”变换器正对“日期”列进行写入，必须在日期字段中选择一个结束日期。结束日期必须晚于起始日期。
 - 若“生成周期表”变换器正对“时间”列进行写入，必须在时间字段中选择一个结束时间。结束时间必须晚于起始时间。

- 若“生成周期表”变换器正向“时间戳记”列进行写入，必须在日期字段中选择一个日期，并在时间字段中选择一个时间。结束日期和时间必须晚于起始日期和时间。
- 要使用日期或 / 和时间指定结束周期，单击**创建结束日期或时间**，并在日期和时间字段中选择日期或 / 和时间。
若没有在**创建行**字段中输入值，必须在“结束”段的日期和时间字段中输入一个或多个值。
在这些字段中，基于下列情况指定想要作为输入到周期表中的最后日期 / 时间值的日期或 / 和时间：
 - 若“生成周期表”变换器正对“日期”列进行写入，必须在日期字段中选择一个结束日期。结束日期必须晚于起始日期。
 - 若“生成周期表”变换器正对“时间”列进行写入，必须在时间字段中选择一个结束时间。结束时间必须晚于起始时间。
 - 若“生成周期表”变换器正向“时间戳记”列进行写入，必须在日期字段中选择一个日期，并在时间字段中选择一个时间。结束日期和时间必须晚于起始日期和时间。
- 可选：在**可用的**列列表中，单击将在该周期表中填充的一个或多个列。然后，单击 **>**。这些列移动至**选择的**列列表中。下表描述可用列：

列	数据类型	注释
公历日	Integer	日期始于公元前 4712 年 1 月 1 日
一天中的小时	Integer	1 - 24
星期中的天	Integer	1 - 7（随语言环境变化，使用语言环境缺省值）
月份中的天	Integer	1 - 31
年份中的天	Integer	1 - 366
月份中的星期	Integer	1 - 5（每月中的星期，从星期天开始）
年份中的星期	Integer	1 - 53
年份中的月	Integer	1 - 12
年份中的季度	Integer	1 - 4
年	Integer	Nnnn
日期名	Varchar	使用语言环境缺省值
月份名	Varchar	使用语言环境缺省值

转换数据

序列号	Integer	此列要求您在“序列号”段中的输入。若选择此选项，则“序列号”控件变为可用。
周期号	Integer	此列要求您在“周期号”段中的输入。若选择此选项，则“周期号”控件变为可用。

- 若将序列号放入选择的列列表中，则在起始值字段中输入数字起始序列号值。然后，在增量值字段中输入数字增量值。
 - 若将周期号放入选择的列列表中，则在起始值字段中输入数字起始周期值。在重复值字段中输入数字重复值。在增量值字段中输入数字增量值。在周期值字段中输入数字周期值。
6. 可选：在“列映射”页上，查看或更改列映射。如果步骤未与目标表链接，可创建一个缺省目标表。有关详情，参见第139页的『定义列映射信息』。
 7. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
 8. 单击**确定**保存更改并关闭步骤笔记本。

反转数据

使用反转数据变换器使表中行和列的次序反转。当使用反转数据变换器时，源表中的行变成输出表中的列，输入表中的列变成输出表中的行。各列中数据从上的次序保持不变，并将它们按从左到右的次序放在行中。

例如，将输入表看作一个矩阵。此变换器围绕着从表的左上角到右下角的对角线交换表中的数据。然后，变换器将转换后的数据写至目标表。

可指定一个附加列，它包含以数字 1 开始的序数数据。此列帮助您在变换器反转表之后标识各行。

还可在源表中指定将用作输出表中的列名的列。此列称作轴心列。

每个轴心组中的列数据必须为相同的数据类型，或者为通过自动提升而彼此相关的数据类型。关于数据类型的自动提升的详情，参见 *IBM DB2 Universal Database: SQL Reference*。

先决条件：在开始此任务之前，必须将仓库数据库的一个源表与步骤连接。还可指定步骤将对其进行写入的目标表，或可指定步骤创建目标表。所需的输出列必须在步骤生成的目标表中手工创建。

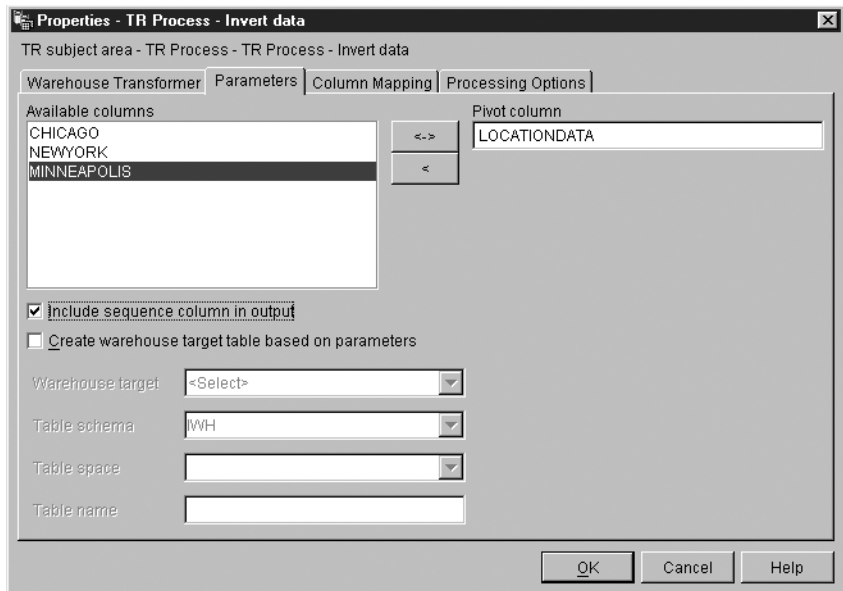
“反转数据”变换器在每次运行期间删除现存的数据库表并重新创建它。每次使用此变换器运行步骤时，会置换现存的数据，但保留表空间和表索引名。

必须将使用“反转数据”变换器的步骤提升为生产方式，才能查看产生的实际数据。

此步骤不使用“列映射”页。

要定义“反转数据”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。要获取更多信息，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的**可用的列**列表中，单击包含目标表列名称的列的名称。然后，单击 > 将所选列移动至**轴心列**字段中。此列中的数据必须为下列数据类型中的一种或多种：
 - CHAR
 - DATE
 - TIME
 - TIMESTAMP
 - SMALLINT
 - INTEGER
 - NUMERIC
 - DECIMAL



轴心列中所有数据必须少于 18 个字符。字母字符在要求大写对象名的 DB2 系统上更改为大写。

若不选择一列，变换器假设该表仅包含原始数据，应反转其中的所有数据。然后，变换器基于列号创建目标表的列名。首列将命名为 1，第二列将命名为 2，依次类推。

4. 可选：在输出复选框中选择**包括顺序列**，以在输出表中创建包含号码的有序列表的列。号码列表从 1 开始并按 1 递增。
5. 若您的步骤未与目标表链接，则选择**基于参数创建仓库目标表**。
6. 如果选择了**根据参数创建仓库目标表**，可从仓库目标列表中选择仓库目标，并提供关于要创建的目标表的信息。要获取用于创建仓库目标表的字段的信息，参见第140页的『创建缺省目标表』。
7. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
8. 单击**确定**保存更改并关闭步骤笔记本。

旋转数据

使用“旋转数据”变换器将源表中所选列（称作轴心列）中的相关数据分组到目标表中的一个列中（称为轴心组列）。可创建多个轴心组列。

可从源表中选择多个列转移至输出表。旋转数据变换器不更改这些列中的数据。

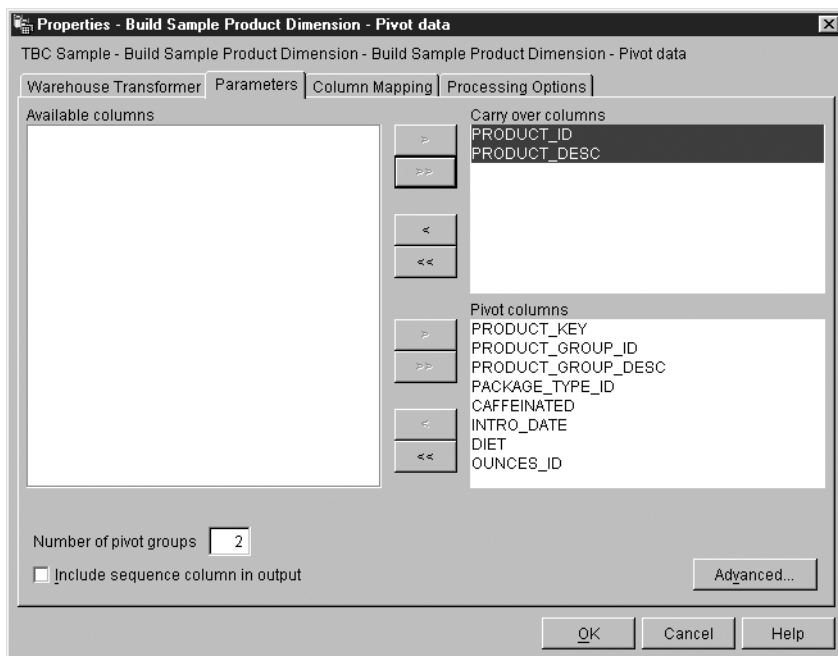
可指定一个附加列，它包含以数字 1 开始的序数数据。此列帮助您在变换器反转表之后标识各行。

每个轴心组中的列数据必须具有相同的数据类型，或者具有通过自动提升而彼此相关的数据类型。关于数据类型的自动提升的更多信息，请参阅 *IBM DB2 Universal Database: SQL Reference*。

在开始此任务之前，在“过程模型”窗口中将仓库源表和步骤连接。“旋转数据”变换器使用同一数据库中的现有目标表，或在包含仓库源的同一数据库中创建一个目标表。仅当步骤处于开发方式时，才可更改步骤定义。

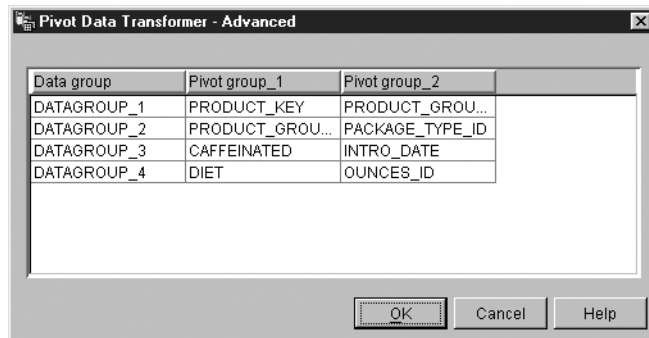
要定义“旋转数据”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。要获取更多信息，参见第139页的『提供关于步骤的一般信息』。
3. 单击**参数选项卡**。



4. 可选：在**可用的列**列表中，单击想要不作任何更改就转移至目标表的列。然后，单击**转移列**列表旁边的 **>**。要选择多个列，在单击列时按住 **Ctrl** 键。
5. 在**可用的列**列表中，单击想要作轴心列的列。然后，单击**轴心列**列表旁的 **>**。

6. 在**轴心组**数字段中，输入想要创建的轴心组的数量。输入的组数必须是选择的轴心列数的约数。例如，若选择 6 个轴心列，可输入数字 1、2、3 或 6。若选择 5 个轴心列，可在此字段中输入 1 或 5。
7. 若选择了轴心列，应对每个轴心列指定数据组和轴心组：
 - a. 在“旋转数据”变换器笔记本的“参数”页上，单击**高级**。“旋转数据变换器 - 高级”窗口打开。



使用此窗口中的表来定义您的轴心组。显示在此表中的行数取决于在“参数”页上选择的轴心列数和指定的轴心组数。例如，若选择了 6 个轴心列和 2 个轴心组，则显示 3 行。

- b. 单击**数据组**字段。然后，为数据组输入一个名称。数据组是目标表中的一列。此列用于描述或标记收集到轴心组列中的信息，以便将输入表中的列标号表示的数据原始含义保留在输出表中。若将此字段留为空白，并且只有一个轴心组，当步骤运行时，将把轴心列的名称旋转到输出列的数据字段中。
- c. 为每个轴心组选择一个轴心列。显示在此表中的轴心组数和您在“参数”页上指定的轴心组数相同。要选择轴心列，双击**轴心组**字段并选择一列。不能为多个轴心组选择一列。必须为轴心组中的每个字段选择一列。
- d. 单击**确定**。“旋转数据变换器 - 高级”窗口关闭。

若对“旋转数据变换器”窗口中选择的轴心列进行更改，则在此窗口中指定的数据组和轴心组值将复位为空白。

8. 若想在输出表中包括已排序的列号，可选择在**输出中包括顺序列**复选框。
9. 可选：在“列映射”页，查看或修改列映射，或创建缺省目标表。要获取更多信息，参见第139页的『定义列映射信息』。
10. 在“处理选项”页上，提供有关步骤如何处理的信息。要获取更多信息，参见第141页的『定义处理选项』。
11. 单击**确定**以保存更改并关闭步骤笔记本。

更改日期字段的格式

使用格式化日期变换器更改源表中日期字段的格式，步骤正将该字段复制到缺省目标表中。可将此变换器与任何其他变换器或 仓库程序一起运行。

格式化日期变换器提供了几种可对输入和输出列指定的标准日期格式。若输入列中的日期与指定的格式不匹配，则变换器将空值写入输出表。

若想要使用的格式未显示在**格式**列表中，可在变换器窗口的**格式字符串**字段中输入格式。例如，若输入列中的日期的结构类似于 Mar 2, 96 或 Jul 15, 83，则输入 MMM D, YY。

输出列字段的数据类型是 VARCHAR(255)。不能通过从“函数参数 - FormatDate”页上的**目录**列表中选择“日期”、“时间”或“日期/时间”来更改数据类型。

要使用格式化日期变换器：

1. 创建在安装变换器的仓库目标中生成缺省目标表的步骤。
2. 与源表链接。
3. 在步骤的“特性”笔记本的“SQL 语句”页上单击**构建 SQL**。
4. 单击**列**选项卡。
5. 单击**添加**。
“表达式构建器”打开。
6. 在“表达式构建器”中的函数名列表中，双击 **FormatDate**。
“函数自变量 — FormatDate”窗口打开。
7. 在**可用的列**列表中，单击包含想要重新格式化的日期或时间信息的列，然后单击 **>** 将所选字段放入**输入列**。
8. 在该窗口的**输入格式**区中指定输入格式：
 - a. 在**类别**列表中为输入列数据选择类别。
 - b. 在**格式**列表中选择日期、时间或时间戳记格式。**示例**列表显示您选择的格式的示例。**格式字符串**字段确认您的选择。还可通过在**格式字符串**字段输入一种格式来指定格式。
9. 在该窗口的**输出格式**区中指定输出格式：
 - a. 从**类别**列表中为输出列数据选择类别。
 - b. 在**格式**列表中选择日期、时间或时间戳记格式。**示例**列表显示您选择的格式的示例。**格式字符串**字段确认您的选择。还可通过在**格式字符串**字段输入一种格式来指定格式。
10. 单击**确定**。

转换数据

“函数自变量 — FormatDate” 窗口关闭，且该表达式显示在“表达式构建器”窗口中。

第8章 计算统计信息

使用统计变换器来执行下列统计功能:

- 方差分析 (ANOVA)
- 计算基本统计信息
- 计算小计
- X 平方测试
- X 平方拟合优度测试
- 相关分析
- 计算滑动平均值
- 回归

ANOVA 变换器

使用“方差分析”(ANOVA)变换器根据少量参数在两个表中进行统计计算。ANOVA 有三种类型: 一向 ANOVA、二向 ANOVA 和三向 ANOVA。

有关 ANOVA 类型的详情, 参见联机帮助。

ANOVA 变换器获取两个独立的方差估计。第一个估计基于各组之间的可变性。第二个估计基于各组内部的可变性。ANOVA 变换器计算这些估计之后, 就计算它们之间的比率。一个分布系列, 即 Fisher-F 分布, 描述此比率的显著性。

此变换器还计算 p 值。p 值表示两个组的平均值相等的概率。小的 p 值将得出平均值不相同的结论。例如, p 值为 0.02 表示样本平均值相等的机会为 2%。同理, 大的 p 值将得出两个组的平均值相同的结论。

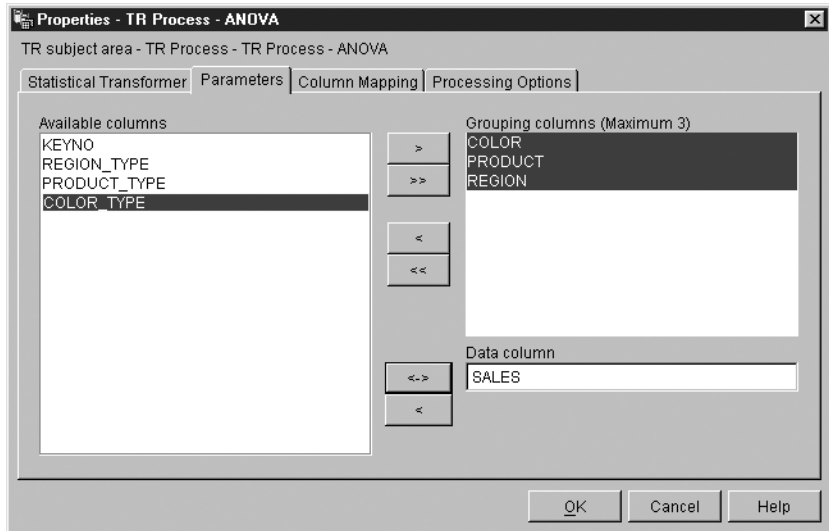
仅可对存在于同一个数据库中的表使用此步骤。使用仓库源或目标表作为 ANOVA 变换器的源, 使用最多两个仓库目标表作为 ANOVA 统计计算的目标。若不想为 ANOVA 转换选择目标表, 可以指定 ANOVA 变换器在目标数据库中创建表。直到在“过程模型”窗口中将该步骤与源链接, “参数”页才可用于此步骤子类型。

每次运行使用此变换器的步骤时, 都会置换现存的数据。ANOVA 变换器会在每次运行时删除现存的数据表并重建它。

仅当步骤处于开发方式时, 才可对步骤进行更改。

要定义 ANOVA 变换器:

1. 打开用于该 ANOVA 变换器的步骤笔记本。
2. 指定关于变换器的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上, 从可用的列字段中选择将作为 ANOVA 变换器的“分组”列的列或多列。根据要创建的 ANOVA 变换器的类型, 选择一列、两列或三列, 并单击**分组列**字段旁边的 **>**。



4. 在**可用的列**字段中, 单击一个列名, 然后单击**数据列**字段旁边的 **>**。此列不能用作**分组列**。它必须包含数字数据。
5. 选择用于变换器的目标表:
 - 若 ANOVA 变换器正使用两个目标表且已与它们链接, 则选择其中一个表来包含 ANOVA 统计信息。从 **ANOVA 统计信息表**列表中, 选择用于该 ANOVA 变换器的统计目标表。
6. 可选: 在“列映射”页上, 可以查看由“参数”页上定义的转换产生的输出列与目标表上的列之间的映射。不能更改这些映射。要获取使用“列映射”页可执行的任务的更多信息, 参见第139页的『定义列映射信息』。
7. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
8. 单击**确定**保存更改并关闭步骤笔记本。

“计算统计信息”变换器

使用“计算统计信息”变换器对来自单个表的任何数目的数据列计算下列描述性统计信息:

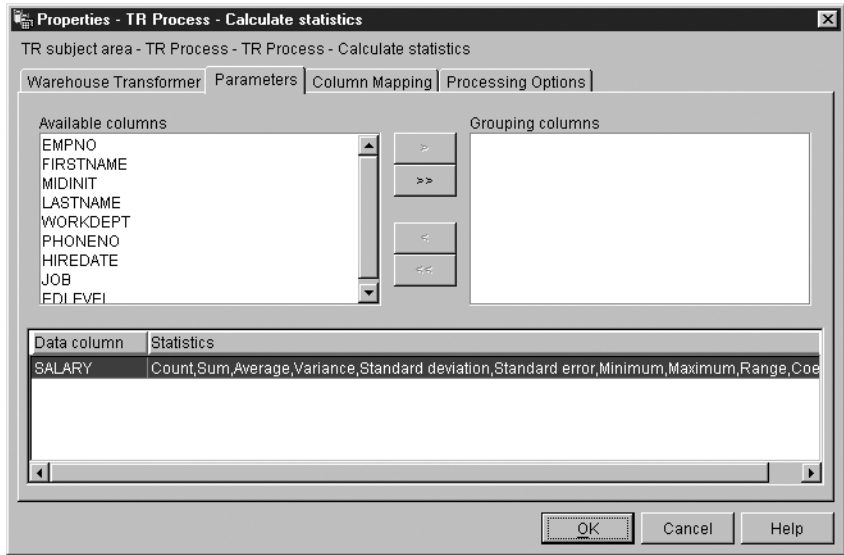
- 计数
- 总和
- 平均值
- 方差
- 标准偏差
- 标准误差
- 最小值
- 最大值
- 范围
- 变异系数

要使用“计算统计信息”变换器, 将该步骤与存在于同一数据库中的仓库源和仓库目标连接。或者, 可以将该步骤与仓库源连接, 并指定该步骤在同一数据库中创建目标表。

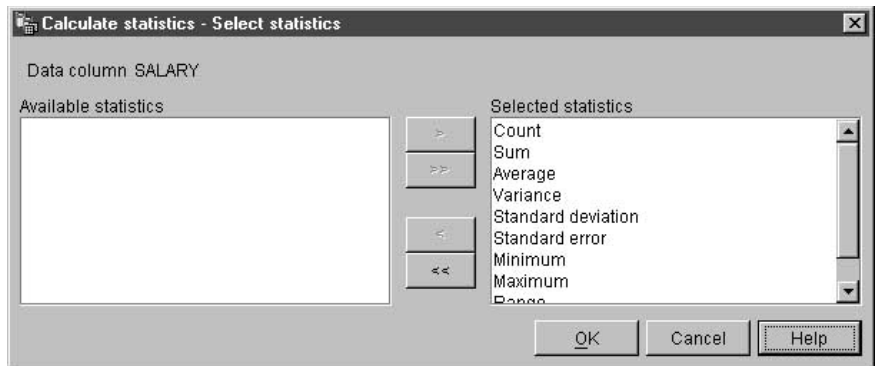
仅当步骤处于开发方式时, 才可更改步骤。

要定义“计算统计信息”变换器步骤:

1. 打开用于该“计算统计信息”变换器的步骤笔记本。
2. 指定关于变换器的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 单击**参数**选项卡。



4. 可选: 从可用的列列表中选择要用作分组列的任何列, 然后单击 **>**。分组列可以包含字符或数字数据。
5. 定义统计信息计算:
 - a. 右键单击“变换器”定义表上的任何地方, 并单击**添加**。表中会添加一行。
 - b. 在刚才创建的行的“数据列”标题下, 单击并选择要对其计算统计信息的列。
 - c. 双击**统计信息**字段。... 按钮变为可用。
 - d. 单击 ... 按钮。“计算统计信息 — 选择统计信息”窗口打开。



- e. 在**可用统计信息**列表中，单击要对所选列运行的一个或多个统计信息。然后，单击 **>**。可选择的统计消息类型取决于输入列的数据类型。对于字符数据类型，仅 **Count** 函数可用。
- f. 根据需要，重复步骤 5a 至 5e。
- g. 单击**确定**。“基本统计信息 — 选择统计信息”窗口打开。

定义了统计信息的计算操作之后，就不能更改它。必须删除不想要的定义并重新定义它们。“计算统计信息”变换器支持部分数据。例如，若选择一列以对它定义统计信息，但不为它选择统计信息，“计算统计信息”变换器会保存您的列选择。然而，不能对选择了部分数据的行映射列，也不能成功地运行选择了部分数据的步骤。

6. 在“列映射”页上，将统计计算产生的输出列映射到目标表中的列。用于统计计算的列名基于“参数”页上选择的数据列和为它选择的统计信息。为对数据列选择的每个统计创建一列。例如，若数据列“销售”定义了统计信息“总和”和“平均值”，则列 **Sales_sum** 和 **Sales_average** 将显示在“列映射”页上。要获取使用“列映射”页可执行的的任务的更多信息，参见第139页的『定义列映射信息』。

如果“参数”页未产生输出列，或此步骤未链接到目标表且您未使用“列映射”页创建目标表，就不能使用此页来映射您的列。某些步骤不允许更改列映射。

7. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
8. 单击**确定**保存更改并关闭步骤笔记本。

“计算小计”变换器

使用“计算小计”变换器来计算按某个时间周期（每周、每半月、每月、每季度或每年）分组的一组数值的运行小计。例如，为了便于记帐，常常需要生成基本时间周期内的数值的小计。在工资单的计算中这种情况最常见，它要求公司生成各种工资单数据的当月和当年小计。

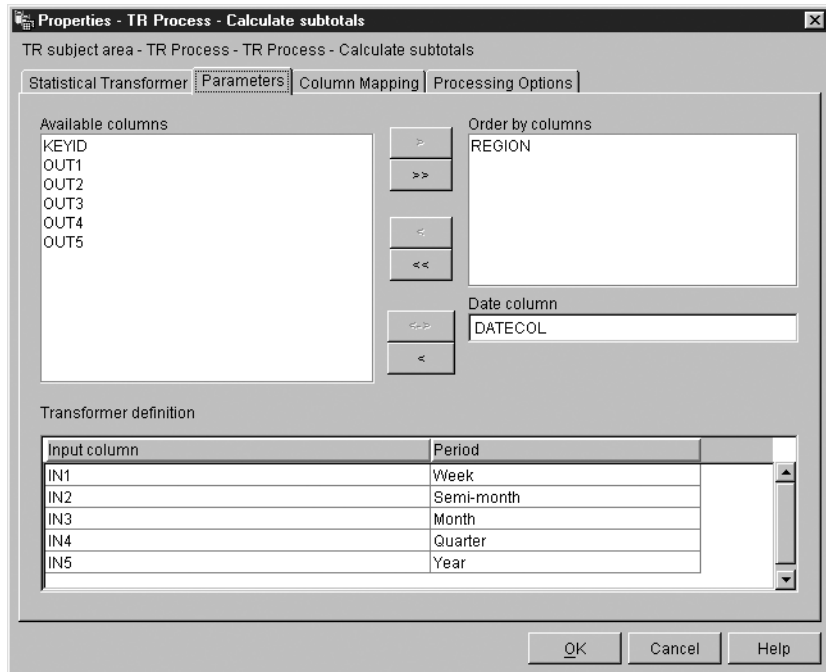
“计算小计”变换器使用仓库目标表作为源。用作源的表必须包含主键。如果使用由“数据仓库中心”生成的目标表，必须在使用该表作为源之前，对它指定一个主键。该变换器对同一个数据库上的表进行写入。在定义此步骤之前，在“过程模型”窗口中用指向该步骤的箭头将仓库目标与该步骤链接。直到在“过程模型”窗口中将此步骤子类型与源链接，“参数页”才可用于该步骤子类型。仅当步骤处于开发方式时，才可对步骤进行更改。

要定义“计算小计”变换器：

1. 打开“计算小计”变换器的步骤笔记本。

计算统计信息

- 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
- 单击**参数选项卡**。



- 可选：在**可用的列**列表中单击列以对输出数据排序。然后，单击**按列排序**列表旁边的 **>**。
- 在**可用的列**列表中，单击输入表中的日期列。列数据类型必须为 DATE 或 TIMESTAMP。然后，单击**日期列**字段旁边的 **>**。
- 定义小计计算：
 - 右键单击“变换器”定义表上的任何地方，并单击**添加**。表中会添加一行。
 - 在刚才添加的行的**输入列**标题下，用鼠标左键单击并选择要对其计算小计的列的名称。仅列示可用的数字列。不会列示其他行写入的或选择为分组列的数字列。若将列用作两行或多行的输入，则不能将该列用作输出列。
 - 在**周期列**标题下，用鼠标左键单击并选择要进行计算的周期。
 - 重复步骤 6a 至 6c，以定义其他小计计算。
- 使用“列映射”页将变换器产生的输出列映射到目标表中的列。如果未将变换器链接到输出目标表，可使用“列映射”页创建一个目标表。要获取使用“列映射”页可执行的任务的更多信息，参见第139页的『定义列映射信息』。

映射列的限制:

下列限制适用于步骤的单个实例。例如，具有步骤 1 和步骤 2。在步骤 1 中，可以使用列 B 作为源列。在步骤 2 中，可以使用列 B 作为目标列。

- 因为计算小计变换器可以向其源列进行写入，所以可以将输入列映射到它自己。例如，可以将源列 A 映射到它自己：

源列	目标列
A	A

仅当列未用作另一个变换器定义行中的输入列，才可将该列映射到它自己。例如，若下列情况为真，不能将列 A 映射到它自己：

源列	目标列
A_week	
A_month	

在此示例中，列 A 用作两个不同的变换器定义中的输入列。

- 不能将映射到目标列的源列用作同一个步骤定义中的另一个映射的目标列。例如，具有列 B、C 和 D。第一行指定如下：

源列	目标列
B	C

因为列 B 已映射为源列，所以不允许如下映射：

源列	目标列
D	B

- 将列映射为目标之后，不能将该列用作此步骤定义中任何其他映射的输入列或目标输出列。例如，具有下列行：

源列	目标列
A	A
B	C

因为将列 A 和列 C 指定为输出列，所以不能将他们用作任何其他行中的输入列或输出列。

如果“参数”页未产生输出列，或此步骤未链接到目标表且您未使用“列映射”页创建缺省目标表，就不能使用此页来映射您的列。某些步骤不允许更改列映射。

8. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
9. 单击**确定**保存更改并关闭步骤笔记本。

“X 平方”变换器

使用“X 平方”变换器对数字数据列执行 X 平方测试和 X 平方拟合优度测试。这些测试是非参数测试。

可以使用这些测试的统计结果来作出下列判断：

- 一个变量的值是否与另一个变量的值相关
- 一个变量的值是否独立于另一个变量的值
- 变量值的分布是否满足您的期望

对于小样本大小情况或当正在考虑的变量可能不是正常分布时，使用这些测试。X 平方测试和 X 平方拟合优度测试都充分利用不能精确测量的数据。

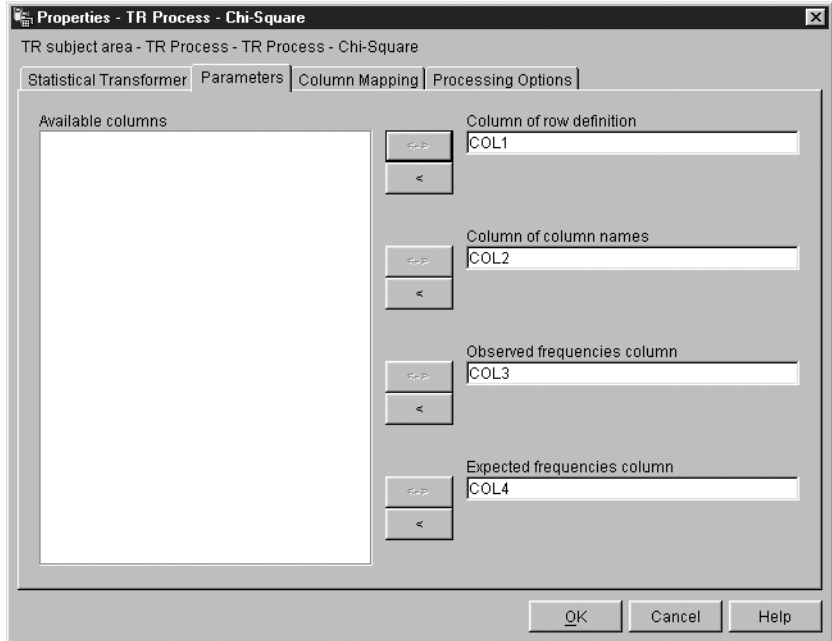
“X 平方”变换器任选地产生称作“期望值”输出表的附加输出表。可以选择一个表用作“期望值”输出表，或指定不产生该表。

当在“过程模型”窗口设置此过程时，将“X 平方”步骤与仓库目标表链接。若想要该步骤产生“期望值”输出表，将此步骤与同一个数据库中的另一个仓库目标表链接。

仅当步骤处于开发方式时，才可更改步骤定义。

要定义“X 平方”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页的**可用的列**列表中，单击一列。然后，单击**行定义列**字段旁边的 **>**。对于拟合优度计算和 X 平方计算，此字段是必需的。若想要步骤作为 X 平方计算运行，则转至步骤 4。否则转至步骤 5。该步骤将作为拟合优度计算运行。



4. 要定义 X 平方计算，单击可用的列列表中的一项，然后单击“列名列”字段旁边的 **>**。
5. 在可用的列列表中，单击包含观察频率数据的列。此列必须为数字类型。然后，单击观察频率列字段旁边的 **>**。
6. 在可用的列列表中，单击包含期望频率数据的列。此列必须为数字类型。然后，单击期望频率列字段旁边的 **>**。对于拟合优度计算，此字段是必需的；而对于 X 平方计算，是可选的。若指定一个期望频率列，则不能使用期望值输出表。
7. 在期望值输出表列表中，为该期望值输出表选择一个目标表。此字段可以是可选的，也可以是必需的，视情况而定：
 - 若“过程模型”窗口中只有一个目标表与 X 平方步骤链接，则此字段是可选的。若想要创建期望值输出表，则选择您的目标表。然后，单击确定保存并关闭步骤。接着，在“过程模型”窗口中，将另一个表与 X 平方步骤链接，以包含一般的 X 平方输出。最后，打开 X 平方步骤并继续定义该变换器的值。
 - 对于 X 平方计算，此字段是可选的。
 - 若“过程模型”窗口中有两个表与 X 平方步骤链接，则此字段是必需的。选择其中一个表作为期望值输出表。

8. 可选: 在“列映射”页上, 可以查看由“参数”页上定义的转换产生的输出列与目标表上的列之间的映射。不能更改这些映射。如果未将目标表与此步骤链接, 可以创建一个目标表。要获取使用“列映射”页可执行的任务的更多信息, 参见第139页的『定义列映射信息』。
9. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。

“相关”变换器

使用“相关”变换器确定一种属性值(例如, 雇佣年限)的更改与另一种属性(例如, 工资)的更改之间的相关程度。用于相关分析的数据包含两个输入列。每列包含您感兴趣的一种属性的值。“相关”变换器可以计算两个输入列之间的各种相关性量度。可以选择多种统计对一对给定的输入列进行计算。

输入列中的数据也可当作从一个较大群体中获取的样本来处理, 且“相关”变换器可用于测试这些属性在该群体中是否相关。在此上下文中, 空假定说明两种属性不相关, 替代假定说明属性是相关的。

对于一对或多对列, “相关”变换器可计算下列任何一个与相关性有关的统计信息:

相关系数 r

相关系数 r 是数据的两种属性或两列之间的线性关系的量度。该相关系数也称作 Pearson 积动差相关系数。值 r 的范围从 -1 到 $+1$, 且与计量单位无关。 r 的值接近 0 表示属性之间的相关性很小; 而该值接近 $+1$ 或 -1 则表示相关性很高。

当两种属性之间的相关系数为正值时, 一种属性的值增加可能会使第二种属性的值也增加。小于 0 的相关系数指示负相关。即当一种属性的值增加时, 另一种属性的值反而减少。

考虑 x 和 y 两个变量:

- 若 $r = 1$, 则 x 和 y 完全正相关。 x 和 y 的可能值都位于 (x,y) 平面上斜率为正的直线上。
- 若 $r = 0$, 则 x 和 y 不相关。它们没有明显的线性关系。但是, 这并不表示 x 和 y 在统计上是独立的。
- 若 $r = -1$, 则 x 和 y 完全负相关。 x 和 y 的可能值都位于 (x,y) 平面上斜率为负的直线上。

协方差 协方差是用来表示数据的两种属性或两列之间的线性关系的量度。协方差

的值的范围可以从负无穷大到正无穷大。但是，若协方差的值太小或太大，以至于不能用一个数来表示，则将用 NULL 来表示该值。

与相关系数不同的是，协方差与计量单位有关。例如，以英寸为单位而不是以英尺为单位来计量两种属性的值，会使协方差增大 144 倍。

T 值 T 值是 T 统计的观察值，T 统计用来测试两种属性相关的假定。T 值的范围可以从负无穷大到正无穷大。T 值接近 0，证明是空假定，即属性之间没有相关性。T 值远离 0（可以是正值，也可以是负值），证明是替代假定，即属性之间有相关性。

T 统计的定义是：

$$T = r * \text{SQRT}((n-2) / (1 - r*r))$$

其中， r 是相关系数， n 是输入值对的数目，而 SQRT 是平方根函数。

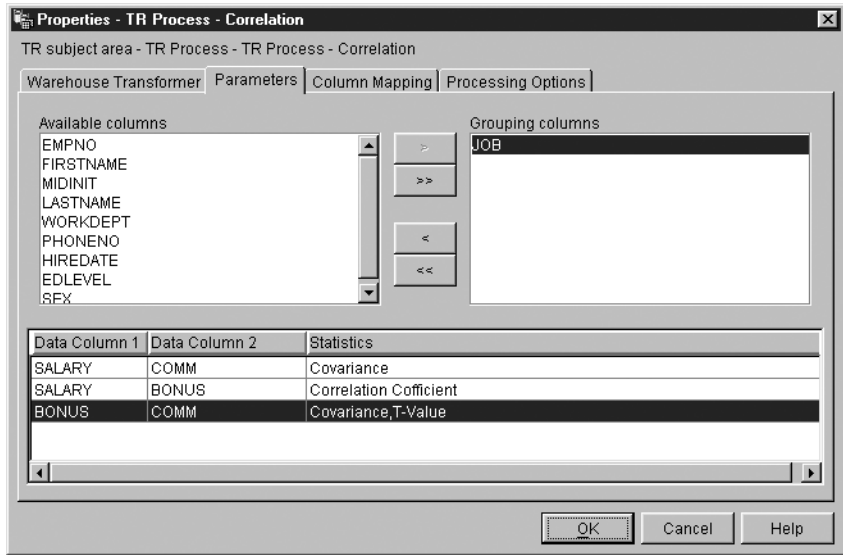
若相关系数 r 是 -1 或 +1，则 T 值用 NULL 表示。若 T 值太小或太大，以至于不能用一个数来表示，则用 NULL 来表示该值。

P 值 P 值是当空假定为真时，“T 统计”的绝对值将等于或大于观察值（T 值）的概率。较小的 P 值证明空假定为假，即属性实际上是相关的。

源表和目标表必须存在于仓库数据库中。根据您的需要，此变换器可在包含源的同时同一仓库数据库中创建目标表。仅当步骤处于开发方式时，才可更改步骤。

要定义“相关”变换器步骤：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 单击**参数选项卡**。



4. 可选: 单击要用作分组列的列, 然后单击 >. 分组列可以包含字符或数字数据。
5. 定义相关统计信息:
 - a. 右键单击“变换器”定义表上的任何地方, 并单击**添加**。表中会添加一行。
 - b. 在刚才添加的行的**数据列 1**标题下, 单击并选择用于计算的第一列。仅列示数字类型的列。
 - c. 在**数据列 2**标题下, 单击并选择用于计算的第二列。仅列示数字类型的列。不能在“数据列 2”中使用在“数据列 1”中使用的同一列。
 - d. 在**统计信息**标题下, 双击刚才创建的行。... 按钮变为可用。
 - e. 单击 ... 按钮。“相关 — 选择统计信息”窗口打开。
 - f. 在**可用统计信息**列表中, 单击一个或多个统计信息。然后, 单击 >. 统计信息移动至**选择的统计信息**列表。
 - g. 根据需要, 重复步骤 5a 至 5f。
 - h. 单击**确定**。“相关 — 选择统计信息”窗口关闭。
“相关”变换器支持部分数据。例如, 若选择要定义统计信息的列, 但不为它选择统计信息, 则“相关”变换器会保存您的列选择。然而, 不能对选择了部分数据的行映射列, 也不能成功地运行选择了部分数据的步骤。
6. 在“列映射”页上, 将从相关统计信息产生的列映射到目标表中的列。如果未将变换器与输出目标表链接, 可创建一个目标表。
相关统计信息的列名基于在“参数”页上选择的数据列项和为它选择的统计信息。为选择的每个统计信息及其相应的数据列, 创建一列。例如, 如果对数据

列 Salary 和 Employment 定义了相关统计信息 Covariance 和 T-value, 则列 Covariance_Salary_Employment 和 T-value_Salary_Employment 将显示在“列映射”页上。要获取使用“列映射”页可执行的任务的更多信息, 参见第139页的『定义列映射信息』。

若目标列映射到一个源列, 则映射保持不变。如果“参数”页未产生输出列, 或此步骤未链接到目标表且您未使用“列映射”页创建缺省目标表, 就不能使用此页来映射您的列。某些步骤不允许更改列映射。

7. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
8. 单击**确定**保存更改并关闭步骤笔记本。

“滑动平均值”变换器

简单滑动平均值和指数平滑滑动平均值通常都可以预测与时间有关的一系列值的未来过程。在商业和金融预测中, 滑动平均值被广泛应用于进行时间序列的分析。在金融领域也广泛使用滚动和。

可使用“滑动平均值”变换器来计算下列值:

- 简单滑动平均值
- 指数滑动平均值
- N 个周期的数据的滚动和, 其中 N 由用户指定

指数滑动平均值也称作指数平滑滑动平均值。

滑动平均值将在短时间内发生的事件重新分布在较长的时间段中。这种重新分布用于除去时间序列数据中的噪声、随机事件、大的波峰和波谷。您可对时间序列数据集应用滑动平均值方法来:

- 除去季节性变化的影响。
- 抽取数据趋势。
- 增强长期循环。
- 在执行更高级别的分析之前平滑数据集。

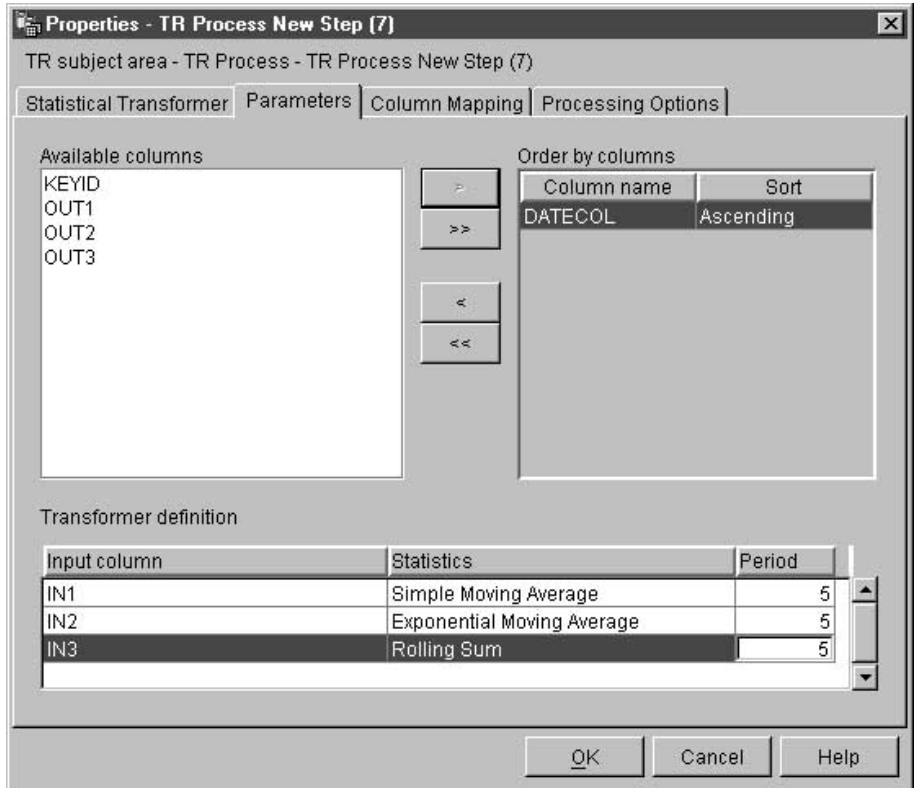
“滑动平均值”变换器使用仓库目标表作为源。用作源的表必须包含主键。如果使用由“数据仓库中心”生成的目标表, 必须在使用该表作为源之前, 对它指定一个主键。该变换器对仓库目标上的表进行写入。在定义此步骤之前, 在“过程模型”窗口中, 使用指向该步骤的箭头将仓库目标与该步骤链接。

要定义“滑动平均值”变换器:

1. 打开步骤笔记本。

计算统计信息

2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页的**可用的列**字段中，单击要用作分组列的列，然后单击 **>**。这些列移动至**按列排序**字段。分组列可以包含字符或数字数据。



4. 对**按列排序**字段中列示的每个列，单击**排序**列表中的排序次序。若为递升次序，则选择**升序**。对降序次序选择**降序**。
5. 定义滑动平均值计算：
 - a. 右键单击“变换器”定义表上的任何地方，并单击**添加**。表中会添加一行。
 - b. 在刚才添加的行的“输入列”标题下，用鼠标左键单击并选择要对其计算滑动平均值的列的名称。仅列示数字列。不会列示其他行写入的或选择为分组列的数字列。若将一列用作两行或多行的输入，则不能将该列用作输出列。
 - c. 在**统计信息列**标题下，用鼠标左键单击并选择要进行计算的滑动平均值的类型。

在“滑动平均值”变换器中，可以选择下列三种统计信息中的一种：

- 简单滑动平均值。简单滑动平均值 (SMA) 定义为：

$$SMA [行 i] = SUM (最后 N 行的数据值) / N$$

- 指数滑动平均值。指数滑动平均值 (EMA) 定义为：

$$EMA [行 i] = (Value [行 i] * K) + (EMA [行 i-1] * (1-K))$$

其中：K = 2/(N+1)

例如：对于四日 EMA，N = 4，K = 2/5 及 1-K = 3/5。指数滑动平均值也称作指数平滑滑动平均值。

- N 个周期的数据的滚动和，其中 N 由用户指定。滚动和 (RS) 定义为：

$$RS [行 i] = SUM (最后 N 行的数据值)$$

- d. 在**周期列**标题下，双击该字段，然后输入计算该滑动平均值所用的天数。例如，若您要计算一周的滑动平均值，输入 7。仅可输入数字。会忽略其他击键。
- e. 根据需要，重复步骤 5a 至 5d。

6. 在“列映射”页上，将变换器产生的输出列映射到目标表中的列。如果未将变换器与目标表链接，可以创建一个目标表。有关详情，参见第139页的『定义列映射信息』。

转换产生的列列示在该页左边。在“参数”页的“输入列”、“统计信息列”和“周期列”中完成选择之后，就会命名这些列。例如，若在“参数”页中选择 SALARY 作为输入列，选择 EMA 作为计算并选择 7 作为周期，将看到 SALARY_EMA_7 显示在“列映射”页的左边。

限制：

这些限制适用于步骤的单个实例。例如，具有步骤 1 和步骤 2。在步骤 1 中，可以使用列 B 作为源列。在步骤 2 中，可以使用列 B 作为目标列。

- 因为滑动平均值变换器可以向其源列写入，所以可以将输入列映射到它自己。例如，可以将源列 A 映射到它自己：

源列	目标列
A	A

仅当列未用作另一个变换器定义行中的输入列，才可将该列映射到它自己。
例如，若下列情况为真，不能将列 A 映射到它自己：

源列	目标列
A_EMA_7	
A_EMA_31	

在此示例中，列 A 用作两个不同的变换器定义中的输入列。

- 不能将映射到目标列的源列用作同一个步骤定义中的另一个映射的目标列。
例如，具有列 B、C 和 D。第一行指定如下：

源列	目标列
B	C

因为列 B 已映射为源列，所以不允许如下映射：

源列	目标列
D	B

- 将列映射为目标之后，不能将该列用作此步骤定义中任何其他映射的输入列或目标输出列。例如，具有下列行：

源列	目标列
A	A
B	C

因为将列 A 和列 C 指定为输出列，所以不能将他们用作任何其他行中的输入列或输出列。

如果“参数”页未产生输出列，或此步骤未链接到目标表且您未使用“列映射”页创建缺省目标表，就不能使用此页来映射您的列。某些步骤不允许更改列映射。

7. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
8. 单击**确定**保存更改并关闭步骤笔记本。

“回归”变换器

使用“回归”变换器来标识一个从属变量与一个或多个独立变量之间的关系，并显示它们的相关有多紧密。可以使用此变换器来显示更改定价对产品需求的影响，位置对广告反应的影响，或显示两个表面上随机的数据集的相关紧密程度。

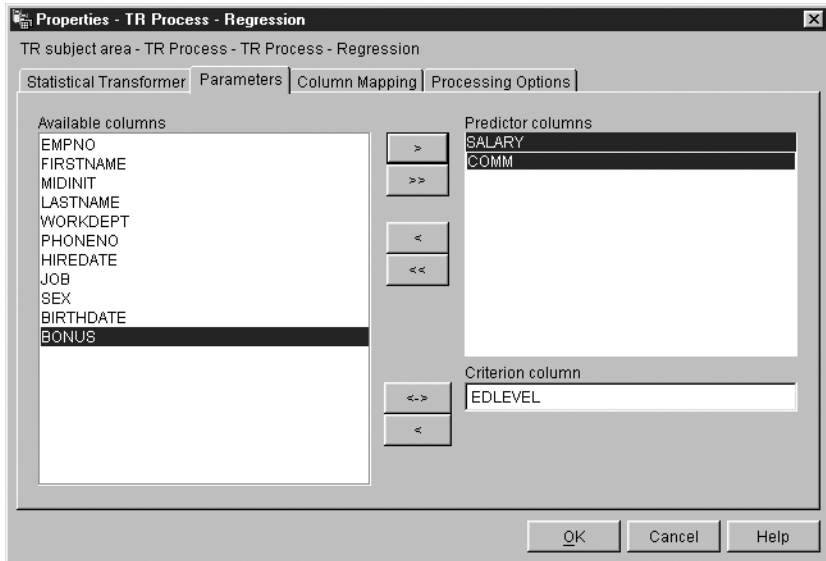
此变换器执行反向的全模型回归。此方法从模型中的所有独立变量开始，每次除去一个最次要的独立变量，直到该模型中仅保留重要的独立变量。

“回归”变换器产生两个附加的输出表：ANOVA 总结表和方程式变量表。

在开始此任务之前，必须在“过程模型”窗口中将此步骤与仓库源表和三个仓库目标表链接。或者，可以将该步骤与一个源链接并指定该步骤创建目标表。这些表必须存在于同一数据库中。“回归”变换器将“回归”转换产生的结果写入一个仓库目标的表中，并分别在第二个和第三个目标上创建 ANOVA 总结表和“方程式变量”表。仅当步骤处于开发方式时，才可对步骤进行更改。

要定义“回归”变换器：

1. 打开步骤笔记本。
2. 指定关于变换器的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，从包含用于回归计算中预测的独立变量数据的可用的列列表中
选择列。然后，单击预测程序列表旁边的 >。这些列会添加至预测程序列表。
仅显示数据类型为数字的列。



4. 在可用的列列表中，单击一个列，该列包含用于回归计算的标准的从属变量数据。然后，单击标准列字段旁边的 >。仅显示数据类型为数字的列。
5. 在总结表列表中，选择要作为 ANOVA 总结表的目标表。
6. 在方程式变量表列表中，选择要作为“方程式变量”表的目标表。

计算统计信息

7. 可选: 在“列映射”页上, 可以查看由“参数”页上定义的转换产生的输出列与目标表上的列之间的映射。不能更改这些映射。如果尚未将变换器与目标表链接, 可以创建一个目标表。要获取更多信息, 参见第139页的『定义列映射信息』
8. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
9. 单击**确定**保存更改并关闭步骤笔记本。

第9章 清理名称和地址数据

使用“数据仓库中心”和 Trillium Software System 来清理名称和地址数据。Trillium Software System 是一种名称和地址清理产品，它可重新格式化、标准化并验证名称和地址数据。通过从用户定义的程序中启动 Trillium Batch System 程序，可在“数据仓库中心”中使用 Trillium Software System。从 Trillium Batch System 脚本或 JCL 导入元数据时，会将用户定义的程序添加到“仓库”树。

“数据仓库中心”已提供了与 Vality and Evolutionary Technologies 公司的工具的集成。

要求

- 必须在仓库代理站点或远程主机上安装 Trillium Software System。
- 在 UNIX 和 Windows 系统上，必须将 Trillium Software System 的 bin 目录的路径添加到 PATH 环境变量中，以启用代理来运行 Trillium Batch System 程序。在 UNIX 上，必须在启动 vwdaemon 进程前添加 PATH 变量到 IWH.environment 文件。
- 用户必须有使用 Trillium 软件的经验。

表22显示了在“数据仓库中心”中进行名称和地址清理所需的软件。

表 22. 进行名称和地址清理所需的软件

操作系统	所需的软件
UNIX	Trillium Software System 版本 4.0 “DB2 通用数据库仓库管理器版本 7.2” 仓库代理
Windows NT 和 Windows 2000	Trillium Software System 版本 4.0 “DB2 通用数据库仓库管理器版本 7.2” 仓库代理 对于远程访问，主机必须具有 ftpd 和 rexecd 守护进程。

清理名称和地址数据

表 22. 进行名称和地址清理所需的软件 (续)

OS/390	安装在远程 OS/390 主机上的 Trillium Software System 版本 4.0
	安装在 UNIX 和 Windows NT 上的“DB2 通用数据库仓库管理器版本 7.2”仓库代理
	必须安装了 TCP/IP 3.2 或更高版本
	仅将 OS/390 操作系统作为远程主机支持

Trillium Software System 组件

Trillium Software System 由四个组件组成: 转换器、分析器、地理编码器和匹配器。可将这些组件作为一组函数来执行名称和地址清理操作。可以从用户定义的程序 Trillium Batch System 中运行组件。

转换器 使用转换器使数据标准化, 并将源数据转换为指定的输出格式。

分析器 使用分析器来解释名称和地址源数据并创建关于源数据的元数据。

地理编码器

使用地理编码器来比较源数据和邮政服务数据, 以提供任何丢失信息, 例如快件服务商或 ZIP+4 代码。地理编码器还可执行与“美国人口普查”数据的匹配操作。

匹配器 使用匹配器来比较类似的名称和地址, 从而识别出双重记录。通过使用匹配器将一个记录与一组记录比较, 可执行引用匹配。

将 Trillium Batch System 和“数据仓库中心”一起使用

在“数据仓库中心”中, 可导入 Trillium Batch System 元数据并创建用户定义的程序步骤。此步骤调用本地仓库代理站点或远程仓库代理站点上的 Trillium Batch System 脚本。在“数据仓库中心”中, Trillium Batch System 脚本是一个有源文件和目标文件的步骤。源文件是用于第一个 Trillium Batch System 命令的输入数据文件。目标文件是由最后一个 Trillium 命令在脚本中创建的输出数据文件。然后该步骤可复制到将与其他步骤一起使用的另一过程。

第239页的图14通过第239页的图17显示了 Trillium Batch System 输入和输出数据文件之间的关系, 以及“数据仓库中心”中的源和目标文件之间的关系。

```

REM 运行转换器
pfcondrv -parmfile c:\tril40\us_proj\parms\pfcondrv.par
REM 运行分析器
pfprsdrv -parmfile c:\tril40\us_proj\parms\pfprsdrv.par
REM 运行匹配器
cfmatdrv -parmfile c:\tril40\us_proj\parms\pfmatdrv.par

```

图 14. 样本 *Trillium* 脚本文件

```

INP_FNAME01 c:\tril40\us_proj\data\convinp
INP_DDL01 c:\tril40\us_proj\dict\input.ddl

```

图 15. *pfcondrv.par* 文件的内容

```

OUT_DDNAME c:\tril40\us_proj\data\maout
DDL_OUT_FNAME c:\tril40\us_proj\dict\parseout.ddl

```

图 16. *pfmatdrv.par* 文件的内容

```

c:\Tril40\us_proj\data\convinp (source file) ->
Trillium Batch System 步骤 -> c:\tril40\us_proj\data\maout (目标文件)

```

图 17. *Trillium Batch System* 步骤定义

导入 **Trillium** 元数据

要将 **Trillium** 元数据导入“数据仓库中心”：

1. 创建 **Trillium Batch System** 脚本或 **JCL**。您可以使用任何脚本或 **JCL** 编写工具来创建脚本或 **JCL** 文件。
2. 右键单击**仓库**，然后单击**导入元数据** → **Trillium** 以打开 **Trillium Batch System** 窗口。
3. 在**脚本**或 **JCL** 字段中，输入要运行的 **Trillium Batch System** 脚本或 **JCL** 文件的名称。
4. 在**输入文件**字段中，输入在指定的脚本或 **JCL** 文件中第一个运行的 **Trillium Batch System** 程序的输入数据文件名。
5. 在**输入 DDL** 字段中，输入描述输入数据文件的输入 **DDL** 文件的名称。此文件必须在仓库代理站点上可用。

6. 在**输出文件**字段中，输入脚本或 JCL 中的最后一个 Trillium Batch System 程序的输出数据文件的名称。
7. 在**输出 DDL** 字段中，输入描述输出数据文件的 DDL 文件的名称。此文件必须在仓库代理站点上可用。
8. 可选：在**输出错误文件**字段中，输入要使用的输出错误文件的名称。此错误文件从 Trillium Batch System 程序中捕获运行时错误。这些错误都记录在 stderr 日志中。对于本地主机，如果不在此处指定名称的话，将创建一个缺省输出错误文件。要获取关于输出错误文件的更多信息，参见第244页的『错误处理』。
9. 单击**连接**选项卡。
10. 如果正在导入的 Trillium 元数据在仓库代理站点上，可单击**本地主机**。
如果正在导入的 Trillium 元数据不在仓库代理站点上：
 - a. 单击**远程主机**，然后输入包含要导入的元数据的远程系统的 TCP/IP 主机名。如果选择**远程主机**，因为远程目标文件不受支持，所以会将目标文件创建为本地文件。可以添加获取远程文件的 FTP 步骤到指定的本地目标文件中。
 - b. 在**远程操作系统**列表中，单击正在访问的远程主机的操作系统。
 - c. 在**远程用户标识**字段中，输入正在访问的远程主机的用户标识。
 - d. 在**密码选项**列表中，选择要对正在访问的远程主机使用的密码选项：

不需要密码

指定访问远程主机上的元数据不需要密码。

检索密码

指定将从用户定义的程序检索密码。

在**密码程序**字段中，输入将检索密码的密码程序的名称。程序必须驻留在仓库代理站点上，并将密码写到第一行中的输出文件中。

在**程序参数**字段中，输入密码程序的参数。第一个参数必须是将密码写至的输出文件。

稍后输入密码

指定将稍后输入的密码。

在运行 Trillium Batch System 程序的步骤的“特性”笔记本中输入密码。

11. 单击**确定**以导入 Trillium 元数据并关闭笔记本。如果脚本或 JCL 不是从缺省代理站点运行的，应在 Trillium Batch System 步骤的“特性”笔记本中指定正在使用的仓库代理站点。

导入操作完成后，以下仓库对象将添加到“仓库”树。

- Trillium Batch System.*scriptName* 模板，其中 *scriptName* 是脚本或 JCL 文件的名称。
- Trillium Batch System 过程。
- 运行用户定义的程序的 Trillium Batch System 步骤。
- 在导入元数据时指定的仓库文件源和仓库文件目标。文件源和文件目标都是固定文件。
- Trillium Batch System 程序组。

映射元数据

要创建源和目标文件的元数据，Trillium Software System 会读取 Trillium DDL 文件。DDL 文件转换为表23中显示的数据类型。

表 23. 映射 DDL 为“数据仓库中心”数据类型

DDL 数据类型	仓库数据类型
ASCII CHARACTER	CHARACTER(<i>n</i>)
ASCII NUMERIC	
EBCDIC CHARACTER	
EBCDIC NUMERIC	
其它类型	NUMERIC

注：只有在 OS/390 操作系统上运行 Trillium Software System 时，才支持 EBCDIC CHARACTER 和 EBCDIC NUMERIC 数据类型。

变量 *n* 是字符串中的字符数。

限制

通过 Trillium DDL 和“数据仓库中心”中的导入元数据操作，可在输入和输出 DDL 文件中指定覆写字段。但是，相应的仓库源和仓库目标文件不能在“数据仓库中心”中与 SQL 步骤或“样本内容”一起使用。因为导入元数据操作会忽略跨整个记录的覆写字段，您仍可以指定这些字段，但它们不会用作产生的源文件和目标文件中的列。

如果指定了错误文件，则脚本的名称不能包含任何空格。

编写 Trillium Batch System JCL 文件

如果编写 Trillium Batch System JCL 文件，必须符合以下需求。

- 作业名必须是用户标识加上一个字符。
- 作业必须发送到保留输出类。
- 运行 Trillium Batch System 程序的每项作业步骤必须包括定义永久数据集的 `SYSTEM DD` 语句。数据集包含了 Trillium Batch System 程序中的错误。此数据集将在提交 JCL 前自动删除。要获取错误处理和错误报告的更多信息，参见第244页的『错误处理』。

输出错误文件必须当在远程主机上运行脚本或 JCL 时指定；否则，错误消息将不会捕获错误消息并将错误消息返回到“数据仓库中心”。在 UNIX 或 Windows 上，捕获错误消息的最简单方法就是编写另一个调用 Trillium Batch System 脚本并将标准错误以管道方式传送到输出文件中的脚本。

```
//SYSTEM DD UNIT=&UNIT,  
//          DISP=(MOD,CATLG,KEEP),  
//          SPACE=(400,(20,20),,,ROUND),  
//          DSN=&PROJPREF.&TRILVER.&PROJECT.STDERR;
```

图 18. 包括了 `SYSTEM DD` 语句的作业步骤的示例

在 UNIX 和 Windows 上编写 Trillium Batch System 脚本文件

如果 Trillium Batch System 脚本或参数文件包含输入文件的相对路径，用户必须在脚本文件的开始放置一条 `cd` 语句，以将目录切换到脚本文件的目录。

定义 Trillium Batch System 步骤

在定义 Trillium Batch System 步骤之前，必须导入要在过程中使用的 Trillium 元数据。要添加 Trillium Batch System 步骤到过程中：

1. 在过程建模器中打开该过程。
2. 在选用板上单击 **Trillium Batch System** 图标。
3. 单击 **Trillium Batch System 程序** → `programName`，其中 `programName` 是要使用的 Trillium Batch System 程序的名称。
4. 在画布上单击希望步骤出现的位置。
5. 在联机帮助中的主题“定义运行用户定义的程序的步骤”中完成这些步骤。

使用 Trillium Batch System 用户定义的程序

Trillium Batch System 用户定义的程序随“DB2 数据仓库中心版本 7.2 和 Windows NT 版和 UNIX 版”附带。在导入 Trillium 元数据时创建的 Trillium Batch System 步骤将运行 Trillium Batch System 用户定义的程序。用户定义的程序将调用 Trillium Batch System 脚本或 JCL。表24包含了 Trillium Batch System 脚本或 JCL 的参数。

表 24. Trillium Batch System 脚本或 JCL 的参数

参数	值
远程主机	<ul style="list-style-type: none"> • localhost 是缺省值。如果 Trillium Batch System 安装在仓库代理站点上，可使用此值。 • 如果 Trillium Batch System 安装在远程操作系统上，则它是远程主机的名称。
脚本或 JCL	脚本或 JCL 的名称
远程操作系统	<p>远程主机上操作系统的名称。如果“远程主机参数”的值是 localhost，则忽略此参数。有效的值是：</p> <ul style="list-style-type: none"> • MVS™，对于 OS/390 操作系统 • UNIX，对于 AIX、Solaris 操作系统、HP-UX 和 NUMA/Q 操作系统 • WIN，对于 Windows NT 和 2000 操作系统
远程用户标识	具有发出远程序命令的权限的用户标识。如果 RemoteHostName 的值是 localhost，则忽略此参数。

清理名称和地址数据

表 24. *Trillium Batch System* 脚本或 *JCL* 的参数 (续)

密码选项	获取密码的方法。有效的值是: ENTERPASSWORD 如果密码在下一个参数中传送, 则使用此值。 PASSWORDNOTREQUIRED 如果不需要密码, 则使用此值。 GETPASSWORD 如果程序名在下一个参数中传送, 则使用此值。 要求: <ul style="list-style-type: none">• 程序必须驻留在代理站点上, 且将密码写到输入文件的第一行上, 并在成功运行后返回 0。• “密码”参数的值必须与密码程序一样。• “程序参数”参数的值必须是一个包括在双引号中的字符串。• 该字符串中的第一个参数必须是将把密码写到的输出文件的名称。
密码	有效的值是密码或密码程序名。密码程序对于仓库代理站点必须是本地的。
程序参数	密码程序的参数。
输出错误文件	输出错误文件的名称。

注: 此表中所有参数的数据类型都是 CHARACTER。

错误处理

Trillium Batch System 程序将错误消息写到 Windows NT 和 UNIX 操作系统上的标准错误 (stderr) 文件中, 以及 OS/390 操作系统上的 SYSTEMM 数据集中。

要捕获 Windows NT 或 UNIX 操作系统上 *Trillium Batch System* 程序中的错误, 必须将标准错误重定向到一个输出错误文件中。

要捕获 OS/390 操作系统上 *Trillium Batch System* 程序中的错误, *JCL* 必须包括 SYSTEMM DD 语句。

如果在“导入元数据”窗口中指定输出错误文件名，必须重定向或存储标准错误输出到错误文件中。“数据仓库中心”将读取该文件并报告所有包含字符串 `ERROR` 作为错误消息的行。所有 Trillium Batch System 程序错误消息都包含字符串 `ERROR`。

如果未在运行在仓库代理站点上的脚本或 JCL 中指定输出错误文件，“数据仓库中心”将自动创建文件名并将标准错误输出重定向到那个文件中。如果遇到错误，将不删除错误文件。错误文件存储在由环境变量 `VWS_LOGGING` 指定的目录中。文件名是 `tbsudp-date-time.err`，其中 `date` 是创建文件时的系统日期，而 `time` 是创建文件时的系统时间。以下文件名显示了输出错误文件名的格式：

```
tbsudp-021501-155606.err
```

错误返回码

表 25. 错误号和描述

错误号	描述
0	成功
4	警告。在 Trillium Batch System 用户定义的程序访问临时文件时，无法擦除密码文件或发生了内部错误。 检查密码文件或由环境变量 <code>VWS_LOGGING</code> 指定的目录下创建的所有临时文件的状态。
8	参数的数量或值不正确。阅读日志文件或文档以获取正确的语法。
12	Trillium Batch System 用户定义的程序通过 FTP 连接到远程主机时发生了问题。检查 FTP 连接或主机名、用户标识和密码。
16	Trillium Batch System 用户定义的程序不能创建日志或内部文件。 检查用户是否有正确的权限。检查磁盘是否已满。
20	不能执行 OS/390 JCL，或在 Trillium Batch System 用户定义的程序通过 FTP 从 OS/390 删除文件或获取文件时出现了错误。 检查 JESLogFile 找出错误原因。
48	找不到环境变量 <code>VWS_LOGGING</code> 或无法创建日志文件。检查日志文件以获取更多信息。
56	不能执行 Windows NT 或 UNIX 脚本，或在 Trillium Batch System 用户定义的程序与远程主机连接时出现了错误。检查连接或主机名、用户标识和密码。

清理名称和地址数据

表 25. 错误号和描述 (续)

500	脚本或 JCL 文件返回了错误，或虽然未返回错误但错误文件中包含了数据。检查日志文件以获取更多信息。在 OS/390 上，还应检查 JESLogFile。
-----	---

日志文件

“数据仓库中心”在 Trillium Batch System 用户定义的程序运行时，将所有诊断信息存储在日志文件中。日志文件的名称是 `tbsudp-date-time.log`，其中 `date` 是创建文件时的系统日期，而 `time` 是创建文件时的系统时间。日志文件创建在由代理站点上的环境变量 `VWS_LOGGING` 指定的目录中。如果 Trillium Batch System 用户定义的程序运行成功，将删除日志文件。

第10章 数据建模

本章描述如何使用 IBM ERwin Metadata Extract Program 从 ER1 文件中抽取元数据并创建“DB2 数据仓库中心”或“信息目录管理器”（以前称为 DataGuide）标记语言文件。

什么是 IBM ERwin Metadata Extract Program?

元数据抽取程序抽取存储在输入 ER1 文件中的所有对象，例如数据库、表和列，并将元数据模型写到“数据仓库中心”或“信息目录管理器”标记语言文件中。还会抽取并创建由实体和属性组成的“信息目录管理器”的逻辑模型，它创建对象之间的所有相应关系标记，例如数据库和表之间以及表和实体之间。对于没有数据库的表，将创建名为 DATABASE 的缺省数据库。对于没有模式的表，将使用缺省模式 USERID。对于模型名，将使用 ER1 文件名。要获取 ER1 属性与“数据仓库中心”或“信息目录管理器”标记之间的映射的信息，参见第253页的『ERwin 到 DB2 “数据仓库中心”的映射』和第254页的『ERwin 到“信息目录管理器”的映射』。

元数据抽取程序支持带有关系数据库的所有 ER1 模型，包括 DB2、Informix、Oracle、Sybase、ODBC 数据源和 Microsoft SQL Server。

软件需求

运行元数据抽取程序，需要以下软件：

- Windows NT 4.0 或更高版本
- ERwin 3.5.2 with Service Pack 3 Build 466

导入 ERwin 标记语言文件，需要以下软件：

对于“数据仓库中心”：

IBM DB2 通用数据库版本 7.2

对于“信息目录管理器”：

IBM DB2 通用数据库仓库管理器 7.2

模板标记语言文件 (.tag) 必须在 VWS_TEMPLATES 环境变量所指向的目录中。

类型标记语言文件 (.typ) 必须在 DGWPATH 环境变量所指向的目录中。

程序文件

元数据抽取程序安装在 IBM DB2 目录的 `sqlib\bin` 子目录中。程序在您的目录中安装以下文件:

flgerwin.exe

主迁移程序

erwext.dll

标记语言文件生成器 DLL

cdmerwsn.dll

ERwin API 包装器类 DLL

要启动抽取程序, 从命令提示符发出 `flgerwin` 命令。

创建标记语言文件

要创建“数据仓库中心”或“信息目录管理器”标记语言文件, 可运行 `flgerwin.exe` 程序并提供两个主参数。第一个参数是 `-dwc` (如果创建的是“数据仓库中心”标记语言文件) 或 `-icm` (如果创建的是“信息目录管理器”标记语言文件)。第一个参数是从中抽取元数据的 ER1 文件的名称。第二个参数是输出标记语言文件的名称。对于“数据仓库中心”标记语言文件, 第二个参数可以是 `-starschema` 或 `-m`。对于“信息目录管理器”标记语言文件, 第二个参数可以是 `-m`、`-u`、`-a` 或 `-d`。缺省情况下, 抽取程序会将 `MERGE` 参数添加到“数据仓库中心”标记语言文件中。

命令语法是:

```
flgerwin inputFile.er1 outputFile.tag [-dwc] [-icm] [ -m] [-u] [-a] [-d]
```

如果要创建星型模式, 命令的语法是:

```
flgerwin inputFile.er1 outputFile.tag [-dwc] [-starschema]
```

-dwc 第一个参数。创建“数据仓库中心”标记语言文件。可选的第二个参数是 `-m` 和 `-starschema`。

-icm 第一个参数。创建“信息目录管理器”标记语言文件。可选的第二个参数是 `-m`、`-u`、`-a` 和 `-d`。

-starschema

创建 ERwin 模式星型标记语言文件。

-m 指定对对象的操作为 `MERGE`。

-u 指定对对象的操作为 `UPDATE`。

-a 指定对对象的操作为 `ADD`。

-d 指定对对象的操作为 DELETE。

元数据抽取程序用于处理元数据而不是数据。在完成 ERwin 标记语言文件的导入后，使用目标表之前，需要使密码与用户标识相匹配。

要将元数据与现有的数据库数据合并：

1. 单击**特性** → **数据库** → **用户标识**。
2. 更改“数据仓库中心”的用户标识与密码，以匹配合并的数据库的用户标识与密码。
3. 单击**确定**保存更改并关闭窗口。

通过元数据抽取程序，可以导入标记语言文件作为目标。在新导入的元数据中，尚未填充表。可以将这些表作为逻辑或物理表来查看，然后创建一个仓库步骤，使用从 ERwin 导入的定义来填充表。

输入 ER1 文件必须处于可写状态。元数据抽取程序运行后，ER1 文件变为只读。要将文件更改为“读/写”方式，可如以下示例中所示使用命令：

```
attrib -r erwinsimplemode.er1
```

其中 erwinsimplemode.er1 是 ERwin 平面文件的名称。

如果 ER1 文件正在当前 ERwin 会话中使用，或检测到某些错误情况，元数据抽取程序将把该文件保存为只读状态。如果 ER1 文件处于只读状态，您可能会收到一个异常程序中断错误消息。元数据抽取程序会显示当前正在处理的表名。当元数据抽取程序处理完成后，您将接收到一条消息性消息。

通过自动连接维表到事实表来创建星型模式时，取决于要使用的表的数量，处理可能会花较长时间。在处理期间，自动连接行变为绿色。保存模式时，这些行变为黑色。

使用自动生成的约束名以确保约束名是唯一的。

在处理期间，可能会收到一个消息“发现了重复的列。将不抽取列。”这是一条消息性消息，它不会影响抽取程序的成功完成。当外键的物理名与当前正在处理的表中列的物理名相同时，将显示该消息。

将标记语言文件导入“数据仓库中心”

可以有两种方式将标记语言文件导入“数据仓库中心”。可以使用“数据仓库中心”或命令行。

要使用“数据仓库中心”来导入标记语言文件:

1. 单击**开始** → **程序** → **IBM DB2** → **控制中心**。“DB2 控制中心”打开。
2. 打开“数据仓库中心”然后登录。
3. 右键单击**仓库**。“导入”窗口打开。
4. 单击**导入元数据** → **ERwin**。导入元数据窗口打开。
5. 在**输入文件**字段中, 输入输入标记语言文件的名称, 然后单击**确定**。
6. 选择**抽取星型模式**复选框, 并定义 ERwin 星型模式元数据模型作为仓库模式。

导入完成后, 可单击**视图** → **刷新**来查看新步骤。

要使用命令行来导入标记语言文件, 可输入以下命令:

```
iwh2imp2 tag-filename log-pathname  
target-control-db userid password
```

tag-filename

标记语言文件的全路径和文件名。

log-pathname

日志文件的全路径名。

target-control-db

导入操作的目标数据库的名称。

userid 用于访问控制数据库的用户标识。

password

用于访问控制数据库的密码。

要更改 DB2 数据库定义, 以使它成为“数据仓库中心”的源, 您可以更改标记语言文件:

- 对要让它成为源的每个数据库, 将 ISWH 标记从 ISWH(Y) 更改为 ISWH(N)。
- 对要让它成为源的每个数据库, 将关系标记从 **:RELTYPE.TYPE(LINK) SOURCETYPE(SCGTARIR) TARGETYPE(DATABASE)** 更改为 **:RELTYPE.TYPE(LINK) SOURCETYPE(SCGSR CIR) TARGETYPE(DATABASE)**。

标记语言文件导入时，可能会收到以下消息：消息：DWC13238E 由 "DBNAME(____) OWNER(____) TABLE(____) COLUMNS(____)" 标识的类型为 "COLUMN" 的对象在标识语言文件中定义了两次。

这只是一条消息性消息，您的导入操作会成功完成。如果实体中有具有相同名称的外键，或者实体中有类似命名的列受到了截断的影响，或有其他类似的环境，您都可能收到此消息。检查您的模型，验证是否有重复的列名称，并作出相应的更改。

将标记语言文件导入“信息目录管理器”

将标记语言文件导入“信息目录管理器”有两种方法。可以使用“信息目录管理员”或使用命令行。

要使用“信息目录管理员”来导入标记语言文件：

1. 单击开始 → 程序 → **DB2** → 信息目录管理器。
2. 单击目录 → 导入。“导入”窗口打开。
3. 单击查找来搜索标记语言文件，然后单击导入。

导入完成后，可以双击主题图符，将打开一个窗口，其中显示了导入的所有模型和数据库。

要用命令界面导入标记语言文件，输入以下命令：

```
DGUIDE /USERID userid /PASSWORD password
/DGNAME dgname /IMPORT filename /LOGFILE
filename /ADMIN /RESTART (B|C)
```

/USERID

用于访问控制数据库的用户标识。

/PASSWORD

此用户标识的密码。

/DGNAME

“信息目录”的名称。

/IMPORT

标记语言文件的全路径和文件名。

/LOGFILE

日志文件的全路径名。

/ADMIN 指定用户具有管理员权限。

/RESTART

指定导入将在标记语言文件的开始处（选项 B）开始或从最后的落实点（选项 C，缺省值）开始。

故障诊断

如果收到错误消息，可查看此处的消息和如何解决该错误的说明。

缺少 ER1 输入文件或标记输出文件。

元数据抽取程序需要具有特定次序的两个参数。第一个参数是 ER1 文件的名称。第二个参数是标记语言输出文件的名称。如果指定一个现有标记语言文件的名称，将覆盖该文件。

Windows 系统异常程序中断。

输入 ER1 文件可能是只读状态。如果在保存 ER1 文件时出现了问题，因此元数据抽取程序将该文件置于只读状态时，可能会发生这种情况。在命令 shell 中发出命令：

```
attrib -r inputFile.er1
```

，将 ER1 文件的状态更改为读 / 写。

无法打开标记语言文件 <file name>。

检查是否存在任何可能阻止文件在当前驱动器上创建或打开的问题。

找不到指向模板文件的路径。

未设置环境变量 VWS_TEMPLATES。检查是否安装了“数据仓库中心”。

未找到指向类型文件的路径。

未设置环境变量 DGWPATH。检查是否安装了“数据仓库中心”。

不支持的服务器版本: <版本>

尝试从中抽取的输入 ER1 文件存储在不受程序支持的目标服务器上。启动 ERwin，打开 ER1 文件，然后单击**服务器** → **目标服务器**以及相应的版本。有关更多信息，参阅第247页的『软件需求』。保存 ER1 文件。

未知的 ERwAPI 错误。

发生了 ERwin API 错误，程序无法获取关于该错误的更多信息。确保安装了 ERwin 3.5.2。必须注册 ERwin API。

要注册 ERwin API，可从安装了 ERwin 程序文件的目录中运行以下命令：**regsvr32 er2api32.dll**。将看到消息“er2api32.dll 中的 DllRegisterServer 成功。可从“数据仓库中心”中启动抽取程序，或从命令 shell 中发出 flgerwin 命令以启动抽取程序。

抽取程序错误: <错误消息>

检查错误消息并采取相应的措施。最有可能的原因是: 这是一个内部抽取程序错误, 需将此问题报告给“IBM 软件支持”部门。

未知的抽取程序错误(n.)。

出现了未知的错误。最有可能的原因是: 这是一个内部错误, 需将此问题报告给“IBM 软件支持”部门。

抽取程序由于出现错误而终止。

由于出现错误而阻止了抽取程序的完成。参考其他错误消息以解决该问题或与“IBM 软件支持”联系。

ERwin 到 DB2 “数据仓库中心”的映射

表26通过表28 显示主 ERwin 对象属性如何与“数据仓库中心”标记相对应:

表 26. 数据库 - WarehouseDatabase.tag 或 SourceDatabase.tag

ERwin	命令行标记	数据仓库中心
图表名称	NAME	仓库源或仓库目标的名称
图表作者	RESPNSBL	联系人
数据库名	DBNAME	数据库名
数据库版本	DBTYPE	数据库类型
图表描述	SHRTDESC	描述

表 27. 表 - Table.tag

ERwin	命令行标记	数据仓库中心
表名	NAME	表名
表名	TABLES	表名
数据库名	DBNAME	不适用
表所有者	OWNER	表模式
表注释	SHRTDESC	描述

表 28. 列 - Column.tag

ERwin	命令行标记	数据仓库中心
列名	NAME	列名
数据类型	NATIVEDT	数据类型
长度	LENGTH	长度
刻度	SCALE	刻度

数据建模

表 28. 列 - *Column.tag* (续)

ERwin	命令行标记	数据仓库中心
Null 选项	NULLABLE	允许 Null (复选框)
位置	POSNO	不适用
主键	KEYPOSNO	不适用
数据库名	DBNAME	不适用
表所有者	OWNER	不适用
表名	TABLES	不适用
列注释	SHRTDESC	描述

ERwin 到“信息目录管理器”的映射

表29通过第255页的表34 显示主 ERwin 对象属性如何与“信息目录管理器”标记相对应:

表 29. 数据库 - *Database.tag*

ERwin	命令行标记	信息目录管理器界面
图表名称	NAME	数据库名
图表作者	RESPNSBL	数据库所有者
数据库名	DBNAME	数据库名
数据库版本	DBTYPE	数据库类型
图表描述	SHRTDESC	简短描述

表 30. 表 - *TableOrView.tag*

ERwin	命令行标记	信息目录管理器界面
表名	NAME	表名
表名	TABLES	表名
数据库名	DBNAME	数据库名
表所有者	OWNER	表所有者
表注释	SHRTDESC	简短描述
ERwin API	TABLVIEW	定义表示视图

表 31. 列 - *ColumnOrField.tag*

ERwin	命令行标记	信息目录管理器界面
列名	NAME	列名
数据类型	DATATYPE	列的数据类型

表 31. 列 - *ColumnOrField.tag* (续)

ERwin	命令行标记	信息目录管理器界面
长度	LENGTH	列的长度
刻度	SCALE	列的刻度
Null 选项	NULLS	列可以为空吗 (?)
位置	POSNO	列位置
主键	KEYPOSNO	主键中列的位置
ERwin API	ISKEY	列是键的一部分吗 (?)
ERwin API	UNIKEY	列是唯一键吗 (?)
数据库名	DBNAME	数据库名
表所有者	OWNER	表所有者
表名	TABLES	表名
列注释	SHRTDESC	简短描述
ERwin	ISTEXT	数据是文本吗 (?)
ERwin API	IDSRES	数据解析

表 32. 模型 - *Model.tag*

ERwin	命令行标记	信息目录管理器界面
ER1 文件名	NAME	模型名称
图表作者	RESPNSBL	要获取进一步的信息...
图表描述	SHRTDESC	简短描述

表 33. 实体 - *Entity.tag*

ERwin	命令行标记	信息目录管理器界面
实体名称	NAME	实体名称
注释	SHRTDESC	简短描述
定义	LONGDESC	详细描述
实体所有者	RESPNSBL	要获取进一步的信息...

表 34. 属性 - *Attribute.tag*

ERwin	命令行标记	信息目录管理器界面
属性名称	NAME	属性名称
注释	SHRTDESC	简短描述
定义	LONGDESC	详细描述

数据建模

表 34. 属性 - *Attribute.tag* (续)

ERwin	命令行标记	信息目录管理器界面
数据类型	DATATYPE	成员的数据类型
长度	LENGTH	成员的长度

第11章 刷新“OLAP 服务器”数据库

可使用仓库程序自动刷新“OLAP 服务器”数据库。“数据仓库中心”可帮助您调度 OLAP 服务器执行三种操作:

- 装入数据
- 更新维数
- 运行计算

要使用“数据仓库中心”将数据装入“OLAP 服务器”数据库:

1. 使用“Essbase 管理器”创建“OLAP 服务器”应用程序和数据库。记下应用程序名、数据库名、用户标识和密码。将需要此信息作为仓库程序的输入。
2. 使用“Essbase 管理器”定义该数据库的轮廓。
3. 定义要从 OLAP 服务器的操作源抽取并装入到 Essbase 数据库中的数据。可使用此数据来更新量度（例如，使用 Essbase IMPORT 命令）和维数（例如，使用 BuildDimension 命令）。

4. 定义从操作数据源抽取数据的步骤，并按步骤 3 中定义的方式构建该数据。

5. 将步骤提升为测试方式，并至少运行它一次。

6. 使用“Essbase 管理器”，编写并测试将数据源装入 Essbase 数据库的装入规则。将这些装入规则保存到数据库，或以文件形式保存在该仓库代理站点。

还可定义在装入数据后运行的计算脚本。将计算脚本保存在仓库代理站点上的文件中。

有关定义装入规则和计算脚本的信息，参见 *OLAP Server Database Administrator's Guide*。

7. 定义使用 Hyperion Essbase 中的一个仓库程序（如 DB2 OLAP: 用装入规则装入文本文件中的数据 (ESSDATA2)）的步骤。使用“过程模型”窗口指定抽取数据的步骤将启动此步骤。

8. 将步骤提升为测试方式，并至少运行它一次。

9. 为抽取数据的步骤定义一个调度表，并将该步骤提升为生产方式。

第258页的图19显示“数据仓库中心”和“OLAP 服务器”之间的数据流。

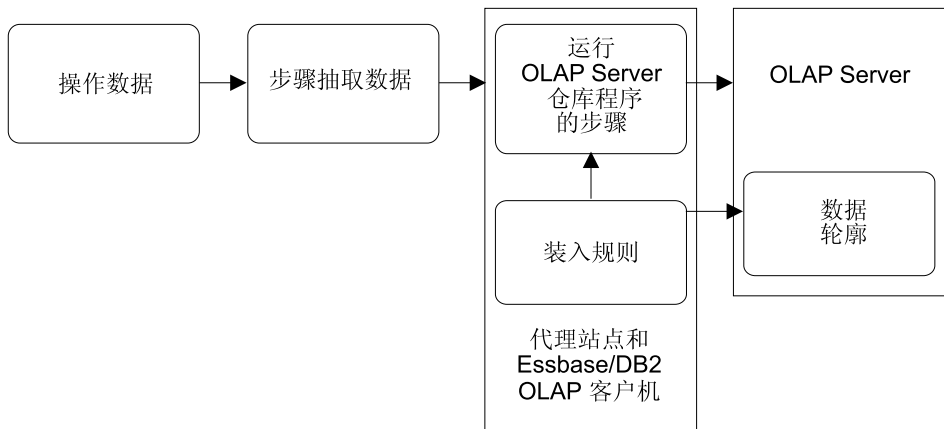


图 19. “数据仓库中心”和“OLAP 服务器”之间的数据流

有关完成此过程的详情，参见“数据仓库中心”联机帮助和《OLAP 服务器：使用 OLAP 服务器》。

定义“OLAP 服务器：缺省 calc” (ESSCALC1) 仓库程序的值

使用“OLAP 服务器：缺省 calc” (ESSCALC1) 仓库程序来调用与目标数据库相关的缺省 calc 脚本。该“OLAP 服务器”可以是一个“OLAP 服务器”或是一个“Essbase 服务器”。

当选择使用“OLAP 服务器：缺省 calc” (ESSCALC1) 仓库程序的步骤时，不要将该步骤与源或目标链接。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要定义运行该仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。

3. 在“参数”页上的“**OLAP 服务器**”系统名字段中，输入要使用的“OLAP 服务器”的名称。
4. 在“**OLAP 服务器**”应用程序名字段中，输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中，输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中，输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中，输入该用户标识的密码。
8. 在验证 **OLAP 服务器**密码字段中，再次输入该密码。
9. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器: 使用计算规则的 Calc” (ESSCALC2) 仓库程序的值

使用“OLAP 服务器: 使用计算规则的 Calc” (ESSCALC2) 仓库程序将指定的 calc 脚本应用于 OLAP 服务器数据库。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

当选择使用此仓库程序的步骤时，不要将该步骤与源或目标链接。

此仓库程序可用于下列操作系统:

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要定义运行此仓库程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中，输入要使用的“OLAP 服务器”的名称。
4. 在“**OLAP 服务器**”应用程序名字段中，输入要使用的 OLAP 服务器应用程序的名称。

刷新“OLAP 服务器”数据库

5. 在“**OLAP 服务器**”数据库名字段中，输入具有要调用的 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中，输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中，输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中，再输入一次该密码。
9. 在 **Calc 脚本文件名** 字段中，输入具有要使用的 calc 脚本文件的名称。若该 calc 脚本的位置标志指示该文件位于 OLAP 服务器系统上，则该文件名必须遵循该客户机或服务器上有关指定文件名的约定。例如，若您正在使用 Essbase 服务器上的文件，则必须指定不带扩展名的文件名。
您还必须使用“Essbase 应用程序管理器”来构建该 calc 脚本文件。为此，您必须将该 calc 脚本存储在使用“Essbase 应用程序管理器”的服务器上。
若程序中的 calc 脚本文件位置标志指示该文件位于一代理站点上，则您必须使用该代理站点的文件系统来指定全限定文件名。
10. 在 **Calc 脚本文件位置** 单选按钮组中，选择 calc 脚本文件的位置（位置标志所指示的）。
11. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
12. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器：自由文本数据装入” (ESSDATA1) 仓库程序的值

使用 OLAP 服务器：自由文本数据装入仓库程序，将数据从逗号定界的文本文件装入多维“OLAP 服务器”数据库中。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

OLAP 服务器：自由文本数据装入仓库程序使用选为该过程中步骤的源的文件。该步骤必须仅选择一个源文件。该源文件中的数据必须与 OLAP 服务器简要表的值完全匹配。选择的源文件必须在该代理站点上。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要定义运行此仓库程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中, 输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中, 输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中, 输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中, 输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中, 输入该用户标识的密码。
8. 在**验证 OLAP 服务器密码**字段中, 再次输入该密码。
9. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
10. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器: 使用装入规则从文件装入数据”(ESSDATA2) 仓库程序的值

使用 OLAP 服务器: 使用装入规则从文件中装入数据 (ESSDATA2)仓库程序, 以使用装入规则将数据从文本文件装入到多维“OLAP 服务器”数据库中。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

该仓库程序使用所选的文件作为该步骤的源。该步骤必须仅选择一个源文件。该源文件名必须遵循 OLAP 服务器系统的客户机或服务器上指定文件名的约定。

若程序中的文件位置标志指示该文件位于一代理站点上, 则您必须在“步骤”笔记本的“参数”页上的**装入规则文件名**字段中指定全限定文件名。

若文件位置标志指示该源文件位于 OLAP 服务器上, 则您必须在“步骤”笔记本的“参数”页上的**装入规则文件名**字段中指定不带扩展名的文件名。

此仓库程序可用于下列操作系统:

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

刷新“OLAP 服务器”数据库

“列映射”页不可用于此步骤。

要定义运行此仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中，输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中，输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中，输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中，输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中，输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中，再输入一次该密码。
9. 在 装入规则文件名 字段中，输入您希望该程序使用的装入规则脚本文件的名称。若该文件在 OLAP 服务器上，则该文件名不能有扩展名。若该文件位于代理站点上，您必须指定全限定文件名。
10. 在装入规则文件位置标志区域中，单击该规则文件的位置。
11. 在源文件位置标志区域中，单击该源文件的位置。
12. 在“**OLAP 服务器**”实用程序异常终止标志区域中，指定若有错误是否终止该 OLAP 服务器实用程序。
13. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
14. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器：使用装入规则从 SQL 表装入数据” (ESSDATA2) 仓库程序的值

使用“OLAP 服务器：使用装入规则从 SQL 表中装入数据” (ESSDATA2)仓库程序，以使用装入规则将数据从 SQL 源装入到多维“OLAP 服务器”数据库。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

OLAP 服务器：使用装入规则从 SQL 表装入数据仓库程序使用在“步骤”笔记本中为用于该步骤的源数据库定义的用户标识和密码。

当您选择一个使用此仓库程序的步骤时，将源表与该步骤链接，但不要将它与目标链接。

此仓库程序可用于下列操作系统:

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要为使用此仓库程序的步骤定义值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中, 输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中, 输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中, 输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中, 输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中, 输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中, 再输入一次该密码。
9. 在装入规则文件名字段中, 输入您希望该程序使用的装入规则脚本文件的名称。若该文件在 OLAP 服务器上, 则该文件名不能有扩展名。若该文件位于代理站点上, 您必须指定全限定文件名。
10. 在装入规则文件位置标志区域中, 单击该规则文件的位置。
11. 在“**OLAP 服务器**”实用程序异常终止标志区域中, 指定若有错误是否终止该 OLAP 服务器实用程序。
12. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
13. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器: 不使用装入规则从文件装入数据”(ESSDATA4) 仓库程序的值

使用“OLAP 服务器: 不使用装入规则从文件中装入数据”(ESSDATA4) 仓库程序, 以不使用装入规则将数据从文本文件装入多维“OLAP 服务器”数据库中。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

刷新“OLAP 服务器”数据库

当选择使用此仓库程序的步骤时，将源文件与该步骤链接，但不要将它与目标表链接。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Sun Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要为使用此仓库程序的步骤定义值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中，输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中，输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中，输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中，输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中，输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中，再输入一次该密码。
9. 在源文件位置标志区域中，单击该源文件的位置。
10. 在“**OLAP 服务器**”实用程序异常终止标志区域中，指定若有错误是否终止该 OLAP 服务器实用程序。
11. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
12. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器：从文件更新轮廓” (ESSOTL1) 仓库程序的值

使用“OLAP 服务器：更新轮廓” (ESSOTL1) 仓库程序，以使用装入规则从源文本文件中更新 OLAP 服务器简要表。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

OLAP 服务器：更新轮廓仓库程序使用所选的文件作为该步骤的源。该步骤必须仅选择一个源文件。该源文件名必须遵循 OLAP 服务器系统的客户机或服务器上指定文件名的约定。

若仓库程序中的文件位置标志指示该文件位于一代理站点上，则您必须在“步骤”笔记本的“参数”页上的**装入规则文件名**字段中指定全限定文件名。

若程序中的文件位置标志指示该源文件位于 OLAP 服务器上，则您必须在“步骤”笔记本的“参数”页上的**装入规则文件名**字段中指定不带扩展名的文件名。

有关详情，参阅 OLAP 服务器的文档。

当选择使用此仓库程序的步骤时，将源文件与该步骤链接，但不要将它与目标表链接。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

刷新“OLAP 服务器”数据库

要为使用此仓库程序的步骤定义值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中, 输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中, 输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中, 输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中, 输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中, 输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中, 再输入一次该密码。
9. 在 装入规则文件名 字段中, 输入您希望该程序使用的装入规则脚本文件的名称。若该文件在 OLAP 服务器上, 则该文件名不能有扩展名。若该文件位于代理站点上, 您必须指定全限定文件名。
10. 在装入规则文件位置标志区域中, 单击该规则文件的位置。在源文件位置标志区域中, 单击该源文件的位置。
11. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
12. 单击**确定**保存更改并关闭步骤笔记本。

定义“OLAP 服务器: 从 SQL 表更新轮廓”(ESSOTL2) 仓库程序的值

使用“OLAP 服务器: 从 SQL 表更新轮廓”(ESSOTL2) 仓库程序, 以使用装入规则从 SQL 源中更新 OLAP 服务器简要表。该 OLAP 服务器可以是一个 OLAP 服务器或是一个 Essbase 服务器。

当您选择使用此仓库程序的步骤时, 将源表与该步骤链接, 但不要将它与目标表链接。

此仓库程序可用于下列操作系统:

- Windows NT
- AIX
- Solaris 操作环境
- AS/400

“列映射”页不可用于此步骤。

要为使用此仓库程序的步骤定义值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上的“**OLAP 服务器**”系统名字段中, 输入要使用的 OLAP 服务器的名称。
4. 在“**OLAP 服务器**”应用程序名字段中, 输入要使用的 OLAP 服务器应用程序的名称。
5. 在“**OLAP 服务器**”数据库名字段中, 输入具有要调用的缺省 calc 脚本的 OLAP 服务器数据库的名称。
6. 在“**OLAP 服务器**”用户标识 字段中, 输入要用于与 OLAP 服务器数据库连接的用户标识。
7. 在“**OLAP 服务器**”密码字段中, 输入该用户标识的密码。
8. 在验证“**OLAP 服务器**”密码字段中, 再输入一次该密码。
9. 在装入规则文件名字段中, 输入您希望该程序使用的装入规则脚本文件的名称。若该文件在 OLAP 服务器上, 则该文件名不能有扩展名。若该文件位于代理站点上, 您必须指定全限定文件名。
10. 在装入规则文件位置标志区域中, 单击这些规则文件的位置。
11. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
12. 单击**确定**保存更改并关闭步骤笔记本。

刷新“OLAP 服务器”数据库

第12章 维护仓库数据库

本章描述如何使用仓库程序来维护仓库数据库:

- 可以使用 DB2 重组仓库程序来重新安排物理存储器中的表。从而可消除碎片并确保表有效地存储在数据库中。还可以使用重组来控制表中的行存储的次序,一般根据索引来进行。
- 可以使用 DB2 runstats 仓库程序来收集关于表及其索引的物理和逻辑特性的统计信息。DB2 通用数据库使用这些统计信息来确定访问数据的最佳方法。

有关对数据库进行调整和配置以获取最佳性能的信息,请参阅《DB2 通用数据库管理指南:性能》。

重组数据

可以重组“DB2 通用数据库”表中的数据或“DB2 UDB OS/390 版”表空间或索引中的数据。

定义 DB2 通用数据库 REORG 程序的值

使用 DB2 UDB REORG 仓库程序的“步骤特性”笔记本来创建可用于重组 DB2 通用数据库表中数据的步骤。

可使用仓库源或目标作为此步骤子类型的源。REORG 程序向源表进行写入。

在此步骤中,“列映射”页不可用。

要定义运行 DB2 UDB REORG 仓库程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。要获取更多信息,参见第139页的『提供关于步骤的一般信息』。
3. 可选:在“参数”页上,指定 REORG 步骤的信息:
 - 在使用临时表空间字段中,输入应在 REORG 步骤运行期间使用的临时表空间的名称。
 - 在使用索引字段中,输入应在 REORG 步骤运行期间使用的索引的名称。
4. 在“处理选项”页上,提供有关步骤如何处理的信息。要获取更多信息,参见第141页的『定义处理选项』。
5. 单击**确定**以保存更改并关闭步骤笔记本。

定义 DB2 OS/390 版 REORG 程序的值

使用“DB2 OS/390 版 REORG 数据仓库中心”程序，可重组表空间以改进访问性能以及对空间进行碎片整理。此外，还可以重组分区表空间的单个分区。

要为使用此仓库程序的步骤定义值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。要获取更多信息，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，指定希望程序在重组前，重组期间及重组后如何处理数据。

要指定重组参数：

- a. 在**表空间**字段中，输入包含您的表（也可以是该表所属的数据库）的表空间的名称。
- b. 可选：选择**重新装入期间记录日志**复选框，以在重组程序的重新装入阶段记录日志。如果未记录日志，表空间将仅在建立了映象副本后才是可恢复的。对 DSNDB01 或 DSNDB06 中的任何表空间（或选择了 SHRLEVEL REFERENCE 或 CHANGE）都不允许在重新装入阶段记录日志。LOB 表空间需要记录日志。
- c. 可选：选择**逻辑复位并重用数据集**复选框，以将数据装入一个非空数据集。
- d. 可选：指定在卸装数据后改进性能的选项：
 - 选择**按表空间扫描卸装数据并以群集顺序排序数据**复选框，以使用顺序扫描来卸装数据。如果不指定此选项，重组程序会使用聚类索引来卸装数据。
 - 选择**按并行方式排序索引键**复选框，以指定在 SORTBLD 阶段索引键将以并行方式排序，以改进性能。
- e. 可选：指定允许的访问类型：
 - 选择**没有应用程序访问**复选框以从正在重组的区域卸装数据（此时，应用程序可从该区域读但不能写），将数据重新装入该区域（此时应用程序没有访问权），然后再次允许读写/访问。此选择是 LOB 表空间所需的。
 - 选择**只读应用程序访问**复选框，从正在重组的区域卸装数据（此时应用程序对该区域可读但不能写），将数据重新装入该区域的影子副本（此时应用程序可读但不能写向该区域的原始副本），通过更换数据集的名称，将应用程序未来的访问权从原始副本切换为影子副本，然后再次允许读/写访问。
 - 选择**读/写应用程序访问**复选框，从正在重组的区域卸装数据（此时应用程序可读写该区域），将数据重新装入该区域的影子副本（此时应用

程序对该区域的原始副本具有读 / 写访问权)，将原始副本的日志应用到影子副本（此时应用程序可读且通常会写到原始副本中），通过更换数据集的名称，将应用程序未来的访问权从原始副本切换到影子副本，然后再次允许读 / 写访问权。

- f. 可选：指定映射表信息：
- 在**所有者**字段中，输入映射表的所有者姓名。
 - 在**名称**字段中，输入映射表的名称，重组程序使用该表在该区域的原始副本中数据记录的记录标识符 (RID) 与影子副本中的相应 RID 之间映射。此表必须已存在并且包括了索引。包含映射表的表空间必须已分段且不能是将重组的那个表空间。
- g. 可选：在**卸装数据集**选项的**数据集名称**字段中，输入包含将要卸装的数据的数据集的数据集名称。缺省名称是 **SYSREC**。数据集必须是一个顺序数据集，可由 **MVS BSAM** 访问方法来访问。
- 指定直接访问存储设备类型，已授权您将在该设备类型中分配空间。将频繁访问的对象指定为快速设备而较少使用的对象指定为慢速设备。也可以指定磁带设备。
 - 为您的设备分配适用的空间。如果存储设备是磁盘，则空间分配就应是柱面。如果设备类型是磁盘，可以保留该项为空白。
- h. 可选：在**本地站点的主复制数据集**选项的**数据集名称**字段中，输入主复制数据集的名称。此数据集包含了映像副本数据集。缺省名称是 **SYSCOPY**。
- 指定直接访问存储设备类型，已授权您将在该设备类型中分配空间。将频繁访问的对象指定为快速设备而较少使用的对象指定为慢速设备。也可以指定磁带设备。
 - 为您的设备分配适用的空间。如果存储设备是磁盘，则空间分配就应是柱面。如果设备类型是磁盘，可以保留该字段为空白。
- i. 对于**输入排序数据集名称**和**输出排序数据集**选项，输入排序输入和输出 (**SYSUT1** 和 **SORTOUT**) 数据集的临时工作文件的名称。当表有索引或此笔记本的选项页上的 **SORTKEYS** 没有使用估计值或使用了估计值 0 指定的时，这些文件是必需的。
- 指定直接访问存储设备类型，已授权您将在该设备类型中分配空间。将频繁访问的对象指定为快速设备而较少使用的对象指定为慢速设备。也可以指定磁带设备。
 - 为您的设备分配适用的空间。如果存储设备是磁盘，则空间分配就应是柱面。如果设备类型是磁盘，可以保留该字段为空白。
- j. 可选：在 **DFSORT 临时数据集的设备类型**字段中，输入将由 **DFSORT** 动态分配的临时数据集的设备类型。它可以是 **DFSORT** 的 **SORT** 或 **OPTION** 选项的 **DYNALLOC** 参数可接受的任何设备类型。

- k. 可选: 选择 **DFSORT 临时数据集数**复选框以使用 DFSORT 临时数据集。输入要由排序应用程序动态分配的数据集的数量。仅当在前一选择中选择了设备类型时才使用 SORTNUM 字段。
- l. 可选: 单击**高级**按钮以打开“重组表空间”笔记本。从此笔记本中, 可指定重组表空间的其他选项。
4. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
5. 单击**确定**保存更改并关闭步骤笔记本。

定义“DB2 OS/390 实用程序”程序的值

使用“DB2 OS/390 版实用程序”仓库程序, 可运行任何 DSNUTILS 支持的实用程序。要获取关于 DSNUTILS 的更多信息, 请参阅 *DB2 Universal Database for OS/390 and z/OS Utility Guide and Reference*。

要定义使用此仓库程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上, 双击**参数值**字段, 为**参数名**字段中显示的相应参数添加值。此处为这些参数及其描述的列表:

UTILITY_ID

指定此实用程序在 DB2 内的唯一标识符 (UID)。它是 1 至 16 个字符的字符串。

RESTART

指定此仓库程序是否重新启动当前的实用程序, 若是, 指定在何处重新启动它。NO 或空值指示该实用程序是新的, 不是重新启动的程序。不得有任何其他的实用程序使用同一个实用程序标识符 (UID)。缺省值为空。

UTSTMT

指定实用程序控制语句。它是一个 VARCHAR(32704) 类型的输入参数。

RETCODE

指定该实用程序的最高返回码。它是一个 INTEGER 类型的输出参数。

UTILITY_NAME

指定要启动的实用程序。仅指定单个需要在 UTSTMT 参数中设置数据

集分配的实用程序。要获取可在此处指定的实用程序的更多信息，请参阅 *DB2 Universal Database for OS/390 and z/OS Utility Guide and Reference*。

RECDSN

指定重组表空间时所需的作为卸装数据集的编目数据集名称。若指定 RECDSN，会将它分配给 SYSREC DDNAME。

RECDEVT

为 RECDSN 中列示的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

RECSpace

指定要用作 RECDSN 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

DISCDSN

指定编目数据集名称，用于在重组表空间时作为废弃数据集，以保存不重新装入的记录。若指定 DISCDSN，会将它分配给 SYSDISC DDNAME。

DISCDEVT

为 DISCDSN 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

DISCSpace

指定要用作 DISCDSN 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

PNCHDSN

用关键字 UNLOAD EXTERNAL 或 DISCARD 指定重组表空间时所用的编目数据集名称。该数据集用于保存生成的 LOAD 实用程序控制语句。若为 PNCHDSN 指定值，会将它分配给 SYSPUNCH DDNAME。

PNCHDEVT

为 PNCHDSN 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

PNCHSpace

指定要用作 PNCHDSN 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

COPYDSN1

指定目标（输出）数据集的名称。若指定 COPYDSN1，会将它分配给 SYSCOPY DDNAME。

COPYDEVT1

为 COPYDSN1 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

COPYSPACE1

指定要用作 COPYDSN1 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

COPYDSN2

指定用作备份副本的目标（输出）数据集的编目数据集的名称。若指定 COPYDSN2，会将它分配给 SYSCOPY2 DDNAME。

COPYDEVT2

为 COPYDSN2 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

COPYSPACE2

指定要用作 COPYDSN2 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

RCPYDSN1

指定用作远程站点主副本的目标（输出）数据集的编目数据集的名称。若指定 RCPYDSN1，会将它分配给 SYSRCPY1 DDNAME。

RCPYDEVT1

为 RCPYDSN1 数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

RCPYSPACE1

指定要用作 RCPYDSN1 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

RCPYDSN2

指定用作远程站点备份副本的目标（输出）数据集的编目数据集的名称。若指定 RCPYDSN2，会将它分配给 SYSRCPY2 DDNAME。

RCPYDEVT2

为 RCPYDSN2 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

RCPYSPACE2

指定要用作 RCPYDSN2 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

WORKDSN1

指定需要用作排序输入和输出的工作数据集的编目数据集的名称。若指定 WORKDSN1，会将它分配给 SYSUT1 DDNAME。

WORKDEVT1

为 WORKDSN1 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

WORKSPACE1

指定要用作 WORKDSN1 所指定数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

WORKDSN2

指定需要用作排序输入和输出的工作数据集的编目数据集的名称。若正在使用重组非唯一类型 1 索引，则它是必需的。若指定 WORKDSN2，会将它分配给 SORTOUT DDNAME。

WORKDEVT2

为 WORKDSN2 所指定数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

WORKSPACE2

指定要用作 WORKDSN2 数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

MAPDSN

指定在带 ENFORCE CONSTRAINTS 的 LOAD 期间，需要用作错误处理的工作数据集所需的编目数据集的名称。对于 LOAD，它是可选的。若指定 MAPDSN，会将它分配给 SYSMAP DDNAME。

MAPDEVT

为 MAPDSN 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

MAPSPACE

指定要用作 MAPDSN 数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

ERRDSN

指定需要用作错误处理的工作数据集的编目数据集的名称。若指定 ERRDSN，会将它分配给 SYSERR DDNAME。

ERRDEVT

为 ERRDSN 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

ERRSPACE

指定要用作 ERRDSN 数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

FILTRDSN

指定需要用作错误处理的工作数据集的编目数据集的名称。若指定 FILTRDSN，会将它分配给 FILTER DDNAME。

FILTRDEVT

为 FILTRDSN 指定的数据集所驻留的设备指定设备地址、类属设备类型或用户指定的组名。

FILTRSPACE

指定要用作 FILTRDSN 数据集的主空间分配的磁道柱面数。次空间分配为主空间分配的 10%。

4. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
5. 单击**确定**保存更改并关闭步骤笔记本。

更新系统编目统计信息

可以更新 DB2 UDB 或 DB2 OS/390 版数据库的系统编目统计信息。

定义 DB2 UDB RUNSTATS 程序的值

使用 DB2 UDB RUNSTATS 仓库程序的“步骤特性”笔记本来创建可用于更新下列数据的系统编目统计信息：表中的数据、表索引中的数据或表及其索引中的数据。优化器使用这些统计信息来选择将使用哪个路径访问数据。

一般，若对表中数据进行了较大的更改，则需要更新统计信息。

RUNSTATS 程序使用仓库目标作为源和目标。在定义某个步骤的值之前，将仓库目标与“处理模型”窗口中的那个步骤链接。

在此步骤中，“列映射”页不可用。

要定义运行 DB2 UDB RUNSTATS 仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 可选：在“参数”页上，指定 RUNSTATS 仓库程序的信息：
 - 通过单击**表统计信息**下的单选按钮，指定想要为表收集的统计信息级别。
 - 通过选择**索引统计信息**下的单选按钮，指定想要为表的索引收集的统计信息级别。

- 使用**共享级别**单选按钮，指定想要其他用户在收集统计信息时对该表具有的访问权类型。
4. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
 5. 单击**确定**保存更改并关闭步骤笔记本。

定义 DB2 OS/390 RUNSTATS 程序的值

使用“DB2 OS/390 版 RUNSTATS”仓库程序可收集关于表空间中数据特征的摘要信息。输出记录在 DB2 OS/390 版的目录表中。DB2 使用此信息选择绑定处理期间到数据的访问路径。该信息可供数据库管理员用于评估数据库的设计，并帮助管理员确定何时必须重组表空间或索引。

“列映射”页不可用于此步骤子类型。

要为使用此仓库程序的步骤定义值：

1. 定义将运行实用程序的仓库目标。
2. 添加 RUNSTATS 步骤到过程中。
3. 将该步骤连接到目标。
4. 打开步骤笔记本。
5. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
6. 在“参数”页上，单击其中一个**共享级别**单选按钮，以指定在收集统计信息时其他程序对该表空间可使用的访问权。
 - 若想要允许其他程序只读访问，则指定**仅引用**。这是缺省值。
 - 若想要允许其他程序更改该表空间或索引，则指定**允许更改**。未落实的数据可以收集进统计摘要。
7. 单击其中一个**报告统计信息**单选按钮，指定是否生成消息集合来报告所收集的统计信息。若想要将消息集合发送到 SYSPRINT 数据集，则指定**是**。指定**否**，则不将消息发送到 SYSPRINT。这是缺省值。
8. 单击其中一个**更新目录表上的选项**单选按钮，以指定是否将收集的统计信息插入到 DB2 OS/390 版目录表：

访问路径和空间统计信息

收集统计信息的结果会更新 DB2 目录中的所有统计信息。

仅访问路径统计信息

收集统计信息的结果仅更新用于选择访问路径的目录表列。

仅空间统计信息

收集统计信息的结果仅更新帮助数据库管理员访问特定表空间或索引的状态的目录表列。

不更新目录表

收集统计信息的结果不更新任何目录统计信息。仅当“报告统计信息”单选按钮为是时才可用。

9. 在为非索引列采样的行的百分比字段中，指定收集非索引列统计信息时要采样的行的百分比。有效值为 1% 至 100%。缺省值为 25%。若您的表空间是 LOB 表空间，则不要使用此字段。
10. 正收集其统计信息的表中的列列示在**可用的列**列表中。单击一列，然后单击 **>** 将该列从**可用的列**列表移动至**选择的列**列表。通过单击 **>>**，可移动表中所有列。单击 **<** 或 **<<** 将列从**选择的列**列表移动至**可用的列**列表。**选择的列**列表中的所有列均为包含收集的统计信息的列。在一个表空间统计信息操作中，最多可以选择 10 列。
11. 在**表空间**字段中，以大写字母输入表空间的名称。如果表空间不在数据库 DSNDB04 中，必须用包含该表空间的数据库的名称来限定该表空间的名称。例如 SAMPLE.EMPLOYEE。
12. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141 页的『定义处理选项』。
13. 单击**确定**保存更改并关闭步骤笔记本。

在“数据仓库中心”中使用 MQSeries

通过“数据仓库中心”，您可以从 MQSeries 消息队列中访问数据就象从 DB2 数据库视图访问数据一样。提供了一个向导帮助您创建 DB2 表函数和 DB2 视图，可通过它们来访问数据。每个 MQSeries 消息都用为定界字符串来对待，根据您的规范对它进行语法分析，然后返回一个结果行。此外，那些是 XML 文档的 MQSeries 消息可作为仓库源来访问。使用“数据库仓库中心”，可以从 MQSeries 消息队列和“DB2 XML Extender 文档访问定义”(DAD) 文件中导入元数据。

创建 MQSeries 消息的视图

创建视图，以从 MQSeries 消息队列访问数据。具有以下需求和限制。

要求

- DB2 通用数据库版本 7.2
- DB2 仓库管理器版本 7.2
- MQSeries 支持。

要获取关于 MQSeries 需求的更多信息，参见“发行说明”中的主题“MQSeries”。

限制

编目仓库源数据库时，数据库别名是在仓库代理站点上编目的。但创建 MQSeries 和 XML 视图时，“数据仓库中心”假设数据库别名也是在客户机工作站上定义的，并将尝试使用仓库源数据库用户标识和密码与该工作站连接。如果连接成功，向导会打开，您可以创建视图。如果连接不成功，会显示警告消息，则您必须在向导中编目或选择不同的数据库别名。

参见“发行说明”中的“SQL 引用”部分，了解 MQ 消息的最大长度。

为 MQSeries 消息创建视图

要为 MQSeries 消息创建视图：

1. 从“数据仓库中心”窗口中，展开仓库源树。
2. 展开要包含视图的那个仓库源。
3. 右键单击视图文件夹，然后单击为 **MQSeries 消息创建...**。

MQSeries 向导打开。完成向导时，会在“数据仓库中心”中创建一个新视图。选择该视图时，将访问 MQSeries 队列，并根据向导中的规范对每条消息作为定界字符串来进行语法分析。

导入 MQSeries 消息和 XML 元数据

如果正在导入 MQSeries 和 XML 元数据，有以下需求和限制。

需求

- DB2 通用数据库版本 7.2
- DB2 XML Extender 版本 7.2
- MQSeries 支持

要获取关于 MQSeries 需求的更多信息，参见“发行说明”中的主题“MQSeries”。参见用户定义的函数的安装部分，以了解设置仓库源的信息。

限制

如果目标表存在并具有主键和外键，导入进程将失败。必须手工删除“数据仓库中心”中这些键的定义，才能导入 MQSeries 和 XML 元数据。

导入 MQSeries 消息和 XML 元数据

要将 MQSeries 元数据导入“数据仓库中心”：

1. 准备仓库目标数据库：
 - 定义仓库目标，然后注册并启用变换器。

- 为 DB2 XML Extender 启用仓库目标。参见“DB2 XML Extender 版本 7.2 发行说明”以获取更多信息。
 - 创建“XML Extender 数据访问定义”(DAD)文件,告诉“数据仓库中心”如何将 XML 文档的内容映射到仓库表。使用 DAD 文件对数据库启用 XML 集合。参见“DB2 XML Extender 版本 7.2 发行说明”以获取更多信息。
2. 右键单击**仓库**,然后单击**导入元数据** → **MQSeries**以打开**导入元数据**窗口。
 3. 在 **AMI 服务**字段中,输入将消息发送至或从其检索消息的服务点。
 4. 在 **AMI 策略**字段中,输入消息传递系统将用于执行操作的策略。
 5. 在 **DAD 文件**字段中,输入 DB2 XML Extender DAD 文件的名称,或通过单击省略号 (...) 搜索要选择的文件。此文件必须是本地的。
 6. 在**仓库目标**字段中,选择将运行步骤的仓库目标的名称。必须已定义了仓库目标。
 7. 在**模式**字段中,输入模式名称,这种模式用于限定在没有限定符的 DAD 文件中的表名。缺省模式是先前选择的仓库目标的登录用户标识。
 8. 选择**目标选项**:
 - 如果希望步骤在运行时替换目标表的内容,可单击**替换表内容**。
 - 如果希望步骤在运行时附加到目标表内容后,可单击**附加表内容**。
 9. 单击**确定**。
 - “导入元数据”窗口关闭。

导入操作完成后,以下仓库对象将添加到“仓库”树。

- 名称为 MQSeries 和 XML 的主题区
- 名称为 MQSeries 和 XML 的过程
- 名称为 MQSeries 和 XML 的用户定义的程序组
- DAD 文件中描述的所有仓库目标表的定义
- 步骤 <ServiceName>.<DAD file base name>.<Warehouse target Name>
- 程序模板 <ServiceName>.<DAD file base name>

如果仓库目标代理站点是远程的,必须更改步骤参数:

1. 右键单击该步骤,然后单击**特性**以打开该步骤的“特性”笔记本。
2. 单击**参数**选项卡。
3. 更改 DAD 文件参数的名称为远程仓库目标代理站点上相同 DAD 文件的名称。
4. 确保“处理选项”页上的**代理站点**字段中有正确的代理站点。

使用 MQSeries 用户定义的程序

MQSeries 和 XML 存储过程称为 **MQXMLXF**，包括在“DB2 数据仓库中心版本 7.2 的 Windows NT 版和 UNIX 版”中。在导入 MQSeries 和 XML 元数据时创建的步骤将运行存储过程。在表35中描述了它的参数。

表 35. MQXMLXF 存储过程的参数

参数	值
MQSeries ServiceName	将消息发送至或从其检索消息的服务点的名称。
MQSeries PolicyName	消息传递系统用于执行操作的策略的名称。
DAD file name	DB2 XML Extender DAD 文件的名称
TargetTableList	由逗号分隔的步骤的目标表列表
Option	REPLACE 或 APPEND
RUN ID	步骤修订版本号（用于记录日志）

注：此表中所有参数的数据类型都是 CHARACTER。

如果选项参数有一个值 REPLACE，存储过程将从目标表中删除所有行。存储过程还调用 DB2 XML Extender 存储过程以对所有现有的 MQSeries 消息填充目标表。

错误返回码

步骤运行时，存储过程可能会返回错误码 SQLCODE -443 和 SQLSTATE 38600。要诊断该错误，参见表36可获取可能的诊断文本。

表 36. 错误码的诊断文本是由 MQXMLXF 用户定义的程序返回的。

错误号	描述
AMIRC=xxxxx;<log file name>	xxxxx 是来自 AMI 层的返回码。请参阅 MQSeries 文档以获取更多详细信息。<log file name> 表示日志文件的位置。
XMLRC=xxxxx;<log file name>	xxxxx 是来自 DB2 XML Extender 的返回码。请参阅 DB2 XML Extender 文档以获取返回码的描述。<log file name> 表示日志文件的位置。
SQLCODE=xxxxx;<log file name>	xxxxx 是在执行 SQL 请求时返回的非零 SQLCODE。<log file name> 表示日志文件的位置。

参见日志文件以获取关于所有错误的更多信息。

错误日志文件

“数据仓库中心”在 MQXMLXF 运行时将所有诊断信息存储在日志文件中。日志文件的名称是 mqxf< nnnnnnnn>.log，其中 < nnnnnnnn> 是发送到存储过程的运行

维护仓库数据库

标识。“数据仓库中心”将在由 `VWS_LOGGING` 环境变量指定的目录中创建该文件。如果未定义环境，日志文件将创建在临时目录中。

在 Unix 系统上，要使 `VWS_LOGGING` 环境变量对存储过程可视，可在 `db2start` 命令之前使用 `db2set` 命令，将 `VWS_LOGGING` 添加到 `DB2ENVLIST` 环境变量中。以下示例显示了环境命令。

```
db2set DB2ENVLIST="AMT_DATA_PATH VWS_LOGGING"
```

如果步骤运行成功，将删除日志文件。

第13章 用用户定义程序扩展“数据仓库中心”

可以使用用户定义程序来使用最佳数据仓库软件以满足您的需要，同时提供单个控制点来管理仓库。“数据仓库中心”将在预定时间启动您定义为用户定义程序的应用程序。

例如，若您有想在仓库表上使用的数据净化程序，可以将该数据净化程序定义为用户定义程序，并对该程序运行一个步骤，该步骤在填充仓库表的步骤之后启动。

本章说明下列任务：

- 『定义用户定义程序』
- 第287页的『使用用户定义程序步骤』
- 第289页的『编写自己的程序以配合“数据仓库中心”使用』

定义用户定义程序

用户定义程序是您创建的程序或由供应商创建的程序。对“数据仓库中心”定义用户定义程序，以便一个或多个步骤可以使用该程序进行处理。若正定义存储过程，必须确保该存储过程在目标数据库中存在。若正定义任何其他类型的程序，则在“数据仓库中心”定义程序之前，必须将该程序安装在一个或多个“数据仓库中心”代理站点上。

要获取关于为 OS/390 或 AS/400 设置存储过程的信息，请参阅《仓库管理器安装指南》。

在对“数据仓库中心”定义用户定义程序之后，该程序定义可在“过程模型”窗口中作为步骤使用。

定义程序组

用户程序组是包含相关用户定义程序的逻辑组。在可对“数据仓库中心”定义用户定义程序之前，必须创建一个用户程序组。

要定义用户程序组：

1. 从“数据仓库中心”中，展开**管理**文件夹。
2. 展开**程序和变换器**文件夹。
3. 右键单击**用户定义程序和变换器**文件夹，并单击**定义组**。

用用户定义程序扩展“数据仓库中心”

4. 指定程序组的信息：可将信息输入这些字段中的任何一个。
 - 在**名称**字段中，为步骤输入一个新名称，或可保留“数据仓库中心”自动为步骤提供的名称。
 - 可选：在**管理员**字段中，输入负责维护该步骤的人员的姓名。
 - 可选：在**描述**字段中，为代理定义输入业务描述。描述最多可以是 254 个字符。
 - 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。
5. 单击**确定**保存更改并关闭笔记本。

指定要使用的程序

要对“数据仓库中心”定义用户定义程序：

1. 从“数据仓库中心”中，展开**管理**文件夹。
2. 展开**程序和变换器**文件夹。
3. 展开**用户定义程序和变换器**文件夹。
4. 若正定义一个新程序，则右键单击程序组，然后单击**定义程序**。
5. 若正编辑现存程序定义，则展开程序组文件夹。然后，右键单击该程序，并单击**特性**。
6. 指定程序的信息：
 - 在**名称**字段中，为步骤输入一个新名称。
 - 在**管理员**字段中，输入负责维护此程序的人员的姓名。
 - 在**描述**字段中，为您的程序输入一个业务描述。描述最多可以是 254 个字符。
 - 在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。
 - 在**程序类型**字段中，选择正定义的程序的类型。
 - 在**程序名字段**中，输入程序的全限定名。
 - 若程序类型为动态链接库，则在**函数名字段**中输入函数名称。

- 若程序类型为 DB2 存储过程，则在**存储过程名**字段中输入模式和过程名称。使用如下语法：

```
schema.procedurename
```

指定要使用的代理站点

在“程序”笔记本的“代理站点”页上，必须选择安装该程序的代理站点。

若定义代理站点时指定了用户标识和密码，则该程序将作为用户过程来运行。若未指定用户标识和密码，则无论如何定义仓库代理，该程序都将运行。

在同一工作站上，可以将一些程序作为用户过程来运行，而将其他程序作为系统过程来运行。为此，在该工作站上定义两个代理站点：一个代理站点具有用户标识和密码，而另一个没有。

若正定义变换器，必须确保与此程序相关的仓库代理与变换器所在的数据库相关。若给此程序定义指定一个仓库代理，该代理与在此页上指定的目标数据库不相关，则在创建使用此程序的步骤时将收到警告。

例如，具有下列仓库代理、源和目标：

代理	源	目标
代理 A	源 1	目标 1
		目标 2

假设“代理 A”与“源 1”和“目标 1”相关。现在，假设变换器存在于“目标 2”上。若在变换器程序定义中选择“代理 A”，则在尝试创建步骤时将收到警告。因为“代理 A”与包含该变换器的目标不相关，所以发出此警告。

要使用“代理 A”，必须使它与“目标 2”相关，或必须将该变换器存储在“目标 1”上，并将“目标 1”添加至目标资源定义中：

1. 根据正定义的程序的类型，单击**代理站点**选项卡或**目标**选项卡。
2. 若正定义程序，则从**可用的代理站点**列表中，选择安装该程序的代理站点，并单击**添加**。该代理站点的名称移动至**选择的代理站点**列表。对安装该程序的每个代理站点重复此步骤。
若正定义变换器，则从**可用目标**列表中，选择存储过程所在的目标并单击**添加**。该目标的名称移动至**选择的目标**列表。
3. 单击**确定**保存更改并关闭笔记本。

有关定义代理站点的详情，参见第13页的『定义代理站点』和联机帮助。

指定参数

您可对某些参数使用预定义记号。“数据仓库中心”在运行时替换记号的值。例如，步骤的目标资源的数据库名有一个记号为 &TDB。若在参数列表中包括此记号，则“数据仓库中心”提供在仓库目标的笔记本中定义的数据库名称，该仓库目标包含与步骤链接的目标表。根据哪个步骤使用该程序，记号允许您更改传送的值。

若想要用户定义程序从与使用该程序的步骤链接的源中抽取数据，可以在参数列表中包括 &STMT 记号。此记号引导“数据仓库中心”将为步骤定义的 SELECT 语句传送给该程序。SELECT 语句包括在双引号中作为一个命令行参数传送给该程序。任何嵌入引号包括在转义字符中。若 SELECT 语句特别长，它可能与最大参数数目的操作系统准则冲突。

有关带预定义记号的参数的完整列表，参见联机帮助中的『选择参数字段和控件』。

要将系统参数或系统记号添加至程序定义：

1. 打开“定义用户程序”笔记本。
2. 单击**参数**选项卡。
3. 单击**添加**。
4. 在**类型**区中，单击**系统参数**或**系统记号**。
5. 在**可用参数**字段中，选择参数或记号。在进行选择时，可通过按 **CTRL** 键选择多个参数或记号。
6. 单击**确定**。

选择的参数或记号列示在“定义用户程序”笔记本的“参数”页上。

示例：您正在 Windows NT 工作站上定义一个用户定义程序，它按固定的时间间隔检查文件。您打算使用此程序查找一个文件，另一步骤将把该文件装入数据库。

该程序使用下列参数：

- 文件名
- 轮询时间间隔
- 超时时间间隔

使用“仓库”目标文件名系统参数 (&TTBN) 来代表该文件名。为轮询时间间隔和超时时间间隔定义您自己的参数。

使用用户定义程序步骤

在定义用户定义程序之后，定义使用该程序进行处理的步骤。

当定义运行用户定义程序的步骤时，可以更改为该程序定义的参数值。若更改该程序的参数值，这些更改只影响正在步骤中使用的程序的实例。这些更改不会影响原始程序定义。

要永久更改用户定义程序的参数值，必须编辑程序定义。

若用户定义程序使用源或目标的记号，必须将此步骤与该源或目标链接。

要定义运行用户定义程序的步骤：

1. 定义一个过程。
2. 右键单击该过程。单击**打开**。“过程模型”窗口打开。
3. 在“过程模型”窗口的左边窗格中单击用户定义程序步骤图符。要查看步骤图符的描述，单击**查看** → **图注**。
4. 选择步骤子类型（若适用的话）。然后，将鼠标移动至“过程模型”窗口中并单击。
5. 将步骤与任何需要的源或目标链接。
6. 右键单击步骤，然后单击**特性**。

要在“过程模型”窗口中编辑现存步骤，单击一个过程，然后单击**打开**。

“过程模型”窗口打开。右键单击步骤，然后单击**特性**。

要在“过程模型”窗口之外编辑现存步骤，在仓库目录树视图中单击过程，包含在该过程中的对象列表就显示在右边框架中。右键单击步骤，然后单击**特性**。

7. 指定步骤的信息：
 - 在**名称**字段中，输入步骤的新名称，或可保留“数据仓库中心”自动为该步骤提供的名称。
 - 可选：在**管理员**字段中，输入负责维护此步骤的人员的姓名。此字段是可选的。
 - 可选：在**描述**字段中，输入步骤的业务描述。描述最多可以是 254 个字符。此字段是可选的。
 - 可选：在**注释**字段中，输入可能对访问此步骤的用户有帮助的详情。此字段是可选的。
8. 在“参数”页上，查看或编辑程序使用的参数和参数值。要编辑参数值，双击参数的**参数值**字段，然后输入新的参数值。若更改参数值，更改只影响此步骤运行的程序的实例。不会更改原始程序参数值。

用用户定义程序扩展“数据仓库中心”

9. 在“处理选项”页上，在**代理站点**列表中，选择想要运行您的步骤的代理站点。此列表中的选项是与源表、目标表和变换器或正定义的程序相关的代理站点。
10. 若想要选择随时运行步骤，选择**按需运行**复选框。步骤必须处于测试或生产方式，才能运行它。
11. 可选：若由外部填充步骤，则选择**外部填充**复选框，意指以不同于“数据仓库中心”的方法启动它。为将方式更改为生产，步骤需要以其他任何方式在“数据仓库中心”中运行。
若未选择**外部填充**复选框，为将方式更改为生产，必须调度步骤与是另一步骤的输入的瞬时表链接，或由另一程序启动该步骤。
12. 若需要重试步骤，在**重试**区中，指定想要该步骤再次运行多少次，以及想要在步骤下一次运行前经过的时间。
13. 单击**确定**保存更改并关闭步骤笔记本。

示例：定义使用您在前述章节中定义的用户定义程序的步骤。该步骤没有源。因为您正使用将作为序列中下一个步骤的源的文件，所以将该文件定义为此步骤的目标。然后定义将该文件用作源的装入步骤。装入步骤将该文件装入数据库。

测试使用用户定义程序的步骤

在定义使用用户定义程序的步骤之后，将该步骤提升为测试方式并运行它。

要将步骤提升为测试方式：

1. 打开“过程模型”窗口。
2. 右键单击要提升的步骤。
3. 单击**方式** → **测试**。

在提升该步骤之后，当右键单击**方式**时，**测试**已选择。

要测试步骤：

1. 打开“过程模型”窗口。
2. 右键单击步骤。
3. 单击**测试**。
4. 单击**确定**。

该步骤显示在“工作进度”窗口中，状态为**正在填充**。该处理应成功完成。但是，若处理失败，有关调试什么出错的信息，参见第143页的『测试步骤』。

在测试步骤之后，可以与其他任何步骤同样的方法调度它。有关调度步骤的详情，参见第146页的『调度仓库过程』。

编写自己的程序以配合“数据仓库中心”使用

可用支持下列程序类型之一的任何语言来编写程序：可执行文件、批处理程序、动态链接库或存储过程。

若用户定义程序的程序类型为可执行文件、批处理程序、批处理命令文件或动态链接库，它必须驻留在代理站点上。仓库代理将在预定时间启动该程序。在 Windows NT 上，缺省情况下，代理将作为系统过程来运行，意指该程序不能访问需要用户标识的资源或程序。另外，该程序需要访问的任何环境变量必须是系统变量。

若使用 Object REXX Windows 版来编写用户定义程序，则完成下列过程，以使这些程序可以在 Windows NT 下运行：

1. 将仓库代理或服务器服务定义为可以与 Windows NT 桌面交互的系统过程：
 - a. 从**服务**列表中选择代理或服务器服务。
 - b. 单击**启动**。
 - c. 单击**系统帐户**。
 - d. 选择**允许服务与桌面交互**复选框。
2. 在仓库代理或服务启动该程序之前，初始化 Object REXX 环境。可以通过从命令行运行任何 Object REXX 程序来初始化该环境。
3. 若 Object REXX 程序发出 DB2 CONNECT 语句，则验证该语句包括用户标识和密码，如以下示例所示：

```
DB2 CONNECT TO testdb USER db2admin USING db2pass
```

若用户定义程序的程序类型为存储过程，则该程序必须驻留在与运行该程序的步骤相关的仓库数据库中。

传送参数

在运行期间，“数据仓库中心”生成一个命令行参数列表，它作为输入传递给用户定义程序。尽可能在步骤中使用您的程序前，从命令行测试它。

示例：在 Windows NT 工作站上编写一个用户定义程序，该程序按固定的时间间隔检查文件。它使用下列参数：

- 文件名
- 轮询时间间隔
- 超时时间间隔

用用户定义程序扩展“数据仓库中心”

该程序获得如图20中所示的参数:

```
char *remoteFile    = argv[1];
int  interval = 5; // check every 5 minutes
int  timeout = 60; // time-out period in minutes

if( (argc==3) || (argc==4) )
{
    if( NumStrValidation( argv[2] ) ) // get polling interval
    {
        interval = atoi( argv[2] );
        if( interval == 0 )
            interval = 5; // used the original one
    }
    else
    {
        trace << "Invalid value for polling interval\n";
        flush(trace);
        return PARAMETER_ERROR;
    }

    if( argc==4 ) // get timeout period
    {
        if( NumStrValidation( argv[3] ) )
        {
            timeout= atoi( argv[3] );
            if( timeout == 0 )
                timeout = 480; // used the original one
        }
        else
        {
            trace << "Invalid value for timeout period\n";
            flush(trace);
            return PARAMETER_ERROR;
        }
    }
}
}
```

图 20. 从命令行读取参数

返回状态信息

在程序运行之后，它必须返回一个返回码给使用该程序的步骤。返回码必须是正整数。若程序不返回返回码，则使用该程序的步骤可能失败。当**错误 RC1** 的值为 8410 时，“数据仓库中心”在“日志详细信息”窗口的**错误 RC2** 字段中显示返回码。若**错误 RC2** 的值是 0，则该程序已成功运行，没有错误。

程序可能返回附加的状态信息给“数据仓库中心”：

- 另一个返回码，它可以与用户定义程序返回的代码相同或不同。

- 一个警告标志，它指示 SQL 返回了警告码，或指示用户定义程序在源表中未找到数据。当设置了此标志时，使用此程序的步骤将在“工作进度”窗口中具有警告状态。
- 一条消息，“数据仓库中心”将在“日志查看器详细信息”窗口的系统消息字段中显示它。
- 用户定义程序已处理的数据行数，“数据仓库中心”将在步骤的“日志查看器详细信息”窗口中显示它。
- 用户定义程序已处理的数据字节数，“数据仓库中心”将在步骤的“日志查看器详细信息”窗口中显示它。
- SQLSTATE 返回码，“数据仓库中心”将在“日志查看器详细信息”窗口的“SQL 状态”字段中显示它。

仓库代理将附加状态信息传送给仓库服务器。

将信息传送给“数据仓库中心”

要将附加的状态信息传送给仓库代理，您的程序必须创建一个称作反馈文件的文件，它包含附加的状态信息。反馈文件的路径和文件名是在 VWP_LOG 环境变量中指定的，该值是仓库代理在调用仓库程序之前设置的。您的程序必须使用恰当的系统调用读取此变量的值，以创建反馈文件。用户定义程序运行完之后，仓库代理将检查是否存在反馈文件。若存在反馈文件，则代理将处理该文件。否则，仓库代理将不做任何处理。若用户定义程序不能创建该文件，它应继续运行。

反馈文件的格式

用户定义程序可将附加状态信息按任何次序写入反馈文件中，但必须使用下列格式来标识信息。在以下列表中，将返回的每个项目放在开始标记 `<tag>` 和结束标记 `</tag>` 中。每个开始标记的后面必须有其结束标记；在一行中不能包括两个开始标记。例如：以下标记格式有效：

```
<RC>...</RC>...<MSG>...</MSG>
```

以下嵌入的标记格式无效：

```
<RC>...<MSG>...</RC>...</MSG>
```

在反馈文件中可以指定以下信息：

返回码 `<RC>return code</RC>`，其中 *return code* 是正整数。

返回码警告标志

`<WARNING>1</WARNING>` 将返回码警告标志设置为“开”。若此标志设置为“开”，则“工作进度”窗口中的步骤的处理状态将为“警告”。

用用户定义程序扩展“数据仓库中心”

“数据仓库中心”系统消息

<MSG>*message text*\n</MSG>

message text

一条或多条消息的文本

\n 换行符。若有多条消息，则在每条消息的末尾包括此字符。

注释 <COMMENT>*comment text*</COMMENT>，其中 *comment text* 是注释的文本。

已处理的数据行数

<ROWS>*number of rows*</ROWS>，其中 *number of rows* 是任何正整数。

已处理的字节数

<BYTES>*number of bytes*</BYTES>，其中 *number of bytes* 是任何正整数。

SQLSTATE

<SQLSTATE>*sqlstate string*</SQLSTATE>，其中 *sqlstate string* 是长度大于 0 且小于或等于 5 位的任何字符串。

图21显示反馈文件的示例。

```
<RC> 20</RC>
<ROWS>2345</ROWS>
<MSG>The parameter type is not correct</MSG>
<COMMENT> Please supply the correct parameter type (PASSWORD
    NOTREQUIRED, GETPASSWORD, ENTERPASSWORD)</COMMENT>
<BYTES> 123456</BYTES>
<WARNING> 1</WARNING>
<SQLSTATE>12345</SQLSTATE>
```

图 21. 反馈文件的示例

反馈如何确定步骤状态

日志查看器中显示的用户定义程序的返回码和步骤状态，因该程序设置的下列值而各异：

- 用户定义程序返回的返回码的值
- 是否存在反馈文件
- 反馈文件中的返回码的值
- 警告标志是否设置为开

表37列示这些值的可能组合以及它们所产生的结果。

表 37. 反馈文件条件和结果

条件				结果	
				步骤状态 ¹	“错误 RC1”和“错误 RC2”的值
用户定义程序返回码为 0	不存在反馈文件 ²			成功	RC1 = 0; RC2 = 0
	存在反馈文件 ²	反馈文件中 <RC> 的值为 0 ³	反馈文件中没有设置 <WARNING>	成功	RC1 = 0; RC2 = 0
			反馈文件中 <WARNING> 的值为 1	警告	RC1 = 0; RC2 = 0
		反馈文件中 <RC> 的值为 non-0 ³	反馈文件中没有设置 <WARNING>	失败	RC1 = 8410 (用户定义程序失败); RC2 = 反馈文件中 <RC> 的值
			反馈文件中 <WARNING> 的值为 1	警告	RC1 = 0; RC2 = 反馈文件中 <RC> 的值

用用户定义程序扩展“数据仓库中心”

表 37. 反馈文件条件和结果 (续)

条件				结果	
				步骤状态 ¹	“错误 RC1”和“错误 RC2”的值
用户定义程序返回码为非零	不存在反馈文件 ²			失败	RC1 = 8410 (用户定义程序失败); RC2 = 用户定义程序返回的代码
	存在反馈文件 ²	反馈文件中 <RC> 的值为 0 ³	反馈文件中没有设置 <WARNING>	成功	RC1 = 0; RC2 = 0
			反馈文件中 <WARNING> 的值为 1	警告	RC1 = 0; RC2 = 0
		反馈文件中 <RC> 的值为非零	反馈文件中没有设置 <WARNING>	失败	RC1 = 8410 (用户定义程序失败); RC2 = 用户定义程序返回的代码
			反馈文件中 <WARNING> 的值为 1	警告	RC1 = 0; RC2 = 反馈文件中 <RC> 的值

注:

1. 步骤处理状态，如“工作进度”窗口中所示。
2. “数据仓库中心”检查是否存在反馈文件，不管用户定义程序的返回码为零或者非零。
3. 反馈文件中 <RC> 的值总是显示为“日志详细信息”窗口中 **RC2** 字段的值。

第14章 维护“数据仓库中心”

本章说明与维护“数据仓库中心”相关的概念和过程。本章包括下列主题：

- 『备份“数据仓库中心”』
- 『扩充仓库』
- 第296页的『导出和导入“数据仓库中心”元数据』
- 第301页的『发布“数据仓库中心”元数据』
- 第306页的『更改活动的仓库控制数据库』
- 第307页的『配置“数据仓库中心”』

备份“数据仓库中心”

定期备份您的仓库控制数据库，以及所有仓库数据库。

要备份仓库控制数据库，使用 DB2 备份和恢复的标准过程。

在备份控制数据库之前，需要停止“数据仓库中心”。要关闭“数据仓库中心”：

1. 在包含仓库服务器的工作站上，单击**设置** → **控制面板** → **服务**。“服务”窗口打开。
2. 选择仓库服务，并单击**停止**。
3. 对“服务”窗口中列示的每个仓库服务重复此步骤。
仓库日志记录器与服务器是相链接的；若停止仓库日志记录器，服务器也会停止。然而，若停止仓库服务器，日志记录器却不会停止。
4. 当停止了所有仓库服务时，单击**关闭**。

要备份仓库数据库，使用数据库管理器提供的备份和恢复过程。

扩充仓库

使用“数据仓库中心”，可随着仓库的增长来扩充网络配置。可添加装有新仓库代理和管理接口的新工作站，或从现存配置中删除代理和管理接口。

还可创建仓库控制数据库。若创建仓库控制数据库，必须再次运行初始化过程。有关详情，参见第306页的『更改活动的仓库控制数据库』。

维护“数据仓库中心”

要将管理接口或仓库代理添加到您的“数据仓库中心”配置中，遵循那些组件的安装过程。要从“数据仓库中心”配置中删除管理接口或仓库代理，除去那些组件。

有关安装和除去“数据仓库中心”组件的详情，参见您的操作系统的《DB2 通用数据库速入门》和《DB2 仓库管理器安装指南》。

导出和导入“数据仓库中心”元数据

可使用“数据仓库中心”的导出和导入能力，在下列操作系统中导出和导入对象定义：

- Windows NT
- AIX
- Solaris 操作环境

导出过程和导入过程将使用大量的系统资源。在导出和导入对象定义时，可能想限制其他程序的使用。

因为导入和导出格式与发行版有关，所以不能使用从先前发行版导出的文件，从“数据仓库中心”的一个发行版迁移到另一个发行版。若想要迁移“数据仓库中心”，参见您的操作系统的《DB2 通用数据库速入门》。

或者，若想要为仓库控制数据库制作副本（如想要分隔测试和生产系统时），可以使用“DB2 通用数据库”导出和装入实用程序来复制该数据。然而，两个数据库都必须具有一个与缺省仓库用户匹配的用户标识和密码，以使用户可以登录到“数据仓库中心”。

将元数据导出到另一个“数据仓库中心”

当将元数据导出到标记语言文件时，“数据仓库中心”查找想要导出的对象，并生成标记语言语句来表示这些对象。然后，它将标记语言语句放入文件中，可将这些文件导入另一个“数据仓库中心”中。

计划导出元数据

若导出一个过程，可能导出大量的元数据。会随过程导出下列对象：

- 过程内包含的所有步骤。不能导出单个步骤。
- 该过程的代理站点。
- 与过程中的步骤相关的任何用户定义程序。
- “数据仓库中心”所需的内部对象，如 SQL 步骤的列映射。
- 过程和步骤的调度信息。可指定不导出此信息。

- 过程和底层步骤使用的任何仓库源或仓库目标定义。可选择不导出仓库源表定义。
- 其输出被正在导出的过程使用的任何其他过程和步骤。
- 所选仓库源或仓库目标中的所有表定义（若导出仓库源或仓库目标）。
- 程序组内的所有程序定义（若导出程序组）。
- 仓库模式内的所有表定义（若导出仓库模式）。

将元数据导出到标记语言文件

要导出标记语言文件:

1. 从“数据仓库中心”中右键单击“仓库”节点，然后单击**导出元数据** → **交换文件**。

“导出元数据”窗口打开。

2. 指定想要用于导出该元数据的标记语言文件。可:

- 使用**文件名**字段中显示的缺省标记语言文件名。

在 Windows NT 系统上，会将该文件写入启动“DB2 控制中心”所用的路径和目录。若您对该目录没有写访问权（例如，若“控制中心”安装在网络驱动器上），会将该文件写入 Windows NT 环境变量指定的主路径和目录，通常为 C 驱动器和根目录。

在 AIX 系统上，会将该文件写入调用“数据仓库中心”所用的路径和目录。若对该目录没有写访问权，会将该文件写入主目录。

- 输入路径和文件名。必须包括正确的文件扩展名。
- 单击 ... 打开一个窗口，可使用该窗口来查看可用的文件并选择一个。会用导出的元数据来覆盖您选择的任何现有文件。

3. 在**可用对象**列表中，选择要导出的对象并单击 >。该对象移动至**选择的对象**列表。对要导出的所有对象重复此步骤。

若要导出一个目录树中的所有可用对象，单击 >>。这些对象移动至**选择的对象**列表。若该目录树上的一个节点不包含对象，则不会移动它。

4. 若不想导出与导出的步骤相关的仓库源、表和列的所有定义，则清除**导出从属源特性**复选框。

5. 若不想导出与正在导出的过程相关的调度信息，则清除**包括调度**复选框。

6. 单击**确定**。

“导出信息”窗口打开并显示一条消息，告知该导出过程是否成功完成。若发生错误，“导出元数据”窗口会保持打开，以便可再试该导出过程。可在《*IBM DB2 通用数据库消息参考*》中查找有关特定错误消息的详情。

维护“数据仓库中心”

可在单个导出过程期间创建多个文件。例如，当导出 BLOB 数据的元数据定义时，会创建多个标记语言文件。导出过程中创建的第一个文件带有扩展名 .tag。若创建了多个文件，则为每个附属文件生成具有与该标记语言文件相同的文件名，并带一个数字扩展名。

例如，若指定的标记语言文件名为 e:\tag\steps.tag，这些附属标记语言文件命名为 e:\tag\steps.1、e:\tag\steps.2，依此类推。该文件扩展名仅用于在基本标记语言文件内标识附属文件，因此可将这些文件移到另一个目录。但是，不应重命名这些文件。必须总是将这些文件保存在同一个目录中，否则将无法成功导入这些文件。

导入元数据

可以导入对象定义，以便在您的“数据仓库中心”系统中使用。可能想要将样本数据导入仓库，或在为新仓库创建原型时导入数据。可导入以下元数据类型：

- 交换文件
- Erwin
- i2
- MQSeries
- Trillium

计划导入对象

当导入标记语言文件时，它包含的元数据存储于仓库控制数据库中。“数据仓库中心”将接受不是由另一“数据仓库中心”系统生成的标记语言文件。例如，可以导入由 ETI 产品生成的标记语言文件。

将忽略非“数据仓库中心”标记语言文件中任何具有未知 OBJTYPE 的对象，而处理继续进行。若对象属于已知的 OBJTYPE，但不包含“数据仓库中心”所需的所有特性，将会出错。

初始化控制数据库： 若正在使用导入实用程序来建立新的“数据仓库中心”，必须在目标系统中初始化新的仓库控制数据库。完成此任务之后，才可以导入任意数目的标记语言文件。

有关如何初始化新控制数据库的详情，参见第306页的『更改活动的仓库控制数据库』。

从一个“数据仓库中心”系统导入到另一个系统： 在完成导出过程和标记语言文件创建之后，就可将该文件导入目标系统。

在导入该文件之前，要考虑下列事宜：

- 命名考虑

若使用导入实用程序来将仓库源从测试环境移至生产环境，要确保生产环境中尚无具有相同仓库源名的仓库源，除非您想要覆盖该仓库源的定义。

若将步骤导入到包含同名步骤的系统中，必须删除想要覆盖的步骤，或将其更改为开发方式。否则，无法更新该步骤，并且会产生错误。

- 安排对象的导入次序

使用逻辑次序来导入对象。必须在目的地仓库控制数据库中，定义被引用但未在标记语言文件中定义的对象，为此，可先导入在单独标记语言文件中引用的对象定义，也可在“数据仓库中心”目录树视图中定义该对象。

例如，可以先导入仓库源定义。当该仓库源定义可供使用时，可以单独地导入主题区、仓库目标定义和任何对应的步骤。

- 若将标记语言文件从一个系统导入到另一个系统，必须随它移动所有相关的文件，且它们必须驻留在同一个目录中。有关在导出过程中可创建的多个文件的详情，参见第297页的『将元数据导出到标记语言文件』。

不要使用导入功能从先前版本的“数据仓库中心”迁移。必须始终使用相同版本的“数据仓库中心”导入和导出标记语言文件。有关从先前版本的“数据仓库中心”迁移的信息，参见您的操作系统的《DB2 通用数据库速入门》。

将标记语言文件导入 数据仓库中心

要将标记语言文件导入 数据仓库中心:

1. 从“数据仓库中心”中右键单击“仓库”节点，然后单击**交换文件**。
“导入元数据”窗口打开。
2. 在**文件名**字段中，输入想要导入的标记语言文件的名称。若想要搜索要导入的文件，单击...。“文件”窗口打开。
3. 选择想要导入的标记语言文件，然后单击**确定**。选择的文件名显示在“导入元数据”窗口的**文件名**字段中。
4. 在**检查点标识**字段中，输入记录在日志文件中的最后一个检查点标识（如果先前有导入指定的标记文件失败的话）。如果先前没有导入标记文件失败，则将检查点标识字段设置为 0。
5. 单击**确定**。该标记语言文件表示的元数据被导入到仓库控制数据库。

在导入过程成功之后，您会接收到带 0 返回码或警告返回码的消息，同时“导入元数据”窗口关闭。当发生错误时，不会有任何更改落实到仓库控制数据库。刷新树视图以显示您导入的对象。

除了导入标记语言文件外，还可以导入 ERwin、MQSeries、Trillium 和 i2 元数据。

维护“数据仓库中心”

- 关于导入 ERwin 元数据的信息，参见第250页的『将标记语言文件导入“数据仓库中心”』。
- 关于导入 MQSeries 元数据的信息，参见第279页的『导入 MQSeries 消息和 XML 元数据』。
- 关于导入 Trillium 元数据的信息，参见第239页的『导入 Trillium 元数据』。
- 关于导入 i2 元数据的信息，参见第399页的『导入 i2 源的元数据』。

导入错误： 若导入过程中发生错误，会接收到带错误返回码的错误消息，且“导入元数据”窗口保持打开。

常见错误包括：

- 导入从属于所用仓库中不存在的另一个对象的对象。
- 导入该仓库中已存在且有其他对象从属于它的对象。
- 导入已存在且不处于开发方式的步骤。
- 导入未正确导出的标记语言文件。

检查日志文件以确定发生错误时正导入什么对象或关系。若信息中未列示该对象，则它应该是列示在该日志文件末尾处的返回码和错误消息之前的最后一个对象。修正该错误，然后再次导入对象定义。若无法确定或修正该错误，与“IBM 软件支持”或提供该标记语言文件的供应商联系。

导入后考虑事项

在成功导入之后，会接收到带返回码为 0 的消息，同时“导入元数据”窗口关闭。

导入元数据时，所有对象都指定到缺省安全性组。除非您属于对该对象具有访问权的仓库组，否则导入的主题区、仓库源或仓库目标将不会在目录树视图中显示。与导出的仓库源或仓库目标相关的仓库组必须与使用那些对象的过程相关；某些仓库组可能不会被导出。若未导入希望仓库源和仓库目标与之相关的仓库组，需要将这些仓库组添加到“数据仓库中心”系统。不会导出个别用户的元数据。必须将这些用户添加到该“数据仓库中心”。

“数据仓库中心”将开发方式赋予所有导入的步骤，以便在将它们提升为测试或生产方式之前，可以验证它们是完整的和准确的。

在使用导入的对象之前： 若有必要，应更新仓库源和仓库目标的数据库信息（数据库名、用户标识和密码）。若导入了任何新代理站点定义，还需要更新它们的密码。若导入任何具有密码类型参数的程序，则必须复位那些值。

将您的用户标识添加到对导入的过程和仓库目标具有访问权的仓库组。

在运行导入的步骤之前：

1. 确保“数据仓库中心”系统中定义和配置了导入的步骤引用的所有仓库源和仓库目标数据库。验证导入的所有定义，并更改任何参数以与新环境匹配。

在导入包括含有连接的步骤的标记语言文件之后，可能注意到未将这些表按您的期望置于“过程模型”或“仓库模式模型”中。不能将某些信息从一个系统导入到另一个系统。将需要使用“过程模型”或“仓库模式模型”窗口来重新放置步骤和星型模式。

2. 按需要提升步骤。

再次导入标记语言文件若再次导入现存的步骤，它必须处于开发方式，否则将发生错误。在再次导入步骤之前，必须将它们降为开发方式。

大多数标记语言文件都可多次导入而不出错。若标记语言文件包括主键和外键，则再次导入该文件之前必须删除这些键。

若表或文件中列的名称或次序已更改（例如，重新定义了仓库源或仓库目标表），则在再次导入同一个表或步骤之前，要从目标系统中删除仓库源表、文件或步骤定义（包括仓库目标中的目标表）。至少，要验证已按您的期望导入了列或字段信息。

发布“数据仓库中心”元数据

发布元数据是将元数据从“数据仓库中心”传送到信息目录管理器的过程。信息目录是信息目录管理器管理的表集合，该集合包含公司元数据，用于帮助用户在该组织中标识和查找可供他们使用的数据和信息。当将元数据发布到该信息目录时，为最终用户和系统分析员提供了查看仓库中的内容的快速途径。

从“数据仓库中心”中，可发布：

- 描述“数据仓库中心”对象的元数据。
- 描述DB2 OLAP 服务器或 Hyperion Essbase 服务器对象的元数据。

本章只讨论发布“数据仓库中心”元数据。有关发布 OLAP 服务器元数据的信息，参见信息目录管理器管理指南。

要发布元数据，必须安装带有信息目录管理器的“DB2 仓库管理器”程序包。信息目录管理器运行在 Windows 32 位操作系统上。有关信息目录可驻留的“DB2 通用数据库”平台的信息，参见信息目录管理器管理指南。

计划发布元数据

在发布元数据之前，需要确保已建立正确的环境。

1. 确保在正确的工作站上安装和配置了所需的仓库组件：

维护“数据仓库中心”

若仓库服务器站点和“数据仓库中心”管理接口组件位于不同的工作站，则二者之上都必须安装“信息目录管理器”管理员功能。

“数据仓库中心”管理接口必须具有与信息目录 API 的 DB2 联网。

2. 确保“信息目录管理器”和“数据仓库中心”的管理员用户标识具有 Windows NT 管理员的特权。

在 AIX 和 Solaris 操作环境中，安装过程将语言设置为发布信息目录时使用的语言，并导出至 OLAP 集成服务器。如果您希望使用这些函数的语言不是安装期间设置的语言，可在一行上输入以下命令来创建以下软链接：

在 AIX 上

```
/usr/bin/ln -sf /usr/lpp/db2_07_01/msg/locale/flgnxolv.str  
/usr/lpp/db2_07_01/bin/flgnxolv.str
```

locale 格式为 *xx_yy* 的本地语言环境的名称

在 Solaris 操作环境中

```
/usr/bin/ln -sf /opt/IBMDB2/V7.1/msg/locale/flgnxolv.str  
/opt/IBMDB2/V7.1/bin/flgnxolv.str
```

locale 格式为 *xx_yy* 的本地语言环境的名称

选择要发布的元数据

表38提供“数据仓库中心”和信息目录中对象类型之间的映射。当将元数据导出到信息目录时，“数据仓库中心”会使用此映射。有关对象类型和对象类型特性的映射的详细信息，参见第329页的『附录B. 元数据映射』。

表 38. “数据仓库中心”和信息目录对象类型之间的映射

“数据仓库中心”对象类型	信息目录对象类型
过程	DWC 过程
步骤	转换（在表或列级别）
列或字段	列或字段
仓库源或仓库目标	数据库，IMS 数据库定义
主题区	业务主题区
表或文件	关系表和视图、文件、IMS 段
仓库模式	星型模式

当发布过程时，下列对象会随之一起发布：

- 过程内包含的所有步骤。不可发布单个步骤。
- 与过程中的步骤相关的任何用户定义程序。

- 过程和步骤的调度信息。
- 过程和底层步骤使用的任何仓库源定义或仓库目标定义。
- 其输出被正在发布的过程使用的任何其他过程和步骤。

通过选择“发布数据仓库中心元数据”笔记本中的**限制目录树中的对象级别**选项，可限制发布的底层过程的数目。

- 仓库模式内的所有表定义（若发布仓库模式）。
- 会发布所选仓库源或仓库目标中的所有表定义。

要求：在发布元数据之前，要确保没有作为管理员登录到该“信息目录管理器”。

要选择想要发布的元数据：

1. 在“数据仓库中心”中，右键单击“仓库”节点，然后单击**发布元数据**。“发布元数据”窗口打开。
2. 右键单击列表中的空白处，然后单击**“数据仓库中心”元数据**。“发布数据仓库中心元数据”笔记本打开。
3. 在**可用对象**列表中，选择想要发布的对象并单击 **>**。
该对象移至**选择的对象**列表。对想要发布的所有对象重复此步骤。
若想要发布所有可用对象，单击 **>>**。若该目录树上的一个节点不包含对象，则不会移动它。
4. 单击**同步设置**选项卡。
5. 在**目录名**字段中，输入想要向其中发布元数据的信息目录的名称。
6. 在**管理员用户标识**字段中，输入对该信息目录具有访问权的用户标识。该用户标识可以是信息目录管理员的标识，或具有某些管理特权的信息目录用户的标识。
7. 在**管理员密码**字段中，输入对应于该用户标识的密码。
8. 在**验证密码**字段中再次输入该密码。
9. 从**将源表映射到目标仓库**列表中选择其中一个选项，以指定在信息目录中发布过程及其源表时想要如何映射它们：
 - 单击**处于表级别**，以在信息目录源表和目标表之间生成信息目录转换对象。在“信息目录管理器”中，当打开数据库对象的“目录树”视图时，您会看到目标表、转换对象和源表。缺省情况下已选择此选项。
 - 单击**处于列级别**，以在信息目录源列和目标列之间生成信息目录转换对象。当打开数据库对象的“目录树”视图时，您会看到目标表、目标列、转换对象、以及源表和源列。
可在列级别映射过程及其源表，以便获取该数据更为详细的视图。
10. 可选：选择下列其中一个复选框以确定将如何更新发布的元数据：

- 选择**更新时在信息目录生成名称和描述**，以使用“数据仓库中心”中同一个对象的名称和描述来置换现存对象名称和描述。若希望保留信息目录中的现存对象和描述，不要选择此复选框。例如，若为信息目录中的对象指定了附加的业务信息，且不想用更新的值覆盖该信息，则不要选择此复选框。若选择**更新前删除信息目录中的现存对象**，则此复选框不可用。
 - 选择**更新前删除信息目录中的现存对象**，以便在从“数据仓库中心”发布更新的对象之前删除现存对象。若想要保留和更新信息目录中的现存对象，不要选择此复选框。若选择此复选框，“信息目录管理器”会删除所有与删除的对象相关的对象，如数据库和表对象。该目录树视图会更新，这样就不会有与其他对象不相关的对象。
11. 可选：可通过限制“目录树”视图中发布和显示的对象级别数，使“信息目录管理器”中显示的目录树结构稍简略一些。
 - a. 选择**限制目录树中的对象级别数**，以限制在“目录树”视图中发布和显示的对象级别数。若不选择此复选框，则会发布和显示所选对象下的所有对象。若指定了限制，则仅发布和显示所选对象下的对象子集。
 - b. 单击向上或向下箭头指定在目录树视图中发布和显示的对象级别数的限制。缺省限制值是 1。指定的限制越低，目录树结构就越简略。

若设置了限制，在目录树视图中发布和显示的对象级别数限制为指定的级别数。
 12. 单击**确定**。

“发布信息”窗口打开并显示一条消息，告知该发布过程是否成功完成。“发布信息”窗口显示日志文件的内容，该文件列示哪些元数据对象已成功发布。若发生错误，在关闭“发布信息”窗口时该笔记本保持打开，以便可以再试发布对象。可在《*IBM DB2 通用数据库消息参考*》中查找有关特定错误消息的详情。

若发布成功，在关闭“发布信息”窗口时该笔记本会关闭，且两个发布对象添加到“发布元数据”窗口中的发布列表。一个发布更新控制元数据，另一个则更新定义元数据。**控制元数据**是描述仓库中数据的更改的元数据。控制元数据的示例有步骤处理更新表的日期和时间。**定义元数据**是描述仓库中对象、数据源和应用用于该数据的转换的元数据。定义元数据的示例有列名、表名和数据库名。第一次发布元数据时，控制元数据和定义元数据都会在信息目录中注册。当更新元数据时，可以选择更新控制元数据或 / 和定义元数据类型。

更新发布的元数据

在发布元数据之后，可定期使用发布对象（发布过程期间创建的）以将“数据仓库中心”中所作的更改传播到信息目录。此过程称作**元数据同步**。

当使元数据同步时，在信息目录中注册的对象元数据会在您运行该发布时被更新，或者基于为该发布创建的调度表进行更新。在下列情况中，元数据不会在信息目录中更新：

- 在“数据仓库中心”中创建新对象时。若对“数据仓库中心”添加新对象，必须先发布该对象。更新过程只更新先前发布的对象。
- 当先前发布到信息目录的对象名称更改时

要求：若计划使元数据同步，必须将“数据仓库中心”仓库控制数据库用作信息目录数据库。

要更新发布的元数据：

1. 在“发布元数据”窗口的**发布列表**中，右键单击想要运行的发布。
2. 单击**运行**。

调度发布的元数据的定期更新

通过创建调度表，可以自动更新先前发布的元数据。

要调度信息目录中的元数据的定期更新：

1. 打开“发布元数据”窗口。
2. 在**发布列表**中，右键单击想要调度的发布，然后单击**调度**。“调度”笔记本的“调度”页打开，以便可以创建用于运行发布的调度表。

有关使用“调度”笔记本的信息，参见第146页的『调度仓库过程』。

查看发布更新的日志文件

可以查看日志文件，它们显示对发布更新的结果。这些文件位于 `VWS_LOGGING` 环境变量指定的目录中。对于 Windows NT，`VWS_LOGGING` 变量的缺省值是 `\sqllib\logging`，其中 `x` 是安装 DB2 通用数据库的驱动器。使用下列日志文件名：

元数据的类型	用途	日志文件名
“数据仓库中心”控制元数据的更新	描述仓库中对象的更改	ICMDWCD.OUT
“数据仓库中心”定义元数据的更新	描述仓库中的对象、数据源和应用用于该数据的转换。	ICMXCHG.OUT

将删除的对象传播到信息目录

当删除“数据仓库中心”中的对象时，有关该删除对象的信息存储在仓库控制数据库中。在同步过程期间更新该元数据时，“数据仓库中心”会在将其他更改导入到该信息目录之前，将这些删除传播到信息目录。当元数据同步成功完成时，“数据仓库中心”会除去仓库控制数据库中的这些项。因为“数据仓库中心”除去这些项，所以仅可将删除传播到一个信息目录。若需要对另一个信息目录进行删除，必须人工删除那些项目。

若更改先前发布到该信息目录中的仓库对象的名称，则必须再次发布该对象以更新信息目录。使用旧名称的对象不会被覆盖，这样元数据同步之后两个对象同时存在于信息目录中；必须人工删除旧对象。

更改活动的仓库控制数据库

安装仓库服务器时，安装期间所指定的仓库控制数据库会进行初始化。在典型安装期间，将创建并初始化一个称为 DWCTRLDB 的缺省控制数据库。初始化是一个过程，在该过程中“数据仓库中心”创建存储“数据仓库中心”元数据所需的控制表。要确定活动的控制数据库的名称，可单击“数据仓库中心 登录”窗口中的高级。

要使用活动的控制数据库以外的控制数据库，可使用“仓库控制数据库管理”工具来切换数据库。“仓库控制数据库管理”工具将要使用的数据库注册为活动的仓库控制数据库。在使用“仓库控制数据库管理”工具前，必须停止仓库服务器。

如果数据库不是已存在于工作站上，“数据仓库中心”将创建您在仓库服务器工作站上指定的数据库。若想要使用远程数据库，则在远程系统上创建该数据库并在仓库服务器工作站上编目它。

“DB2 控制中心”或“命令行处理器”可能会指出仓库控制数据库处于不一致状态。此消息是预料中的，因为仓库服务器未向仓库记录器落实它的初始启动消息。

要初始化新的仓库控制数据库：

1. 停止仓库服务器。参见第7页的『启动仓库服务器和日志记录器』以获取如何启动和停止仓库服务器的信息。
2. 单击**开始** → **程序** → **IBM DB2** → **仓库控制数据库管理**以打开“仓库控制数据库管理”窗口。
3. 在**新控制数据库**字段中，输入要用作仓库控制数据库的数据库的名称。
4. 在**模式**字段中，输入用于该数据库的模式名称。

5. 在**用户标识**字段中，输入对数据库有访问权的用户标识。
6. 在**密码**字段中，输入该用户标识的密码。
7. 在**验证密码**字段中，再次输入该密码。
8. 单击**确定**。
该窗口保持打开。**消息**字段会显示消息，指示创建和迁移过程的状态。
9. 该过程完成之后，单击**取消**以关闭该窗口。
10. 重新启动仓库服务器。
11. 启动“数据仓库中心”管理界面。
12. 在“数据仓库中心登录”窗口中单击**高级**，然后在**控制数据库**字段中输入新仓库控制数据库的名称。
13. 单击**确定**以保存更改。

在安装期间初始化仓库控制数据库

要初始化在安装期间创建的缺省控制数据库以外的仓库控制数据库，必须执行定制安装并更改“定义本地仓库控制数据库”窗口中的名称。安装过程将用指定的名称创建该数据库，初始化与“数据仓库中心”一起使用的数据库，并将该数据库注册为活动的仓库控制数据库。

注：如果应用修订包或安装 DB2 或 DWC 的新发行版，则必须迁移现有的控制数据库以更新它包含的对象。

要在安装期间在工作站上创建仓库控制数据库（该工作站不是安装了仓库服务器的那个工作站），可在定制安装期间选择**仓库本地控制数据库**。安装过程将创建数据库。安装后，必须遵循步骤 306，在仓库服务器工作站上使用“仓库控制数据库管理”窗口。指定在安装期间指定的数据库名。将初始化数据库与数据仓库中心一起使用，并注册为活动的仓库控制数据库。

配置“数据仓库中心”

可以使用“数据仓库中心特性”笔记本来更改用于“数据仓库中心”安装的全局设置。可覆盖使用这些全局设置的对象中的多个全局设置。例如，可使用“特性”笔记本来指定当仓库代理在源表或文件中找不到任何行时，正在处理的步骤的缺省行为。可以覆盖特定的步骤中的这个全局设置。

仅当“数据仓库中心”服务器安装在工作站（以及管理客户机）上时，才可使用配置工具。

警告：“特性”笔记本中的某些字段应仅根据“IBM 软件支持”的建议来更改。有关详情，参见“数据仓库中心特性”笔记本的联机帮助。

维护“数据仓库中心”

更改“数据仓库中心”特性

使用“数据仓库中心特性”笔记本来更改“数据仓库中心”配置参数。例如，可更改缺省通知消息的参数以适应特定的业务情况。

必须具有管理员特权，才能更改“数据仓库中心”特性。

要更改“数据仓库中心”特性:

打开“数据仓库中心特性”笔记本:

1. 右键单击**仓库**对象，并从弹出菜单中选择**特性**。

2. 更改下列任何设置:

- 在“跟踪级别”页上:
 - 控制数据库
 - ODBC
 - 服务器
 - 代理
 - 记录器
- 在“服务器”页上:
 - 代理启动 / 停止超时
 - 当达到总记录数时清除日志
 - 重新启动中断的作业
- 在“处理选项”页上:
 - 重试
 - 不返回任何行
 - SQL 警告
- 在“通知消息”页上:
 - 情况
 - 发送方
 - 主题
 - 系统参数
 - 消息
 - 缺省邮件服务器

邮件服务器必须支持 ESMTP。

3. 单击**确定**以保存更改。

有关该笔记本中的设置的详细描述，参见“数据仓库中心特性”笔记本的联机帮助。

也可使用“数据仓库中心特性”笔记本来运行“数据仓库中心”组件跟踪。有关详情，参见第326页的『组件跟踪数据』。

第15章 在“数据仓库中心”中创建星型模式

本章描述如何使用“数据仓库中心”来创建星型模式。可使用 DB2 OLAP Integration Server 中的星型模式来定义支持 OLAP 客户所需的多维立方。多维立方是定义多维数据库的数据和元数据的集合。

在使用数据来构建多维立方之前，应用净化后的数据来填充“数据仓库中心”中的星型模式。

OLAP 模型是一个逻辑结构，描述如何计划计量您的业务。该模型采用星型模式的形式。星型模式是一个专门的设计，它由多个维表和一个事实表组成，维表描述业务的各个方面，事实表包含关于该业务的事实。例如，若您有一个销售书的邮购业务，一些维表是客户、书、目录和财政年度。事实表包含有关在该财政年度期间每个用户从每个目录中订购的书籍的信息。在“数据仓库中心”中定义的星型模式称作仓库模式。

表39描述使用“数据仓库中心”和 DB2 OLAP Integration Server 创建仓库模式和用数据装入结果多维立方时涉及的任务。该表列示任务并告知使用哪个产品或组件来执行每个任务。本章描述每个任务。

表 39. 创建星型模式和填充多维立方的任务

任务	完成任务从	
	数据仓库中心	DB2 OLAP Integration Server
定义关系数据的仓库目标，将使用这些数据作为 OLAP 多维立方的源。	X	
定义仓库模式。	X	
将仓库模式导出到 DB2 OLAP Integration Server。该仓库模式将用作 DB2 OLAP Integration Server 中的模型。	X	
使用该模型（仓库模式）为模型定义层次结构。		X
使用该模型定义多维立方所基于的元轮廓。		X

表 39. 创建星型模式和填充多维立方的任务 (续)

任务	完成任务从	
	数据仓库中心	DB2 OLAP Integration Server
创建命令脚本，它将用于将数据装入多维立方。		X
创建运行该命令脚本的批处理文件。		X
从“数据仓库中心”中导出定义要运行的批处理文件的元数据。自动生成对象，它们使从“数据仓库中心”调度装入和测试多维立方成为可能。		X
调度由导出过程创建的步骤，以便能填充多维立方。	X	

在“数据仓库中心”中设计仓库模式

使用“仓库模式模型”窗口生成和存储与仓库相关的仓库模式模型。可以方便地将仓库模式模型作为元数据导出到 DB2 OLAP Integration Server (作为 OLAP 模型)。

定义仓库模式

在定义仓库模式之前，必须定义仓库目标表，将使用它们作为仓库模式的源表：

- 当定义将用于仓库模式的目标表时，为目标表选择 **OLAP 模式的部件** 复选框（在“定义仓库目标表”笔记本中），您计划把该目标表用作维表或事实表。
- **要求：** 当为仓库模式定义仓库目标时，仓库目标名必须与在其上定义该仓库目标的物理数据库名准确匹配。

有关定义仓库目标的详情，参见第101页的『第4章 设置对仓库的访问权』。

任何仓库用户可在模式中定义表，但仅属于对包含表的仓库目标具有访问权的仓库组的仓库用户可更改这些表。有关详情，参见第19页的『“数据仓库中心”安全性』。

要定义仓库模式：

1. 在“数据仓库中心”中，右键单击**仓库模式**文件夹。
2. 单击**定义**。

“定义仓库模式”笔记本打开。

3. 在**名称**字段中，输入模式名。
4. 可选：在**管理员**字段中，输入该仓库模式的联系人的姓名。
5. 可选：在**描述**字段中，输入该仓库模式的描述。若将仓库模式的元数据发布到信息目录，则此字段中的信息将用作该仓库模式的“简短描述”特性的值。
6. 可选：在**注释**字段中，输入想要添加的任何注释。若将仓库模式的元数据发布到信息目录，则此字段中的信息将用作该仓库模式的“详细描述”特性的值。
7. 可选：要仅显示一个数据库中的表，选择**仅使用一个数据库**复选框，然后从**仓库目标数据库**列表中选择数据库名。
仅可将那些由一个数据库中的表组成的仓库模式导出到 DB2 OLAP Integration Server。
8. 单击**确定**定义仓库模式。
新仓库模式添加到**仓库模式**文件夹下的目录树中。

在定义仓库模式之后，可通过打开“仓库模式建模器”窗口来更新它。

将表和视图添加至仓库模式

使用“添加数据”窗口将仓库目标表、源表或源视图添加至所选仓库模式。直到如第315页的『将仓库模式导出到 DB2 OLAP Integration Server』所述将仓库模式导出到 DB2 OLAP Integration Server 中，才可为帐户和时间维定义表。

要将维表和事实表添加至仓库模式：

1. 打开“添加数据”窗口：
 - a. 展开对象目录树直到找到**仓库模式**文件夹。
 - b. 右键单击仓库模式，然后单击**打开**。“仓库模式建模器”窗口打开。
 - c. 在选用板中单击**添加数据**图符，然后在画面上想要放置这些表的位置单击该图符。“添加数据”窗口打开。
2. 展开“仓库目标”目录树直到看见**表**文件夹下的表列表。
3. 要添加表，从**可用的表**列表中，选择想要包括在该仓库模式中的表，然后单击 **>**。**选择的表**列表中的所有表在“仓库模式建模器”画面上包含表图符。
单击 **>>** 将所有表移动到**选择的表**列表中。要从**选择的表**中除去表，单击 **<**。
要除去**选择的表**中的所有表，单击 **<<**。
4. 要创建新源表和目標表，在**可用的表**目录树中右键单击**表**文件夹，然后单击**定义**。“定义仓库目标表”或“定义仓库源表”窗口打开。
5. 单击**确定**。选择的表显示在窗口中。

自动连接表

使用“仓库模式模型”窗口自动连接表。若定义了仓库主键和外键，且单击**自动连接**，则会自动连接所选表。有关在目标表上定义键的详情，参见第111页的『定义仓库目标』。

1. 展开对象目录树直到找到**仓库模式**文件夹。
2. 右键单击仓库模式，然后单击**打开**。“仓库模式建模器”窗口打开。
3. 按住 **Ctrl** 键并单击每个表选择想要自动连接的表。

每个表的主键与作为该键的一部分的列的图符一起显示。使用**视图**菜单隐藏或显示表上的列。要创建链接，列必须在两个表上都可见。当您在组中选择表时，所有表不需要处于同样的状态。

4. 在工具栏上单击**自动连接**图符，或单击 **仓库模式** → **自动连接**。主键和外键关系以绿色显示。
5. 在工具栏上单击**保存图符**，或单击**仓库模式** → **保存**以保存您的工作。

在非键列之间添加连接关系

使用“仓库模式模型”窗口，在显示在画面上的表的非键列之间添加连接关系。可连接任意两列（不允许多列连接）。

1. 展开对象目录树直到找到**仓库模式**文件夹。
2. 右键单击仓库模式，然后单击**打开**。“仓库模式建模器”窗口打开。确保在该画面上至少定义了两个表。
3. 在工具栏上单击**连接**图符，然后按住鼠标按钮并单击第一个表中的列。
4. 将光标移至第二个表的列以建立连接关系。连接关系以黑色显示。
5. 在工具栏上单击**保存图符**，或单击**仓库模式** → **保存**将连接关系保存为仓库模式的一部分。

在定义仓库模式之后，第315页的图22会显示它的外观。

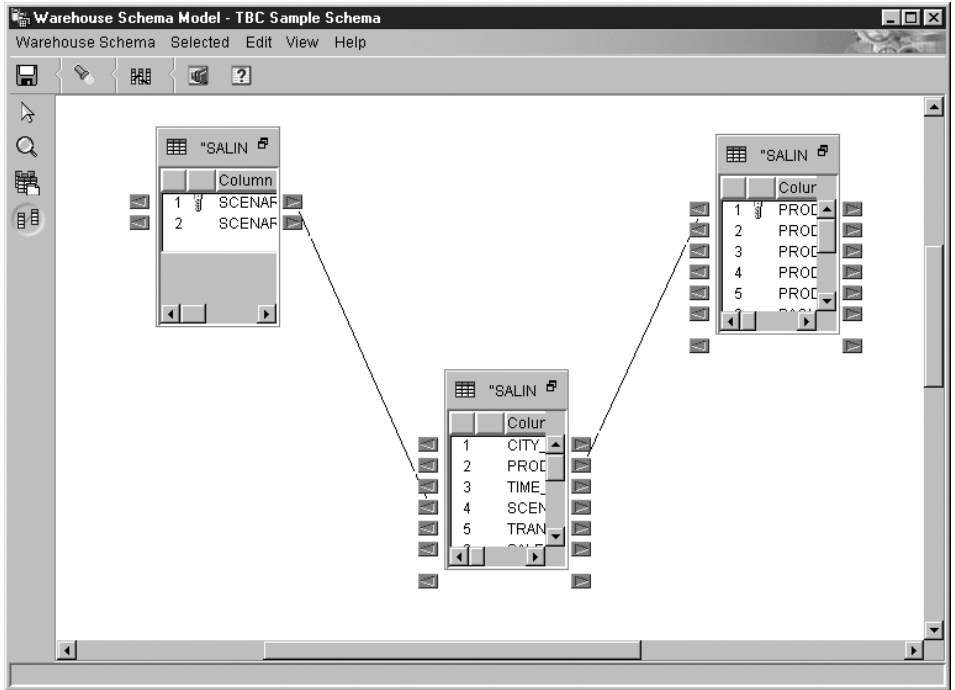


图 22. “仓库模式模型”窗口

将仓库模式导出到 DB2 OLAP Integration Server

使用“导出仓库模式”笔记本将仓库模式导出到 DB2 OLAP Integration Server（作为 OLAP 模型）。

要导出仓库模式：

1. 在“数据仓库中心”中，右键单击仓库节点并单击**导出元数据** → **OLAP Integration Server**。
“导出仓库模式”笔记本打开。
2. 从**可用列表**中，选择想要导出的仓库模式。仅那些用**仅使用一个数据库**选项定义的仓库模式可供导出。
3. 单击 **>**。
所选模式移动至**选择的列表**。
4. 单击**集成服务器**选项卡。

5. 在**目录名字段**中，输入 DB2 OLAP Integration Server 目录的名称，您想要将仓库模式导出到该目录。缺省名称是上次用来导出仓库模式的目录。
6. 在**目录表模式名字段**中，输入 DB2 OLAP Integration Server 目录使用的目录表模式的名称。
7. 可选：在**时间维表字段**中，选择“时间”维表的名称。若不选择名称，或列表中没有任何表，将使用“常规”维类型。若选择导出多个模式，则仅当所选模式具有公共的维表时，此字段才可用。

将模式存储在 DB2 OLAP Integration Server 目录中之后，不能更改维类型。

8. 可选：在**帐户维表字段**中，选择“帐户”维表的名称。若不选择名称，或列表中没有任何表，将使用“常规”维类型。若选择了多个模式，则仅当所选模式具有公共的维表时，此字段才可用。

将模式存储在 DB2 OLAP Integration Server 目录中之后，不能更改维类型。

9. 在**用户标识**字段中，输入将用来访问 DB2 OLAP Integration Server 目录的用户标识。
10. 在**密码**字段中，输入该用户标识的密码。
11. 在**验证密码**字段中，再次输入该密码。
12. 若 OLAP 模型已经存在于 DB2 OLAP Integration Server 目录中，则选择下列其中一个选项来指定想要采取的操作。

- 若 OLAP 模型名与仓库模式名相同，则单击**将新数据添加至现存 OLAP 模型**。

若选择此选项，则仅将仓库模式中的新信息添加至现存 OLAP 模型。不修改现存信息。例如，若选择此选项，则事实表保持不变且不能更改任何维。可添加新维。

- 若 OLAP 模型名与仓库模式名相同，则单击**显示消息且不执行导出**。

若选择此选项，则显示一条错误消息，并且不修改现存模型。缺省情况下已选择此选项。

- 若 OLAP 模型名与仓库模式名相同，则单击**覆盖现存模型**。

若选择此选项，则删除关于 OLAP 模型的所有现存元数据，并用仓库模式中的新元数据置换它们。

13. 单击**确定**将所选仓库模式导出到 DB2 OLAP Integration Server 目录。

笔记本关闭，并显示一个进度指示器直到导出完成。当已导出所有指定的仓库模式时，“导出信息”窗口打开，显示关于该导出的成功信息或失败的信息。单击**确定**关闭此窗口。

可查看存储关于该导出过程的跟踪信息的日志文件。该文件位于 `VWS_LOGGING` 环境变量指定的目录中。对于 Windows

NT, VWS_LOGGING 变量的缺省值是 \sqllib\logging, 其中 x 是安装 DB2 通用数据库的驱动器。日志文件的名称是 FLGNXHIS.LOG。

在 DB2 OLAP Integration Server 中使用仓库模式

将在“数据仓库中心”中设计的仓库模式导出之后, 使用 DB2 OLAP Integration Server 来完成多维立方的设计。

要查看导出的仓库模式, 可使用在“数据仓库中心”中使用的仓库模式名来打开 OLAP 模型(仓库模式)。确保指定仓库目标, 它已用来将仓库模式定义为模型的数据源。当您在 DB2 OLAP Integration Server 桌面上打开模型时, 图23会显示它的外观。会显示您在事实表和维表之间定义的连接关系。

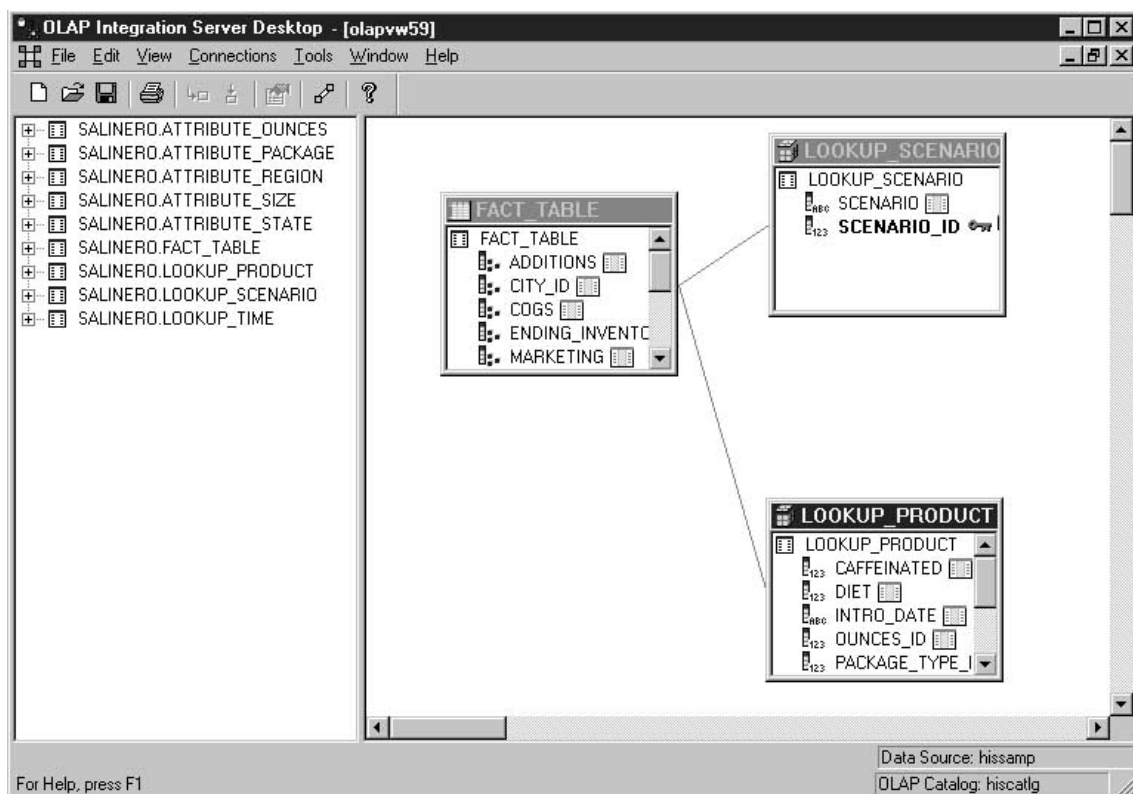


图 23. 仓库模式显示为 OLAP 模型

从 DB2 OLAP Integration Server 中, 必须完成下列任务:

1. 创建 OLAP 元轮廓，它是 OLAP 模型（仓库模式）的一个子集，您的多维立方将基于该模型。在该元轮廓中，定义 OLAP 模型各维中的层次结构。有关基于 OLAP 模型创建元轮廓的详情，参见 *OLAP Integration Server Model and User's Guide*。
2. 创建一个轮廓，它描述 Essbase 数据库所需的所有元素，您的多维立方定义在该数据库中。例如，轮廓将包含成员和维的定义、成员和公式。还将定义用来将数据装入该立方的脚本。然后将定义从中调用该脚本的批处理文件。有关详情，参见『在 DB2 OLAP Integration Server 中创建轮廓和为多维立方装入数据』。
3. 将定义该批处理文件的元数据导出到“数据仓库中心”，以便您可在常规基础上调度立方的装入。参见第320页的『将元数据导出到“数据仓库中心”』。

在 DB2 OLAP Integration Server 中创建轮廓和为多维立方装入数据

本节描述如何创建一个轮廓，并使它与将数据装入多维立方的脚本相关。在将数据装入轮廓之后，可通过电子表格程序（比如 Lotus® 1-2-3® 或 Microsoft Excel）访问结果立方，以便对数据进行分析。

有关窗口中的字段和控件的详情，参见 DB2 OLAP Integration Server 的联机帮助。

要从 DB2 OLAP Integration Server 桌面创建数据库轮廓：

1. 打开基于 OLAP 模型（仓库模式）创建的元轮廓。
2. 单击**轮廓** → **成员和数据装入**。“Essbase 应用程序和数据库”窗口打开。
3. 在**应用程序名字段**中，选择 OLAP 应用程序的名称，它将包含想将数据装入的 Essbase 数据库。也可输入一个名称。
4. 在**数据库名字段**中，输入想将数据装入的 OLAP 数据库的名称。
5. 在其余字段中输入其他任何选项，然后单击**下一步**。
6. 在“命令脚本”窗口中输入其他任何选项，然后单击**下一步**。
7. 在“调度 Essbase 装入”窗口中单击**现在**。
8. 单击**完成**。

OLAP 轮廓已创建。下一步，必须创建装入脚本。

要创建装入脚本：

1. 打开仓库模式的元轮廓。
2. 单击**轮廓** → **成员和数据装入**。“Essbase 应用程序和数据库”窗口打开。
3. 在**应用程序名字段**中，选择 OLAP 应用程序的名称，它将包含想将数据装入的数据库。也可输入一个名称。

4. 在**数据库名**字段中，输入想将数据装入的 OLAP 数据库的名称。
5. 在其余字段中输入其他任何选项并单击**下一步**。
6. 在“命令脚本”窗口中输入其他任何选项并单击**下一步**。
7. 在“调度 Essbase 装入”窗口中单击**仅保存装入脚本**。
8. 单击**保存脚本**。“将脚本另存为”窗口打开。
9. 为该命令脚本文件输入名称和文件扩展名。
10. 单击**完成**。

在 `..\IS\Batch\` 目录中创建新命令脚本，它将数据装入多维立方。该命令脚本包含下列项目：

- 包含立方的源数据的 DB2 数据库名称。
- 将存储立方的 Essbase 数据库。
- 将用于立方的 OLAP 目录名。
- 将数据装入立方的指示信息。
- 在定义脚本时指定的任何计算选项。

图24显示名为 `my_script.script` 的命令脚本的示例。LOADALL 项上的换行符不重要。可在一行上输入所有项。

```

LOGIN oisserv
SETSOURCE "DSN=dbc;UID=user;PWD=passwd;"
SETTARGET "DSN=essserv;UID=user;PWD=passwd"
SETCATALOG "DSN=TBC_MD;UID=user;PWD=passwd;"
LOADALL "APP=app1;DBN=db1;OTL=TBC Metaoutline;FLT_ID=1;OTL_CLEAR=N;
CALC_SCRIPT=#DEFAULT#;"
STATUS

```

图 24. OLAP 命令脚本: `my_script.script`

在创建轮廓和命令脚本之后，必须创建调用该脚本的批处理文件。该批处理文件用作“数据仓库中心”步骤的参数，它运行该脚本来装入多维立方。

要创建批处理文件，使用文本编辑器并输入命令以调用该脚本。可创建类似于图 25中的文件的文件来运行 `my_script.script`。不要在此示例中输入换行符。

```

"C:\IS\bin\olapicmd" < "C:\IS\Batch\my_script.script" >
"C:\IS\Batch\my_script.log"

```

图 25. 调用命令脚本的批处理文件: `my_script.bat`

my_script.log 日志文件显示关于导出到“数据仓库中心”的元数据的信息。它还显示该导出过程是否成功。

将元数据导出到“数据仓库中心”

使用“DB2 OLAP Integration Server 管理器”将批处理文件（它装入多维立方）的元数据导出到“数据仓库中心”。该导出过程在“数据仓库中心”中创建对象，这些对象使装入和测试该立方成为可能。

在导出元数据之前，确信已经如第312页的『在“数据仓库中心”中设计仓库模式』所述为仓库模式定义了表。

要将元数据导出到“数据仓库中心”，从 DB2 OLAP Integration Server 桌面开始：

1. 单击**工具** → **管理器**。“管理器”窗口打开。

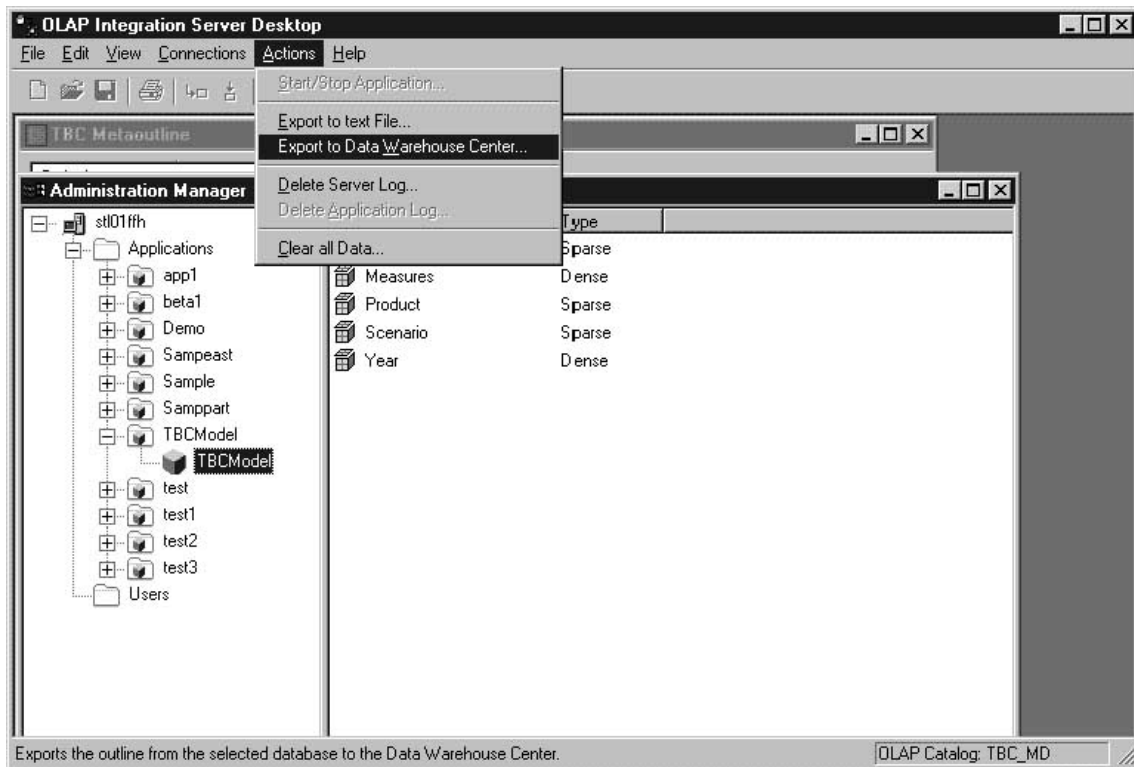


图 26. “管理器”窗口

2. 展开**应用程序**文件夹，然后右键单击想要导出其元数据的数据库。
3. 单击**操作** → **导出到“数据仓库中心”**。“导出到数据仓库中心”窗口打开。

4. 在 **OLAP 模型名字段**中，输入元轮廓所基于的模型（仓库模式）的名称。
5. 在**目录 DSN** 字段中，输入 DB2 OLAP Integration Server 中的目录数据库源名的名称，它包含元轮廓所基于的模型。
6. 在**用户名字段**中，输入用来与目录数据库连接的用户标识。
7. 在**密码**字段中，输入该用户标识的密码。
8. 在**表模式**字段中，输入对 DB2 OLAP Integration Server 目录中的表使用的表模式。
9. 在**控制数据库**字段中，输入仓库控制数据库的名称，您想将元数据导出到该数据库。
10. 在**用户名字段**（“数据仓库中心”的）中，输入用来与该仓库控制数据库连接的用户标识。
11. 在**密码**字段中，输入该用户标识的密码。
12. 在**表模式**字段中，输入对仓库模式的表使用的表模式。
13. 在**调用 OLAP Integration Server 脚本的批处理文件**字段中，输入批处理文件的全限定名，该文件运行命令脚本以装入多维立方。
例如，输入 `c:\is\batch\my_script.bat`。
14. 单击**确定**。

该批处理文件的元数据导出到“数据仓库中心”。

有关已导出到“数据仓库中心”的元数据的信息，参见日志文件。

有关已在“数据仓库中心”中创建的对象的信息，参见『从“数据仓库中心”中装入多维立方』。

从“数据仓库中心”中装入多维立方

当从 DB2 OLAP Integration Server 中导出元数据时，会创建下列“数据仓库中心”对象，且这些对象与仓库模式中的目标表相关：

- 名为 OLAP 立方的主题区
- 该主题区中按如下格式命名的过程：

`servername.applicationname.databasename.outlinename`

servername

OLAP 服务器的名称。

applicationname

OLAP 服务器应用程序的名称，它包含 *databasename* 标识的数据库。

databasename

OLAP 服务器数据库的名称，它包含 *outlinename* 标识的轮廓。

outlinename

已导出其元数据的 OLAP 服务器简要表的名称。

- 以与该过程相同的格式命名的步骤。

该步骤使用批处理文件（已导出其元数据）作为参数。当在该步骤的“特性”笔记本中单击**参数**选项卡时，**参数值**栏显示批处理程序的全限定名，该程序调用您在 DB2 OLAP Integration Server 中创建的命令脚本。**参数值**栏中的值与批处理文件匹配，该文件是您在“DB2 OLAP Integration Server 导出到数据仓库中心”窗口的**调用 OLAP Integration Server 脚本的批处理文件**字段中指定的。例如，“参数”值栏可能显示 `c:\is\batch\my_script.bat`。

当运行步骤时，该批处理文件会调用该脚本来装入多维立方。

当选择过程时，构成仓库模式的表显示在“数据仓库中心”的右窗格中。当运行步骤时，仓库模式表用作构建和填充多维立方的源表。维表用作 OLAP 模型的成员的源，而事实表用作量度（多维立方中的数据）的源。

有关 DB2 OLAP Integration Server 元数据如何映射至“数据仓库中心”元数据的详情，参见第339页的『DB2 OLAP Integration Server 和“数据仓库中心”之间的元数据映射』。

创建填充多维立方的调度表

可调度装入多维立方的步骤，并且可提升它以使它按常规运行。要调度和提升步骤：

1. 右键单击步骤，然后单击**调度**。“调度”笔记本打开。
2. 使用“调度”页定义想要该步骤运行的时间间隔。
3. 使用“调度”笔记本的“任务流程”页指定此步骤仅当另一个步骤成功完成时才运行，后者是定义为仓库模式创建仓库目标表（如第312页的『在“数据仓库中心”中设计仓库模式』所述）的步骤。有关调度步骤来基于另一个步骤的成功或失败来运行的详情，参见第146页的『调度仓库过程』。
4. 右键单击步骤，然后单击**方式**。可单击**测试**或**生产**来运行该步骤。

在成功地调度和测试步骤之后，就会填充使用您的仓库模式构建的多维立方。当正在填充多维立方时，第323页的图27会显示“工作进度”窗口。

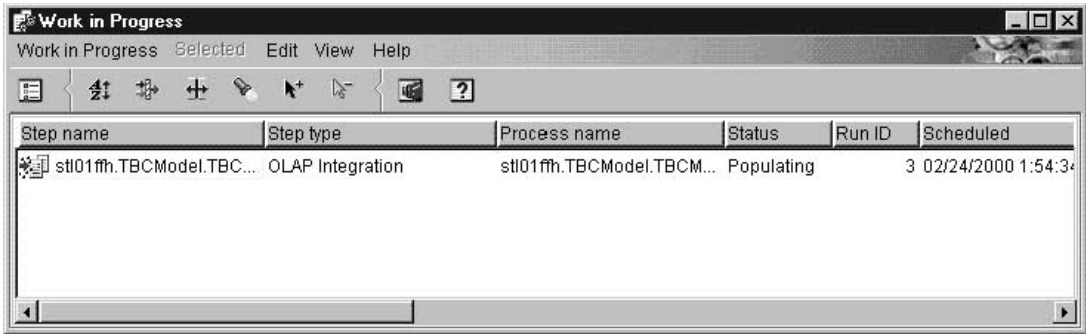


图 27. “工作进度” 窗口有一个正在填充立方的步骤

发布关于您的仓库模式的元数据

可使用“发布数据仓库中心元数据”笔记本将元数据发布至信息目录，该元数据描述您的仓库模式中的表。有关详情，参见第301页的『发布“数据仓库中心”元数据』。

仓库模式映射至信息目录管理器中的星型模式。使用命令行界面将 DB2 OLAP Integration Server 元数据发布至信息目录。当发布 DB2 OLAP Integration Server 元数据对象时，会在“多维数据库中的维”对象类型和信息目录中的表对象之间创建链接关系。若将仓库模式发布至信息目录并发布 DB2 OLAP Integration Server 元数据，您将拥有 OLAP 模型的元数据的完整描述。有关发布 DB2 OLAP Integration Server 元数据的详情，参见 *Information Catalog Manager Administration Guide*。

附录A. “数据仓库中心” 记录和跟踪数据

“数据仓库中心” 提供三个级别的记录能力:

- 基本记录功能
- 程序跟踪
- 启动错误跟踪

为了帮助您诊断问题，这些记录级别中的每一种都包含多种功能部件。本附录描述这些“数据仓库中心”记录级别和功能部件。

基本记录功能

“数据仓库中心” 中的基本记录功能捕捉在运行时和构建期发生的所有事件和错误。此信息存储在仓库控制数据库中。日志查看器显示这些运行时和构建期事件和错误，以及与它们相关的任何消息或返回码。

要查看运行时（步骤处理）错误:

1. 打开“数据仓库中心”桌面。
2. 单击 **数据仓库中心** → **工作进度**。

“工作进度”窗口打开。

3. 选择想要查看其错误的步骤。
4. 单击**记录**。

“日志查看器”窗口打开，并显示所选步骤的运行时错误。

要查看构建期（表导入、对象创建和步骤提升）错误:

1. 打开“工作进度”窗口。
2. 单击**工作进度** → **显示日志**。

“日志查看器”窗口打开并为“数据仓库中心”显示构建期错误。

有关日志查看器显示的返回码和错误消息的详细解释，参见《DB2 信息和原因码》。

日志保留记录，直到达到指定的计数限制。当达到计数限制时，“数据仓库中心”自动删除日志。

要更改日志的计数限制，参见“数据仓库中心特性”笔记本的联机帮助。

“数据仓库中心”记录和跟踪数据

注：建议：将日志记录计数的大小设置为可保留 3 至 4 天的记录。

不能关闭基本记录功能。

还可通过“Windows NT 应用程序事件”（可通过“事件查看器”应用程序进入）查看错误消息。

组件跟踪数据

在“IBM 软件支持”的指示信息下，运行“数据仓库中心”跟踪以产生程序执行的记录。可以在仓库控制数据库上运行运行 ODBC 跟踪和在仓库服务器、代理和日志记录器组件上运行跟踪。

运行跟踪时，“数据仓库中心”将信息写入文本文件。从步骤调用的“数据仓库中心”程序也将任何跟踪信息写入此目录。这些文件位于 *VWS_LOGGING* 环境变量指定的目录。

对于 AS/400 系统，许多“仓库中心”跟踪文件都存储在“iSeries 集成文件系统”中。要编辑这些跟踪文件，可以使用 FTP 将这些文件移动到工作站上，或使用 Client Access for iSeries。要获取更多信息，请参阅《仓库管理器安装指南》。

VWS_LOGGING 的缺省值为：

组件跟踪数据功能提供关于下列组件的信息：

Windows NT 和 OS/2

x:\program files\sqlib\logging

UNIX 和 OS/390

/var/IWH

AS/400

/QIBM/UserData/IWH

在 Windows NT 上，“数据仓库中心”写入下列文件：

AGNT $nnnn$.LOG

包含跟踪信息。其中 *nnnn* 是仓库代理的数字过程标识，它可以为 4 或 5 个字符，视操作系统而定。

AGNT $nnnn$.SET

包含代理的环境设置。其中 *nnnn* 是仓库代理的数字过程标识，它可以为 4 或 5 个字符，视操作系统而定。

IWH2LOG.LOG

包含对日志记录器组件跟踪的结果。

IWH2SERV.LOG

包含对仓库服务器跟踪的结果。

IWH2DDD.LOG

包含对仓库控制数据库跟踪的结果。

若运行的是 UNIX 代理，“数据仓库中心”在 UNIX 工作站上写入下列文件：

startup.log

包含有关仓库代理守护进程的启动的跟踪信息。

vwdaemon.log

包含有关仓库代理守护进程的处理的跟踪信息。

要运行“数据仓库中心”组件跟踪：

1. 右键单击仓库对象，并单击**特性**。
2. 根据“IBM 软件支持”的指示信息，指定仓库控制数据库、ODBC 连接、服务器、代理或日志记录器的跟踪级别。
3. 单击**确定**。
4. 根据请求重新启动这些服务。
5. 执行失败操作。
6. 重复步骤 1 至步骤 4，以将跟踪级别设置回 0。

仓库程序和变换器

所提供的 仓库程序和变换器对日志文件写入错误。

仓库程序

提供的仓库程序将数据写入 VWS_LOGGING 环境变量指定的目录。在发送日志文件到“IBM 软件支持”后，清除日志文件目录。

有关详情，参见特定 仓库程序的联机帮助。

变换器

可以记录当使用变换器时生成的错误。要启用记录，在“步骤”笔记本的“处理选项”页上指定日志表名，并给该日志表名添加后缀 *.n*。*n* 的值指示记录级别：

- | | |
|----------|----------------------------|
| 0 | 不记录 |
| 1 | 仅记录错误 |
| 2 | 记录错误和警告（这是缺省记录级别） |
| 3 | 记录错误、警告和资料性消息（例如，启动和停止变换器） |

“数据仓库中心”记录和跟踪数据

例如，要指示一个名为 MyLogTable 的记录表（它包含处于记录级别 3 或更低级别的记录项），指定 MyLogTable:3。

可以在记录表名后包括一个表空间名。若想这样做，将记录级别追加到表空间名。

例如，要指示一个名为 MyLogTable 的记录表（它位于 MyTableSpace 表空间中，并且包含处于记录级别 3 或更低级别的项），指定 MyLogTable,MyTableSpace:3。

Apply 程序

可以跟踪使用 Apply 程序时生成的错误。要对 Apply 程序启用跟踪，可在“仓库特性”页中设置“代理跟踪”值 = 4。“代理跟踪” = 4 时，“代理”对 Apply 打开完整跟踪。

如果在 CD 表中看不见任何数据，表示 Capture 程序尚未启动，或尚未通过更新源表来创建更改的数据。

启动错误跟踪文件

当日志记录器未运行时，“数据仓库中心”自动创建三个日志文件。日志文件名为 IWH2LOGC.LOG、IWH2LOG.LOG 和 IWH2SERV.LOG。“数据仓库中心”将这些文件存储在 VWS_LOGGING 环境变量指定的目录中。

这些日志为：

IWH2LOGC.LOG

当日志记录器不运行时，过程会将消息写入此文件。“数据仓库中心”服务器和 OLE 服务器对此文件进行写入。仅当日志记录器停止时，该文件才存在。该文件包含未能发送的所有消息的全部内容。

IWH2LOG.LOG

当日志记录器不能自行启动或当跟踪被激活时，日志记录器创建此文件。当日志记录器不能自行启动且不能写入至“数据仓库中心”日志时，将关键诊断信息写入此文件。在 DB2 未启动或文件系统已满的情况下，这很重要。若在日志记录器停止时听到五次嘟嘟声，应查看此文件。若日志记录器不能启动，则服务器不能启动。

IWH2SERV.LOG

服务器日志包含启动消息，并且当服务器跟踪打开时它将增大。

附录B. 元数据映射

本附录列示下列元数据的对象类型和对象类型特性:

- 信息目录管理器元数据到“数据仓库中心”元数据, 在下面的『信息目录管理器和“数据仓库中心”之间的元数据映射』中描述。
- 信息目录管理器元数据到 OLAP 服务器元数据, 在第338页的『信息目录管理器和 OLAP 服务器之间的元数据映射』中描述。
- “数据仓库中心”元数据到 DB2 OLAP Integration Server 元数据, 在第339页的『DB2 OLAP Integration Server 和“数据仓库中心”之间的元数据映射』中描述。

信息目录管理器和“数据仓库中心”之间的元数据映射

下表显示信息目录管理器和“数据仓库中心”之间每种对象类型的元数据映射。信息目录管理器列显示对象类型特性, 如同它们在对象的“描述”视图中显示的那样。“数据仓库中心”列显示对象特性的名称, 如同它们在各种对象特性笔记本中显示的那样。在某些情况下, “数据仓库中心”特性信息(如步骤的处理时间戳记)取自“工作进度”窗口。

表 40. 信息目录管理器和“数据仓库中心”的数据库对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	仓库源名或仓库目标名
简短描述	描述
详细描述	注释
访问数据所用的 URL	N/A
操作	N/A
数据库或子系统名	数据库名
数据库类型	此特性的值可为 RELATIONAL 或 FILE。 该映射是从仓库源或仓库目标类型派生出来的。
代理类型	N/A
数据库位置	N/A
数据库主机服务器名	系统名
系统代码页	此元数据是“数据仓库中心”内部的。 ¹

表 40. 信息目录管理器和“数据仓库中心”的数据库对象之间的元数据映射 (续)

数据库服务器和数据库扩展类型。	该映射是从仓库源或仓库目标类型派生出来的。 例如，若仓库目标为“DB2 通用数据库 Windows NT 版”数据库，则数据库服务器类型为“DB2 系列”。数据库扩展类型为 DB2 NT。
数据库所有者	N/A
上次更改源定义的时间戳记	数据库定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
数据库状态	N/A
数据库扩展类型	数据库子类型和数据库版本。 该映射是从仓库源或仓库目标类型派生出来的。例如，若仓库目标为“DB2 通用数据库 Windows NT 版”数据库，则数据库扩展类型为 DB2 NT。
创建源定义的时间戳记	N/A
有关详情	管理员
注:	
1. 若将特性指定为“数据仓库中心”的内部特性，则不会显示它。	

表 41. 信息目录管理器和“数据仓库中心”的 IMS DBD (数据库描述定义) 对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	仓库源名
简短描述	描述
详细描述	注释
操作	N/A
数据库上次刷新时间	N/A
有关详情	管理员
数据库所有者	N/A
数据库主机服务器名	系统名
数据库服务器类型	数据库类型和数据库版本。 该映射是从仓库源类型派生出来的。IMS DBD 的特性值是 IMS。

表 41. 信息目录管理器和“数据仓库中心”的 IMS DBD (数据库描述定义) 对象之间的元数据映射 (续)

数据库或子系统名	数据源名
数据库类型	此特性设置为 HIERARCHICAL。 该映射是从仓库源类型派生出来的。
数据库扩展类型	数据库子类型和数据库版本。 该映射是从仓库源类型派生出来的。IMS DBD 的特性值是 IMS。
数据库状态	N/A
IMS 访问方法	N/A
操作系统访问方法	N/A
共享索引名	N/A
访问数据所用的 URL	N/A
创建源定义的时间戳记	N/A
上次更改源定义的时间戳记	IMS DBD 的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
注:	
1. 若将特性指定为“数据仓库中心”的内部特性, 则不会显示它。	

表 42. 信息目录管理器和“数据仓库中心”的关系表或视图对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	表名
简短描述	描述
详细描述	N/A
访问数据所用的 URL	N/A
操作	N/A
目录注释	N/A
本地数据库别名	N/A
表数据上次刷新时间	运行和将表用作目标表的步骤的上次完成时间戳记。 此信息显示在“工作进度”窗口中。

表 42. 信息目录管理器和“数据仓库中心”的关系表或视图对象之间的元数据映射 (续)

转换程序类型	此特性的值为数据仓库中心。 此特性在“数据仓库中心”没有特定元数据。
数据库或子系统名	包含该表的仓库源或仓库目标数据库的数据库名
表所有者	表模式
表名	表名
上次更改源定义的时间戳记	表定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
基表所有者名称	N/A
基表名	N/A
转换程序运行方式	N/A
转换程序的上次运行时间	N/A
转换程序运行频率	N/A
部分或全部表复制 / 更新	N/A
复制 / 更新的数据状态一致	N/A
目录刷新 / 更新频率	N/A
转换程序的上次更改时间	N/A
转换程序编译时间	N/A
表类型	该映射是从包含该表的数据库的仓库源或仓库目标子类型派生出来的。 例如, 若仓库源或目标为“DB2 通用数据库 Windows NT 版”数据库, 则数据库表类型为 DB2 NT。
定义表示视图	N/A
表的内部 IDS 名	N/A
表用作维表	维表
数据库主机服务器名	包含表的仓库源或仓库目标数据库的系统名。
创建源定义的时间戳记	N/A
有关详情	包含该表的仓库源或仓库目标数据库的管理员。

表 42. 信息目录管理器和“数据仓库中心”的关系表或视图对象之间的元数据映射 (续)

注:

1. 若将特性指定为“数据仓库中心”的内部特性，则不会显示它。

表 43. 信息目录管理器和“数据仓库中心”的列或字段对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	列或字段名
简短描述	描述
详细描述	N/A
访问数据所用的 URL	N/A
操作	N/A
目录注释	N/A
列或字段的数据类型	数据类型
主键中列或字段的位置	N/A
列或字段的长度	长度或精度（取决于数据类型）
列或字段的刻度	刻度
列或字段可空	可空
列或字段位置	仓库源或仓库目标的表或文件笔记本中显示的列或字段在列表中的位置。
数据库或子系统名	包含该列的表所在的仓库源或仓库目标的数据库名
表所有者	包含该列的表的表模式
表名	包含该列的表名。
包含维	N/A
列或字段名	列名
文件名	包含该字段的文件的文件名（仅“数据仓库中心”文件）
列或字段距开始处的字节偏移	此字段在固定类型文件中的偏移。 此元数据是“数据仓库中心”内部的。 ¹
列或字段是否是键的一部分	N/A
列或字段是否是唯一键	N/A
数据是前映象、后映象还是计算的映象	N/A
用于填充列的源列 / 字段名或表达式	N/A

表 43. 信息目录管理器和“数据仓库中心”的列或字段对象之间的元数据映射 (续)

上次更改源定义的时间戳记	列定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
用于表示空值的字符串	N/A
日期的分辨率	N/A
列或字段的精度	N/A
数据是否是文本	是文本吗？ 此特性的值是 Y 或 N。
数据库主机服务器名	包含该列的表所在的数据库的系统名。
列或字段的上次刷新时间	N/A
创建源定义的时间戳记	N/A
有关详情	包含该列的表所在的数据库的管理员。
列次序	N/A
注:	
1. 若将特性指定为“数据仓库中心”的内部特性，则不会显示它。	

表 44. 信息目录管理器和“数据仓库中心”的文件对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	此特性的值是从文件名派生的。
简短描述	描述
详细描述	N/A
访问数据所用的 URL	N/A
操作	N/A
信息的上次刷新时间	运行和将该文件用作目标文件的步骤的上次完成时间戳记。
转换程序类型	此特性的值为数据仓库中心。 此特性在“数据仓库中心”没有特定元数据。
数据库主机服务器名	包含该文件的仓库源或仓库目标的系统名。
数据库或子系统名	包含该文件的仓库源或仓库目标的数据库名。
文件所有者	N/A
文件路径或目录	文件路径或目录的特性值是从文件名派生的。

表 44. 信息目录管理器和“数据仓库中心”的文件对象之间的元数据映射 (续)

文件名	该特性值是从文件名派生的。
文件类别或类型	文件类型
源定义的上次更改时间	文件定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
转换程序的上次运行时间	N/A
转换程序运行频率	N/A
部分或全部文件复制/更新	N/A
复制/更新的数据状态一致	N/A
转换程序的上次更改时间	N/A
转换程序的上次编译时间	N/A
创建源定义的时间戳记	N/A
有关详情	包含该文件的仓库源或仓库目标的管理员。
注:	
1. 若将特性指定为“数据仓库中心”的内部特性, 则不会显示它。	

表 45. 信息目录管理器和“数据仓库中心”的 IMS 段对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	表名
简短描述	描述
详细描述	N/A
访问数据所用的 URL	N/A
操作	N/A
数据库或子系统名	数据源名
段名	N/A
段最大长度	N/A
段最小长度	N/A
实际逻辑子段源	N/A
逻辑父代连接键源	N/A
转换程序的上次运行时间	N/A
转换程序运行频率	N/A
上次更改源定义的时间戳记	段定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
数据库主机服务器名	IMS 数据库定义 (DBD) 的系统名

表 45. 信息目录管理器和“数据仓库中心”的 IMS 段对象之间的元数据映射 (续)

段所有者	N/A
段的上次刷新时间	N/A
创建源定义的时间戳记	N/A
有关详情	包含该段的 IMS DBD 的管理员。
注:	
1. 若将特性指定为“数据仓库中心”的内部特性, 则不会显示它。	

表 46. 信息目录管理器和“数据仓库中心”的转换对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	步骤名
简短描述	描述
详细描述	N/A
访问数据所用的 URL	N/A
操作	N/A
转换标识符	转换的唯一标识符。 此元数据是“数据仓库中心”内部的。 ¹
转换程序名	程序名
转换类别或类型	程序类型
源列 / 字段名、表达式或参数	对于 SQL 步骤, 此特性的值为 SQL 语句。 对于非 SQL 步骤, 该值是步骤的任何“参数”值的结合。
源定义的上次更改时间	步骤定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
数据库主机服务器名	目标数据库系统名
转换所有者	N/A
源顺序	N/A
转换次序	N/A
转换双向性	N/A
创建源定义的时间戳记	N/A
有关详情	管理员
注:	
1. 若将特性指定为“数据仓库中心”的内部特性, 则不会显示它。	

表 47. 信息目录管理器和“数据仓库中心”的业务主题区对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	主题区名
简短描述	描述
详细描述	注释
操作	N/A
数据刷新频率	N/A
访问数据所用的 URL	N/A
文件名	N/A
有关详情	管理员

表 48. 信息目录管理器和“数据仓库中心”的星型模式之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	仓库模式名
简短描述	描述
详细描述	注释
操作	N/A
有关详情	管理员
访问数据所用的 URL	N/A
上次更改源定义的时间戳记	仓库模式的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹
注:	
1. 若将特性指定为“数据仓库中心”的内部特性，则不会显示它。	

表 49. 信息目录管理器和“数据仓库中心”的过程对象之间的元数据映射

信息目录管理器元数据	“数据仓库中心”元数据
名称	过程名
简短描述	描述
详细描述	过程注释
操作	N/A
有关详情	管理员
访问数据所用的 URL	N/A
上次更改源定义的时间戳记	过程定义的上次更新时间戳记。 此元数据是“数据仓库中心”内部的。 ¹

表 49. 信息目录管理器和“数据仓库中心”的过程对象之间的元数据映射 (续)

注:

1. 若将特性指定为“数据仓库中心”的内部特性，则不会显示它。

信息目录管理器和 OLAP 服务器之间的元数据映射

表50显示 OLAP 服务器元数据到信息目录管理器公共对象类型的映射。OLAP 服务器元数据是指“DB2 OLAP 服务器”、DB2 OLAP Integration Server 或“Hyperion Essbase 服务器”的元数据。

当将 DB2 OLAP Integration Server 元数据发布到信息目录时，会在“多维数据库内的维”对象类型和表对象之间创建链接关系。

表的左列显示 Essbase API 结构的名称。右列显示信息目录管理器对象和对象类型特性。

表 50. OLAP 服务器元数据到信息目录管理器公共对象类型的映射

OLAP 服务器元数据	信息目录管理器元数据
轮廓	多维数据库
OLAP 对象的四部分名称以如下格式： server.application.database.outline	名称
指示宽度和深度限制的消息	详细描述
OLAP 服务器（名称的第一部分）	数据库主机服务器名
OLAP 数据库（名称的第三部分）	数据库或子系统名
N/A	数据库类型 此特性的值为 MULTIDIMENSIONAL。
ESB_OUTLINEINFO_T 中的 usOutlineType	数据库扩展类型 此特性的值可为 NORMAL 或 CURRENCY。
N/A	数据库状态 此特性的值是 PRODUCTION。
轮廓中的维数	多维数据库内的维数
EssOtlGetMemberAlias 中的维别名或名称	名称
OLAP 服务器	数据库主机服务器名
OLAP 数据库	数据库或子系统名
OLAP 应用程序	使用应用程序名称

表 50. OLAP 服务器元数据到信息目录管理器公共对象类型的映射 (续)

维名	维名
ESS_MBRINFO_T 中的 usCategory	维类别或类型
维中的成员	多维数据库内的成员
EssOtlGetMemberAlias 中的成员别名或名称	名称
OLAP 服务器	数据库主机服务器名
OLAP 数据库	数据库或子系统名
OLAP 应用程序	使用应用程序名称
维名	维名
成员名	成员名
最后一个 calc 字符串或 EssGetMemberCalc 中的 calc 字符串	派生自
ESS_MBRINFO_T 中的 usShare	将此特性作为共享成员（具有多个父代的成员）处理。

DB2 OLAP Integration Server 和“数据仓库中心”之间的元数据映射

表51显示 DB2 OLAP Integration Server 元数据到“数据仓库中心”元数据的映射。

表 51. DB2 OLAP Integration Server 元数据到“数据仓库中心”对象和标记语言的映射

DB2 OLAP Integration Server 元数据	“数据仓库中心”元数据标记语言
N/A	主题区 – OLAP 立方
OLAP 立方名 此特性值是以如下格式的四部分名称： server.application.database.outline	过程名
OLAP 立方名 此特性值是以如下格式的四部分名称： server.application.database.outline	步骤名
数据源	源数据库（预先存在于控制数据库中）
此特性基于 DB2 OLAP Integration Server 目录（OA_INFO 表）的查询，该查询基于 PHYSICAL_TABLE 列。	步骤输入表 为源数据库中的每个表创建“步骤输入表”对象。

表 51. DB2 OLAP Integration Server 元数据到“数据仓库中心”对象和标记语言的映射 (续)

<p>此特性基于 DB2 OLAP Integration Server 目录 (OM_INFO 表) 的查询, 该查询基于以下列:</p> <p>列 MODEL_NAME 和 MODEL_ID get MODEL_DATA_SOURCE</p>	数据库名
<p>此特性基于 DB2 OLAP Integration Server 目录 (OA_INFO 表) 的查询, 该查询基于以下列:</p> <p>MODEL_NAME 和 MODEL_ID get PHYSICAL_TABLE</p> <p>仅使用句点右边的名称 (没有模式)。</p>	表物理名
<p>OLAP 立方名</p> <p>此特性值是以如下格式的四部分名称: server.application.database.outline</p>	立方名
N/A	StepViewVWPOutputTable
<p>OLAP 立方名</p> <p>此特性值是以如下格式的四部分名称: server.application.database.outline</p>	步骤名
N/A	VWPPProgramTemplate – 构建 OLAP 立方
N/A	VWPPProgramTemplateParameter
N/A	VWPPProgramInstance
N/A	VWPPProgramInstanceParameter
DB2 OLAP Integration Server 批处理文件名	VWPPProgramInstanceParameterData
N/A	VWP 组 – OLAP 集成

附录C. “数据仓库中心” 如何迁移版本 5.2 定义

“数据仓库中心” 具有与 Visual Warehouse 版本 5.2 不同的对象模型。阅读本附录以理解“数据仓库中心” 如何将您的版本 5.2 定义迁移至“数据仓库中心”。

Visual Warehouse 对象和“数据仓库中心” 对象之间的映射

表52显示 Visual Warehouse 对象和“数据仓库中心” 对象之间的映射。

表 52. Visual Warehouse 对象和“数据仓库中心” 对象之间的映射

Visual Warehouse 版本 5.2	数据仓库中心
主题	主题区
信息资源或信息源	仓库源
仓库	主题区和仓库目标
业务视图	步骤

除了这些对象之外，“数据仓库中心” 添加了新的过程对象，过程对象包含完成某个仓库任务的一组步骤。在迁移之后，查看创建的过程，并确定需要创建的其他过程。可在过程之间移动步骤，并可在主题区之间移动过程。移动步骤和过程将步骤分成更小的可理解为单个任务的过程。

主题

Visual Warehouse 版本 5.2 主题将在“数据仓库中心” 中有为它创建的对应主题区。主题区的名称将与版本 5.2 主题的名称相同。每个主题区将包含一个过程。过程将命名为“主题区名 - 过程 1”，其中主题区名是包含该过程的主题区的名称。

仓库

Visual Warehouse 版本 5.2 仓库将在“数据仓库中心” 中有为它创建的对应主题区。主题区的名称将与版本 5.2 仓库的名称相同。每个主题区将包含一个过程。过程将命名为“主题区名 - 过程 1”，其中主题区名是包含该过程的主题区的名称。

业务视图

每个业务视图将在主题区和过程中具有一个对应的步骤，该过程对应包含该业务视图的主题或仓库。若业务视图包括在主题和仓库中，则对应该业务视图的步骤将包括在 **VW 至 DWC 迁移主题**主题区和 **VW 至 DWC 过程**过程中。必须仔细查看这些步骤并为它们确定适当的主题区和过程。

“数据仓库中心”如何迁移版本 5.2 定义

每个步骤将具有与它相关的类型和子类型。若业务视图使用了程序或变换器，则步骤将具有该程序或变换器的类型和子类型。若业务视图使用了 SQL，则步骤的类型将为 SQL。

触发器程序和并行级联关系不再存在。触发器程序已迁移为单独的步骤。并行级联关系已转换为调度表。

如果要将“可视仓库”元数据同步业务视图迁移到“数据仓库中心”，应在迁移仓库控制数据库前将业务视图提升到生产状态。如果业务视图处于生产方式，它们的调度也会迁移到“数据仓库中心”。如果业务视图未处于生产方式，它们将迁移到测试方式，但不迁移它们的调度。不能将迁移的步骤迁移到生产方式。必须在“数据仓库中心”中再次创建同步步骤，然后删除迁移的步骤。

Visual Warehouse 程序

Visual Warehouse 程序在“数据仓库中心”中已按下列方式更改：

- “数据仓库中心发布”设施置换了元数据同步 Visual Warehouse 程序。
- 有用于“DB2 通用数据库”、DB2 AS/400 版和 DB2 OS/390 版的新版 DB2 程序。Visual Warehouse 版本 5.2 中存在的 DB2 程序版本包含在称作“Visual Warehouse 版本 5.2 DB2 程序”的程序组中，仍可供您使用。
- “DB2 OLAP 程序”程序组现在称作“OLAP 服务器程序”。

安全性更改

下列 Visual Warehouse 安全性特权不再存在：

- 业务视图定义
- 业务视图维护
- 资源定义
- Visual Warehouse 程序定义

仅“管理”和“操作”特权保留在“数据仓库中心”中。

不再有用于业务视图的“更新安全性组”。与 Visual Warehouse 仓库或主题相关的安全性组仍将与某些过程相关，这些过程包含从仓库或主题迁移的业务视图。

附录D. 定义 Visual Warehouse 5.2 兼容程序的值

本附录为想要在“数据仓库中心”中运行某些 Visual Warehouse 5.2 程序的 Visual Warehouse 5.2 用户提供信息。

定义“Visual Warehouse 5.2 DB2 UDB 数据导出”(VWPEXPT1) 程序的值

使用此步骤在“数据仓库中心”中运行“Visual Warehouse DB2 UDB 数据导出”(VWPEXPT1) 仓库程序。使用“数据导出”程序从本地 DB2 数据库导出数据。

在为此步骤子类型定义值之前，在“过程模型”窗口中将仓库源与此步骤连接。将基于您的源定义自动定义此步骤子类型的参数值。

若导出过程出现警告，则程序返回为成功完成。

此程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- OS/2

“列映射”页不可用于此步骤。

要定义使用此程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，双击 SELECT 语句的**参数值**字段，并为该程序指定一个 SELECT 语句。
4. 可选：双击“文件类型修饰词字符串”参数的**参数值**字段，然后输入文件类型修饰词字符串。例如，“文件类型修饰词字符串”参数 "coldel," 指示必须用逗号分隔列。不要在 coldel 和逗号之间插入空格。确保将此参数包括在双引号中。否则，命令行处理器会将某些字符解释为文件重定向字符。
5. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
6. 单击**确定**保存更改并关闭步骤笔记本。

定义 Visual Warehouse 5.2 兼容程序的值

定义“Visual Warehouse DB2 UDB 数据插入装入”(VWPLOADI) 程序的值

使用此步骤在“数据仓库中心”中运行“Visual Warehouse 5.2 DB2 UDB 数据插入导入”(VWPLOADI) 仓库程序。使用 VWPLOADI 将文本文件中的数据装入 DB2 表，追加至现存数据之后。

必须在“过程过程”窗口中将该步骤与仓库源和仓库目标连接。

“Visual Warehouse 5.2 DB2 UDB 插入装入”仓库程序从“过程模型”窗口和步骤定义抽取下列步骤和仓库源参数值：

- 选择作为该步骤的源的文本文件。该步骤必须仅选择了一个源文件。源文件必须包含与目标表相同的字段数目和次序。仅支持定界 ASCII (ASCII DEL) 源文件。有关定界文件的格式的信息，参见 *DB2 Command Reference*。
- 仓库目标数据库名。必须对 DB2 数据库具有 SYSADM 或 DBADM 授权。“DB2 UDB 插入装入”程序不支持多节点数据库。对于多节点数据库，使用“DB2 UDB 扩展企业版”的“将文本文件装入 DB2 EEE”(VWPLDPR)。
- 该仓库目标的用户标识和密码。
- 为该步骤定义的目标表。

这些参数是预定义的。不用为这些参数指定值。另外，步骤还传送其他您提供值的参数。在程序将新数据装入表之前，它将该表导出到备份文件，该备份文件可用于恢复。

建议：在目标表自己的专用 DB2 表空间中创建它。缺省情况下，创建的任何专用表空间将由所有未指定表空间的新表使用。若处理失败，DB2 可能使整个表空间处于保持状态，使该表空间不可访问。要避免这种保持问题，为不使用装入程序的步骤创建另一个专用表空间。

要创建表空间：

```
CREATE TABLESPACE tablespace-name MANAGED BY SYSTEM USING ('d:directory')
```

其中 *directory* 是将包含数据库的目录。DB2 为您创建该目录。

限制：正在运行该程序的仓库代理站点的“数据仓库中心”定义必须包括用户标识和密码。不能由名为 SYSTEM 的用户运行 DB2 装入实用程序。确保在使用该程序的步骤的仓库源和仓库目标中选择相同的仓库代理站点。数据库服务器不需要在该代理站点上。然而，源文件必须在数据库服务器上。按在 DB2 服务器上的定义指定源文件的全限定名。

若仓库程序在处理期间检测到故障，该表将被清空。若装入出现警告，则程序返回为成功完成。

仓库程序不收集数据库统计信息。在完成相当大的装入之后，运行 DB2 UDB RUNSTATS 程序。

此仓库程序可用于下列操作系统：

- Windows NT
- AIX
- Solaris 操作环境
- OS/2

“列映射”页不可用于此步骤。

要定义使用此仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 双击“备份文件名”参数的**参数值**字段，并输入备份文件的全限定名。
4. 双击“文件类型修饰词字符串”参数的**参数值**字段，然后输入文件类型修饰词字符串。例如，“文件类型修饰词字符串”参数 "coldel," 指示必须用逗号分隔列。不要在 coldel 和逗号之间插入空格。确保将此参数包括在双引号中。否则，命令行处理器会将某些字符解释为文件重定向字符。

文件类型修饰词可使用下列修饰词：

修饰词	描述
Chardel <i>x</i>	<i>x</i> 是单字符串字符串定界符。缺省值是双引号 (")。您指定的字符用于置换包括字符串的双引号。可指定单引号 (') 作为字符串定界符，如下所示: Modified by chardel ''
Coldel <i>x</i>	<i>x</i> 是单字符串列定界符。缺省值为逗号 (,)。您指定的字符用于代替逗号来标识列的结束。不要在 coldel 和逗号之间插入空格。将此参数包括在双引号中。否则，命令行处理器会将某些字符解释为文件重定向字符。在以下示例中，coldel ; 导致导出实用程序将它遇到的任何 (;) 解释为列定界符: Db2 'export to temp of del modified by coldel; select * from staff where dept = 20'
Dateisiso	Dateisiso 导致以 ISO 格式导出所有日期数据值。

定义 Visual Warehouse 5.2 兼容程序的值

Decplusblank	Decplusblank 导致在正小数值前加一个空格，而不是加号 (+)。缺省操作是在正小数值加一个加号。
Decpt <i>x</i>	<i>x</i> 是代替句点作为小数点字符的单个字符。缺省值为句点 (.)。· 您指定的字符用于置换句点作为小数点字符。

有关这些修饰词的详情，参见 *DB2 Utilities* 一书。

5. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
6. 单击**确定**保存更改并关闭步骤笔记本。

定义“Visual Warehouse DB2 UDB 数据置换装入”(VWPLOADR)程序的值

使用此步骤在“数据仓库中心”中运行“Visual Warehouse 5.2 DB2 数据置换装入”(VWPLOADR)仓库程序。使用 VWPLOADR 将数据从文本文件装入 DB2 表，置换现存数据。

必须在“过程过程”窗口中将该步骤与仓库源和仓库目标连接。

“Visual Warehouse 5.2 DB2 UDB 置换装入”仓库程序从“过程模型”窗口和步骤定义抽取下列步骤和仓库源参数值：

- 选择作为该步骤的源的文本文件。该步骤必须仅选择一个源文件。源文件必须包含与目标表相同的字段数目和次序。仅支持定界 ASCII (ASCII DEL) 源文件。有关定界文件的格式的信息，参见 *DB2 Command Reference*。
- 仓库目标数据库名。必须对 DB2 数据库具有 SYSADM 或 DBADM 授权。此程序不支持多节点数据库。对于多节点数据库，使用“DB2 UDB 扩展企业版”的“将文本文件装入 DB2 EEE”(VWPLDPR)。
- 该仓库目标的用户标识和密码。
- 为该步骤定义的目标表。

这些参数是预定义的。不用为这些参数指定值。

建议：在目标表自己的专用 DB2 表空间中创建它。创建的任何专用表空间将由所有未指定表空间的新表使用。若处理失败，DB2 可能使整个表空间处于保持状态，使该表空间不可访问。要避免这种保持问题，为不使用装入程序的步骤创建另一个专用表空间。

要创建表空间:

```
CREATE TABLESPACE tablespace-nameMANAGED BY SYSTEM USING ('d:/directory')
```

其中 *directory* 是将包含数据库的目录。DB2 创建该目录。

限制: 正在运行该程序的代理站点的“数据仓库中心”定义必须包括用户标识和密码。不能由名为 **SYSTEM** 的用户运行 DB2 装入实用程序。确保在使用该仓库程序的步骤的仓库源和仓库目标中选择相同的代理站点。数据库服务器不需要在该代理站点上。然而,源文件必须在数据库服务器上。按在 DB2 服务器上的定义指定源文件的全限定名。

若仓库程序在处理期间检测到故障,则该表被清空。若装入出现警告,则程序返回为成功完成。

“DB2 UDB 置换装入”程序在装入期间收集数据库统计信息,因此不需要在此程序之后运行 DB2 UDB RUNSTATS (VWPSTATS) 程序。

此仓库程序可用于下列操作系统:

- Windows NT
- AIX
- Solaris 操作环境
- AS/400
- OS/2

要定义使用此程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情,参见第139页的『提供关于步骤的一般信息』。
3. 可选: 在“参数”页上,双击“文件类型修饰词字符串”参数的**参数值**字段,然后输入文件类型修饰词字符串。例如,“文件类型修饰词字符串”参数 "coldel,"指示必须用逗号分隔列。不要在 coldel 和逗号之间插入空格。确保将此参数包括在双引号中。否则,命令行会将某些字符解释为文件重定向字符。

文件类型修饰词可使用下列修饰词:

定义 Visual Warehouse 5.2 兼容程序的值

修饰词	描述
Chardel <i>x</i>	<i>x</i> 是单字符字符串定界符。缺省值是双引号 (")。您指定的字符用于置换包括字符串的双引号。可指定单引号 (') 作为字符串定界符, 如下所示: Modified by chardel ''
Codel <i>x</i>	<i>x</i> 是单字符列定界符。缺省值为逗号 (,)。您指定的字符用于代替逗号来标识列的结束。不要在 codel 和逗号之间插入空格。将此参数包括在双引号中。否则, 命令行处理器会将某些字符解释为文件重定向字符。在以下示例中, codel ; 导致导出实用程序将它遇到的任何 (;) 解释为列定界符: Db2" export to temp of del modified by codel; select * from staff where dept = 20"
Dateisiso	Dateisiso 导致以 ISO 格式导出所有日期数据值。
Decplusblank	Decplusblank 导致在正小数数值前加一个空格, 而不是加号 (+)。缺省操作是在正小数数值加一个加号。
Decpt <i>x</i>	<i>x</i> 是代替句点作为小数点字符的单个字符。缺省值为句点 (.)。您指定的字符用于置换句点作为小数点字符。

有关这些修饰词的详情, 参见 *DB2 Utilities* 一书。

4. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
5. 单击**确定**保存更改并关闭步骤笔记本。

定义 Visual Warehouse 5.2 DB2 UDB REORG (VWPREORG) 程序的值

使用此步骤在“数据仓库中心”中运行 Visual Warehouse 5.2 DB2 UDB REORG (VWPREORG) 程序。此步骤在目标表上运行 DB2 UDB REORG 实用程序。

可调度此步骤在过程完成之后在过程的目标表上运行。在“过程模型”窗口中, 从该步骤拖出一个数据链接到仓库目标。

Visual Warehouse 5.2 DB2 UDB REORG 程序从“过程模型”窗口和步骤定义抽取下列步骤和仓库源参数值:

- 仓库目标数据库名
- 仓库目标的用户标识和密码

- 为该步骤定义的目标表

这些参数是预定义的。不用为这些参数指定值。

要定义使用此程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
4. 单击**确定**保存更改并关闭步骤笔记本。

定义 Visual Warehouse 5.2 DB2 UDB RUNSTATS (VWPSTATS) 程序的值

使用此步骤在“数据仓库中心”中运行 Visual Warehouse 5.2 DB2 UDB RUNSTATS (VWPSTATS) 程序。

此步骤在目标表上运行 DB2 UDB RUNSTATS 实用程序。可调度此步骤在过程完成之后在过程的目标表上运行。在“过程模型”窗口中, 从该步骤拖出一个数据链接到仓库目标。

Visual Warehouse 5.2 DB2 UDB RUNSTATS 仓库程序从“过程模型”窗口和步骤定义抽取下列步骤和仓库源参数值:

- 仓库目标数据库名
- 仓库目标的用户标识和密码
- 为该步骤定义的目标表

这些参数是预定义的。不用为这些参数指定值。

要定义使用此仓库程序的步骤的值:

1. 打开步骤笔记本。
2. 指定关于仓库程序的一般信息。有关详情, 参见第139页的『提供关于步骤的一般信息』。
3. 在“处理选项”页上, 提供有关步骤如何处理的信息。有关详情, 参见第141页的『定义处理选项』。
4. 单击**确定**保存更改并关闭步骤笔记本。

定义 Visual Warehouse 5.2 兼容程序的值

定义“Visual Warehouse 5.2 将文本文件装入 DB2 UDB EEE” (VWPLDPR) 程序的值 (仅 AIX)

使用此步骤在“数据仓库中心”中运行“Visual Warehouse 5.2 将文本文件装入 DB2 UDB EEE (仅 AIX)” (VWPLDPR) 程序。此步骤将数据从定界文本文件装入“DB2 通用数据库扩展企业版”数据库中，置换现存数据。

在使用此仓库程序之前，必须熟悉并行系统概念和并行装入。

当 VWPLDPR 程序将数据装入到并行数据库时，它执行下列步骤：

1. 与目标数据库连接。
2. 获取该数据库的目标分区映射图。
3. 分割输入文件，以便可在节点上装入每个文件。
4. 在所有节点运行远程装入。

若装入步骤在任何节点上失败，VWPLDPR 程序执行下列操作：

1. 在每个节点上构建一个空的装入数据文件。
2. 装入空的数据文件。

VWPLDPR 程序从“过程模型”窗口和步骤定义抽取下列步骤和仓库源参数值：

- 选择作为该步骤的源的文本文件。该步骤必须仅选择一个源文件。仅支持定界 (DEL) 文件。输入文件和分割的文件必须在由数据库装入中涉及的所有节点共享的文件系统上。必须将共享文件系统安装在所有节点的同目录上。该目录必须大到足够包含分割之前和分割之后的输入文件。
- 仓库目标数据库名。
- 该仓库目标的用户标识和密码。
- 为该步骤定义的目标表。

这些参数是预定义的。不用为这些参数指定值。另外，有几个参数必须提供值。

“将文本文件装入 DB2 UDB EEE”在装入之后不运行 DB2 RUNSTATS 实用程序。若想要在装入之后自动运行 RUNSTATS 实用程序，给运行 RUNSTATS 的过程添加一个步骤。

建议：在目标表自己的专用 DB2 表空间中创建它。创建的任何专用表空间将由所有未指定表空间的新表使用。若处理失败，DB2 可能使整个表空间处于保持状态，使该表空间不可访问。要避免这种保持问题，为不使用装入程序的步骤创建另一个专用表空间。

要创建表空间：

定义 Visual Warehouse 5.2 兼容程序的值

```
CREATE TABLESPACE tablespace-name MANAGED BY SYSTEM USING (d:/directory')
```

其中 *directory* 是将包含数据库的目录。DB2 创建该目录。

此程序在 AIX 上运行。

要定义使用此仓库程序的步骤的值：

1. 打开步骤笔记本。
2. 指定关于程序的一般信息。有关详情，参见第139页的『提供关于步骤的一般信息』。
3. 在“参数”页上，提供下列参数的值：
 - a. 双击“列定界符”参数的**参数值**字段，然后输入列定界符。最常用的值是逗号 (,) 或分号 (;)。
 - b. 双击字符串参数的**参数值**字段，然后输入字符串参数。最常用的值是双引号 (")，输入为 "\"。
 - c. 双击“小数定界符”参数的**参数值**字段，然后输入小数定界符。最常用的值是句点 (.) 或逗号 (,)。
 - d. 双击本地非共享文件目录参数的**参数值**字段，然后输入本地非共享文件目录。
 - e. 双击该参数的“路径名和前缀”的**参数值**字段，然后输入分割文件的路径名和前缀。每个文件的名称将由前缀加一个数字标识符组成。
 - f. 双击“分区键”参数的**参数值**字段，然后为每个分区键输入一个参数。分区键必须以 `db2split` 数据库实用程序使用的格式。格式通常如下：
`col1,1,,N,integer` 后跟 `col3,3,,5N,character`
4. 在“处理选项”页上，提供有关步骤如何处理的信息。有关详情，参见第141页的『定义处理选项』。
5. 单击**确定**保存更改并关闭步骤笔记本。

定义 Visual Warehouse 5.2 兼容程序的值

附录E. 数据入库样本

“DB2 通用数据库”提供数据入库样本，可用来熟悉“数据仓库中心”。它包括样本数据和元数据，可运行它们在仓库数据库中创建表。

安装和运行该样本以了解源数据和目标数据之间的关系，以及“数据仓库中心”中移动和转换数据的定义。

该样本在“数据仓库中心”中定义一个星型模式。该星型模式具有三个维表：

- PRODUCT
- SCENARIO
- TIME

它具有一个事实表，名为 FACT TABLE。

安装样本

仅可在 Windows NT 上安装“数据入库”样本。该样本在 Windows NT 上安装为“DB2 通用数据库”的典型安装的一部分。也可通过在定制安装中选择入门组件，或入门组件的**第一步和样本数据库**子组件，来安装该样本。

必须在仓库服务器所在的同一工作站上安装该样本。

创建样本数据库

在安装该样本的文件之后，必须创建样本数据库。要创建数据库：

1. 打开“第一步”窗口。
2. 单击**创建样本数据库**。
“创建样本数据库”窗口打开。
3. 选择**数据入库样本**复选框。
4. 单击**确定**。
5. 若正安装数据入库样本，则打开一个窗口，供您输入用于访问该样本的 DB2 用户标识和密码。
 - a. 输入想要使用的用户标识和密码。必须指定有效的 DB2 用户标识和密码。
 - b. 单击**确定**。

进度窗口打开。创建数据库完成时，单击**确定**。

向 ODBC 注册创建的数据库。

创建了下列数据库：

DWCTBC

包含教程的数据入库部分所需的操作源表。

TBC_MD

包含“数据仓库中心”的元数据。

创建仓库数据库

需要为运行样本时生成的数据创建一个数据库。

要创建数据库：

1. 启动“DB2 控制中心”：

- 在 Windows NT 上，单击**开始** → **程序** → **IBM DB2** → **控制中心**。
- 在 AIX 或 Solaris 操作环境中，输入如下命令：

```
db2jstrt 6790
db2cc 6790b
```

2. 右键单击**数据库**文件夹，并单击**使用向导创建数据库**。“创建数据库向导”打开。

3. 在**数据库名**字段中，输入数据库的名称：

SAMPWHS

若使用不同的名称，则必须更改“TBC 样本目标”仓库目标中的数据库的名称。否则，样本将不能运行。

4. 从**缺省驱动器**列表中，选择数据库的驱动器。

5. 在**注释**字段中，输入数据库的描述：

样本仓库数据库

6. 单击**完成**。此向导中的所有其他字段和页是可选的。创建 SAMPWHS 数据库并将它列示在“DB2 控制中心”中。

向 ODBC 注册数据库有多种方法。可使用 Windows NT 上的“客户机配置辅助”、“DB2 命令行处理器”或 Windows NT 上的“ODBC32 数据源管理员”。下列指示信息适用于“客户机配置辅助”。

有关“命令行处理器”的详情，参见 *DB2 Universal Database Command Reference*。

有关“ODBC32 数据源管理员”的详情，参见“管理员”中的联机帮助。

要在 Windows NT 上向 ODBC 注册 SAMPWHs 数据库:

1. 通过单击**开始** → **程序** → **IBM DB2** → **客户机配置辅助**, 启动“客户机配置辅助”。“客户机配置辅助”窗口打开。
2. 从数据库列表中选择 **SAMPWHs**。
3. 单击**特性**。“数据库特性”窗口打开。
4. 选择对 **ODBC 注册此数据库**。使用缺省选择**作为系统数据源**, 意指该数据可供系统上的所有用户使用。
5. 单击**确定**。其他所有字段是可选的。已向 ODBC 注册 SAMPWHs 数据库。

使用样本内容

使用“样本内容”可查看包含在仓库表中的数据样本。从包含要采样的表的“表”文件夹中, 或从过程建模器中的表图符运行“样本内容”。

从过程建模器在仓库目标上运行“样本内容”时, 仓库源、仓库目标和步骤必须有一个公共的代理站点, 否则请求将失败。要避免这种问题, 可对要使用的代理站点在“特性”笔记本中的“程序”页上选择步骤。

对目标表运行“样本内容”时, 将使用所选列表上的第一个代理站点。但是, 数据库维护操作可能会影响到代理站点列示的次序。如果选择的代理站点与源或目标文件不在同一系统中, 样本内容将失败。

如果“样本内容”请求以平面文件源的形式使用 AS/400 代理, 该请求会失败。

查看样本数据

DWCTBC 数据库包含样本仓库的源表。它包含下列表:

- SALES
- INVENTORY
- PRODUCTION_COSTS
- GEOGRAPHIES
- SCENARIO
- TIME
- PRODUCT

要查看这些表中的数据:

1. 从“DB2 控制中心”中, 展开 DWCTBC 数据库中的对象, 直到看到表文件夹。
2. 单击该文件夹。在右窗格中, 会看到该数据库的所有表。

3. 找到想要查看的表。右键单击它，并单击**样本内容**。

最多可显示该表的 200 行。列名显示在窗口的顶部。可能需要向右滚动以查看所有列，和向下滚动以查看所有行。

查看和修改样本元数据

要访问样本，必须登录到“数据仓库中心”，指定 TBC_MD 为仓库控制数据库。

若 TBC_MD 数据库不是包含仓库服务器的工作站的本地数据库，必须在该工作站上将它编目为远程数据库。若它不是包含“数据仓库中心”管理客户机的工作站的本地数据库，必须也在该工作站上编目它。

要登录到“数据仓库中心”：

1. 在“DB2 控制中心”中单击**工具 → 数据仓库中心**。系统将使用您的“DB2 控制中心”用户名和密码自动尝试登录到“数据仓库中心”。“数据仓库中心登录”窗口打开。
2. 单击**高级**。
“高级”窗口打开。
3. 在**控制数据库**字段中，输入 TBC_MD，它是样本中包括的仓库控制数据库的名称。
4. 在**服务器主机名字段**中，输入安装仓库服务器的工作站的 TCP/IP 主机名。
5. 单击**确定**。

“高级登录”窗口关闭。

下次登录时，“数据仓库中心”将使用您在“高级登录”窗口中指定的设置。若先用对“数据仓库中心”定义的用户标识登录到“DB2 控制中心”，当单击**工具 → 数据仓库中心**时，“数据仓库中心”将使用同一用户标识自动进行登录。

6. 在“登录”窗口的**用户标识**字段中，输入创建“数据入库”样本数据库时指定的用户标识。
7. 在**密码**字段中，输入该用户标识的密码。
8. 单击**确定**。

“数据仓库中心登录”窗口关闭，现在您已登录到“数据仓库中心”。

登录到“数据仓库中心”之后，可查看和修改样本元数据的特性：

1. 展开**仓库源**文件夹。
显示“TBC 样本源”仓库源。
2. 右键单击 **TBC 样本源** 仓库源，并单击**特性**以查看该仓库源的特性。

3. 在“数据库”页的**系统名**字段中，指定样本数据库所在工作站的主机名。
此步骤是可选的，且仅当使用名称与不同工作站上的另一数据库或文件相同的数据库或文件时，才使用该步骤。
4. 在**用户标识** 字段中，输入创建样本数据库时指定的用户标识。
5. 在**密码**字段中，输入该用户标识的密码。
6. 在**验证密码**字段中再次输入该密码。
7. 单击**确定**关闭笔记本。
8. 单击**表**文件夹。仓库源中的表显示在窗口的右边。

仓库源包含下列表:

- GEOGRAPHIES
- INVENTORY
- PRODUCT
- SCENARIO
- TIME
- PRODUCTION_COSTS
- SALES

9. 右键单击一个表，并单击**特性**以查看它的特性。
10. 单击**确定**或**取消**关闭笔记本。
11. 展开**仓库目标**文件夹。
显示“TBC 样本目标”仓库目标。
12. 右键单击 **TBC 样本目标**仓库目标，并单击**特性**以查看该仓库目标的特性。
13. 在“数据库”页的**系统名**字段中，指定样本数据库所在工作站的主机名。
此步骤是可选的，且仅当使用名称与不同工作站上的另一数据库或文件相同的数据库或文件时，才使用该步骤。
14. 在**用户标识** 字段中，输入创建样本数据库时指定的用户标识。
15. 在**密码**字段中，输入该用户标识的密码。
16. 在**验证密码**字段中再次输入该密码。
17. 单击**确定**关闭笔记本。
18. 展开**主题区**文件夹。
显示 **TBC 样本**主题区的图符。
19. 右键单击 **TBC 样本**主题区，并单击**特性**以查看该主题区的特性。
20. 查看完特性之后，单击**确定**或**取消**关闭笔记本。
21. 展开 **TBC 样本**主题区目录树。

22. 展开**过程**文件夹。

下列四个过程显示在该文件夹下：

- 构建样本事实表
- 构建示例产品维
- 构建样本方案维
- 构建样本时间维

23. 右键单击**构建样本事实表**过程，然后单击**打开**。

“过程模型”窗口打开。

该过程包含下列对象：

- 三个源：
 - SALES
 - PRODUCTION_COSTS
 - INVENTORY
- 名为“事实表 SQL”的 SQL 步骤。该 SQL 步骤连接该过程中的源，并添加三列：CITY_ID、TIME_ID 和 SCENARIO_ID。
- 名为 TARGET FACT TABLE 的目标表。
- “选择时间”步骤的快捷方式。此快捷方式在调度此样本中运行的步骤顺序时使用。

要查看过程中的对象的元数据，右键单击该对象并单击**特性**。单击**确定**或**取消**关闭笔记本。

24. 右键单击**构建示例产品维**过程，然后单击**打开**。

“过程模型”窗口打开。

该过程包含下列对象：

- 名为 PRODUCT 的源。
- 名为“选择产品”的 SQL 步骤。该 SQL 步骤选择源中的所有列。
- 名为 TARGET PRODUCT 的生成目标表。
- “选择方案”步骤的快捷方式。此快捷方式在调度此样本中运行的步骤顺序时使用。

要查看过程中的对象的元数据，右键单击该对象并单击**特性**。单击**确定**或**取消**关闭笔记本。

25. 右键单击**构建样本方案维**过程，然后单击**打开**。

“过程模型”窗口打开。

该过程包含下列对象：

- 名为 SCENARIO 的源。
- 名为“选择方案”的 SQL 步骤。该 SQL 步骤选择源中的所有列。不能使用“SQL 助手”编辑此步骤中的 SQL。SQL 是在由“SQL 助手”创建后编辑的。
- 名为 TARGET SCENARIO 的生成目标表。
- “选择时间”步骤的快捷方式。此快捷方式在调度此样本中运行的步骤顺序时使用。

要查看过程中的对象的元数据，右键单击该对象并单击特性。单击确定或取消关闭笔记本。

26. 右键单击构建样本时间维过程，然后单击打开。

“过程模型”窗口打开。

该过程包含下列对象：

- 名为 TIME 的源。
- 名为“选择时间”的 SQL 步骤。该 SQL 步骤选择源中的所有列。
- 名为 TARGET TIME 的生成目标表。
- “事实表 SQL”步骤的快捷方式。此快捷方式在调度此样本中运行的步骤顺序时使用。

要查看过程中的对象的元数据，右键单击该对象并单击特性。单击确定或取消关闭笔记本。

27. 展开仓库模式文件夹。

28. 右键单击该仓库模式，并单击打开。

“仓库模式模型”窗口打开。它包含下列表的连接：

- TARGET FACT TABLE
- TARGET PRODUCT
- TARGET SCENARIO
- TARGET TIME

提升步骤

在运行步骤之前，必须将它们提升为测试方式，再提升为生产方式。必须按它们运行的次序提升它们：

1. 选择产品
2. 选择方案
3. 选择时间

4. 事实表 SQL

要将步骤提升为测试方式:

1. 右键单击**构建示例产品维**过程，然后单击**打开**。
2. 右键单击“选择产品”步骤，并单击**方式** → **测试**。
“数据仓库中心”显示进度窗口。在启动下一过程之前，等待直到“数据仓库中心”完成处理。
3. 对其余步骤重复步骤1和2，以将这些步骤提升为测试方式：
 - “构建样本方案维”过程中的“选择方案”步骤。
 - “构建样本时间维”过程中的“选择时间”步骤。
 - “构建样本事实表”过程中的“事实表 SQL”步骤。

要将步骤提升为生产方式:

1. 右键单击**构建示例产品维**过程，然后单击**打开**。
2. 右键单击“选择产品”步骤，并单击**方式** → **生产**。
“数据仓库中心”开始创建目标表，并显示一个进度窗口。在启动下一过程之前，等待直到“数据仓库中心”完成处理。
3. 对其余步骤重复步骤1和2，以将这些步骤提升为生产方式：
 - “构建样本方案维”过程中的“选择方案”步骤。
 - “构建样本时间维”过程中的“选择时间”步骤。
 - “构建样本事实表”过程中的“事实表 SQL”步骤。

运行步骤

要运行这些步骤，运行“样本产品”步骤。其余步骤在“样本产品”步骤完成处理之后依次启动。

要运行“样本产品”步骤:

1. 在“数据仓库中心”主窗口中，单击**仓库** → **工作进度**。
“工作进度”窗口打开。使用“工作进度”窗口监视“数据仓库中心”中所有正运行或已调度的步骤的进度。
2. 单击**工作进度** → **运行新步骤**。
“运行新步骤”窗口打开。
3. 选择“样本产品”步骤，并单击 **>**。
4. 单击**确定**。

应看见正在运行的步骤项。当步骤正在运行时，它的状态为“正在填充”。当它完成运行时，它的状态应为“成功”。当序列中的一个步骤完成处理时，下一个步骤开始运行并变为“正在填充”状态。

查看样本仓库数据

您创建的 SAMPWHS 数据库包含步骤处理的结果。它包含下列表：

- TARGET FACT TABLE
- TARGET PRODUCT
- TARGET SCENARIO
- TARGET TIME

要查看这些表中的数据：

1. 从“DB2 控制中心”中，展开 SAMPWHS 数据库中的对象，直到看到**表文件夹**。
2. 单击该文件夹。在右窗格中，会看到该数据库的所有表。
3. 找到想要查看的表。右键单击它，并单击**样本内容**。

最多可显示该表的 200 行。列名显示在窗口的顶部。可能需要向右滚动以查看所有列，和向下滚动以查看所有行。

附录F. 在“数据仓库中心”中使用 Classic Connect

可在“数据仓库中心”中使用 Classic Connect 来访问 IMS 和 VSAM 数据库中的数据。

什么是 Classic Connect?

Classic Connect 提供了对存储在“信息管理系统”(IMS)数据库和 OS/390 上“虚拟存储访问方法”(VSAM)数据集中的非关系数据的读访问权。它提供通信、数据访问和数据映射功能,因此可使用关系查询读取非关系数据。

本章包含以下几节:

- 『它起什么作用?』
- 『它可访问哪些数据源?』
- 第364页的『如何使用它?』
- 第364页的『它有哪些组件?』

它起什么作用?

Classic Connect 允许通过从“数据仓库中心”步骤发出标准的 SQL 查询访问非关系数据。就象数据在 DB2 数据库中一样访问这些数据。

它可访问哪些数据源?

Classic Connect 对 IMS 数据库和 VSAM 数据集提供只读关系访问。它创建一个逻辑关系数据库,由映射至 IMS 或 VSAM 数据库中的实际数据的逻辑表组成。尤其是:

- 对于 VSAM,每个逻辑表对应一个 VSAM 数据集。数据集中的每个字段映射为逻辑表的一列;数据集的每个记录映射为一行。Classic Connect 可从 VSAM KSDS、RRDS 和 ESDS 数据集中读取数据。
- 对于 IMS,每个逻辑表对应 IMS 全功能数据库中的路径的一段或多段。路径中多个段的字段对应逻辑表的列。给定路径的每个唯一的段实例集对应逻辑表中的一行。

使用此关系结构, Classic Connect 对 IMS 数据库和 VSAM 数据集解释用户提交的关系查询。

在“数据仓库中心”中使用 Classic Connect

可对单个数据源（如一组 VSAM 数据集或单个 IMS 数据库）定义多个逻辑数据库。可在逻辑数据库中定义多个逻辑表。

可对单个数据实体（如一个 VSAM 数据集或 IMS 段）定义多个逻辑表。例如，单个 VSAM 数据集可有为其定义的多个逻辑表，每个逻辑表以新的方法映射数据。

如何使用它？

若仓库源使用 IMS 或 VSAM 数据库中的操作数据，则在“数据仓库中心”中使用 Classic Connect。使用 Classic Connect 将非关系数据映射为伪关系数据。然后使用 CROSS ACCESS ODBC 驱动程序访问伪关系数据。然后可在“数据仓库中心”中定义对应伪关系数据的 IMS 或 VSAM 仓库源。

它有哪些组件？

在“数据仓库中心”中使用 Classic Connect 由下列主要组成部分组成：

- 第365页的『仓库代理』
- 第366页的『CROSS ACCESS ODBC 驱动程序』
- 第366页的『Classic Connect 数据服务器』
- 第368页的『企业服务器』
- 第370页的『数据映射器』

第365页的图28显示 Classic Connect 及其组件如何融入整个“数据仓库中心”体系结构。

在“数据仓库中心”中使用 Classic Connect

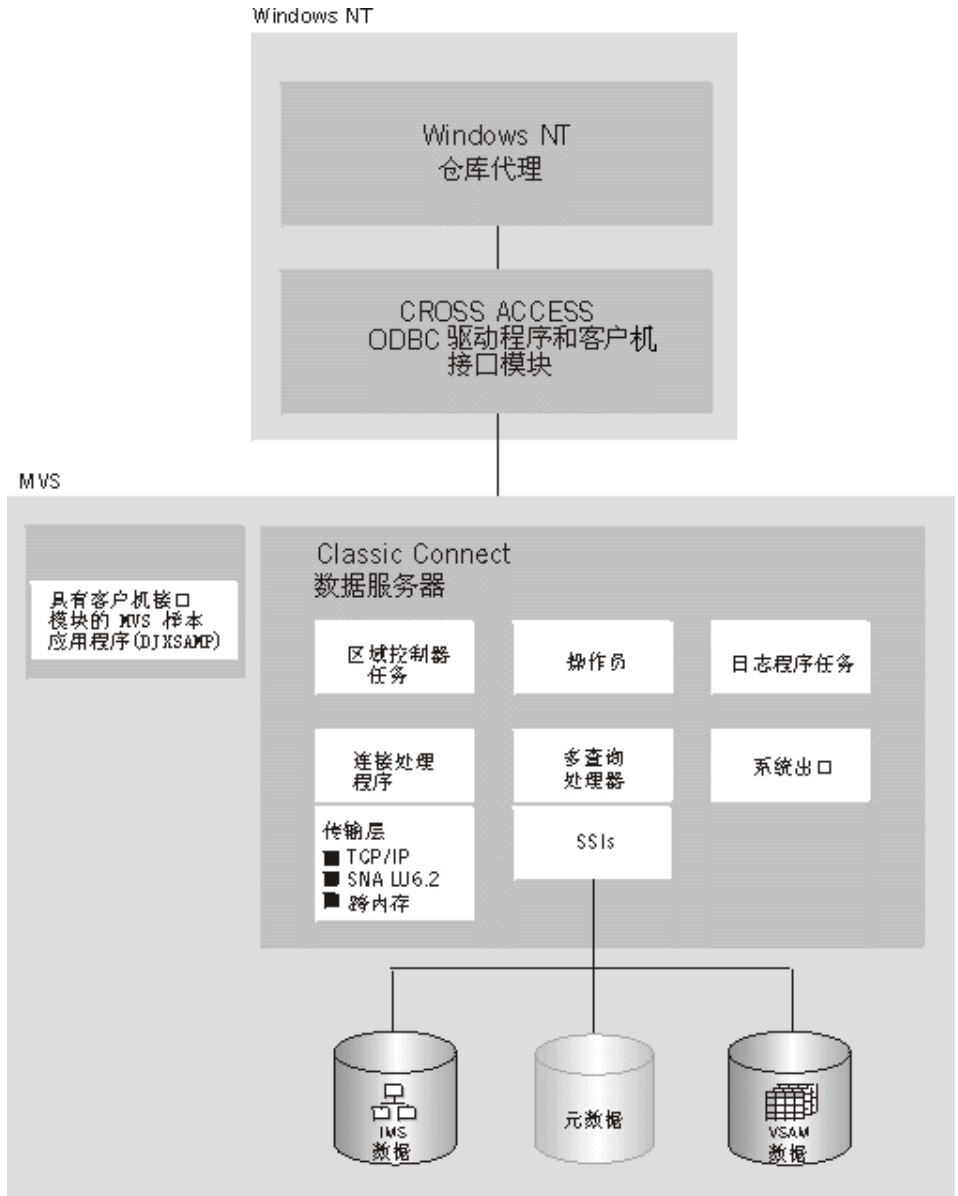


图 28. Classic Connect 体系结构

仓库代理

仓库代理管理数据源和目标仓库之间的数据传送。仓库代理使用“CROSS ACCESS ODBC 驱动程序”与 Classic Connect 通信。

在“数据仓库中心”中使用 Classic Connect

CROSS ACCESS ODBC 驱动程序

“开放式数据库链接” (ODBC) 接口允许应用程序使用“结构化查询语言” (SQL) 语句访问关系和非关系数据库管理系统中的数据。关于设置 ODBC 接口 z/OS 版和 OS/390 版的信息, 请参阅《仓库管理器安装指南》。

ODBC 体系结构由四部分组成:

- 依照 ODBC 的应用程序执行处理, 并调用 ODBC 函数提交 SQL 语句和检索结果。
- “驱动程序管理器”为应用程序装入驱动程序。
- 驱动程序处理 ODBC 函数调用, 将 SQL 请求提交给特定的数据源, 并将结果返回给应用程序。
- 数据源由用户想要访问的数据组成。数据源名与 CROSS ACCESS 数据服务器配置文件中的数据源名等效。

“驱动程序管理器”和驱动程序充当一个处理 ODBC 函数调用的单元。

Classic Connect 数据服务器

所有数据访问由 Classic Connect 数据服务器执行。数据服务器负责下列功能:

- 接受来自“数据仓库中心”的 SQL 查询。
- 确定要访问的数据类型。
- 将 SQL 查询重写为所需的本机文件或数据库访问语言。单个 SQL 访问可转换为多个本机数据请求。
- 基于类属 SQL 查询重写和文件或数据库特定的优化, 优化查询。
- 查询 JOIN 的多个数据源。
- 将结果集转换为一致的关系格式, 涉及将非关系数据重构为列和行。
- 根据需要对结果集进行排序; 例如, ORDER BY。
- 对 Classic Connect 元数据目录发出所有客户机目录查询。

Classic Connect 数据服务器接受来自 CROSS ACCESS ODBC 驱动程序和 OS/390 上的样本程序的连接请求。

可在数据服务器中运行五种服务:

- 区域控制器服务, 包括 MTO 操作员界面
- 初始化服务
- 连接处理程序服务
- 查询处理器服务
- 日志记录器服务

区域控制器服务: 数据服务器的主要组件是区域控制器。区域控制器负责启动、停止和监控数据服务器的其他所有组件。这些不同的组件称作服务。服务作为单个装入模块来实现，这些装入模块在数据服务器地址空间内作为单独的 OS/390 任务运行。服务可具有多个实例，且每个实例可支持多个用户。

区域控制器基于 SERVICE INFO ENTRY 参数设置确定启动哪些服务。

区域控制器服务中包括“OS/390 主终端操作员”(MTO)界面，它允许您显示和控制数据服务器提供的服务正在接受服务的用户。使用 MTO 界面，还可动态地配置数据服务器:

初始化服务: 初始化服务是一些特殊服务，用于初始化和终止与底层数据库管理系统或 OS/390 系统组件的不同类型的接口。目前，提供三种初始化服务:

IMS BMP/DBB 初始化服务

用于初始化 IMS 区域控制器，以使用 BMP/DBB 接口访问 IMS 数据

IMS DRA 初始化服务

用于初始化 Classic Connect DRA 接口和与 IMS DBCTL 区域连接，以使用 DRA 接口访问 IMS 数据

WLM 初始化服务

用于初始化和注册“OS/390 工作负荷管理器”子系统(使用“WLM 系统出口”)。这允许以 WLM 目标方式处理单个查询。

连接处理程序服务: 连接处理程序(CH)服务任务负责收听来自“数据仓库中心”的连接请求。将连接请求按路径发送到适当的查询处理器任务进行后续处理。

Classic Connect 提供三种典型的可由 CH 任务装入的传输层模块:

- TCP/IP
- SNA LU 6.2
- OS/390 跨内存服务。

OS/390 客户机应用程序 DJXSAMP 可使用这些方法中的任一种与数据服务器连接;然而,对本地客户机建议的方法是使用 OS/390 跨内存服务。“数据仓库中心”可使用 TCP/IP 或 SNA 与远程数据服务器通信。

查询处理器服务: 查询处理器是数据服务器的一个组件,负责将客户机 SQL 转换为数据库和文件特定的数据访问请求。查询处理器将 IMS 和 VSAM 数据当作单个数据源处理,并能够处理访问 IMS 或 / 和 VSAM 的 SQL 语句。多个查询处理器可用于分别控制配置参数,如影响跟踪和主管程序的那些参数,以满足个别应用程序的需要。

在“数据仓库中心”中使用 **Classic Connect**

查询处理器可对 SELECT 语句提供服务。查询处理器调用一个或多个子系统接口 (SSI) 来访问 SQL 请求中引用的目标数据库或文件系统。支持下列 SSI:

IMS BMP/DBB 接口

允许通过 IMS 区域控制器访问 IMS 数据。将区域控制器限制为数据服务器的单个 PSB，限制了数据服务器可处理的并行用户的数目。

IMS DRA 接口

允许使用 IMS DRA 接口访问 IMS 数据。DRA 接口支持多个 PSB，并支持大量用户的唯一方法。这是建议的接口。

VSAM 接口

允许访问 VSAM ESDS、KSDS 或 RRDS 文件。此接口还支持使用备用索引。

Classic Connect 提供 RUNSTATS 实用程序，用于更新逻辑表及其相关索引和键的填充统计信息。查询处理器可使用此信息优化 JOIN。

日志记录器服务: 日志记录器服务是用于系统监控和疑难解答的任务。单个日志记录器任务可在数据服务器内运行。在正常操作期间，不需要考虑日志记录器服务。

企业服务器

企业服务器是一个可选组件，可用于管理跨多个数据源的大量并行用户。除查询处理器服务和初始化服务之外，企业服务器包含数据服务器使用的相同任务。

在“数据仓库中心”中使用 Classic Connect

图29显示企业服务器如何融入 Classic Connect 配置:

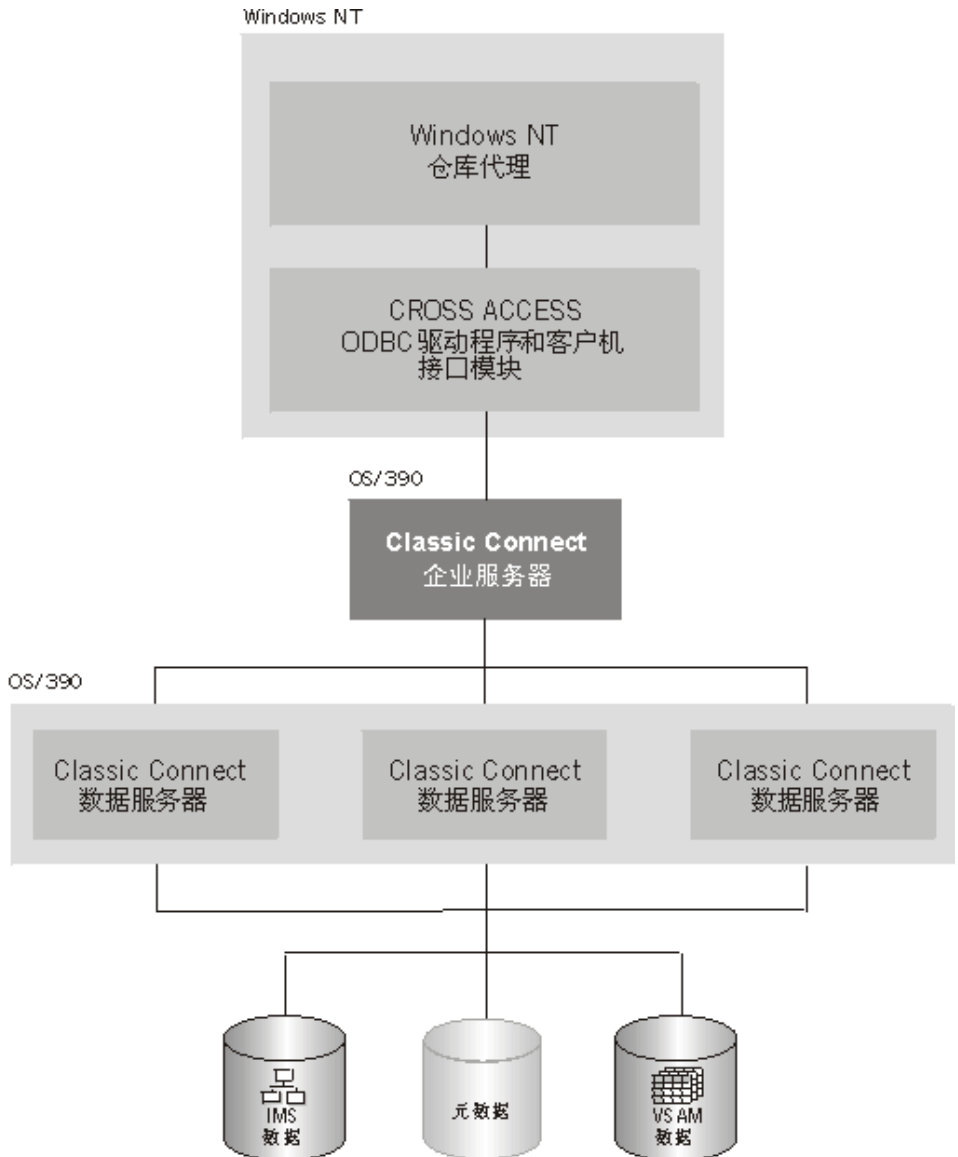


图 29. 装有企业服务器的 Classic Connect 体系结构

象数据服务器一样，企业服务器的连接处理程序负责收听客户机连接请求。然而，当接收到连接请求时，企业服务器不将该请求转送到查询处理器任务进行处理。而是将连接请求转送到数据源处理程序 (DSH)，然后转送到数据服务器用于处

在“数据仓库中心”中使用 Classic Connect

理。企业服务器维护客户机应用程序和目标数据服务器之间的端对端连接。它负责在客户机应用程序和数据服务器之间发送和接收消息。

企业服务器还用于执行负载平衡。使用配置参数，企业服务器确定将与它通信的数据服务器的位置，以及那些数据服务器是否正在与该企业服务器相同的平台上运行。

若没有活动的实例，企业服务器可自动启动本地数据服务器。当当前活动的实例已达到它们可服务的最大并行用户数，或当前活动的实例都忙时，企业服务器还可启动本地数据服务器的附加实例。

数据映射器

Classic Connect 非关系数据映射器是一个基于 Microsoft Windows® 的应用程序，它可使为非关系数据结构创建逻辑表定义所需的许多任务自动进行。其目标是将单个文件或文件的一部分作为一个或多个关系表查看。在完成映射的同时，必须维护底层数据库或文件的结构完整性。

数据映射器解释定义非关系数据的内容和结构的现存物理数据定义。使用缺省定义的方法，该工具设计用来使管理工作最小化。

通过用现存非关系数据定义（COBOL 记录定义文件）创建元数据语法，数据映射器完成创建非关系数据结构的逻辑表定义。元数据语法用作 Classic Connect 元数据实用程序的输入，以创建定义非关系数据结构如何映射为等效的逻辑表的元数据目录。查询处理器任务使用元数据目录，使访问非关系数据结构中的数据和将非关系数据结构中的数据转换为关系结果集更为便利。

数据映射器导入实用程序用 COBOL 记录定义文件创建初始的逻辑表。可视的指向和单击环境用于使这些初始逻辑表细化，以满足站点和用户特定的需求。可利用数据映射器自动创建的初始表定义，或根据需要定制那些定义。

可创建映射至单个物理文件或数据库的多个逻辑表。例如，站点可选择创建全部映射至一个员工 VSAM 文件的多个表定义。一个表供部门经理使用，他们需要访问有关他们部门的员工的信息；另一个表供 HR 经理使用，他们有权访问所有员工信息；另一个表供 HR 职员使用，他们有权访问非机密信息；还有一个供员工自己使用，他们可查询有关他们自身利益结构的信息。根据用户的需要定制这些表定义不仅对最终用户有益，而且建议这样做。

第371页的图30显示使用数据映射器的数据管理工作流程。

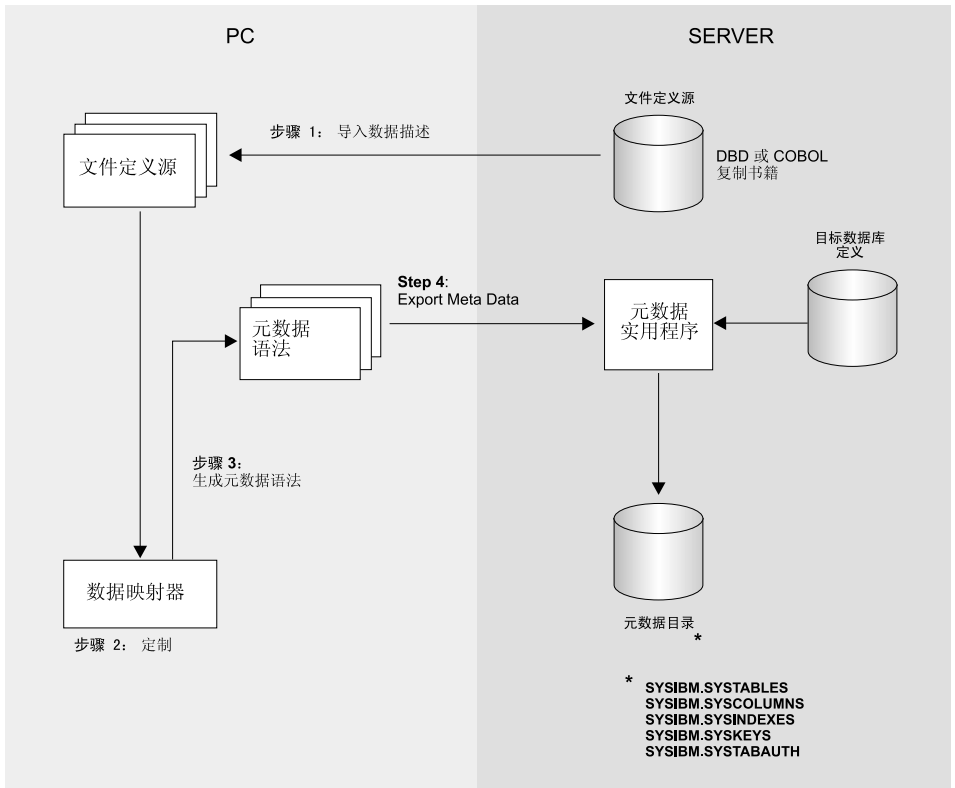


图 30. 数据映射器工作流程

数据映射器包含嵌入的 FTP 支持，使与大型机主机之间的文件传送更为便利。

图30中的步骤描述如下:

1. 将非关系数据的现存描述导入数据映射器。COBOL 记录定义文件和 IMS数据库定义 (DBD) 均可导入数据映射器。

数据映射器用 COBOL 记录定义文件信息创建缺省逻辑表定义。若这些缺省表定义可接受，则跳过下一步直接转至步骤3。

2. 根据用户需要细化或定制缺省表定义。例如，导入 VSAM 客户主文件的记录布局创建缺省 Customer_Table。还可用原始数据创建两个附加的表：
 - Marketing_Customer_Table, 仅包含市场营销部门所需的那些数据项
 - Service_Customer_Table, 仅包含支持代表所需的那些数据项
3. 生成将供 OS/390 上的元数据实用程序使用的实际元数据语法。
4. 将逻辑表定义导出到数据库或文件驻留的大型机。这些定义然后用作创建元数据目录的元数据实用程序的输入。

在“数据仓库中心”中使用 Classic Connect

完成这些步骤之后，就可与您的工具和应用程序一起使用 Classic Connect 操作组件访问非关系数据。

建立环境

本节概述在 Classic Connect 和“数据仓库中心”之间建立集成的需求。

硬件和软件需求

集成需要下列软件：

- DataJoiner Classic Connect 版本 2 发行版 1
- DB2 通用数据库版本 7

任选地，可使用 DataJoiner Classic Connect 数据映射器生成元数据语法。可从下列 Web 站点获取数据映射器：

<http://www.software.ibm.com/data/datajoiner/news.html#newcxa>

安装和配置的先决条件产品

完成表53中概述的任务，以在 Classic Connect 和“数据仓库中心”之间建立集成。有关详情，参见每个任务中列示的文档。

表 53. 安装和配置任务摘要

任务	内容	位置
了解集成	什么是 Classic Connect?	第363页的『什么是 Classic Connect?』
	概念和术语	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第二章

表 53. 安装和配置任务摘要 (续)

安装和配置数据服务器	系统需求和计划	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第三章
	在 OS/390 上安装 Classic Connect	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第四章
	数据服务器安装和验证过程	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第六章
	数据服务器设置介绍	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第六章
	配置 OS/390 和 Windows NT 之间的通信协议	第374页的『配置 OS/390 和 Windows NT 之间的通信协议』
安装和配置客户机工作站	配置 Windows NT 客户机	第382页的『配置 Windows NT 客户机』
	定义代理站点	第13页的『定义代理站点』
使用 IMS 或 VSAM 仓库源	映射非关系数据和创建查询	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第十三章和 <i>DataJoiner Classic Connect: Data Mapper Installation and User's Guide</i>
	优化	<i>DataJoiner Classic Connect: Installation, Configuration, and Reference Guide</i> 的第十四章
	定义仓库源	第95页的『在“数据仓库中心”中定义非 DB2 数据库仓库源』
从“Visual Warehouse 主机适配器”迁移	从“Visual Warehouse 主机适配器”迁移到 Classic Connect	第393页的『从“Visual Warehouse 主机适配器”迁移到 Classic Connect』

配置 OS/390 和 Windows NT 之间的通信协议

Classic Connect 支持 TCP/IP 和 SNA LU 6.2 (APPC) 通信协议，以在 Visual Warehouse 代理和 Classic Connect 数据服务器之间建立通信。第三方协议“跨内存”用于 OS/390 上的本地客户机通信。

本章描述在可配置 Classic Connect 之前需要对 TCP/IP 和 SNA 通信协议进行的修改，包括以下各节：

- 通信选项
- 配置 TCP/IP 通信协议
- 配置 LU 6.2 通信协议

通信选项

Classic Connect 支持下列通信选项：

- 跨内存
- SNA
- TCP/IP

跨内存

“跨内存”应该用于配置本地 OS/390 客户机应用程序 (DJXSAMP) 访问数据服务器。不象对 SNA 和 TCP/IP 那样，使用 OS/390 跨内存接口没有设置需求。此接口使用 OS/390 数据空间和 OS/390 记号命名服务在客户机应用程序和数据服务器之间进行通信。

每个“跨内存”数据空间可最多支持 400 个并行用户，虽然由于资源限制此数目在实际情况中可能要低一些。要在数据服务器上支持超过 400 个用户，配置多个连接处理程序服务，每个具有不同的数据空间名。

以下示例说明通信复合地址字段：

```
XM1/DataSpace/Queue
```

因为不需要修改任何跨内存配置设置，所以在此不详细讨论此协议。

SNA

SNA 是一个更完善的协议并支持硬件压缩，可大大减少线路上实际传送的数据量。美中不足的是，与使用 TCP/IP 相比，通常使用 SNA 的基础结构需求更高，设置时间更长。

TCP/IP

单个 TCP/IP 连接处理程序服务最多可对 255 个并行用户提供服务。取决于您的特定 TCP/IP 子系统，可能有进一步的限制。

可在指定的端口号上创建多个对话。端口上传送的对话数是将支持的并行用户数加一个收听对话，连接处理程序使用收听对话接受来自远程客户机的连接。若正使用的 TCP/IP 实现要求您指定可在单个端口上传送的对话数，必须确保已定义正确的对话数。未那样做将导致在客户机应用程序尝试与数据服务器连接时出现连接故障。

配置 TCP/IP 通信协议

本节描述在 OS/390 系统和 Windows NT 系统上配置 Classic Connect 的 TCP/IP 通信接口所必须执行的步骤。本节还包括设计用来说明 TCP/IP 参数关系的 TCP/IP 计划模板和工作表。

与 Classic Connect 一起工作的 TCP/IP CI 有两种类型：IBM 的 TCP/IP 和“Berkeley 套接字”。您的配置可能因使用的 TCP/IP CI 类型不同而各异。若您的站点不使用 TCP/IP，则前进至 第378页的『配置 LU 6.2 通信协议』。

两个接口均允许 Classic Connect 与 OS/390 TCP/IP 堆栈通信。“Berkeley 套接字”允许使用主机和服务名，而 IBM 的 TCP/IP 需要数字 IP 地址和端口号。

“Berkeley 套接字”可使用本地主机文件代替调用 DNS。然而，“Berkeley 套接字”和 IBM 的 TCP/IP 都需要 TCP/IP 地址空间名。

在 OS/390 上配置 TCP/IP

Classic Connect 的 TCP/IP 与 IBM 的和 Interlink 的“Berkeley 套接字”TCP/IP 兼容。本节描述如何使用 IBM 的 TCP/IP 配置 Classic Connect。有关 IBM 的或 Interlink 的 TCP/IP 的详情，参考适当产品的文档。

IBM 和 Interlink 支持“Berkeley 套接字”。“Berkeley 套接字”版本在 DJXDSCF 成员中需要一个称作 TASK PARAMETER 的附加参数，它标识 Interlink 子系统名并标识 IBM 的配置数据集的位置。在配置数据集内，用户必须指定用于启动 TCP/IP 地址空间名的启动任务过程的名称，还可指定 TCP/IP DNS IP 地址。若未传递环境变量，则对地址空间名和标准配置文件的高级限定符 (hlq) 使用缺省值 TCPIP:

- hlq.TCPIP.DATA
- hlq.ETC.HOSTS
- hlq.ETC.PROTOCOLS
- hlq.ETC.SERVICES
- hlq.ETC.RESOLV.CONF

在“数据仓库中心”中使用 Classic Connect

Classic Connect 使用搜索次序查找数据集，不管 Classic Connect 是否设置 hlq。

为安装 Classic Connect 的 OS/390 系统确定下列值，并在第378页的图31 的工作表部分中输入这些值。

IP 地址或主机名

获取 OS/390 系统的主机名或 IP 地址。

使用主机名需要有已配置的本地 HOSTS 文件或域名服务器可用。若涉及域名服务器，则需要一些额外开销来将 HOST 名称解析为正确的 IP 地址。然而，为了可靠和便于将来对配置进行更改，建议在远程客户机配置文件中使用主机名。

使用主机名还使在环境变化时更改 IP 地址更容易。若使用主机名，则不必频繁重新配置数据服务器 / 远程客户机。可将 Classic Connect 当机，网络管理员可更改 OS/390 和客户机域名服务器中的主机名的 IP 地址。当重新启动数据服务器时，它将自动在新的 IP 地址上收听来自远程客户机的连接请求。当远程客户机与数据服务器连接时，它将自动使用已指定给主机名的新 IP 地址，无需更改 Classic Connect 配置文件。

对于 IBM 的 TCP/IP，确定安装 Classic Connect 的主机的 IP 地址或主机名。若正在没有负载的设备上运行 OS/390 TCP/IP，则指定 OS/390 映象上的 TCP/IP 堆栈的地址，而不是没有负载的网关的 IP 堆栈的地址。

端口号

对将从客户机访问的每个数据服务器，获取大于 1024 的唯一端口（套接字）号。

端口号不能与已定义为供任何其他应用程序使用的任何端口相同，包括同一 OS/390 系统上的其他 Classic Connect 数据服务器使用的端口。使用非唯一的端口号会导致数据服务器不能启动。要确定是否已将一个端口号指定给另一应用程序，从“假脱机显示和搜索设施” (SDSF) 日志发出如下命令：

```
TSO NETSTAT SOCKETS
```

因为某些站点限制对特定应用程序使用某些端口号，所以还应与网络管理员联系以确定您选择的端口号是否唯一和有效。

任选的，可替换给对您的系统定义的端口号指定的服务名。

IBM 的 TCP/IP 的服务名、地址和调整值包含在一系列数据集中。

- hlq.TCPIP.DATA
- hlq.ETC.HOSTS

- hlq.ETC.PROTOCOLS
- hlq.ETC.SERVICES
- hlq.ETC.RESOLV.CONF

其中“hlq”表示这些数据集的高级限定符。可接受缺省高级限定符 TCPIP, 或可特别为 Classic Connect 定义高级限定符。

当已确定这些值时, 使用第378页的图31完成您的 TCP/IP 通信的 OS/390 配置。

在 Windows NT 上配置 TCP/IP

必须配置 Windows NT 机器以查找 OS/390 上的数据服务器。

1. 解析客户机上的主机地址。

若在客户机配置文件中正使用 IP 地址, 可跳过此步骤。

客户机工作站必须知道它正尝试连接的主机服务器的地址。有两种方法解析主机地址:

- 通过在网络上使用名称服务器。这是建议的方法。有关将 TCP/IP 配置成使用名称服务器的信息, 参见您的 TCP/IP 文档。

若已在网络上使用名称服务器, 前进到步骤2。

- 通过在本地的 HOSTS 文件中指定主机地址。在 Windows NT 客户机上, HOSTS 文件位于 %SYSTEMROOT%\SYSTEM32\DRIVERS\ETC 目录中。

在客户机上为服务器的主机名给 HOSTS 文件添加一项, 如下:

```
9.112.46.200 stplex4a # host address for Classic Connect
```

其中 9.112.46.200 是 IP 地址, stplex4a 是“主机名”。若服务器驻留在与客户机相同的 Internet 域中, 则此名称可以是一个平面主机名。若服务器不在相同的域中, 该名称必须是全限定域名, 如 stplex4a.stl.ibm.com, 其中 stl.ibm.com 是域名的示例。

注:

- a. 必须用注释 (# comment) 结束最后一行, 或在该行的末尾按 Enter 键插入行结束字符。
 - b. 有关解析主机地址的特定信息, 应参考您的 TCP/IP 产品的文档。
- #### 2. 更新客户机上的 SERVICES 文件。

若在客户机配置文件中正使用端口号, 可跳过此步骤。

必须将以下信息添加到客户机上的 SERVICES 文件中, 以获取 TCP/IP 支持:

```
ccdatser 3333 # CC data server on stplex4a
```

SERVICES 文件位于 %SYSTEMROOT%\SYSTEM32\DRIVERS\ETC 目录中。

在“数据仓库中心”中使用 Classic Connect

TCP/IP 通信模板和工作表

图31的左边提供 OS/390 配置的 TCP/IP 值的示例集；在以后步骤中的数据服务器和客户机配置期间将使用这些值。使用该图的右边作为输入自己的值的模板。

TCP/IP-MVS	
MVS 主机名	not used *
或 IP 地址	9.112.46.200
子系统名	tcp
数据服务器的端口号 或服务名	3333

* 此示例使用 IP 地址
而不是主机名

TCP/IP-MVS	
MVS 主机名	
或 IP 地址	
子系统名	
数据服务器的端口号 或服务名	

图 31. TCP/IP 通信模板和工作表

配置 LU 6.2 通信协议

本节描述为配置 Classic Connect 的 LU 6.2 (SNA/APPC) 通信，在 OS/390 系统和 Windows NT 系统上必须确定的值和必须执行的步骤。

需求： 对于 Classic Connect 和 DataJoiner Windows NT 版之间的 APPC 联网，需要 Microsoft SNA Server 版本 3.0 带服务包 3 或更高版本。

本节中的信息特定于 Microsoft SNA Server 版本 3.0。有关配置 Microsoft SNA 服务器简要表的详情，参见适当的产品文档。本节还包括通信模板和工作表，它们是为理解 OS/390 和 Windows NT 上的 LU 6.2 参数关系和辅助您进行 LU 6.2 配置设计的。

在 OS/390 上配置 LU 6.2

若将使用 LU 6.2 从 DataJoiner 访问 Classic Connect，需要在 OS/390 系统上配置 VTAM[®] 表定义，这些表定义包括：

- “方式”表项
- 应用程序 ID

“应用程序 ID”必须对数据服务器唯一。使用非唯一的值会导致数据服务器不能启动。

不象 TCP/IP 那样，可对通过 SNA 网络的传输层传送的数据指定包大小。然而，应由网络管理员做出此决定，因为它涉及考虑复杂的路由和机器 / 节点能力。通常，通信媒体或管道的带宽越大，RU 大小也越大。

在 Windows NT 上配置 LU 6.2

本节说明要对 Windows NT 客户机使用 SNA LU 6.2 协议，需要配置的值。

对于每个 Windows NT 系统，配置下列值：

- SNA 服务器特性简要表

- SNA 服务器的连接

本示例假设已安装 SNA DLC 802.2 链接服务。与网络管理员联系以获取本地和远程节点信息。

- 本地 APPC LU 简要表

LU 名和网络必须与 SNA 服务器简要表的“连接”中的本地节点值匹配。LU 必须为独立的 LU 类型。

- 远程 APPC LU 简要表

对于将访问的每个 Classic Connect 数据服务器或企业服务器，必须具有这些简要表中的一个。它必须支持并行对话。它将与第378页的『在 OS/390 上配置 LU 6.2』中讨论的 OS/390 上 VTAM 表定义中配置的“应用程序 ID”匹配。

- APPC 方式

参见第380页的图32中的 CX62R4K 方式。它的最大 RU 大小为 4096。

- CPIC 符号名

对于将访问的每个 Classic Connect 数据服务器和企业服务器，必须具有这些名称中的一个。此简要表中引用的“TP 名”必须对给定 OS/390 系统上的 Classic Connect 数据服务器或企业服务器唯一。

在输入这些值之后，保存配置，然后停止并重新启动您的 SNA Server。当 SNA Server 和“连接”（本例中分别为 OTTER 和 SNA OS/390）“活动”时，该连接可用于测试应用程序。

APPC 通信模板和工作表

第380页的图32提供 VTAM[®] 和 SNA 值的示例集，已包括进来供您参考。使用第381页的图33（它是图32没有填写值的副本）作为工作表，可在其中输入特定于您的 LU 6.2 配置的 VTAM 和 SNA 值。将需要在此工作表中输入的值来完成后续章节中的配置步骤。

在“数据仓库中心”中使用 Classic Connect

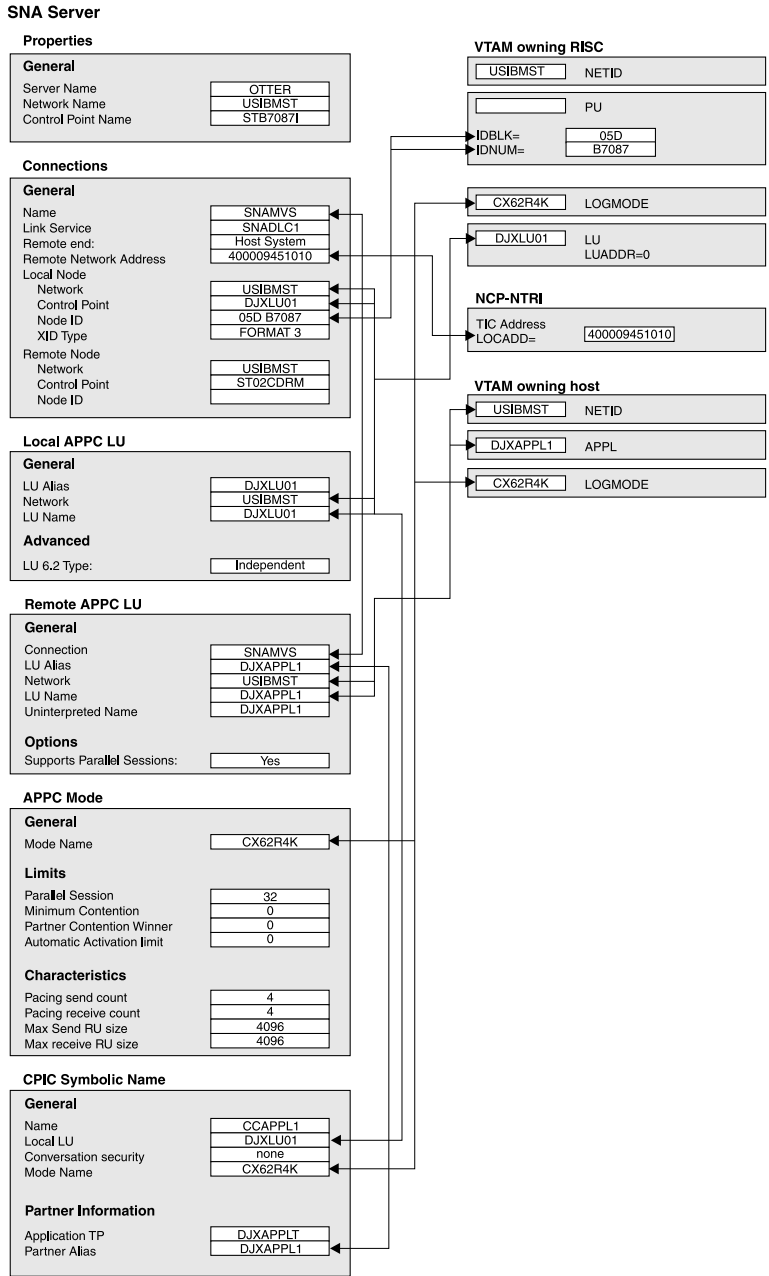


图 32. LU 6.2 配置模板

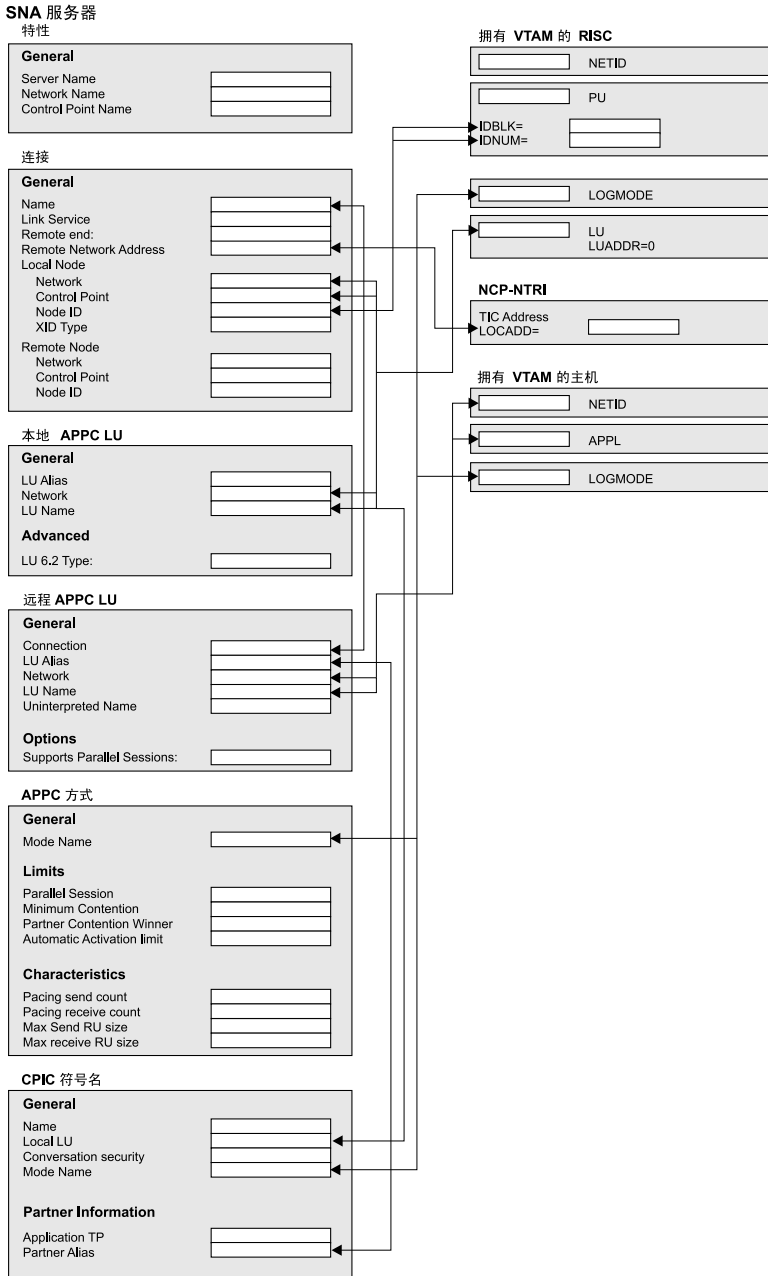


图 33. LU 6.2 配置工作表

配置 Windows NT 客户机

本节描述如何安装“Classic Connect 驱动程序”组件和 CROSS ACCESS ODBC 驱动程序，以及如何使用 CROSS ACCESS ODBC 驱动程序来配置数据源。

安装“Classic Connect 驱动程序”组件

安装 Classic Connect ODBC 驱动程序:

1. 将“仓库管理器”CD-ROM 插入 CD-ROM 驱动器中。
2. 在启动板中单击**安装**。
3. 在“选择产品”窗口中，验证是否选择了 **DB2 仓库管理器**复选框，然后单击下一步。
4. 在“选择安装类型”窗口中，选择**定制**，然后单击下一步。
5. 在“选择组件”窗口中，选择**Classic Connect 驱动程序**，然后选择**仓库代理**，清除所有其他的复选框，然后单击下一步。
6. 在“开始复制文件”窗口中，复查选择。
如果要更改任何选择，单击**上一步**按钮，返回可以更改选择的窗口。
7. 单击下一步以开始安装。

安装 CROSS ACCESS ODBC 驱动程序

通过执行“DB2 仓库管理器版本 7”的定制安装来安装 CROSS ACCESS ODBC 驱动程序，然后选择“Classic Connect 驱动程序”组件。驱动程序不会作为“DB2 仓库管理器”的典型安装的一部分来安装。

CROSS ACCESS ODBC 驱动程序将安装在 SQLLIB 目录的 ODBC32 子目录中。安装完成后，必须手工添加驱动程序的路径（例如，C:\Program Files\SQLLIB\ODBC32）到 PATH 系统环境变量中。如果已安装了另一版本的 CROSS ACCESS ODBC 驱动程序，应将 ...\SQLLIB\ODBC32\ 路径放在另一版本的路径之前。操作系统将使用包含 CROSS ACCESS ODBC 驱动程序的路径中的第一个目录。

配置数据源

CROSS ACCESS ODBC 数据源是使用“ODBC 管理员”注册和配置的。通过此实用程序维护对每个数据源唯一的配置参数。

可在单个系统上定义许多数据源。例如，单个 IMS 系统可具有称作 MARKETING_INFO 的数据源和称作 CUSTOMER_INFO 的数据源。每个数据源名应提供数据的唯一描述。

配置先决条件

在尝试配置 ODBC 驱动程序之前，必须具备下列信息。若丢失此信息中的任何部分，与系统管理员联系。

- 要在“ODBC 管理员”中定义的 CROSS ACCESS 数据源的名称
- 若正使用 TCP/IP:
 - 运行数据服务器的主机系统的 IP 地址。
 - 在数据服务器的 TASK INFO ENTRY 参数中指定给 TCP/IP 连接处理程序的端口号。
- 若正使用 LU 6.2:
 - SNA 服务器中定义的出站端信息记录（SIR 出站）的名称。
 - 主机组件的数据压缩级别。

在配置 ODBC 驱动程序之前，确信 Windows 客户机已设置想要使用的连接处理程序，TCP/IP 或 LU 6.2。

对于 Classic Connect 和 DataJoiner Windows NT 版之间的 APPC 联网，需要 Microsoft SNA Server 版本 3 带服务包 3 或更高版本。

指定数据源

为当前安装的所有 ODBC 驱动程序定义的数据源列示在“ODBC 数据源管理员”窗口中。从此窗口，您可以：

- 添加和配置数据源。
- 修改数据源的配置。
- 删除数据源。

要打开“ODBC 数据源管理员”窗口：

1. 在桌面上单击**开始**，并单击**设置**。
2. 单击**控制面板**。
3. 打开 ODBC 图符。“ODBC 数据源管理员”窗口打开。

在“数据仓库中心”中使用 Classic Connect

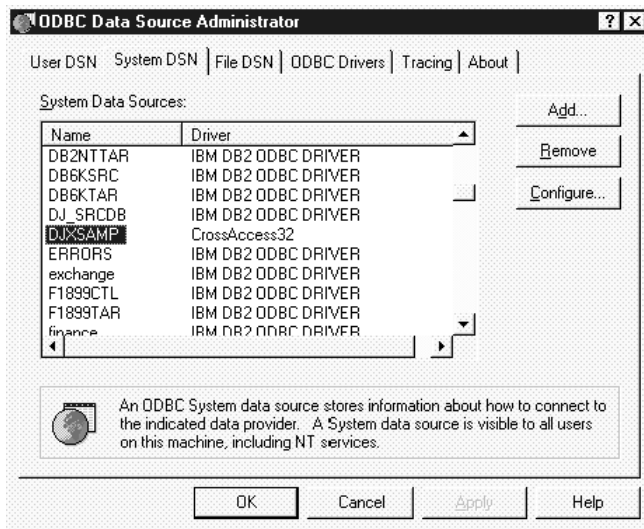


图 34. “ODBC 数据源管理员” 窗口

此窗口在“系统 DSN” 页上显示数据源和驱动程序的列表。

添加和配置数据源

要添加和配置数据源：

1. 打开“ODBC 数据源管理员” 窗口。
2. 在“系统 DSN” 页上，单击**添加**。“创建新的数据源” 窗口打开。



图 35. “创建新的数据源”窗口

- 单击 CrossAccess32。
- 单击完成。“CROSS ACCESS 通信协议”窗口打开。
- 选择要对正配置的数据源使用的通信接口。

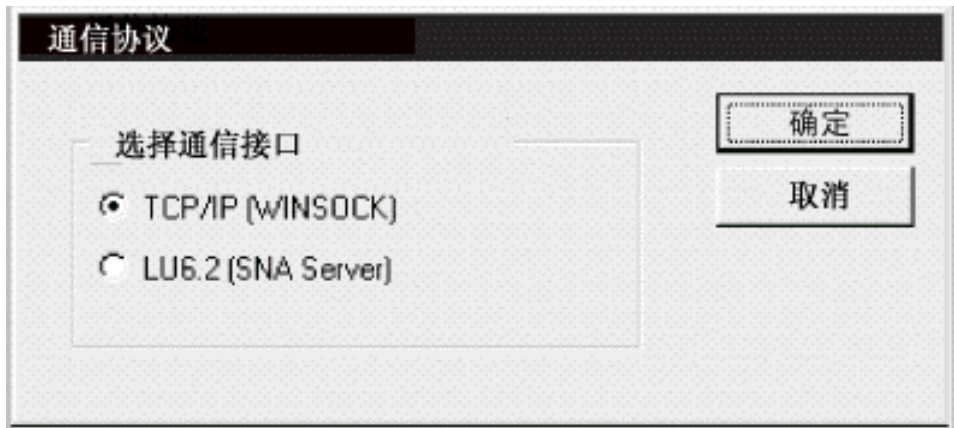


图 36. “通信协议” 窗口

6. 单击**确定**。“CROSS ACCESS ODBC 数据源配置”窗口打开。

在此窗口中，可输入新数据源的参数，或修改现存数据源的参数。许多参数必须与在服务器配置中指定的值匹配。若不知道这些参数的设置，与 Classic Connect 系统管理员联系。

在此窗口中输入的参数因是在使用 TCP/IP 还是 LU 6.2 通信接口而各异。

- 对于 TCP/IP 通信接口，参见『配置 TCP/IP 通信』。
- 对于 LU 6.2 通信接口，参见第389页的『配置 LU 6.2 通信』。

配置 TCP/IP 通信

使用“CROSS ACCESS ODBC 数据源配置”窗口以：

- 命名数据源。
- 配置 TCP/IP 通信设置。
- 指定必要的授权。

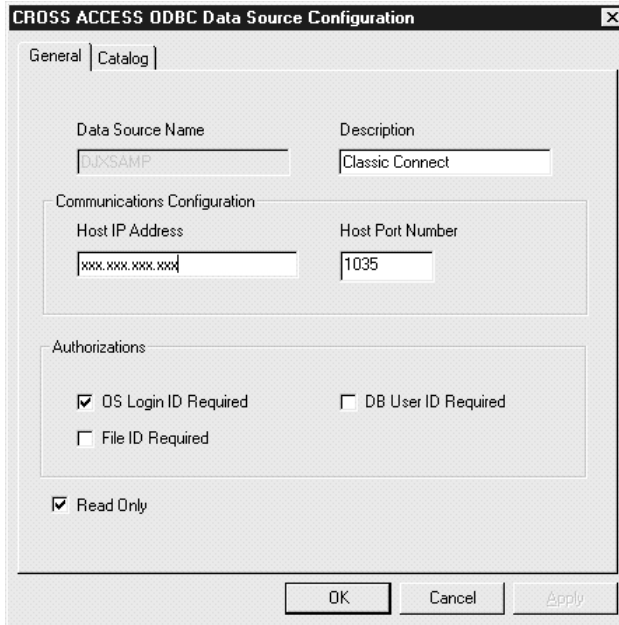


图 37. TCP/IP 的“CROSS ACCESS ODBC 数据源配置”窗口

要配置 TCP/IP 通信:

1. 在**数据源名**字段中输入数据源的名称。此名称必须与数据服务器配置文件的 QUERY PROCESSOR SERVICE INFO ENTRY 的“字段 2”匹配。(有关示例, 参见 *DataJoiner Classic Connect: Installation, Configuration, and Reference Guide* 的第九章。)
2. 在**描述**字段中输入数据源的简要描述。
3. 在**主机 IP 地址**字段中输入数据服务器的 IP 地址。此字段以主机名或点分十进制表示法指定安装数据服务器的主机的 IP 地址。
4. 在**主机端口号**字段中, 输入给主机组件 TCP/IP 通信指定的端口号(套接字)。此号码必须与数据服务器配置文件的 TCP/IP SERVICE INFO ENTRY 的“字段 10”匹配。(有关示例, 参见 *DataJoiner Classic Connect: Installation, Configuration, and Reference Guide* 的第九章。)
5. 选择下列复选框中的一个或多个:
 - **必需的 OS 登录标识**。选择此框以提示您输入登录到操作系统的用户标识和密码。
 - **必需的 DB 用户标识**。选择此框以提示您输入登录到数据库系统(如 DB2 或 Sybase)的用户标识和密码。

在“数据仓库中心”中使用 Classic Connect

- **必需的文件标识。** 选择此框以提示您输入访问数据库所需的文件标识和密码。对于某些数据库（如 Model 204），文件标识和密码是必需的。

6. 指定数据源是否具有更新能力。缺省值是只读访问。

设置数据库目录选项： 使用“目录”页执行下列任务：

- 指定目录表选项。
- 指定游标管理。

要设置数据库目录选项：

1. 单击“CROSS ACCESS ODBC 数据源配置”窗口中的目录选项卡。



图 38. “CROSS ACCESS ODBC 数据源配置”窗口

2. 在目录所有者姓名字段中输入数据库目录所有者的姓名。
3. 若想要 ODBC 驱动程序在应用程序发出 CLOSE CURSOR 调用之后自动发出 COMMIT 调用，选择在关闭之后落实复选框。在某些数据库系统上，在打开游标期间将出现资源锁定。仅可由 COMMIT 调用和 CLOSE CURSOR 调用释放这些锁定。

若保留此框为空白，则无需发出 COMMIT 调用就可释放游标。

4. 单击**确定**。

保存 TCP/IP 通信信息。

配置 LU 6.2 通信

使用“CROSS ACCESS ODBC 数据源配置”窗口以：

- 标识数据源。
- 配置 LU 6.2 通信设置。
- 指定必要的授权。

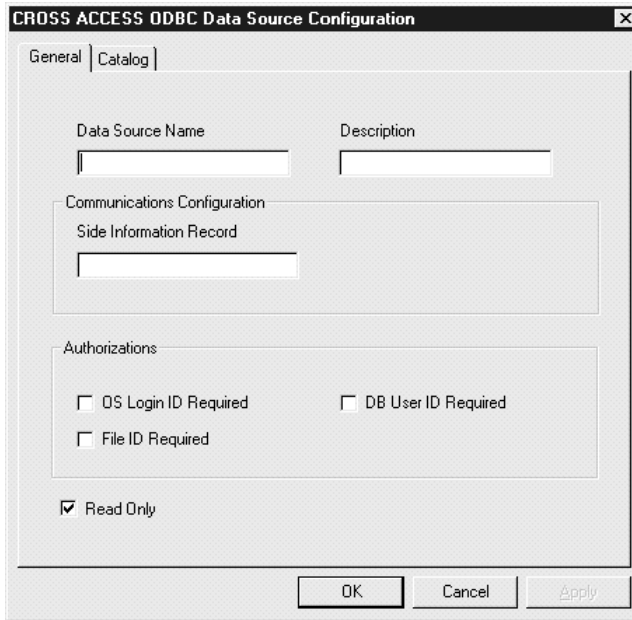


图 39. LU 6.2 的“CROSS ACCESS ODBC 数据源配置”窗口

要配置 LU 6.2 通信：

1. 在**数据源名**字段中输入数据源的名称。此名称必须与数据服务器配置文件的 QUERY PROCESSOR SERVICE INFO ENTRY 的“字段 2”匹配。（有关示例，参见 *DataJoiner Classic Connect: Installation, Configuration, and Reference Guide* 的第九章。）
2. 在**描述**字段中输入数据源的简要描述。
3. 在**补充信息记录**字段中输入补充信息记录（SIR）名。
SIR 名指 SNA 服务器中定义的补充信息记录（在第381页的图33中也称作 CPIC 符号名）。此 SIR 必须包括代表数据服务器的配置参数。
4. 选择下列复选框中的一个或多个：
 - **必需的 OS 登录标识**。选择此框以提示您输入登录到操作系统的用户标识和密码。

在“数据仓库中心”中使用 Classic Connect

- **必需的 DB 用户标识。** 选择此框以提示您输入登录到数据库系统（如 DB2 或 Sybase）的用户标识和密码。
 - **必需的文件标识。** 选择此框以提示您输入访问数据库所需的文件标识和密码。对于某些数据库（如 Model 204），文件标识和密码是必需的。
5. 清除**只读**复选框以指示该数据源具有更新能力。缺省值是只读访问。

设置数据库目录选项： 使用“目录”页执行下列任务：

- 指定目录表选项。
- 指定游标管理。

要设置数据库目录选项：

1. 单击“配置”窗口中的**目录**选项卡。

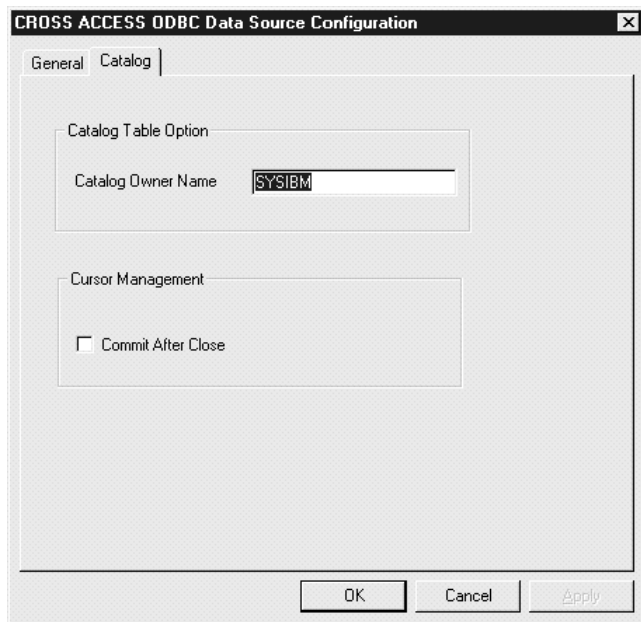


图 40. “数据库目录选项”窗口

2. 在**目录所有者姓名**字段中输入数据库目录所有者的姓名。
3. 若想要 ODBC 驱动程序在应用程序发出 CLOSE CURSOR 调用之后自动发出 COMMIT 调用，选择**在关闭之后落实**复选框。在某些数据库系统上，在打开游标期间将出现资源锁定。仅可由 COMMIT 调用和 CLOSE CURSOR 调用释放这些锁定。

若保留此框为空白，则无需发出 COMMIT 调用就可释放游标。

4. 单击**确定**。

保存 LU 6.2 通信信息。

配置 ODBC 驱动程序

CROSS ACCESS ODBC 驱动程序维护所有 CROSS ACCESS 数据源的公共配置参数集。在“CROSS ACCESS 管理员”窗口中执行对这些参数的配置。下列步骤显示如何配置 ODBC 驱动程序参数。

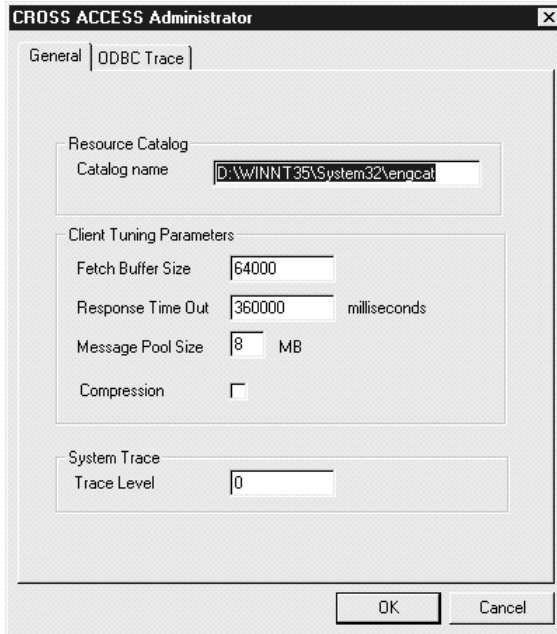


图 41. “CROSS ACCESS 管理员”窗口的“常规”页

1. 在“CROSS ACCESS 管理员”窗口的“常规”页中，在**目录名**字段中输入语言目录的全路径名。此值是必需的。

语言目录包含特定语言中的消息，并由 CROSS ACCESS 配置文件内包含的文件指出。

2. 任选地，在**取装缓冲区大小**字段中输入 CROSS ACCESS 的大小。

此值通过控制单个取装请求中返回的数据量调整消息分块。系统将尽可能多的数据行装进指定大小的取装缓冲区。例如，若取装缓冲区设置为 10,000 字节且每行 2,000 字节，则每个取装请求系统装进 5 行。若返回单个行但指定的缓冲区装不下该行，则取装缓冲区内部增加以便能装下该行数据。要关闭消息分块，将此参数设置为 1。该值必须为 1 至 64,000，包括 1 和 64,000 在内。缺省值是 10,000。

在“数据仓库中心”中使用 **Classic Connect**

3. 任选地，在**响应时间超时**字段中，输入此服务将等待期望的响应的最大时间量（在终止连接前）。

可指定下列时间间隔：

- nMS = 毫秒数
- nS = 秒数
- nM = 分钟数

指定 0 和 1000MS、60S 或 60M 之间的值。缺省值为 6M。

4. 在**消息池大小**字段中，输入用于所有内存分配的内存大小。此值是必需的。以字节为单位指定该数目。实际的可工作最大值应设置为比堆大小小 2 MB。若指定的值小于 1 MB，则使用 1 MB。若可获取的存储器量小于指定的值，则获取最大可用量。最大允许值为 2,097,152,000 字节 (2 GB)。缺省值为 1,048,575 字节 (1 GB)。
5. 任选地，指定对所有任何任务（内部的和外部的）之间和启动程序与 ODBC 驱动程序之间的数据传送，是打开还是关闭数据压缩。在**压缩**字段中输入下列其中一个值：

- 1 打开数据压缩。
- 2 关闭数据压缩。

缺省值为 2（关）。

6. 任选地，在**跟踪级别**字段中，输入对应 ODBC 驱动程序写入跟踪日志的信息量的值。该值必须为 0 和 4 之间的整数，其中：
 - 0 不记录跟踪信息。
 - 1 记录最少的跟踪信息。
 - 4 记录最多的跟踪信息。

这种跟踪不同于 ODBC 跟踪；后者特定于 Visual Warehouse 使用的 ODBC 驱动程序。

7. 单击“CROSS ACCESS 管理员”窗口中的 **ODBC 跟踪**选项卡。

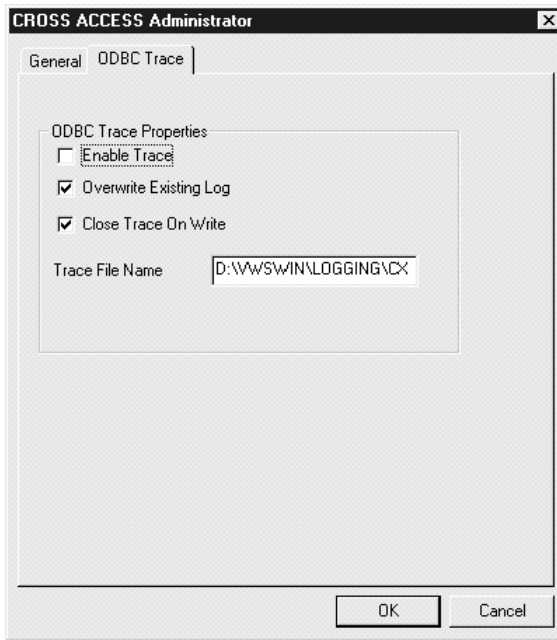


图 42. “CROSS ACCESS 管理员”窗口的“ODBC 跟踪”页

8. 选择允许跟踪框以生成 ODBC 跟踪。
9. 选择覆盖现存日志框覆盖现存日志跟踪。
10. 若想要驱动程序在写入每条消息之后关闭跟踪日志，选择 写入时关闭跟踪框。
11. 在跟踪文件名字段中，指定跟踪文件的名称。
若未指示目录，将在对 ODBC 数据源发出查询的工具的相应“程序文件”目录的子目录中创建跟踪文件。
12. 单击确定。

从“Visual Warehouse 主机适配器”迁移到 Classic Connect

要从“Visual Warehouse V1.3 主机适配器”迁移到 Classic Connect:

1. 编辑“元数据”实用程序。在 SDJXSAMP 成员 DJXMETAU 中可找到样本“元数据实用程序”JCL。有关详情，参见 *DataJoiner Classic Connect: Installation, Configuration, and Reference Guide* 的附录 A。
2. 提供有效的作业卡并修改 DJX 高级限定符。
3. 若正映射 IMS 数据，修改 IMS 高级限定符和 DBDLIB DD 语句。
4. 若用 DD 名称引用 VSAM 文件，则在“元数据”实用程序 JCL 中添加具有相同 DD 名称的 DD 语句。

在“数据仓库中心”中使用 **Classic Connect**

5. 第一次运行“元数据”实用程序，需要创建元数据目录文件。与“Visual Warehouse 主机适配器”不同的是，IMS 和 VSAM 元数据可存在于同一目录中。
6. 将与“Visual Warehouse 主机适配器”一起使用的元数据语法文件定义为“元数据”实用程序的输入。
7. 验证对 JCL 中引用的数据集的执行和访问权限。
8. 提交“元数据”实用程序 JCL 供执行并复查输出。
9. 若有必要，安装并配置 CROSS ACCESS ODBC 驱动程序。
10. 修改“Visual Warehouse 主机适配器”设置以使用 Classic Connect:
 - a. 打开“仓库源”笔记本。
 - b. 从**仓库源类型**列表中，选择 **IMS** 或 **VSAM**。
 - c. 选择**数据库**选项卡。
 - d. 将数据源名更改为与配置 CROSS ACCESS ODBC 驱动程序时使用的数据源名相同。
 - e. 关闭“仓库源”笔记本。
11. 运行使用已修改的仓库源的步骤以确保它成功运行。

附录G. “数据仓库中心” 环境结构

本附录中的信息说明“数据仓库中心”是如何构造的，以及它如何与操作系统交互作用。

“数据仓库中心” 环境变量

在表54中，驱动器盘符 C 指示本地驱动器。

表 54. 为 Windows 进行的“数据仓库中心” 环境变量更新

环境变量:	被添加或修改, 以包括:
PATH (用于访问“数据仓库中心” 代码)	C:\Program Files\SQLLIB\BIN
LOCPATH (供“数据仓库中心主机适配器客户机” 使用)	C:\Program Files\SQLLIB\ODBC32\LOCALE
VWS_TEMPLATES	C:\Program Files\SQLLIB\TEMPLATES
VWS_LOGGING	C:\Program Files\SQLLIB\LOGGING
VWSPATH	C:\Program Files\SQLLIB

“数据仓库中心” 注册表更新

将下列值添加到 Windows NT 或 Windows 2000 注册表中的 HKEY_LOCAL_MACHINE\Software\IBM\DB2\DataWarehouseCenter\ServiceParams 中。

值名	值数据
数据库名	<控制 DB 名>
日志目录	<disk:\dir\>
密码	<密码>
限定符	<表限定符>
用户标识	<DB2 用户标识>

“数据仓库中心” 配置文件

对 TCP/IP 配置信息进行了下列更新:

“数据仓库中心” 环境变量

TCP/IP 服务:
vwkernel 11000/tcp
vwlogger 11002/tcp

附录H. Microsoft OLE DB 和“数据事务服务”支持

可以使用“数据仓库中心”从 OLE DB 供应商（作为一个 DB2 数据库视图）来访问数据。可以使用“数据仓库中心”提供的“OLE DB 辅助”向导来创建 DB2 OLE DB 表函数和 DB2 视图，通过它们可访问数据。

“Microsoft 数据转换服务” (DTS) 允许您导入、导出以及转换 OLE DB 源和目标之间的数据，以构建数据仓库和数据集市。DTS 是随 Microsoft SQL Server 一起安装的。所有的 DTS 任务都存储在可使用用于“DTS 数据包”的 Microsoft OLE DB 供应商来运行并访问的“DTS 数据包”中。因为可从 DTS 将数据包作为 OLE DB 源来访问，所以也可以采用对 OLE DB 数据源相同的方式，通过“OLE DB 辅助”向导对 DTS 数据包创建视图。在运行时访问视图时，执行 DTS 数据包而 DTS 数据包中任务的目标表则成为创建的视图。

在“数据仓库中心”中创建视图后就可以使用它，象对其他任何视图一样。例如，可将 DB2 表与 SQL 步骤中的 OLE DB 源连接。在 SQL 步骤中使用新视图时，将调用 DTS 供应商然后 DTS 数据包才运行。

软件需求

具有以下软件需求。

- 使用 DB2 通用数据库 Windows NT 版的版本 7.2 作为仓库目标数据库
- DB2 仓库管理器版本 7.2

限制

- 如果仓库目标数据库是在版本 7.2 之前创建的，则必须在安装 DB2 通用数据库 Windows NT 版的版本 7.2 之后运行 db2upd7 命令。
- 编目仓库源数据库时，数据库别名是在仓库代理站点上编目的。但启动向导时，“数据仓库中心”假设数据库别名也是在客户机工作站上定义的，并将尝试使用仓库源数据库用户标识和密码与该工作站连接。如果连接成功，向导会启动，您可以创建视图。如果连接不成功，会显示警告消息，您必须在向导中编目或选择不同的数据库别名。
- 要标识 DTS 数据包中的特定表，必须在创建目标表的 DataPumpTask 的“工作流特性”窗口的“选项”页上选择 **DSO 行集供应商** 复选框。如果打开了多个 DSO 行集供应商属性，只会使用第一个所选步骤的结果。选择了一个视图后，将返回它的目标表的行集，而您在后续步骤中创建的所有其他行集将忽略。

Microsoft OLE DB 和“数据事务服务”支持

- 为向导输入表名时，使用在该任务的“ workflow 特性”笔记本的“选项”页上显示的步骤名。
- 为向导输入表名时，使用在该任务的“ workflow 特性”笔记本的“选项”页上显示的步骤名。
- DTS 数据包连接字符串具有与 dtsrun 命令相同的语法。

为 OLE DB 表函数创建视图

要为 OLE DB 表函数创建视图：

1. 从“数据仓库中心”窗口中，展开仓库源树。
2. 展开要包含视图的那个仓库源。
3. 右键单击视图文件夹，然后单击 **OLE DB 表函数创建**。

“OLE DB 辅助”向导打开。向导将引导您逐步完成在仓库源数据库中创建新视图的任务。

为 DTS 数据包创建视图

要为 DTS 数据包创建视图：

1. 从“数据仓库中心”窗口中，展开仓库源树。
2. 展开要包含视图的那个仓库源。
3. 右键单击视图文件夹，然后单击 **DTS 数据包的 Microsoft OLE DB 供应商**。

“OLE DB 辅助”向导打开。向导将引导您逐步完成在仓库源数据库中创建新视图的任务。

要获取关于 DTS 的更多信息，请参阅 Microsoft Platform SDK 2000 文档，其中包括了如何构建向导连接到 DTS 供应商所需的供应商字符串的详细描述。

附录I. 创建 i2 步骤

通过 i2 TradeMatrix BPI 连接器，可以调度对 i2 商务模型、分布式数据集市和 OLAP 立方的更新。但必须安装“DB2 仓库管理器连接器”。要获取更多信息，请参阅《DB2 仓库管理器安装指南》。

i2 TradeMatrix Business Process Intelligence (BPI) 产品是 i2 Technologies 公司的一种企业范围内的分析和报告解决方案。通过 BPI，您可以使用“i2 TradeMatrix BPI 体系结构”来创建维、量度（事实）、层次结构、级别等的“联机分析处理技术” (OLAP) 定义。这些定义存储在“i2 商务库”的公司模型中。您再从存储在“i2 商务库”的公司模型中的定义来创建商务模型。可以部署商务模型，从而在分布式数据集市 (DDM) 中创建星型模式维和事实表。从这些星型模式表中，可以创建 OLAP 立方结构以进行报告和分析。

通过 i2 TradeMatrix BPI 连接器，您可以：

- 将更改导入到已部署的 BPI 商务模型的元数据中。
- 更新商务模型的 BPI 分布式数据集市和 i2 OLAP 立方。

i2 TradeMatrix BPI 连接器生成一些步骤，用于管理从现有的输入文件（以标准 BPI 逗号分隔的变量 (CSV) 格式）装入 i2 维和事实表，以及装入 OLAP 立方。在导入 i2 源的元数据后，数据仓库中心为 i2 装入步骤创建并填充过程文件夹。因为 i2 装入步骤是由数据仓库中心控制的，所以您可以控制何时运行它们以及运行频率，或按需要运行它们。

i2 TradeMatrix BPI 连接器在 Microsoft Windows NT 或 Windows 2000 上运行。
i2 TradeMatrix 商务模型源可在任何受支持的平台上。

导入 i2 源的元数据

要将 i2 源的元数据导入数据仓库中心：

1. 右键单击**仓库文件夹**，单击**导入元数据**，然后单击 **i2**。
只可以将元数据从部署的 i2 商务模型导入到数据仓库中心。
2. 在“导入元数据”窗口的“连接”页中，输入包含“i2 TradeMatrix BPI 商务库”的 DB2 数据库的数据库名、“i2 商务库”的模式名以及用于连接数据库的用户标识和密码。
3. 在“导入元数据”窗口的“目标”页上，选择代理站点、主题区、过程、仓库目标和目标的模式名。选择想从它导入元数据的部署商务模型。

创建 i2 步骤

仓库代理站点是 i2 TradeMatrix BPI 连接器运行的工作站、安装 i2 TradeMatrix BPI 引擎的位置以及编目 DDM（仓库目标）的位置。

创建 i2 步骤

在导入 i2 源的元数据后，数据仓库中心为 i2 装入步骤创建并填充过程文件夹。打开“过程模型”窗口时，步骤已添加到画布中，可用作运行 i2 装入步骤的模板。

限制：用于更新“BPI 商务模型”事实表和 OLAP 立方的 i2 步骤只能引用与单个事实表和 OLAP 立方相关的部署的“BPI 商务模型”。i2 步骤不支持已部署到多个事实表和 OLAP 立方的“商务模型”。

要定义步骤的特性，右键单击画布上的“i2 步骤”，然后单击**特性**或双击这些“i2 步骤”之一。

在装入 i2 维和事实表的那个步骤的“特性”窗口中，必须指定以下参数的值：

维名 维的名称。

维类型 维的类型。可以是“一致”或“不一致”。

成员文件名

“BPI 导入”目录中包含对维成员的更改的文件名称。

层次结构文件名

“BPI 导入”目录中包含对维的层次结构更改的文件名称。

别名文件名

“BPI 导入”目录中包含对维的别名更改的文件名称。

属性文件名

“BPI 导入”目录中包含对维属性的更改的文件名称。

商务模型

与步骤相关的商务模型的名称。

身份类型

导入成员数据的方式：

0 按代码

1 按唯一名

2 按系列

层次结构格式

层次结构的成员之间的关系：

0 父 / 子关系

1 平面层次结构关系

2 雪花关系

字符串的引号

输入文件中的数据是否有引号:

真 字符串有引号。

假 字符串没有引号。

固定长度

输入文件中的数据有固定长度:

真 数据有固定长度。

假 数据没有固定长度。

定界符 输入文件中的数据是如何分隔的:

0 数据由逗号 (,) 分隔。

1 数据由管道符分隔。

2 数据由制表符分隔。

在装入 i2 OLAP 立方的那个步骤的“特性”窗口中，必须指定以下参数的值:

商务模型

与步骤相关的商务模型的名称。

定义文件名

“BPI 导入”目录中的某个文件的名称，该文件描述包含在数据文件中的数据
的维和量度。

数据文件名

“BPI 导入”目录中包含对事实表数据的更改的文件名称。

身份类型

导入成员数据的方式:

0 按代码

1 按唯一名

2 按系列

数据文件格式

输入文件的格式:

0 数据用于单列事实表。

1 数据用于多列事实表。

字符串的引号

输入文件中的数据是否有引号:

True 字符串有引号。

False 字符串没有引号。

创建 i2 步骤

固定长度

输入文件中的数据有固定长度:

True 数据有固定长度。

False 数据没有固定长度。

定界符 输入文件中的数据是如何分隔的:

0 数据由逗号 (,) 分隔。

1 数据由管道符分隔。

2 数据由制表符分隔。

重要事项: 如果不想对特定参数指定值, 必须对它指定值 NULL。不能对它保留空白。

i2 TradeMatrix BPI 连接器不会创建用作 i2 步骤的源数据的、以逗号分隔的变量 (CSV) 文件。CSV 文件在可以运行 i2 步骤之前, 必须存在于仓库代理站点的“BPI 导入”目录中。

声明

IBM 可能未在所有国家或地区中提供本文档中讨论的产品、服务或功能部件。关于您所在区域目前可用的产品及服务的信息，请向当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并不明示或默示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以用来代替 IBM 产品、程序或服务。然而，评估和验证任何非 IBM 产品、程序或服务均由用户自行负责。

IBM 可能已经申请或正在申请与本文档有关的各项专利权。提供本文档并不表示允许您使用这些专利。 您可以用书面方式将许可证查询寄往：

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

有关双字节（DBCS）信息的许可证查询，请与您的国家或地区的“IBM 知识产权部”联系，或用书面方式将查询寄往：

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

以下段落不适用于联合王国或该条款与当地法律不一致的任何国家或地区：国际商业机器公司以“仅此状态”的基础提供此出版物，不附有任何形式的（无论是明示的还是默示的）保证，包括（但不限于）不侵犯、适销性或适用于某特定用途的默示保证或条件。一些国家或地区允许否认某些事务中的明示或默示保证，因此，此声明可能不适用于您。

本资料可能会包含技术错误或印刷错误。此处的信息会定期得到更改；这些更改将编入本出版物的新版本中。IBM 可能随时对此出版物中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

在此信息中对非 IBM Web 站点的任何引用仅是出于方便起见，不以任何方式提供对这些 Web 站点的保证。这些 Web 站点中的资料不是此 IBM 产品资料的一部分，使用这些 Web 站点时风险自负。

IBM 对于您以任何方式提供的任何信息，有权利以任何它认为适当的方式使用或分发，而不必对您负任何责任。

为了以下目的：（i）允许在独立创建的程序和其他程序（包括本程序）之间进行信息交换（ii）允许对已经交换的信息进行相互使用，而希望获取本程序有关信息的被许可方请与以下地址联系：

IBM Canada Limited
Office of the Lab Director
8200 Warden Avenue
Markham, Ontario
L6G 1C7
CANADA

只要遵守适当的条款和条件，包括某些情形下的一定数量的付款，都可获取这方面的信息。

本资料中描述的许可程序和它可用的全部许可证材料均由 IBM 根据“IBM 客户协议”、“IBM 国际程序许可证协议”或任何与客户之间的等效协议中的条款提供。

此处包含的所有性能数据都是在受控环境中确定的。因此，在其他操作环境中获得的结果可能与之相差很大。某些测量可能是在开发级的系统上进行的，不能保证这些测量方法在通用系统上同样可用。此外，某些测量方法可能是通过外推法归纳来估计的。实际结果可能会有所不同。此文档的用户应针对他们的特定环境验证数据是否适用。

关于非 IBM 产品的信息是从那些产品的供应商、他们发布的声明或其他公用来源获得的。IBM 未测试那些产品，不能确认与非 IBM 产品相关的性能、兼容性或任何其他声明的准确性。如有关于非 IBM 产品的功能的问题，应向那些产品的供应商提出。

所有关于 IBM 未来方向或意向的声明都可能随时更改或撤销，而不作任何通知，并且仅代表发展目标。

本资料中可能包含用于日常业务运作的的数据或报表的示例。为了尽可能完整地说明问题，这些示例可能包含个人、公司、品牌和产品的名称。所有这些名称都是虚构的，如与实际商业企业所使用的名称和地址相似，纯属巧合。

版权许可证：

本资料可能包含源语言的样本应用程序，它们举例说明各种操作平台上的编程技术。为了开发、使用、市场营销或分发符合编写这些样本程序所针对的操作系统的的应用程序编程接口的应用程序，您可以以任何形式复制、修改和分发这些样本

程序，而不必向 IBM 付款。尚未在所有条件下彻底测试这些示例。因此，IBM 不能保证或默示这些程序的可靠性、适用性或功能。

这些样本程序或任何派生产品的每个副本或任何部分都必须包括如下版权声明：

©（您的公司名）（年份）。本代码的某些部分是从“IBM 公司样本程序”派生的。

© Copyright IBM Corp. _输入年份_。 All rights reserved.

商标

以下术语是国际商业机器公司在美国和 / 或其他国家或地区的商标。

ACF/VTAM	IBM
AISPO	IMS
AIX	IMS/ESA
AIX/6000	LAN Distance
AIXwindows	MVS
AnyNet	MVS/ESA
APPN	MVS/XA
AS/400	Net.Data
BookManager	OS/2 OS/390
CICS	OS/400
C Set++	PowerPC
C/370	QBIC
DATABASE 2	QMF
DataHub	RACF
DataJoiner	RISC System/6000
DataPropagator	RS/6000
DataRefresher	S/370
DB2	SP
DB2 Connect	SQL/DS
DB2 Extender	SQL/400
DB2 OLAP Server	System/370
DB2 Universal Database	System/390
Distributed Relational	SystemView
Database Architecture	VisualAge
DRDA	VM/ESA
eNetwork	VSE/ESA
Extended Services	VTAM
FFST	WebExplorer
First Failure Support Technology	WIN-OS/2

下列各项是其他公司的商标或注册商标:

Microsoft、Windows 和 Windows NT 是 Microsoft Corporation 的商标或注册商标。

Java 和所有基于 Java 的商标和徽标以及 Solaris 是 Sun Microsystems, Inc. 在美国和 / 或其他国家或地区的商标。

Tivoli 和 NetView 是 Tivoli Systems Inc. 在美国和 / 或其他国家或地区的商标。

UNIX 是经 X/Open Company Limited 唯一许可的在美国和 / 或其他国家或地区的注册商标。

其他公司、产品或服务名称, 可能是其他公司的商标或服务标记。

索引

[A]

安全性 19
安装先决条件产品 372

[B]

备份“数据仓库中心” 295
笔记本
 程序
 参数页 286
 代理站点页 285
编写
 用户定义的程序 289
变换器步骤 4
变换器, 设置 123
变量 395
 环境 395
标记语言文件 296
标记语言文件、DWC、ICM、ERwin
 文件和 ER1 文件 248
标准配置文件 375
表空间名 328
步骤
 测试 142, 143
 从“数据仓库中心”外部启动
 150
 打印 122
 代码转换 195
 调度 146
 反转数据 212
 方式
 测试 142
 定义 142
 开发 142
 分级表 186
 更改聚集 183
 汇总 200
 基本聚集 181
 将源列映射到目标列 139
 开发 142

步骤 (续)
 清理数据 202
 生成键表 207
 生成周期表 209
 时间点 181
 使用用户定义程序 287
 使用 FTP 复制文件 173
 使用 ODBC 将数据导出到文件中
 160
 提交 OS/390 JCL 作业流 176
 提升
 测试方式 143
 生产方式 150
 添加 122
 外部填充 149
 旋转数据 214
 用户复制 181
 用于过滤源数据 199
 与源和目标链接 136
 运行 142
 运行 FTP 命令文件 175
 状态
 和用户定义程序反馈 292
 子类型 124
 “滑动平均值”变换器 231
 “回归”变换器 234
 “计算统计信息”变换器 221
 “计算小计”变换器 223
 “相关”变换器 228
 “DB2 通用数据库”导出 159
 “DB2 通用数据库”装入 161
 “OLAP 服务器: 使用计算规则的
 Calc” (ESSCALC2) 仓库程序
 259
 “OLAP 服务器: 使用装入规则从
 文件中装入数据” (ESSDATA2)
 仓库程序 261
 “X 平方”变换器 226
 ANOVA 变换器 219
 DB2 AS/400 版数据插入装入
 162

步骤 (续)
 DB2 AS/400 版数据置换装入
 166
 DB2 OS/390 版装入 170
 i2 122
 OLAP 服务器: 不使用装入规则从
 文件中装入数据 (ESSDATA4)
 仓库程序 263
 OLAP 服务器: 从文件更新轮廓
 (ESSOTL1) 264
 OLAP 服务器: 从 SQL 表更新轮
 廓 (ESSOTL2) 266
 OLAP 服务器: 缺省 calc
 (ESSCALC1) 仓库程序 258
 OLAP 服务器: 使用装入规则将数
 据从 SQL 表中装入
 (ESSDATA3) 仓库程序 262
 OLAP 服务器: 自由文本数据装入
 仓库程序 260
 SAP 122
 SQL
 选择和插入数据 155
 Web 流量 122
步骤, 创建
 i2 400

[C]

参数替换 286
仓库 1
 程序
 定义 119
 用于数据净化 193
 设置 7
 设置访问权 101
 受支持的数据库 111
 选择 101
 映射到源数据 139
 “DB2 公共服务器”产品 101
 DB2 AS/400 版
 特权 102

- 仓库 (续)
 - DB2 AS/400 版 (续)
 - DB2 Connect 网关站点 103
 - DB2 EEE
 - 特权 106
 - DB2 OS/390 版
 - 特权 105
 - 仓库变换器
 - 反转数据 212
 - 清理数据 202
 - 生成键表 207
 - 生成周期表 209
 - 旋转数据 214
 - 仓库步骤 2
 - 变换器 4
 - 程序 4
 - 用户定义的程序 4
 - SQL 3
 - 仓库程序
 - 参数 286
 - 使用 FTP 复制文件 173
 - 使用 ODBC 将数据导出到文件中 160
 - 提交 OS/390 JCL 作业流 176
 - 运行 FTP 命令文件 175
 - 在步骤中使用 287
 - “DB2 通用数据库”导出 159
 - “DB2 通用数据库”装入 161
 - DB2 AS/400 版数据插入装入 162
 - DB2 AS/400 版数据置换装入 166
 - DB2 OS/390 版装入 170
 - OLAP 服务器 257
 - OLAP 服务器: 不使用装入规则将数据从文件中装入 (ESSDATA2) 263
 - OLAP 服务器: 从文件更新轮廓 (ESSOTL1) 264
 - OLAP 服务器: 从 SQL 表更新轮廓 (ESSOTL2) 266
 - OLAP 服务器: 缺省 calc (ESSCALC1) 258
 - OLAP 服务器: 使用装入规则将数据从文件中装入 (ESSDATA2) 261
 - 仓库程序 (续)
 - OLAP 服务器: 用 calc 规则进行计算 (ESSCALC2) 259
 - OLAP 服务器: 自由文本数据装入 260
 - OLAP Server: 使用装入规则将数据从 SQL 表中装入 (ESSDATA3) 262
 - 仓库代理 2
 - 仓库代理守护进程
 - 停止 11
 - AS/400
 - 启动 8
 - 验证活动 9
 - OS/2
 - 启动 8
 - OS/390
 - 启动 9
 - Windows NT
 - 启动 8
 - 仓库代理, 本地 2
 - 仓库代理, 远程 2
 - 仓库管理器
 - i2 TradeMatrix BPI 连接器
 - 创建 i2 步骤 400
 - 导入 i2 源的元数据 399
 - 仓库过程 2
 - 仓库模式
 - 导出到 DB2 OLAP Integration Server 315
 - 定义 312
 - 发布元数据关于 323
 - 连接表 314
 - 添加表和视图 313
 - 在 DB2 OLAP Integration Server 中使用 317
 - 自动连接表 314
 - 仓库目标 2
 - 定义 111
 - 仓库源 2
 - 定义 51
 - 受支持的版本和发行版级别 25
 - 操作数据 1
 - 测试
 - 步骤 142
 - 用户定义的程序 288
 - 测试步骤 143
 - 测试方式 142
 - 插入数据 155
 - 查询处理器 367
 - 程序笔记本
 - 参数页 286
 - 代理站点页 285
 - 程序步骤 4
 - 程序组
 - 定义 283
 - 初始化
 - 参数 306
 - 定义 306
 - 类型 306
 - 初始化服务 367
 - 初始化辅助控制数据库 306
 - 窗口
 - 登录 11
 - 从“Visual Warehouse 主机适配器”
 - 迁移 393
 - 存储过程 289
 - 错误处理 144
 - 错误消息 144
 - 错误, 记录 327
- [D]
- 打开过程 121
 - 打印
 - 步骤 122
 - 代理 365
 - 代理站点 2
 - 定义 13
 - 配置 13
 - 代理站点, 缺省 2
 - 代理, 仓库 2
 - 代码转换 195
 - 导入
 - 初始化控制数据库 298
 - 从一个“数据仓库中心”系统到另一个系统 298
 - 对象定义 298
 - 导入, 标记语言文件, ICM, 信息目录管理器 251
 - 导入, 导入, 标记语言文件, DWC 250

- 登录窗口 11
- 登录到“数据仓库中心”桌面 11
- 调度步骤 146
- 定义
 - 安全性 19
 - 步骤 122
 - 反转数据 212
 - 分级表 186
 - 更改聚集 183
 - 基本聚集 181
 - 清理数据 202
 - 生成键表 207
 - 生成周期表 209
 - 时间点 181
 - 使用 FTP 复制文件 173
 - 使用 ODBC 将数据导出到文件中 160
 - 提交 OS/390 JCL 作业流 176
 - 旋转数据 214
 - 用户复制 181
 - 运行 FTP 命令文件 175
 - “DB2 通用数据库”导出 159
 - “DB2 通用数据库”装入 161
 - DB2 AS/400 版数据插入装入 162
 - DB2 AS/400 版数据置换装入 166
 - DB2 OS/390 版装入 170
 - 仓库
 - 供 DataJoiner 使用 107
 - 仓库目标
 - 在“数据仓库中心”中 111
 - DB2 EEE 版 107
 - DB2 OS/390 版 106
 - 程序组 283
 - 代理站点 13
 - 过程 120
 - 特权
 - DB2 公共服务器 101
 - DB2 通用数据库 AS/400 版 47
 - DB2 通用数据库 OS/390 版 47
 - DB2 系列 46
- 定义 (续)
- 特权 (续)
 - DB2 AS/400 版 102
 - DB2 EEE 版 106
 - DB2 OS/390 版 105
 - DB2 VM 版 47
 - DB2 VSE 版 47
- 信息资源 51
- 用户定义的程序 284
- 源
 - 供 DataJoiner 使用 98, 107
- SQL
 - SQL 155
- 定义步骤
 - 用户复制 181
- 定义源
 - 复制 180
- 端口号 376
- 多维立方
 - 用数据装入 318
- [F]**
 - 反馈文件 291
 - 反向回归 235
 - 反转数据 变换器 212
 - 方差分析 219
 - 非参数测试 226
 - 非仓库源
 - 受支持的版本和发行版级别 25
 - 非关系数据映射器 370
 - 非关系数据源 363
 - 服务器调度方式 307
 - 覆盖现存日志 393
 - 负载平衡 370
 - 复制表 179
 - 复制数据
 - 在 z/OS 和 OS/390 表之间 172
- [G]**
 - 格式化日期 变换器 217
 - 跟踪
 - 级别 307, 392
 - 目录 307
 - 文件名 393
- 更改配置 295
- 更新环境变量 395
- 更新远程数据库中的现存表 110
- 公司元数据 4
- 关系查询 363
- 管理界面, 启动 11
- 滚动和 231
- 过程
 - 打开 121
 - 定义 120
 - 添加
 - 目标 121
 - 源 121
- 过滤数据 199
- [H]**
 - 环境变量 395
 - 更新 395
 - 恢复
 - 备份“数据仓库中心” 295
 - 使用日志文件来恢复 325
 - 回归, 反向 235
 - 回归, 全模型 235
 - 汇总 200
- [J]**
 - 基本统计信息 221
 - 记号 286
 - 记录错误 327
 - 记录级别 328
 - 计算的列
 - 定义 200
 - 计算脚本 257
 - 简单滑动平均值 231
 - 键
 - 外部 199
 - 主 199
 - 建立联网 49
 - 键列 207
 - 键值 207
 - 将步骤映射至数据源 139
 - 将目标表从 DataJoiner 移动至远程数据库 110
 - 警告消息 144

聚集列 214

[K]

开发步骤 142

开发方式 142

客户机连接请求 369

控制配置参数 367

控制数据库

安装新的控制数据库 306

初始化 306

导出考虑事项 296

控制数据库的密码 307

跨内存 374

[L]

类属 ODBC 驱动程序 25

连接

全外连接 196

外连接 196

星型连接 198

右外连接 196

与仓库

DB2 公共服务器 101

DB2 AS/400 版 103

DB2 EEE 107

DB2 OS/390 版 105

与源

DB2 通用数据库 AS/400 版
48

DB2 通用数据库 OS/390 版
48

DB2 系列 46

DB2 VM 版 48

DB2 VSE 版 48

左外连接 196

连接处理程序 367

连接器

i2 399

连接源列 193

联网

建立

AS/400 代理 49

OS/390 代理 50

联网需求

仓库服务器和仓库代理之间的 16

在远程数据库之间 49

轮廓 257

逻辑表 364

逻辑数据库 364

[M]

名称和地址清理, Trillium 237

目标 13

受支持的版本和发行版 111

添加 121

与步骤链接 136

目标表

更新远程数据库中的现存表 110

移动 110

用 DataJoiner 创建 109

目标列

映射到源列 139

目标数据库, 导出考虑事项 296

目录名 391

[N]

内核主机名 307

拟合优度测试 226

[P]

配置 13

本地 OS/390 客户机 374

数据源 382

先决条件产品 372

IMS 85

Informix

OS/2 91

UNIX 85

Windows NT 60

LU 6.2 通信 389

Microsoft Access 78

Microsoft SQL Server

UNIX 90

Windows NT 74

MicrosoftExcel 82

ODBC 驱动程序 391

配置 (续)

Oracle

OS/2 94

UNIX 88

Windows NT 67

OS/390 上的 LU 6.2 378

OS/390 上的 TCP/IP 375

Sybase

OS/2 92

UNIX 87

Windows NT 64

TCP/IP 通信 386

VSAM 85

Windows NT 上的 LU 6.2 379

Windows NT 上的 TCP/IP 377

配置文件 395

配置先决条件 383

配置“数据仓库中心”

安装 307

更改 295

以使用“OLAP 服务器” 257

[Q]

启动

仓库代理守护进程

AS/400 8

OS/2 8

OS/390 9

Windows NT 8

启动“数据仓库中心”

服务器 7

管理界面 11

日志记录器 7

企业服务器

定义 368

实现 369

清理类型 203

清理数据 变换器 202

趋势预测 231

区域控制器 367

取装缓冲区大小 391

全模型回归 235

全外连接 196

缺省代理站点 2

[R]

- 任务
 - 入库 4
- 日志 325
- 日志表 144, 328
- 日志记录器服务 368
- 日志文件
 - MQSeries 281
- 入库
 - 对象 1
 - 概述 1
 - 任务 4
- 入库的概述 1
 - (入库中的) 步骤 2
 - (入库中的) 过程 2

[S]

- 设置
 - 日志目录 328
- 设置仓库 7
- 设置数据库目录选项 388, 390
- 生成键表 变换器 207
- 生成周期表 变换器 209
- 时间序列分析 231
- 事件查看器 326
- 示例
 - 将源数据映射到目标表 139
 - 使用仓库程序用于数据净化 193
 - 外部触发程序 152
 - 星型连接 199
 - 星型模式 198
 - 摘要 200
 - 左外连接 196
 - ANOVA 变换器 219
 - DataJoiner 服务器映射 99
 - SQL DATE 函数 200
 - WHERE 子句 199
- 事实表 198
- 输出日志表 144
- 数据
 - 操作 1
 - 插入 155
 - 查看
 - 目标表 117

数据 (续)

- 查看 (续)
 - 源表 52
- 过滤 199
- 信息 1
- 选择 155
- 转换 193
- 数据仓库中心 1
 - 安全性, 定义 19
 - 备份 295
 - 从外部启动步骤 150
 - 登录 11
 - 服务器, 启动 7
 - 环境变量 395
 - 记号 286
 - 配置文件 395
 - 日志记录器, 启动 7
 - 特性, 更改 308
 - 映射到“信息目录管理器”对象类型 302
- 状态表
 - 创建 117
 - 和 DataJoiner 108
- OLE DB 支持 397
- 数据服务器 366
- 数据清理, 名称和地址清理 237
- 数据映射器
 - 工作流程 371
 - 描述 370
- 数据源处理程序 369

[T]

- 特权
 - DB2 仓库的 101
 - DB2 公共服务器
 - 仓库 101
 - DB2 通用数据库 AS/400 版
 - 源 47
 - DB2 通用数据库 OS/390 版 47
 - DB2 系列
 - 源 46
 - DB2 AS/400 版
 - 仓库 102
 - DB2 EEE 版
 - 仓库 106

特权 (续)

- DB2 OS/390 版
 - 仓库 105
- DB2 VM 版 47
- DB2 VSE 版 47
- 提升步骤
 - 测试方式 143
 - 生产方式 150
- 添加数据源 384
- 通信复合地址字段 374
- 通信选项 374
- 统计变换器
 - 计算统计信息 变换器 221
 - “滑动平均值”变换器 231
 - “回归”变换器 234
 - “计算小计”变换器 223
 - “相关”变换器 228
 - “X 平方”变换器 226
 - ANOVA 变换器 219
- 统计信息, 计算 221

[W]

- 外部触发程序 150
 - 返回码 153
 - 示例 152
 - 状态 153
- 外部填充的步骤 149
- 外键 199
- 外连接 196
- 维表 198
- 伪关系数据 364
- 文件, 配置 395

[X]

- 系统过程 145
- 相关系数 228
- 响应时间超时 392
- 项目计划
 - 选择仓库 101
 - 选择源 25
- 消息 144
- 消息池大小 392
- 消息传递
 - MQSeries 278

- 消息类型 144
- 小计, 计算 223
- 协方差 228
- 写入时关闭跟踪 393
- 信息目录 4
- 信息目录管理器
 - 发布元数据 301
- 信息数据 1
- 星型连接 198
- 星型模式 4, 198
 - 创建 311
 - 描述 311
- 旋转数据 变换器 214
- 选择仓库 101
- 选择数据 155
- 选择源 25

[Y]

- 压缩 392
- 验证通信 10
- 样本数据
 - 查看
 - 目标表 117
 - 源表 52
- 移动数据
 - 复制 179
- 映射非关系数据 364, 370
- 用户定义程序步骤 4
- 用户定义的程序
 - 编写 289
 - 参数 289
 - 测试 288
 - 定义 283, 284
 - 返回码 290
 - 反馈 290
 - 和步骤状态 292
 - 和 SELECT 语句 286
 - 将代理更改为用户过程 145
 - MQSeries 281
 - Object REXX Windows 版 289
- 用户定义函数 217
- 用户过程 145
- 用于数据净化的公式和表达式 193
- 用 DataJoiner 创建目标表 109
- 右外连接 196

- 源 13
 - 非 DB2
 - 定义 95
 - 非 DB2, 联网 60
 - 复制 179, 180
 - 类型 25
 - 设置
 - 访问 25
 - DB2 Connect 网关站点 48
 - 添加 121
 - 与步骤链接 136
 - DB2
 - 定义 51
 - DB2 通用数据库 AS/400 版 47
 - DB2 通用数据库 OS/390 版 47
 - DB2 系列产品 45
 - DB2 VM 版 47
 - DB2 VSE 版 47
 - EOF
 - 定义 57
 - IMS 28, 35, 39, 43, 85
 - Informix
 - OS/2 91
 - UNIX 85
 - Windows NT 60
 - Informix 7.2 32, 37, 41, 45
 - LAN 文件 33, 37, 42, 45
 - Microsoft Access 78
 - 联网需求 33
 - Microsoft Excel 82
 - 联网需求 33
 - Microsoft SQL Server
 - UNIX 90
 - Windows NT 74
 - Microsoft SQL server 32, 37, 41
 - Oracle
 - OS/2 94
 - UNIX 88
 - Windows NT 67
 - Oracle 7 31, 36, 41, 44
 - OS/390 29, 35, 40, 44
 - Sybase 30, 36, 40, 44
 - OS/2 92
 - UNIX 87
 - Windows NT 64
 - VM 文件 29, 35, 40, 44

- 源 (续)
 - VSAM 29, 35, 39, 43, 85
- 源列
 - 映射到目标列 139
- 元数据 4
 - 导出到标记语言文件 296
 - 导出和导入, 概述 296
 - 发布到信息目录 301
 - 将删除的对象传播到信息目录 306
 - 设置发布到信息目录的环境 301
 - 信息目录管理器和 OLAP 服务器 的映射 338
 - 用信息目录管理器映射“数据仓库中心” 329
 - 用 DB2 OLAP Integration Server 映射“数据仓库中心” 339
 - 与“信息目录管理器”同步 304
 - 在信息目录中更新 304
 - 在信息目录中进行更新 创建的日志文件 305
- 元数据语法 370
- 元数据, 导入 i2 源 399
- 允许跟踪 393
- 运行小计 223

[Z]

- 增量落实 157
- 诊断 325
- 指数平滑滑动平均值 231
- 周期表 209
- 主机名 376
- 主键 199
- 主题区 2
- 转换
 - 代码 195
 - 数据 193
- 装入规则 257
- 子系统接口 368
- 左外连接 196

[特别字符]

- “分级表”步骤 186
- “更改聚集”步骤 183

- “滑动平均值”变换器 231
- “回归”变换器 234
- “基本聚集”步骤 181
- “计算统计信息”变换器 221
- “计算小计”变换器 223
- “时间点”步骤 181
- “使用 FTP 复制文件”仓库程序 173
- “使用 ODBC 将数据导出到文件中”仓库程序 160
- “数据仓库中心”程序位置 289
- “提交 OS/390 JCL 作业流”仓库程序 176
- “相关”变换器 228
- “信息目录管理器”对象类型映射到“数据仓库中心”对象类型 302
- “用户复制”步骤 181
- “运行 FTP 命令文件”仓库程序 175
- “DB2 通用数据库”导出仓库程序 159
- “DB2 通用数据库”装入仓库程序 161
- “DB2 AS/400 版数据插入装入”仓库程序 162
- “DB2 AS/400 版数据置换装入”仓库程序 166
- “DB2 OS/390 版装入”仓库程序 170
- “OLAP 服务器”仓库程序 257
- “OLAP 服务器: 使用计算规则的 Calc” (ESSCALC2) 仓库程序 259
- “OLAP 服务器: 使用装入规则从文件中装入数据” (ESSDATA2) 仓库程序 261

A

ANOVA 变换器 219

B

Berkeley 套接字 375

BVBESTATUS 表
 创建 117
 和 DataJoiner 108
 OS/390 版数据库 106

C

CASE 语句 196
 Classic Connect
 仓库代理 365
 非关系数据映射器 370
 数据服务器 366
 “数据仓库中心”步骤 363
 Client Access/400 103
 CROSS ACCESS ODBC 驱动程序 366

D

DataJoiner
 创建目标表 109
 定义
 仓库 107
 源 98, 107
 访问源和目标数据 98, 107
 服务器映射示例 99
 更新远程数据库中的现存表 110
 移动目标表 110
 DB2
 仓库 101
 在主机系统上, 作为源访问 46
 DB2 公共服务器
 特权 101
 与仓库连接 101
 作为目标访问 101
 DB2 通用数据库 AS/400 版
 设置 DB2 Connect 网关站点 48
 特权 47
 与源连接 48
 作为源访问 46
 DB2 通用数据库 OS/390 版
 设置 DB2 Connect 网关站点 48
 特权 47
 与源连接 48
 作为源访问 46

DB2 系列
 特权 46
 与源连接 46
 作为源访问 45
 DB2 AS/400 版
 特权 102
 与仓库连接 103
 作为目标访问 102
 DB2 Connect 102
 DB2 Connect 网关站点 48
 DB2 EEE
 仓库目标 107
 DB2 EEE 版
 特权 106
 与仓库连接 107
 DB2 OLAP Integration Server
 多维立方
 根据调度表从“数据仓库中心”中装入 321
 多维立方, 用数据装入 318
 将元数据导出到“数据仓库中心” 320
 “数据仓库中心”的元数据映射 339
 DB2 OS/390 版
 仓库目标 106
 授权到公共特性 106
 特权 105
 与仓库连接 105
 DB2 VM 版
 设置 DB2 Connect 网关站点 48
 特权 47
 与源连接 48
 作为源访问 46
 DB2 VSE 版
 设置 DB2 Connect 网关站点 48
 授权到公共特性 106
 特权 47
 与源连接 48
 作为源访问 46
 DJXSAMP 367
 DRDA 46
 DSNUTILS 122
 DWC
 OLE DB 支持 397

F

Fisher-F 分布 219

H

Hyperion Essbase 仓库程序 257

I

i2

步骤 122

IMS 联网需求 28, 35, 39, 43

IMS 逻辑表 363

IMS BMP/DBB

初始化服务 367

接口 368

IMS DRA

初始化服务 367

接口 368

Informix 7.2 联网需求 32, 37, 41, 45

IP 地址 376

IWH2LOGC.LOG 328

IWH2LOG.LOG 328

IWH2SERV.LOG 328

L

LAN 联网需求 33, 37, 42, 45

LU 6.2 配置

工作表 381

示例 380

M

Microsoft Access

联网需求 33

Microsoft Excel

联网需求 33

Microsoft SQL Server 联网需求 32, 37, 41

MQSeries

安装需求 279

创建视图 278

错误日志文件 281

MQSeries (续)

导入 279

用户定义的程序 281

与“数据仓库中心”一起使用
278

XML 元数据 279

MTO 界面 367

O

Object REXX Windows 版 289

ODBC

类属驱动程序 25

ODBC 驱动程序 366

OLAP 服务器

信息目录管理器的元数据映射
338

OLAP 服务器: 不使用装入规则从文
件中装入数据 (ESSDATA4) 仓库程
序 263

OLAP 服务器: 从文件更新轮廓
(ESSOTL1) 264

OLAP 服务器: 从 SQL 表更新轮廓
(ESSOTL2) 266

OLAP 服务器: 缺省 calc
(ESSCALC1) 仓库程序 258

OLAP 服务器: 使用装入规则将数据
从 SQL 表中装入 (ESSDATA3) 仓
库程序 262

OLAP 服务器: 自由文本数据装入仓
库程序 260

OLE DB 支持 397

Oracle 7 联网需求 31, 36, 41, 44
OS/390

复制表数据 172

OS/390 客户机应用程序 367

OS/390 文件联网需求 29, 35, 40, 44

P

P 值 229

p 值 219

Pearson 积动差相关系数 228

S

SAP

步骤 122

SELECT 语句 286

SNA 协议 374

SQL

聚集函数 200

CASE 语句 196

GROUP BY 子句 200

JOIN 子句 195

OUTER JOIN 关键字 196

SELECT 语句 286

SUM 函数 200

WHERE 子句 199

WHERE 子句和数据净化 193

SQL 步骤 3

SQL 代码 144

Sybase 联网需求 30, 36, 40, 44

T

T 值 229

TCP/IP

配置工作表 378

配置示例 378

配置信息 395

协议 375

U

UDP

MQSeries 281

V

VM 文件联网需求 29, 35, 40, 44

VSAM

接口 368

逻辑表 363

VSAM 联网需求 29, 35, 39, 43

W

Web 流量

步骤 122

WHERE 子句 199
WHERE 子句, 在数据净化中使用
193
Windows NT
事件查看器 326
WLM 初始化服务 367

X

X 平方 变换器 226
XTClient
语法 151

Z

z/OS
复制表数据 172

与 IBM 联系

在联系 DB2 客户支持之前，请先查阅 *Troubleshooting Guide* 以获取关于您的特定技术问题的帮助。如果仍然需要帮助，请在联系 DB2 客户支持前完成收集所需信息的步骤。

要获取信息或订购任何“DB2 通用数据库”产品，与当地分支机构的 IBM 代表联系或与任何授权的 IBM 软件经销商联系。

您如果住在美国，请致电下列其中一个号码：

- 1-800-237-5511，可获得客户支持
- 1-888-426-4343，可了解所提供的服务项目

产品信息

可通过电话或万维网网址：

<http://www.ibm.com/software/data/db2/udb/winos2unix/support> 获取 DB2 通用数据库的产品信息。

此站点包含了关于技术资料库、订购书籍、客户机下载、新闻组、修订包、新闻以及与 web 资源的链接的最新信息。

您如果住在美国，请致电下列其中一个号码：

- 1-800-IBM-CALL (1-800-426-2255) 以订购产品或获取一般信息。
- 1-800-879-2755，可订购出版物。

有关如何联系美国以外的 IBM 的信息，请访问 IBM 全球网页：
<http://www.ibm.com/planetwide>

在某些国家，IBM 授权的代理商应联系他们的代理商支持机构以获取信息。



SB84-0226-01



Spine information:



IBM® DB2® 通用数据库

数据仓库中心系统管理员指南

版本 7