

DB2 Platform: UNIX, Windows

DataLinks: Manage Enterprise Content

*Kiran Mehta and Joshua Hui
IBM Almaden Research Center*

Wednesday, May 17, 2000

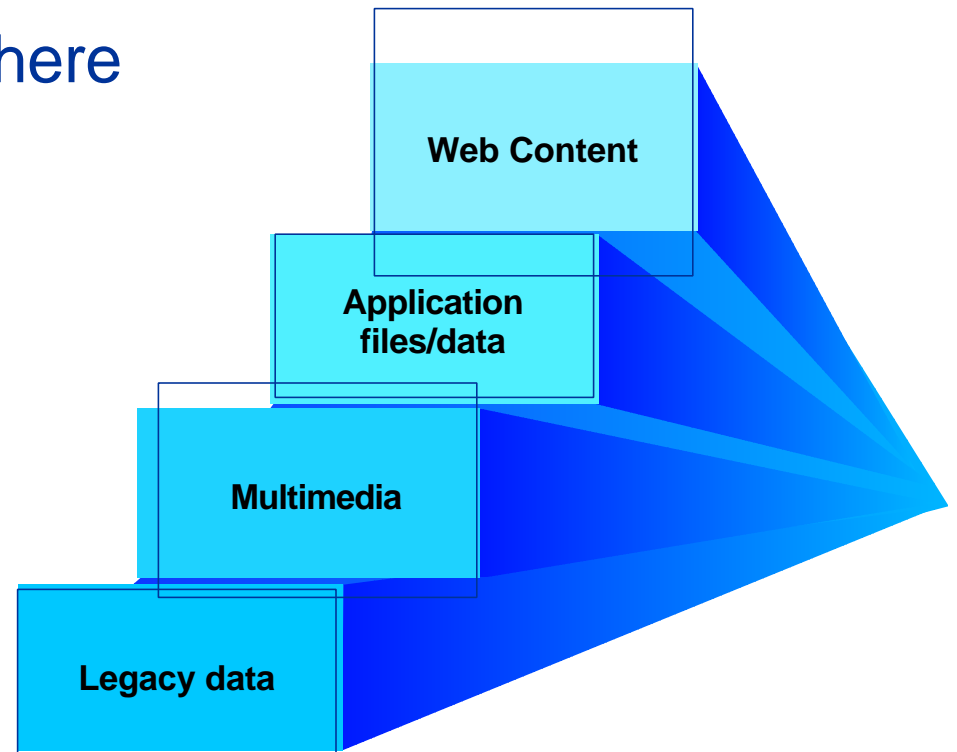
**INTERNATIONAL
DB2 USERS GROUP**



Independent • Not-for-Profit • User Run

Enterprise Data

- Most of Enterprise Data lives in Files
- It will continue to live and grow there
- It is typically unmanaged



Managed Enterprise Content?

- Content = (Metadata, Data)
- Enterprise Content = Lots of (Content)
- Managed (Enterprise Content) = DataLinked (Metadata, Data)

Current DBMS Limitations

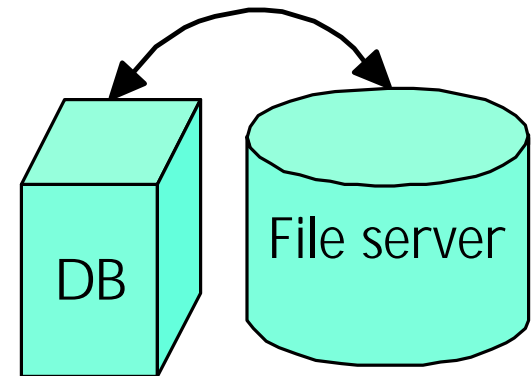
- DBMS provides referential integrity, recovery and access control, but:
 - owns only a fraction of world's data
- Access methods are specific to the DBMS (SQL, ODBC), but:
 - most applications store and reference data in files.

Current DBMS Limitations (*continued*)

- LOBs provide support for large objects in a database, but:
 - cannot deliver data isochronously
 - data has to be copied back and forth from DBMS files
 - It is best to let application use file system primitives
- Oracle iFS provides a file system feel and look to DBMS objects, but:
 - implemented on top of BLOBs
 - performance degradation

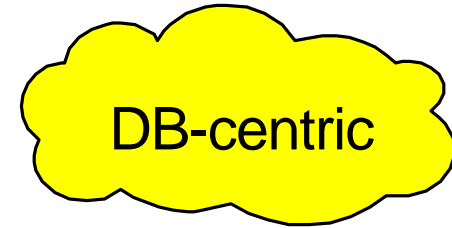
DataLinks: DBMS and File Systems

- Extends DB's management to include file data in file systems with respect to
 - Referential Integrity
 - File Data Integrity
 - Coordinated Backup and Recovery
 - Access Control
- DataLinks allows management of files **as if** they are stored in the database
- Retain file system performance, APIs and look and feel



Fusion of database and file system technologies

DataLinks Applications



Example - Insurance Application

Policy #	Customer Name	Insurance Type	Date of Accident

Schema (now)

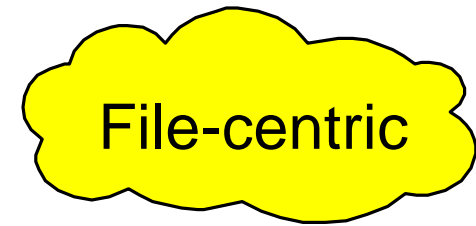
Extend DB schema to incorporate new data types, e.g., text, image, video

Policy #	Customer Name	Insurance Type	Date of Accident	Video Clip of Accident
				pointer to file 1 →
				Null
				pointer to file 3 →
				Null

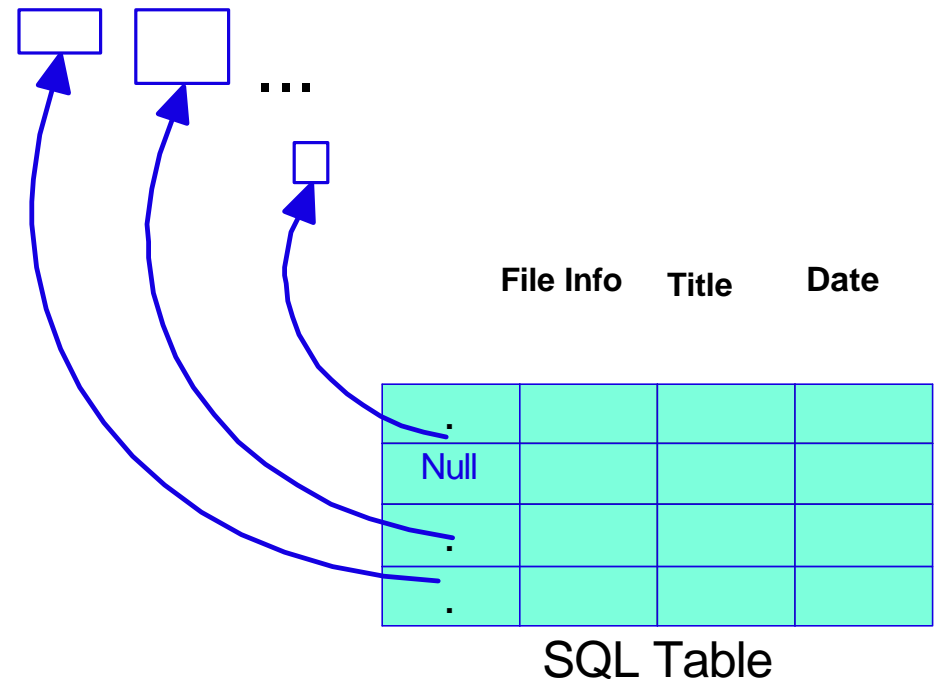
New Schema

Can extend query to include search on video clip

DataLinks Applications

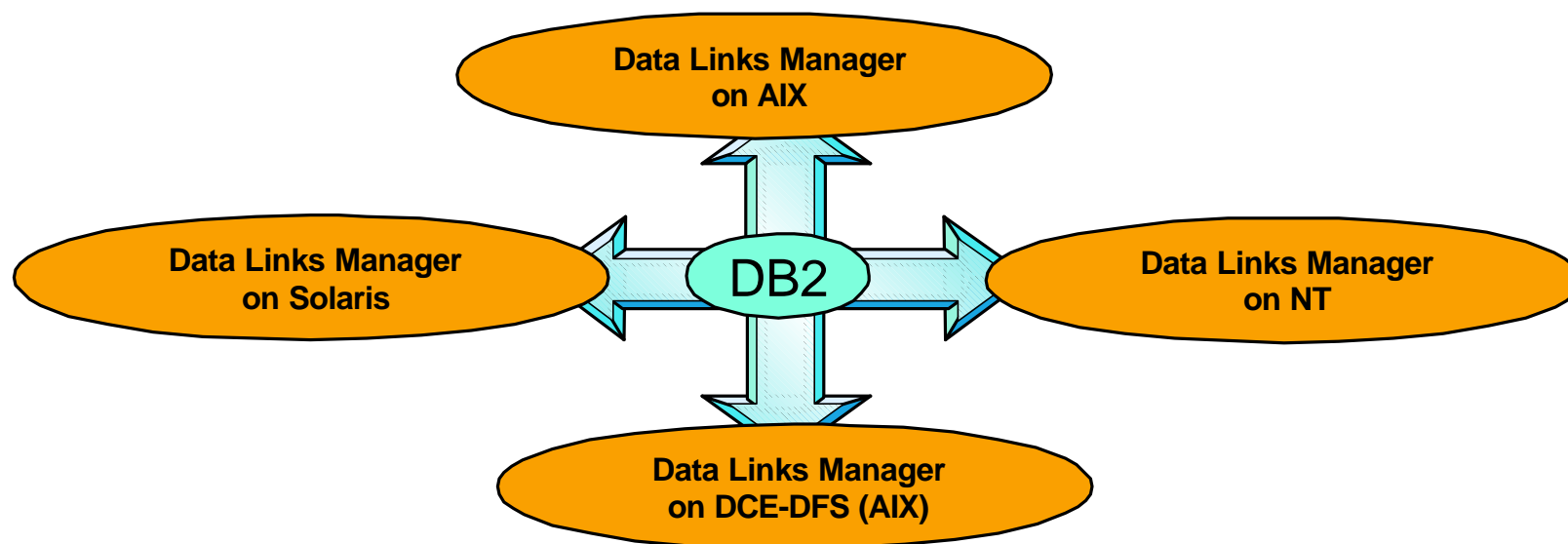


- Scenario: lots of unmanaged files
- Possible solution
 - Establish metadata and store it in DBMS
 - Use SQL to search metadata to find files of interest
 - Access files using legacy applications



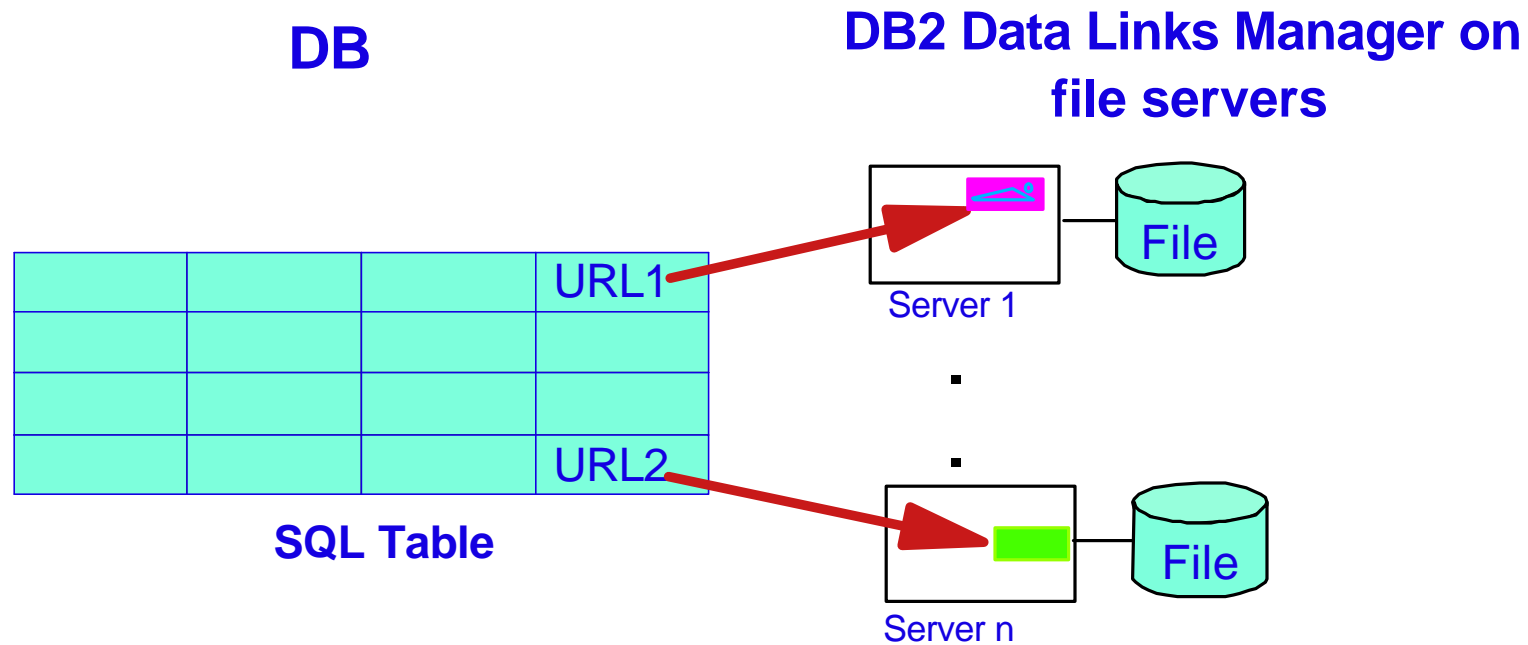
Leave file data where it is and manage it

DataLinks: Scalable Heterogeneous Solution



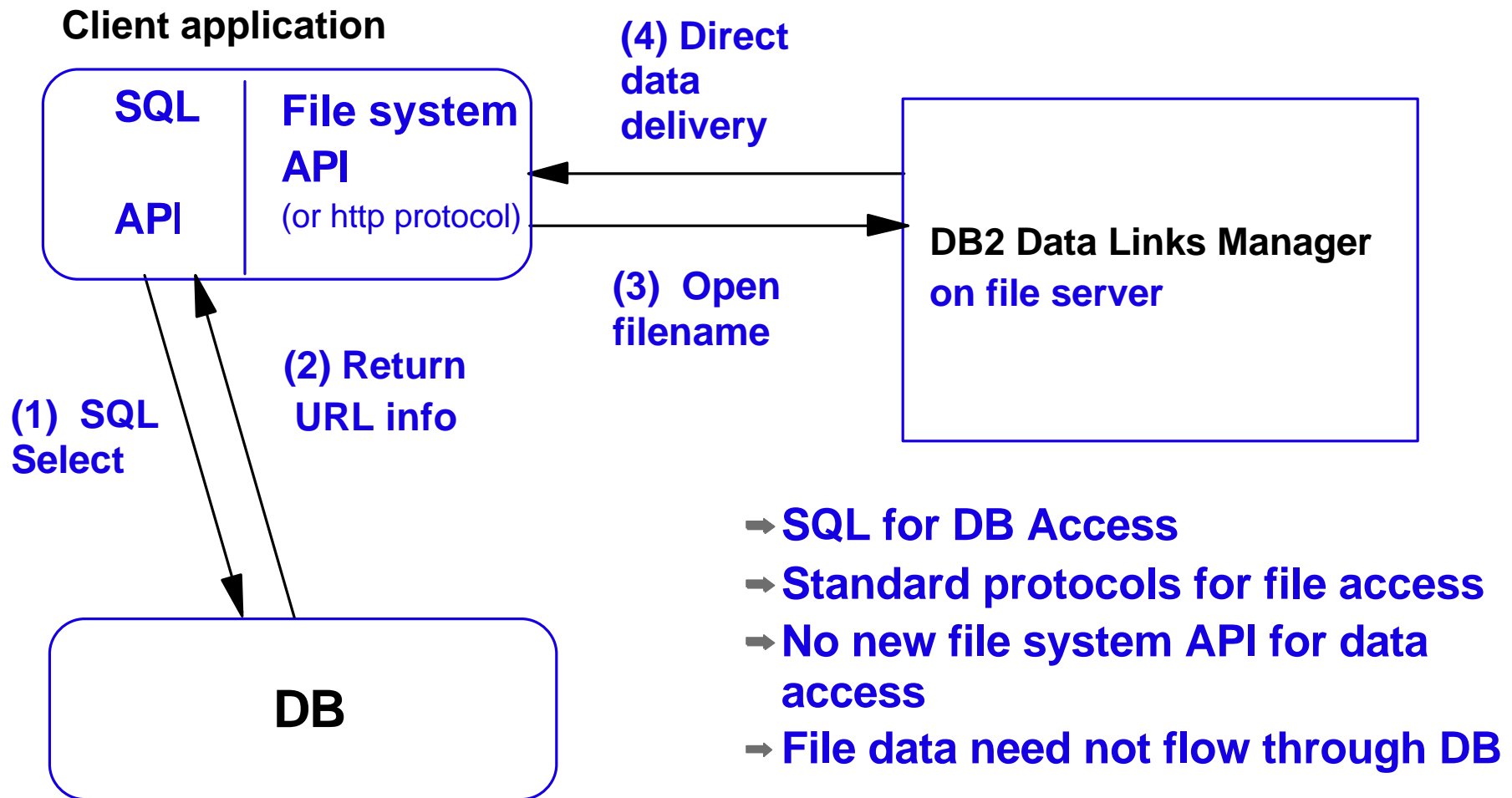
Product/Release	Platforms
DB2 UDB 5.2	AIX
DB2 UDB 6.1	AIX, NT
DB2 UDB 7.1 (06/2000, GA)	AIX, NT, DCE-DFS (AIX)
DB2 UDB 7.1 (08/2000 FP)	AIX, NT, DCE-DFS (AIX), Solaris
DB2/400 V4 Release 4 September 1999	AS/400

DataLinks Storage Model

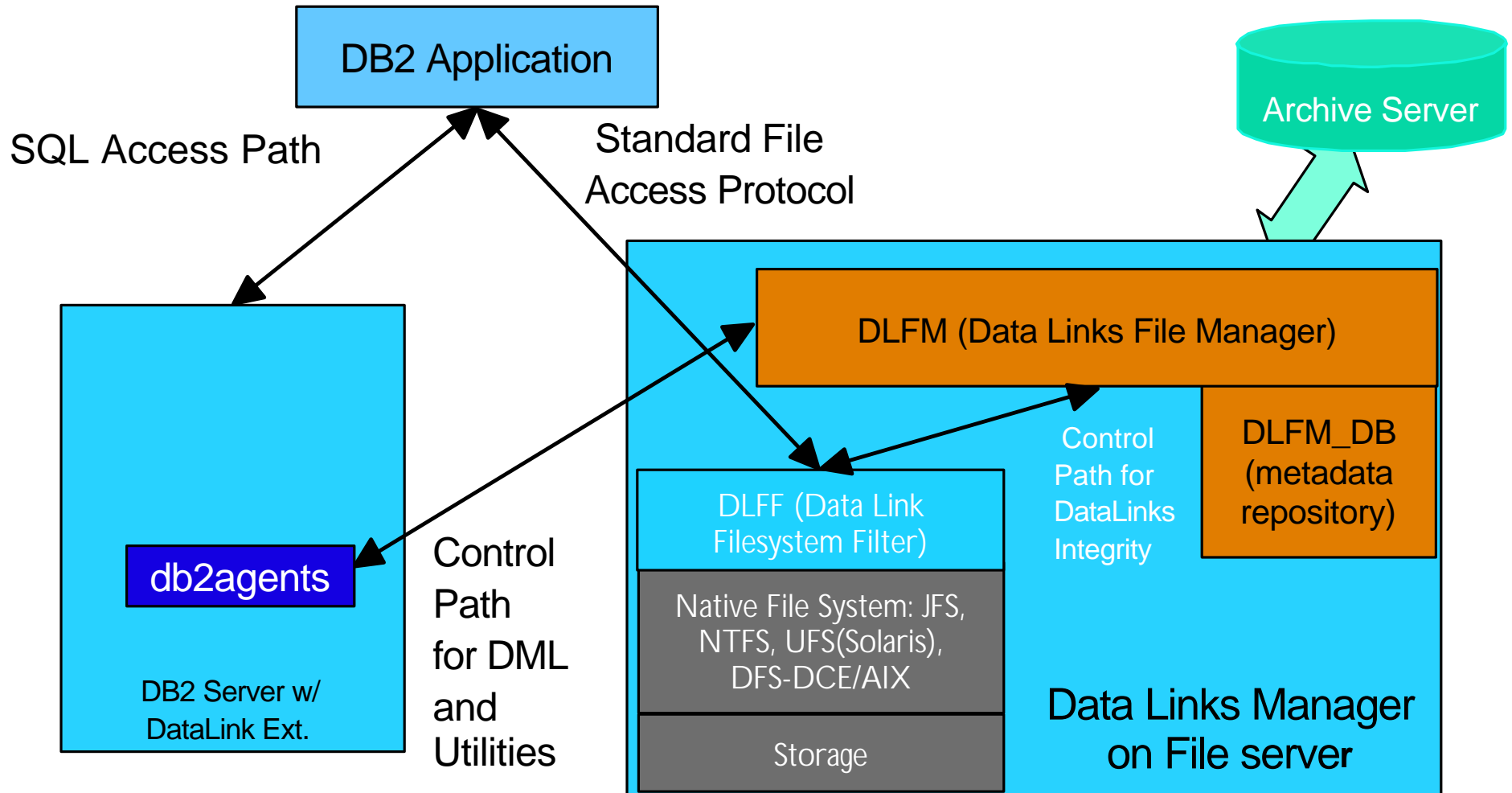


DB data linked to files in file systems in a network with Integrity, access control and coordinated backup & recovery

DataLinks Programming Model



DataLinks Architecture



Referential Integrity

Goal: No dangling file references

Enforcement:

- Disallow file rename and delete after file is linked to database

File Data Integrity

Goal: Maintain file data consistency

Enforcement:

- Disallow file updates after it is linked to the database

Coordinated Backup/Recovery

Goal: Keep file metadata in database in sync with file data

Enforcement:

- DLFM makes archive of file at INSERT time
- Materialize appropriate "version" of the file at DB media recovery time

File Access Control

Goal: Allow file access based on database authorization

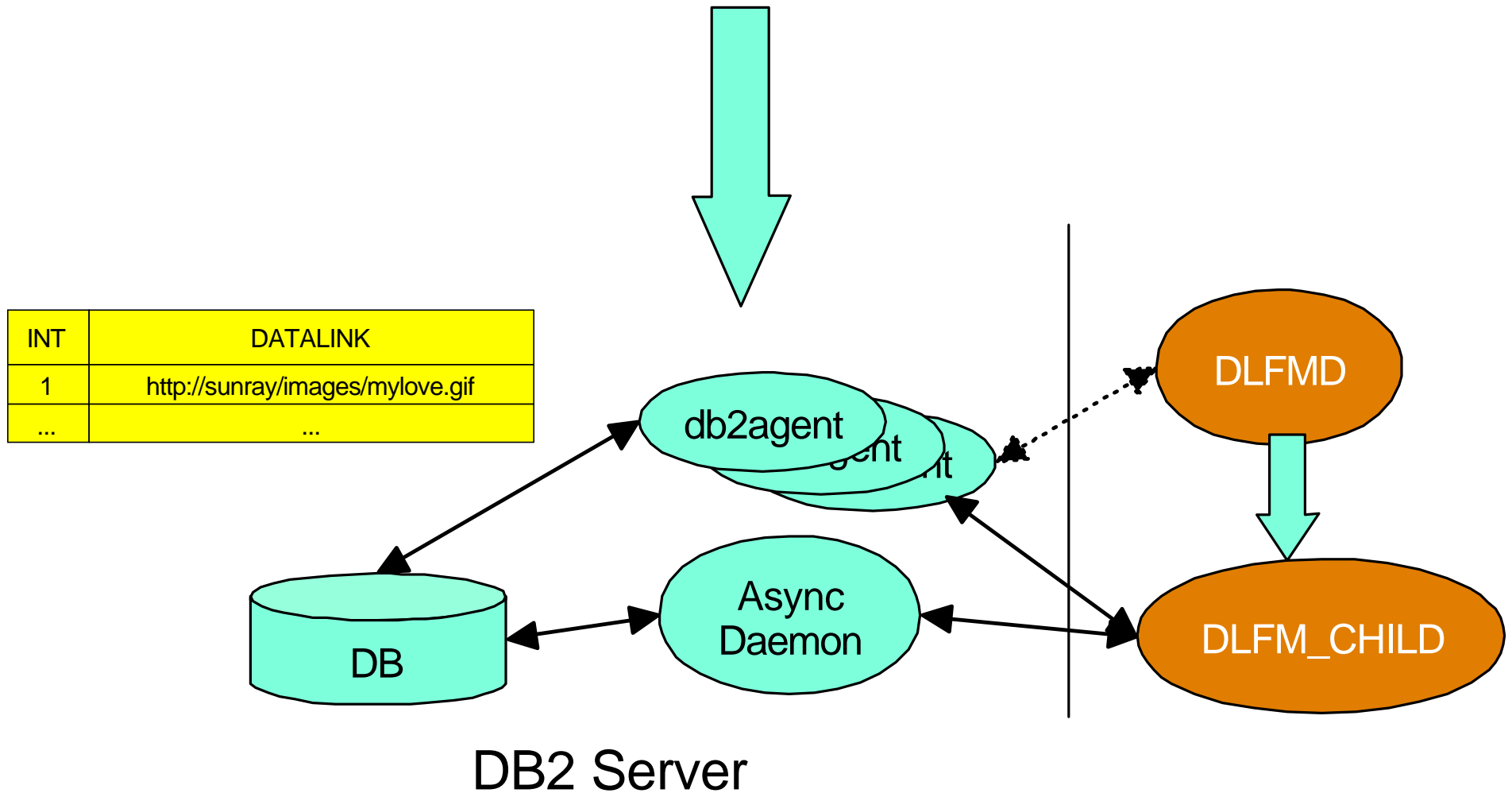
Enforcement:

- File ownership changed to database
- Access token generated on query and embedded in the filename
- Token validated by DLFF during filesystem open() call

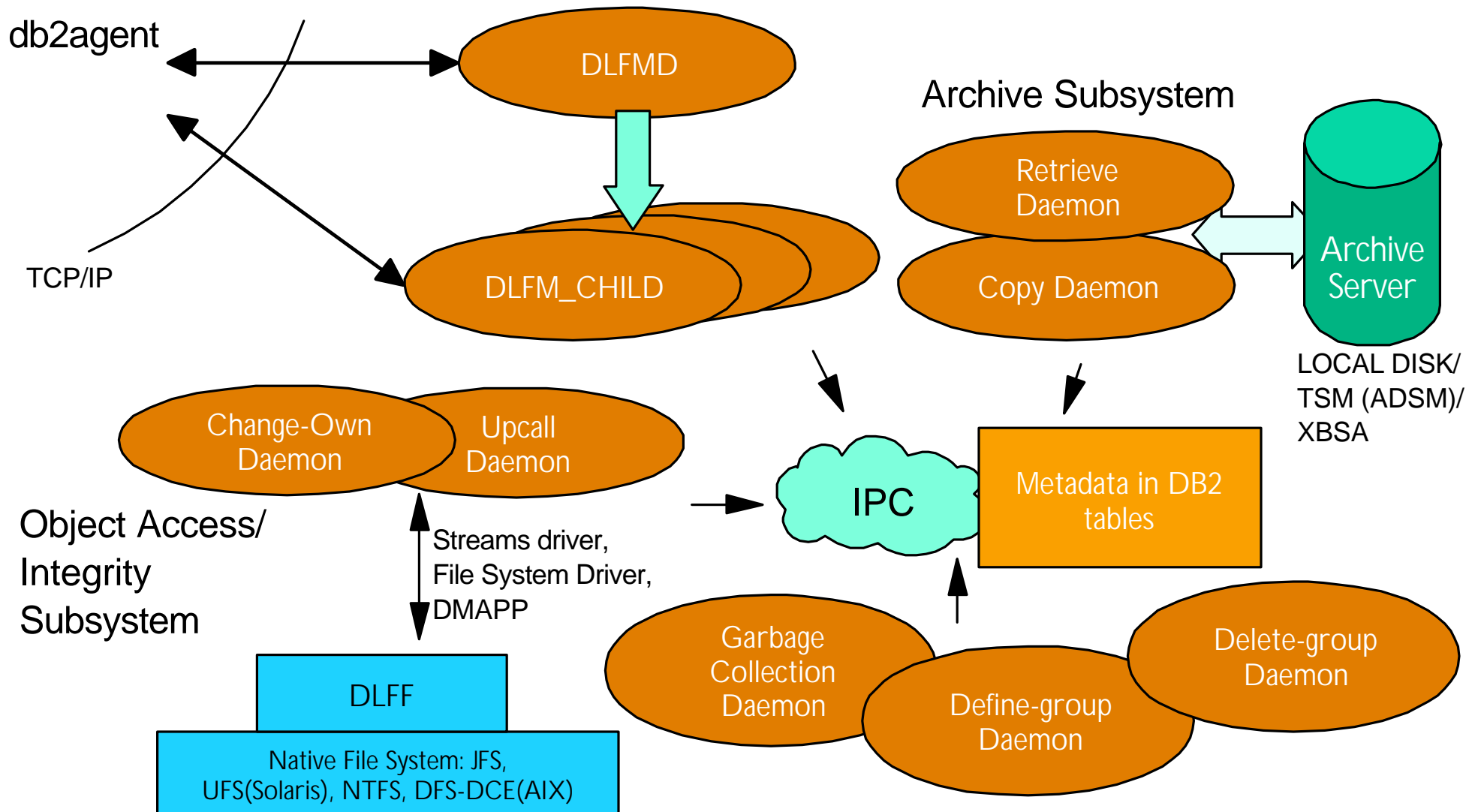
Process Model: DB2 Server

SQL: SELECT, INSERT, DELETE

UTILITIES: BACKUP/RESTORE, LOAD, IMPORT, EXPORT



Process Model: Data Links Manager



Data Links Filesystem Filter (DLFF)

- DLFF is a control layer on top of a filesystem
- DLFF is stateless
- Intercepts specific filesystem calls
 - open -- validates the embedded token if present
 - rename file and directory -- disallows for linked files
 - delete file -- disallows for linked files
- Token generation and validation

Example: `/videos/french.mpg` ==> `/videos/8QGdxt2fB7Ux5QHAU;french.mpg`

 - Shared secret between DB2 and Data Links Manager (algorithm and key)
 - Two levels of security
 - MAC_0 : encryption based on filename
 - MAC_1 : encryption based on full path name

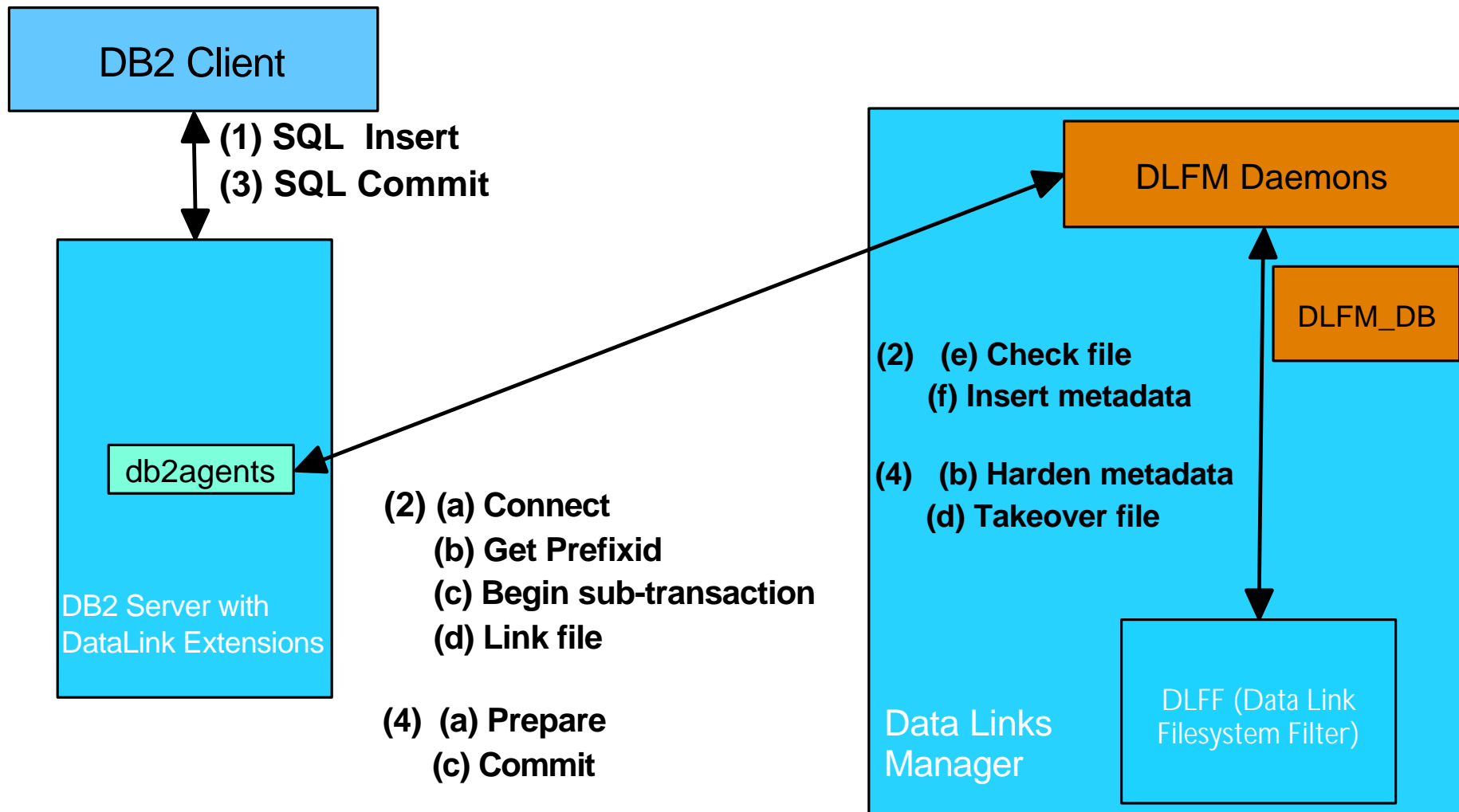
DataLinks: File System Performance

- DLFF is **NOT** in read/write path. File access performance not impacted (compared to native file system)

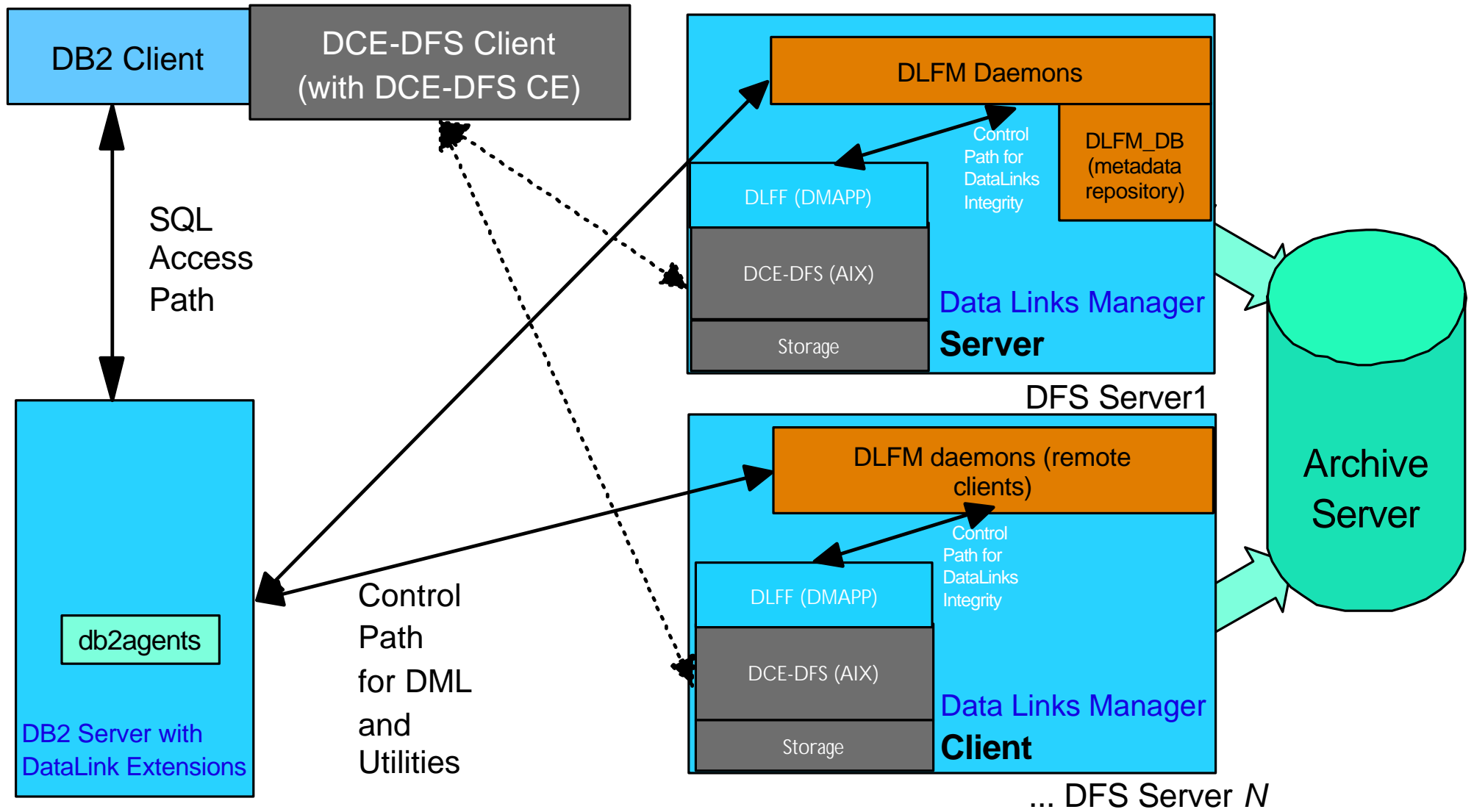


No performance loss
AND
I get referential integrity.

Control Flow on SQL Insert



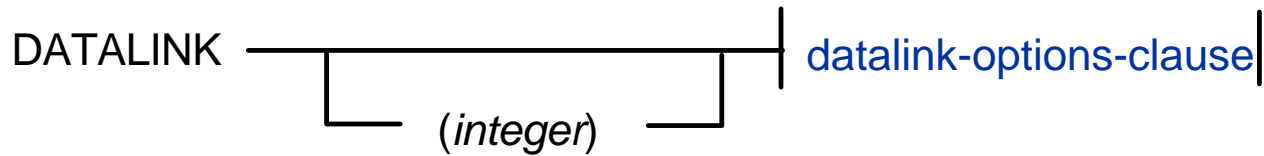
Architecture on DCE-DFS for AIX



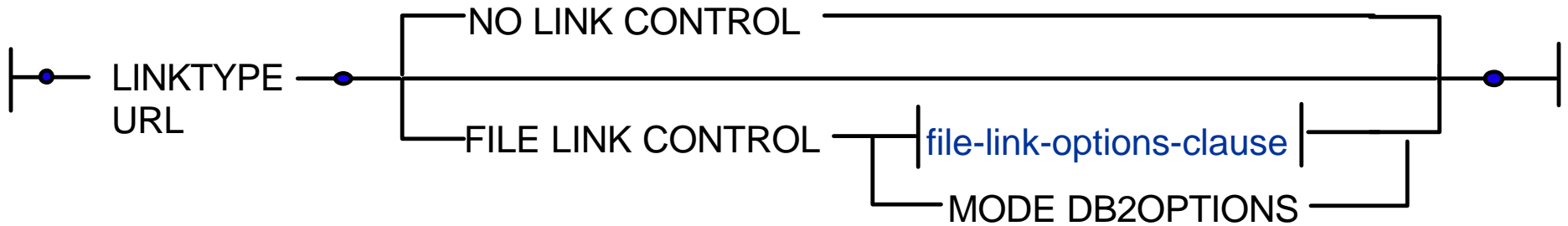
DATALINK - Datatype

- Standardization in progress
 - With ISO/ANSI: Database Language: SQL - Part 9: SQL/MED (Management of External Data). Final Committee Draft status
- DATALINK is a base datatype
- DATALINK value is a URL

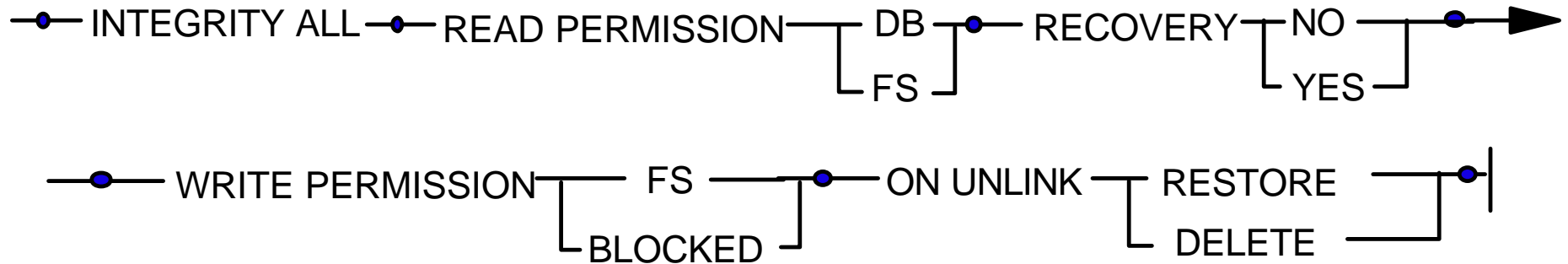
Datatype Grammar



datalink-options-clause:



file-link-options-clause:



Datatype Features

Opt #	Read	Write	Recovery	Unlink	Referential Integrity	DB Access
1	FS	FS	No	N/A	✓	
2	FS	Blocked	No	N/A	✓	
3	FS	Blocked	Yes	N/A	✓	
4	DB	Blocked	No	Delete	✓	✓
5	DB	Blocked	Yes	Delete	✓	✓
6	DB	Blocked	No	Restore	✓	✓
7	DB	Blocked	Yes	Restore	✓	✓

Valid Combinations for FILE LINK CONTROL Options

Datatype Semantics: INSERT, SELECT

- SQL **INSERT** INTO T1 (C1, C2) VALUES (:hvc1, DLVALUE (:HVURL1));
 - hvurl1 is a host variable declared as VARCHAR
 - On UNIX:
 - :hvurl1 = "http://server1/x/y/a.b" or "file://server1/x/y/a.b"
 - File /x/y/a.b is linked on server1
 - On NT: UNC format is supported as well
 - :hvurl1 = "unc:\\server2\gdrive\x\y\a.b"
 - *gdrive* is the sharename for G:, File G:\x\y\a.b is linked on server2
 - In DCE-DFS environment:
 - :hvurl1 = "dfs://.../almaden.ibm.com/fs/x/y/a.b"
 - File /x/y/a.b is linked in cell almaden.ibm.com
- SQL **SELECT** C1, DLURLCOMPLETE (C2) INTO :hvc1, :hvurl1
FROM T1 WHERE C1=5;
 - If DATALINK column is specified with Read Permission (DB), an access-token is embedded in the path name of the url

Datatype Semantics: DELETE, DROP, UPDATE

- **DELETE / DROP** - unlink existing file(s) in DATALINK column(s) of the row
- **UPDATE** - unlink existing link and create a new link
 - Example: "http://dlnk/some/file.avi" is original value
 - update video set clip = dlvalue('http://breeze/another/file.avi')
where video_id = 5
 - commit
 - File /some/file.avi on server dlnk returned to user
 - File /another/file.avi on server breeze is linked
 - As part of one TRANSACTION.
- **SQL Restrictions** for DATALINK type columns
 - cannot be part of an index
 - cannot be part of a constraint
 - cannot be compared

Datatype - Application Support

- Scalar functions

- Construct a datalink value - dlvalue
- Get parts of a datalink value
 - dllinktype, dlurlscheme, dlurlserver, dlurlpath, dlurlpathonly, dlurlcomplete, dlurlcomment
- Example - DATALINK column in predicates
 - select dlurlcomplete(clip) from video where dlurlscheme(clip) = 'FILE' AND dlurlserver(clip) like '%ALMADEN%'

- CLI

- Functions - SQLBuildDataLink, SQLGetDataLinkAttr
- See example in Call Level Interface Guide and Reference

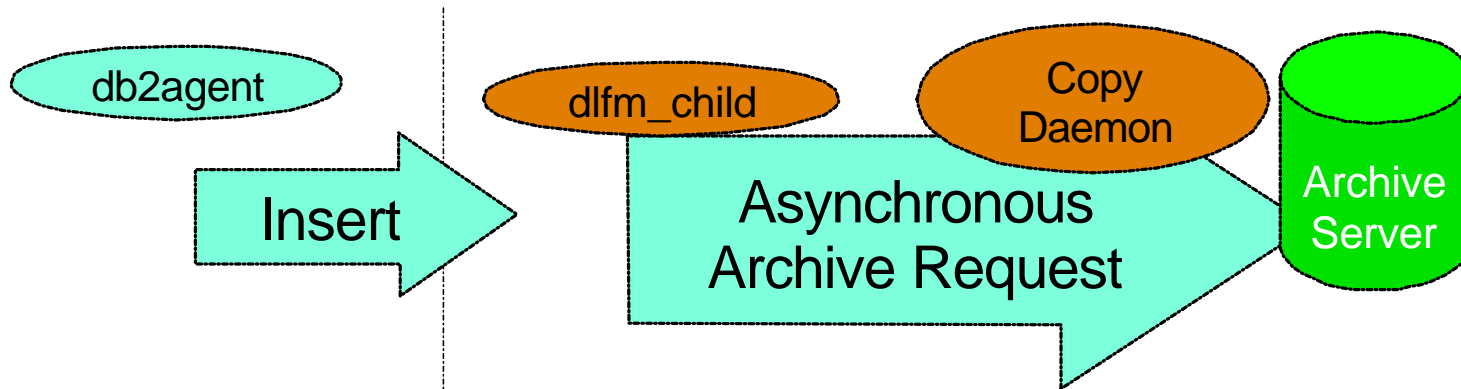
- CLP

- DATALINK column value shown on 'SELECT *'
 - (b represents a blank)
 - URLbb<url value>b<comment value>
 - NULL value in DATALINK column is shown "URL -"

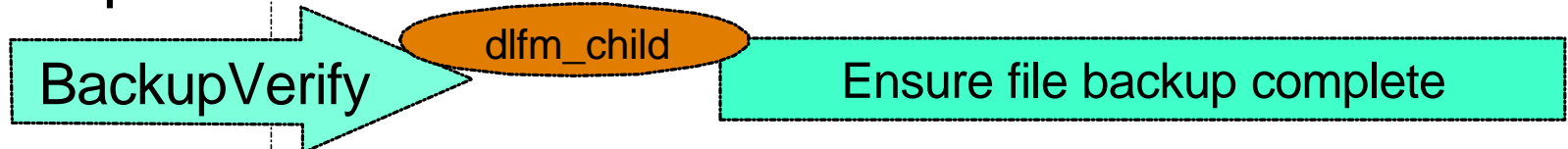
- JDBC Support: Java equivalent of DATALINK datatype is being proposed

Coordinated Backup and Recovery

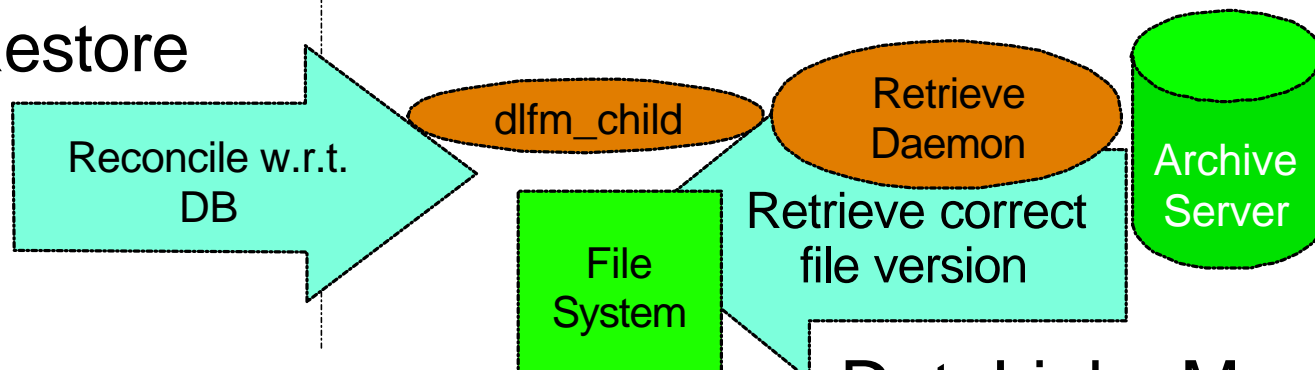
- Insert



- Backup



- Restore



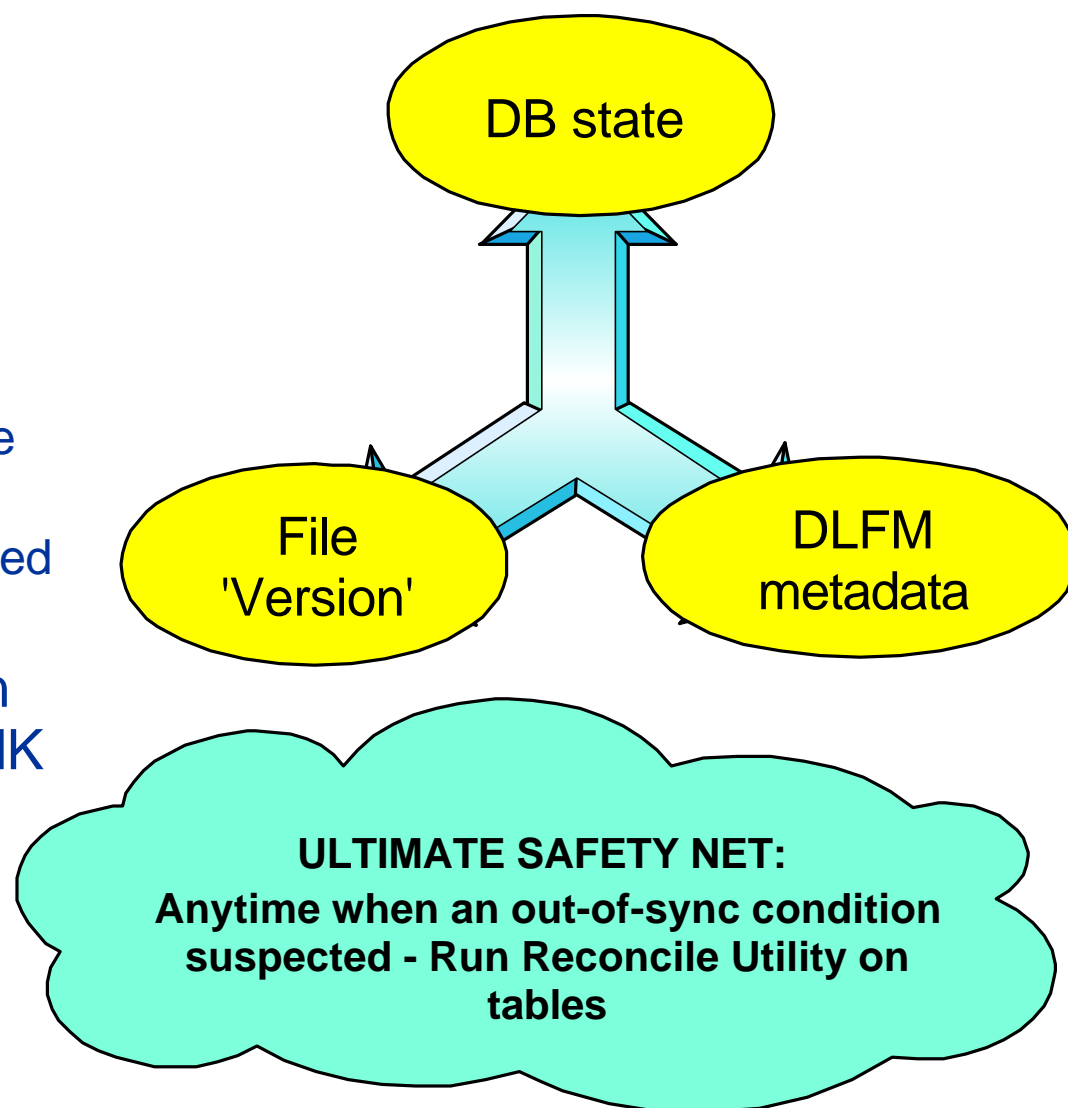
DB2

DataLinks Manager

Reconcile: Keeping Things in Sync

- **RECONCILE utility:**

- Operates at the table level
- Table is scanned and a list of files (+ version identifier) sent to Data Links Manager
- Data Links Manager, verifies and if required retrieves the file from archive server
- Unresolved file references are recorded in an exception table on DB2 side
- db2_recon_aid provides ability to run reconcile on all tables with DATALINK column
- Fast Reconcile at DB restore w/o rollforward (internal)
 - Sync done based on LSN



Bulk File Processing: Load, Export, Import

- All input formats to LOAD and IMPORT are supported
 - DATALINK SPECIFICATION provides flexibility for transforming DATALINK values in data files
- SAVECOUNT in LOAD causes consistency point for files linked in Data Links Manager(s)
- Exceptions for DATALINK column(s) are reported in the exception table
 - The DATALINK columns in the exception table must specify NO LINK CONTROL
- Load Copy and Load Replace options are not supported
- Export:
 - DB2 EXPORT command generates control file containing file references
 - DLFM-side utility '*dlfm_export*' generates a TAR (ZIP on NT) file based on control file
- Import:
 - DLFM-side utility '*dlfm_import*' uses the TAR (ZIP on NT) file to materialize files prior to running IMPORT command on DB2 side
- See "DB2 Data Movements Guide" for details

DLFM/DB2 Crash Recovery - Salient Features

- DB2 server drives resolution of in-doubt transactions on DLMs
 - DLM is passive in this regard
- Asynchronous daemon on DB2 server resolves in-doubts on DLMs unavailable during restart
 - DB2 restart is NEVER held up
- DB2 does not allow normal processing on that DLM until in-doubts are resolved
- In DCE-DFS DLM, restart recovery can proceed through any DFS server as long as DLFM_DB is available

Data Links Manager: Disk Crash Recovery

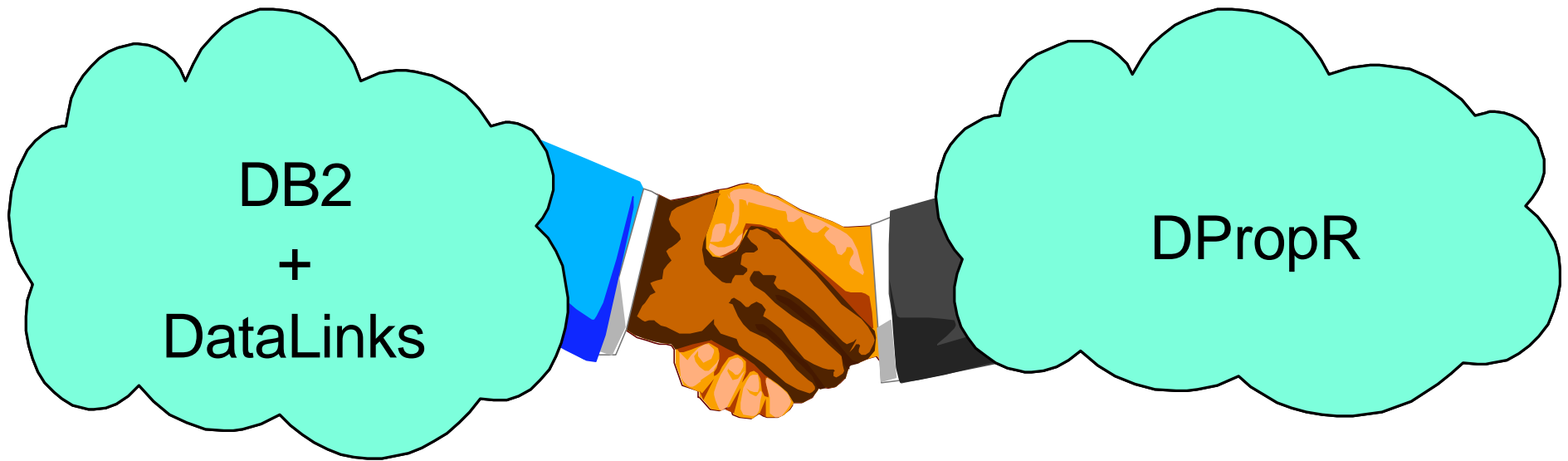
Problem statement

- DL file system OR DLFM_DB metadata OR both can be lost as a result of disk crash
- Need to recreate: files in filesystem, DLFM_DB metadata

Logic

- Non recreatable DLFM_DB metadata (prefix-id, dbid) always backed up
- DB2 contains file (+version) information. This is flowed to DLM
- Appropriate file version retrieved from archive server
- Rest of DLFM_DB metadata recreated from file archive header

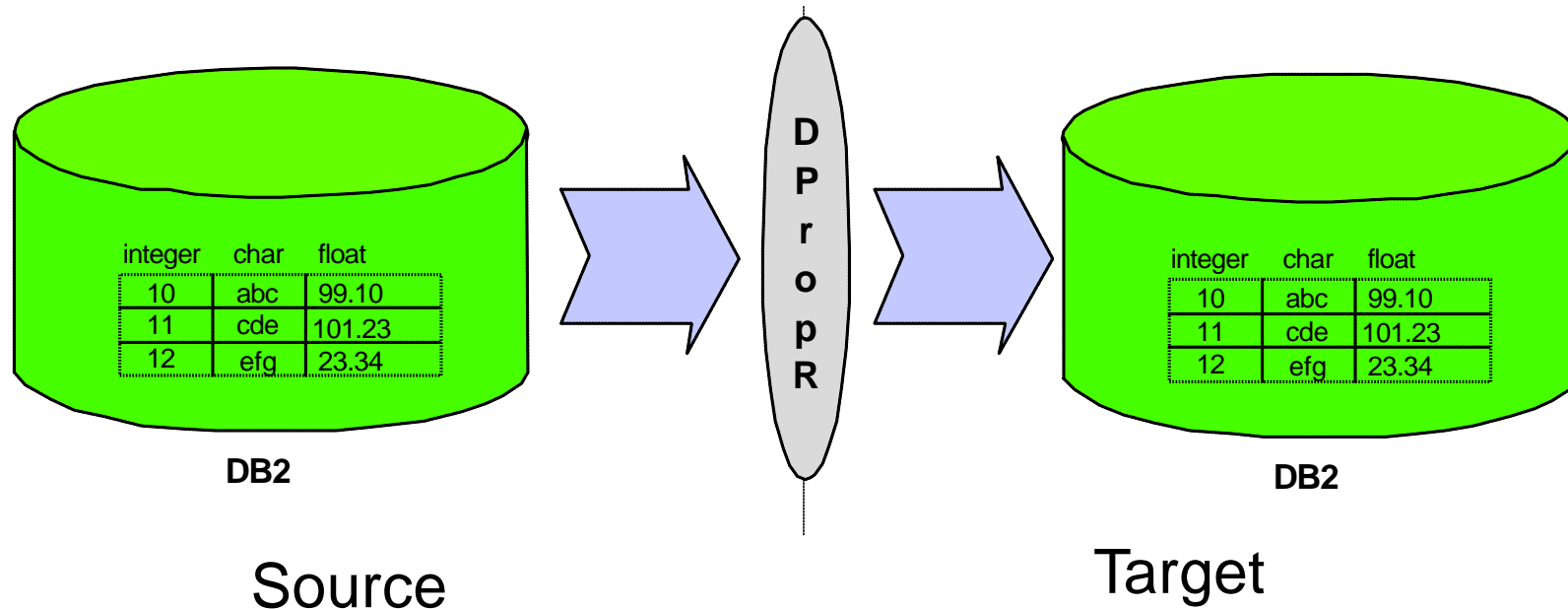
DataLinks Replication



Why Replication ?

- Performance
 - allow data to be stored close to application
 - reduce network traffic
- Availability
 - provide failover in case of system failure
- Security
 - in B2B, replicate only data which the business partners need
 - prevent external access to the primary system
- Load balancing
 - provide multiple copies of a system image for better load balancing

Traditional Database Replication

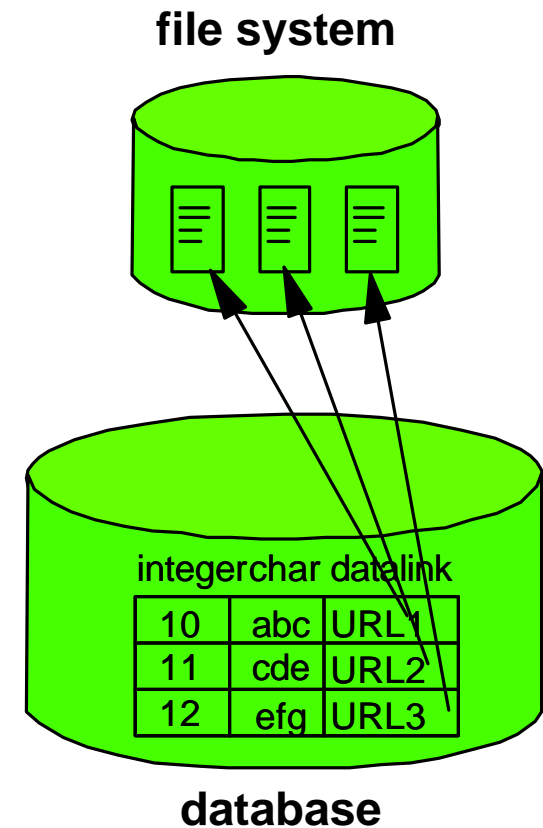


- Solution for data replication between DBMS has existed for a long time
- DB2 DataPropagator (DPropR) is an IBM solution for data replication
- This solution works for **traditional** business data

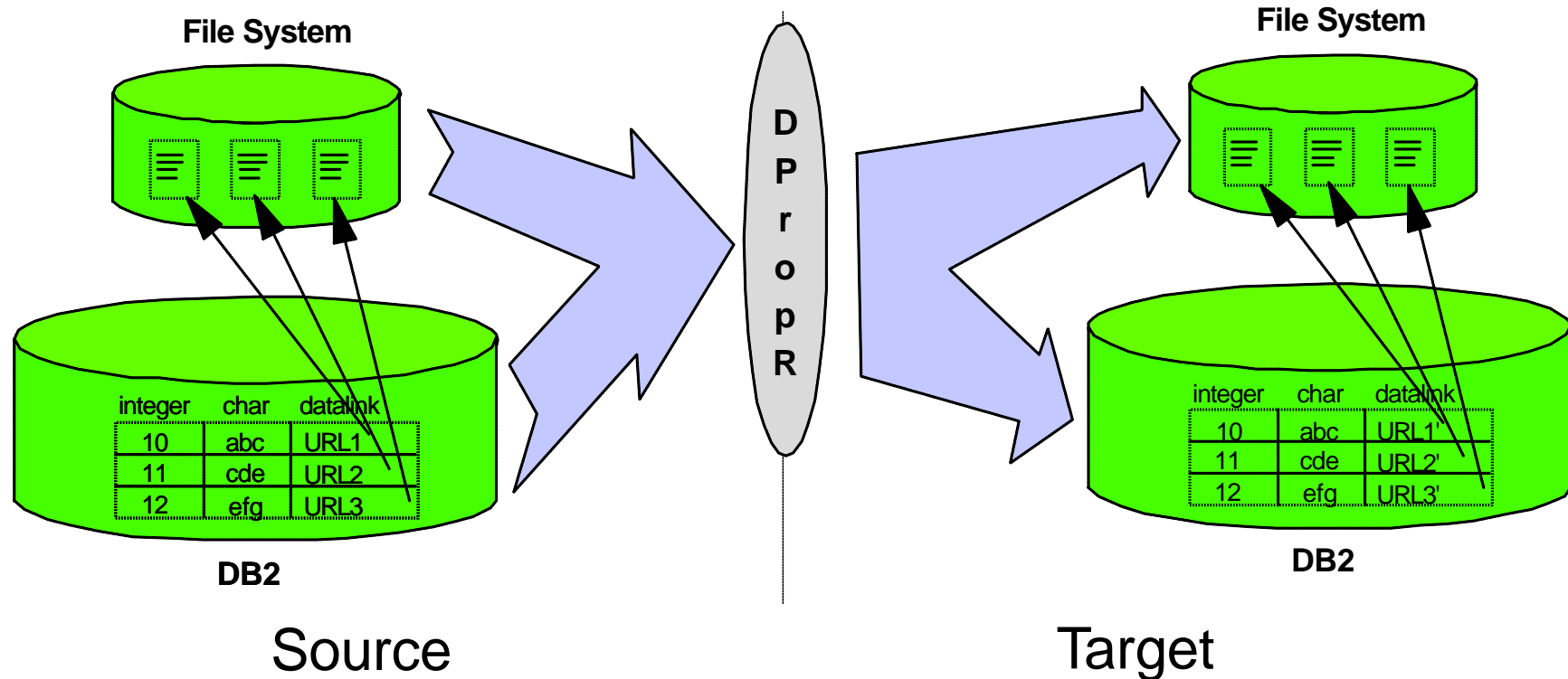
How about data in the external files?

- How would we replicate the file data along with the metadata stored in the database in an **automatic** and **consistent** way?

→ *DataLinks Replication is the key*

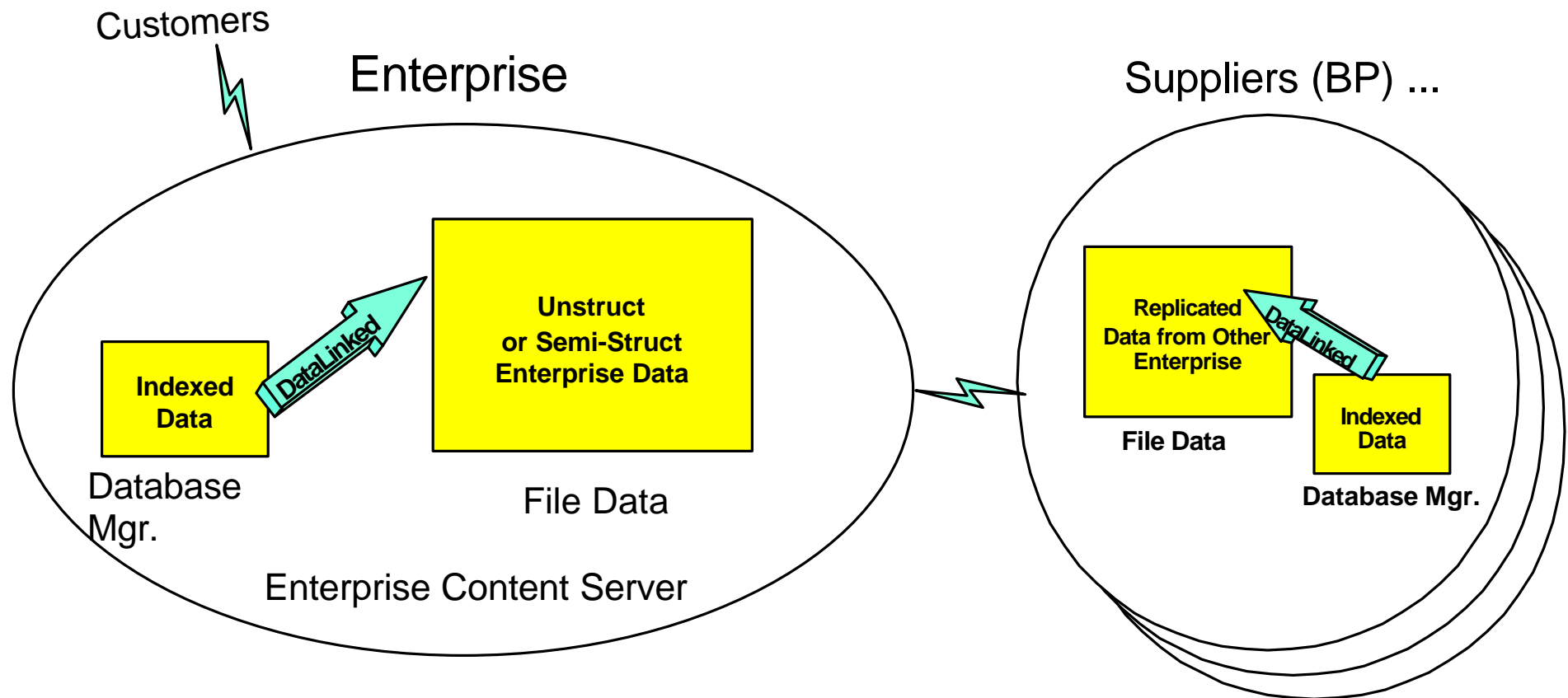


What DataLinks Replication Can Do?



Both database data and external referenced files will be replicated **together** in an **automatic** and **consistent** way.

DataLinks Replication Enables B2B



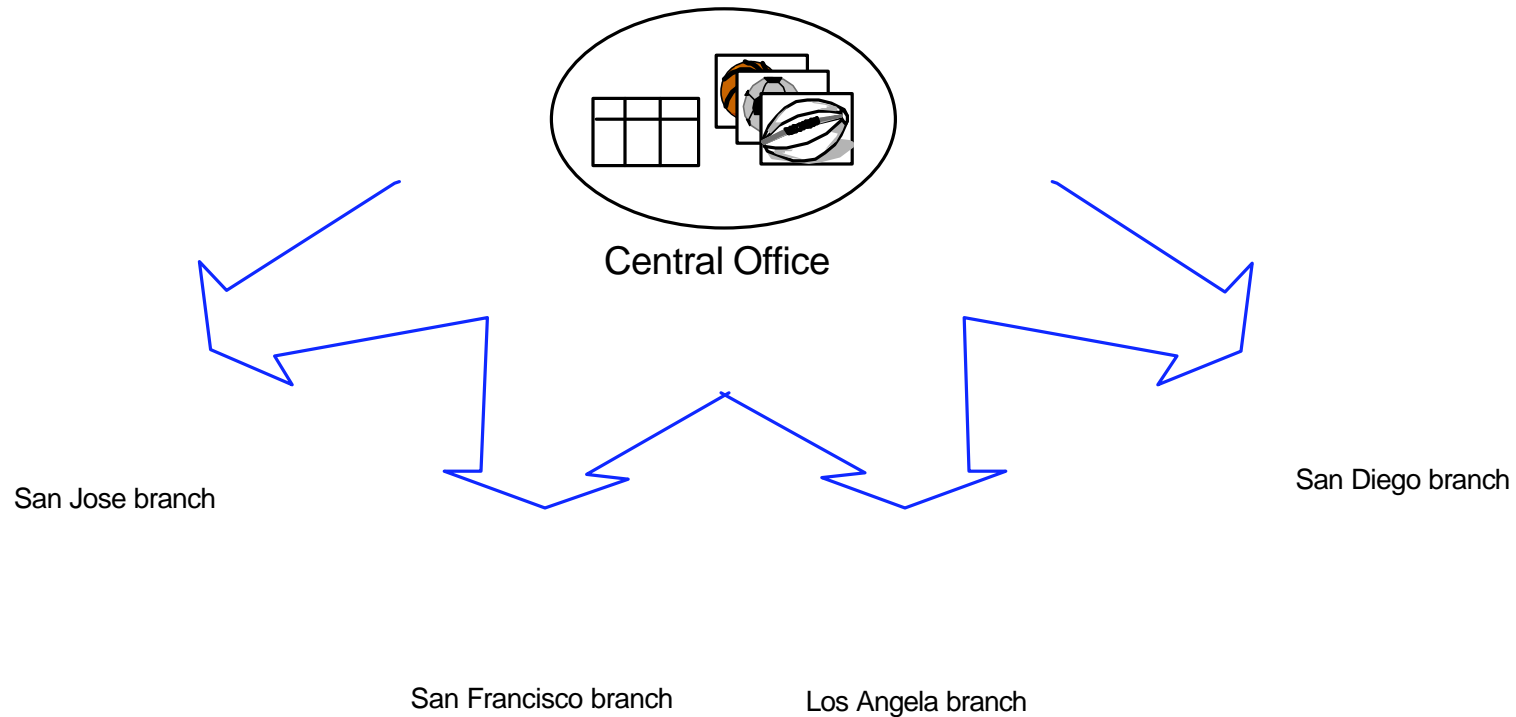
Automated and Integrated Solution for Content Sharing in B2B

- Replication of file data and metadata between enterprise and business partners
- Supports selective file data and metadata replication

B2B Solution (contd.)

- Supply Chain (SC) model is common in automotive and aerospace industries for engineering designs
 - Example: A large automotive manufacturer outsources 70% of a vehicle design
 - Requires content sharing between different enterprises
 - Needs replication of both engineering drawings (files) and metadata (database)
- B2B Customer Support Document System
 - Example: A large airplane manufacturer needs to deliver airplane maintenance documents in common format to airlines

Another Example: Catalog Distribution System

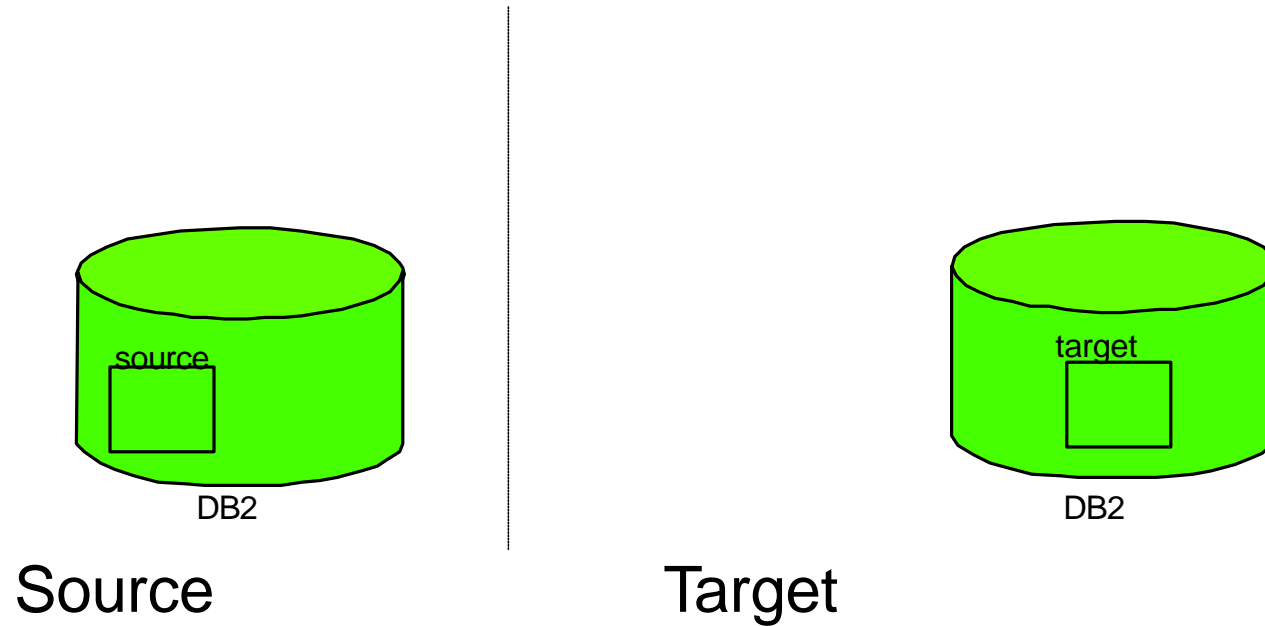


- Catalog creation is done in a central office and distributed to branches for availability, performance and cost savings
- Both the catalog drawings and metadata are periodically distributed

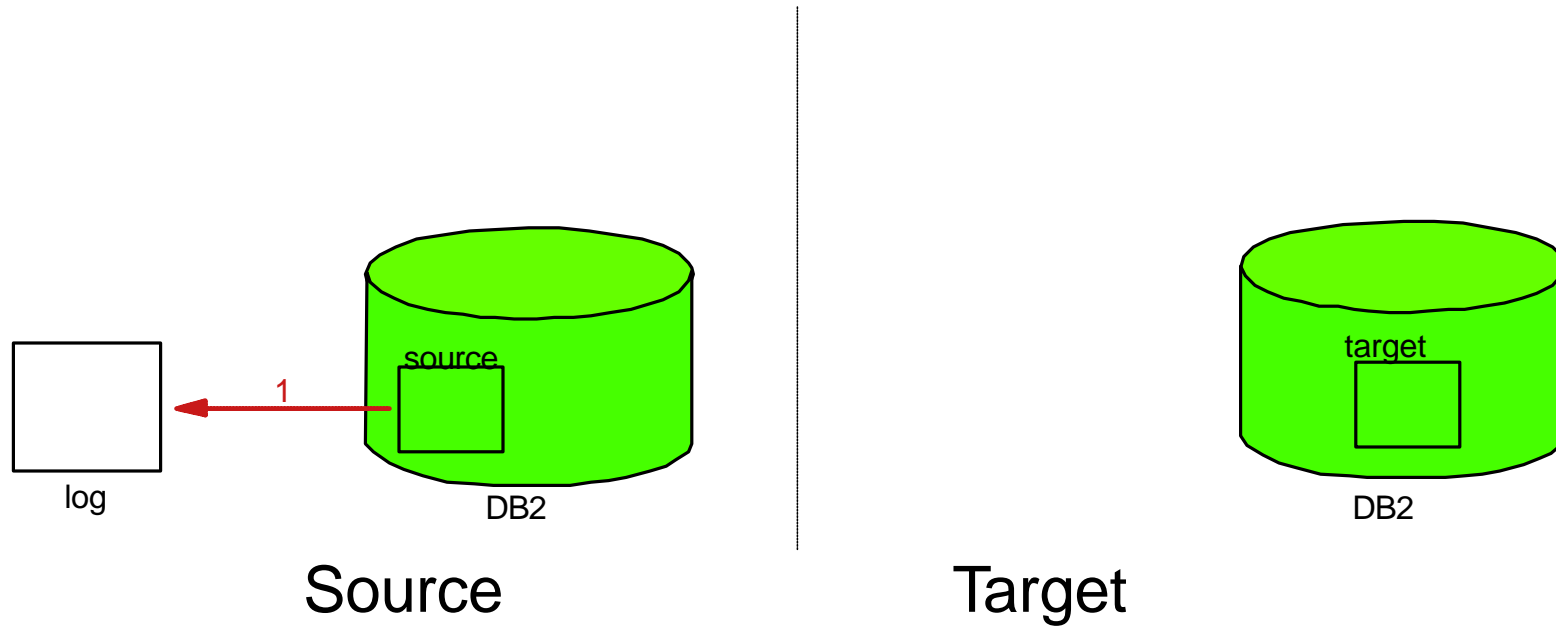
DataLinks Replication Technology

- Integrated with DB2 DataPropagator (DPropR)
- Replicate external files first and then metadata
- DPropR is extended to invoke a customizable user exit for file replication
- Update-anywhere (bi-directional) replication is also supported, however with NO CONFLICT detection
- Available in DB2 V7.1 (06/2000, GA) on AIX, NT & AS/400

DPropR Architecture

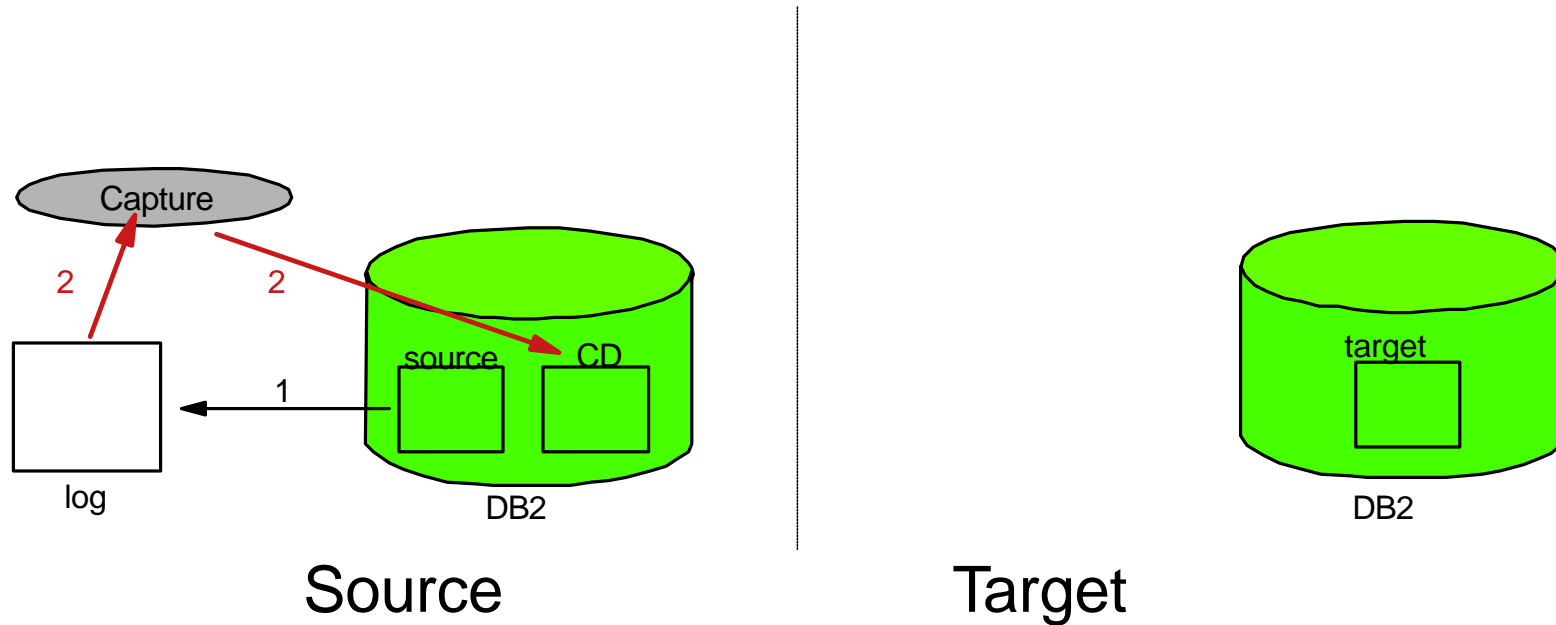


DPropR Architecture (cont.)



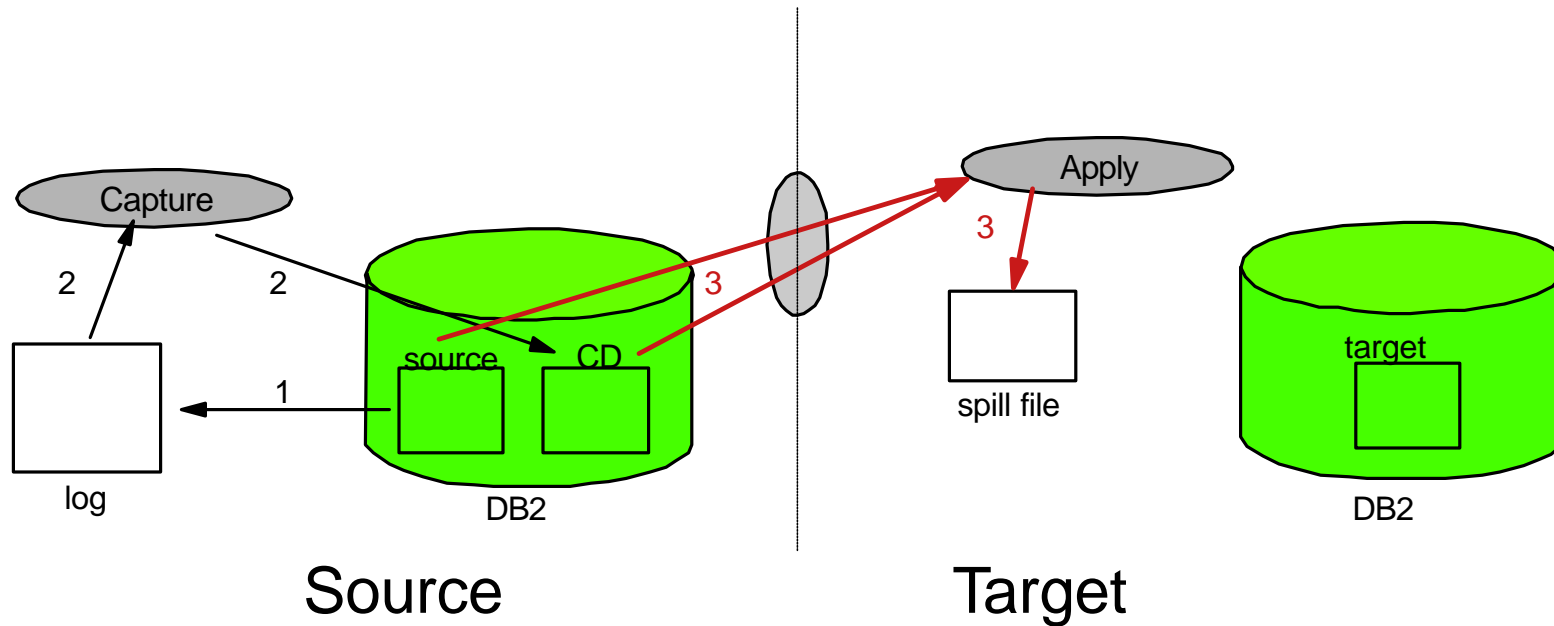
1. INSERT/UPDATE/DELETE are recorded in the database log.

DPropR Architecture (cont.)



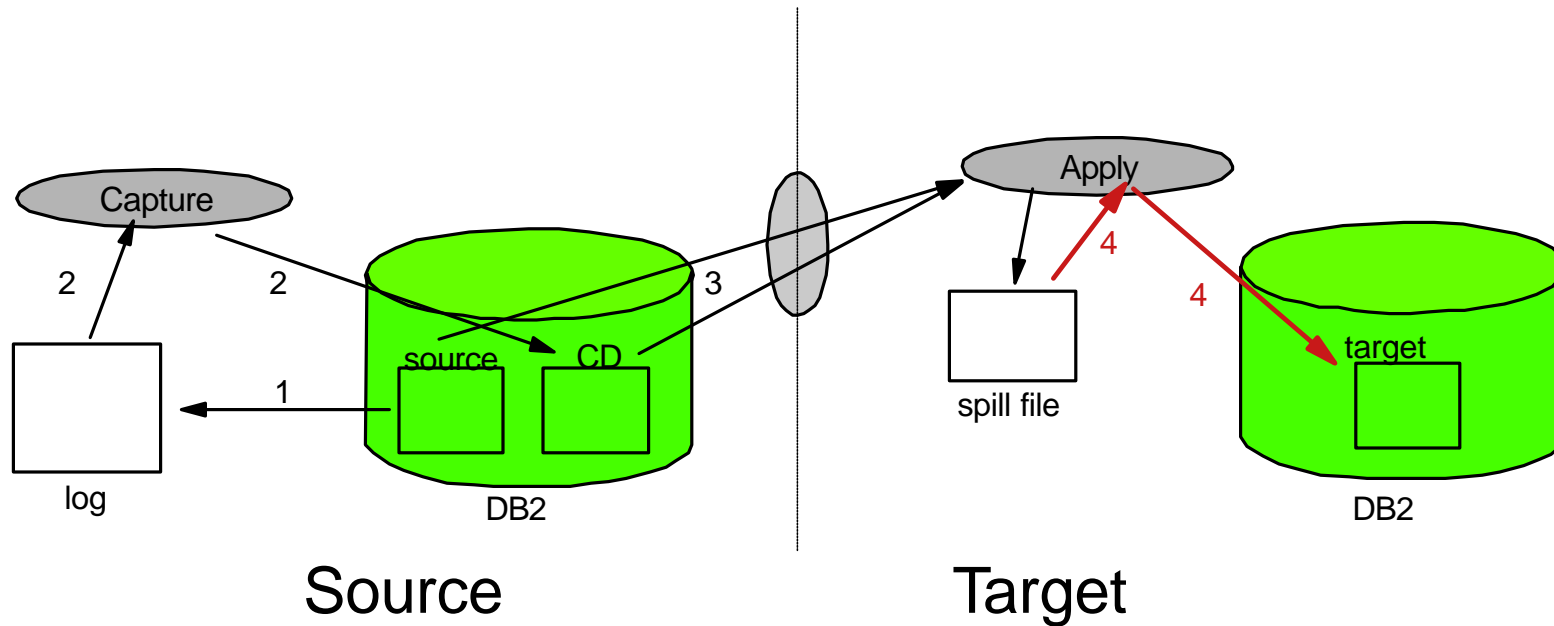
1. INSERT/UPDATE/DELETE are recorded in the database log.
2. Capture reads the log and stores the changes to the CD table.

DPropR Architecture (cont.)



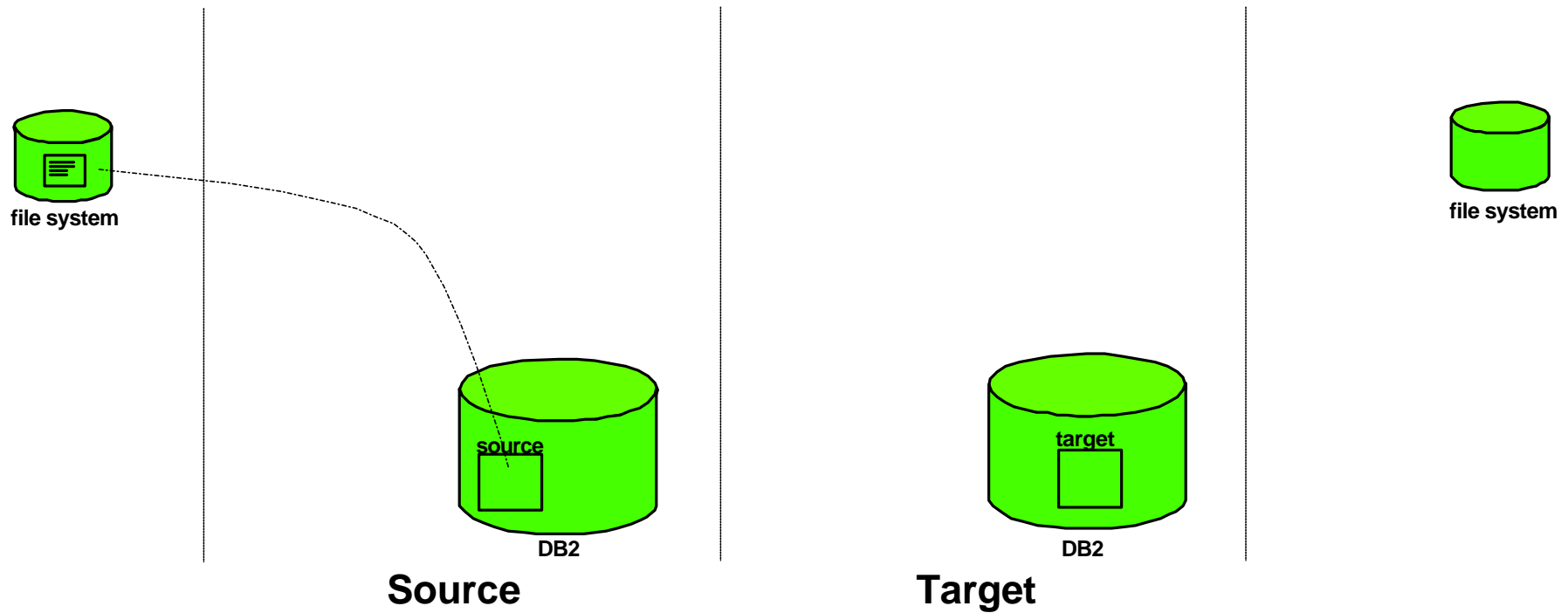
1. INSERT/UPDATE/DELETE are recorded in the database log.
2. Capture reads the log and stores the changes to the CD table.
3. Apply reads the CD table and copies the changes to a spill file.

DPropR Architecture (cont.)

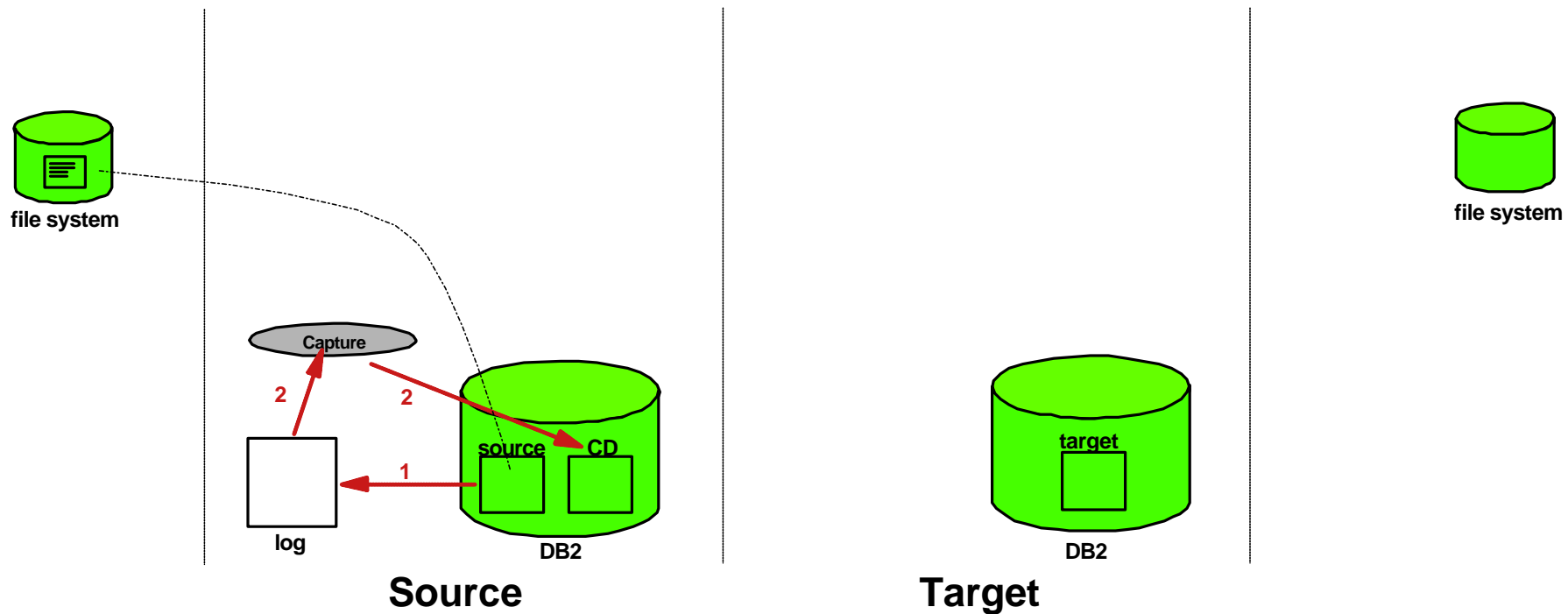


1. INSERT/UPDATE/DELETE are recorded in the database log.
2. Capture reads the log and stores the changes to the CD table.
3. Apply reads the CD table and copies the changes to a spill file.
4. Apply propagates the changes from the spill file to the target table.

DPropR with DataLinks support

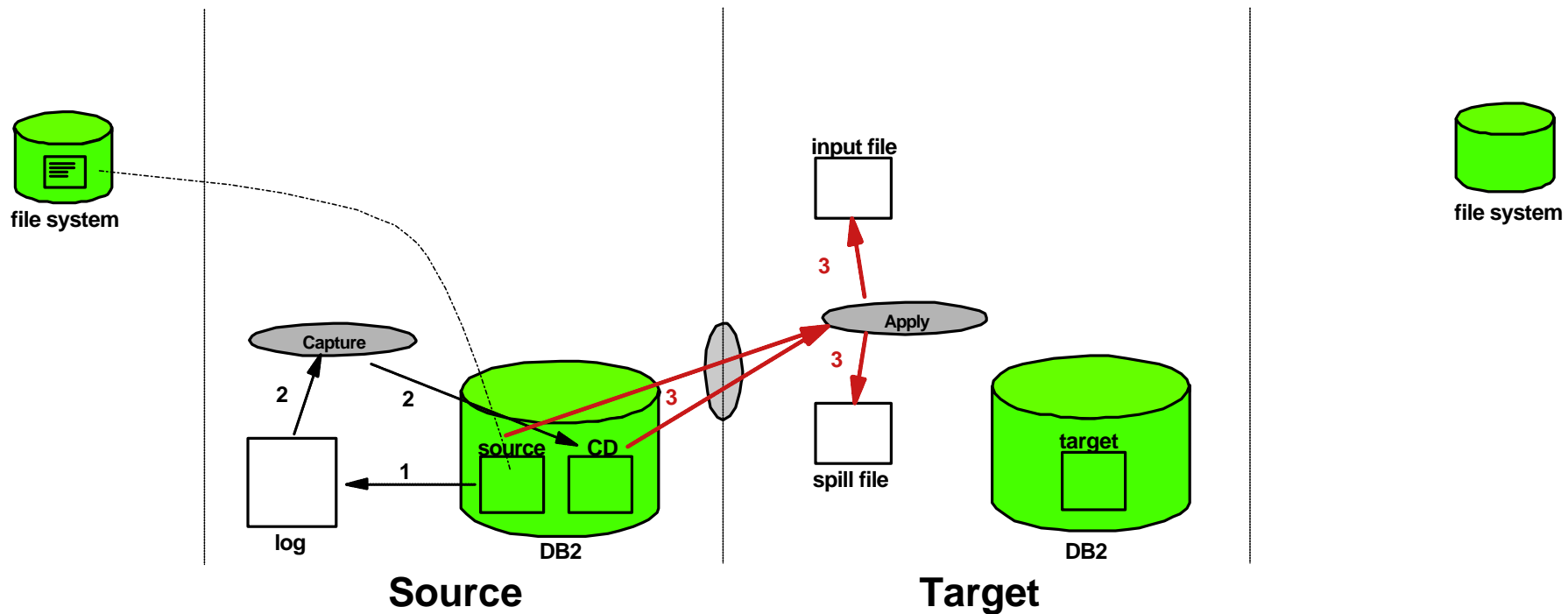


DPropR with DataLinks support



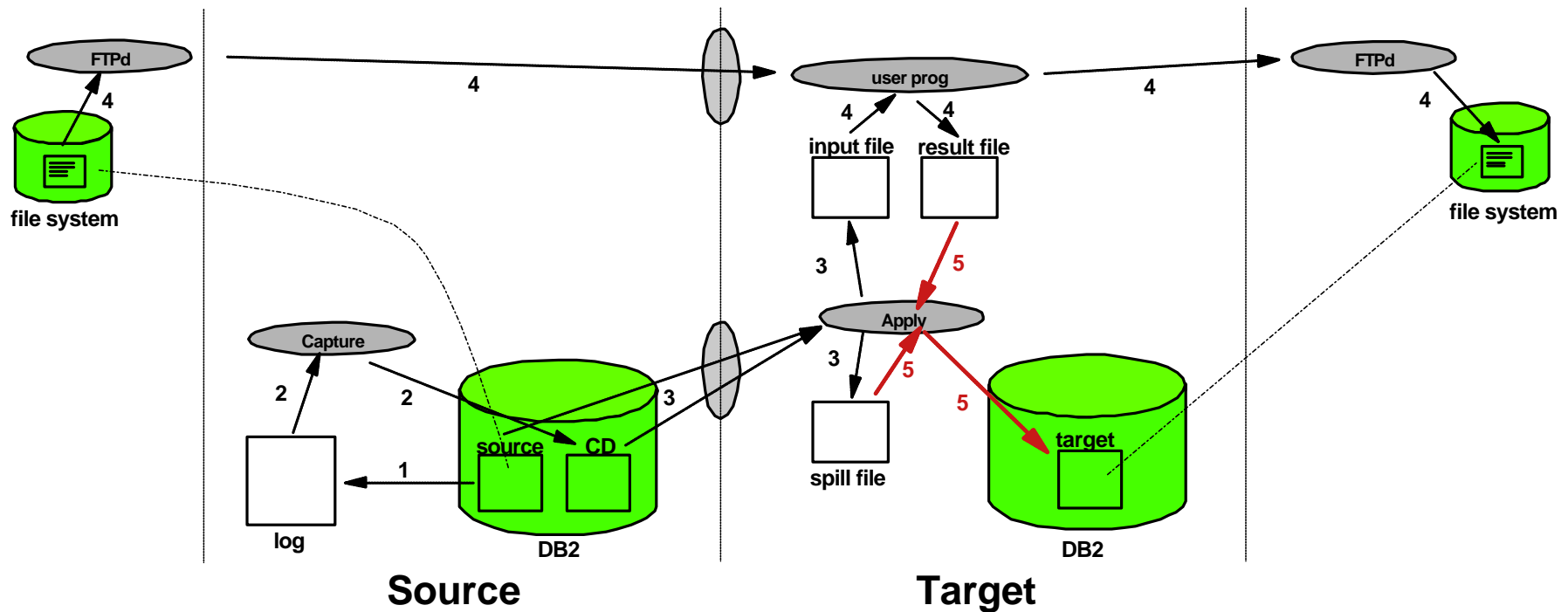
1. Datalink changes are recorded in the database log.
2. Capture reads the database log and stores the changes to the CD table.

DPropR with DataLinks support



1. Datalink changes are recorded in the database log.
2. Capture reads the database log and stores the changes to the CD table.
3. Apply copies the change data from the CD table to a spill file. At the same time, stores the Datalink file references in a separate file (input file).

DPropR with DataLinks support



4. User exit program maps the file references and copies them from the source file system to the target file system through FTP.
5. Apply propagates the metadata and the new DATALINK file reference from the "result file" to the target table.

DataLinks: Customers/Applications

- Dassault - CAD/CAM ISV
- Boeing - Manage Millions of Engineering Drawings
- Large Automotive Manufacturer in US/Europe
- Foundation for IBM's Content Management Solution
- Potential applications
 - Mail Server - manage attachments to mails
 - Version Control
 - Web Asset Management (eLinks v2.0)
 - Scientific Data Management

Summary

- Explosive growth in data stored in files
- Management of file data critical for eBusiness
- DataLinks extends DB management of data to files
- DataLinks: simple, scalable, multi-platform solution
- DB2 DataLinks Replication automatically and consistently replicates metadata and file data

Appendix A: Terminology

- Access Token
 - Embedded token in the filename that can be used to open files that are owned by the DB
- DATALINK
 - The datatype, the technology is referred to as DataLinks
- DLFF
 - DataLinks Filesystem Filter. Sub-component of DLM. Kernel/user level subsystem that sits atop the native filesystem to provide RI and access control
- DLFM
 - Data Links File Manager. Sub-component of DLM dealing with file metadata processing, user-process level daemons. Interacts with DLFF and DB2
- DLM
 - Data Link Manager. DataLinks application that is installed on the file server
- DPropR
 - IBM's replication technology
- Prefix
 - The mount point of the DLFF monitored filesystem

Appendix B: Publications

- DataLinks White paper:

DataLinks: Managing External Data with DB2 Universal Database, Judy R. Davis.
Prepared for the IBM Corp. by Database Associates International, February 1999
<http://www.software.ibm.com/data/pubs/papers/datalink.html>

- "IBM's DB2 6.1 Strengthens Web Appeal", Maggie Biggs, Infoworld Test Center, June 28, 1999, issue of Infoworld
- "How to Improve RDBMSes -- Seven long-term requirements for managing complex data", Nagraj Alur and Judy R. Davis, Byte Magazine, April 1997
- "IBM emerges as early leader in burgeoning content management market", ContentWatch, September 1997, www.contentgroup.com
- "DataLinks - Linkage of database and filesystems", Inderpal Narang and Robert Rees, Proceedings of the Sixth High Performance Transaction Systems (HPTS), September 1995
- "An Interview with Don Haderle", David Stodder, DB2 Magazine, Summer 1998
- "Universal Database Management", Chapter 9, Cindy Saracco, Morgan Kaufmann Publishers, Inc.
- "Reaching Beyond the Database", Peter Gwynne, IBM Research Magazine, Number 3, 1998, <http://www.research.ibm.com/resources/magazine>
- "A distributed Scientific Data Archive Using the Web, XML and SQL/MED", Mark Papiani et al, ACM SIGMOD Record, Vol 28, Number 3, September 1999
- "DLFM: A Transactional Resource Manager", H. Hsiao and I. Narang, to appear in ACM SIGMOD/PODS 2000
- DB2 Administration Guide, DB2 SQL Reference, DB2 Data Links Manager Quick Beginnings, DB2 Data Movement Utilities Guide.

Appendix C: Table States

- A table with DATALINK column can have additional states
 - DRP (DataLink Reconcile Pending)
 - Data Links Manager metadata is out-of-sync with table data
 - DRNP (DataLink Reconcile Not Possible)
 - Data Links Manager metadata is missing for table
- RESTORE and ROLLFORWARD utilities may set these states.
- Table access is restricted in these states
 - Select is allowed. Insert/delete is disallowed. Update is allowed selectively in DRNP state
 - User may set state to DRP if integrity compromise is suspected
 - Select access may also be disallowed by setting the CHECK PENDING state in addition to DRP state by
 - SET INTEGRITY FOR table TO DATALINK RECONCILE PENDING
- RECONCILE utility should be run to bring table out of DRP state (i.e. consistent with DLM)
- See SQL and DB2 Administration Guides for detail

Appendix D:

DB2 Server Setup

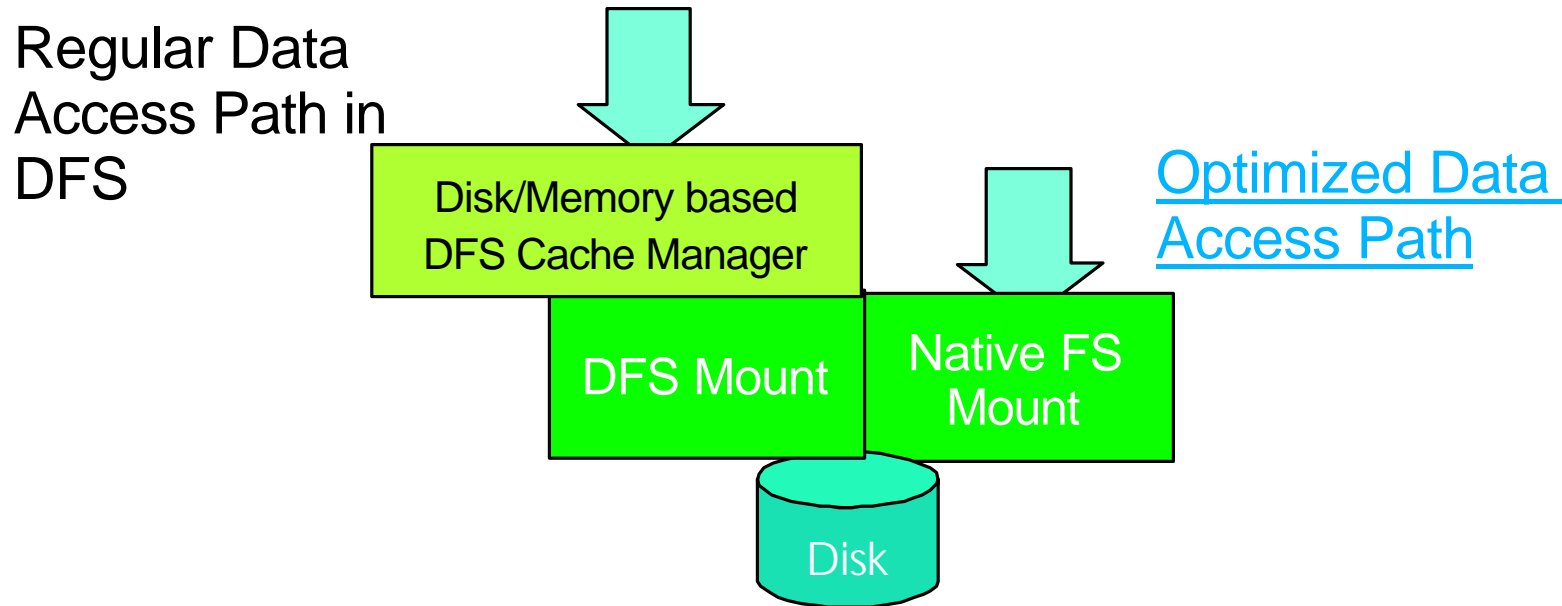
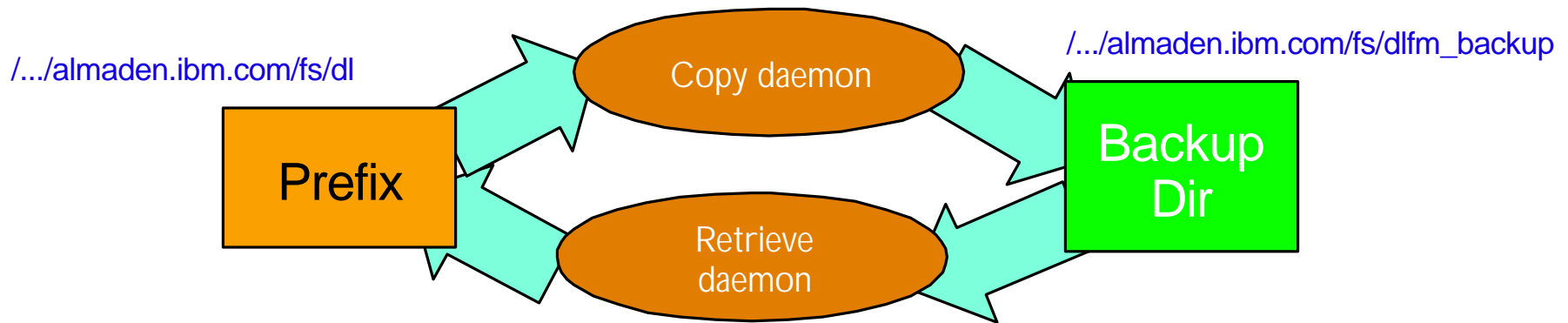
- Set database manager config parameter DATALINKS to Yes
 - db2 update dbm cfg using datalinks yes
- Add each Data Links Manager to a database config file - datalink.cfg
 - Example: ADD DATALINKS MANAGER FOR DB sample USING NODE dlnk.almaden.ibm.com PORT 50001
 - For DCE-DFS, add the cell and dlm administrator instance name
- Database Config Parameters Relevant to DataLinks
 - DL_EXPINT, DL_TOKEN, DL_UPPER, DL_TIME_DROP, REC_HIS_RETENTN, NUM_DB_BACKUPS

Appendix D (contd.):

Data Links Manager Install and Setup

- DB2 Data Links Quick Beginnings (AIX, NT, DCE-DFS for AIX)
- Install
 - installs DLFM, DLFF, and a local DB2 with DLFM_DB
 - For DCE-DFS:
 - on DLFM Server a DB2 server is installed, on DLFM Clients, only a DB2 client is installed
 - installs DFS-CM on DCE clients
- Register filesystems which are under the control of DLFF using dlfm command
 - `dlfm add_prefix filesystem_name`
 - Prefix: mount point or sharename
 - Example: `dlfm add_prefix /test`
- Register DB2 server(s) which would use this Data Links Mgr. using dlfm command
 - `dlfm add_db database instance hostname`
 - Example: `dlfm add_db sample narang kwb.almaden.ibm.com`
- Specify Backup Mode (Local directory, ADSM, XBSA)

Appendix E: File Archive Optimization in DCE-DFS



`../../almaden.ibm.com/fs/dl/kiran.pic <==> /localmount/dl/kiran.pic`