

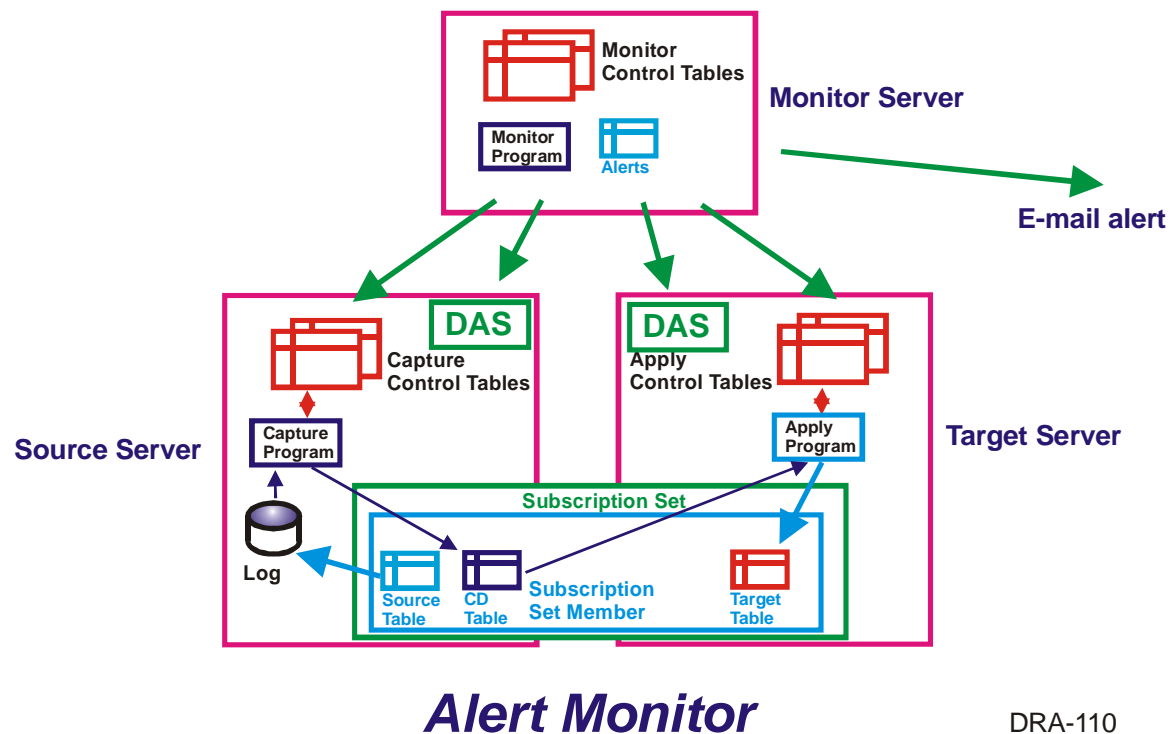
SQL Replication Tasks

SQL replication tasks	2-2
Defining source and target servers	2-3
Registering a source table	2-6
Defining a subscription set	2-8
Defining subscription set members	2-10
Capture threads	2-13
Apply threads	2-14
Log sequence number	2-15
Apply operation	2-16
Alert monitor operation	2-17
New in Version 8	2-18
Memory utilization	2-24
Storage	2-26
ASNCLP commands	2-28

SQL Replication Tasks

The SQL Replication tasks are:

1. Defining database servers that will be replication sources and targets
2. Defining replication source tables
3. Defining replication subscriptions
4. Operating Capture and Apply
5. Monitoring Capture and Apply



DRA-110

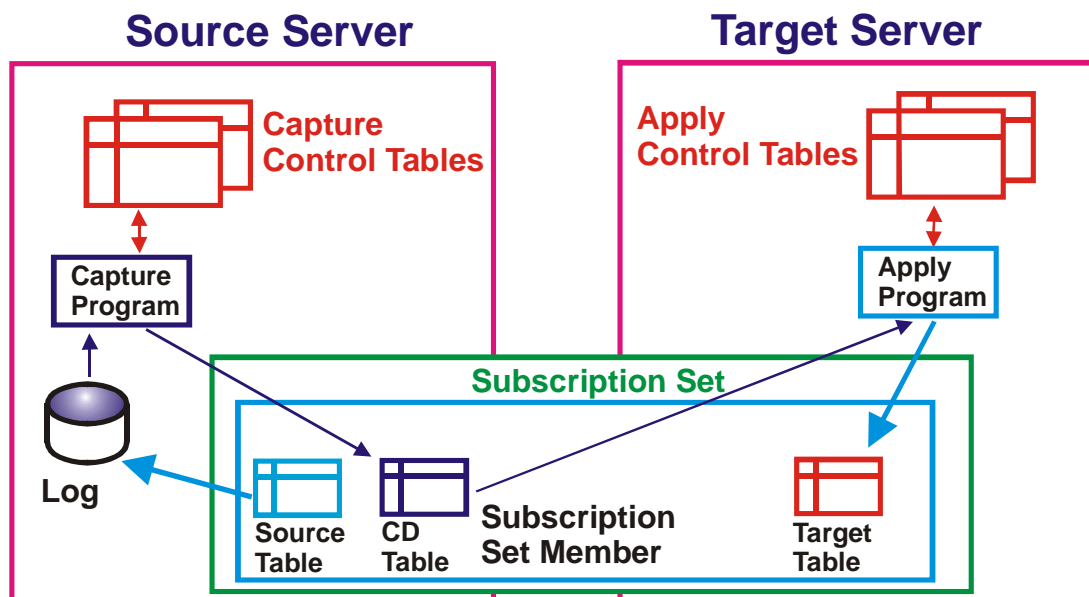
Defining source and target servers

A database server is one of the following:

- a DB2 for z/OS subsystem or data sharing group
- a DB2 for iSeries database
- a DB2 for Windows and Unix database
- an Informix Dynamic Server database

A server is known as a federated database if it contains nicknames which point to non-DB2 tables (e.g. Informix).

A database where Capture runs is called a capture control server. A source server is the same database as the Capture control server unless remote journaling is used.



Control Tables for Capture and Apply

DRA-20

You define a capture control server when you use the Replication Center to create the capture control tables in the database. Each set of capture control tables has its own schema. There can be multiple sets of capture control tables.

Each replication source and subscription and each Capture program is associated with a specific capture schema. There is one global table, ASN.IBMSNAP_CAPSCHEMAS, which has an entry for each schema you define within a database. There are twelve capture control tables for DB2 capture control servers.

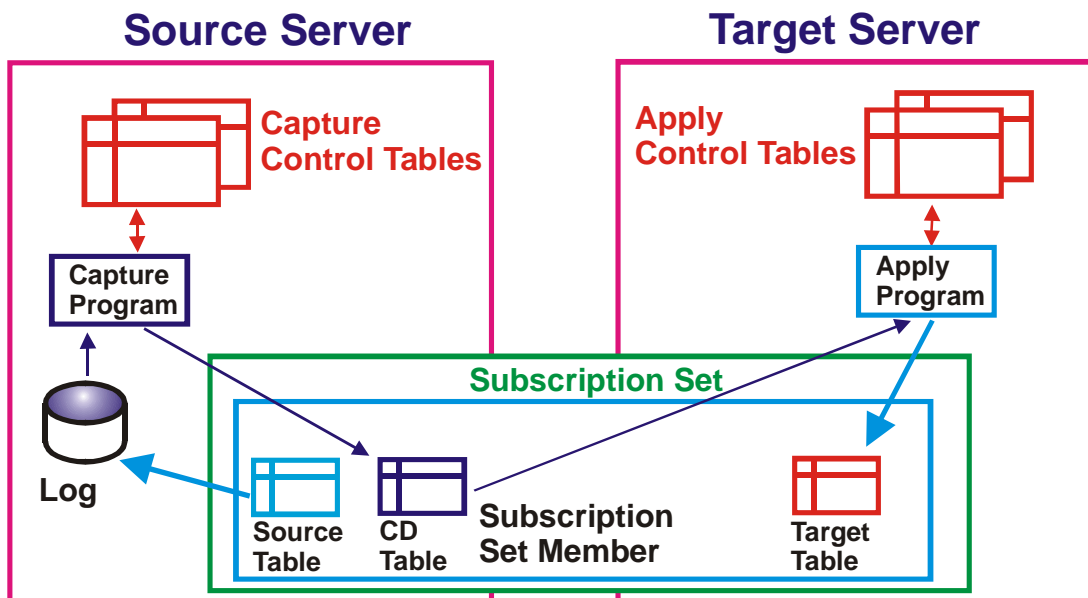
A database that holds a replication target is called a target server.

The control tables for Apply can be located on the capture control server, on the target server, or any other server in your enterprise that can be accessed by the Apply program.

The database where the apply control tables are located is called the apply control server.

Apply control tables always have the schema ASN.

There are 10 apply control tables.



Control Tables for Capture and Apply

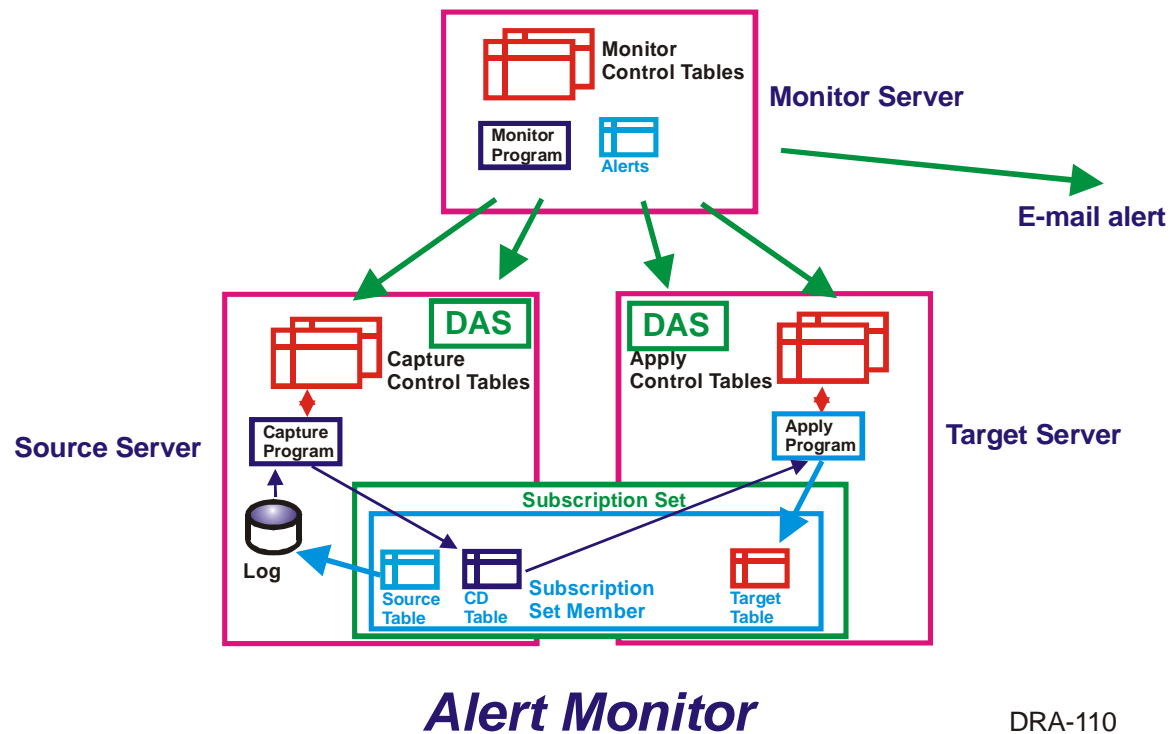
DRA-20

A database that performs monitor replication is a monitor control server.

There are nine monitor control tables. The monitor control tables are defined in the monitor server database using the replication center.

The monitor server can be located anywhere in your enterprise. It must have access to the capture and apply control servers that you want to monitor.

The monitor control tables have the schema ASN.



DRA-110

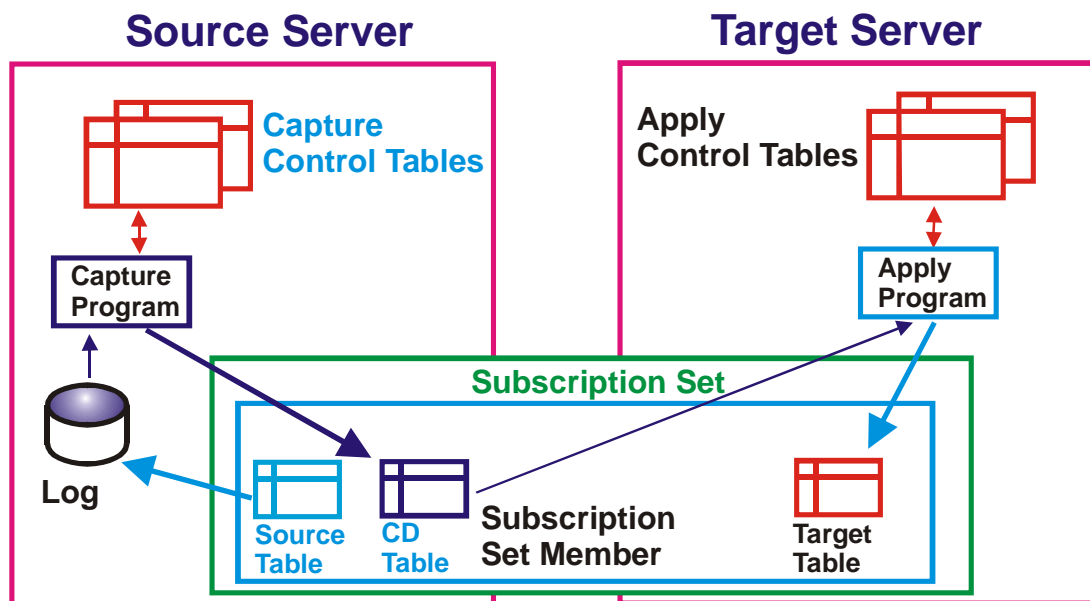
Registering a source table

Source table definition is called registering a source table.

The capture control tables must already exist in the database where the source table is located.

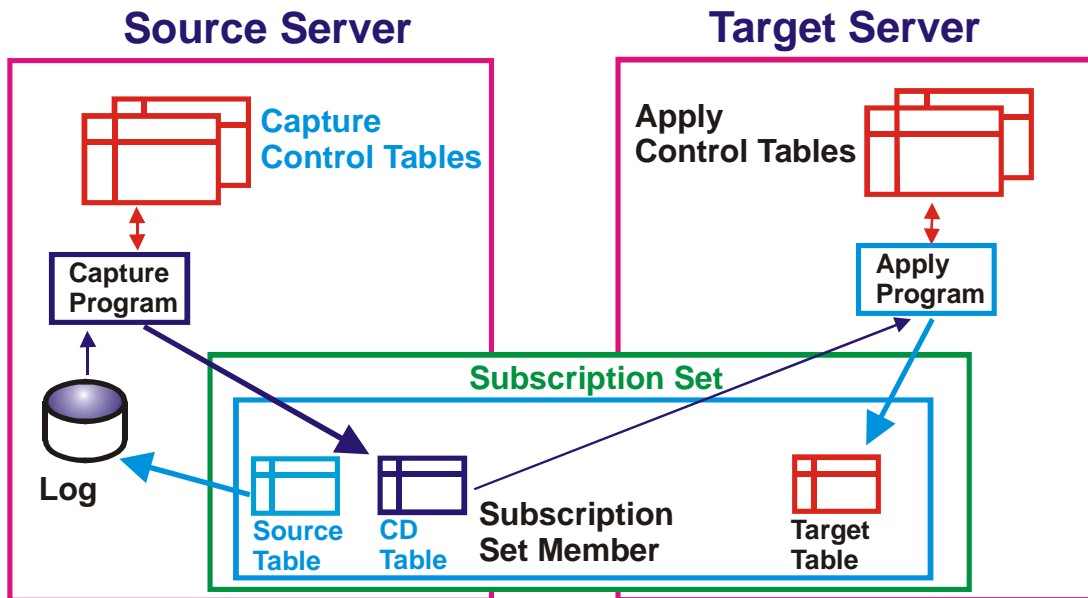
You select a source table from a list of tables in the source database and a capture schema.

You can choose to register a source table for full refresh only or for change capture.



Registering a Source Table

DRA-30



Registering a Source Table

DRA-30

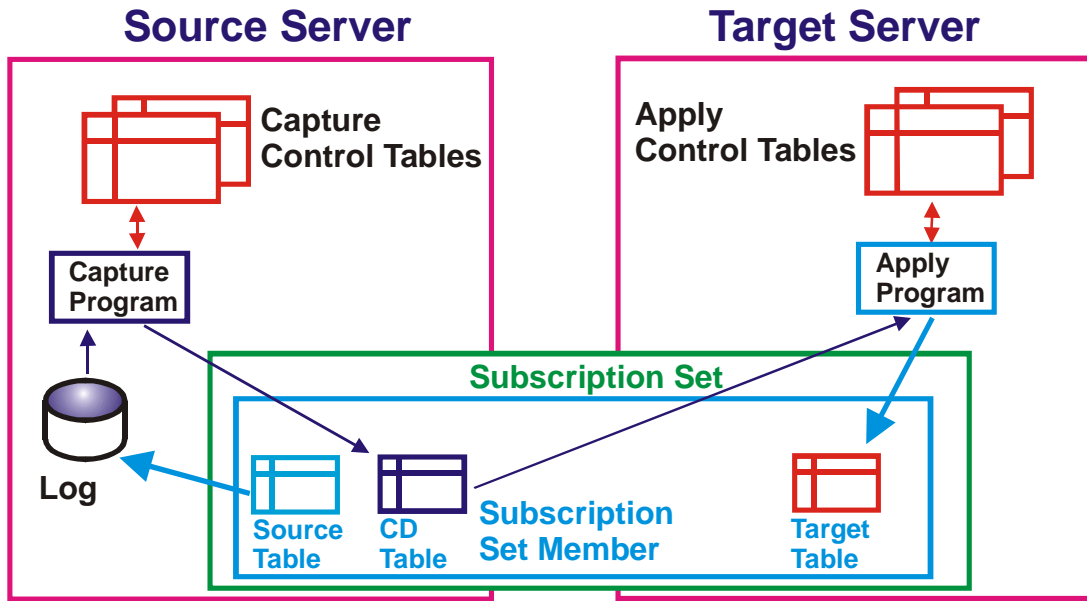
When you register a DB2 source table, the following occur:

- The source table is altered to have the `DATA CAPTURE CHANGES` attribute (if necessary). This allows full-row logging for the source table and enables the Capture program to capture the before and after images of changes to the table.
- A change data table (CD table) is created to hold the captured changes. The CD table has the columns which you selected when registering the table. The CD table columns for after-image values have the same names and data types as the source table (except for LOBs and datalinks). If you decide to capture the before-image values, the CD table columns will have the column names of the source prefixed with a character you choose. The default before-image prefix is X. There are also control columns in the CD table which identify the type of change and the log record sequence number of the change.
- A row is inserted into the control table `capschema.IBMSNAP_REGISTER`. This row contains details of the registration, such as the name of the CD table. The `STATE` column in this row is set to I.

Defining a subscription set

A subscription set contains the mappings from the source tables to the target tables.

Each source to target mapping is called a subscription member.



Creating a Subscription Set

DRA-40

A subscription set has a single source server and only one target server.

Each subscription set has an apply qualifier associated with it.

The Apply program is started with an apply qualifier and then processes all the sets with a matching apply qualifier.

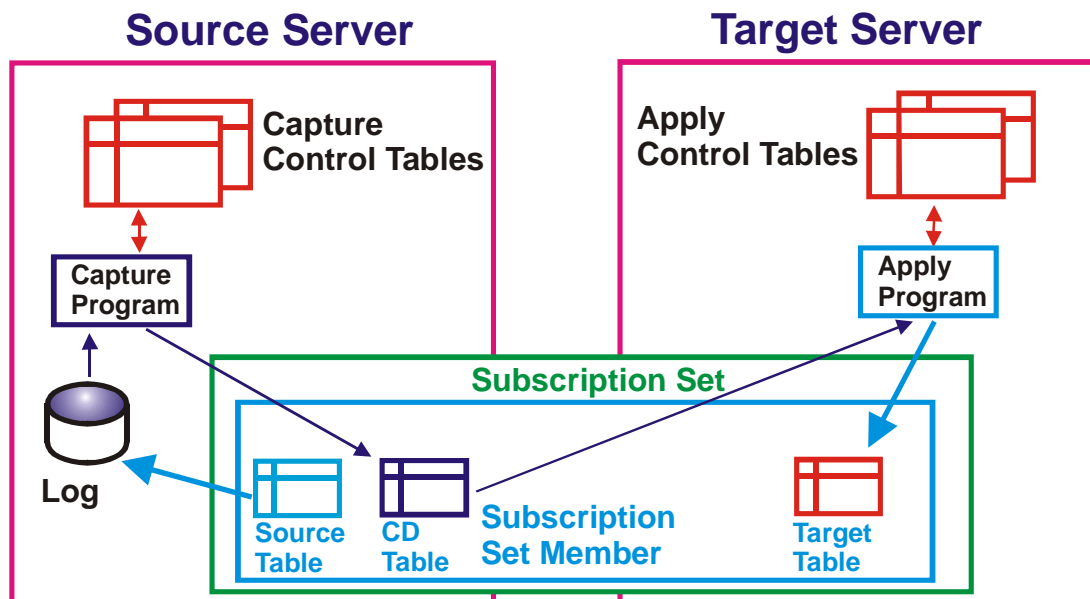
An apply qualifier can have many subscription sets. Each set has zero or more members.

Defining an empty subscription set

When you create a subscription set, you specify the following:

- the apply control server
- the apply qualifier
- the set name
- the capture control server and schema
- the target server
- scheduling information

You have the option of creating an empty subscription set and adding members later or creating both the subscription set and members at the same time.



Creating a Subscription Set

DRA-40

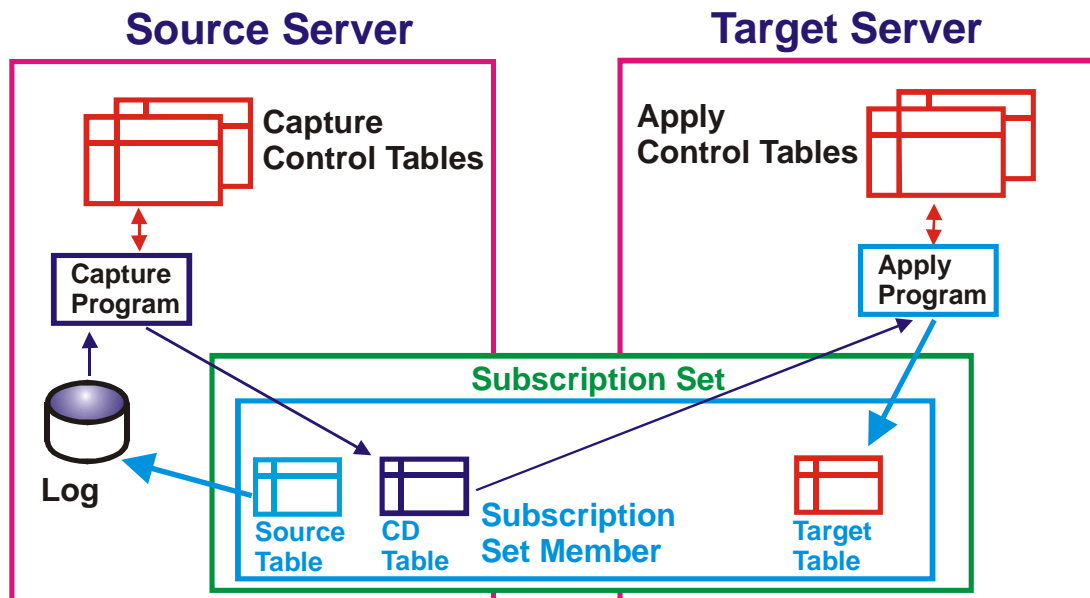
You can schedule the subscription set on a time interval and/or when an event occurs.

With event-based replication, you must specify an event name for the subscription set. The `ASN.IBMSNAP_SUBS_EVENTS` table holds rows containing details of the event's start time for Apply processing and, optionally, the end point for the changes processed.

Defining subscription set members

When you add a member to a subscription set, you must specify the following:

- the source and target tables
- the mappings from source columns to target columns
- the unique target table columns for the target table index
- row filters (predicates)



Creating a Subscription Set

DRA-40

There are 6 target table types:

User copy	a complete or partial copy of a source table, optionally with transformed columns or new columns
Point-in-time	a user copy with an IBMSNAP_LOGMARKER column that contains the timestamp when each source row was changed
Consistent change data (CCD)	a complete or partial copy of the changes to a source table
Replica	an updateable copy of all or a portion of a source table (DB2 only)
Base aggregate	an aggregate of a source table based on SQL column functions and GROUP BY filters that you define
Change aggregate	an aggregate of the changes to a source table based on SQL column functions and GROUP BY filters that you define

A CCD target table includes:

- columns for the type of change (insert, update, delete)
- the log record sequence number of the change
- the log record sequence number of the commit for that change
- the timestamp of that commit
- optionally, the authorization id that made the change

CCD Table types can be:

Complete	initialized with a CCD row for each row in the source table
Non-complete	no initialization, only a CCD row for each change
Condensed	each CCD row represents the last change to a source table row
Non-condensed	one CCD row for each change to a source table row

When Apply processes a base aggregate target table, it issues a select against the source table and inserts new rows in the target table.

When Apply processes a change aggregate table, it issues a select against the CD table and inserts new rows in the target table.

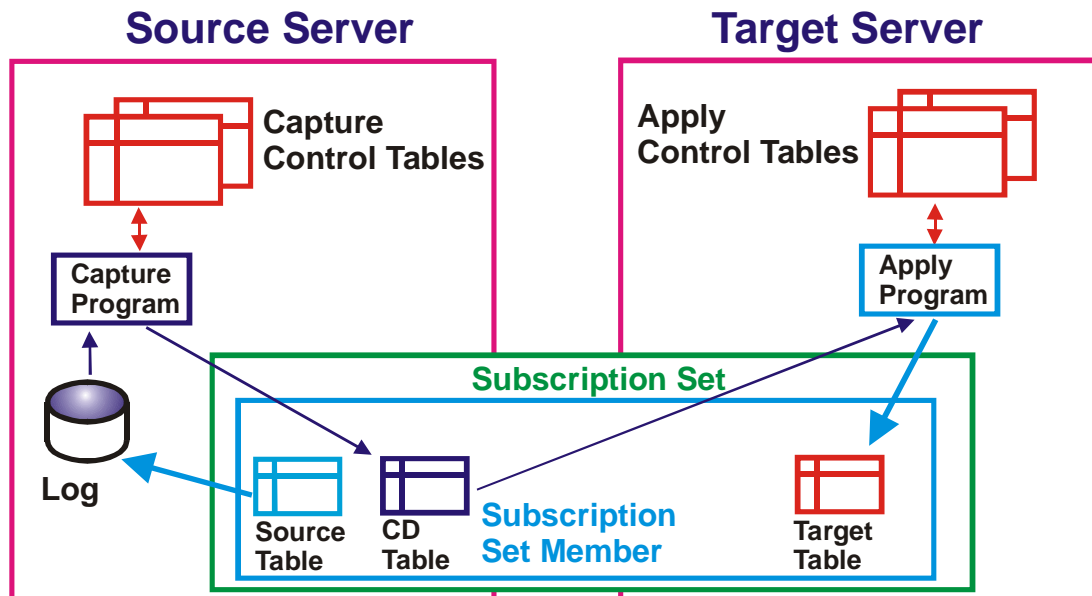
When you define a subscription member, the following occur:

- A row is inserted into the capschema.IBMSNAP_PRUNE_SET table at the capture control server (if this is the first member added to a set). The row contains set information which is used when pruning change data after it has been applied.
- A row is inserted into the capschema.IBMSNAP_PRUNCNTL table at the capture control server for each subscription member. Each member is uniquely identified by a MAP_ID number. The MAP_ID starts at 0 and increases by 1 for every subscription to this capture control server.
- A row is inserted into the ASN.IBMSNAP_SUBS_MEMBR table at the apply control server. This row contains the apply qualifier, set name, and source and target table names. This provides the source table to target table mapping. The MEMBER_STATE for each member is N for new. The PREDICATES value is the row filter that you specified for this mapping.
- Each source table column to target table column mapping is recorded in one row of the ASN.IBMSNAP_SUBS_COLS table at the apply control server. Each row defines a mapping from a source column or SQL expression to a target table column. The row also includes a flag which identifies the target table columns which are part of the target table primary key or unique index.
- If the target table does not already exist, it is created, along with a unique index.
- If the target table is a replica (Update Anywhere), then additional processing defines the replication from the replica to the source server:
 - A second set is created in ASN.IBMSNAP_SUBS_SET, and rows are inserted into both the ASN.IBMSNAP_SUBS_MEMBR and ASN.IBMSNAP_SUBS_COLS tables.
 - The replica is registered as a replication source at the apply control server.

Capture Threads

Capture has five threads:

INIT	this prepares the environment and starts the other threads
ADMIN	this manages administrative tasks, including error messages and traces, and stores monitoring statistics at the capture control server
HOLDL	this holds an exclusive lock on capschema.IBMSNAP_CAPENQ. It prevents a second instance of Capture with the same capture schema and capture control server.
WORKER	this requests log records from DB2, manages log information and inserts rows in CD tables and the UOW table. This thread also updates the capture control tables.
PRUNE	this deletes (prunes) rows from the CD tables and UOW tables after they have been applied. This thread also prunes: <ul style="list-style-type: none"> - capschema.IBMSNAP_TRACE - capschema.IBMSNAP_CAPMON - capschema.IBMSNAP_SIGNAL



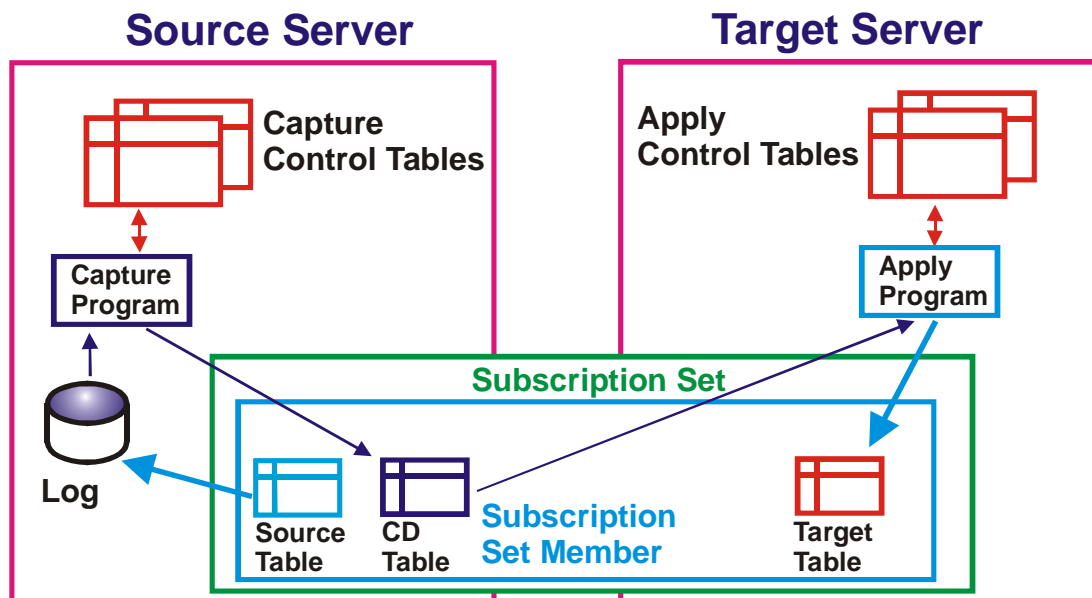
SQL Replication

DRA-10

Apply Threads

There are three Apply threads:

ADMIN	this manages administrative tasks
HOLDL	this holds an exclusive lock on the ASN.IBMSNAP_APPENQ row with a matching apply qualifier. It prevents a second instance of Apply with the same apply control server and apply qualifier.
WORKER	this processes the subscription sets based on the apply control server and apply qualifier specified when Apply is started. This thread accesses: <ul style="list-style-type: none"> – capture control tables – source tables – CD tables – UOW tables – apply control tables – target tables



SQL Replication

DRA-10

Use the Query Status option on the Replication Center to view the status of each of the Capture and Apply threads.

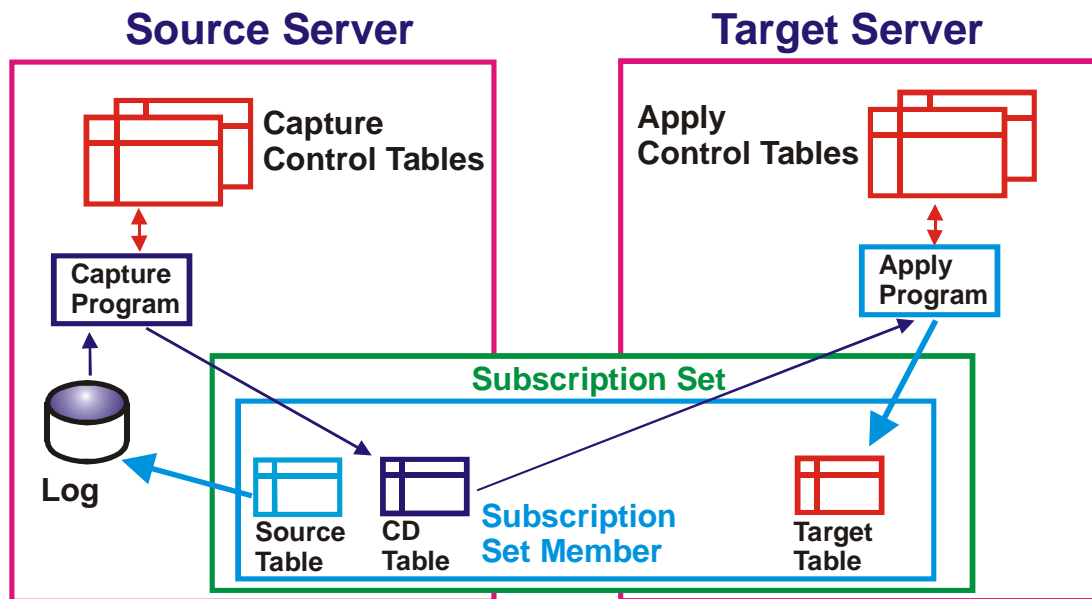
LSN: Log Sequence Number

DB2 SQL replication makes use of the log record sequence number, the LSN.

The LSN for each change is stored in the CD table row for that change, along with the LSN of the commit statement for that change's unit of work. These values are used to ensure that changes are replicated in the order they occurred at the source.

Capture and apply control tables have columns named SYNCHPOINT, which are used as progress indicators to control replication and restart processing.

Recovery logging must be enabled for any DB2 for Windows and UNIX capture control servers before Capture is started the first time.



SQL Replication

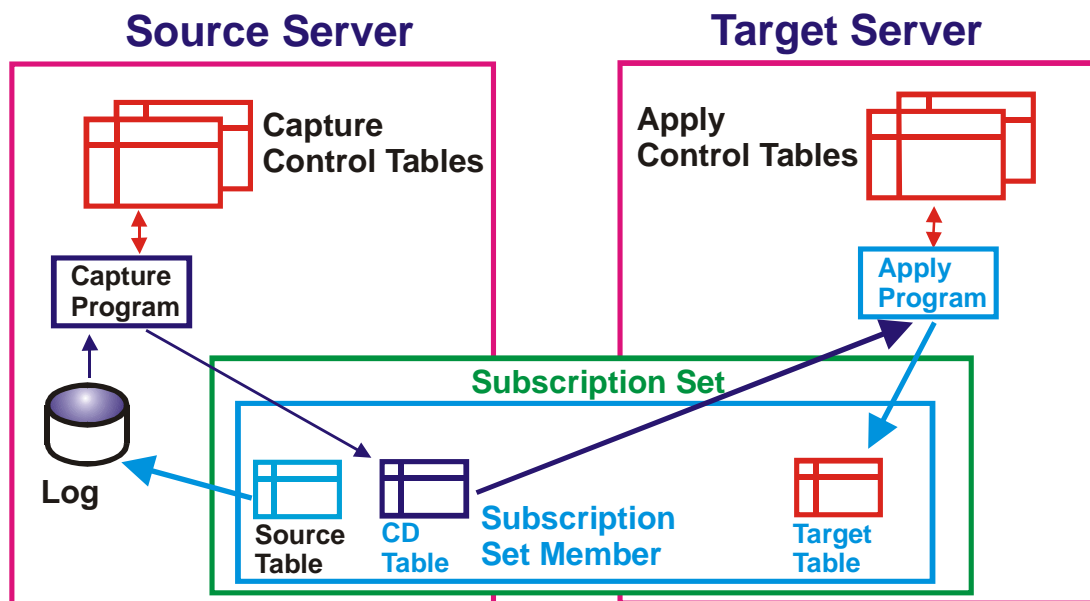
DRA-10

Apply operation

The Apply program selects rows from the source table or view. This process includes any transformations you have defined with SQL expressions and any row filters you specified for each member.

Apply stores the result set in a spill file at the server where Apply was started. There is one spill file for each member in the subscription set.

Apply issues a DELETE against the target table to delete existing rows and then INSERTs to the target table using the data from the spill file.



Applying data to a Target

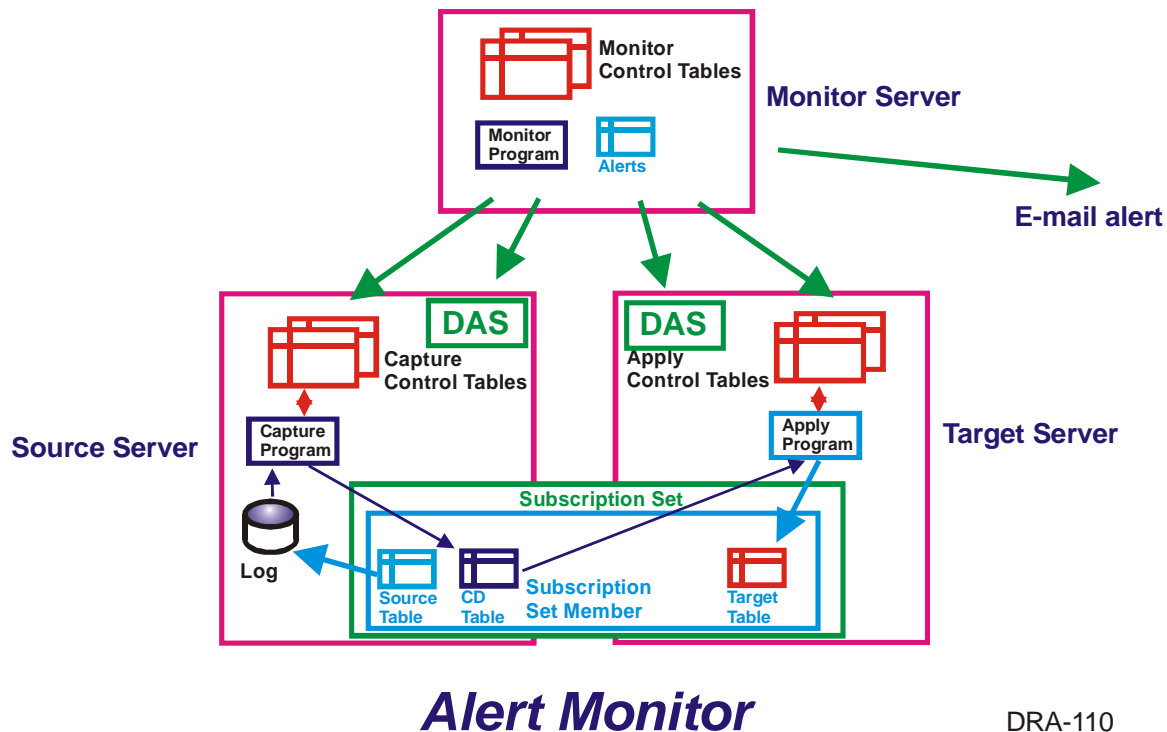
DRA-60

You can run user exit program with Apply. It is called ASNLOAD and is controlled by the LOADXIT Apply parameter. ASNLOAD uses DB2 export and import (or load) commands to do the initial full refresh, leading to better performance.

Alert Monitor Operation

The Alert Monitor can check for:

- Status of Capture and Apply programs
- Error and warning messages
- Latency thresholds
- Memory usage
- Subscription set failures or full refreshes
- Transactions rejected due to update-anywhere conflicts
- Transactions reworked by Apply



DRA-110

The `monitor_interval` is the number of seconds in a monitor cycle. It is specified when the Alert Monitor is started.

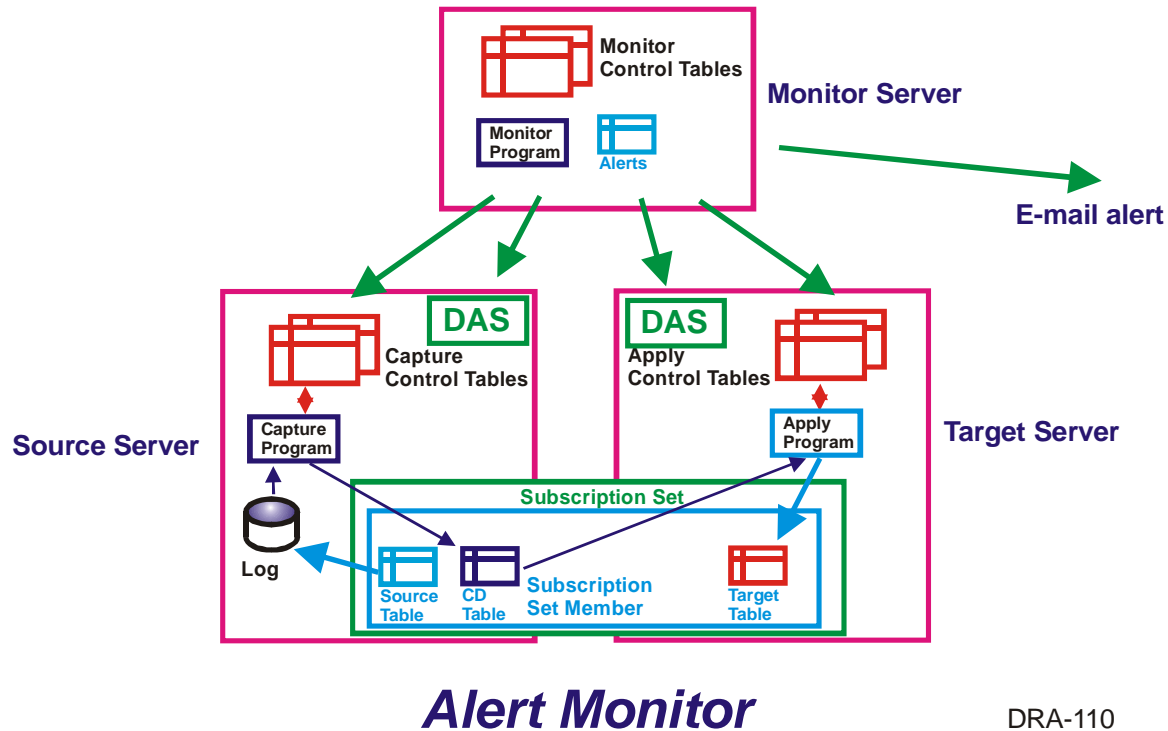
The Alert Monitor checks for any alert conditions by selecting values from capture and apply control tables and issuing system commands to the DAS running on the capture or apply server.

If the Alert Monitor detects an alert condition, it sends an e-mail describing the condition to the contact or contact group defined in the monitor control tables. The alert is also inserted into the `ASN.IBMSNAP_ALERTS` table at the monitor control server.

The `MAX_NOTIFICATIONS_PER_ALERT` startup parameter can be used to prevent flooding the contacts with alerts for the same problem.

The `MAX_NOTIFICATIONS_MINUTES` startup parameter is used to control the number of minutes between notifications for the same alert.

The `ASN.IBMSNAP_ALERTS` table is pruned by the Alert Monitor at startup, and is controlled by the `ALERT_PRUNE_LIMIT` startup parameter.



New in Version 8: Administration

- The Replication Center is a graphical user interface which supplies the administrative functions found in the previous tools (DataJoiner Replication Administration (DJRA) and replication functions in the Control Center).
- Source and target table names can be up to 128 characters.
- Column names can be up to 30 characters.
- Replication definitions can be added without re-initializing the Capture program.
- New columns can be added to replication sources while Capture is running (except for iSeries).
- There is a migration utility which is used to convert existing DB2 Replication V5, V6, and V7 environments to DB2 Replication V8.
- Stored passwords are encrypted.
- The Replication Center is supported on 64-bit Windows and Unix systems with a 64bit Java Virtual Machine.

New in Version 8: Capture

- The `IBMSNAP_SIGNAL` table is used for communications between Capture and Apply. Once Capture has been started successfully, there is no longer a need to always start Capture before starting Apply. Also, Capture need not be running before Apply processes a new subscription.
- Changes are not inserted in change data and UOW tables until they have been committed. Changes which are rolled back are therefore no longer put in the CD table. User copies do not require a join of the CD and UOW tables.
- If you choose a subset of source columns when defining a replication source, you can specify that no changes should be captured for that source unless the change affects your selected columns.
- When defining a replica, you can specify that the changes processed by Apply from the master site should not be recaptured at the replica site.
- Multiple Capture programs can run on the same DB2 database, subsystem, or data sharing group.
- Capture prunes changes concurrently with the capture of changes on DB2 for Windows, UNIX, and z/OS. This means that pruning no longer affects replication latency. Pruning does not require a join of each Change Data table with the Unit of Work table. Pruning is cursor-based with interim commits.
- The `IBMSNAP_SIGNAL` table provides a way to communicate with Capture through log records. Apply inserts records into this table to signal that capturing should start for a table. It is also used to signal update anywhere replication and to provide precise end points for Apply events.
- Capture start-up parameters can be modified while Capture is running.
- Capture is supported on 64-bit platforms — Windows, UNIX, z/OS.

New in Version 8: Apply

- When applying changes to user copies, joins of the Change Data table and the Unit of Work table are no longer needed. This is because all the required information is contained in the Change Data table.
- Changes to target table primary key values are handled without converting all captured updates to delete/insert pairs. You must capture the before-image of the source table columns used for the target table primary key. Apply will use the before-image values to locate the target table row for an update.
- Faster full refreshes of target tables can be done using the Apply exit ASNLOAD.
- Apply password files are encrypted. Use the asnpwd command to create and maintain the password file.
- Apply is supported on 64-bit platforms — Windows, UNIX, z/OS.

New in Version 8: Monitor

Replication Center Monitoring

You can check the status of replication processes and display historical information including:

- Capture messages
- Capture throughput analysis
- Capture latency
- Apply messages
- Apply throughput analysis
- End-to-end latency

Status details

The following commands show the status of Capture and Apply processes:

- Capture asncmd with the status parameter
- Apply asnacmd with the status parameter

On the iSeries, users can use the WRKSBSJOB QZSNDPR command.

Replication Alert Monitor

The Replication Alert Monitor can monitor one or many Capture/Apply processes.

New in Version 8: Troubleshooting

There are several serviceability improvements including:

- Trace facilities for Capture and Apply are based on the db2trc model. The asntrc command starts and stops a trace while the Capture and Apply programs are running. This trace is available on Windows, UNIX, and z/OS.
- Additional trace points have been added to DataPropagator for iSeries to provide more debugging information.
- The Replication Analyzer Program now works with DB2 replication V8 control tables.

Memory Utilization: Capture

When the Capture program reads the DB2 log, it stores individual transaction records in memory until it reads the associated commit or abort record. Data associated with a commit record is written to the CD table and the UOW table.

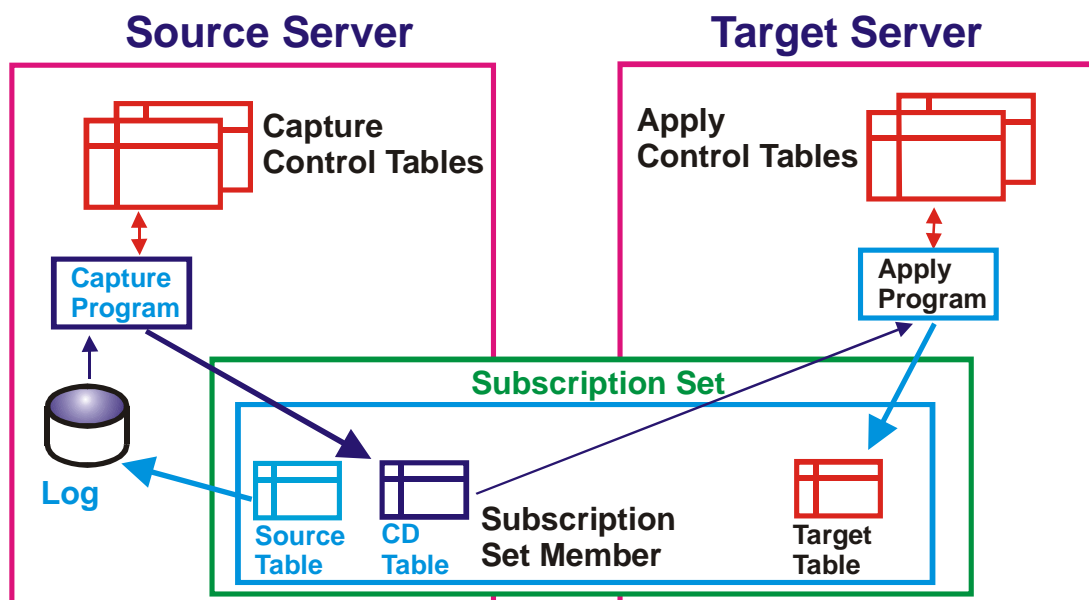
The committed transactions stay in memory until the Capture program commits its work when it reaches its commit interval.

You can monitor how much memory the Capture program is using by looking in the CURRENT_MEMORY column of the Capture monitor (IBMSNAP_CAPMON) table.

The memory_limit parameter ensures that Capture uses a specified amount of memory for storage that is associated with transactions. This parameter can be changed at run-time.

If Capture reaches the memory limit, it writes some transactions to a spill file.

The log records are placed in a memory buffer. The default size of the buffer on Linux, UNIX and Windows is 50 x 4 KB pages. The default size for z/OS is 66 x 1 KB pages.



Capturing data from a Source Table

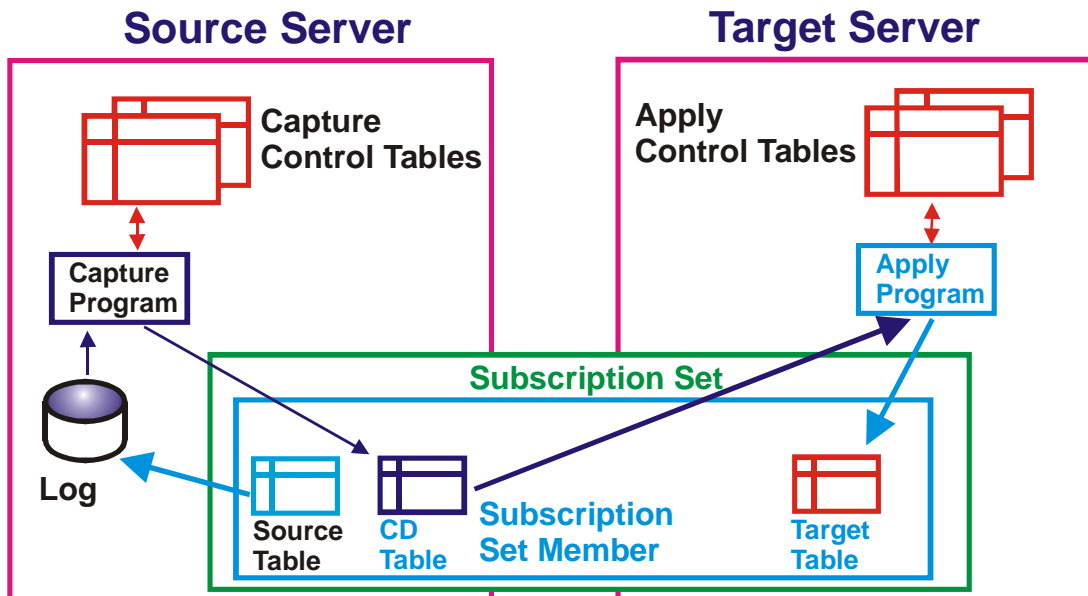
DRA-50

Memory Utilization: Apply

When the Apply program fetches data, it normally uses a small amount of memory for fetching individual rows.

The amount of memory used depends on the size of the table columns and the number of rows fetched at one time. This means that if the Apply program is fetching a LOB column, it could potentially use 2 GB of memory.

The Apply program also uses memory to hold details of the active subscription sets. The amount of memory used at one time by the Apply program is generally proportional to the amount of memory required to process the subscription set that has the most members.



Applying data to a Target

DRA-60

Storage

Disk storage will be needed for the following:

Database log and journal data	Extra data is logged to support the replication.
Target tables and control tables	The replicated data and control tables (including CD tables).
Temporary files	The data stored by replication programs in spill files and diagnostic log files (for example, *CAP.log and *APP.log).

The location of the temporary files depends on the values of the capture_path, apply_path, and monitor_path start-up parameters.

There is one spill file per transaction.

The size of the Capture spill files depends on the following factors:

Memory limit	The bigger the memory limit, the less likely the Capture program will spill to files.
Size of transactions	Larger transactions normally increase the need to spill to file.
Number of concurrent transactions	If there are multiple Capture transactions at the same time, the Capture program needs to store more information in memory or on disk.
Commit interval	Lower commit intervals reduce the need for storage because Capture has to store information in memory for a shorter period of time before committing it.

Replicating data between databases with compatible code pages

If you need SQL statements and data to go between systems with differing code pages, the underlying DB2 protocols such as DRDA will handle any code page translation.

If data is passed between DB2 and non-DB2 relational databases, SQL replication expects the underlying database products to handle any code page translation.

If you plan to replicate data between databases with differing code pages, check that the code pages you have are compatible. If your databases do have compatible code pages, check if the databases use code pages differently.

For example, if one database product allows a different code page for each column in a table while another database product does not allow different code pages per column, you will have to specify the code page only at the database level.

A table with multiple code pages in the first product cannot be replicated to a single database in the second product.

ASNCLP commands for SQL Replication

ASNCLP is a command tool program which enables you to create scripts that control replication. Its commands are:

Command	Command action
CREATE CONTROL TABLES	Create control tables
DROP CONTROL TABLES	Drop control tables
CREATE REGISTRATION	Create registration
ALTER REGISTRATION	Alter registration
DROP REGISTRATION	Drop registration
PROMOTE REGISTRATION	Promote registration
CREATE SUBSCRIPTION SET	Create subscription set
ALTER SUBSCRIPTION SET	Alter subscription set
DROP SUBSCRIPTION SET	Drop subscription set
PROMOTE SUBSCRIPTION SET	Promote subscription set
CREATE MEMBER	Create member
ALTER MEMBER ADD COLS	Add columns to an existing member
DROP MEMBER	Drop member
CREATE STMT	Create a statement for an existing subscription set
DROP STMT	Drop statements for an existing subscription set
OFFLINE LOAD	Control a manual full refresh for offline load procedures
SET SESSION	Establish a session for your chosen type of replication
SET SERVER	Assign a database alias for a logical replication server or specify a user ID and password to use when connecting to the database
SET PROFILE	Set up customization rules for creating table space objects
SET DROP	Determine whether to drop the table space when you drop the database object that it contains
SET OUTPUT	Define output files for the replication command line interface
SET LOG	Define the log file for the replication command line interface
SET CAPTURE SCHEMA	Set a default source and target Capture schema for all task commands
SET TRACE	Enable and disable the trace for the ASNCLP commands
SET RUN SCRIPT	Specify whether to automatically run each task command from an input file before the ASNCLP commands process the next task command