Kelly Smith - Business Executive

22nd November 2011

Fast, Safe and Effective Methods for Solving Data Copy Challenges

Agenda

- **Customer challenges, solutions, and scenarios when copying large amounts of data**
  - Data warehouses, SAP, PeopleSoft, read-only
- **Database and storage integration**
  - Operational advantages
- **Cloning terminology**
- **Options for copying data**
- **How to clone DB2 sub-systems quickly with large amounts of data**
  - New subsystem features
- **How to refresh DB2 table and index spaces quickly**
  - New table space features

Customer challenges, solutions, and scenarios when copying large amounts of data

- Customer 1: DB2 Cloning Tool is able to perform the complete SAP cloning function in approximately 15 minutes, a saving of over 11 hours and 45 minutes. Because the time to clone a DB2 subsystem was reduced so drastically they are cloning weekly instead of monthly.
- Customer 2: We create HSC copies primarily for development and for some unit testing. When a lot of development was going on, we had up to 93 SAP instances. Creating SAP homogenous system copies is a requested process and can be anywhere from once or twice a month to three to eight times a week during heavy development times. We went from two to three days to clone an SAP instance to only minutes.
- Customer 3: The end users had access to the warehouse data 1 day a week. Now, the users have unlimited access to the data.
- Customer 4: It took us 48 hours to clone our PeopleSoft subsystem. Now it takes 30 minutes.

Database and Storage Integration - Storage Aware Tools

- **Large database systems require high availability**
  - Fast and non-intrusive backup and cloning facilities are required
  - Fast recovery capabilities are required to minimize downtime and promote high availability
  - Most backup, recovery and cloning solutions do not leverage storage-based fast-replication facilities
- **Storage-based fast-replication facilities are under-utilized**
  - Tend to be used by storage organizations
  - Tend not to be used by database administrators (DBAs)
- **Storage aware database products allow DBAs to use fast-replication in a safe and transparent manner**
  - Provides fast and non-intrusive backup and cloning operations
  - Simplifies recovery operations and reduces recovery time
  - Simplifies disaster recovery procedures

Database and Storage Integration - Operational Advantages

- **Reduces time and cloning administration costs**
- **Simplifies cloning strategies by using automation to coordinate database system operations with fast-replication facilities**
- **Reduces host CPU and I/O resource utilization**
- **Creates cloned copies without affecting running applications**
- **Leverages storage processors and fast-replication investments**
  - IBM, EMC, HDS, STK
- **Exposes fast-replication capabilities to the DBAs *safely and transparently* using "*storage-aware*" database utilities**
- **Provides a sophisticated infrastructure and metadata to manage the DBMS and storage processor coordination**

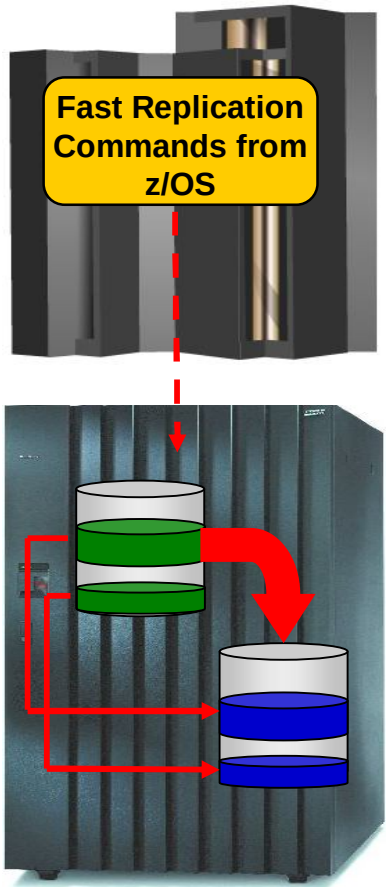
Cloning Terminology

- **A clone is an exact replica**
  - Clone a DB2 system by volume
  - Clone a table space by data set
- **DB2 system cloning and table space refresh**
  - The act of replicating the data, making the replica accessible, and then using the replica in lieu of the original data
- **DB2 system cloning automation**
  - Clones a complete DB2 system including all its databases **by VOLUME**
- **DB2 table and index space refresh automation**
  - Refreshes specific table and index spaces **by data set**
  - Lowest level is by data set

Fast Replication Data Copy Options

Fast copy processes offloaded to the storage processor
No host CPU or I/O resources

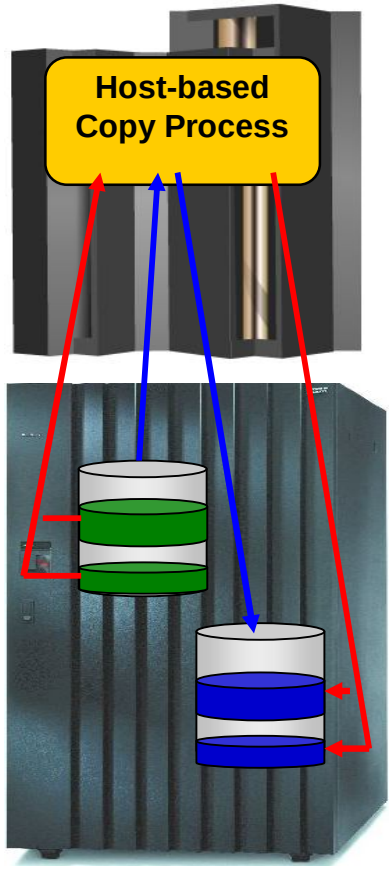
- **Volume based fast replication options for DB2 system cloning**
  - FlashCopy (IBM,EMC,HDS)
  - SnapShot (IBM,STK)
  - TimeFinder/Clone Volume Snap (EMC)
  - TimeFinder/Snap (EMC)
  - Mirror processes
    - PPRC (IBM,EMC,HDS)
    - TimeFinder/Mirror, SRDF (EMC)
    - ShadowImage HUR (HDS)
    - TDMF (software based)
- **Data set based fast replication options for DB2 table space refresh**
  - Data Set FlashCopy (IBM,EMC,HDS)
  - Data set SnapShot (IBM,STK)
  - TimeFinder/Clone Data set Snap (EMC)



Host Based Data Copy Options

Data copy processes use host based CPU and I/O facilities
Slower than storage-based fast replication

- **Volume copy options for DB2 system cloning**
  - TDMF (IBM)
  - FDRPAS (Innovation Data Processing)
  - DFSMSdss (IBM)
  - FDR (Innovation Data Processing)
- **Data set copy options for DB2 table space refresh**
  - Any traditional data set copy processes



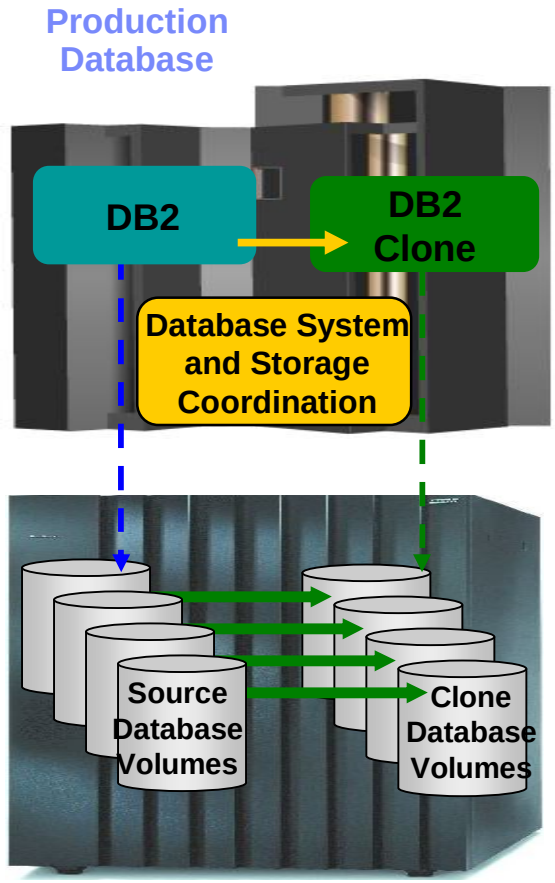
Fast, safe and effective methods ...

DB2 Subsystem Cloning

Copy complete DB2 subsystems

Clone DB2 Systems Using DB2 Cloning Tool
DB2 System Cloning Automation

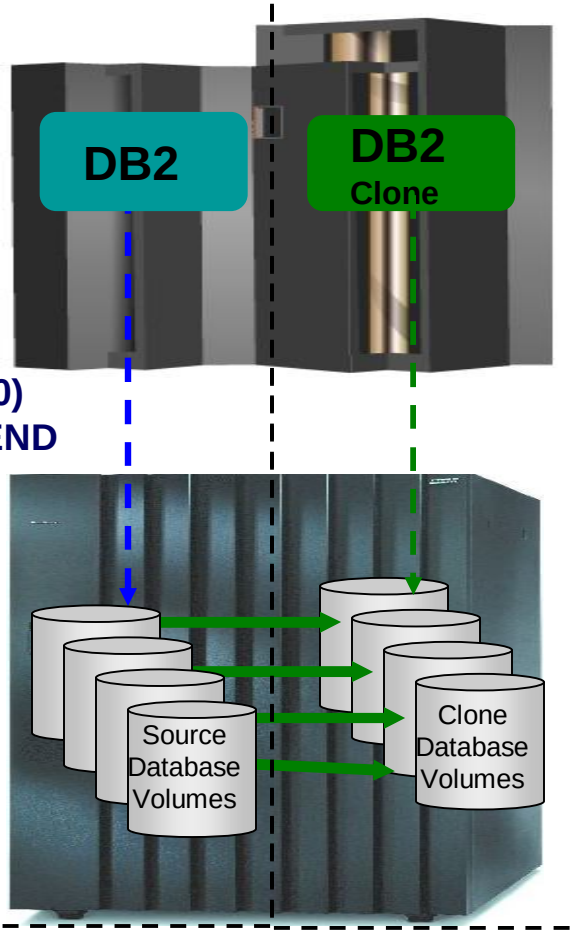
- **Performs automated cloning of DB2 systems**
- **DB2 data copied using storage-based dataset fast-replication**
  - Data can be cloned while online or offline
- **Performs rapid volume reconditioning and data set renaming on cloned volumes to solve the data access challenges**
  - Target volumes retain their target volume label
  - Renames the VTOC, VTOCIX, and VVDS to match the target volume
  - Renames and catalogs all data sets to a new HLQ
- **Adjusts target DB2 system to accommodate and accept the cloned data**
  - DB2 catalog, directory, BSDS, active / archive Log
  - Makes data accessible on the same or shared LPAR



DB2 System Cloning Steps

Production DB2 'Source'

1. DB2 volume selection
2. A. SET LOG LOAD(0) SET LOG SUSPEND
   B. Consistency Group
3. Volume copy
4. SET LOG RESUME if 2A



Target DB2

5. Rename
6. Update DB2 directory and BSDSs
7. Start DB2 in maintenance mode for metadata management
8. Correct DB2 catalog and directory page spaces
9. Update DB2 catalog
10. Correct application page spaces
11. Stop target in maintenance mode
12. Start DB2 clone in normal mode

NEW in v3.1 – Subsystem Cloning Enhancements

- **A DB2 stored procedure that works with the DB2 Administrative Task Scheduler to clone an entire DB2 subsystem**
- **FlashCopy Preserve Mirror support to remove the necessity of copying cloned data over the Metro Mirror/Peer to Peer Remote Copy (PPRC) link**
- **ISPF interface can be used to create a clone from a DB2 System Level Backup**
- **The ability to clone a DB2 Data Sharing Group without having to drop and re-create the work database for each DSHARE member**

NEW in 3.1 – A DB2 stored procedure that works with the DB2 Administrative Task Scheduler to clone an entire DB2 subsystem

- **Stored procedure is an application interface (API) to DB2 Cloning Tool**
- **The stored procedure will:**
  - Generate the necessary jobs to do the subsystem cloning
  - Schedule the jobs in the DB2 Administrative Task Scheduler and cause the jobs to be submitted
  - Monitor the execution of the jobs
- **The stored procedure will return to the caller, when the requested cloning has ended, either in success or failure**
- **It can be called from a platform other than z/OS to clone an entire z/OS DB2 subsystem**

NEW in 3.1 – A DB2 stored procedure that works with the DB2 Administrative Task Scheduler to clone an entire DB2 subsystem

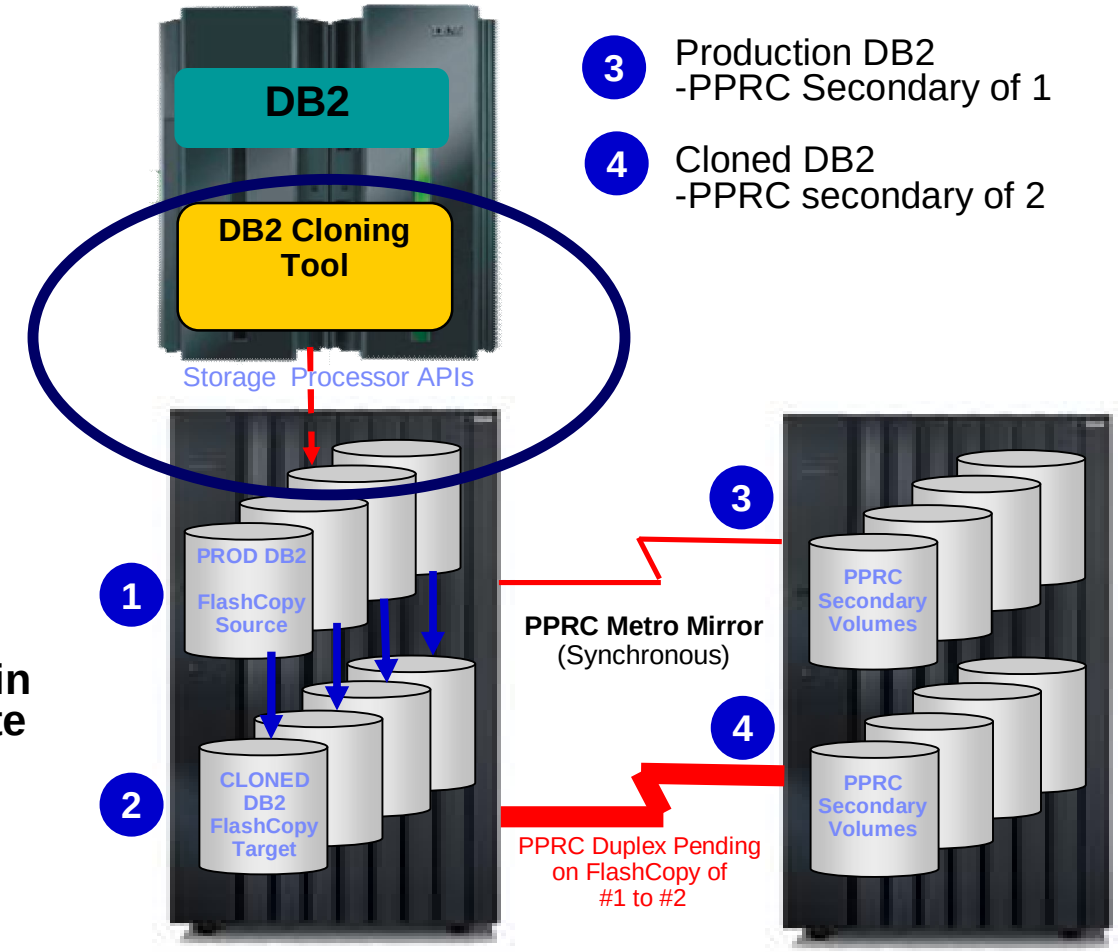
- **Input to the stored procedure:**
- **Product Parameter File**
  - DSName of the product load library
  - DSName of the CKZINI file
- **DB2 Systems Parameter File**
  - SSID of each DB2 subsystem
  - DSName of the SDSNLOAD library
  - SYSAFFINITY for the LPAR where the jobs will execute
  - VCATNAME for the DB2 catalog
  - DSNames of the BSDS data sets
- **Cloning Parameter File**
  - JOBCARD, USERID and PASSWORD for the executing jobs
  - DSName of the output JCL library
  - HLQ of the work data sets
  - Source and Target volumes, ICF catalogs and DB2 subsystems
  - Rename masks

NEW in 3.1 – A DB2 stored procedure that works with the DB2 Administrative Task Scheduler to clone an entire DB2 subsystem

- **Output of the stored procedure:**
  - Output JCL PDS
    - The jobs created by the stored procedure are written here
  - Output Status file
    - Used by the stored procedure to record and monitor the jobs submitted by the DB2 Administrative Task Scheduler

IBM FlashCopy of DB2 data where volumes are being mirrored
******* Clone ***Without*** Preserve Mirror FlashCopy *******

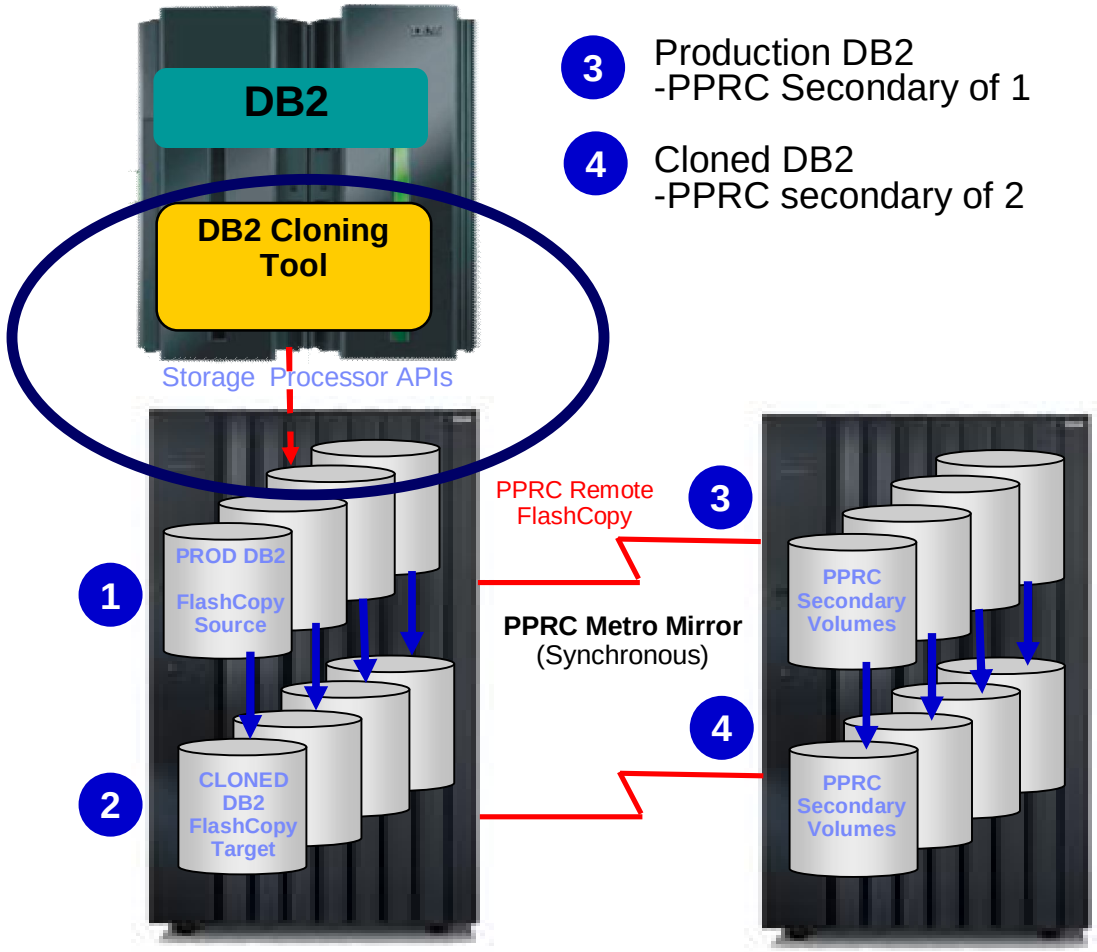
1. Production DB2
   -FlashCopy source volumes
   -PPRC primary
2. Cloned DB2
   -FlashCopy target volumes
   -PPRC primary
3. Production DB2
   -PPRC Secondary of 1
4. Cloned DB2
   -PPRC secondary of 2



- **FlashCopy from #1 to #2 <u>WILL</u> cause data to be copied through PPRC link**
- **It will cause PPRC duplex pending state because the PPRC primary #2 has been changed. After the data is in synch, it is in a duplex state**
- **Results in a significant amount of traffic**

NEW in 3.1 - IBM FlashCopy of DB2 data where volumes are being mirrored ******* Clone ***With*** Preserve Mirror *******

1. Production DB2
   -FlashCopy source volumes
   -PPRC primary
2. Cloned DB2
   -FlashCopy target volumes
   -PPRC primary
3. Production DB2
   -PPRC Secondary of 1
4. Cloned DB2
   -PPRC secondary of 2



- **FlashCopy from 1 to 2 will NOT be copied over PPRC links**
- **DB2 Cloning Tool drives Preserve Mirror In-band FlashCopy operation**
  - –Remote PPRC production secondary volumes #3, are Flashed to remote PPRC secondary volumes #4

NEW in 3.1 – ISPF interface can be used to create a clone from a DB2 System Level Backup

- **ISPF interface can generate the jobs necessary to clone a DB2 subsystem from a:**
  - DB2 BACKUP SYSTEM utility
  - A clone created by DB2 Recovery Expert is currently only supported in batch

NEW in 3.1 – Clone a DB2 Data Sharing Group without having to drop and re-create the work database for each DSHARE member

- **Simplifies the process of creating a data sharing DB2**
- **Updates of the GROUP_MEMBER column in the cloned DB2 catalog table SYSIBM.SYSDATABASE with the new member name**

DB2 Support

- **DB2 Support**
  - DB2 offline
  - DB2 online
    - Clone from an executing DB2 subsystem
    - Clone from a System Level Backup
  - DB2 data sharing
  - DB2 data sharing with many to less members
  - DB2 data sharing to non-DB2 data sharing

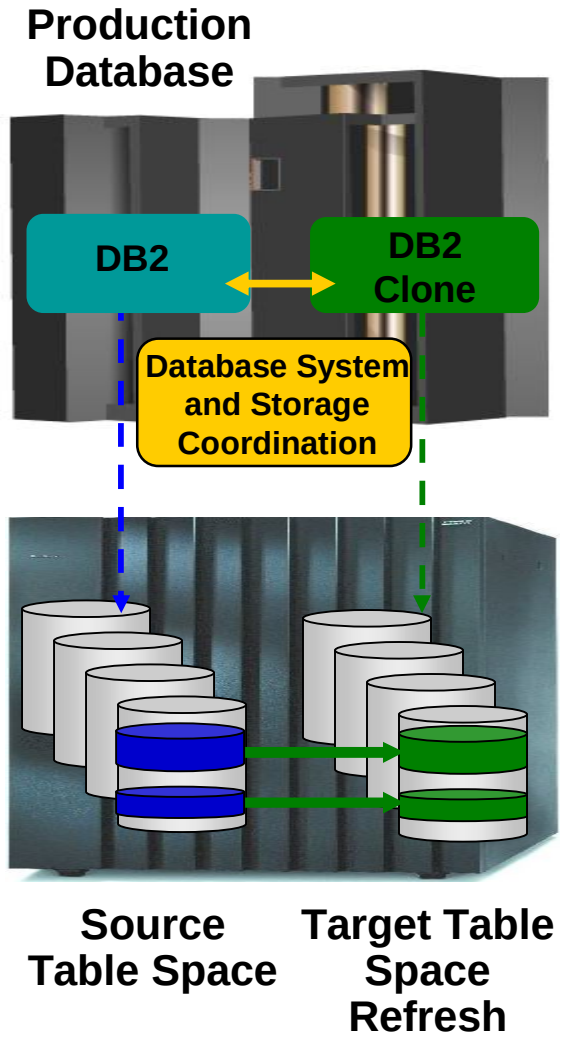
Fast, safe and effective methods ...

Table Space and Index Space Cloning

Propagate objects to test environments to help speed application deployment in production

Refresh DB2 Table and Index Spaces Using DB2 Cloning Tool
Table and Index Space Refresh Automation

- **Performs automated DB2 table and index space refresh operations**
  - DB2 RI relationships, LOBS, and Identity columns
  - XML on DB2 V9 or greater
- **Verifies source target database compatibility**
- **DB2 data copied using storage-based dataset fast-replication**
  - Data can be cloned while online or offline
  - Slow copy mechanism can be used
- **Performs object ID translations, data masking, log apply, and target system meta-data**

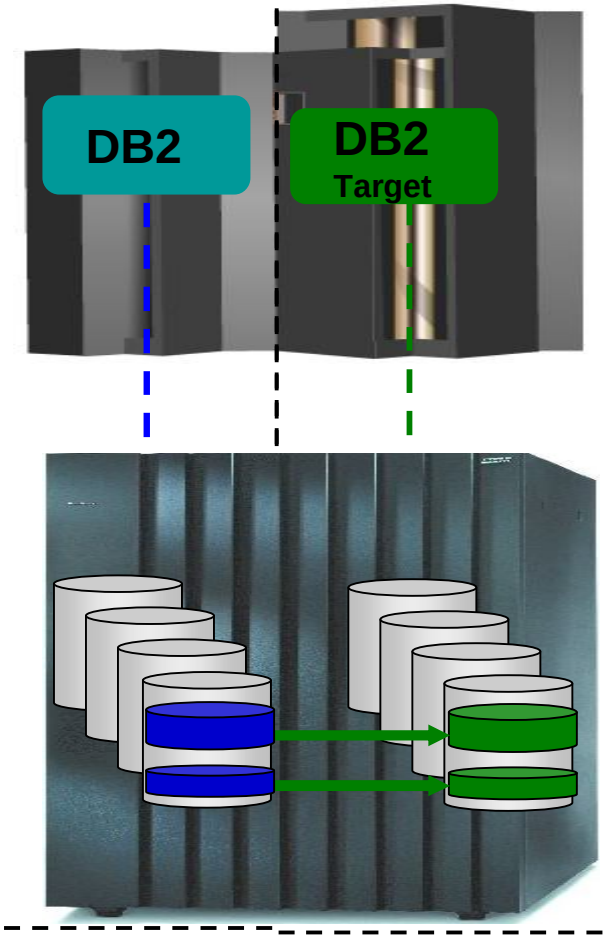


DB2 Table and Index Space Refresh Steps

Production DB2 'Source'

Source Job

1. LISTDEF selection
2. Verify object compatibility
3. Create target DB2 DDL and table and index spaces if they don't exist
4. Stop table space or fuzzy copy
5. Perform copy process
6. Start, if stopped



Target DB2

Target Job

7. Object ID translation, data masking and log apply
8. Update identity columns
9. Start target table and index spaces

Target locations that DB2 Cloning Tool supports

- **Within the same DB2 system or to another DB2 system**
- **To the same database name or to another database**
- **To the same table space name, or to another table space name**
- **To the same Creator ID or to another Creator ID**

DB2 Cloning Tool does the Catalog Research for you

- **Determines the source and target data sets**
- **Determines the source and target object IDs**
- **Determines compatibility**
- **Creates the XLATE parameters**
- **Tracks source DB2 extents**

Data Masking

- **Option to mask one or more columns during the table space refresh process**
- **The masking changes are made during OBID translation step based on masking rules that are enabled during the copy**
- **All referential integrity (RI) columns can have same masking function applied**
  - Examples of fields that a user might change are US--- >Social Security numbers< --- credit card numbers, names and addresses
- **NEW in 3.1 - DECFLOAT column data is now supported for STATIC, SCRAMBLE and USEREXIT rules.**

Data Masking

- **Types of data masking functions are provided, such as:**
  - *PATTERN RULE (Sir | Mr) Bill*
  - *SCRAMBLE RULE , SCRAMBLE(FIELD)*
  - *USER EXIT RULE , FIELD = USER_EXIT()*
  - STATIC RULE , FIELD = CONSTANT VALUE
  - MASK RULE, FIELD = [a-z0-9]*10
  - RANDOM RULE, FIELD = RAND(1, 100)
  - SEQUENCE RULE , SEQ(1, 1)
  - CURRENT DATE, CURRENT TIME, CURRENT TIMESTAMP RULES
  - CURRENT USER RULE

NEW in v3.1 - Table Space Cloning Enhancements

- **Create Target DDL and target table spaces and index spaces that do not already exist**
- **Step restart a table refresh process that ends in error**
- **LOG-APPLY to provide transactional consistency without stopping the source objects**
- **Clone objects by specifying only their DB2 STOGROUP**
- **Increased MAX-SUBTASKS limit to 99 to increase parallelism and decrease elapsed time**
- **Added ISPF option to check the index keys of indexes (was only in batch)**
- **LONGVAR (v8) to VARCHAR v(9) supported**
- **WARN-ON-SIMPLE-TABLESPACE**
- **DECFLOAT column data masking support**

NEW in 3.1 – Create target DDL and target table spaces and index spaces that do not already exist

- **Optionally creates and/or executes DDL for target objects that don't exist**
- **CREATE DDL is generated for the following objects:**
  - Databases
  - Table spaces
  - Tables
  - Indexes
- **LOB and XML spaces are supported**
- **All referenced STOGROUPS, distinct types and other supporting objects must exist on the target to be able to execute the DDL generated**

NEW in 3.1 – Step restart a table space refresh process that ends in error

- **A new repository keeps track of all data sets processed by Target job**
- **The repository allows a failed Target Job to be restarted**
  - Skips data sets that were successfully processed
- **New Target Job Report**
- **Many of the columns in the report may *optionally* be added to the Target Job runtime repository**

NEW in 3.1 – LOG APPLY

- **Log Apply brings objects to consistent point when fuzzy copy is used**
  - Updates the target tables with source table DB2 log records
  - Log records are from before and **after** the completion of the data set copies
  - Commands:
    - **QUIESCE-POINT(Y/N)** - Issues a QUIESCE command for all table spaces (in the source job) unless they have a utility status that prevents it. This can reduce the amount of time to find the consistent point by reducing the number of log records to interrogate.
    - **COMMON-CONSISTENT-POINT(Y)** - Find a consistent point common to the source DB2 objects (all objects will be consistent with each other)
    - **COMMON-CONSISTENT-POINT(N)** - Specify consistency but not common across all source tables

NEW in 3.1 – LOG APPLY

- **Mini-logs are created from the source active and archive logs**
- **The mini-logs are dynamically created data sets containing extractions from the DB2 logs for the cloned tables**
- **The mini-log entries are applied to the target DB2 data sets when the object IDs are translated**
- **The target indexes must be rebuilt afterwards**

NEW in 3.1 – Clone objects by specifying only their DB2 STOGROUP

- **LISTDEF STOGROUP selects all objects in all databases that have STOGROUP as their default.**

- **For Example:**
  - If there are three DATABASEs: DB11, DB21 and DB31
  - And the default STOGROUP names are: STG1, STG2 and STG3

  LISTDEF STOGROUP STG% selects all objects in DB11, DB21 and DB31

Thank You for Joining Us today!

Go to www.ibm.com/software/systemz/events/calendar **to:**

- Replay this teleconference
- Replay previously broadcast teleconferences
- Register for upcoming events