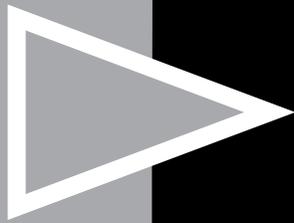




ACTIVATE BUSINESS WITH THE POWER OF I.T.™



DB2 Data Replication Using BMC Software Technologies

By Rick Weaver, BMC Software

Table of Contents

- Introduction 1

- Data Replication Using BMC Technologies 1

- Standard DBA Techniques 1

- Creative DBA Techniques 2

- Choosing the Right Technique 2
 - > Technique 1 2
 - > Technique 2 2
 - > Technique 3 2
 - > Technique 4 3
 - > Technique 5 3
 - > Technique 6 3
 - > Technique 7 3
 - > Technique 8 3
 - > Technique 9 3
 - > Technique 10 4
 - > Technique 11 4
 - > Technique 12 5

Introduction

Customers often need to replicate data from one DB2® z/OS subsystem to another DB2 system or RDBMS. They may want to update a QA environment with a copy of production data, or replicate or propagate data changes to a data warehouse. Because replication frequently is done at an application or object level, granularity and flexibility in the extract process is important. Solutions aimed at capturing all transactions and replicating them to another site for disaster recovery failover do not address this challenge.

During the replication process, users can face additional challenges such as:

- > Getting consistent data (no in-flight updates on the replicated output data set)
- > Reducing or eliminating any outage on the source DB2
- > Reducing resource requirements (CPU, I/O, and storage) to extract and apply the replicated data

Replication can be accomplished using native IBM DB2 utilities, service aids, or tools (DSNUTILB, unload, load, copy, DSN1COPY, Data Propagator), but meeting the above criteria for consistency, availability, and resource utilization constraints is increasingly problematic.

Some of the large enterprise resource planning (ERP) applications (e.g., SAP on DB2 z/OS) have a subsystem-cloning requirement. In these environments, “cloning” implies the DB2 datalog and directory are copied, as well as the application data. That requirement is not the focus of this paper. The focus here is on extracting a subset of DB2 data at an application level.

Data Replication Using BMC Technologies

BMC Software offers several techniques to help replicate data from a DB2 z/OS source to a target system (usually also a DB2 z/OS system, but not always). The various techniques build on the following products:

- > BMC CATALOG MANAGER for DB2
- > BMC CHANGE MANAGER for DB2
- > BMC COPY PLUS for DB2
- > BMC LOADPLUS for DB2
- > BMC Log Master for DB2
- > BMC RECOVER PLUS for DB2
- > BMC SNAPSHOT UPGRADE FEATURE
- > BMC UNLOAD PLUS for DB2
- > Online Consistent Copy (OCC), a feature of the BMC Recovery Management for DB2 solution

This paper presents BMC techniques (sometimes in conjunction with IBM tools) that at a higher level can aid DB2 z/OS data replication. This list assumes the presence of the native IBM DB2 utilities (DSNUTILB) and service aids (DSN1COPY). Because some operations can be conducted by either IBM or BMC utilities, the BMC brand name is specified on all BMC products. The use of specific BMC utility syntax or special function is noted in bold.

Standard DBA Techniques

The following recommended techniques use standard database administration practices for DB2 data replication, which is currently in widespread use across the industry.

1. DSNUTILB COPY or BMC COPY PLUS (both SHRLEVEL REF) on the source, followed by DSN1COPY or BMC RECOVER PLUS on the target (both may require OBIDXLAT¹)
2. BMC Snapshot Copy (BMC COPY PLUS and BMC SNAPSHOT UPGRADE FEATURE) on the source, followed by DSN1COPY or BMC RECOVER PLUS on the target (both may require OBIDXLAT)
3. BMC CHANGE MANAGER **Data Only Migration**, using DSNUTILB COPY or BMC COPY PLUS on the source, and BMC RECOVER PLUS **OBIDXLAT** on the target (this is an automated version of technique 1)
4. DSNUTILB UNLOAD or BMC UNLOAD PLUS (specify **FORMAT BMCLOAD** if using BMC LOADPLUS) on the source, followed by DSNUTILB LOAD or BMC LOADPLUS (specify **FORMAT BMCUNLOAD** if input is from BMC UNLOAD PLUS) on the target
5. BMC Snapshot Unload (BMC UNLOAD PLUS and BMC SNAPSHOT UPGRADE FEATURE, specify **FORMAT BMCLOAD** if using BMC LOADPLUS) on the source, followed by DSNUTILB LOAD or BMC LOADPLUS **FORMAT BMCUNLOAD** on the target
6. BMC CHANGE MANAGER **Data Only Migration**, using DSNUTILB UNLOAD or BMC UNLOAD PLUS on the source, and DSNUTILB LOAD or BMC LOADPLUS on the target (this is an automated version of technique 4)
7. BMC Log Master **MIGRATE SQL** on the source, followed by a **SQL High Speed Apply** engine on the target (target can be DB2 z/OS or any RDBMS such as Oracle on UNIX)
8. BMC CATALOG MANAGER Copy Table function (generates SQL SELECTs on the source, and INSERTs on the target)

¹ A note on OBIDXLAT: For replication techniques based on image copy input, the source and target DB2 data sets will have internal identifiers (DBID, PSID, and OBID) that must match or be translated. It is possible to CREATE TABLE with OBID specification, and it is possible to create PSIDs within a database in the same sequence between two DB2 systems. It is technically possible to have DBIDs match between two subsystems if all of the objects in both DB2s are created in the exact sequence, or if the target DB2 is a clone of the source DB2 (the DB2 catalog and directory are copied from the source and used to define the target). However, it is nearly impossible to ensure DBID, PSID, and OBID all match between two non-cloned DB2 subsystems. So, for those techniques that state “may require OBIDXLAT,” unless the DB2 catalog and directory are cloned, it may be assumed that OBIDXLAT will be required.

Creative DBA Techniques

The following techniques make creative use of BMC recovery technology to aid in DB2 data replication.

9. BMC RECOVER PLUS OUTCOPY ONLY on the source DB2 (possibly coupled with Timestamp Recovery feature of BMC Recovery Management for DB2 solution to render a consistent output image copy from a SHRLEVEL CHANGE input), followed by DSN1COPY or BMC RECOVER PLUS INCOPY on the target (both may require OBIDXLAT)
10. BMC RECOVER PLUS RECOVER INDEP OUTSPACE OBIDXLAT on the source DB2 with target DB2 data set name output (possibly coupled with Timestamp Recovery feature of BMC Recovery Management for DB2 solution to render a consistent output with SHRLEVEL CHANGE input)
11. Online Consistent Copy (OCC, a feature of BMC Recovery Management for DB2 solution) on the source, followed by BMC RECOVER PLUS using OCC input with OBIDXLAT on the target
12. Online Consistent Copy (OCC, a feature of BMC Recovery Management for DB2 solution), followed by BMC UNLOAD PLUS with OCC input on the source, followed by DSNUTILB LOAD or BMC LOADPLUS on the target DB2. The unload file created with input from the OCC is a consistent sequential file. It could be ported to some other platform and loaded there, using appropriate utilities

Choosing the Right Technique

When deciding which technique is best for your replication needs, the decision can be partially based on the tools already in house. However, depending on the three drivers (data consistency, source and target availability, and resource utilization), some of the techniques may be a better fit in certain environments. This section presents some of the benefits and considerations of each of the techniques presented above, as well as field “war stories” where relevant.

Technique 1

COPY on the source, followed by RECOVERY target (OBIDXLAT)

Benefit: Simple, can be supported by DSNUTILB COPY and DSN1COPY, or by BMC COPY PLUS and BMC RECOVER PLUS **OBIDXLAT**

Consideration: This is a DB2-to-DB2 solution. Source and target object schema must match exactly (e.g., segmented table space to segmented table space, all table and column definitions the same). Consistent copy must be created using

SHRLEVEL REFERENCE. This causes an application outage that can last several hours. DSN1COPY OBIDXLAT translation must be coded manually. BMC RECOVER PLUS can automate the OBIDXLAT. Both BMC COPY PLUS and BMC RECOVER PLUS are more efficient than the native utilities and use less CPU and elapsed time.

Technique 2

BMC Snapshot Copy on the source, followed by RECOVER on the target (OBIDXLAT)

Benefit: Simple process using BMC snapshot technology on the source DB2, resulting in a very brief outage on the source DB2

Consideration: Same as technique 1, except the availability impact of creating the consistent copy is drastically reduced. BMC Snapshot Copy essentially issues a QUIESCE for all the source objects, which flushes the buffer pools and establishes a recovery point. Once that is done, the objects are opened for read/write (R/W). BMC Snapshot Copy then uses one of several techniques to move the source pages to the output copy, making sure that any updates to the R/W objects are not put through until a “clean” version of the page is protected. The output copy always gets the clean pages, and therefore is registered as a SHRLEVEL REFERENCE copy in SYSCOPY.

Technique 3

BMC CHANGE MANAGER, Data Only Migration (automated technique 1 or 2)

Benefit: Automates the build and maintenance of the data migration process, thus reducing the overhead on DBA time

Consideration: BMC CHANGE MANAGER is a powerful tool to help the DBA in the alteration and migration of DB2 schema as applications go through development cycles.

Field war story: A customer used BMC CHANGE MANAGER with BMC RECOVER PLUS for DB2 to synchronize data warehouse and query environments with production. BMC Snapshot Copy was used for input. The complete process was automated, synchronizing both structures and data migration with OBIDXLAT. This product combination enabled automatic resolution of the multi-table table space OBID translation (a major advantage of using BMC RECOVERY PLUS). Further, the customer enjoyed fast and efficient recovery (including index builds) and dynamic allocation of all data sets, including multi-volume/multi-data set table spaces.

Technique 4

UNLOAD on the source, followed by LOAD on the target

Benefit: No need for OBIDXLAT processing; physical schema constraints allowed (e.g., segmented source to simple target table space allowed). Column definitions (sequence, data types) must match.

Consideration: Outage, CPU, and I/O on both source and target are required for consistent replication.

Technique 5

BMC Snapshot Unload on the source, followed by LOAD on the target

Benefit: A minimal outage on the source DB2 will create consistent unload data sets. Specifying **FORMAT BMCLOAD** reduces CPU time. If using BMC LOADPLUS, specifying **FORMAT BMCUNLOAD** reduces CPU time.

Consideration: BMC UNLOAD PLUS will need to QUIESCE all of the objects in the application before proceeding to create unload jobs. In a multiple-table partitioned table space environment, this can mean waiting for thousands of objects to clear. Further, the unload output data sets will be one per table space (even if there are 255 parts), which can lead to massive sort impact and excessive CPU.

Technique 6

BMC CHANGE MANAGER, Data-Only Migration (automated technique 4)

Benefit: Automates the build and maintenance of the data migration process, thus reducing the overhead on DBA time

Consideration: BMC CHANGE MANAGER is a powerful tool to help the DBA in the alteration and migration of DB2 schema as applications go through development cycles.

Technique 7

BMC Log Master MIGRATE SQL on the source, followed by SQL High-Speed Apply engine on the target

Benefit: This technique works well for propagating source environment changes to a target. BMC Log Master will browse the DB2 log, based on the filter criteria, and extract all DML (INSERT, UPDATE, and DELETE) for a selected set of objects. The DML will be copied to a file that can then be applied to the target DB2 via the SQL **High-Speed Apply engine**. This can be faster and cheaper than replicating the entire source application every time. It is especially useful for data warehouse applications. The BMC Log Master

Ongoing Process will ensure that the output file is consistent (only committed updates). In-flight updates are noted and carried forward to the next cycle. There is no outage to either source or target DB2 — all the work is done by reading log data and applying SQL. There will be CPU and I/O consumption on both the source and target.

Consideration: This technique assumes a one-time (or periodic) replication, using one of the other full-system replication techniques from the source to establish a base on the target system. The source and target TABLE schema must match (column definitions must be the same across source and target). The physical structures can be different (e.g., segmented table space to simple table space in DB2, or DB2 z/OS to Oracle).

Field war story 1: A large agency wants to keep 60 days of data online to DB2. It extracts updates to production daily DB2 with BMC Log Master **MIGRATE SQL Ongoing Process**, and ports the data to update an enterprise data warehouse (EDW) via the **High-Speed Apply engine**. Data is pulled from 73 of 202 DB2 tables. After using this technique in production for 22 months, the agency replicated almost 23 **billion** rows of data to the EDW.

Field war story 2: A large European telephone company wants to replicate data for disaster recovery purposes, but does not want to invest in expensive storage-based solutions. The company established a copy of the DB2 environment, and periodically transmits the SQL generated by BMC Log Master for application at a remote site. Based on the level of SQL activity, this technique is cheaper for the company to implement and sustain than a complete storage mirroring solution.

Technique 8

BMC CATALOG MANAGER Copy Table function

Benefit: No outage on source or target DB2. This technique is great for moving a subset of data from one DB2 to another without replicating the entire application.

Consideration: Generates SQL SELECTS on the source, and SQL INSERTs for the target. This technique generates CPU, I/O, and DB2 system overhead.

Technique 9

BMC RECOVER PLUS OUTCOPY ONLY on the source, followed by OBIDXLAT on the target (possibly using Timestamp Recovery)

Benefit: No outage on the source DB2

This technique makes use of normal image copy as input to replication process. This is a useful technique if the source and target DB2s are in different sites (or on different systems in the same site). The OUTCOPY ONLY syntax tells BMC RECOVER PLUS to use the existing source DB2 image copy as input; but instead of recovering the table space, the utility creates a new data set that can be ported to the target DB2. If the input image copy is SHRLEVEL CHANGE, the Timestamp Recovery process can be used on the source to render the output data set consistent (by applying log records up to a specified time stamp to the input copy).

Consideration: Both the source and target systems will experience CPU and I/O. An outage will be required on the target DB2 during the RECOVER OBIDXLAT process.

Technique 10

BMC RECOVER PLUS RECOVER INDEP OUTSPACE OBIDXLAT on the source DB2, with target DB2 Data set Name output (possibly using Timestamp Recovery)

Benefit: No outage on the source DB2. Uses existing image copy as input to replication. There is CPU and I/O consumption on the source DB2. There is an outage on the target DB2, but *no* CPU or I/O on the target system.

Consideration: This is a good technique when both source and target DB2 are in the same sysplex (or at the least on shared catalog and DASD environments). The **INDEP OUTSPACE OBIDXLAT** syntax tells BMC RECOVER PLUS to read the existing image copy on the source. Then, instead of recovering the source table space, the utility creates a new output data set with the naming conventions of the target DB2. OBIDXLAT is done on the fly. If the source image copy is SHRLEVEL CHANGE, then the **Timestamp Recovery** feature can be used to render the output data set consistent.

Field war story: A large financial company bought a part of another business that was in DB2. The company wanted to bring this new application into its existing DB2 environment, which meant the entire subsystem would not be “recovered.” The company decided to create the tables in its DB2 subsystem, take the image copies from the new business, and do the BMC RECOVER PLUS **INDEP OUTSPACE** recover using **OBIDXLAT**. This enabled the company to quickly move the data to its new data center, while using fewer resources from both the old and new environments.

Technique 11

Online Consistent Copy on the source, followed by BMC RECOVER PLUS using OCC input with OBIDXLAT on the target

Benefit: No outage on the source system to create the OCC. There is little CPU or I/O to create the OCC (it is driven by storage data set snap technology). OCC renders the OCC consistent with log apply technology. The OCC is then input to the BMC RECOVER PLUS **OBIDXLAT** utility on the target, where replicated data sets will be created.

Consideration: OCC requires the appropriate storage technology. IBM, EMC, HDS, and STK all support data set snap capability and can be used by OCC. OCC is a feature of the BMC Recovery Management for DB2 solution. It cannot be unbundled; a full product license is required to use OCC. BMC Recovery Management for DB2 includes all the functionality of BMC COPY PLUS, BMC RECOVERY PLUS, BMC SNAPSHOT UPGRADE FEATURE, BMC Log Master for DB2, and BMC RECOVERY MANAGER for DB2, as well as exclusive solution features such as OCC and Timestamp Recovery. Per data set, OCC is very fast. Most of the work is done by the storage device. However, there is a level of serialization to the process, so conducting OCC on a large number of data sets may take a while (there's no outage, but process time can climb). There is an outage on the target system during the OBIDXLAT recovery. The recovery will require CPU and I/O on the target system. The source and target schema definitions must match. Specifically, the table space, table, and column definitions must match (e.g., segmented TS to segmented TS).

Field war story: A customer had a requirement to periodically replicate production data to a QA environment. No outage was allowed for source, and source CPU was constrained. BMC successfully installed and configured BMC Recovery Management for DB2 and EMC Snap to exercise the OCC function of the solution. We created a 4+M row table with an index, and replicated it from one subsystem to another.

The stats:

Before: Unload .45 CPU, Load .77 CPU

After: OCC .03 CPU, R+ OBIDXLAT .13
(86 percent reduction in CPU)

The OCC time is expected to remain somewhat flat, even for large objects (the bigger the source objects, the more impressive the CPU savings on the source environment). The customer is very happy with the results and will put together a program that lets application QA staff use the OCC-based replication at will, since it produces a consistent copy with no impact on source availability and little on resource consumption.

Technique 12

OCC followed by BMC UNLOAD PLUS, with OCC input on the source, followed by LOAD on the target

Benefit: No outage on the source system to create the OCC. There is little CPU or I/O to create the OCC (it is driven by storage data set snap technology). OCC renders the OCC consistent with log apply technology. The OCC is then input to BMC UNLOAD PLUS to create sequential output data sets. These can then be loaded into another DB2, using any DB2 LOAD utility, or ported to another platform for loading. Consideration: OCC requires the appropriate storage technology. IBM, EMC, HDS, and STK all support data set snap capability and can be used by OCC. OCC is a feature of the BMC Recovery Management for DB2 solution. It cannot be unbundled; to use OCC, a full product license is required.

OCC is very fast on a per data set basis. Most of the work is done by the storage device. However, there is a level of serialization to the process, so conducting OCC on a large number of data sets may take a while (there's no outage, but processing time can climb).

The unload process, using OCC as input, requires CPU and I/O on the source system.

There is an outage on the target system for the load process. The load process will require CPU and I/O on the target system. The source and target table and column definitions must match.

Field war story: A company has the following requirement and environment:

- > Company objective is to extract all of the data from a production DB2 source and replicate it to the target environment (not a DB2 system). The target environment is used for customer reporting purposes and must be refreshed daily.
- > The source DB2 should experience no outage.
- > The extract file must be transactionally consistent (no uncommitted updates).
- > There should be minimal impact on source CPU for data extract.
- > Source data is a DB2 application with 24 DB2 tables, each in a partitioned table space, with 255 partitions (more than 6,000 data sets).
- > Source data is approximately 3TB (compressed).
- > The target is on another platform. The ported file must be a sequential file.
- > The extract and file format operation cannot exceed six hours. The extract file is then ported to the target environment and loaded there, allowing for next-day customer reporting.

Native utilities could not begin to meet the company's "consistent with no outage" objective. While BMC Snapshot Unload could approach that objective, the snapshot unload process drove a massive sort that exceeded the CPU and process time parameters. **The combination of OCC and UNLOAD PLUS created more than 6,000 consistent sequential data sets totaling more than 3TB within the company objective of six hours.**

All of the options included here assume that a consistent replicate is required. Some of the techniques require that the DB2 source be stopped to get consistency, while others have technology that can render a consistent replicate without an outage.

For clarification of the relative impact on both source and target in these techniques, see Table 1.

Technique	Options	Source availability impact for consistent replicate extract	Target availability impact for replicate apply	Source resource consumption
1	IBM DSNUTILB COPY, IBM DSN1COPY OBIDXLAT	Yes (High)	Yes (Medium)	Medium
	BMC COPY PLUS, IBM DSN1COPY OBIDXLAT	Yes (Lower)	Yes (Medium)	Lower
	BMC COPY PLUS, BMC RECOVER PLUS OBIDXLAT	Yes (Lower)	Yes (Lower)	Lower
2	BMC Snapshot Copy, IBM DSN1COPY OBIDXLAT	Yes (Brief)	Yes (Medium)	Lower
	BMC Snapshot Copy, BMC RECOVER PLUS OBIDXLAT	Yes (Brief)	Yes (Lower)	Lower
3	BMC CHANGE MANAGER Data Only Migration	As 1,2 above	As 1,2 above	As 1,2 above
4	IBM DSNUTILB UNLOAD, IBM DSNUTILB LOAD	Yes (High)	Yes (High)	High
	BMC UNLOAD, IBM DSNUTILB LOAD	Yes (Lower)	Yes (Lower)	Lower
	BMC UNLOAD, BMC LOAD	Yes (Lower)	Yes (Lower)	Lower
5	BMC Snapshot Unload, IBM LOAD	Yes (Brief)	Yes (Medium)	Lower
	BMC Snapshot Unload, BMC LOADPLUS	Yes (Brief)	Yes (Lower)	Lower
6	BMC CHANGE MANAGER Data Only Migration	As 4,5 above	As 4,5 above	As 4,5 above
7	BMC Log Master REDO SQL gen from Log	No	No	Medium
8	BMC CATALOG MANAGER Copy Table SELECT/INSERT SQL	No	No	High
9	BMC RECOVER PLUS OUTCOPY ONLY, DSN1COPY OBIDXLAT	No (if Timestamp Recovery is used)	Yes	Medium
	BMC RECOVER PLUS OUTCOPY ONLY, BMC RECOVER PLUS INCOPY OBIDXLAT	No (if Timestamp Recovery is used)	Yes	Medium
10	BMC RECOVER PLUS RECOVER INDEP OUTSPACE OBIDXLAT	No (if Timestamp Recovery is used)	Yes	Medium
11	BMC Online Consistent Copy, BMC RECOVER PLUS with OBIDXLAT	No	Yes	Medium
12	BMC Online Consistent Copy, BMC UNLOAD, IBM DSNUTILB LOAD	No	Yes	Medium
	BMC Online Consistent Copy, BMC UNLOAD, BMC LOAD	No	Yes	Medium

Table 1. Choosing a BMC Replication Technology



ACTIVATE BUSINESS WITH THE POWER OF I.T.™

About BMC Software

BMC Software delivers the solutions IT needs to increase business value through better management of technology and IT processes. Our industry-leading Business Service Management solutions help you reduce cost, lower risk of business disruption, and benefit from an IT infrastructure built to support business growth and flexibility. Only BMC provides best-practice IT processes, automated technology management, and award-winning BMC Atrium technologies that offer a shared view into how IT services support business priorities. Known for enterprise solutions that span mainframe, distributed systems, and end-user devices, BMC also delivers solutions that address the unique challenges of the mid-sized business. Founded in 1980, BMC has offices worldwide and fiscal 2007 revenues of \$1.58 billion. Activate your business with the power of IT. www.bmc.com.

About the Author

Rick Weaver is a product manager for DB2 solutions at BMC Software.

