Faster + Easier = Improved DB2 Disaster Recovery (using FlashCopy)

Judy Ruby-Brown

For years users have known that DFSMSdss could not be used for taking disaster recovery backups for DB2 data for two reasons. First, to insure good performance, DB2 does not update data base pages at the completion of a transaction. Instead, it writes the log records for the updates and holds the updated pages in its buffer pools in anticipation of re-reference. Pages are written to DASD when buffer pool thresholds are reached and, eventually, when DB2 is stopped. Second, there was no way to stop the writing of the log buffers to the DB2 log. Customers who wanted to employ DFSMSdss had to wait until DB2 was stopped, and this caused a lengthy outage as well as operational problems.

Most users implemented DB2-only disaster recovery procedures, which have been documented in the _DB2 Administration Guide_, under the topic "Remote Site Recovery from Disaster at the Local Site". Image copies of all tables and indexes and the archive logs covering the image copy interval were transported to a disaster recovery site. Following a conditional restart of DB2, the entire subsystem and application data were recovered, a process which consumed many hours and required deep recovery skills on the part of DB2 systems and DBA staffs.

DB2 UDB for OS/390 V6 provided the SET LOG SUSPEND command. This command allows user to temporarily "freeze" a DB2 subsystem by stopping updates to the log. The logs and database can be quickly copied using IBM's Enterprise Storage Server FlashCopy or equivalent vendor products to produce 'instant' copies. The outage can be reduced to a few minutes and minimizes operational disruptions to current users and operations staff. After shipping tapes, created from the backup copy, to a remote site, DB2 users can implement a simplified disaster recovery scenario, which also reduces the recovery time to that necessary to restore the tape dumps.

This scenario assumes a single local site. No remote copy technology is used. No transmission capabilities are employed. The following procedure should be performed at a point of low activity, since there is a short outage to obtain the 'instant' copy.

1. Stop DB2 updates.
    Enter –SET LOG SUSPEND on the OS/390 console
2. FlashCopy all DB2 volumes. When DSNJ372I is displayed on the OS/390 console, the DFSMSdss dumps can be started. Include any ICF Catalogs used by DB2, as well as active logs and BSDSs. When the jobs end, the logical relationship has been established
3. Resume normal DB2 activity.
    Enter -SET LOG RESUME on the OS/390 console
4. Make tape copies and transport off site. The tape copy of the secondary volumes can begin as soon as the FlashCopy relationship is established.

When the -SET LOG SUSPEND is issued from the OS/390 console, DB2 suspends logging and update activity for the current DB2 subsystem until a subsequent -SET LOG RESUME request is issued.   A highlighted DSNJ372I message is issued and will remain on the console until update activity has been resumed*. Note*: All work in progress will remain in its current state as of that moment; activity does not come to a commit point. There will be in-flight units of recovery, but they will be handled at the recovery site when DB2 is restarted.

This function is in the base DB2 V7 code. It was added to DB2 V6 as PQ31492 (on F9911 as UQ36695).   *Note*: It is available as a usermod for DB2 V5. It can be requested from the IBM Support Center. Since there is no -SET LOG command in V5, the suspend is implemented as -SET TIME(0) to suspend logging and -SET TIME(1) to resume logging.

These commands have member scope. For Data Sharing, the command must be issued on each data sharing member.

After DSNJ372I is issued, logging is suspended and the FlashCopy can be started.

Step 2: FlashCopy all DB2 volumes

All DB2 volumes in the DB2 subsystem can be dumped using DFSMSdss. Hardware which supports 'instant' copy is required because the copies must complete in only a few minutes. Units of recovery (UR), which are waiting on the log latch, hold locks. Transactions, which arrive later, may time out waiting on locks held by the log latch waiters.

Sample DFSMSdss JCL is shown:

```
//COPYJOB JOB...
//INSTIMG EXEC PGM=ADRDSSU
//SYSPRINT DD SYSOUT=*
//SYSUDUMP DD SYSOUT=V,OUTLIM=3000
//SYSIN DD *
 COPY FULL INDYNAM (srcvol) OUTDYNAM (tgtvol)  DUMPCONDITIONING
 /*
```

This step is ended when the copies are logically complete (the relationship established). At that point the copies from 'srcvol' to 'tgtvol' proceed independently in the DASD subsystem (not in the operating system environment).

Note: The copies may be written directly to tape by using the TSO interface to FlashCopy and specifying a mode of NOCOPY. The relationship persists until the tape copy is complete. This option is intended for disaster recovery backup, as no copy is available for local FlashBack. A description of this option may be referenced in the document listed at the end of this paper.

Step 3: Resume normal DB2 activity

When –SET LOG RESUME operator command is issued, normal DB2 update activity resumes.

Step 4: Make tape copies for off site transport

The copies must be dumped to tape and taken offsite. This is not a trivial task and can take hours. The amount of time required is a function of the number of volumes to be dumped and the availability of tape devices.

The effect of DFSMSdss COPY statement options
1. COPYVOLID – It is used for SMS-managed volumes, but tape dumps must be made on another OS/390 image. The secondary device will be varied offline when the VOLID is copied, as requested in the COPYVOLID parameter.
2. NOCOPYVOLID without DUMPCONDITIONING – It can be used for non-SMS managed volumes. The tape dumps can be made on the same OS/390 image as the original because the volume serial is different. At the recovery site, the volume serial must be clipped to the original one.
3. NOCOPYVOLID with DUMPCONDITIONING – It can be used for SMS managed volumes if OW45674 and OW48234 have been applied. The secondary copy becomes an 'interim' copy. The tape dumps can be made on the same OS/390 image as the original volume and the resulting dump tape will appear to have been made with the source volume serial.

   A copy of a source volume (VOL001 for example) to a target volume (VOL002 for example) with DUMPCONDITIONING specified, followed by a tape dump of the target volume (VOL002), results in a dump data set that looks like it was created by dumping the source volume (VOL001). For more information on this option, please see WSC Flash 10092 at
   http://www-1.ibm.com/support/techdocs/atsmastr.nsf/PubAllNum/Flash10092

These considerations are well known to the storage systems staff

At the Recovery Site

The goal of this scenario is to make it appear that DB2 crashed locally.

1. Use DFSMSdss to restore the FlashCopy data sets to DASD. Restoration will probably take in excess of an hour, depending on the number of volumes, which must be restored, and the hardware configuration available at the recovery site (tape drives, channels etc.).

   If NOCOPYVOLID without DUMPCONDITIONING had been specified on the COPY statement, the volume label must be clipped to the original volume serial following restoration of the tape.

For either COPYVOLID or NOCOPYVOLID with DUMPCONDITIONING, the volume is already labeled correctly.

2. Issue the START DB2 command using the local DSNZPARM member. A normal restart is performed. All in-flight and in-abort URs will back out at restart. In-doubts will remain until the coordinator is restarted. This is standard function for DB2 restart.
3. When restart is complete, new work can begin.
4. All utilities must be terminated, since they cannot be restarted.

Data Sharing Considerations

Local Site: The –SET LOG SUSPEND command must be issued on each member of the data sharing group. FlashCopy cannot be performed until all members have DSNJ372I displayed on their consoles.

Recovery Site: All DB2 structures in the Coupling Facility must be forced using SETXCF commands <u>before</u> DB2 is restarted. All members may be then be started, and DB2 will perform a group restart to populate the SCA and LOCK structures; the Group Buffer Pools will be allocated. Following restart, all data sets, which were in the Group Buffer Pools at the time of the FlashCopy, are set to GRECP exception state (Group Buffer Pool Recovery Pending). The user must issue –START DB commands to perform GRECP recovery before any new work involving those data sets can begin.

Comparison to Traditional Disaster Recovery

Since this scenario assumes there is no transmission capability, could it be compared to the traditional DB2 D/R scenario, where image copies and archive logs are transported from which the entire subsystem is recovered?

Issues	FlashCopy	Traditional DB2 D/R
Computer Infrastructure at recovery site?	None	None
Transmission capability	None	None
Transport	Daily	Daily
Image Copies and Archives required?	No, except inline copies for active REORG/LOAD jobs	Yes, and inline image copies for active LOAD/REORG jobs
DFSMSdss tape dumps	Yes	No
Recovery Point Objective	<24 hours	<24 hours
Recovery Time Objective	**1-2 hours** (# of volumes)	**<24 hours**
Procedures required?	Simple – DFSMSdss restore and DB2 restart	Complex – DB2 conditional restart and recovery
Skills needed	DASD – storage systems	Skilled DB2 systems/DBA
Outage incurred for backup	Few minutes	None

Faster, Easier DB2 Disaster Recovery using FlashCopy

Conclusion: The FlashCopy restore and restart provides an attractive alternative to the traditional DB2 recovery, at comparable financial cost. The recovery point objective (RPO) is the same, but the recovery time objective (RTO) is greatly reduced, to perhaps an hour.

- It requires no image copy decisions (full vs. incremental vs. mergecopy) for D/R.
- It requires suspending log for a couple of minutes to provide backup.
- It eliminates labor costs of manual procedures. It can eliminate software costs for vendor tools for recovery of the DB2 subsystem, application table spaces, and indexes.
- It transfers recovery responsibility from DB2 systems and DBA staffs to the storage systems team, for whom disaster recovery is a standard priority.

Additional Reference:
Implementing ESS Copy Services on S/390, SG24-5680, a red book downloadable from www.ibm.com/redbooks. See pages 163-165 or 4.19 "Using FlashCopy with DB2". For NOCOPY option see pages 148-149 or 4.5.2 "TSO commands".
.

The End