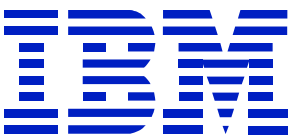
IBM SAP Joint White Paper:

Backup / Recover APO liveCache with ESS FlashCopy



September 4, 2003

Carol Davis	IBM
Cay-Uwe Kulzer	IBM
Norbert Pistoor	IBM
Christian Reith	IBM
Peter Jäger	SAP
Mona Mohan George	SAP
Dnyaneshwar Shelke	SAP

Table of Contents

Preface

The purpose of this white paper is to document the usage of ESS FlashCopy functionality and SAP DB liveCache (will be referred as SAP DB in this documentation) functions. The goal of this proof of concept is, to demonstrate that the combination of FlashCopy and the SAP DB functions in a customer environment provide a solution for quick SAP DB backup and recovery.

Furthermore FlashCopy Version 2 has been used also during the proof of concept. This version of FlashCopy dramatically improves the flexibility of this state of the art data replication function. It will for example allow consistent copies of data, without having to place applications in a consistent state. It also allows instant access to data after an FlashCopy has been made, even when background copy or nocopy option is still running. This option will make a backup and recovery with FlashCopy much easier to handle.

On the other hand, for users that would not like to upgrade their ESS to the new FlashCopy version, SAP DB will provide new functions that will allow SAP DB to be placed in a consistent state. These functions are called LOGWRITER SUSPEND and LOGWRITER RESUME. LOGWRITER SUSPEND will temporarily suspend IOs to the LOG volumes of SAP DB , allowing a fast replication tool like FlashCopy to run copies of a large amount of database volumes in a consistent state.

Therefore, this white paper consists of two parts:

1. Use of SAP DB LOGWRITER SUSPEND and ESS FlashCopy Version 1 to backup SAP DB
2. and FlashCopy Version 2, which does not need Writelog Suspend to create a consistent state SAP DB SAP DB.

Both methods have been used to create point-in-time consistent copies with FlashCopy. A SAP functional test team has verified that the SAP DB was not only consistent after a database recovery, but also checked that all data of SAP DB that was copied with FlashCopy was intact.

Introduction

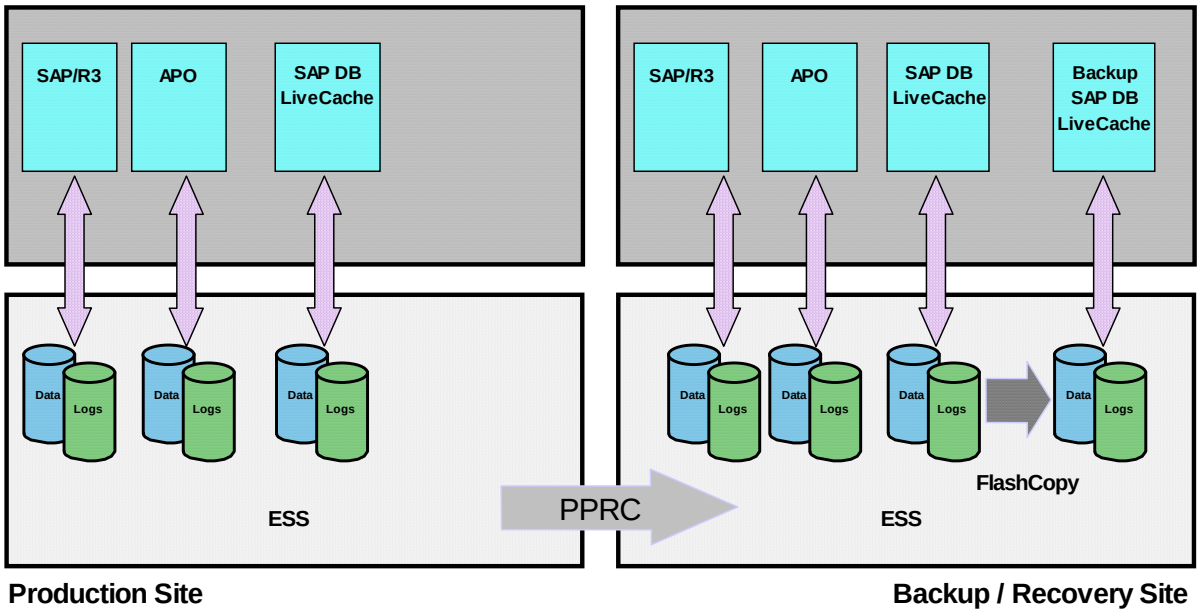
Recent new products such as SAP's B2B Procurement, CRM, mySAP.com Portals and Trading Exchange markets require an ever increasing need for continuous system availability. This paper provides information on an essential component of advanced infrastructure solutions – the High Availability Split Mirror Backup/Recovery (SMBR) for Systems of a mySAP landscape on SAP DB SAP DB operating in an AIX environment. In the following R/3 will be used as a synonym for SAP products.

The solution described in this paper is intended to deliver a backup process with no impact on the live R/3 system ("server less") using the advanced functions of the IBM Enterprise Storage Server (ESS). This "zero" downtime for the live R/3 system means that SAP users typically will not see a disruption of activity while the backup of the live database takes place. No transactions are canceled typically during the copy process /backup process. "Instant" availability of consistent copies of the database provides the ability to place an emergency system at the user's disposal while recovering the live database from a disaster. Beyond Backup / Recovery, a consistent copy of the database may be used for various purposes, such as creation of a reporting instance, production-fix system or repository for building Business Warehouse (BW) system.

This white paper, written from the DBA's perspective, addresses the infrastructure design, implementation tasks and techniques required for complex Enterprise Application Integration landscapes for high availability SAP R/3 applications consisting of a database (SAP DB), an operating system (AIX) and a storage subsystem (ESS) which all inter operate to deliver an easy-to-manage backup & recovery solution for SAP customers. Backup & Recovery solutions are mission critical activities in today's world of 7 x 24 computing and are a major focus for IT personnel, application management and DBAs. With the exploding growth in storage requirements for SAP application environments, this work touches on major elements of each area of technology spanning critical operating requirements and how this storage-centric solution delivers compelling value for SAP customers.

SAP Split Mirror Backup / Recovery (SMBR) Methodology

In order to deliver a solution that matches the complex requirements for high availability backup, recovery and performance, the SAP administrative workload should be taken off the live production system and performed on a copy of the system. The recommended environment is depicted in the figure below:

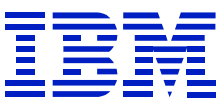


While a production server is connected to a primary fault tolerant storage controller, a mirrored remote copy of the production database is created without the computing support of the production server, in a separate secondary fault tolerant storage controller located at a spatially separated site, providing high availability and disaster recovery capability for business continuance. This copy can be accessed by a recovery/backup server should IT management procedures require its intervention in the event the primary site experiences business interruption. A local copy function within the secondary storage controller produces a consistent point-in-time copy of the production database. This consistent copy, at the option of the user, can be either held inside the secondary storage controller or be transferred to tape for remote vaulting. In the event that the primary database or its mirror is not available, the production server or the backup server can be connected to this disk resident copy, helping to dramatically reduce downtime. The database and SAP Basis administration can be augmented by providing a homogeneous split-mirror-based system copy for the purposes of test upgrades, support packages and database recovery routines. Intelligent storage controllers need to deliver this capability to create near-instant, consistent copies of the database.

Thus a number of key benefits can accrue to users from this recommended environment:

- Backup & recovery time is database size independent
- Minimum impact on production environment
- Physical and logical disaster prevention
- Enhanced database administration
- Remote data vaulting
- Offloading backup from production database server
- Reduction of backup-recovery-time to minutes

In order to realize this concept, the database management system needs to cooperate with storage subsystems to deliver the results. It should support the creation of a consistent database copy during the application READ/WRITE processing in a manner that exerts negligible impact on the production system, database or the live production storage server.

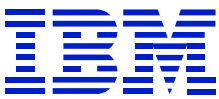


Considerations for a successful usage of SMBR

All objects in the R/3 environment need to be backed up. These objects are:

1. R/3 data objects:
   a. R/3 archiving objects
   b. R/3 Interfaces
   c. SAP Executables
2. Computing center data such as:
   a. The Operating System
   b. Third Party Programs connected to R/3
   c. Database objects

SAP DB supports database backups in ADMIN or ONLINE mode. Given a properly implemented database backup cycle, all forms of backup attain the same end goals. The specific circumstances at each SAP R/3 installation determine which of the methods is appropriate. The SAP tools for database backup support all options.



SAP DB and ESS Features to support SMBR

This section describes the key SAP DB features required to create a consistent backup database image of the production database using a set of specially developed SAP DB commands. The backup image includes all database relevant information, like logs, offline logs, traces, etc. The advanced functions of the ESS such as FlashCopy will be needed to create consistent point-in-time copies for backup and recovery purposes, while Peer-to-Peer Remote Copy (PPRC) may be included for disaster recovery procedures.

ESS Advanced Copy Services Features

- ***FlashCopy Version 1:*** The ESS administration tool (ESS Specialist) identifies the Logical Unit Numbers (LUNs) by their ESS internal serial numbers. FlashCopy, ESS's "near-instant" local copy function, can be used for all systems that have volumes or LUNs within the same Logical Subsystem (LSS) of an ESS. FlashCopy is set up using the Web interface of the ESS Specialist. Then, task selections on the volume pair – "FlashCopy" with Full or No Copy, and "WITHDRAW" options can be made.

  No Copy: This option in FlashCopy is useful if the copy operation has to complete with in a short time so that the source database/application are returned to their normal usage from a quiesced state. The relationship between source and target will automatically end if a physical copy should complete, which is an unusual situation for FlashCopy volumes in NOCOPY state.

  Copy: This parameter will create a point-in-time logical copy of a pair of ESS volumes that are defined as source and target volumes. Afterwards, the ESS will start a physical copy of the point-in-time version of the data on the source volumes to the target volumes. This option is of interest in case that a user would like to use the target volumes as clones of the source, or as backups. The relationship between source and target will end automatically when the physical copy is completed.

  Withdraw: command enables the removal of source and target relationships from a previously specified NOCOPY command.

Using the ESS Specialist, FlashCopy tasks - with either No Copy or Copy options are created. These tasks are initiated from the Unix command line using Command Line Interface (CLI) rsExecuteTask. The status of those tasks can be queried using the CLI, rsExecuteQuery. FlashCopy, when set up by the ESS Specialist, creates an identical copy of the source volume onto target volumes when an appropriate task is initiated using CLI. Volume identification or disk signatures need to be validated with respect to the host that is connected to the ESS in order for that host to start using the target ESS FlashCopy volumes. In order for a single host to mount both source and target volumes of FlashCopy pairs, AIX provides the RECREATEVG command, which is packaged as a PTF for AIX 4.3.3 in APAR IY10456. It is officially available in:

- AIX 4.3.3 Recommended Maintenance Level 05 (RML05)
- AIX 4.3.3 RML06
- AIX 5.1
- AIX 5.2

- ***FlashCopy Version 2:*** This version of FlashCopy enhances the flexibility of this unique point-in-time replication function of the ESS, especially when used as a backup and recovery function. Enhancements have been added to FlashCopy Version 1, and some are shown in the following list:

  Freeze Consistency Group: This option will allow a user to FlashCopy a set of application related volumes in a consistent state, without the need to quiesce, stop or suspend IOs to the ESS volumes. Therefore it is interesting for all types of data bases that do not provide the capability to temporary stop IOs to its volumes. When an ESS volume is replicated with FlashCopy FREEZE CONSISTENCY GROUP, the related volume will present a long busy state to attached hosts. The default setting is to freeze the volume for 120 seconds. If FlashCopy is run with this option against a DB LOG volume, it will freeze the DB, giving atomation enough time to FlashCopy the related data devices, since the DB application will not continue with IOs.

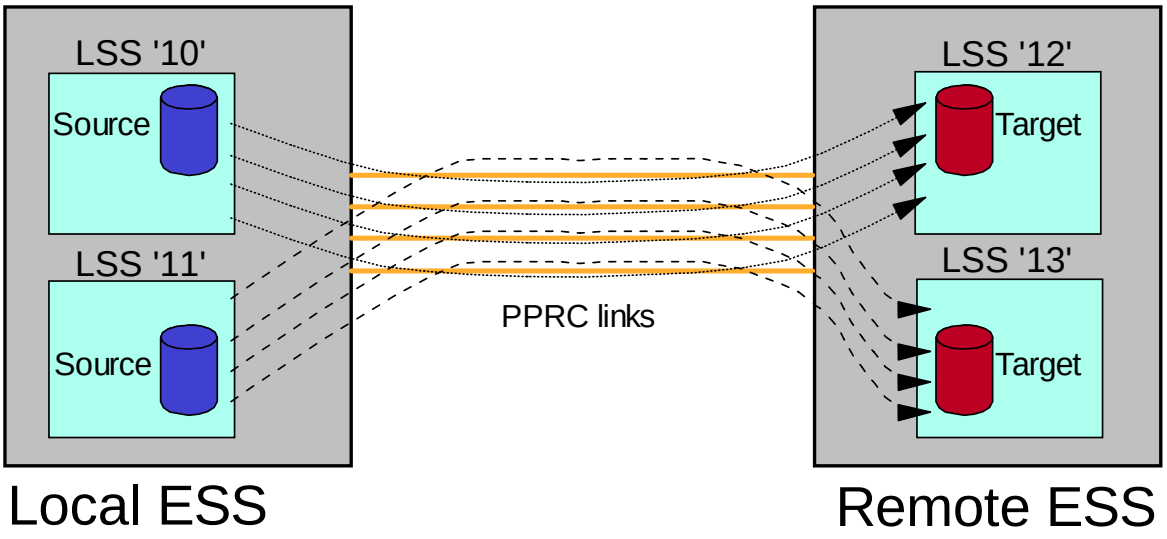
  Restore: This parameter will allow an immediate FlashCopy from target volumes to their corresponding source volumes, without having finished a background copy. This is an option that may be of interest if one needs to recover quickly from target volumes. In some instances this is referred to “Flashback”. The purpose of this operation is that a user does not need to recover volumes from tape or other media. This on the other hand results in a much less time consuming recovery operation.

  Incremental: This option can be used to FlashCopy only changed data since the last time a FlashCopy has been made to a FlashCopy pair. The advantage of this option is that only changed data needs to be copied to the target volumes, resulting in less IO traffic in the backend of the ESS disk.

PPRC (Peer to Peer Remote Copy) was not part of this proof of concept, but it will be briefly described for completeness purposes, since it offers additional advantages, especially when one considers it for disaster recovery at a remote location.

- ***Peer to Peer Remote Copy (PPRC) Version 1:*** PPRC is a hardware-assisted synchronous remote copy or synchronous mirroring function that can help preserve data integrity. Synchronous mirroring means that an I/O is only completed after it is acknowledged from the remote site. PPRC is set up on a volume or LUN basis in two or more ESS's, which are connected by ESCON as in the figure below.

Updates to a PPRC volume on the local or primary site (source or primary volume)



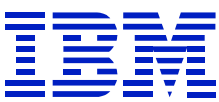
go first into cache and nonvolatile storage (NVS) in the primary storage. The updates are then sent over the ESCON links via larger ESCON frame transmission to the remote or secondary storage control. When the data is in cache and NVS on the secondary site, the receipt of the data is acknowledged and the local storage control signals to the application the completion of the update I/O. PPRC can be run up to 103 km distance and optimized ESCON protocols are used to ensure a fast transmission of data to the remote ESS. PPRC setups are made with the ESS Copy Services GUI or CLI and the following steps are required:

Establish links: This functions will establish communication from the local ESS LSSs to the remote ESS LSSs.
Establish pairs: Once links are configured, the user will be able to create the mirrored PPRC pairs. The following options are available:
Copy: A full physical copy of the source volume is performed.

RESYNC: Is used, if a pair was suspended for a while. In such a case, that local ESS will keep a bitmap of the changed data of the source volume. Just changed data will be transmitted to the target.
NOCOPY: Will setup a PPRC pair immediately without copying data. This may be useful in situation where a user would like to reverse a pair temporarilly.

- **PPRC Version 2:** This new version of PPRC offers some new functionality that enhances the disaster recovery scenario at a customer site. The main new option is briefly described in the next paragraph.

  Cascading PPRC: This PPRC setup will allow a user to create a third copy of data out of the PPRC target volumes configured in a remote ESS. This methodology will allow a user to mirror data to a far away location with PPRC-XD (Peer to Peer Remote Copy Extended Distance), which is an asynchronous PPRC solution capable of mirror data to far away locations without major performance impact. It can be temporarilly set to synchronous mode, ensuring, that from time to time a user will find a consistent state of data at the far away location.



SAP DB Features

For very large SAP DB databases in SAP R/3 production environments, the Split Mirror Backup Restore capability is essential for the creation of consistent database backup copies without stopping the production system. In order to make this possible, SAP DB created a set of special commands that enable temporary suspension of writes to the database:

- ***Suspend Logwriter:*** This command suspends the I/O writes to the SAP DB log volumes. However, read only transactions are allowed to continue. Any changes to the data, have to wait until the writes are subsequently resumed. Once the *SUSPEND LOGWRITER* state is in effect for the database, SAP DB does not complete a Savepoint. The last Savepoint is the reference for a restart or recovery.

- ***Resume Logwriter:*** This command resumes I/O writing to the log volume, allowing the kernel to complete a Savepoint and resume updates to the data volumes.

  Note: Naming and availability of this function in SAP DB is still to be defined.

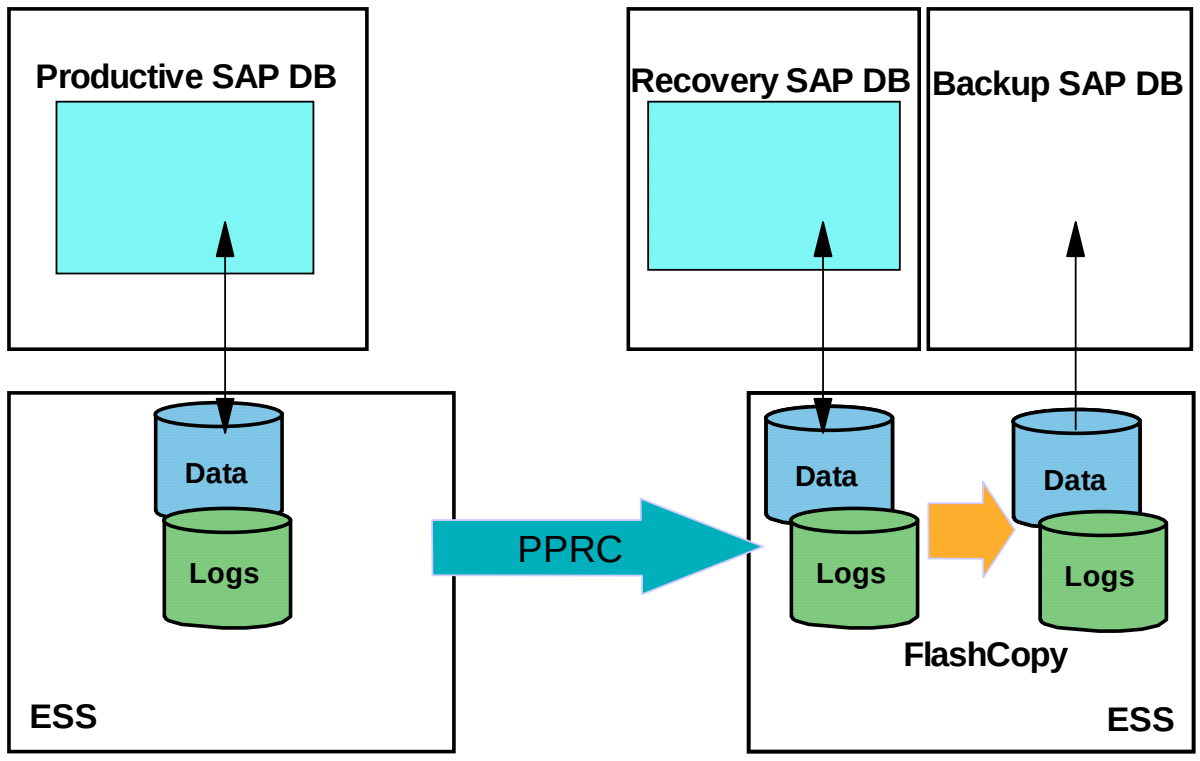
The advantages of these commands are, that the SAP DB will not perform updates to its corresponding data and log volumes. This on the other hand allows a disk system like the ESS to run a consistent split with FlashCopy of the data and log volumes, even when the FlashCopy executes sequentially one volume after the other. However, it should be noted that future improvements will be implemented in the ESS FlashCopy in form of a Consistency Group FlashCopy function. This function will allow a user, to run a FlashCopy without application (i.e. SAP DB) preparation. The user just needs to define a Consistency Volume Group and the ESS will take care that no updates will be made to the defined volumes while FlashCopy is executed. With this function, the SAP Split Mirror process will be much easier to perform, since no application preparation in order to create a consistent copy of data is required.

Basic Understanding of SAP SMBR

The SMBR setup involves the physical database design from the SAP installations guide. This includes file systems definitions according to sizing (based upon planned usage statistics, I/O forecasts, numbers of users etc.). The file systems requirements follow standard SAP installation procedures. In a full-featured high availability SMBR solution we recommend to use two physically separated database hosts and two ESS,each containing two copies of the production database. SAP testing procedures have been designed to verify the results. This testing suite includes:

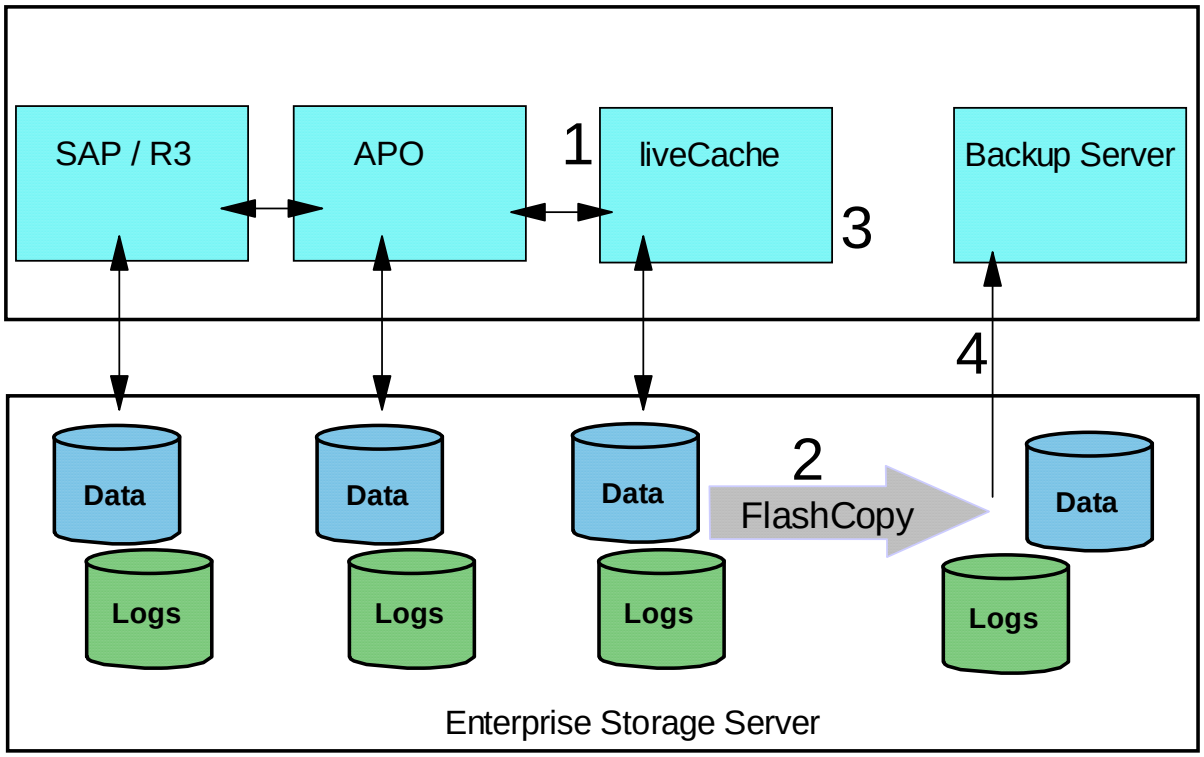
Updates of data volumes
SAP transactions

The ESS LUNs design is based on logical addressing of striped pysical volumes in a RAID-5 or RAID-10 array. The LUN definitions are based on file system requirements and are translated into volume groups via the IBM Subsystem Driver (SDD) mapping of virtual paths installed on the AIX host. SDD is a multipath software that provides higher LUN availability to a UNIX / NT host. LUNs may be configured also to be accessible via one path. In such a case, LUNs will be presented to AIX as hdisk without vpath assigment. The following figure shows the full scenario:



Test Scenario 1: SAP DB SUSPEND WRITELOGER and ESS FlashCopy Version 1

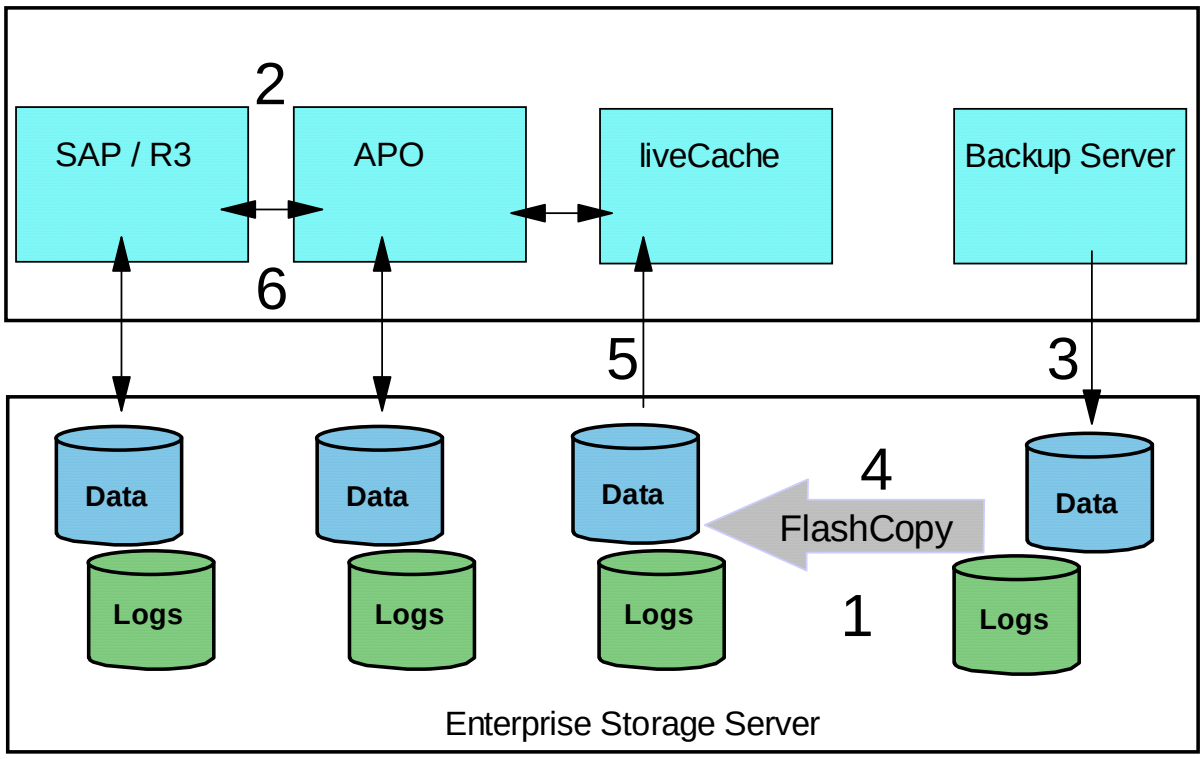
In many previous testings and proofs of concept, the ESS has successfully demonstrated its synchronous remote copy capability with PPRC. Therefore, in order to simplify the testings, the team has decided to document the important steps to backup and recovery SAP DB only. The figure below shows the basic setup:



To backup SAP DB perform the following steps:

1. Set SAP DB in SUSPEND LOGWRITER mode.
2. FlashCopy data and log volumes, archived logs, etc.
3. Set SAP DB in RESUME LOGWRITER mode.
4. Access the target FlashCopy volumes from the Backup Server. Mount file systems or raw devices and backup the data as “crashed image”. This means, do not restart SAP DB at the Backup Server, since it would recover the database. To backup the data one may use:
   a. DD statements from AIX
   b. Full image backup procedures with TSM
   c. or any equivalent method provided i.e. by Veritas, Legato, etc.

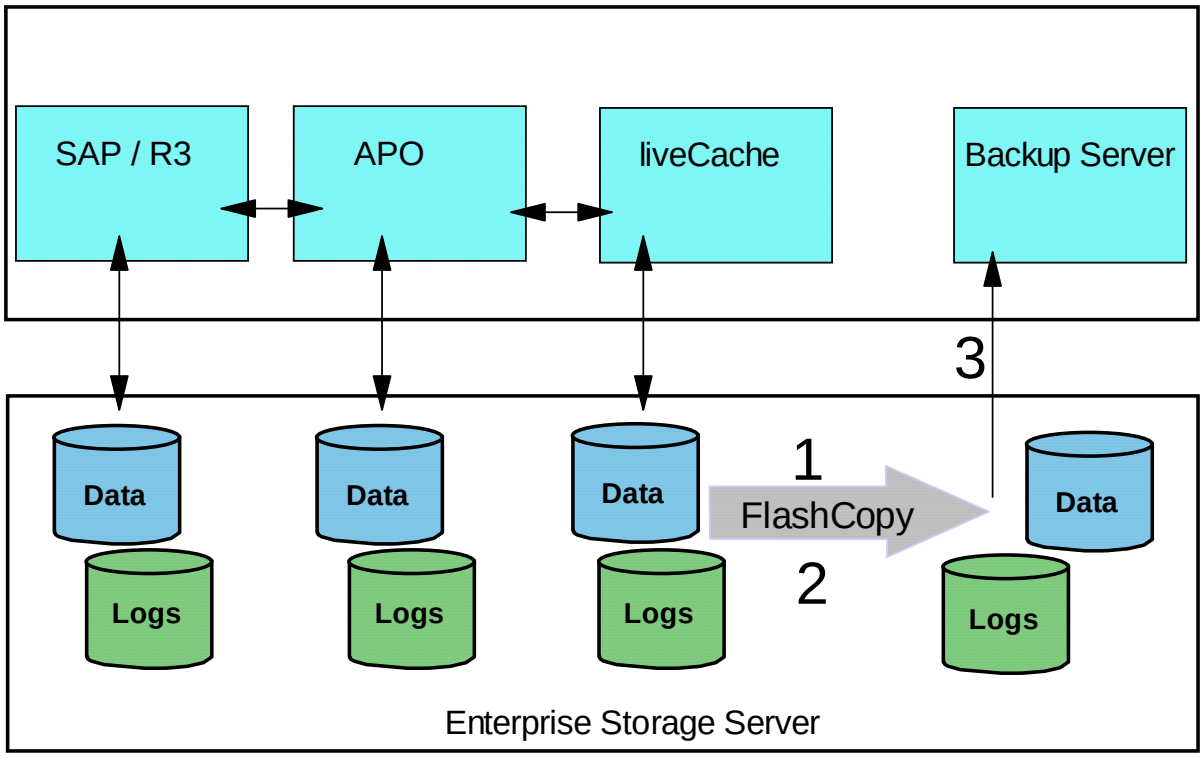
If a recovery should be needed, then proceed as follows:



1. Check for any previously running FlashCopy. If necessary, withdraw it.
2. Stop SAP/R3 and APO SAP DB systems.
3. Restore the backup via the Backup Server to their corresponding file systems or raw devices.
4. Run FlashCopy to the corresponding LUNs. Usually this is referred to as a Flashback.
5. Access the SAP DB from its corresponding DB server. SAP DB will now find the “crashed image” instance that was backedup previously. Since the image shows a non closed or stopped DB, SAP DB will run all necessary recoveries to check SAP DB data consistency.
6. As a last step, run any application dependent procedure to synchronize APO SAP DB to SAP/R3 (see SAP Note 425825, 481281 for further information).

Test Scenario 2: SAP DB and ESS FlashCopy Version 2

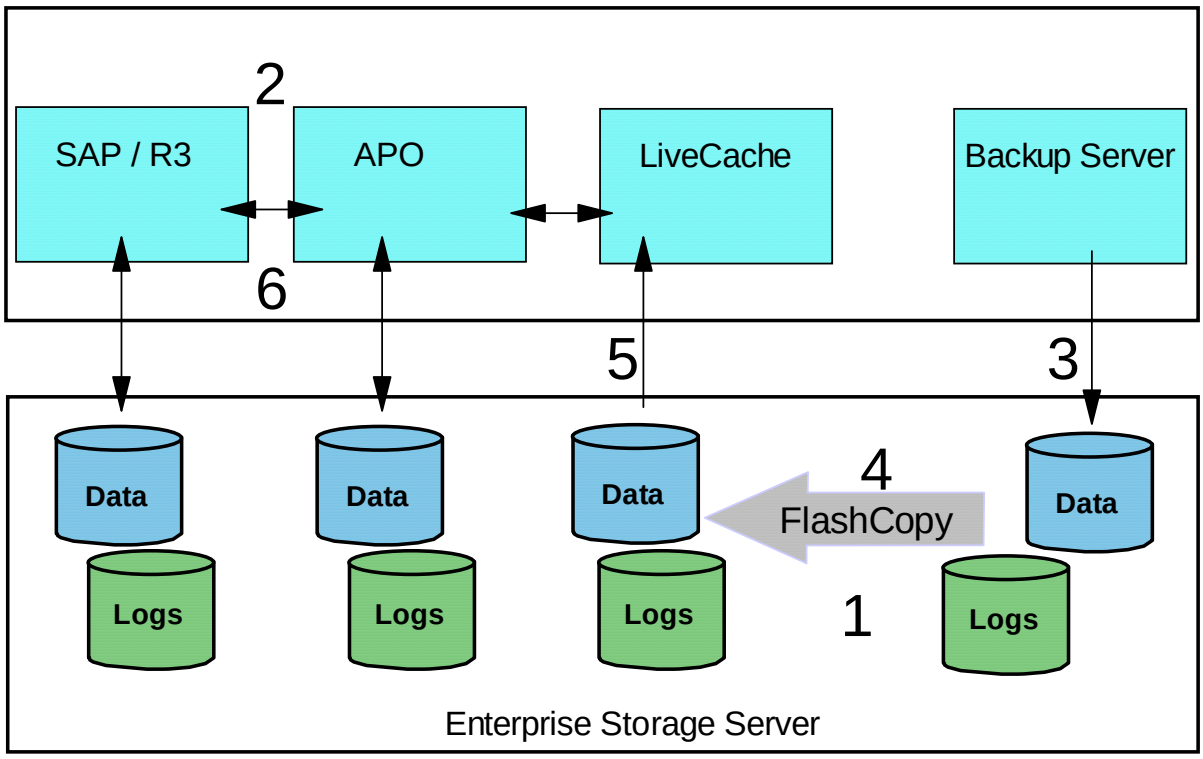
In many previous testings and proofs of concept, the ESS has succesfully demontrated its synchronous remote copy capability with PPRC. Therefore, in order to simplify the testings, the team has decided to document the important steps, to backup and recover SAP DB only. The figure below shows the basic setup:



To backup SAP DB perform the following steps:

1. FlashCopy data and log volumes, archived logs, etc. All volumes of the database must use the FREEZE CONSISTENCY GROUP option of FlashCopy.
2. Once FlashCopy FREEZE CONSISTENCY Group for all volumes has finished successfully, the LONG BUSY condition can be released.
3. Access the target FlashCopy volumes from the Backup Server. Mount file systems or raw devices and backup the data as “crashed image”. This means, do not restart SAP DB at the Backup Server, since it would recover the database. To backup the data one may use:
   a. DD statements from AIX
   b. Full image backup procedures with TSM
   c. or any equivalent method provided i.e. by Veritas, Legato, etc.

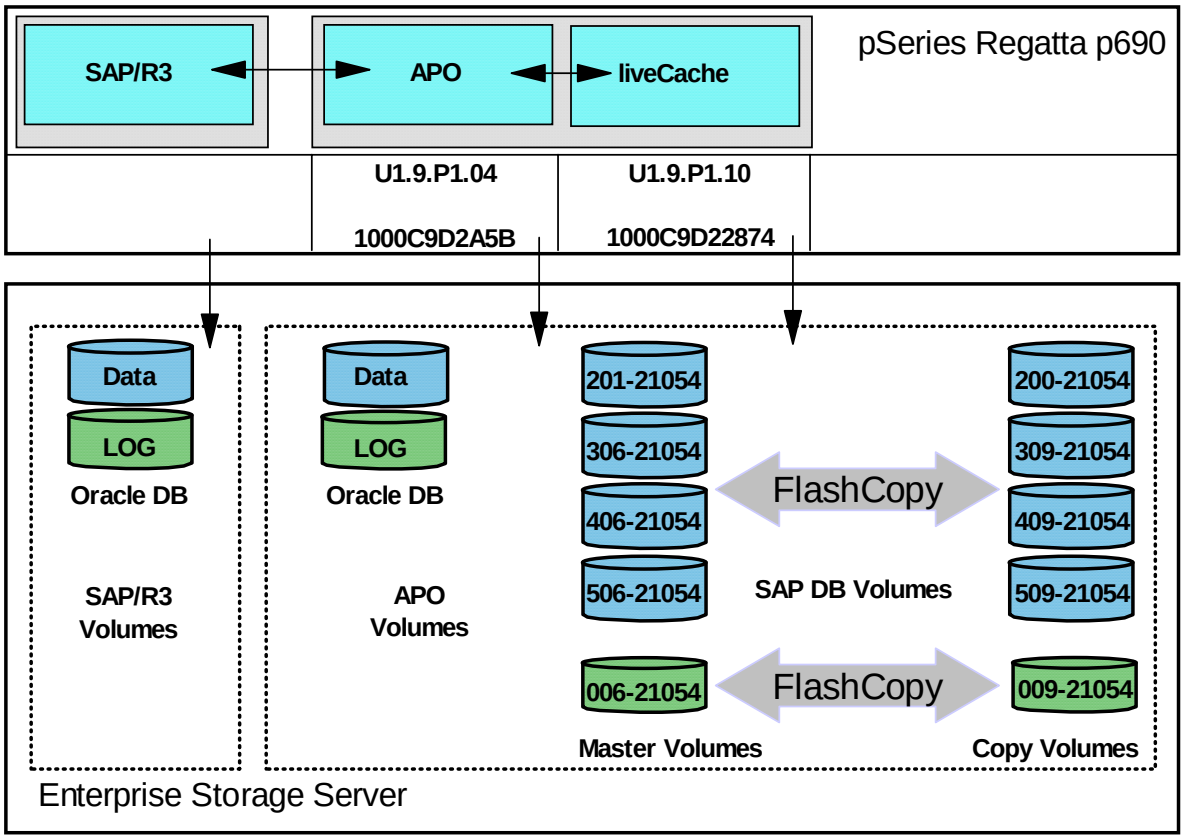
If a recovery should be needed, then proceed as follows:



1. Stop SAP/R3 and APO SAP DB systems and unmount all SAP DB relevant file systems or vary off the corresponding volume groups if one is using raw devices.
2. If the target volumes of a previous FlashCopy contain the latest backup, then you can proceed to the next step. Otherwise, recover the file systems or raw devices from the desired backup via the Backup Server.
3. Run FlashCopy RESTORE option to the corresponding LUNs. Usually this is referred to as a Flashback.
4. Access the SAP DB SAP DB from its corresponding DB server. SAP DB will now find the “crashed image” instance that was backup previously. Since the image shows a non closed or stopped DB, SAP DB will run all necessary recoveries to check SAP DB data consistency.
5. As a last step, run any application dependent procedure to synchronize APO SAP DB to SAP/R3 (see SAP Note 425825, 481281 for further information).

Testing Environment for SAP DB and ESS FlashCopy

The picture below gives the reader an idea of the testing environment that has been used for this project.



APO SAP DB will have access to their corresponding volumes in the ESS either via the fiber channel adapter denoted as U1.9.P1.04 or U1.9.P1.10. This will depend on which set of volumes are used for testing. The volumes 201-21054, 306-21054, 406-21054, 506-21054 are used as data volumes, while the volume 006-21054 is used as online LOG volume and it contains also a set of offline LOG volumes. Subsystem Device Driver (SSD), a multipath software delivered with the ESS, was not used in this environment. The data volumes were striped with the AIX LVM, while the LOG uses a separate volume. This setup will allow an independent FlashCopy of the volume containing the data devices and log devices.

The following table shows an assigment from AIX and APO SAP DB point of view.

ESS Volume	AIX VG Assigment	APO SAP DB
21054-006 or 21054-009	LClog	Online, Offline LOG volume
21054-201 or 21054-200	LCvg	Data volume
21054-306 or 21054-309	LCvg	Data volume
21054-406 or 21054-409	LCvg	Data volume
21054-506 or 21054-509	LCvg	Data volume

To find out the relation of ESS ID to AIX hdisk assignment, it is recommended to use the ESS CLIs. A useful CLI is the rsList2105s command. It gives a quick overview of how AIX hdisk relates to the ESS volumes. This on the other hand will simplify the search for the correct vg assignment.

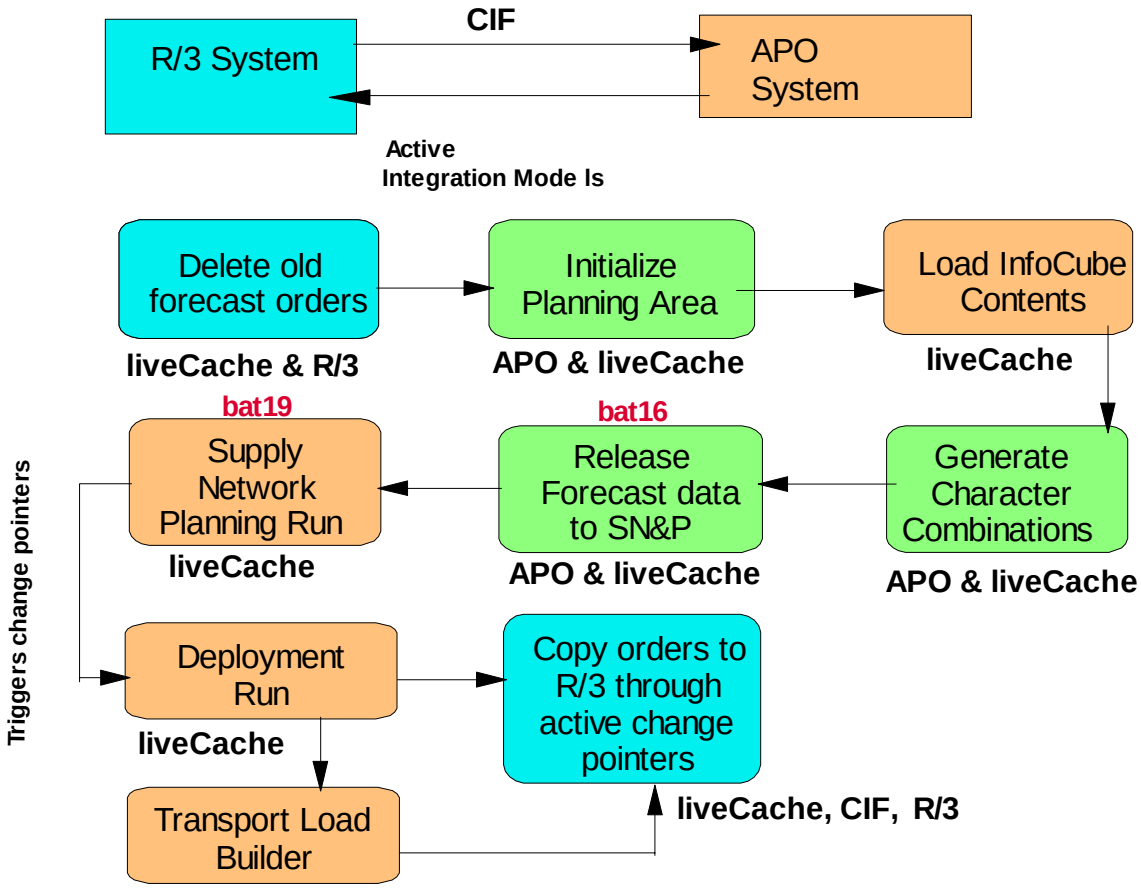
SAP DB was configured according to SAP guidelines with the following file systems:

/sapdb/NSL/data/DISKD0001
/sapdb/NSL/data/DISKD0002
/sapdb/NSL/data/DISKD0003
/sapdb/NSL/data/DISKD0004
/sapdb/NSL/data/DISKD0005
/sapdb/NSL/data/DISKD0006
/sapdb/NSL/data/DISKD0007
/sapdb/NSL/data/DISKD0008
/sapdb/NSL/data/DISKD0009
/sapdb/NSL/data/DISKD0010
/sapdb/NSL/data/DISKD0011
/sapdb/NSL/data/DISKD0012
/sapdb/NSL/data/DISKD0013
/sapdb/NSL/data/DISKD0014
/sapdb/NSL/data/DISKD0015

/sapdb/NSL/log/DISKL001
/sapdb/NSL/log/DISKL002

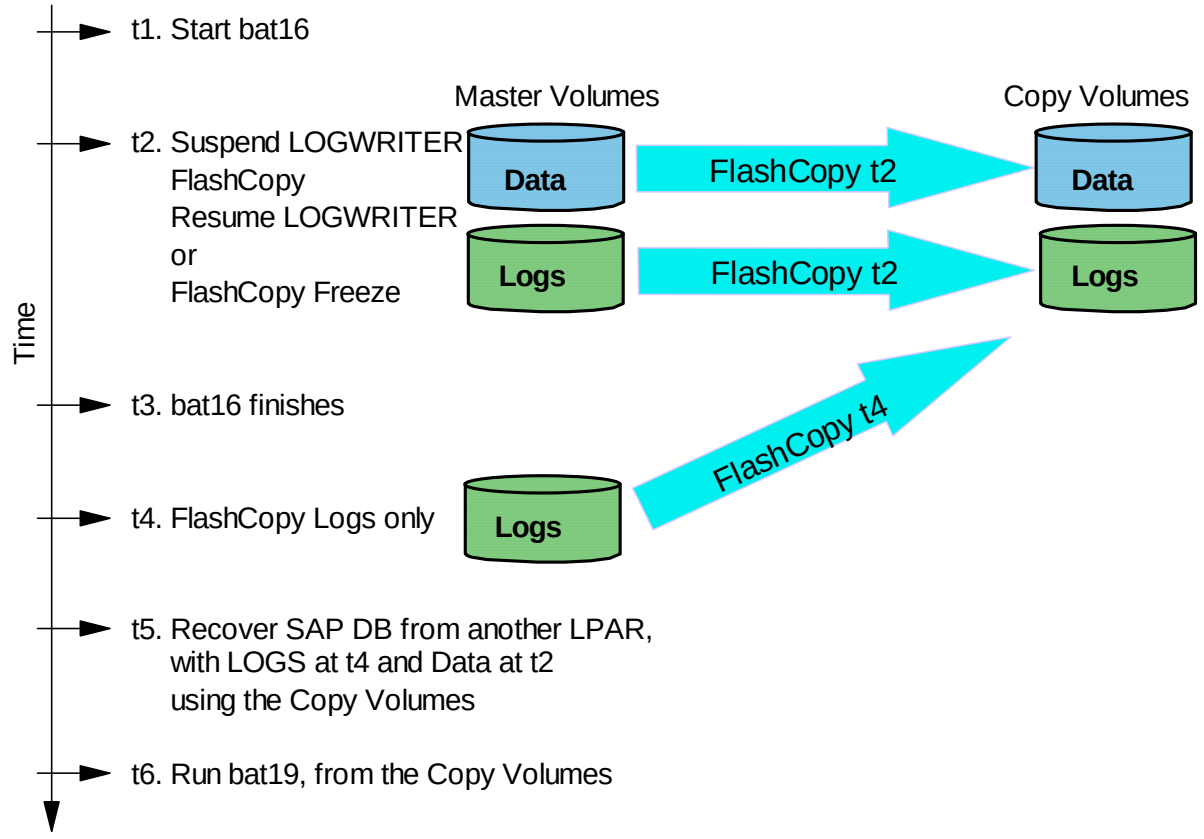
Test Verification Procedure

The picture below shows a flow chart of the batch jobs that have been used to put load on APO SAP DB. These procedures will do typical operations for this type of setup.



The test was positioned such that the FlashCopy would occur during the runtime of a job which was moving data down to the SAP DB. This data was required for the next batch job to run successfully. As these jobs are part of a well known scenario within the test box, the results of their complete and successful processing could be verified by the APO SN&P experts assisting the tests. The FlashCopy backup was run during the job **bat16** which releases the forecast data and transfers down to SAP DB from APO. The following important job battery, **bat19**, uses this data to run the SN&P Heuristics in SAP DB. This method allows the verification of the data consistency on a logical level as well as at database recovery level following the SAP DB recovery and roll-forward. In order to prove logical consistency, the tests were done during an actual customer test box SN&P batch load.

Furthermore, to be able to run the verification procedures as described before, the tests needed to run in a different way than described in Test Scenario 1 and in the Test Scenario 2. The following description should be used to understand how the SAP verification team has checked the consistency of SAP DB after FlashCopy. The picture below is used for further explanations.



The actual test procedure used for Test scenario 1 was as follows:

Establish the FlashCopy with running load

Start Bat16
Wait until the jobs were halfway done
Put SAP DB in ***Suspend Logwriter***
Wait for inactivity
Sync the file systems (OS command sync;sync)
Establish the FlashCopy across all volumes of SAP DB
Release Writelog Suspend with ***Resume Logwriter***

Continue the load, increasing the discrepancy between the FlashCopy and the master

Complete bat16

Simulate the crash

FlashCopy the current log volumes
(the result of this simulates the state of the Log volumes if the running system should have failed it is in a mounted and active state, suddenly disconnected from its server)
FlashCopy the device containing the archive logs
(both the archive-log volumes, and the log volumes are simulated as LVM copies, but rather than build actual LVM mirrors, FlashCopy was used to produce a copy for recovery purposes)

Recovery

Shutdown the LPAR and restart it using the FlashCopy SAP DB :
- Data volumes at point of FlashCopy
- Log and archived log volumes from the end of bat16

Recover the SAP DB to the end of Bat16 status

Verification

Verification at APO logical level using APO to SAP DB consistency checks (OM03 and OM13)
Application level verification of status by SN&P experts
Run the bat19 SN&P Heuristics
Application level verification of the Heuristics run by SN&P experts

Recommence Production

Start the follow-on job battery (bat19), SN&P Heuristics, to process the data prepared prior to the failure and recovery (prepared in bat16).

Reverification

Application level verification of the Heuristics run by SN&P experts to ensure that the output of the heuristics run is correct.

The actual test procedure used for Test scenario 2 was as follows:

Establish the FlashCopy at Time1 with running load

Start Bat16
Wait until the jobs were halfway done
FlashCopy Freeze Consistency Group across all volumes of SAP DB
Wait until the ESS releases the Long Busy State (this was used during the test, In a production environment, one would release Long Busy after FlashCopy terminates. The idea was to check if SAP and APO would have a problem with a Long Busy State).

Continue the load, increasing the discrepancy between the FlashCopy and the master

Complete bat16

Simulate the crash

FlashCopy the current log volumes
(the result of this simulates the state of the Log volumes if the running system should have failed it is in a mounted and active state, suddenly disconnected from its server)
FlashCopy the device containing the archive logs
(both the archive-log volumes, and the log volumes are simulated as LVM copies, but rather than build actual LVM mirrors, FlashCopy was used to produce a copy for recovery purposes)

Recovery

Shutdown the LPAR and restart it using the FlashCopy SAP DB :
- Data volumes at point of FlashCopy
- Logs and archived volumes from the end of bat16

Recover the SAP DB to the end of Bat16 status

Verification

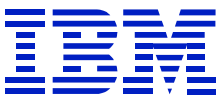
Verification at APO logical level using APO to SAP DB consistency checks (OM03 and OM13)
Application level verification of status by SN&P experts
Run the bat19 SN&P Heuristics
Application level verification of the Heuristics run by SN&P experts

Recommence Production

Start the follow-on job battery (bat19), SN&P Heuristics to process the data prepared prior to the failure and recovery (prepared in bat16).

Reverification

Application level verification of the Heuristics run by SN&P experts to ensure that the output of the heuristics run is correct.



SAP DB Recovery Procedure

In the tests, but also in the documented scenarios, SAP DB will be backedup as a “crashed image”. That means, after a restore SAP DB needs first to be recovered. The following steps document the actions required for that purpose. The procedure described here may differ from the scenarios described earlier, since for this test, the latest available LOG and archived LOGs have been replayed with FlashCopy at the end of battery 16 run. Should one not have the chance to replay the latest LOG to the last backup done, one will need to follow additional procedures to synchronize APO and SAP/R3 with SAP DB (see SAP Note 425825, 481281 for further information).

The method described here will use the dbmcli for recovery purposes. Once the production server has access to the disk, the user will have to go through the following procedure, in order to bring back liveCache operational.

1. Connected to dbmcli and set the database to “COLD mode” and review the SAP DB restart information with the db_restartinfo command as shown below:

```
dbmcli on NSL>db_restartinfo
OK
Used LOG Page           1427372
First LOG Page          1156632
Restartable             0
Id Restart Record       is02d3:NSL_20030402_182810
Id LOG Info             is02d3:NSL_20030402_182810
Consistent              1
```

   As expected, SAP DB was in a consistent state, but not restartable, since the log volumes have a different time stamp as the data volumes. This can be seen when looking at the file systems used by SAP DB:

```
17 2097152 -rw-rw----  1 nsladm   sapsys   2147482624 Apr  7 14:15 /sapdb/NSL/data/DISKD0001
18 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0002
19 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0003
20 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0004
21 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0005
22 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0006
23 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0007
24 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0008
25 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0009
26 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0010
27 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0011
28 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0012
29 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0013
30 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0014
31 2097024 -rw-rw----  1 nsladm   sapsys   2147336192 Apr  7 14:15 /sapdb/NSL/data/DISKD0015
17 2029312 -rwxrwxrwx  1 nsladm   sapsys   2078015488 Apr  7 14:32 /sapdb/NSL/log/DISKL001
18 1000064 -rw-rw----  1 nsladm   sapsys   1024008192 Apr  7 14:32 /sapdb/NSL/log/DISKL002
```

2. As a next step, one will need to identify the medium to which SAP DB has archived the offline logs. Proceed as follows:

```
dbmcli on NSL>util_connect
OK
dbmcli on NSL>>backup_history_open
OK
dbmcli on NSL>>backup_history_listnext
OK
CONTINUE

3DEB33580003|          |HISTLOST|2002-12-02 11:18:00|                   |2003-04
-14 10:11:00|                   |          |          |   |
                                    |          |          |         0|
.
.
.

3E9AE47E000A|LOG_00015|SAVE WARM|2003-04-14 10:38:02|2003-04-14 10:49:30|2003-04
-14 10:49:30|2003-04-14 10:49:33|   3176767|   3189344|   |logauto|     12520|
   1|         0|

3E9BC7C1000B|LOG_00016|SAVE WARM|2003-04-14 10:40:30|2003-04-14 10:50:09|2003-04
-14 10:50:09|2003-04-14 10:50:13|   3189345|   3202087|   |logauto|     12520|
   1|         0|

3E9BC7EE000C|LOG_00017|SAVE WARM|2003-04-14 10:50:09|2003-04-14 10:50:54|2003-04
-14 10:50:54|2003-04-14 10:50:57|   3202088|   3214911|   |logauto|     12520|
   1|         0|
```

The last line shows LOG_00017 as the first offline log to start the recovery with. Additionally to that, it also indicates the archieved media, which is: logauto

3. To start the recovery, one will need to enter the dbmcli the following command:

```
dbmcli on NSL>recover_start logauto LOG 017
OK
Returncode              -8020
Date                    20030408
Time                    00151904
Server                  is02d3
Database                NSL
Kernel Version          Kernel    7.4.2    Build 017-000-037-470
Pages Transferred       12520
Pages Left              0
Volumes                 1
Medianame               logauto
Location                /sapdb/NSL/log/log_archive/logauto.017
Errortext
Label                   LOG_00017
Is Consistent
First LOG Page          2056526
Last LOG Page           2069248
DB Stamp 1 Date         20030408
DB Stamp 1 Time         00132223
DB Stamp 2 Date         20030408
DB Stamp 2 Time         00132308
Page Count              12722
Devices Used            1
Database ID             is02d3:NSL_20030407_191924
Max Used Data Page
```

4. The return code -8020 indicates that there are more offline logs that need to be applied to SAP DB. Because of the architecture of SAP DB, the next offline log to be used is the next sequential one, which would be **`logauto.018`**. To continue the recovery enter the following command from the dbmcli:

```
dbmcli on NSL>recover_replace logauto /sapdb/NSL/log/log_archive/logauto 018
OK
Returncode              -8020
Date                    20030408
Time                    00161312
Server                  is02d3
Database                NSL
Kernel Version          Kernel    7.4.2    Build 017-000-037-470
Pages Transferred       12520
Pages Left              0
Volumes                 1
Medianame               logauto
Location                /sapdb/NSL/log/log_archive/logauto.018
Errortext
Label                   LOG_00018
Is Consistent
First LOG Page          2069249
Last LOG Page           2082030
DB Stamp 1 Date         20030408
DB Stamp 1 Time         00132308
DB Stamp 2 Date         20030408
DB Stamp 2 Time         00132706
Page Count              12781
Devices Used            1
Database ID             is02d3:NSL_20030407_191924
Max Used Data Page
```

5. As seen, the return code is still -8020, and one will need to proceed with the next offline log until the returncode shows 0. At this point, only a summary will be shown.

```
dbmcli on NSL>recover_replace logauto /sapdb/NSL/log/log_archive/logauto 019
OK
Returncode              -8020
.
.
.

dbmcli on NSL>recover_replace logauto /sapdb/NSL/log/log_archive/logauto 023
OK
Returncode              0
.
.
.
```

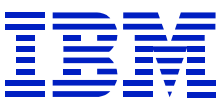
6. As said previously, Returncode 0, will indicate, that SAP DB is fully recovered. Now it is ready to be placed in warm mode to make it available. From the dbmcli perform:

```
dbmcli on NSL>db_warm
OK

dbmcli on NSL>db_state
OK
SERVERDB: NSL
The database state is WARM
```

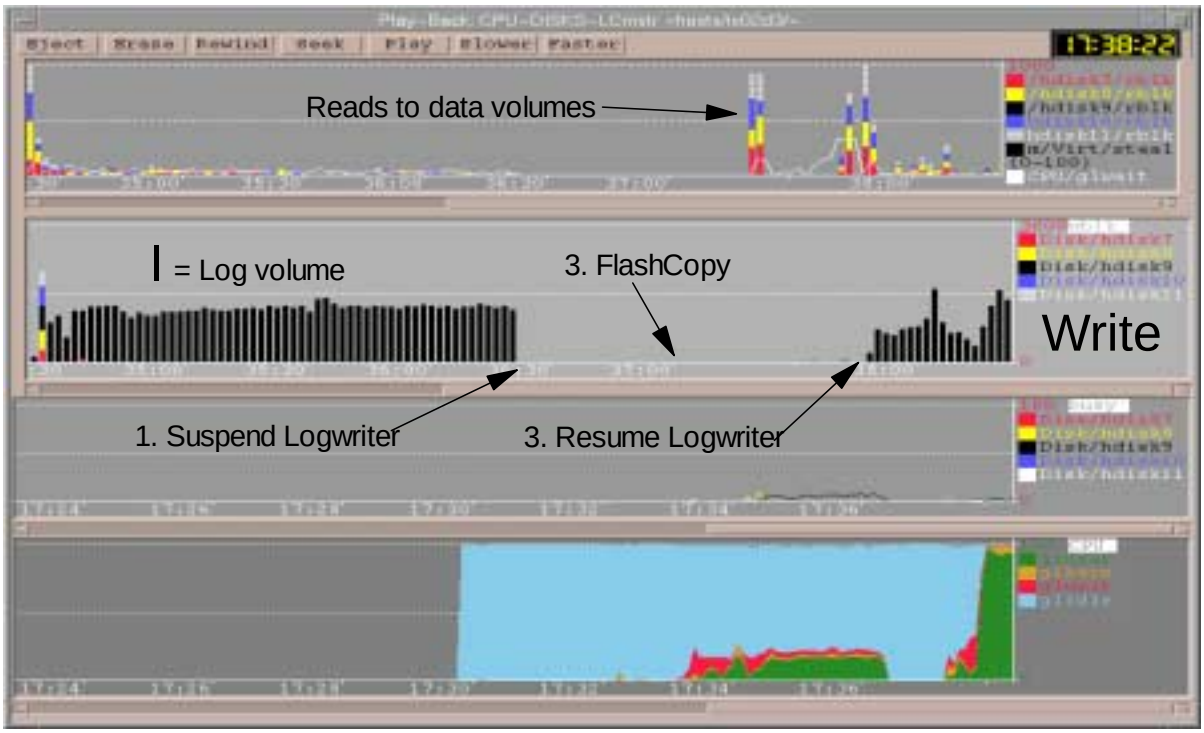
   This finishes the recovery of SAP DB.

7. As stated before, this procedure documents the specific scenario of our environment, in which we have made available the latest log and archieved log volumes with FlashCopy to SAP DB. If one does not have this capability, then additional recovery from APO and SAP/R3 may be required. These procedures are described in SAP Note 425825 and SAP Note 481281.



Disk and CPU Behavior during SAP DB Suspend Logwriter and Resume Logwriter

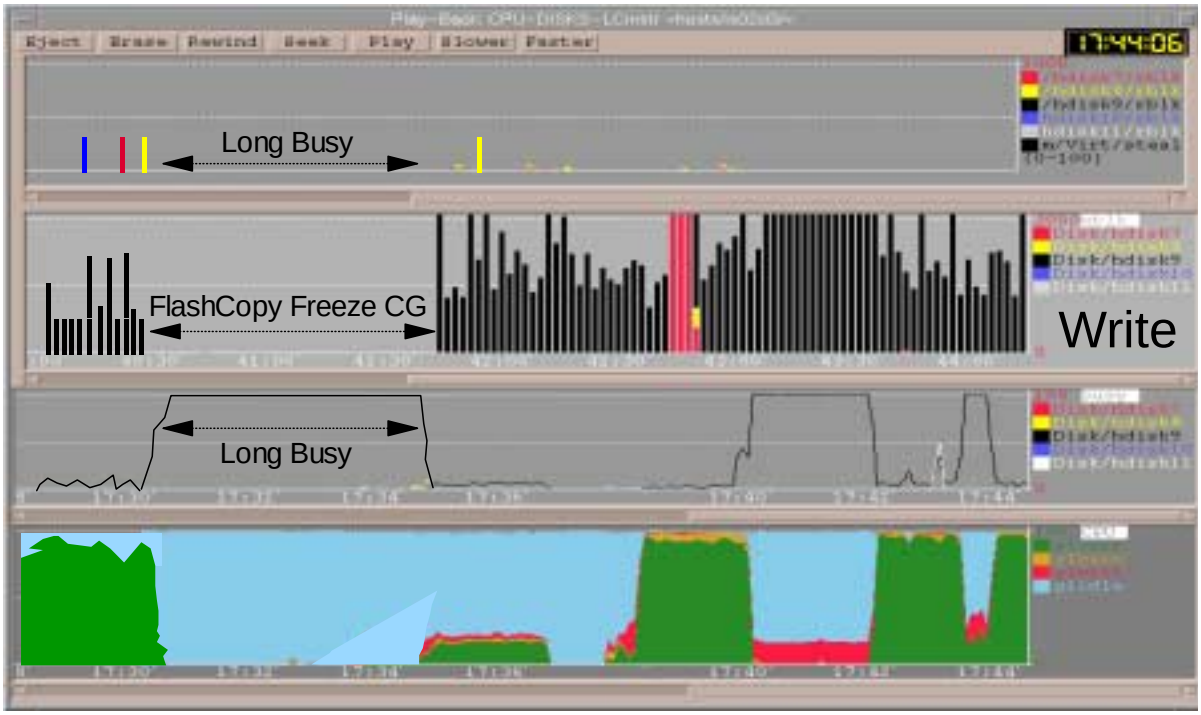
As described before in Test Scenario 1, SAP DB Suspend Logwriter will be used from the application point of view to create a consistent state of the database prior to be replicated by FlashCopy. While SAP DB is in Suspend Logwriter, no writes will be made to the entire database, but read operations will be possible. This was observed during the test, and the picture below shows this behavior.



The picture shows clearly that Suspend Logwriter stops all writes to its corresponding log volumes. After SAP DB is in Suspend Logwriter state it can be copied with FlashCopy in a consistent state. One can also see, that SAP DB may perform reads to the database, if it would be necessary. After the FlashCopy has run, SAP DB full operations will be resumed with Resume Logwriter. Shortly after issuing the command, one can observe write activity to the log devices again.

Disk and CPU Behavior during FlashCopy Freeze

As described before in Test Scenario 2, SAP DB will be replicated by FlashCopy with the option FREEZE CONSISTENCY GROUP. While FlashCopy FREEZE is running, one will observe a different behavior as it is with SAP DB Suspend Writeloger. The behavior of disk and CPU is shown in the next picture.



The picture shows the different behavior from SAP DB Suspend Logwriter. During the time the database volumes are in long busy state, neither reads nor writes are visible in the diagrams. The long busy is stated clearly by the 100% busy of the log volume, which was very active before. Once the long busy time is over (default 120 seconds), one can observe again log volume activity. One will not have to wait until FlashCopy Freeze resets the long busy state after the timer is over. The long busy state can be released immediately after FlashCopy has run. For the purpose of testing and demonstration it was not released during the test, to document the behavior shown above.

AIX and FlashCopy Freeze Setup Considerations

At the time when a FlashCopy Freeze Consistency Group is performed, the affected ESS volumes will present a Long Busy State to the operating system. During a Long Busy, no IOs , neither writes nor reads, are issued to the disk. As long as this happens in between the timeout specification of the disk, an operating system will not see that as a failure. Once the timeout settings are exceeded, the operating system will report failures to the application using the specified disk. This may cause applications to crash. The long busy duration presented by the ESS is set to a default of 120 seconds. Because of that, it is necessary either to:

- Change AIX timreout settings of the 2105 disk
- or change the long busy values in the ESS.

To check and change the timeout values of AIX proceed as follows:

1. Check the current value of the rw time out: "lsattr- El hdiskX"
   An example of the output of this command is shown below:

```
ibmr0010:/ > lsattr -El hdisk4
reserve_policy single_path Physical volume identifier        True
PR_key_value   none               400 MB SCSI Disk Drive     True
scsi_id        0x610913           N/A                        True
lun_id         0x530a000000000000 N/A                        True
location                          N/A                        True
ww_name        0x5005076300cf9ae6 N/A                        False
pvid           none               Physical volume identifier False
q_type         simple             Queuing TYPE               True
queue_depth    20                 Queue DEPTH                True
start_timeout  180                START unit time out value  True
Rw_timeout     60                 READ/WRITE time out value  True
```

2. Vary Off the vg, the disk is part of: "varyoffvg hdiskxx"
3. Change the value: "chdev -l hdiskX -a rw_timeout=yy"
   This value could be set to > 120s to match the default
   long busy settings of the ESS.

If one would like to adapt the ESS long busy time to the AIX timeout setting, then execute the following steps:

1. Start the Copy Services GUI and login.
2. Select Logical Subsystems (LSS).
3. From the next panel, select the ESS serial number you want to work with. Once an ESS has been selected, all available LSS will be shown.
4. Click with the left mouse button on the desired LSS and then select properties. A panel with the actual long busy time will be displayed.
5. Overwrite the long busy time with the desired value. It should be lower than the AIX time out value. This setting applies to all volumes in the selected LSS.

It is also important to understand that once a long busy is running, applications may go into a wait state. Depending on how critical the application is, a long busy may have to selected in such a way, that it matches the application requirements.

During the proof of concept there was no indication found that SAP/R3 APO SAP DB in this environment would have a problem even with long busy up to the defaults of the ESS (120 seconds). No transaction made by the different batches has been lost during all of the tests made.

ESS FlashCopy Version 1 Tasks

The following tasks have been used to run FlashCopy Version 1 in Test Scenario 1:

Group Task(t20)
Task(t21)
Task type: FlashCopy establish
Task options: none
Source: Target:
21054:12 21054:12
volume: 001 (20121054) volume: 000 (20021054)
Task(t22)
Task type: FlashCopy establish
Task options: none
Source: Target:
21054:14 21054:14
volume: 006 (40621054) volume: 009 (40921054)
Task(t23)
Task type: FlashCopy establish
Task options: none
Source: Target:
21054:13 21054:13
volume: 006 (30621054) volume: 009 (30921054)
Task(t24)
Task type: FlashCopy establish
Task options: none
Source: Target:
21054:15 21054:15
volume: 006 (50621054) volume: 009 (50921054)

Task(tlog2_CG)
Task type: FlashCopy establish
Task options: none
Source: Target:
21054:10 21054:10
volume: 006 (00621054) volume: 009 (00921054)

ESS FlashCopy Version 2 Tasks

The following tasks have been used for FlashCopy Consistency Group Freeze in Test Scenario 2:

Group Task(t20)
Task(t21)
Task type: FlashCopy establish
Task options: Freeze CG
Source: Target:
21054:12 21054:12
volume: 001 (20121054) volume: 000 (20021054)
Task(t22)
Task type: FlashCopy establish
Task options: Freeze CG
Source: Target:
21054:14 21054:14
volume: 006 (40621054) volume: 009 (40921054)
Task(t23)
Task type: FlashCopy establish
Task options: Freeze CG
Source: Target
21054:13 21054:13
volume: 006 (30621054) volume: 009 (30921054)
Task(t24)
Task type: FlashCopy establish
Task options: Freeze CG
Source: Target:
21054:15 21054:15
volume: 006 (50621054) volume: 009 (50921054)

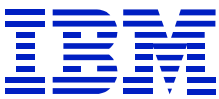
Task(tlog2_CG)
Task type: FlashCopy establish
Task options: Freeze CG
Source: Target:
21054:10 21054:10
volume: 006 (00621054) volume: 009 (00921054)

Disclaimer

The information provided in this document is distributed "AS IS" basis without any warranty either express or implied. IBM AND SAP DISCLAIM ALL EXPRESS AND IMPLIED WARRANTIES WITH RESPECT TO SUCH INFORMATION, INCLUDING ANY WARRANTIES OF MERCHANTIBILITY OR FITNESS FOR A PARTICULAR PURPOSE. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into their operating environment.

While the information contained in this paper has been reviewed by IBM and SAP for accuracy, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk. The performance data contained herein was obtained in a controlled environment based on the use of specific data. Actual results that may be obtained in other operating environments may vary significantly. These values do not constitute a guarantee of performance.

References in this document to IBM and SAP products, programs, or services do not imply that IBM or SAP intend to make such products available in all countries in which each company operates. Neither IBM nor SAP warrants the others products. Each company’s products are warranted in accordance with the agreements under which they are provided.



Acknowledgments

It is obvious that a proof of concept like this cannot be run by a single person. Many people have helped to make this project a success and special thanks is given to all people that paticipated in one or the other way. These persons are:

1. The group the has run the proof of concept:

Cay-Uwe Kulzer	TotalStorage EMEA Sales Support	IBM Germany
Carol Davis	Senior pSeries / AIX Specialist	IBM Germany
Christian Reith	Business Recovery Services	IBM Germany
Norbert Pistoor	pSeries / AIX Support Central Region	IBM Germany
Peter Jäger	SAP Technical Consultant	SAP Germany
Mona Mohan George	SAP Technical Consultant	SAP India
Dnyaneshwar Shelke	SAP Applications Consultant	SAP India

2. The sponsors, that helped with manpower and budget:

Eva Lau	Business Development & Alliances	IBM USA
Rey Rodriguez	IBM eServer Technology Solutions	IBM USA
Keith Murray	Manager ISSIC Walldorf	IBM Germany
Paul Hopkins	SAP Quality Manager	SAP Germany
Dr. Christian H Bartels	Manager Executive Briefing Center	IBM Germany

3. The technical support functions:

Frank Kaisinger	IBM SSG Product Engeneering Mainz	IBM Germany
Rainer Zielonka	IBM Director Competence Center	IBM Germany
Sam Werner	IBM SSG Product Engeneering Tucson	IBM USA
Steve Broadbent	IBM SSG Product Engeneering Tucson	IBM USA
Werner Thesing	SAP DB Lab Berlin	SAP Germany
Uwe Hahn	SAP DB Lab Berlin	SAP Germany

4. And last but not least for reviewing the document:

Dr. Immo Wegner	TotalStorage EMEA Sales Support	IBM Germany