

NetBackup Whitepaper Protecting SAP HANA

This document explains how to backup and recovery the SAP HANA in-memory database application with Symantec NetBackup.

This document applies to the following versions of NetBackup: 7.5.0.6, 7.6

If you have any feedback or questions about this document please email them to IMG-TPM-Requests@symantec.com stating the document title.

This document is provided for informational purposes only. All warranties relating to the information in this document, either express or implied, are disclaimed to the maximum extent allowed by law. The information in this document is subject to change without notice. Copyright © 2013 Symantec Corporation. All rights reserved. Symantec, the Symantec Logo, and the Checkmark Logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.



Document Control

Revision History

Version	Date	Changes
1.0	19 th July 2013	
1.1	12 th Sept. 2013	Clarification on where to find more information about hdbsql

Related Documents

Title	Date
SAP HANA Administration Guide – SAP HANA Appliance Software	
NetBackup for SAP 7.5.0.6 Administrator's Guide	
Scheduling SAP HANA backup from NetBackup	

Table of Contents

Introduction	1
Intended Audience	1
SAP HANA Overview	2
What is backint and what are the benefits that NetBackup provides?	2
The SAP HANA Structure	4
Simple data protection with HANA	7
How backint provides a simpler and better approach for data protection	7
Overview of Backup and Recovery	8
Backing up SAP HANA with Backint and NetBackup	9
Backup performance considerations	10
Running backups	10
Recovering instances	11
Data Backup	12
How the Backint communication works for backup	12
How a Backint backup is performed in SAP HANA Studio	17
Log Backup	20
The Backint Benefit	22
Log file storage considerations	22
Data Recovery	24
How Backint Communication Works For Restore	26
How a recovery via Backint is performed in SAP HANA Studio	32

Appendices

APPENDIX A – Other useful information

Introduction

This paper describes how to back up and recovery the SAP in-memory solution SAP HANA using Symantec NetBackup. It provides a conceptual overview of the process for data and log backup. It also describes the benefits of the backint implementation and how this differs from other methods of backup.

Any guidelines provided in this paper are general recommendations and do not target any specific size or type of SAP HANA implementation.

Intended Audience

This paper is intended for SAP and NetBackup administrators and technical solution architects. It is also suitable reading for Symantec Business Partners and Symantec employees who want to know more about the NetBackup for SAP HANA offering. This document can be used in the following ways:

- Architects can use this document to overcome perceived hurdles and challenges to SAP HANA backup.
- Engineers and administrators can use this document as a catalog of benefits.
- DBAs can use this document to gain an understanding of how database backup through backint can benefit them.
- Application owners can use this document to gain an understanding of how the SAP HANA infrastructure is protected in the background through backup.
- Business staff and process owners can use this document to get a conceptual overview how SAP HANA backup integrates in existing investments in Enterprise data protection with NetBackup for other parts of the company.

SAP HANA Overview

Big Data and Data Analytics are tremendously valuable tools for businesses today, delivering insights into every aspect of the business from customer interactions to network security. SAP created a new offering to address the requirements of today's businesses to access and mine information in near real time. SAP HANA is a hardware and software combination, delivered in an appliance form factor through SAP partners (HP, IBM, Fujitsu, Dell, Hitachi, Huawei) that integrates a number of SAP components. The SAP HANA database is a hybrid in-memory database that combines row-based, column-based, and object-based database technology. It is optimized to exploit the parallel processing capabilities of modern multi-core CPU architectures. With this architecture, SAP applications can benefit from current hardware technologies, helping customers to improve their operational efficiency, agility, and flexibility.

Like any other business application, Big Data can be impacted by downtime and disasters. A hardware failure or a data loss can have a serious impact on this valuable data source and possibly also affect a huge amount of information.

SAP has worked with partners, including Symantec, to create a new interface and certification process for backup solutions. The result is a "backint" API for SAP HANA. NetBackup integrates with this new API to provide a fast, effective data protection solution for SAP HANA and has been certified by SAP under the SAP Integration and Certification Center (SAP ICC). This document explains how NetBackup integrates with SAP HANA, how NetBackup leverages the newly created backint as well and the advantages this has.

Note: For information on how to configure NetBackup to protect a SAP HANA appliance please refer to chapter 6 of the [NetBackup for SAP 7.5.0.6 Administrator's Guide](#).

What is backint and what are the benefits that NetBackup provides?

Backint is the interface through which SAP HANA can interact with certified 3rd party backup products such as NetBackup. Backup, restore, delete and inquiry operations are specified by the backint interface, which also manages parallel operations.

In the current release (SPS05) SAP HANA orchestrates the backup and restore for the data and log files. This is true for the file backup method to disk as well as the newly created backint implementation. This is important to understand from a process flow perspective. Over time this will change and Symantec is

already working with the SAP HANA team to allow scheduled jobs to be run which have been initiated through NetBackup.

One benefit of this initial implementation is that there is nothing new the database administrator has to learn, thus allowing easier adoption of this new backint method. The tasks of backup and recovery are natively integrated into the SAP HANA Studio User Interface.

SAP HANA connects to NetBackup and transparently leverages the Media and Device Management capabilities (disk, tape, deduplicating storage). NetBackup provides extended capabilities to SAP HANA backup and recovery including encryption, creation and management of multiple backup copies, and cross site Disaster Recovery support as well as configuration flexibility to adapt to the different workloads for data and log file backup – all centrally managed in the NetBackup catalog database.

With the power of SAP HANA and the protection capabilities of NetBackup, businesses can enjoy an extra element of confidence working with Big Data. The combination of a highly resilient backup infrastructure and high performance architecture ensures SAP HANA backups and recoveries are done in the fastest possible way. NetBackup also complements the backup process with rich Alerting and Reporting capabilities through OpsCenter (not covered in this document).

Table 1 shows some of the benefits of the integration between SAP HANA and NetBackup.

	Available with SAP HANA backup to disk	Available with NetBackup for SAP HANA (backint)
Data backup	Yes	Yes
Log backup	Yes	Yes
Backup scheduling	Not available in SAP HANA Studio. External schedulers can be used in conjunction with scripts (SQL interface)	Yes, NetBackup SAP policies can be configured to initiate SAP HANA backups (refer to TECH209343 for details).
Manual backups	Yes: SAP HANA Studio, SQL commands (hdbsql), external scheduler triggering	Yes: SAP HANA Studio, SQL commands (hdbsql), external scheduler triggering
Backup of configuration files	No	Yes
Point in Time recovery	Yes	Yes
Recovery to a specific data	Yes	Yes

	Available with SAP HANA backup to disk	Available with NetBackup for SAP HANA (backint)
backup		
Backup Media and Capacity Management	No	Yes
Ability to encrypt data	No	Yes
Multi Copy and Retention Management	No	Yes

Table 1 – Backup approach comparison table

The SAP HANA Structure

In-memory computing is a technology that allows the processing of massive quantities of data in main memory to provide immediate results from analysis and transaction. To achieve the desired performance, in-memory computing follows several concepts:

- Keep data in main memory to speed up data access.
- Minimize data movement by leveraging the columnar storage concept, compression, and performing calculations at the database level.
- Leverage the multi-core architecture of modern processors and multi-processor servers.

With modern enterprise class servers is it now possible to hold several terabytes of data in main memory. The increase in capacity and reduction in cost makes of server main memory it a viable approach to keep huge amounts of business data in memory.

With processing happening completely in memory it is necessary to safeguard against the case of a power loss. It is important to guarantee that a transaction that was committed in memory remains committed. In order to achieve this SAP HANA captures all changes made by transactions via database logs stored on persistent storage such as hard drives, SSD, or Flash devices. Additionally, at regular intervals copies of the database memory content are written to the persistent storage, this is called a savepoint.

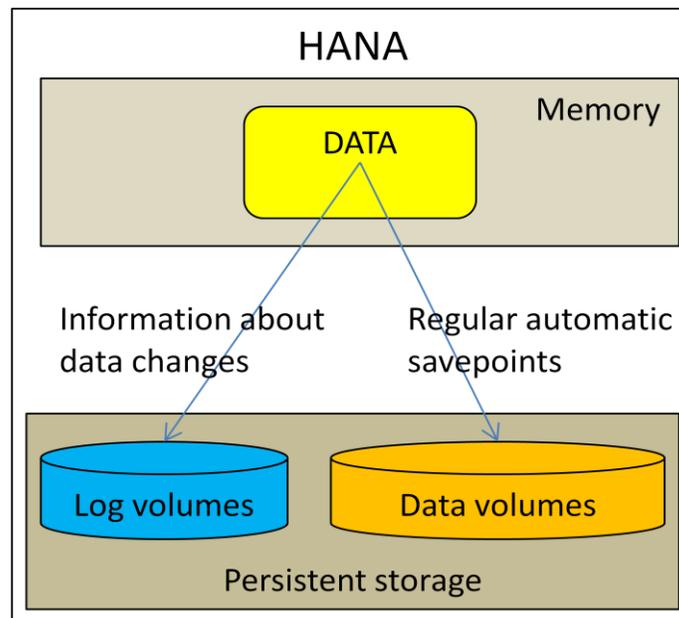


Figure 1 – SAP HANA components

Figure 1 gives a high level view the components involved. While SAP HANA itself operates in memory, the data and log volumes on the persistent storage layer allow the database to start quickly following a reboot, a node failover or a power outage. The backup process we talk about in the follow on sections is responsible for creating an additional independent copy to allow restores to the last state of the database or a point in time. This is aimed for protection against logical errors and damage to the persistent storage layer affecting the data or log volumes.

Some terminology and explanations:

Data

- Consists of SQL data and undo log information.
- Is kept in memory.
- The auto-protect write process of the data to the data volumes and log volumes happens asynchronously.
- Data backups create independent, consistent data copies to a secondary backup target.

Log

- Keeps track of all data changes.

- Log Buffers in memory are directly saved to the log segments stored onto the log volume when a transaction is committed.
- Log backups create independent copies of the log segments.
- Only if marked as backed up a log segment can be overwritten.

Savepoint

- Is a consistent state of the database.
- Is written automatically to the data volumes at least every 5 minutes (customizable).

The data in memory is also called payload. This payload is written into service-specific persistent storage files located in the file system in form of the data file and the redo log file.

The relevant content of all data volumes of all services (name server, index server, statistics server, etc.) that belong to a single or multi-host/node instance are marked by an instance-wide savepoint. In other words, the savepoint marks a consistent state of the data volume from which the database can be restarted.

The log segments contain all data changes that were committed to the database transactions in memory.

Any clean shutdown will sync the redo logs with the data files and triggers a savepoint.

However, in unclean shutdown situations (e.g. after a power failure) the database restart will return the database to its last consistent state by replaying the log segments since the last savepoint. The relationship between savepoints and logs can be seen in Figure 2.

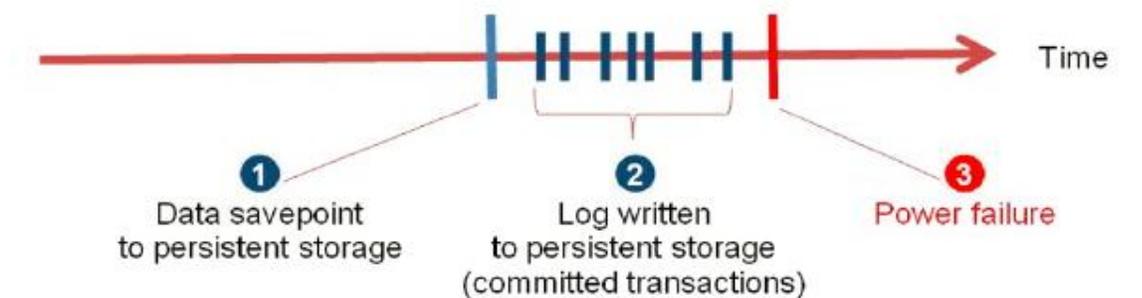


Figure 2 - Savepoint to log relationship

Simple data protection with HANA

In a holistic protection strategy it is necessary to think through the whole process of creating the secondary backup copy in terms of responsibilities, automation and control. For the built in method “backup to file”:

- All management aspects of the availability of the backup storage, proper storage management, providing enough space, cleaning out older backups, performance optimization, firewall settings, keeping track of retentions and additional copies become the database administrator’s responsibility.
- There is no linkage between an additional data copy of the “backup to file” data and the SAP HANA management that was created to balance available disk space for backup. Using this approach makes backup and recovery for SAP HANA a two-stage process.
- Manually managing the log backup copy can quickly become a challenge. See the section in this whitepaper about log backups for more detail on this.

How backint provides a simpler and better approach for data protection

Backint provides an easier method of protecting the data and takes away the entire administrative burden. It provides the additional intelligence to tell the backup application when it e.g. can delete old backups. The data sent through backint is additionally checked for integrity. NetBackup on the other hand manages the storage aspects for SAP HANA through the built in Media Manager. Through NetBackup tape libraries and tape drives can be shared between SAP HANA and any other backup in the environment (NDMP, File, Database backup). Disk Storage for log backups is managed with advanced features like Storage Grouping, High/Low Level Watermarks, I/O Limit settings, Load balanced access etc. Additionally the process of creating secondary copies for the backup (e.g. for storage in a remote location) is now fully taken care of.

Overview of Backup and Recovery

A SAP HANA node has a set of services that keep independent data and log files. A typical installation would consist of at least 4 services on a node: Index Server, Statistics Server, XS Engine, Name Server (see Figure 3) but there could be additional services depending on the specific database.

Each service has its own data and log portion that needs to be backed up.

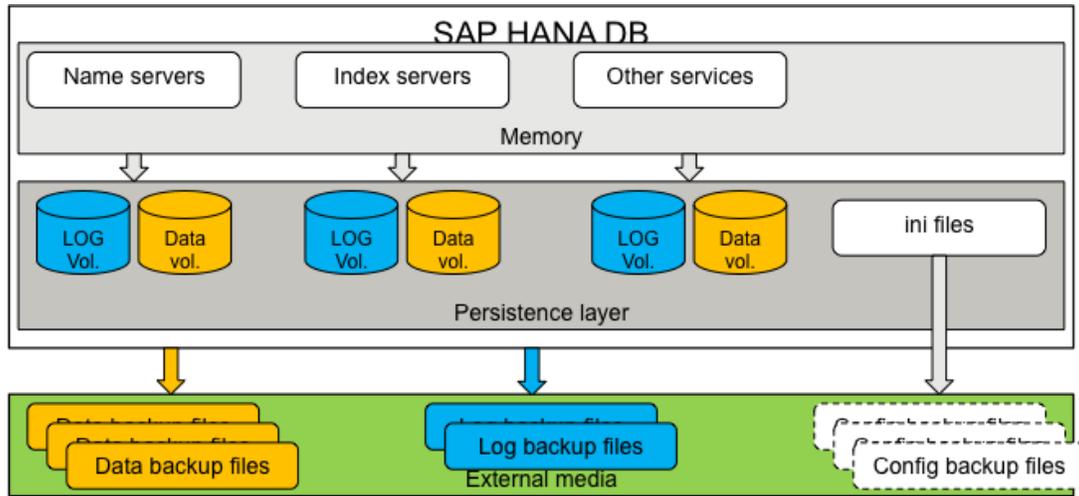


Figure 3 - Single node view of SAP HANA

In a multi node environment the database system spans multiple nodes as can be seen in Figure 4.

Each SAP HANA database system consists of one or multiple nodes. The system is represented through its SID. A SID can consist of one or multiple instances. On each node there is a set of services. The name service, although present on all nodes, is only active on one node at a time creating log and data information on the node it is running on.

A multi node SAP HANA installation consisting of 4 nodes (as shown in Figure 4) will typically 4 x 4 services. As it is only active at one node at a time, the name server and associated configuration files are only backed up from the active node. This equates to 13 data sources plus 13 log sources which require backup.

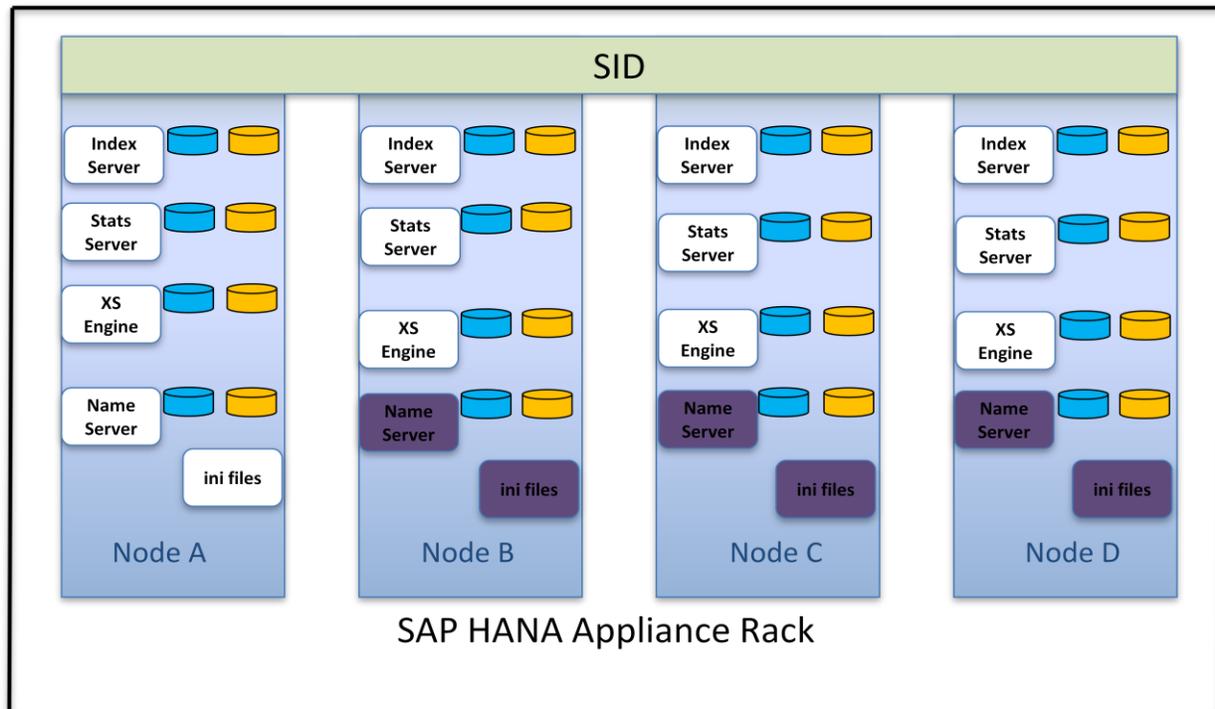


Figure 4 - A Multi Node SAP HANA environment

In a SAP HANA appliance a single SID can span multiple nodes and, in test & development environments, multiple SIDs can run on a single node (in production environments only one SID is allowed per node).

Backing up SAP HANA with Backint and NetBackup

Backint backups create data stream pipes that get filled from SAP HANA and read from by NetBackup. Each service creates an independent pipe. The Statics Server, the XS Engine and the Name Server only contain small amounts of data compared to the Index Server. In a backup situation the Index Server has the longest run time while the other services complete quickly.

A backup should not impact the database performance, as the main operations of SAP HANA are memory based. However, log files and savepoints are written to the file system at regular intervals. To minimize contention the stream pipe buffer area should be placed on a dedicated set of storage whenever possible.

To further clarify this, there is a regular savepoint where the SAP HANA data portions are written down from the memory into the persistent storage. The savepoint is not the backup discussed here.

A savepoint is first established in the process for the backup but then an independent copy onto an independent storage location (e.g. a tape under the control of NetBackup) is created through the backint mechanism.

SAP HANA overwrites the persistent storage savepoint inside the SAP HANA appliance with the next cycle. The section about log files explains the role they play here.

The backup copy in NetBackup is a real independent recovery point. The backup is kept until it expires (default is infinite retention), is manually expired or called from SAP HANA through backint for expiration.

Backup performance considerations

The backup is a highly efficient, compressed data stream delivered to the backup application. There is generally very little I/O involved in the data backup because of the stream pipes but the stream pipes will get buffered in the shared persistent storage.

It is also worth noting that there are currently no incremental backup or single file restore capabilities in SAP HANA.

SAP HANA does not currently support placing NetBackup media servers on the appliance nodes themselves so there is always a network connection between the appliance nodes at the NetBackup media server. The speed of these network connections will also impact the overall throughput rates.

The configuration of the SAP HANA appliance (single node vs. multi node, shared network vs. dedicated backup network) will also influence the achievable backup throughput.

Running backups

Backups can be initiated from automatically from NetBackup, using a NetBackup SAP backup policy to schedule the start time of the backup, or manually from the SAP HANA Studio User Interface or from the command line using the hdbsql command. This hdbsql command can be configured in a script to run in batch mode, controlled by an external scheduler such as “cron”. Information about how to script hdbsql commands and create batch files can be found in the [SAP HANA Administration Guide](#). An example script that can be used with a NetBackup SAP backup policy together with instructions on how to implement backups scheduled from NetBackup can be found in [TECH209343](#).

Recovering instances

Part of the backup process through backint is the process of passing on a catalog file at the end of the backup. SAP HANA stores a copy of the information about the completed backup, the involved services, the timestamps etc. in a file along with the backup. SAP HANA itself keeps a copy in its own database but can recall the catalog information from NetBackup should the need arise to recovery from a fresh install.

In a recovery SAP HANA shuts down the database and recovers the data and log files in one recovery operation. Once the recovery is running, a recovery cannot be canceled. If an error occurs during a recovery, the complete recovery must be repeated. This is regular operation in SAP HANA and is not related to the backup method (“backup to file” or using backint).

Data Backup

The following pages explain the data backup process first. The section beginning on page 20 explains the log segment backup mechanism in detail.

How the Backint communication works for backup

Figure 5 is a diagram of a SAP HANA multi node environment with NetBackup. The NetBackup Client is installed on each node and contains the Symantec backint implementation. This is like a lock and key concept. NetBackup comes with the key to unlock the available SAP HANA backint functions. A configuration file (.uti file) can be kept per node or as a single central copy in the shared location. The configuration file contains information about which NetBackup Policy and Schedule to use.

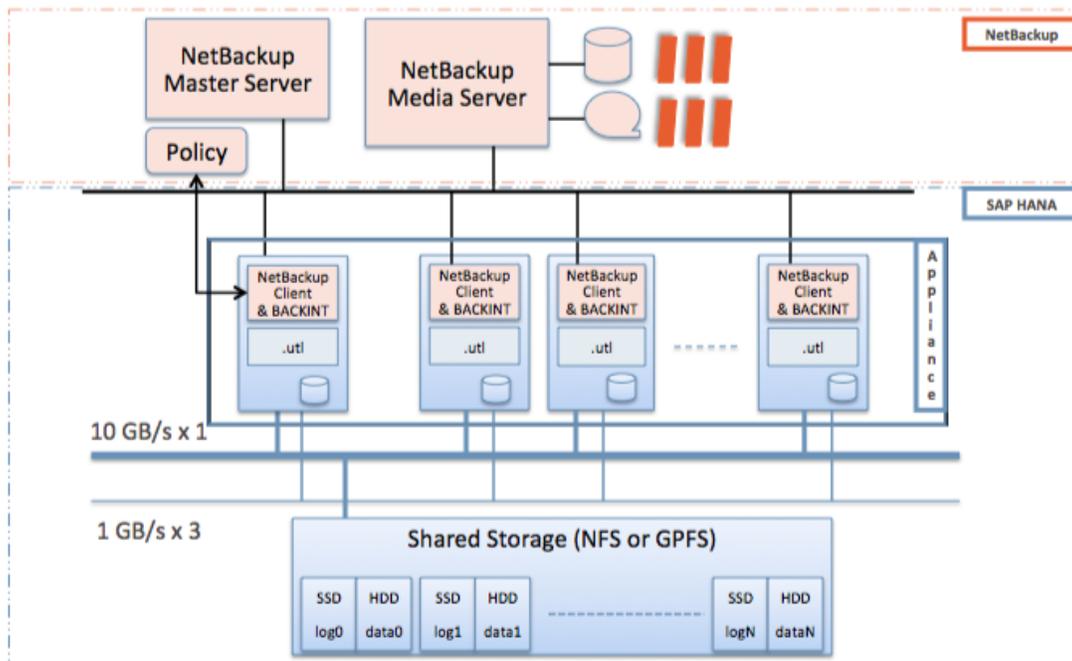


Figure 5 - Overview of the backint implementation with NetBackup

Step1: Figure 6 shows how the backup is initiated through the SAP HANA User Interface (SAP HANA Studio) or through the SAP HANA Command Line via hdbsql. The backint communication is initiated in this phase.

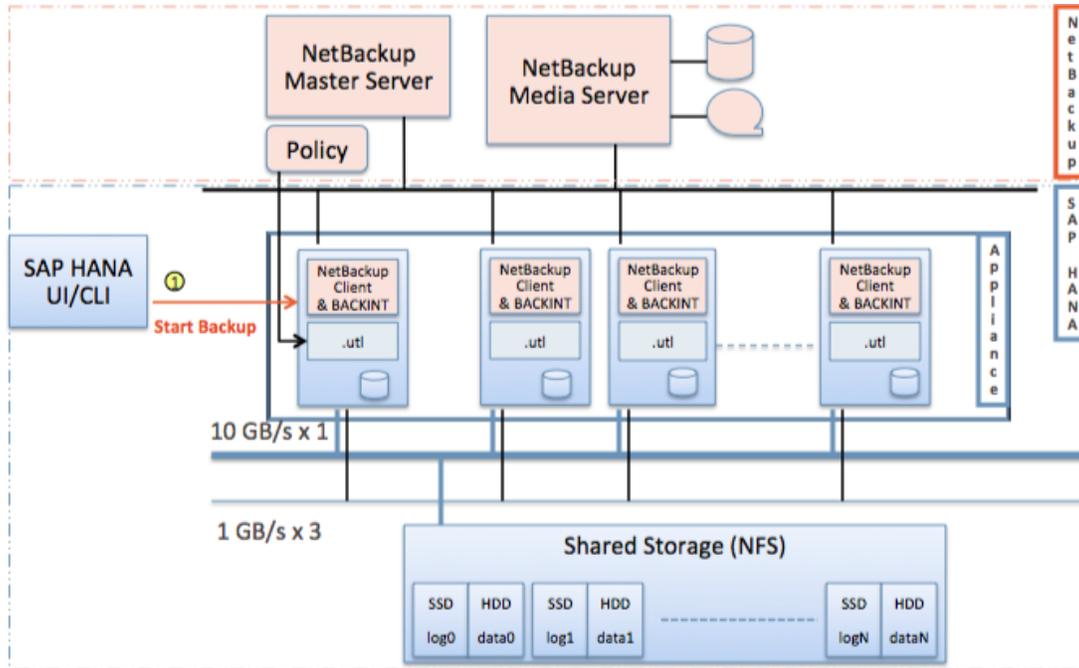


Figure 6 - Data backup initialization

Step 2: SAP HANA brings the database into backup mode, seen in Figure 7.

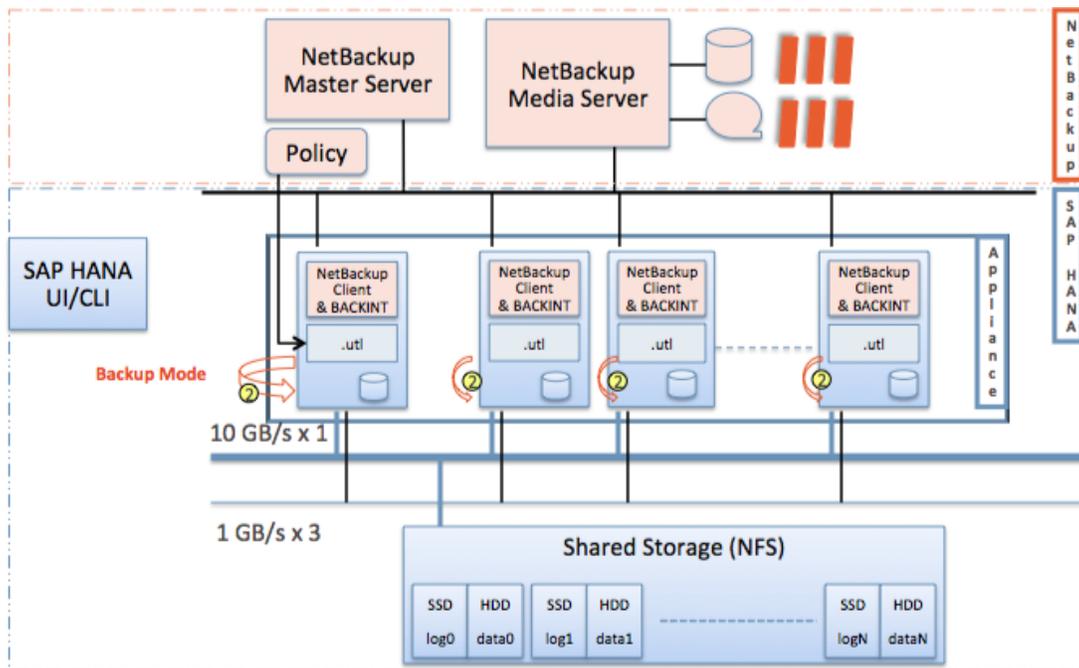


Figure 7 - Establish backup mode

Step 3 and 4: Figure 8 shows the handshake process with NetBackup. SAP HANA will create the backint required .in files on all nodes for communication with NetBackup. The NetBackup Client backint component will read the .utl configuration file for instructions about which NetBackup policy to use for the backup. The .in files contain version information like the backint and SAP HANA revision as well as information where the stream pipes will get created plus name information. At the same time NetBackup allocates the Storage Unit resources (mounts a tape, pre-allocates storage on disk).

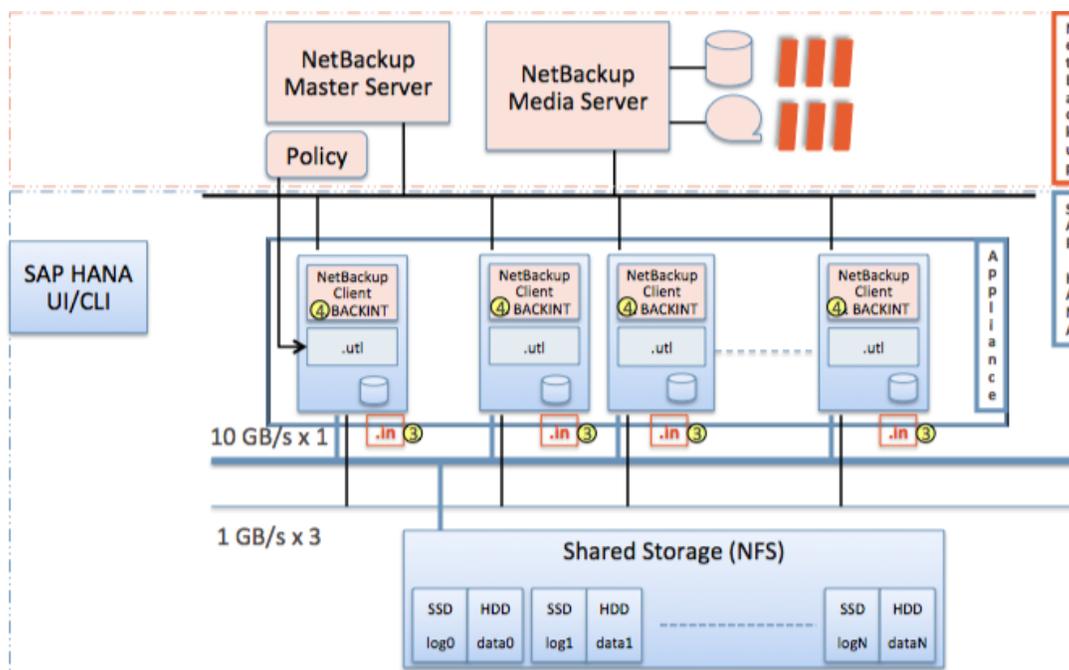


Figure 8 - Handshake with NetBackup

A sample .in file:

```
#SOFTWAREID "backint 1.04" "hdb indexserver 1.00.32"

#TOOLOPTION "Key1:value, Key2:othervalue"

#PIPE

"/usr/sap/BR1/HDB01/backup/data/COMPLETE_DATA_BACKUP_databackup_2_1"
```

Step 5: Figure 9 shows how the data pipes are created. In the process the backint creates 1 stream per pipe defined in .in file. Each service will have its own stream pipe e.g. one for the Index Server, one for the Name Server, one for the Statistics Server and one for the XS Engine. Through this the data backup can be executed in parallel. While the Name Server, the Statistics Server and the XS Server usually do

not contain a lot of data and finish quickly the Index Server holds the bulk of information and will run longest as a job.

NetBackup allows the user to control how many backup jobs per Client / SAP HANA Node are run in parallel. It is also possible to define how many nodes in parallel should run active backups (e.g. only 2 nodes at a time out of 4 nodes should be sending data in parallel). Other means of controlling the throughput are restricting the allowed usable bandwidth, limiting the available resources (e.g. number of tape drives allowed for use) or putting limits on disk I/O throughput for back to disk storage units in NetBackup.

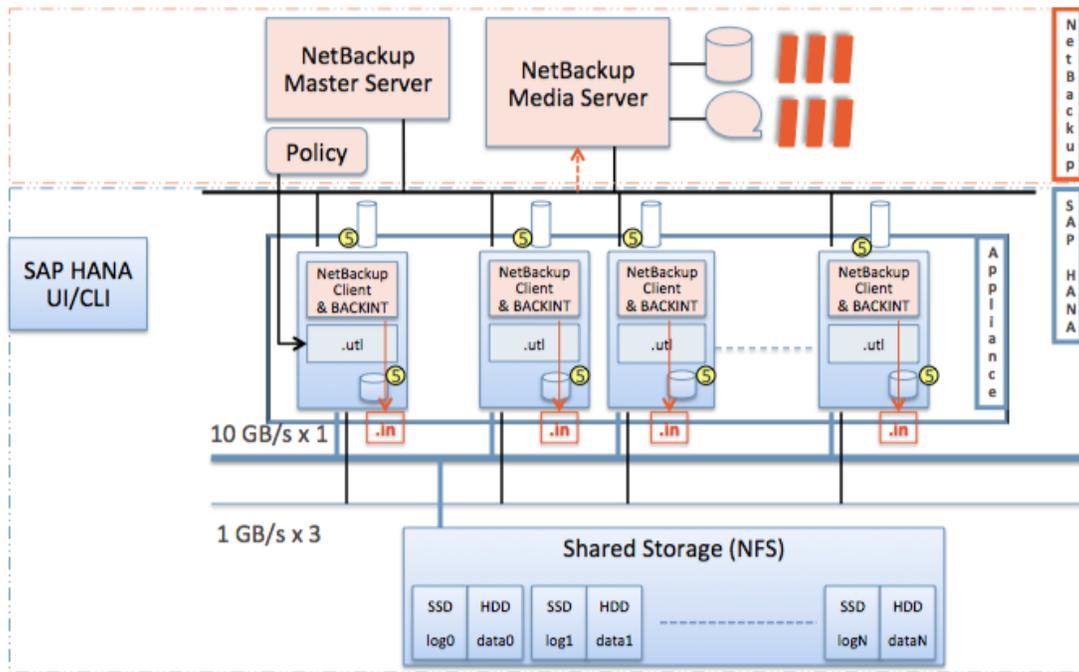


Figure 9 - Pipe generation

Step 6: The NetBackup Client backint component passes the data through into NetBackup at maximum available speeds (see Figure 10). NetBackup Media Servers are I/O machines, which are tuned for performance. The data is written to the NetBackup Storage Unit location that was defined in the backup policy – Disk or Tape. The datastream sent by SAP HANA is highly compressed and optimized. NetBackup can be used to additionally encrypt the datastream at the client, the media server or on the tape drive.

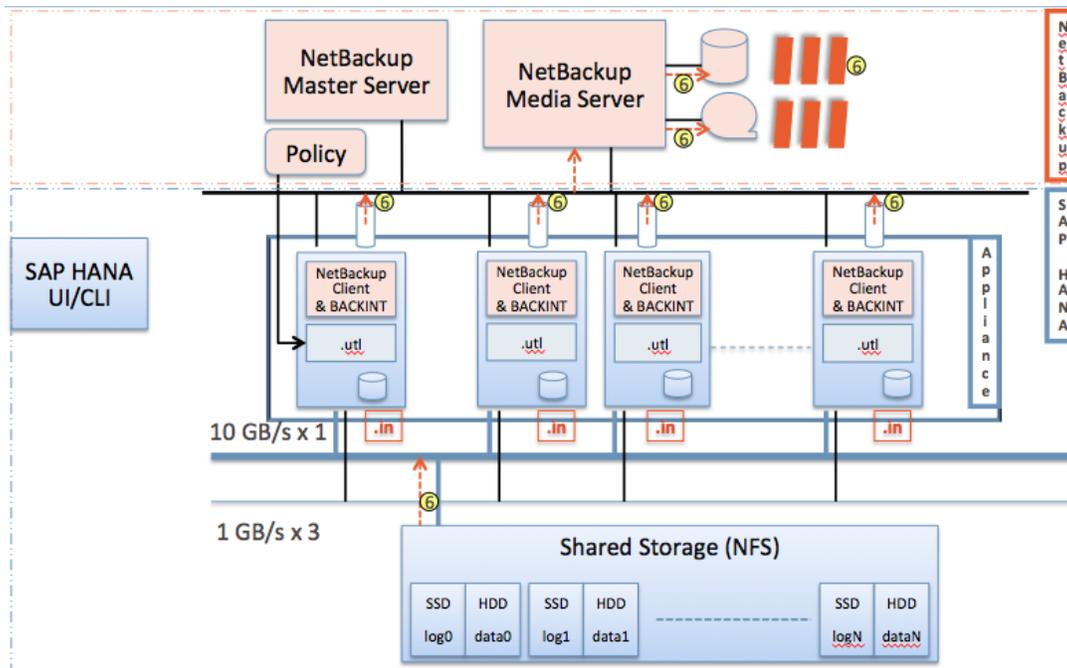


Figure 10 - Actual data backup

Step 7: The SAP HANA catalog file is also backed up (Figure 11). This contains information about all pieces that belong to this backup.

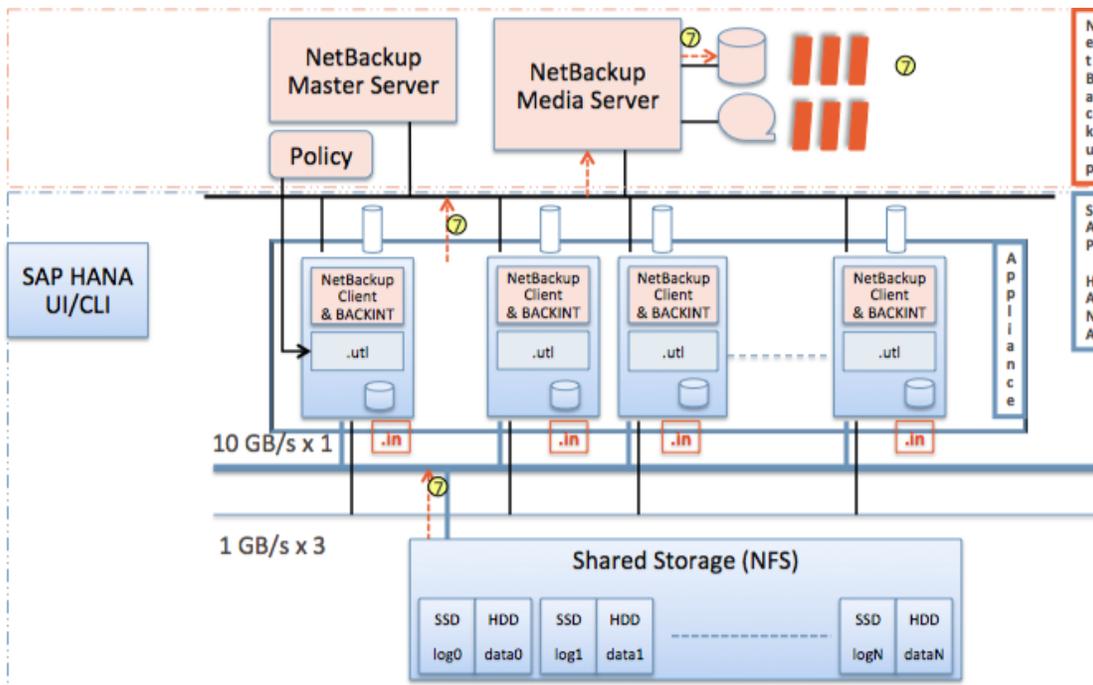


Figure 11 - SAP HANA catalog file backup

Step 8: Figure 12 shows the final step in the backup process. The NetBackup Client backint communicates the backup status back to the .out file. The .out file is used for error handling as well as exchanging relevant information about NetBackup backup id's and NetBackup image id's. This information is then also made available in SAP HANA Studio in the backup log files.

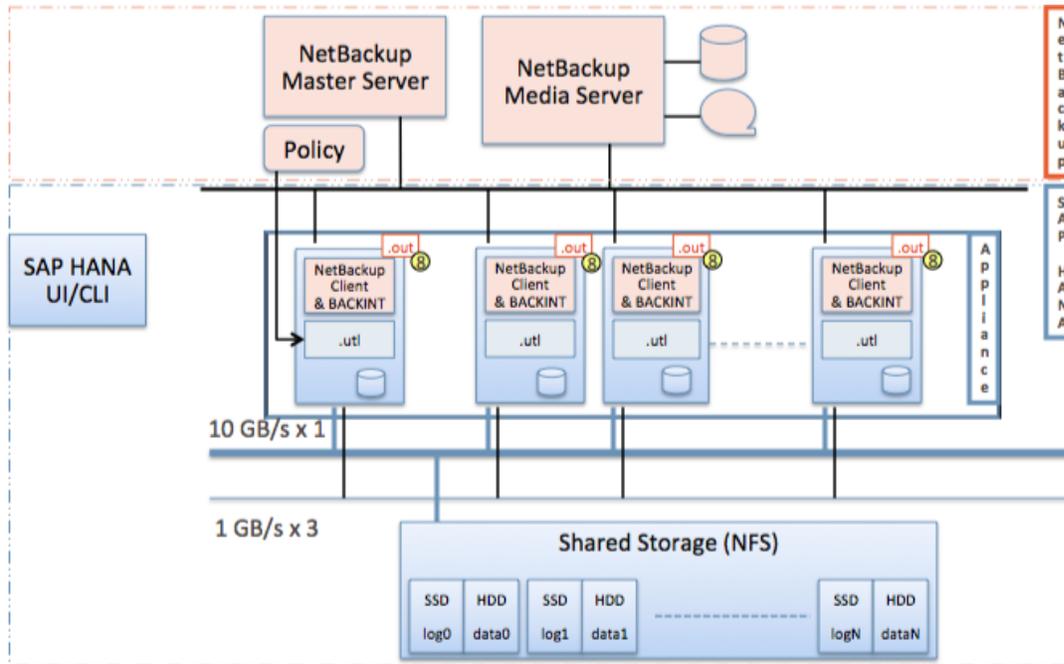


Figure 12 - Completing the backup

How a Backint backup is performed in SAP HANA Studio

To backup a database instance using SAP HANA Studio, right click the Instance Name and select “Back Up” from the drop down menu as shown in Figure 13.

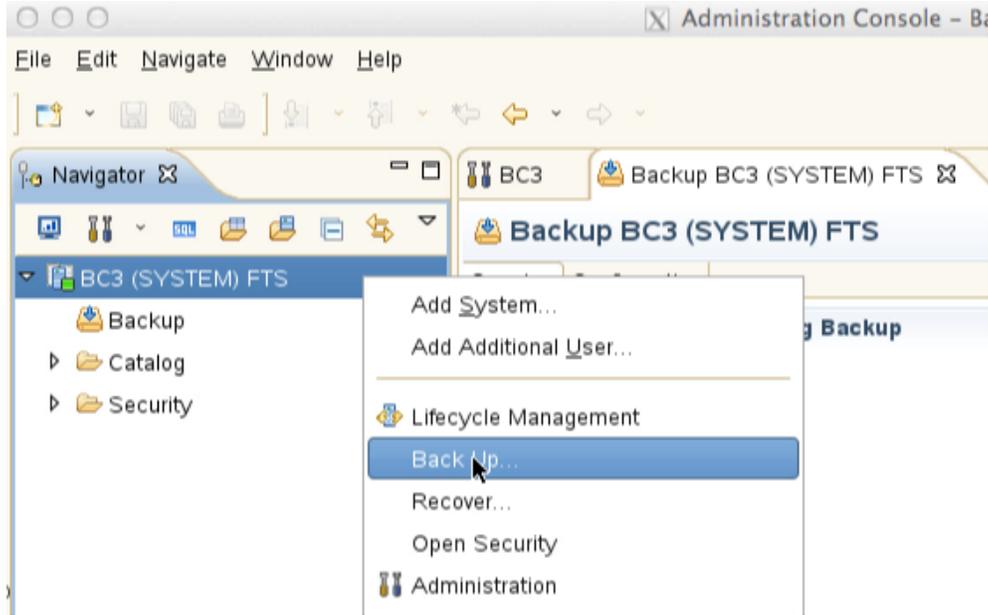


Figure 13 - SAP HANA Studio backup context menu

Select “**Backint**” from the drop down menu for the destination type field (Figure 14).

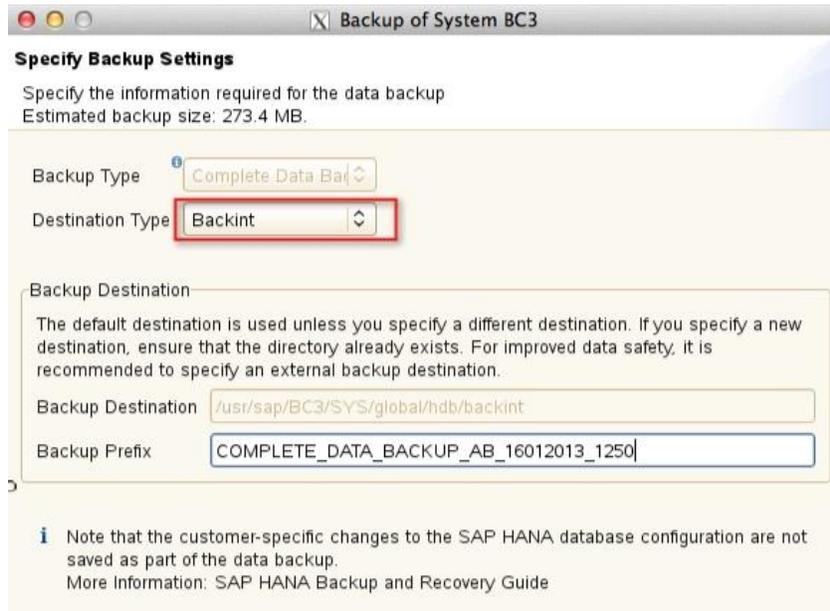


Figure 14 - Backint drop down for Destination Type

Note: Custom text can be added to the backup prefix section, for example a date and time. The date / time value is also stored independently from the backup and is available during the browsing for a restore.

Once the backup is started the progress can be monitored in SAP HANA Studio for each service data stream. SAP HANA initiates the backups. In the NetBackup Activity Monitor the backup jobs can be monitored as well. Through the backint communication the status of the jobs and error handling is dealt with.

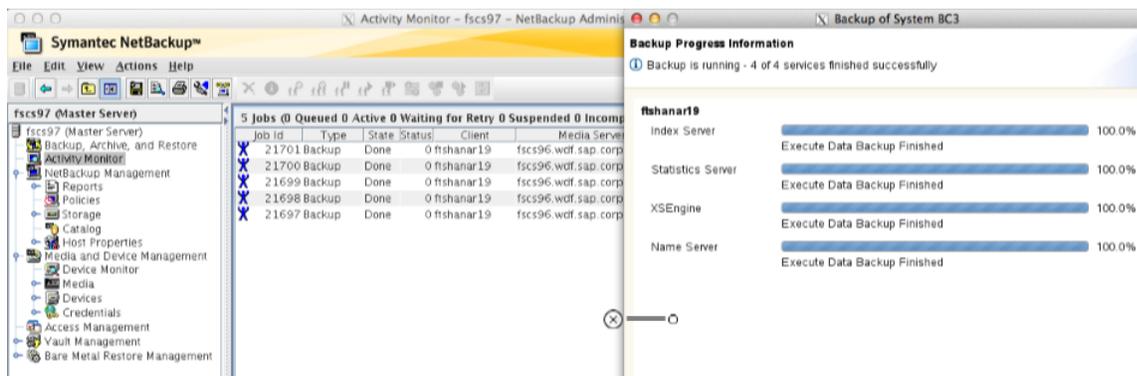


Figure 15 - SAP HANA and NetBackup Activity Monitor view

Each backup is recorded in the SAP HANA Studio backup logs. Once the backup has completed the logs can be opened and viewed at any time in the backup log section in SAP HANA Studio.

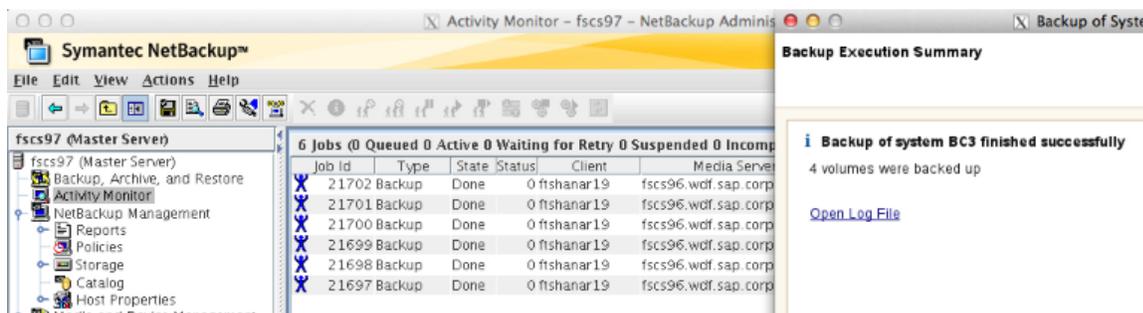


Figure 16 - Backup Success Report and Log availability in SAP HANA Studio

The SAP HANA Administrator requires no special knowledge about NetBackup to run a backup. No extra training is involved and management tasks are reduced. The SAP HANA administrator finds relevant information about backups in the native toolset.

However NetBackup is much more than just an unintelligent storage dump location with its complementing offerings for encryption and backup data handling. Automated Image Replication allows placing a backup copy to an independent DR site for immediate restore availability.

Log Backup

In order to understand the log backup better a deeper look at the log mechanism is necessary.

As previously stated, each service on each node tracks all operations its own log. Figure 17 shows the portion of the local persistent storage layer copy (for restarting the database) and the backup copy on external media to protect against hardware failure or logical errors. But there is more to it to understand.

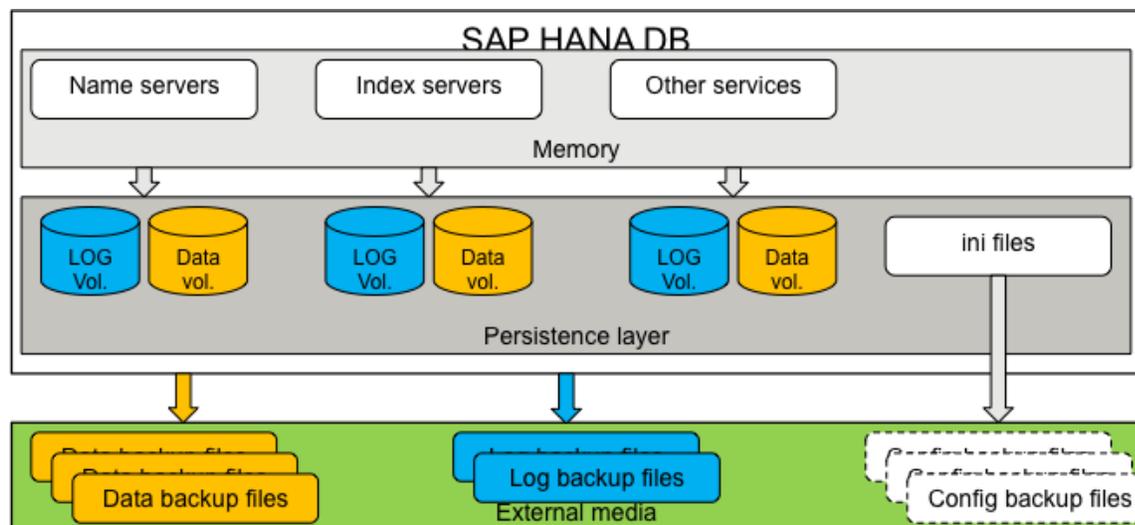


Figure 17 - SAP HANA Single Node Data Structure view

SAP HANA performs all running operations in main memory. A portion of this memory is allocated as a log buffer per service. Log buffers are allocated in 1 MB chunks and rotated. Whenever a log buffer becomes full OR a commit entry is written, the log buffer is written to the assigned log segment in the persistent storage layer.

From there a backup of the services log segments to an independent storage location is required. Log segments in the persistent storage layer can only be overwritten if a backup copy of the log segment exists. This means that the log volume in the persistent storage layer can grow until a disk full condition is hit if no backups of the logs happen. In a log disk full condition the SAP HANA database will freeze.

Note: This is also true if the log volume of just one service on any of the participating nodes will experience a disk full condition.

The backed up log segments remain on disk until they have been released automatically after a savepoint. After the log segments have been released, the oldest unused log segment can be overwritten with new log entries. If there are no unused log segments, new log segments are created.

Note: The amount of log segments on disk is not automatically reduced. A database restart will also not reclaim disk space from old log segments that have been released after a savepoint. The database administrator has to manually issue “alter system reclaim log”.

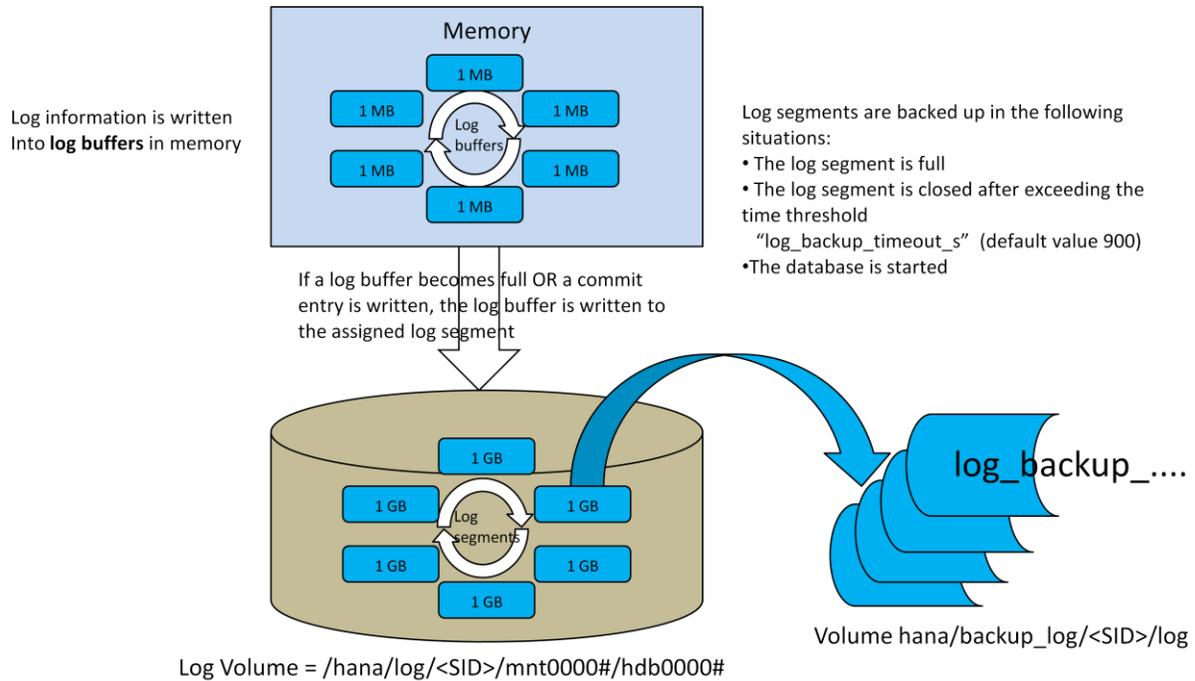


Figure 18 - The memory to persistent storage to backup destination relationship

Points to note:

- An initial data backup has to be completed in order to start the log backup process.
- Log backup is configured in SAP HANA Studio inside the global.ini parameter.
- The parameter “enable_auto_log_backup = yes” is set in the global.ini.

The log_mode setting defines the behavior of the write process from memory to persistent storage. Multiple log_mode setting options are available but for production usage “log_mode=normal” should be used. The log mode setting is automatically switched from “overwrite” to “normal” after the first data backup. Log_mode normal is to be used in productive environments in order to provide support for point-in-time recovery.

The log write process from memory to persistent storage happens on a continuous basis. In log_mode normal a log segment can only be overwritten if a log backup of that segment has happened. Other log modes allow overwriting of log files, even if they have not been backed up which is why those log modes are only applicable for test installations. Whether an automatic log backup will occur is defined by the parameter “enable_auto_log_backup=yes”.

If the “backup to file” method is configured in SAP HANA studio this leads to a steady increase in the amount of backup disk space required for these log backups. It is advised to avoid disk full conditions also for the backup area in any circumstance. When the backup area where the backed up log files are stored fills up (disk full condition) no further log backups are possible. This then leads to a situation where the log segment allocation grows until all the persistent storage space is used. Once this space is used the database will freeze.

The Backint Benefit

The backint integration with NetBackup eliminates the risk of the database freezing due to disk full conditions by leveraging NetBackup’s media management layer to provide effectively unlimited backup storage.

Backint passes the log backups to NetBackup, which then automatically takes care about storage management. NetBackup automatically organizes all background data management aspects and most importantly tracks the location of all log backups in the NetBackup Catalog. SAP HANA has transparent full access to all log backups through the backint, wherever the physical storage location of the required log backups may be.

Once configured and connected through backint NetBackup is invisible to the SAP HANA admin but provides infinite Storage to the log backup process.

All data recovery operations are handled in the background through backint requests to NetBackup.

Log file storage considerations

Unlike the data backup, the log backups are very small and happen frequently. Each service on each node writes its own log. In a recovery situation all logs are needed, the recovery will fail if a single log is missing. Backint has built in check mechanism to first recall the SAP HANA catalog file to ensure all available parts from the data backup and corresponding log files can be retrieved.

Log file backups are best stored on disk. The log files are comparable small and only contain unique data. For best performance in tape out or restore operations a consecutive write order for the logs of a service should be kept. NetBackup Advanced Disk has the concept of pre-allocating the storage in consecutive blocks to avoid a lot of spindle head movement built in.

Disk staging for log backups can be applied if long term storage of log files is required. Within NetBackup disk staging mechanisms are available, space management with high level/low level watermarks, failover mechanisms for storage, I/O Limits for the disk storage etc.

Data Recovery

In the event of data corruption or loss the SAP HANA database can be recovered using data and log backups (and if available log entries from the log volumes). A recovery may be necessary in one of the following situations:

- Data area is unusable e.g. disk crash of the data area.
- Log area is unusable e.g. disk crash of the log area.
- Logical error e.g. requirement to reset the system to certain point in time for special recovery purposes.

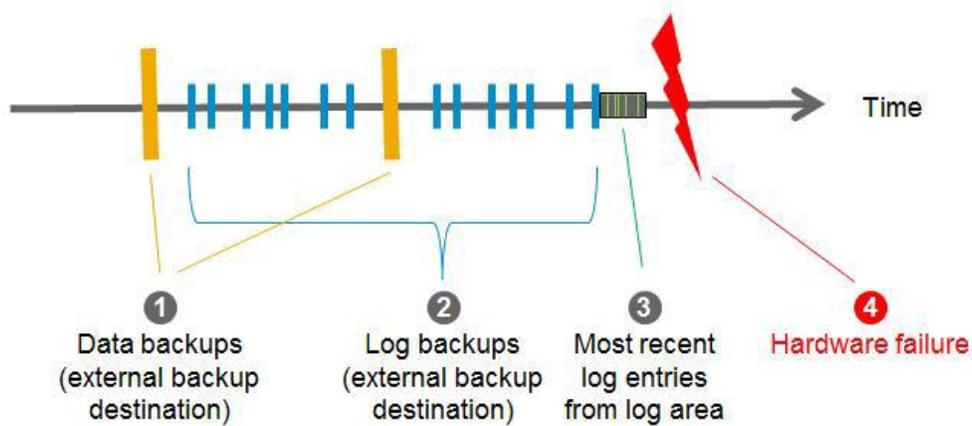


Figure 19 - Backups mapped against a timeline prior to a failure

Additionally backups can be recovered to a different SAP HANA appliance system and used for disaster recovery testing. (This approach can also be used to move SAP HANA instances between production and test/dev environments.)

Recovery Type	Description
Recover the database to its most recent state	<p>This option recovers the database to a state as close as possible to the current time.</p> <p>This recovery option uses the following data:</p> <ul style="list-style-type: none"> • The most recent data backup • Log backups made since the most recent data backup • Log area
Recover the database to the following point in time	<p>This recovery option uses the following data:</p> <ul style="list-style-type: none"> • The last data backup available before

Recovery Type	Description
	<p>the specified point in time</p> <ul style="list-style-type: none"> • Log backups made since the data backup to be used • Log area <p>Note:</p> <p>If a point in time in the future is specify, the effect will be the same as recovering the database to the most recent state.</p>
Recover the database to a specific data backup	<p>This recovery option uses the following data:</p> <ul style="list-style-type: none"> • The specified data backup <p>Note: Log entries are not replayed, either from the log backups or from the log area. All log entries that still exist in the log area are deleted.</p>
Recover the database to the following log position	<p>Note:</p> <p>This recovery type is an advanced option that can be used in exceptional cases if a previous recovery failed.</p> <p>This recovery option uses the following data:</p> <ul style="list-style-type: none"> • The most recent data backup available before the specified log position • Log backups made since the data backup to be used • Log areas

Table 2 - Recovery Types

Figure 20 shows the points on the time line from Figure 19 that equate to:

- (A) Recover the database to a specific data backup.
- (B) Recover the database to a specific point in time.
- (C) Recover the database to its most recent state.

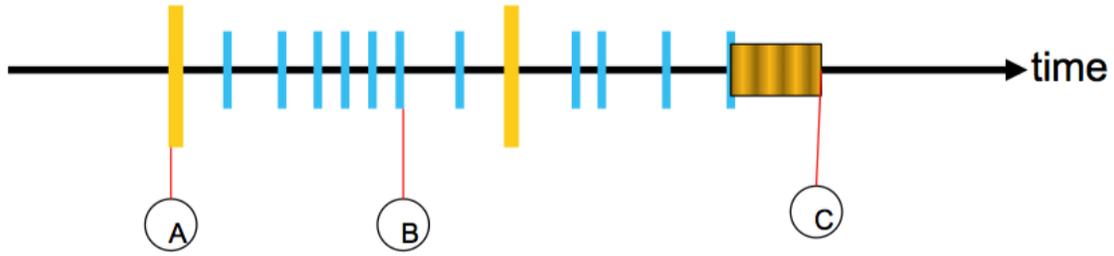


Figure 20 - Recovery points on the backup time line

It is important to note that the SAP HANA recovery process causes the database to be shut down. The system is, therefore, not accessible for end-users or applications while recovery is in progress. This is independent of the use of NetBackup.

Additional notes in regards to Recovery:

- At the beginning of the recovery, all data and log backups to be recovered must be available. The backint implementation with NetBackup automatically cares about this aspect.
- To recover the database, the database needs to shut down.
- If recovery fails, the complete recovery needs to be repeated.
- The SAP HANA database software version used during the recovery, must be the same or higher than the version used to create the backup.
- The restore environment has to have the same setup in regards to the amount of nodes/services.

How Backint Communication Works For Restore

Step1: Figure 21 shows how the restore is initiated through the SAP HANA Studio user interface. The restore must be performed by the <sid>adm user (the database administrator account which is automatically generated when the instance is created). SAP HANA shuts down the database in the very first step. As an alternative the restore can also be initiated via command line through hbdsq.

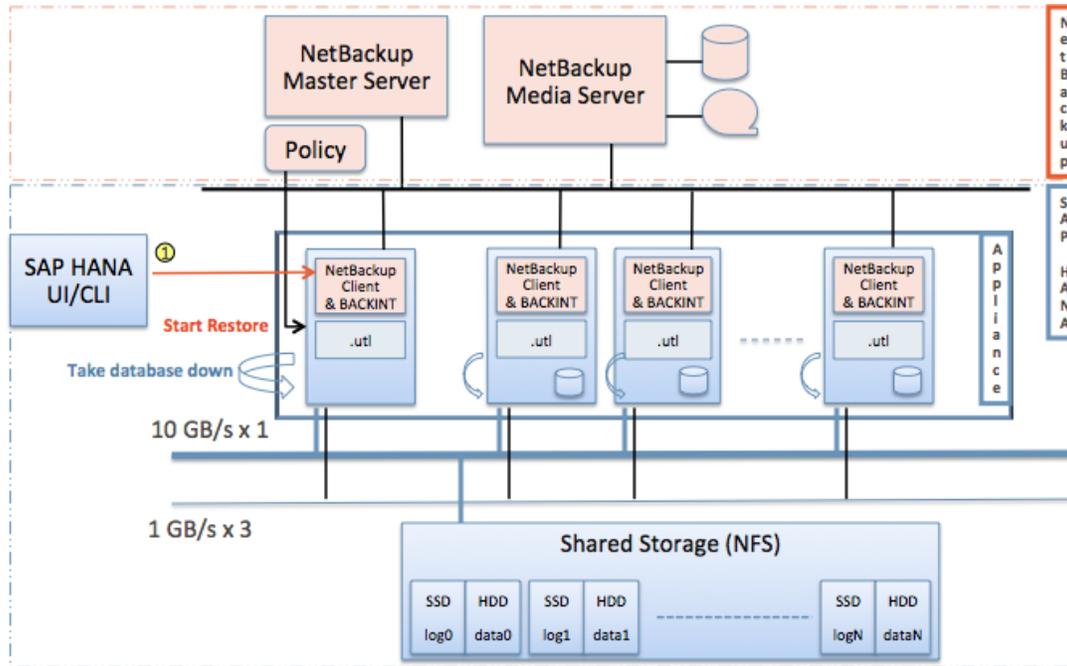


Figure 21 - Restore initialization

Step2 & 3: The backint communication is established, shown in Figure 22. SAP HANA will create one or more .in files on all nodes. Each .in file is used to coordinate the list of backup IDs and corresponding pipes in the next steps. The backint processes on all nodes are started.

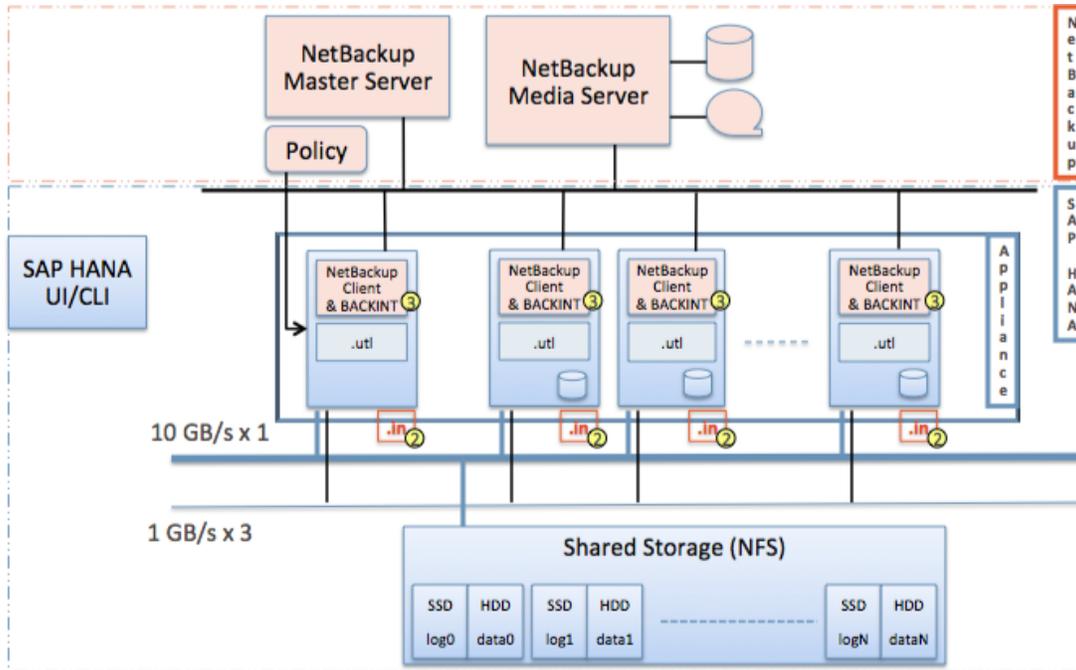


Figure 22 - Handshake with NetBackup

Step 4: Figure 23 shows how the backint requests the latest SAP HANA catalog file from NetBackup. SAP HANA performs a cross check function and validation operation to compare the SAP HANA database information with the backup information. Should backups in NetBackup no longer be available this would be detected and only available backups are listed for selection for restore in the SAP HANA Studio user interface.

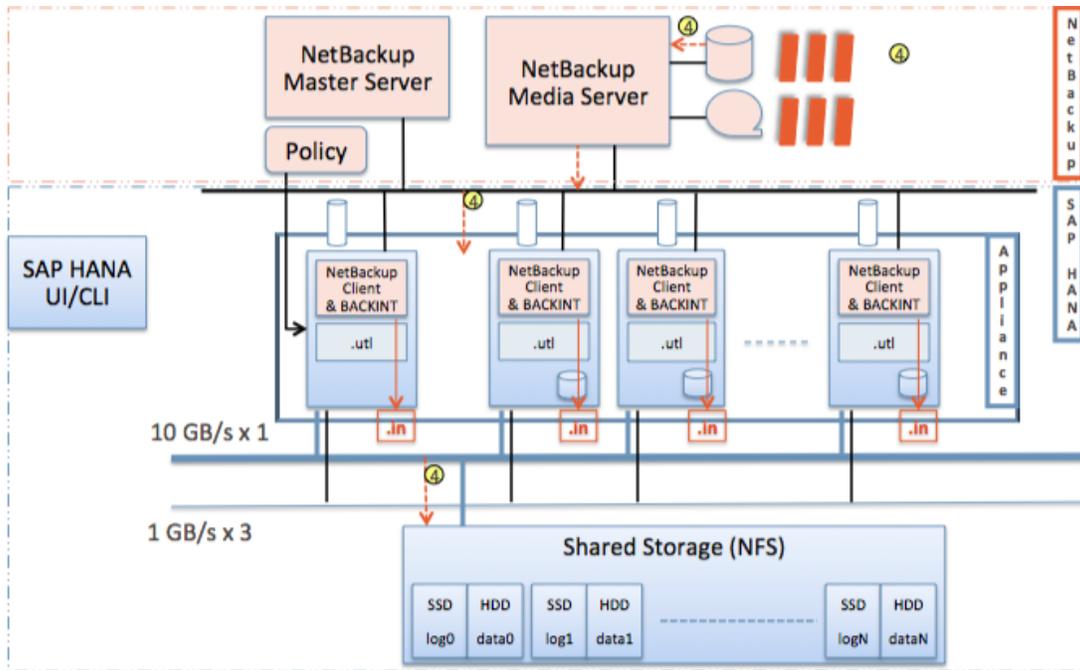


Figure 23 - SAP HANA catalog file inquiry

Step 5: The backint establishes the pipe communication and uses the .in file for placing the backup image ID information which it had derived from the catalog (see Figure 24). Backint then requests the data of the backup images from the NetBackup Master.

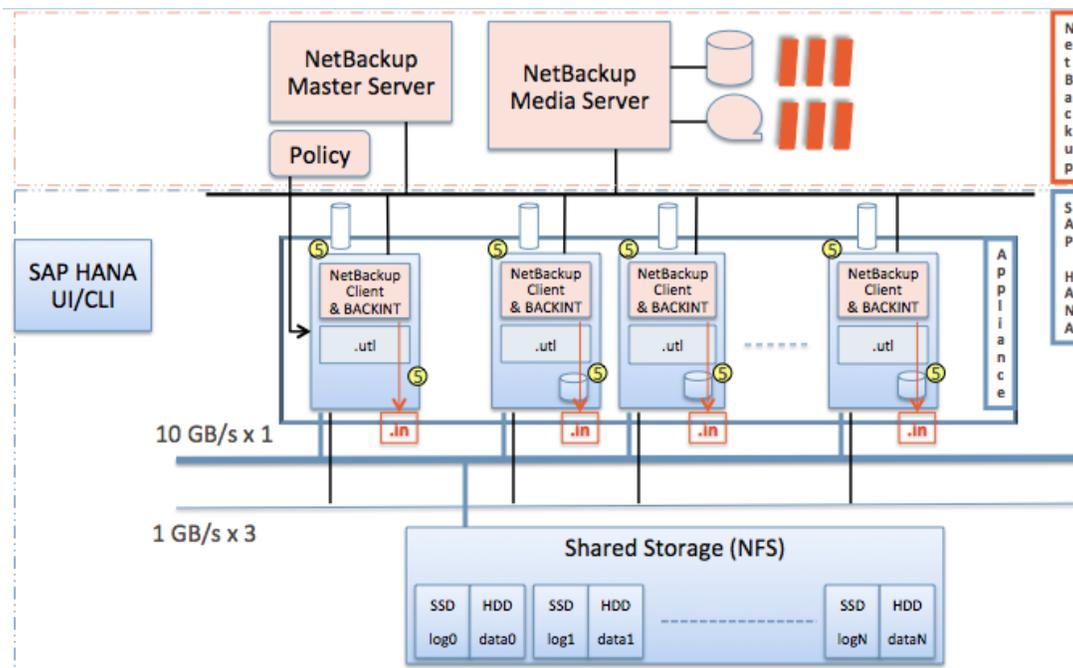


Figure 24 - The restore request process

Step 6: Figure 25 shows how the actual data restore happens. The restore data, which is transferred from the media servers, will be pushed into the specified pipe. The .ini file coordinates the activities as each pipe is associated with a service. This also includes the recovery to the spare node in a multi-node environment in a double failure scenario (one node failed and the failure resulted in corruption of the database).

It's important to note that SAP HANA has several conditions that need to be met:

- The restore environment has to have the same setup in regards to the amount of nodes.
- The instance name cannot be changed and the SID needs to be the same.
- If the configuration files have been modified (which are not part of the regular SAP HANA data and log backups) from the standard these modifications must be manually applied following a re-install before the recovery can be attempted.
- The SAP HANA database software version used during the recovery, must be the same or higher than the version used to create the backup.

These conditions exist regardless of whether “backup to file” or backint is used and are not related to NetBackup.

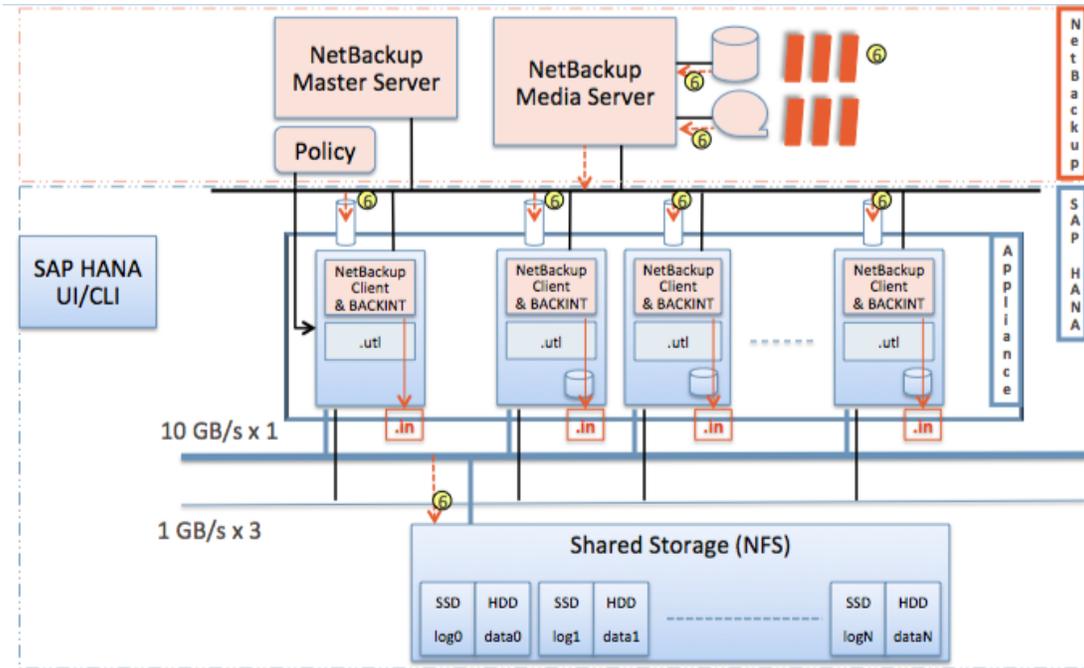


Figure 25 - Actual data restore

Step 7: At the end of the data transfer (depending on the recovery scenario chosen the restore did include data and log files) an .out file is generated. This can be seen below in Figure 26. Each backint per node will have 1 line per image id from .in file listed in the .out mentioning success or failure for proper error handling and communication with SAP HANA. This information is automatically transferred into the backup logs and available from within the SAP HANA Studio for review.

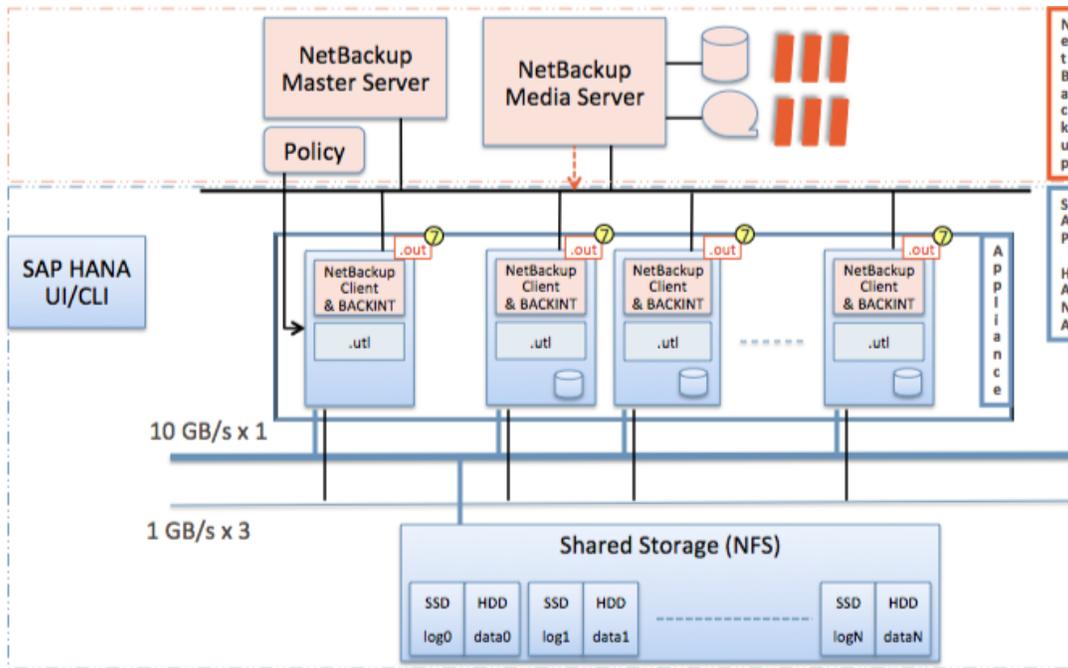


Figure 26 - Finalizing the restore

How a recovery via Backint is performed in SAP HANA Studio

Right click the Instance Name and select “Recover” from the drop down menu.

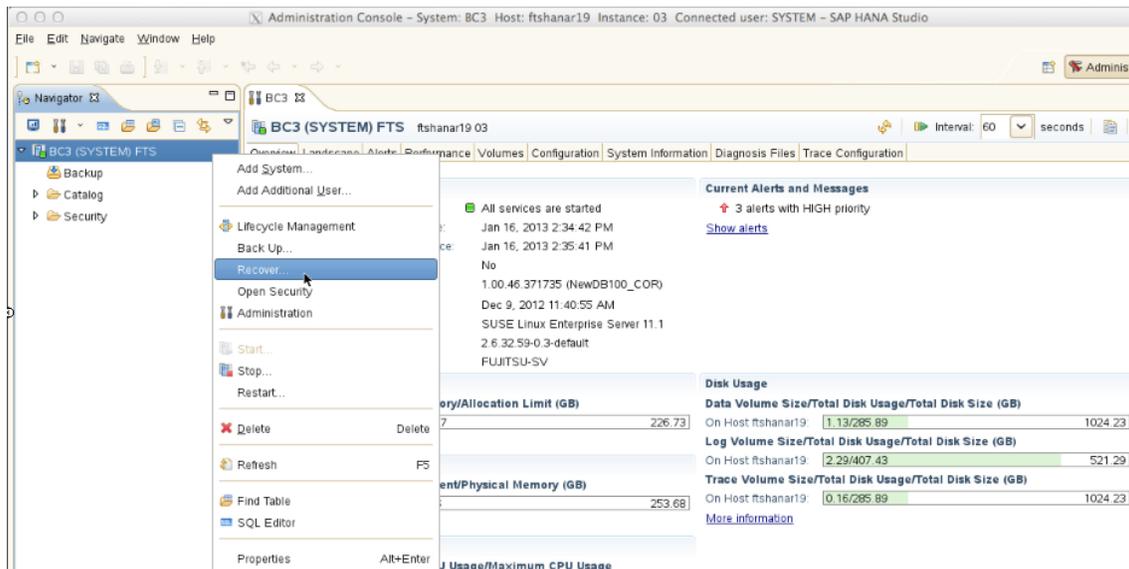


Figure 27 - SAP HANA Studio recovery context menu

Select the method for recovery, e.g. to the latest Point in Time. Select the backup to be recovered.

Once the process is started SAP HANA restores all necessary data and log backups and recovers the database.

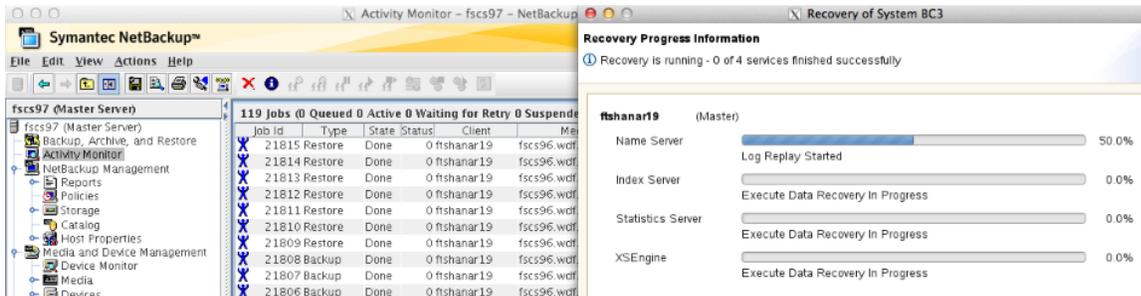


Figure 28 - SAP HANA and NetBackup Activity Monitor recovery progress view

Starting the services as part of the recovery operation automatically triggers a new log backup for the service.

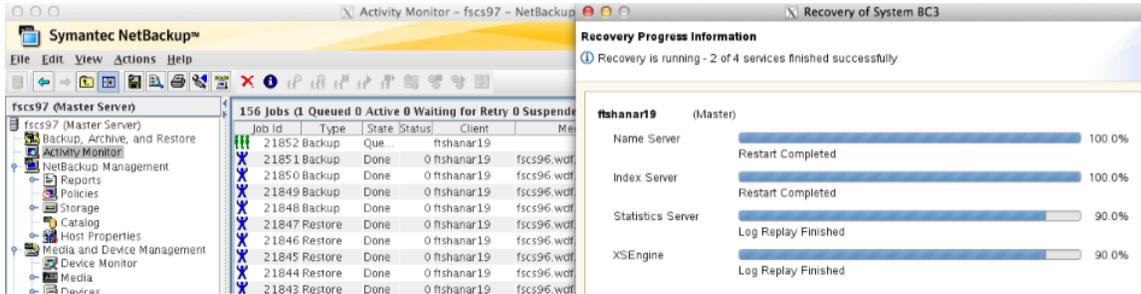


Figure 29 - Service restart interaction

A recovery is also recorded in the SAP HANA Studio backup logs and can be viewed once the recovery is complete.

APPENDIX A – Other useful information

Backup User and Privileges

During installation of SAP HANA the database user SYSTEM is created. For security reasons SAP recommends creating a dedicated user for backup purposes and not to use the SYSTEM or <sid>adm user.

The backup user needs to have following system privileges:

Privilege	Explanation
Backup Admin	Authorization to perform backups
Catalog Read	Authorization to collect the information needed by the backup wizard

Relevant Log Modes and additional switches

The log_mode parameters are located in the global.ini under persistence.

log_mode = normal is recommended to provide support for point-in-time recovery.

log_mode = overwrite

Log segments are freed by savepoints and no log backup is performed. This can be useful for test installations that do not need to be backed up or recovered.

Note: log_mode = overwrite is not recommended for production systems because no point-in-time recovery is possible. For recovery, only data backups are used; the logs must be disregarded.

enable_auto_log_backup

This parameter enables or disables automatic log backup.

Default: enable_auto_log_backup = yes

Note: In the default log_mode normal, if automatic log backup is disabled, the log area grows until the file system is full. At that stage, the database will freeze.

Data backup size estimation

The size of a data backup can be retrieved from following locations.

- Using sqlhdb:

Sum in bytes for all services

```
select sum(allocated_page_size) from M_CONVERTER_STATISTICS;
```

Example:

```
SUM(ALLOCATED_PAGE_SIZE)
```

31917608960

List of volumes with size of each volume in bytes

```
select volume_id, sum(allocated_page_size) from  
M_CONVERTER_STATISTICS group by volume_id;
```

Example:

```
VOLUME_ID,SUM(ALLOCATED_PAGE_SIZE)  
5,7780433920  
2,19243372544  
3,1215135744  
4,68943872  
6,3609788416
```

Note: The first digit is the volume ID and the value after the comma separator is the allocated page size in bytes is listed.

- At the beginning of a SAP HANA Studio backup:

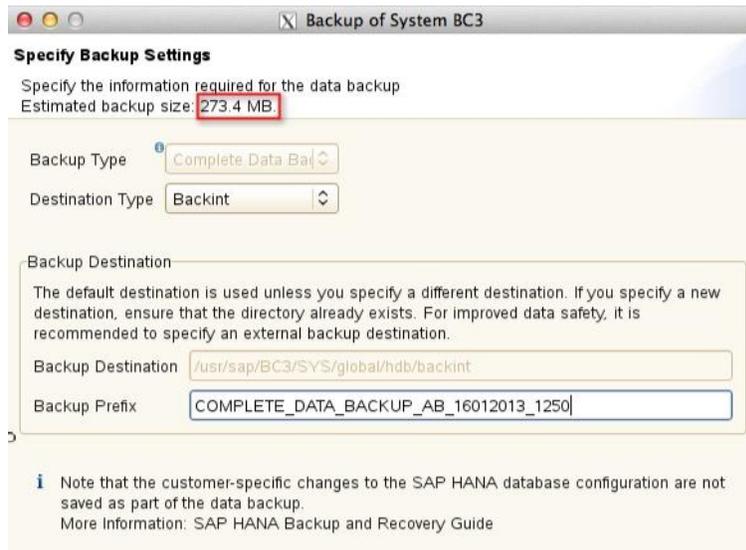


Figure 30 – Estimated backup size

About Symantec:

Symantec is a global leader in providing security, storage, and systems management solutions to help consumers and organizations secure and manage their information-driven world. Our software and services protect against more risks at more points, more completely and efficiently, enabling confidence wherever information is used or stored.

Headquartered in Mountain View, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site: www.symantec.com

Symantec Corporation
World Headquarters
350 Ellis Street
Mountain View, CA 94043 USA
+1 (650) 527 8000
+1 (800) 721 3934

Copyright © 2013 Symantec Corporation. All rights reserved. Symantec and the Symantec logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.