Exchanging the central Storage System during Operations

Bernd Holzhauer[1] and Dr. Osvaldo L. Peinado.[2]
DLR – GSOC, Muenchner Strasse 20, 82334 Wessling, Germany

The ISS Columbus Ground Segment is a complex system with several subsystems using a central Storage Area Network (SAN). The Project operates 24/7. Therefore the migration from old to new hardware had to be performed without any interruption to operations.

The migration became necessary because the subsystems' software has to be updated together with an update of the operating system. With the complexity of the old SAN that would only have been possible with 4 to 6 week interruption to ongoing operations.

To avoid such a situation in the future, it was decided to replace the system completely and to use the K.I.S.S. approach (Keep It Simple and Stupid) for the new system design. The Columbus system uses a variety of different applications running on different operating system versions. Projecting this into the future, it is very important to be able to update the storage system while maintaining the subsystems independently.

Nomenclature

GSOC = German Space Operation Center
Col-CC = Columbus Control Center
SAN = Storage Area Network
SANng = New installed SAN
HSM-SAN = Old replaced SAN
Interims SAN = Intermediate SAN system for tests

[1] SAN System Engineer, Telespazio Deutschland GmbH, c/o DLR – GSOC, Muenchner Str. 20, 82334 Wessling, Germany, no AIAA Member.
[2] Ground Operations Manager, DLR – GSOC, Muenchner Str. 20, 82334 Wessling, Germany, no AIAA Member

I. Introduction

Lifetime in the world of computers is quite different to the understanding of time in space industries. The hardware components of the SAN system in COL-CC were running out of service time much sooner than expected. Also with the increasing amount of data aquired during real operations the SAN became more and more unstable.

In addition a planned upgrade of application software was impractical because this required a newer version of the Linux operating system to SLES 10. But the SAN software drivers omitted the upgrade from SLES 8 to SLES 10 on the application servers.

All together combined with a complex system design of the old HSM-SAN lead to a deadlock situation. The final result of the investigation was to design and replace the old SAN with a new one and to migrate the total Columbus Control Center to this new infrastructure.

The storage network in the Columbus Control Center (Col-CC) is an infrastructure type component. All major subsystems are based on it. So the SAN system could not easily be switched off and replaced offline. Instead it should be exchanged with a minimum, ideally no interruption to Col-CC services.

Defining, planning and testing the new SAN (called SANng) was major undertaking. The experience made with the "old SAN" redefined some major requirements.

A. Columbus Control Center – Col-CC

The Columbus Control Center is a part of the German Space Operation Center (GSOC) in Oberpfaffenhofen (near Munich). Col-CC is the European central interface to the Columbus module connected to the ISS (International Space Station). All local European User Operation Control Centers (USOC) controlling experiments in the Columbus space lab are routing their information through Col-CC. Col-CC is operating 365 days a year 24 hours a day. So there is almost no time to replace a basic infrastructure style system like the SAN as a central storage system. All servers in the Columbus ground segment are connected to the SAN and run on it.

If this SAN fails, the control center will go down until the SAN is restored for operations. The old Columbus HSM-SAN caused that some times by severe failures. This and the obsolescence of the installed SAN components drove an exchange to a more practical and stable solution/system.

Col-CC is divided into several Subsystems. Some of the Subsystems are also split into independent instances for Operations (Ops), Simulation (Sim) and Test (Tst). Here is a list of the Col-CC subsystems and instances which have data (file systems) on the SAN and therefore need access to it.

- DaSS – Data Services Subsystem
  Instances: Ops – Sim – Tst
- MCS – Monitoring and Control Subsystem
  Instances: Ops – Sim – Tst
- OST – Operating Support Tools
  Instances: Ops – Sim
- IMS – Integrated Management Subsystem
  Instances: Prime – Backup
- VIDS – Video Subsystem
  single instance only
- VOCS – Voice Conferencing System
  single instance only
- Infra CM – Infrastructure and Configuration Management

The subsystems and instances shall be independent from each other as much as possible. Cross access between subsystems and instances are not allowed whereas sharing files within a subsystem and instance is mandatory.

II. The "old" HSM SAN

The Columbus ground segment was designed as a complex computer system based on a SAN using a Hierarchal Storage Management (HSM) approach. This was designed and implemented during Columbus mission planning. Due to a large delay in the Columbus launch the system became obsolete quite early after the Columbus module was docked at the space station.

Each subsystem had its own file system and was separated from each other. Data on disk was shared between servers of the same subsystem.

B. Three Tiered Storage

The original designed Columbus SAN was designed as a Three Tiered Storage (also called HSM SAN). The basic idea was to have an almost unlimited SAN where the data is written to a low capacity but fast and expensive Fibre storage. For more capacity the data was moved after a while to a cheaper ATA storage and afterwards moved to a much less expensive but large tape robotic system.

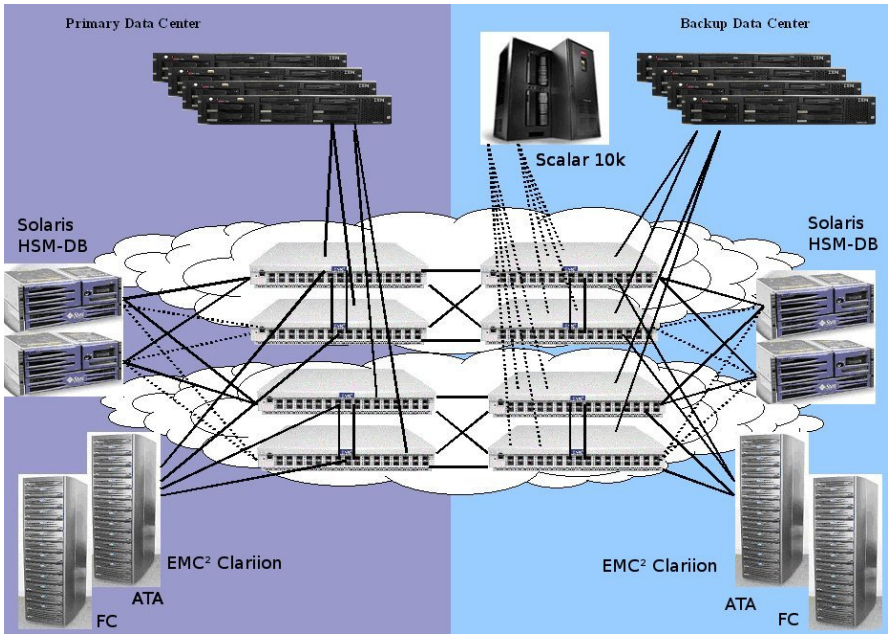

Figure 1 HSM SAN Overview

The servers in the HSM-SAN structure where operating on an EMC Clariion with low capacity but fast Fibre Channel disks. After 10 days of no access the data was moved to an ATA-Clariion (larger capacity with less expensive drives). If data was not accessed more then 180 days the files where moved to a tape robotic system with almost unlimited capacity.

A database system (the HSM-DB), running on Solaris servers, managed the positions and automatic migrations of the files. Due to the separation into subsystems and their instances 14 different file systems were in use. The vendor recommended a maximum of seven (7) file systems per HSM-cluster which caused the installation of two Solaris HSM-DB clusters.

The HSM-filesystem was actually a client/server system. The HSM-server resides as a database in the Solaris systems and communicated to HSM-client software installed on each application server system. This HSM-client driver software existed as a layer between Fibre Channel adapter card and the file system layer of the operating system. These drivers supported SuSE Linux Enterprise (SLES 8) and MS Windows 2000 plus Windows 2003, but there was no support for SLES 10 nor for other Windows versions. For each file access the client needed to ask via a separate HSM-network to the HSM-DB on Solaris where the data exists and then load it respectively.

The Solaris systems were setup as redundant clusters to avoid single points of failures, but during operations the

HSM-DB itself showed up as a single point of failure. The database was shared between the two cluster heads and failed over to the remaining one. Before installation and during simulations of the Columbus project, the system was working fine, but as soon as the project was started and loaded the system with a constant data flow some major problems arose.

1. As more files where stored during real operations, the database increased in size and became instable. The database itself shows itself to be a single point of failure for all file systems on this database cluster.
2. Since the HSM was designed as a client/server model upgradeability was very limited. SLES 9 was skipped in subsystem software development. Software should be upgraded to SLES 10 versions. But no HSM-client/server version was available which could handle SLES 8 and SLES 10 clients at the same time. An upgrade procedure worked out with the vendor of the HSM system would have caused a 4 to 6 week downtime of the complete Col-CC.
3. Also accessing multiple files which are stored on tape only caused the HSM database to crash. This caused all subsystems on the HSM-cluster to fail.
4. The Clariion storage systems became end of life and the service contract was not continued by the vendor.

All these together led to the decision to replace this old dinosaurian by a more modern and more flexible system.

III. K.I.S.S. – Keep It Simple an Stupid

Since one of the major concerns – this finally brought the awareness: "the system is not upgradeable online" – was the complexity of the old system and the client/server construct, the new system should avoid this. So the slogan "K.I.S.S. – keep it simple and stupid" was born. A simple system is easier to maintain. Also the components should be as much as possible independent from each other to make them exchangeable in small chunks if some parts become obsolete in the future.

Due to "lessons learned" and this K.I.S.S. approach new requirements where defined. The new SAN should avoid major problems, as there are:

- update problems
  each component of SAN environment, incl. SAN attached servers should be replaceable without influence to other items. This results in "no special drivers should be necessary" to become independent from operating system versions. This finally ends in
- no more special (vendor proprietary) software at application servers
  About 90% of the servers are running on Linux (SLES 8 and SLES 10) and the rest uses MS Windows. As standard protocols NFS for Linux and CIFS for MS Windows shares should be used.
- single point of failure
  System should not have a hidden single point of failure

But there are also some points which needed to be fulfilled like

C. File Sharing between Servers in a Subsystem

Some subsystems are using multiple servers to store and process the data. Therefore the system must provide file sharing. A simple SAN will not do this without vendor specific software (which should not be used).

But file sharing can be supported by using standard network protocols like NFS and/or CIFS.

The deadlock situation in the old system was mainly caused by the special driver software handling the shared access to the SAN disks. Since operating systems and software will be updated quite often in the future, there is an absolute "no-go" for special driver software from a third party vendor.

A solution was searched providing file sharing between servers with standard operating system drivers.

Network protocols like NFS for UNIX/Linux systems and CIFS for the MS Windows world can handle such file shares. A set of file servers or a NAS (Network Attached Storage) can provide this.

D. Use "out of the box" Hard- and Software

COTS (Common Of The Shelf) products should be used as much as possible. This keeps investment cost low. Also COTS products are usually easy to replace after end of product life time. The more complex components should be supported until end of mission (2018 during planning period – currently 2020) to avoid end of life time

and other data migrations before end of project. If ISS and Columbus project will be extend beyond 2020 an upgrade path must exist.

IV. System Concept

With all the above in mind a computer consultant (system house) was met at a computer exhibition. In a lot of following meetings and discussions they understood the special space business requirements and the Col-CC specifications/concerns in particular. They also brought Hitachi Data System (HDS) into the game. As a team we developed the specific Col-CC system concept.

Hitachi offers an appliance called HNAS (for Hitachi Network Attached Storage). This HNAS resides on top of a standard SAN storage and provides amongst other services disk space as NFS and CIFS network shares. This HNAS hardware platform is capable of emulating up to 64 virtual file servers (named EVS = Enterprise Virtual Server) on a single hardware. The HNAS is not just a PC style system. It is specialized in file services and operates in specially designed hardware (FPGAs = Field Programmable Gate Arrays). Therefore it is very fast and (nearly) without problems on the operating system level.

The basic HNAS features are:

- 24 Hours times 365 days operations possible in cluster configuration with automatic failover.
- Independency between subsystems and instances can be guarantied by the EVS concept. EVSs will automatically migrate to the remaining cluster part of HNAS in case of a HNAS failure.
- Split brain between HNAS heads is omitted by using a SMU (System Management Unit) as quorum device.
- Each EVS will have its own file system and disk quotas may be used.
- A failover (at client site) is less than 30 seconds.
- To keep system topology simple, as much as possible storage should be used via EVS, i.e. used on NFS or CIFS exports.
- For special applications, like Oracle RAC or cluster quorum disks, direct LUNs from storage are also available but should be used rarely.

Storage capacity was calculated by the incoming and generated data flow per year (app. 40-50 TB) times the remaining project time (10 Years). Also Hitachi promised to support at least the USP storage towers until end of project time (2018 at time of project design).

V. Financial Concerns

The first system design was made having the idea of keeping all project data over 10 years online. This ended up in a storage system with the capacity of 500 TB (net) online mirrored between Prime Data Center (PDC) and Backup Data Center (BDC). The storage towers itself should be “Enterprise Class” Storage to guaranty the 365 by 24 operation.

From cost perspective this “high end” system was too expensive and alternatives where discussed. The most costly items in the concept where

- storage capacity and
- the capacity license for online mirroring (pretty expensive at Hitachi)

But with less capacity the target of 10 year online data could not be reached. So the alternative was to (re)define a Data Retention Policy.

E. Data Retention Policy (DRP)

Redefining the Data Retention Policy opened the door for significant cost savings. The DRP basically allows now to move data older than 12 month to a tape archive. System log-files do not need to be archived which means they can be deleted after 12 months. This seems to be not far away from old system design, but the process is now completely different and the new system is still “simple”.

As a result from this new definition the total capacity of the storage was stripped down to 55 TB (net). Furthermore the HNAS itself is able to replicate data from prime to backup storage. This replication is almost comparable to a mirroring function except mirroring is done without (visible) time delay whereas replication has a time scheduler behind.

As a result, if replication is done for example on an hourly basis, data on backup storage may be almost aged by one hour until the system fails. This may cause a loss of the last hour data in case of a failure.

Since the used Hitachi USP-VM storage system is an enterprise class storage it is very unlikely this system will fail. But anyhow, even a small possibility of a one hour data loss is not acceptable for the Ops instances. Video, Test and Sim instances are not this critical. So they where defined to run on replicated instead than mirrored storage.

This ends up in mirroring just 10 TB and replicate the remaining 45 TB of storage. It is a little conflict with the K.I.S.S. strategy but the financial concerns forces this.

VI. SANng System Overview

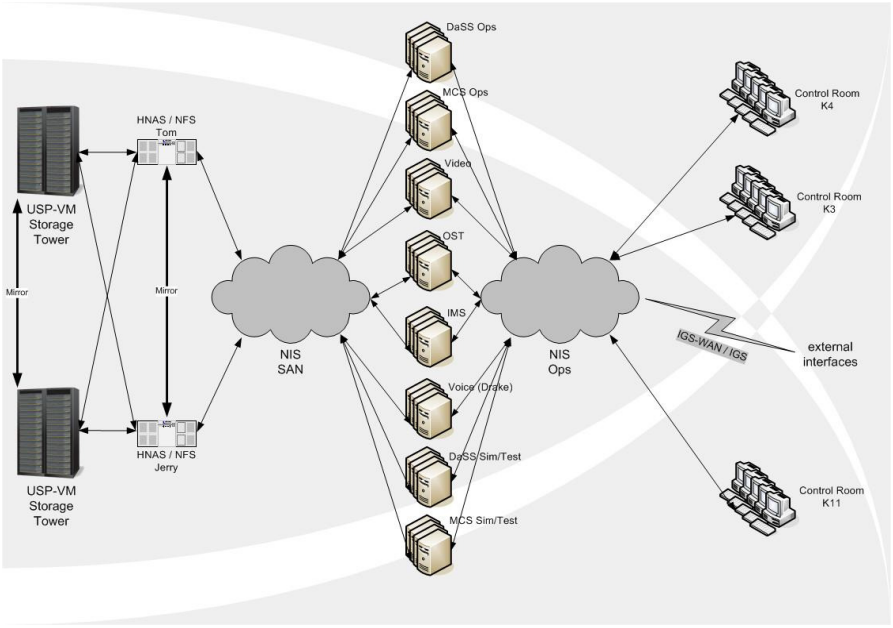

Figure 2 Col-CC SANng Overview

The system consists of two identical USP-VM storage towers, one in Prime Data Center (PDC) and one in Backup Data Center (BDC). The Link between PDC and BDC is an optical cable about 400 meters in length. Storage towers and HNAS systems are connected via redundant Fibre Channel Networks (2 Fibre fabrics). The connection between HNAS and the NIS-SAN cloud is a 10GB redundant link.

The TCP network NIS (Network Infra Structure) is split into two parts. The newly installed storage network NIS SAN (replaces the old HSM-LAN) and the already existing OPS network (NIS Ops). All application servers are connected redundant to both NIS clouds. Two clouds are designed to avoid bottlenecks and cross connections between OPS and Storage Network.

VII. Proof of concept

During design of the new system it was not clear if speed for the data processing would be sufficient. One possible bottleneck could be the data flow through the TCP/IP network instead of using high speed Fibre Channel connections.

Before going deeper into planning and running the complete order processing procedure a proof of concept was requested. Therefore a smaller system, a single AMS 2300 with 20 TB net capacities was borrowed from Hitachi. This system was also equipped with a complete HNAS cluster (later on named Max & Moritz) on top.

The plan was to proof the system under real operation conditions. Therefore some Test and Sim instances should be ported to the borrowed test system. After getting familiar with the usual failover scenarios and other testing scenarios, the system should be used also during real space (operating) simulations (JMST, etc).

This should proof where the total system may run into performance problems and also should verify if and how the subsystem software will react to failovers and movements of EVSs within the HNAS cluster.

F. Interims SAN

During the first tests on this loaned system, the old HSM-SAN became more weak and failed too often. We became familiar and satisfied with the test system and found it much more stable than the old operating HSM-SAN. So the loan system was turned into the "Interim SAN". Which means after getting familiar with the system some OPS instances where migrated to this Interim SAN to drop the stress on the old database based HSM-SAN system.

Since the interims SAN became more and more to be a productive OPS system a backup system was purchased and installed to avoid data losses. The backup system is a single Linux server which can mount all network shares as local drives. The Backup system works like that for the Interim SAN but later on also for the SANng.

The server itself will also control the tape archive in the future. Therefore it is setup as a very powerful system which can easily handle data migrations forward and backward between the two SAN systems.

VIII. Migrating Subsystem by Subsystem to Interims SAN

For data migration from the old to the new SAN storage a Linux server was installed and connected to both SAN systems – the old HSM-SAN and the new Interim SAN. This server had full access to all file systems on both sites. With the Linux RSYNC command it was a quick and easy job to copy the immediately needed data from the old system to the new one.

Then changing the mount points in the appropriate application servers from the old HSM-SAN disks to the new provided network shares and a subsystem was operating (migrated) on the new Interim SAN.

This was a quick and dirty solution to keep Col-CC alive but it stopped the real proof of concept in the middle because the system was now in use for real time OPS and could not be tested seriously anymore.

G. Linux Rsync does the main Data Migration Job

Rsync, a standard UNIX/Linux command for remote file synchronization, was used to do the migration job. It is a very handy tool to copy files from a source to a destination location. By command line parameters it can be forced to copy data only if files do not exit on destination and to copy just newer files. Rsync can also work recursive through whole directory trees and complete file systems.

It does not matter if the initial copy takes hours, weeks or month. Rsync can copy all files during and without disturbing operations. Only files which are currently modified need to be copied again later on. Rsync is able to find such files and synchronize them in a second or third stage. So with each run of Rsync less data will be copied.

For the final move (migration) the application needs to be stopped. The application server must be disconnected from old and connected to the new mount points. Also the final Rsync job has to be executed. But this are tasks which need only a few minutes.

That's all. The final subsystem move (migration) can be done within a 15 to 20 minute LOS slot.

H. Moving during LOS times and coordination with other subsystem maintenance

Copying the data with Rsync could be done during and without affecting operations. More then 99% of the data could be migrated without interruption of services.

The ultimate movement i.e. remounting the shares and the final Rsync of single subsystem or instance was coordinated within other regular subsystem maintenance or LOS slots. So the movement from the old HSM-SAN to the Interim SAN was almost not visible to the Flight Control Team (FCT) – except the total system became more stable/reliable.

IX. Installing the Columbus SAN

Since the Proof of Concept was interrupted by using the Interim SAN for operations, it was resumed later after installing the final SANng and combined with the large Site Acceptance Test (SAT). Real time OPS on the Interim SAN were something like a reverse proof of concept. It showed up the functionality of the system concept and even the "smaller" system never run into performance bottlenecks.

But still some tests were missing and could not be performed on the Interim SAN. Part of the proof of concept was testing various error conditions during running real simulations (like JMST, European Simulation, etc.) in the control rooms.

I. Extensive Testing and Site Acceptance Test

After the installation of the final SANng a very extensive Site Acceptance Test (SAT) was required. So the final tests where done within the SAT. The SAT took more then 3 weeks and included all possible failure scenarios, a complete site failover from PDC to BDC and back. Also traffic simulation systems where used to generate a system load up to 10 times higher than today's standard data rates in Col-CC. This was done to future-proof extensions like upgrading from standard video to high definition video (HDVCA) which will triple at least the existing video data rates.

J. Tom & Jerry vs. Max & Moritz

All the testing, migration and operations was done in parallel at the two different Hitachi HNAS clusters. In order not to get confused with similar names of which HNAS cluster with is related to what job at a time the SANng HNAS were named "Tom and Jerry" and the nodes from Interim SAN were called "Max and Moritz". This caused some smile at the beginning but practice proved it to be a useful naming convention. So this naming was kept.

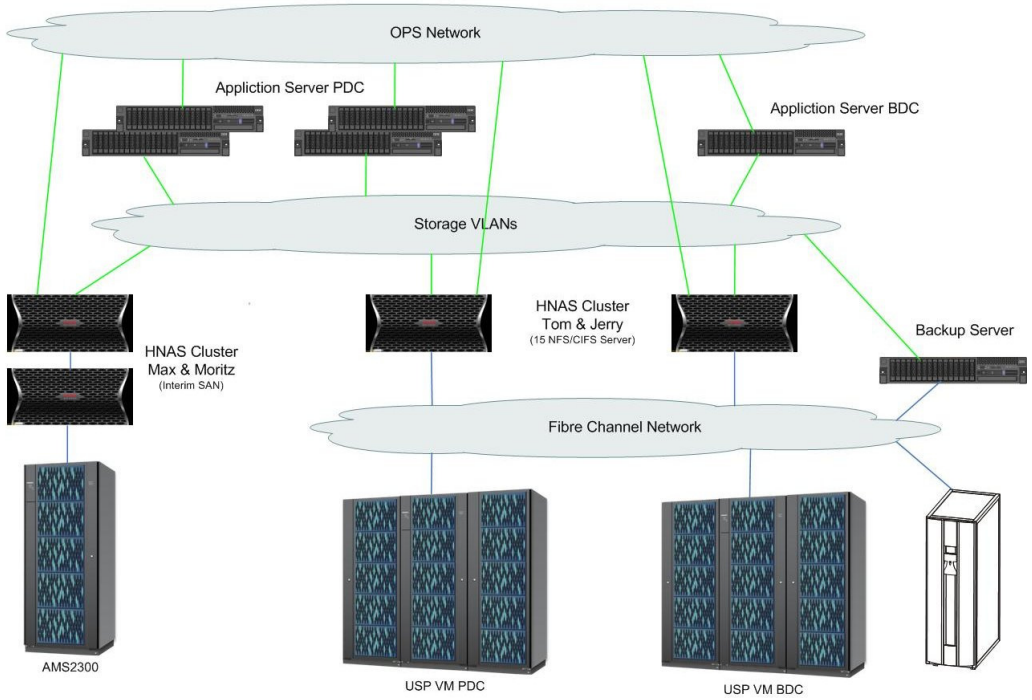

Figure 3 - GSOC SAN Overview

Figure 3 gives an overview of how the two Hitachi SANs are connected. The HNAS Cluster, Max & Moritz of the Interim SAN is placed in a single Rack. The other cluster Tom & Jerry (the SANng) is controlling the USP-VM storage split over two different buildings (PDC and BDC) to support a more secure setup.

Both SANs are connected to both clouds of NIS. Mainly the data is transported via the Storage VLANs but also some shares are exported directly to the OPS Network for example supporting user "Home" directories for the FCT Team. With old HSM-SAN this feature was setup as an extra NFS server cluster.

The Fibre Channel connections of the (very few) servers supported with direct LUNs are not drawn here.

X. Data Migration from Interims SAN to SANng

As described above (see VIII) the data migration using Rsync is very easy and can mainly be done without interrupting operations. The backup server can be connected easily to both SAN clusters. So it is very easy to move (migrate) subsystems between Tom & Jerry = SANng and Max & Moritz = Interim SAN. This made the total system extermly flexible and provided a good opportunity for the extensive testing.

XI. Migration of "historical" data from the old HSM SAN

Migration of the old data residing on the HSM tapes caused us some extra headaches. Just doing an Rsync like with the old "online" data ended up in a mess. The HSM was a basic system and inserted tape cassette after tape cassette into the drives which resulted in a very slow data transfer, too much delay caused by the tape robot and also broke the tape drives itself.

Writing a script, which proofs data availability on disk or on tape and invoking a semi automated restore on directory level instead of reading file by file speed up the process dramatically. But even with that script, reading back all old data from tape took about 18 month in total. Also some old files where lost during this process because of broken AIT tapes and invalid (unusable) second copies due to some more broken tapes or database errors. Most of the missing data could be recovered by getting Path-TM data on tape cassettes from NASA and merging this into our data sets.

XII. Migrating native LUNs

During the system migration most of the old native LUNs were converted into network shares. But some servers still need "native" LUNs. These are for example Microsoft Windows Cluster Server.

These servers where connected redundant to the old Clariions via a dual Fibre link. For migration the redundancy was broken. This means that one Fibre link was set free. So the servers could be connected to both storage systems – the old FC network with Clarrion and the new Fibre fabrics attached to USP VM – at the same time. This allowed copying data from old LUNs to the new LUNs. By adding a second Quorum disk during operations most of the migration could be done without shutting the applications down.

With this preparation the real outage for final switching was just a few minutes, stopping for example a MSQL database and restarting them on the new disks. With the application working on the new disks the old disks could be deactivated and disconnected.

Only de-installing the old EMC Powerpath software and installing the new Hitachi HDLM (Multipath Software) – to get back to redundant data paths – required some cluster reboots. But this could be handled without big interruptions of operations in some LOS periods.

XIII. Backup

Also for the backup system a very simple strategy is used. Applications and operating system itself are standards and can be reinstalled from an installation server. So there is no need for backing up those "local" data.

Only operational data is important to backup. But all operational data is written to the file systems on SAN anyhow. These file systems are visible completely to the backup server. The backup server mounts them and all the data is backed up (like backup PC local data) to a tape robot system. The robot library is split into two parts, a small one for backup and a larger part for the archive.

XIV. Archive

The tape archive is under installation and test, but not yet in use. It will operate similar to the backup scenario and will use the same infrastructure.

In contrast to the backup the archiving process will not be automatically executed. Data from current year and the year before are defined to be available at the online storage (means: on disk). From time to time data will be written to the archive and deleted from online storage.

For availability (safety) reasons archive data will be written to at least to two (currently under consideration three) tape copies. One tape will be kept in the robot library. The others will be taken to a safe location(s).

Data stored in the archive will not be available directly to the user. A DART (Data Archive Retrieval Request) request has to be made and the SAN team will restore the requested data for the defined period. After processed by the user this data will be archived and removed from online storage again.

XV. Conclusion

After all, the migration took nearly two years in total, but the actual switching times were hidden within other service times and/or during LOS periods. So far the migration from old HSM-SAN to the new SANng was (almost) not visible to the FCT. The real visible effect for FCT is that the SAN is now much more stable for operations. No more outages were caused by SAN in the last two years of operation.

The availability of the Interim SAN was good luck and very helpful since data could be moved very smoothly from HSM-SAN to Interim SAN and to SANng later.

No system is perfect, but …

Two years later we found servers with uptime values of 300 days and more. We had of course some minor failures with the Hitachi systems but we are proud not having any to FCT visible outages due to SAN failures like in the years before. Application servers with uptime values of more than 300 days are never seen with old HSM SAN but with SANng this is very usual.

Timeframe:

- 2008 – Planning and testing different SAN concepts at vendor site, mainly: Hitachi, IBM and NetApp
- December 2009 – Installation of Interim SAN as a test system
- March 2010 – Migration of the first subsystems to Interim SAN
- July 2010 – Installation of SANng
- Aug/Sept. 2010 – SAT
- Sept 2011 – final HSM-SAN shutdown

Appendix A
Acronym List

BDC	Backup Data Center
Col-CC	Columbus Control Center
COTS	Common Of The Shelf products
DRP	Data Retention Policy
EVS	Enterprise Virtual Sever – a virtual file server inside a HNAS
GSOC	German Space Operation Center
FC	Fibre Channel
FCT	Flight Control Team
HDS	Hitachi Data Systems
HNAS	Hitachi NAS (Appliance – virtualizes up to 64 EVS)
HSM	Hierarchical Storage Management
HSM-SAN	The old Col-CC SAN based on HSM technology
LUN	Logical Unit (see below)
NAS	Network Attached Storage
NIS	Network Infra Structure
PDC	Prime Data Center
SAN	Storage Attached Network
SANng	SAN Next generation

Appendix B
Glossary

EVS	Enterprise Virtual Server. This is a virtualization of a file server offering network file services. From the possible file services only NFS and CIFS are used. An EVS does not need to be redundant. Redundancy is provided by the HNAS cluster itself. The EVS is migrated to the other node in case of an error.
HNAS Cluster	Hitachi NAS Appliance Tom & Jerry are the two HNAS Heads for SANng Max & Moritz are the two HNAS Heads for Interim SAN
HSM-SAN	The old SAN in Col-CC based an a Hierarchical Storage Management

	Was installed in 2003 and running out of Support in 2009/2010
Interims SAN	A loaned system for proof of concept which was actually used to keep operations up and running.
LUN	Logical Unit. In a SAN environment many disks are combined to some kind of a large disk pool. The pool will be partitioned into practical slices (parts). This parts exported to a server system is typically called LUN.
Multipath	Multipath Software combines 2 separate connections to the storage media and may use them in failover and/or load sharing combination. Without this software the server will see the same Disk or LUN twice and handle the disk like 2 independent ones.
SANng	The term “SAN Next Generation” was defined for the new SAN to have an easy understanding which SAN was meant … because 3 SANs where operated in parallel during migration.