



RUS

*VTLs mit TSM am RUS*

# Einsatz von NetApp VTLs mit TSM am RUS

**Hermann Frasch**

**09.06.2008**



## Agenda

### ■ Ausgangslage

- „Kontinuierlicher Backup in ein Schwarzes Loch?“
- „Wer gewinnt das Hase-Igel-Rennen?“ -> Beschaffungsantrag HBF G

### ■ Aufbau und Abnahme

- Hektik in der Vorweihnachtszeit

### ■ Implementations- und Testphase

- Ringen um die richtige Konfiguration
- Hell of Cases?

### ■ Migration

- Wie verteilen wir 1100 Nodes auf 5 TSM-Server?

### ■ Betrieb

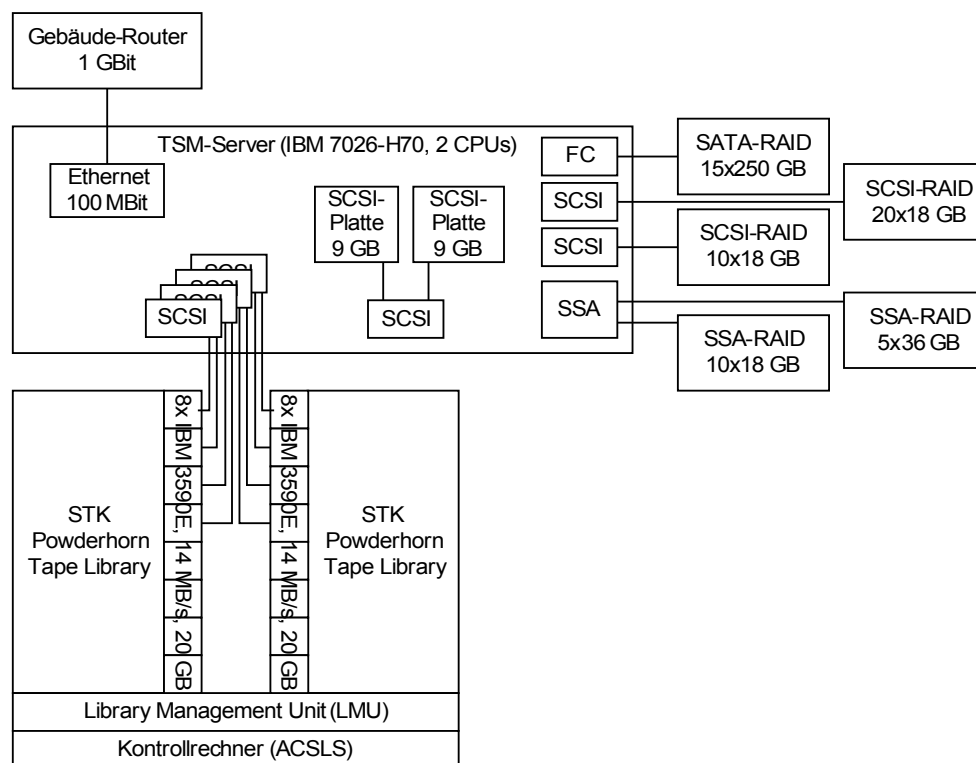
- Allseitige Zufriedenheit oder alltäglicher Frust?



RUS

# Backup - HW Stand Ende 2006

VTLs mit TSM am RUS





## Ausgangslage Ende 2006

### ■ Engpässe

- 100 Mbit-Interface
- STK-Bandbibliothek, 8 Laufwerke, von 11500 Plätzen nur 30 frei, 5055 3590 20GB-Bänder für TSM
- Zahl und Dauer der täglichen Backups, z.T. länger als 24 Stunden, kontinuierlicher Backup -> 2 alternierende 2-tägige Schedules
- Recovery Log wird voll, viele Transaktionen, Log Pins, häufige Datenbankbackups
- Wartungsprozesse (Expiration, Space Reclamation, Migration) laufen häufig parallel und werden tagsüber nicht fertig -> Space Reclamation nachts!
- Lange Restorezeiten, bis zu mehrere Tage bei Full Restore
- Exponentielles Wachstum der Datenmenge/Zahl der Objekte, „Hase-Igel-Rennen“
- Datenbank überlastet, bis 160 GB, Full Backup > 3h, Database Cache Hit Ratio < 90%!
- Viele inaktive Nodes (ca. 150) und Filespaces, Silos als „Schwarzes Loch“

### ■ TSM-Konfiguration

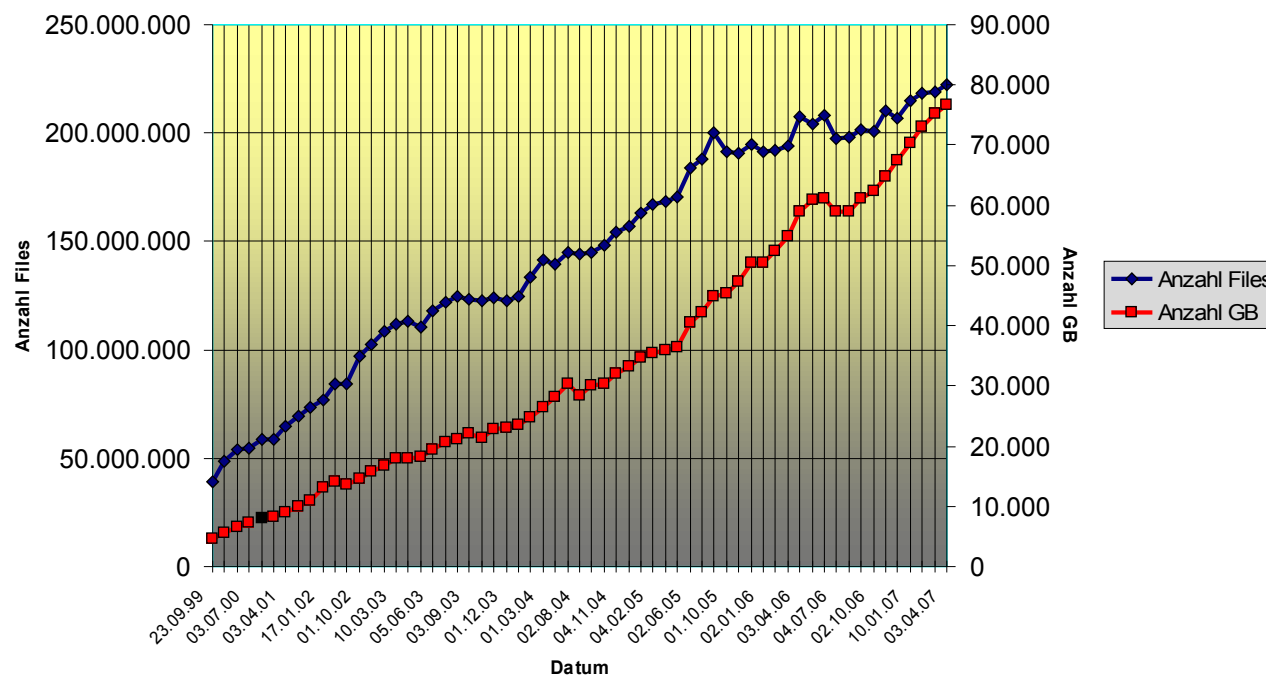
- Policy Domain STANDARD/COLLOCATION (ca. 20 Nodes)
- Migration von Disk- zu Band-Pools (ARCH\_MIG, BKP\_MIG, COLLOC\_MIG)
- Kopie nur von Archivdateien (ARCH\_COPY)
- Datenbankbackups auf Band
- Max. ca. 100 TB Bandkapazität



# Entwicklung der TSM-Auslastung bis 2007

VTLs mit TSM am RUS

Entwicklung der TSM-Auslastung





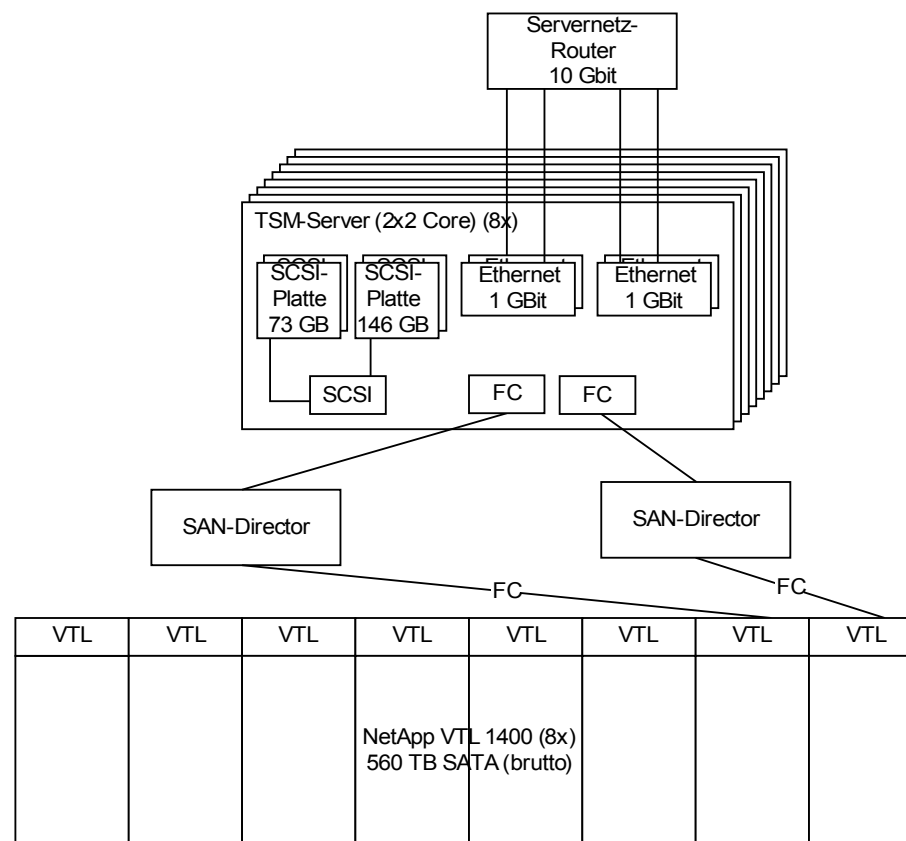
## Backup - Beschaffte Geräte

- **Disk basiertes Speichersystem für TSM, Virtuelle Tape Library „NetApp VTL 1400“**
  - Datenspeicherung auf Festplatten (RAID-Systeme)
  - 1120 Platten (S-ATA) zu 500 GB brutto
  - Virtuelle Bandkassetten: 500 pro VTL
  - Virtuelle Bandlaufwerke: 200 pro VTL
  - Virtuelle Libraries: 2 pro VTL
  - kein Zeitverlust beim Laden und Positionieren der virtuellen Bandkassetten
- **8 TSM-Server FSC RX300/S3**
  - 2 \* DualCore CPU 3 GHz
  - 8 GB RAM, 2 St. 4 Gb FC-HBA, RAID1
  - Neueste Version TSM 5.4 unter Linux SLES9 vereinheitlicht die Administration
- **FC-SAN**
  - 2 Fabrics zu je zwei 32-Port FC-Switches
- **Redundante Netzwerkanbindung**
  - 2 St. CISCO 6509



# Backup: Neue Konfiguration

VTLs mit TSM am RUS





## Aufbau und Abnahme

### ■ Aufbau

- Lieferung der je 8 TSM-Server und VTLs am 15.12.2007
- Aufbau und Verkabelung der TSM-Server, Switches, Router, VTLs – Arbeit wie im Ameisenhaufen

### ■ Konfiguration und Tests

- 18.12.-21.12.2007
- Plan: 2 virtuelle Libraries pro VTL mit je 100 Laufwerken, 1. Library am Port 0a, 2. am Port 0b
- Auswahl Library- und Band-Typ
  - Welches ist der passende Typ für Passthru-Treiber tsmscsi?
- Anbindung an TSM
  - Erkennt der TSM-Server die VTLs? Verkabelung, SAN Zoning
  - Wie definiere ich am schnellsten 200 Laufwerke und die Pfade dazu und wie entferne ich sie wieder? Entwicklung von entsprechenden Shell-Skripten
  - Erster Case wegen „Hardware or media error“ (Drive or media failure)

### ■ Abnahme

- 21.12.2007 – Geräte können bezahlt werden



## Implementations- und Testphase

### ■ Konfiguration VTLs

- 7 VTLs mit 10 Disk Shelves a 14 500GB-SATA-Platten, 1 mit 1 Disk Shelf, Compression Card
- 7 mit je 28 Raid Groups mit 4 Hot Spare Disks, 1 mit 3 Raid Groups /1 Hot Spare
- 6 VTLs für Produktion mit je 2 virtuellen Libraries mit je 250 100GB-Tapes, 300 Slots, 100 Drives vom Typ AIT300 (300 GB Raw Capacity, 100 GB Disk Allocation, SONY SDX-700C), 1 VTL für NDMP mit 500/500/25, 1 VTL für Tests mit 2 virt. Libraries mit 20/20/10

### ■ Konfiguration TSM-Server

- 8 Server (RUSTSMA-H) mit TSM 5.4.0.0 (5 Produktion, 1 NDMP, 1 Reserve, 1 Test)
- Keine Migration, Daten werden direkt auf virtuelle Bänder geschrieben
- 2 Device Classes vom Typ 8mm (AIT3PRI, AIT3SEC)
- 2 virtuelle Libraries (RUSTSMA: primär: RUSVTLA0, sekundär: RUSVTLB1) mit je 100 virtuellen 8mm-Laufwerken für Primary und Copy Storage Pools für Backup- und Archivdaten

### ■ Tests

- Mangelnder Durchsatz oder langsame Netzanbindung der Clients?
- Cases bei NetApp NOW wegen VTL-Hängern, Physical end of media, Device or media failure, SCSI adapter failure; mehrere Firmware Updates/Upgrades; Tausch eines HBAs
- Bandende nicht bei 100 GB Disk Used, sondern bereits bei 100 GB Data Written



RUS

## Migration

### ■ Vorbedingungen

- Nur Nodes, keine Daten werden migriert
- Nur aktive Nodes (letzter Zugriff <180 Tage) werden gleich übernommen
- Archivdateien müssen per Retrieve vom alten TSM-Server zurückgeholt werden
- Gleichmäßige Verteilung der Nodes auf 5 TSM-Server (Round Robin?)
- Neues Namensschema unter Verwendung der Institutsnummer
- Neue Standard Policy Domains (S-STANDARD, H-STANDARD)
- Neues Contact-Feld wegen TSM Operational Reporting Tool
  - Inst\_bezeichnung, tel\_nr, mail\_addr, kontaktperson, flag

### ■ Migration Stuttgarter Nodes

- Skript zur Benachrichtigung der Kontaktperson und Erstellung von Macros mit Register Node und Define Assoc-Befehlen
- 27.03.2007 Registrierung von 125 RUS-Nodes
- 24.04.2007 537 Nicht-RUS-Nodes, institutsweise auf 5 TSM-Servern entspr. Datenmenge
- Probleme mit TSM-Passwort, nicht erreichbaren Kontaktpersonen
- Keine Löschung von Nodes bzw. Daten auf altem Server bis Abschaltung



## Betrieb - TSM

- 7 TSM-Server in Produktion, am 25.09.2007 Update auf TSM 5.4.1.1
- Selftunebufpoolsize, Bufpoolsize 1048575, Database Cache Hit Ration >99%
- Datenbank zw. 35 bis 65 GB, Recovery Log 10 GB
- 1 Full Database Backup/Tag auf /srv/tsm/backup (per NFS vom NAS-Filer)
- Kopie von Devconfig und VolumeHistory auf /srv/tsm/backup
- Shell-Skripte, Makros, Scripts auf /srv/tsm/common
- Schedules
  - Administrativ: Datenbank Backup (06:00), Storage Pool Backup (08:00), Expiration (10:00), Space Reclamation (11:30), DELETE\_VOLHIST (05:59, >3 Tage)
  - Client: 18:00-08:00: S-TAEGlich, S-2-TAEGIG\_1, S-2-TAEGIG\_1, So 02:00-12:00: S-WOECHENTLICH; Hohenheim: 18:00-04:00, 12:00-15:00, spezielle
- Client Options Sets (SCHEDMODE POLLING, COMPRESSION NO, Excludes)
- Wenige Langläufer (Idle Timeout von 30 Min. wird ignoriert (Bug), langsame Netzverbindung, Ressourcenmangel, z.T. große Differenz zw. Network und Aggregate Data Transfer Rate)
- Mangel an Scratch-Bändern -> Export von Nodes, Verzicht auf 2. Kopie
- Administration über Befehlszeile (ssh zum Server), Server-Group tsmgroup1
- Fehler-Meldungen über TEC-Console
- Tägliche HTML-Reports vom TSM Operational Reporting Tool



## Betrieb - VTLs

- Administration über Web Interface
  - Monitoring
  - Konfiguration (Raid Groups, Libraries, Tapes)
  - Diagnostics Download (z.T. umfangreiche Dateien)
  - Firmware Upgrade
  - Spezielle Zusatzbefehle zur URL (diag, debug, cat)
  - Diagnostics Port 8888 (Reboot, Diagnostics)
- Console Interface über KVM-Switch oder per ssh
- Wöchentliche Email Notification mit Statusberichten und komprimierten Systemdateien
- Action Requests nach Neustart, bei Problemen, Firmware Upgrade → Resolve
- Action Notification per Email
- Meldung und Bearbeitung von Cases bei NetApp NOW
- letzter Firmware Upgrade 5.2.3P2 am 29.10.2007, Patch D30 am 03.03.2008
- Austausch von 3 Platten, 1 Compression Card, 1 Power Supply



## Fazit

### ■ Vorteile:

- Wartungsprozesse werden tagsüber fertig
- Datenbankengpässe behoben
- Wesentlich kürzere Backup- und Restorezeiten
- Höherer Datendurchsatz (früher max. 800 GB pro Tag, heute bis zu 277 GB pro Stunde)

### ■ Nachteile und Probleme:

- Ungeklärte temporäre I/O- und sonstige Mediafehler (jedoch selten mit Datenverlust)
- Gelegentliche Hänger der VTLs, Schedule Backups bleiben hängen, Neustart der VTLs und von jeweils 2 TSM-Servern notwendig
- Bänder gelegentlich lange im Zugriff von Client-Sitzungen, Konkurrenz mit Prozessen
- Kapazität
  - TSM rechnet mit Basis 2, VTL mit Basis 10
    - $1 \text{ GB Basis } 10 = 0,93 \text{ GB Basis } 2$ ,  $1 \text{ TB Basis } 10 = 0,91 \text{ TB Basis } 2$
  - Gesamtkapazität  $(7 * 50000 \text{ GB } (10) * 0,93 = 325500 \text{ GB } (2))$  wegen 2. Kopie zu knapp
  - 2. Kopie auf VTLs zu kostspielig (notwendig?)
- Exponentielles Datenwachstum geht ungehindert weiter – Wer wird Hase? Wer Igel?



# Entwicklung Gesamtbelegung TSM 2007/2008

Datum	Gesamt in MB	2. Kopie in MB
13.07.2007	207.755.297	103.877.648
03.09.2007	231.978.020	115.912.334
17.10.2007	254.070.625	126.808.895
07.11.2007	264.839.644	132.413.239
19.12.2007	274.598.815	136.362.357
07.01.2008	282.553.169	138.009.240
13.02.2008	290.652.607	136.935.213
03.06.2008	211.257.838	37.593.193

VTLs mit TSM am RUS