

# Einsatz von Sam-FS HSM am Max-Delbrück-Centrum

Vortrag am 18.6.2015  
auf der 7. Sam-FS  
Nutzerkonferenz

Christian Steinborn

Dipl.-Inf./FH-Systembetreuer

IT-Abteilung am Max-Delbrück-Centrum für  
molekulare Medizin in der Helmholtz-Gemeinschaft  
in Berlin-Buch

# Max-Delbrück-Centrum in Stichpunkten...

- In Berlin-Buch angesiedelt, Mitglied der Helmholtz-Gemeinschaft
- 1992 gegründet; aus der Akademie der Wissenschaften der ehem. DDR hervorgegangen
- Ca. 70 Arbeitsgruppen, ca. 1600 Mitarbeiter, wachsend
- Forschung im Bereich der molekular(genetisch) bedingten Ursachen von Herz- Kreislauf-, Stoffwechsel- und onkologischen sowie Erkrankungen des Nervensystems, Behandlungsansätze und Prävention
- Angeschlossen Berlin Institute of Health (Systembiologie)
- Zusammenarbeit mit der Charite
- Weitere Informationen unter <https://www.mdc-berlin.de>

# IT-Abteilung am MDC

- Mehrere Bereiche, 23 Mitarbeiter
- High Performance Computing
- Web-Anwendungen
- SAP und kaufmännische Anwendungen
- Netzwerk
- Virtualisierung
- Storage und Backup
- User Support
- Bildlabor

# Sam-FS am MDC

- Seit 1998 im Einsatz, zunächst im Fileservice bis 2012, danach ZFS Open Storage
- Ab 2005 dann auch im Backup
- Ab 2009 im Bereich der PACS (Picture Archiving and Computing System) - Archivierung von im MRT erstellten Bildern
- Ab 2011 im Archivbereich

# I. Ausfallsicheres Backup - die Hardware

- Insgesamt vier SPARC-Server, ein Master, ein Mediaserver sowie zwei weitere Server in anderem RZ, die als Ausfallserver dienen
- Brocade FC-SAN mit zwei Fabrics
- Sun/Oracle Storage STK6140, mehrere Luns in Hostgruppen, gemappt an die Server
- Eine VTL3600 sowie eine SL500 (ACSLs) mit acht LTO4-Laufwerken

# Backup Funktionsweise - die Software

- Solaris 10, SamFS 4.6.85
- Symantec Netbackup, noch 6.5.4
- Luns des HSM-Caches sind in Metasets organisiert; diese sind das Ziel (STU) von Netbackup
- Hostbasierte Spiegelung zwischen Standorten, Teile der Installation im schwenkbaren Storage, virt. IP-Adresse
- Gesamte Sam- und Netbackup-Installation kann verschoben werden, Teile der Sam- sowie der NBU-Installation liegen in den Luns des Metasets

# Erfahrungen

## Vor- und Nachteile

- Netbackup und SamFS vertragen sich gut
- Performance am Anfang unbefriedigend, Grund: Gleichzeitiges Lesen und Schreiben, Folge: Vermehrte Ausfälle im Tape-Bereich
- Abhilfe: Verbessertes Streaming durch getrenntes Backup und Archivierung (Crontab-Steuerung)
- Danach kaum noch Tape-Probleme
- Cache muss groß genug sein (6 und 11TB)
- Bestimmte Dateien dürfen nicht released werden
- Tape-Devicepfade sollten gleich sein
- Komplexität des Aufbaus hoch, Einblick in Interna notwendig

## II. Archiv, installiert 2011

- Nicht ausfallsicher ausgelegt
- Ein Oracle X4200 Server, HBA an zwei FC-SAN-Fabrics angeschlossen
- Solaris 10 x86, SamFS 5.2
- Mehrere Luns einer STK6140 sind der HSM-Cache (15TB)
- Linux-Server schreibt über NFS (hausprogrammierte Archiv-Anwendung) gezippte Dateien in den HSM-Cache
- Zwei Kopien werden auf eine SL500 mit sechs LTO4-Laufwerken (ACSLs) archiviert
- Nach einigen Tuningmaßnahmen (Netz) im Wesentlichen störungsfreier Betrieb, nur ein Restore nötig (Raid5 Double Fault)
- z.Z. Ca. 70 TB, Archivaufkommen unterschiedlich



### III. MRT Bereich (PACS-Archivierung)

- Solaris 10, SamFS 5.2, Oracle, PACS
- Zwei T5220-SPARC-Server, ein PACS-Datenbankserver (Oracle und Carestream) sowie ein SamFS-Server
- Datenbankserver und SamFS-Server greifen gemeinsam über QFS auf das HSM-Filesystem zu
- Eine Kopie auf eine VTL1200, eine weitere auf eine SL500 (ACSLs), angebunden mit sechs LTO4-Laufwerken über eine FC-SAN-Fabric
- Betrieb bislang bis auf wenige Ausnahmen störungsfrei

# Ausblick Backup

- Grundsätzlicher Ansatz HSM mit Kopplung zu Netbackup soll beibehalten werden
- Es soll jedoch auf x86-Hardware unter Linux laufen
- Konsolidierung, weniger Server und Komplexität
- Anderes HSM
- Fazit: HSM und Netbackup zusammen kann sich lohnen, wenn SAM- bzw. HSM-Lizenz ohnehin vorhanden ist

# Ausblick Archiv

- Es soll weiter auf X86-Hardware, aber unter Linux laufen
- Als Archiv-Software soll iRods eingesetzt werden, da die vorhandene Softwarelösung zu eingeschränkte Möglichkeiten bietet
- Migration der alten Daten notwendig, da iRods anderes Format verwendet als unsere Anwendung
- Anderes HSM notwendig, evtl. derselbe HSM-Server wie beim Backup
- Höhere Cache-Größen als bisher, um Daten sicher aufnehmen zu können, andere Storage-Hardware

# Ausblick MRT

- PACS soll noch weiter auf der alten Hardware laufen
- SamFS soll noch weiter laufen
- Konzept hat sich grundsätzlich bewährt
- Planung auch abhängig davon, ob altes PACS-System weiter verwendet wird und ein evtl. neues System unter Solaris läuft

# Zwei der Tape-Roboter am MDC



# Virtual Tape Library (200TB)

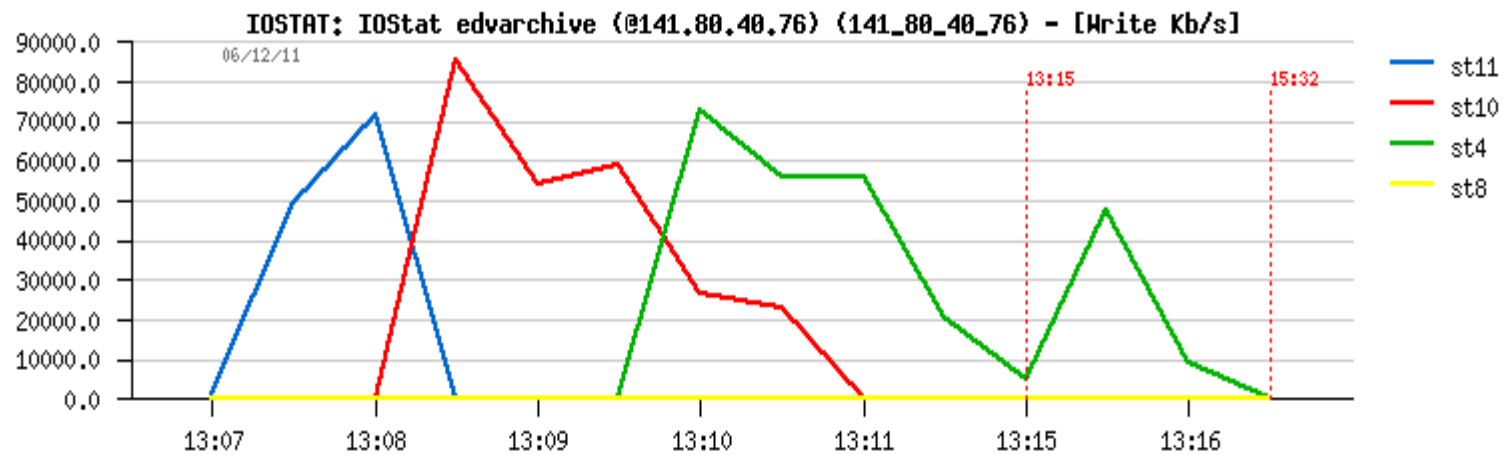
Tapes werden auf Disk emuliert



# Performance Monitoring

*dim\_STAT* Analyzer

Iostat edvarchive (@141.80.40.76) (141\_80\_40\_76)



# Tape-Überprüfung...

Crossroads®  
**ReadVerify Appliance™**

**Library : View Drives**

Drives (8)

- View errors for last 24 hours
- View statistics for last 24 hours
- View drive properties

	Drive	Product	Location ▾	WWN	Tape	Pools
<input type="checkbox"/>	<a href="#">HU19487TJF</a>	HP Ultrium 4-SCSI rev: H58S	0,0,3,2	500110a001328fd6		
<input type="checkbox"/>	<a href="#">HU19487TW9</a>	HP Ultrium 4-SCSI rev: H58S	0,0,3,1	500110a00129df52		
<input type="checkbox"/>	<a href="#">HU19487TFB</a>	HP Ultrium 4-SCSI rev: H58S	0,0,2,4	500110a00132a8e6		
<input type="checkbox"/>	<a href="#">HU195082EM</a>	HP Ultrium 4-SCSI rev: H58S	0,0,2,3	500110a00132c4be		
<input type="checkbox"/>	<a href="#">HU19487TFA</a>	HP Ultrium 4-SCSI rev: H58S	0,0,2,2	500110a00132a036		
<input type="checkbox"/>	<a href="#">HU1047DP5C</a>	HP Ultrium 4-SCSI rev: H58S	0,0,2,1	500110a0014461f2		
<input type="checkbox"/>	<a href="#">HUE1101P7E</a>	HP Ultrium 4-SCSI rev: H58S	0,0,1,2	500143801214958a	<b>400160</b>	
<input type="checkbox"/>	<a href="#">HU19497YBH</a>	HP Ultrium 4-SCSI rev: H58S	0,0,1,1	500110a00132b5ea	<b>400159</b>	



# Kleiner Teil der Disk-Systeme (HSM Cache)



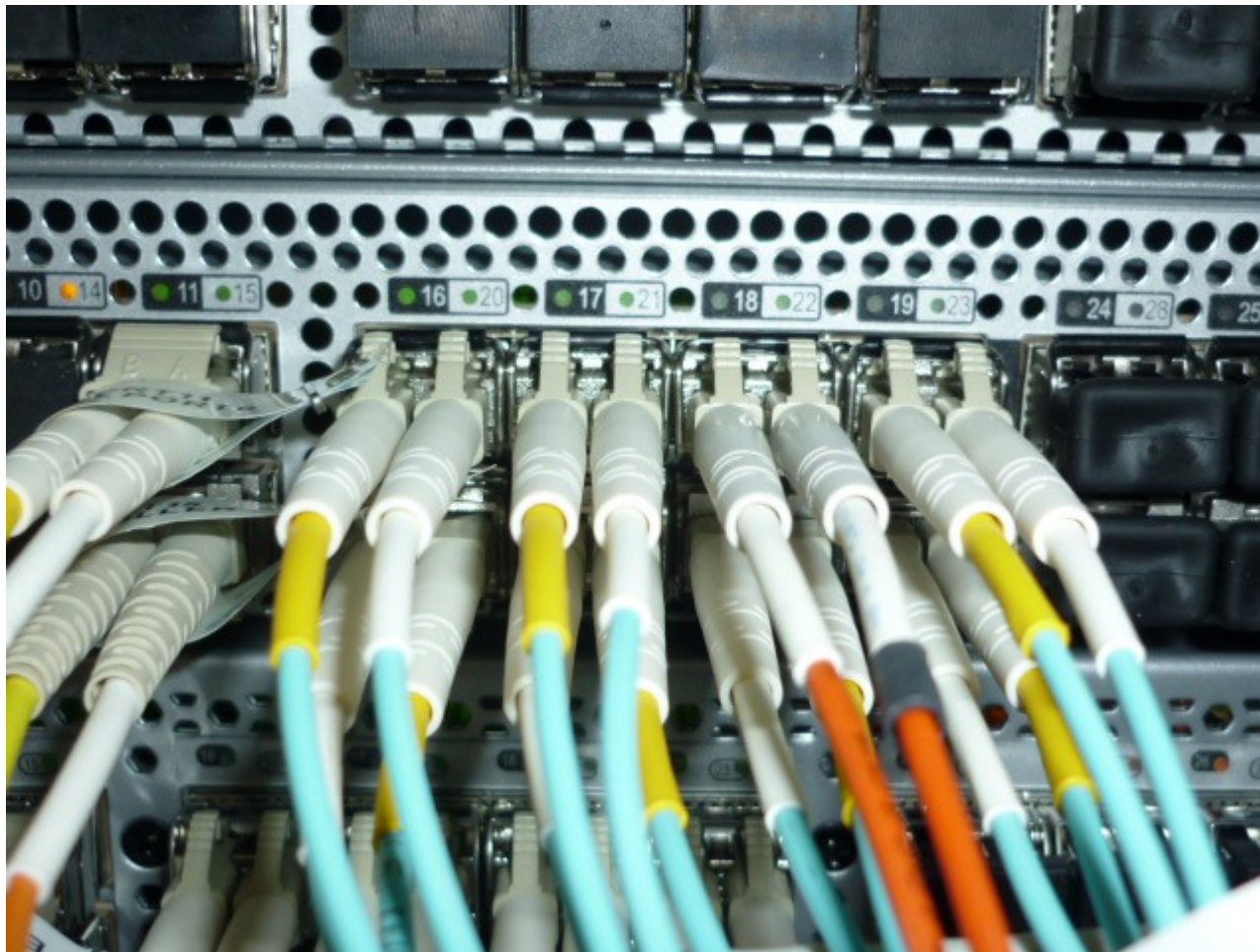
# Teil der FC-Zonenkonfiguration...

## Quellen und Ziele werden definiert

edvbcks1\_PCI2P2\_edvbckm\_a1\_cB\_ch1  
edvbcks1\_PCI2P2\_edvbckm\_a2\_cB\_ch1  
edvbcks1\_PCI2P2\_vtl3600\_x4600oben\_PCI3P1  
edvbcks1\_PCI2P2\_vtl3600\_x4600unten\_PCI3P1  
edvbcks1\_PCI4P2\_edvbckm\_a1\_cB\_ch1  
edvbcks1\_PCI4P2\_edvbckm\_a2\_cB\_ch1  
edvbcks1\_PCI4P2\_vtl3600\_x4600oben\_PCI3P1  
edvbcks1\_PCI4P2\_vtl3600\_x4600unten\_PCI3P1  
edvbcks1\_PCI5P2\_edvgl500b\_tape2  
edvbcks1\_PCI5P2\_edvgl500b\_tape3  
edvbcks1\_PCI5P2\_edvgl500b\_tape6  
edvbcks1\_PCI5P2\_edvgl500b\_tape7

edvbcks2\_PCI2P2\_edvgl500b\_tape2  
edvbcks2\_PCI2P2\_edvgl500b\_tape3  
edvbcks2\_PCI2P2\_edvgl500b\_tape6  
edvbcks2\_PCI2P2\_edvgl500b\_tape7  
edvbcks2\_PCI4P2\_edvbckm\_a1\_cB\_ch1  
edvbcks2\_PCI4P2\_edvbckm\_a2\_cB\_ch1  
edvbcks2\_PCI4P2\_vtl3600\_x4600oben\_PCI3P1  
edvbcks2\_PCI4P2\_vtl3600\_x4600unten\_PCI3P1  
edvbcks2\_PCI5P2\_edvbckm\_a1\_cB\_ch1  
edvbcks2\_PCI5P2\_edvbckm\_a2\_cB\_ch1  
edvbcks2\_PCI5P2\_vtl3600\_x4600oben\_PCI3P1  
edvbcks2\_PCI5P2\_vtl3600\_x4600unten\_PCI3P1

# Teil der FC-Verkabelung



# Eindruck und Erfahrungen

- Mit Sam-FS grundsätzlich gute Erfahrungen; leider nur unter Solaris verfügbar
- Einblick in die Administrations-Problematik notwendig; kann nicht mal eben von anderen mit administriert werden
- Gute Planung und Designphase notwendig
- Gewisses Spezialwissen notwendig, da mehrere Bereiche ineinander greifen
- Nur Quantität bei Ausstattung reicht nicht aus
- Konfig muss angepasst werden, wenn sich Anforderungen ändern

# Eindruck und Erfahrungen

- Manchmal scheinen andere Lösungen einfacher; deren Nachteile werden dann in Kauf genommen. Dies hängt auch mit sinkenden Disk-Speicher-Preisen zusammen
- Beim Fileservice gab es Vorteile (Backup) beim HSM gegenüber der jetzigen Lösung (ZFS Storage), aber auch Nachteile (keine Snapshots, ADS-Anbindung)
- Fazit für mich: Grundsätzlich optimale Lösung für “Big Data” gibt es wohl nicht
- Vorteile HSM manchmal schwer nahe zu bringen

Vielen Dank für Ihre Aufmerksamkeit.  
Hat jemand Fragen ?