

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CSCS

Swiss National Supercomputing Centre



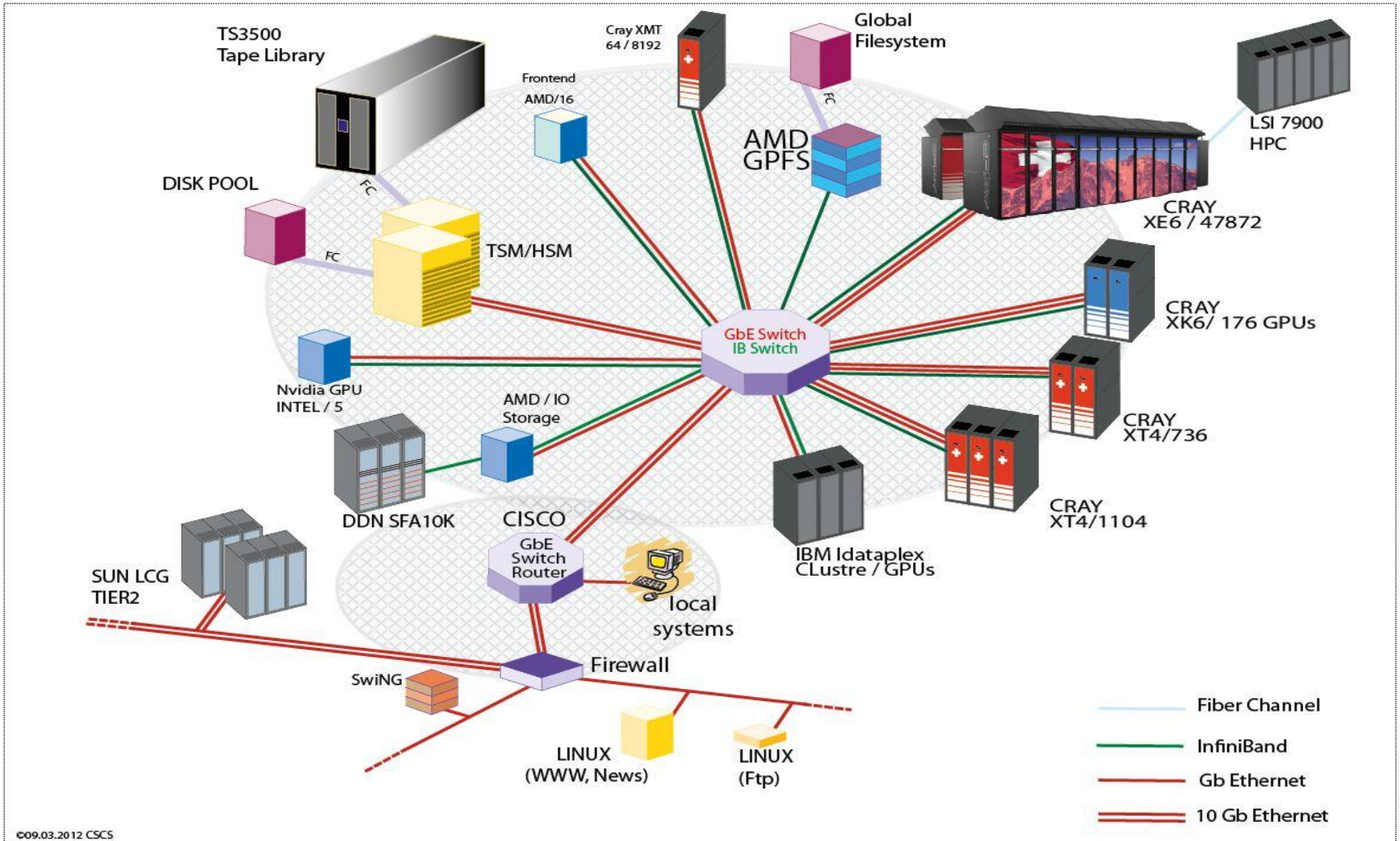
HPC Disaster Recovery at CSCS

Hussein N. Harake

Points to Cover

- Overview CSCS resources
- Disaster Recovery Requirements
- Data & File Systems
- Archiving and Backup Solution
- Backup & Restore
- The challenging part

CSCS Resources



©09.03.2012 CSCS

Disaster Recovery Requirements



Requirements

- Support up to 1 PB GPFS file system with ~ 20M files (/project)
- Daily changes of ~ 3 TB
- GPFS metadata recoverable within 2 hours
- Critical files recoverable within 2 days (10% to 20% of total file system)
- Non critical files recoverable within 2 to 3 weeks

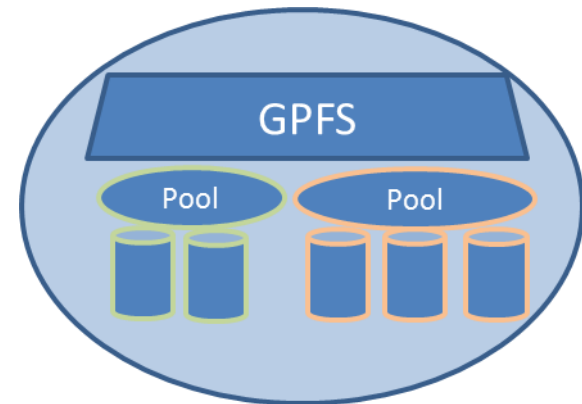
Backup & Disaster Recovery Requirements



Requirements

- Must support migration of 1 PB archive from SAM-FS (Adula)
- Transparently to CSCS users
- To be completed by November 2011, with assistance from IBM-HMK

Storage systems

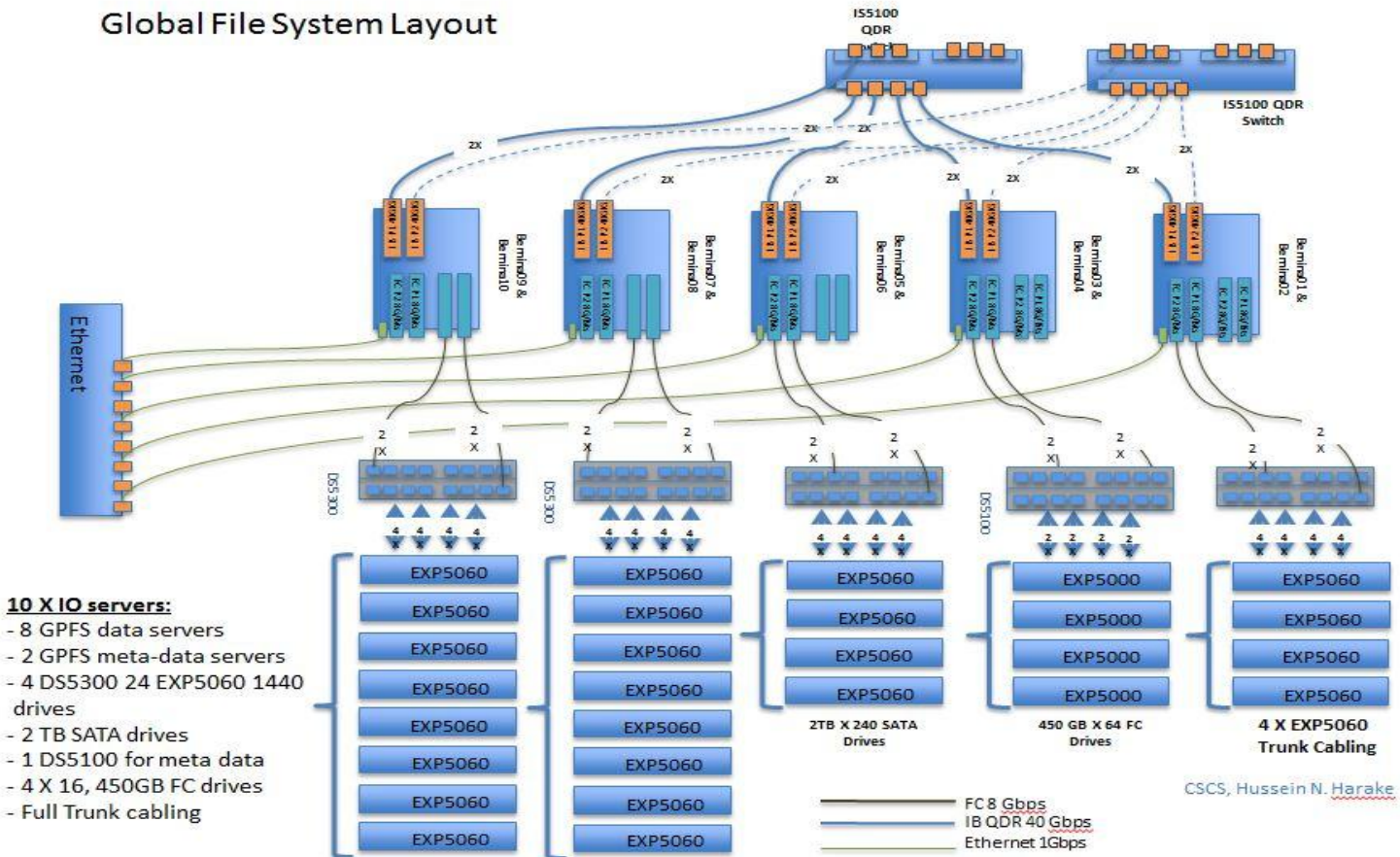


Requirements

- Solution must scale to support future data growth and storage services:
- /project file system expected to grow to ~ 3 PB, 100M files within the next 2 years
- 2H11 addition of /store file system (multi-PB GPFS)

Data & File Systems

Global File System Layout



Data & File Systems

Project file system

- GPFS File system
- 1PB of used space
- ~ 35M files
- Average ~30MB per file

Store file system

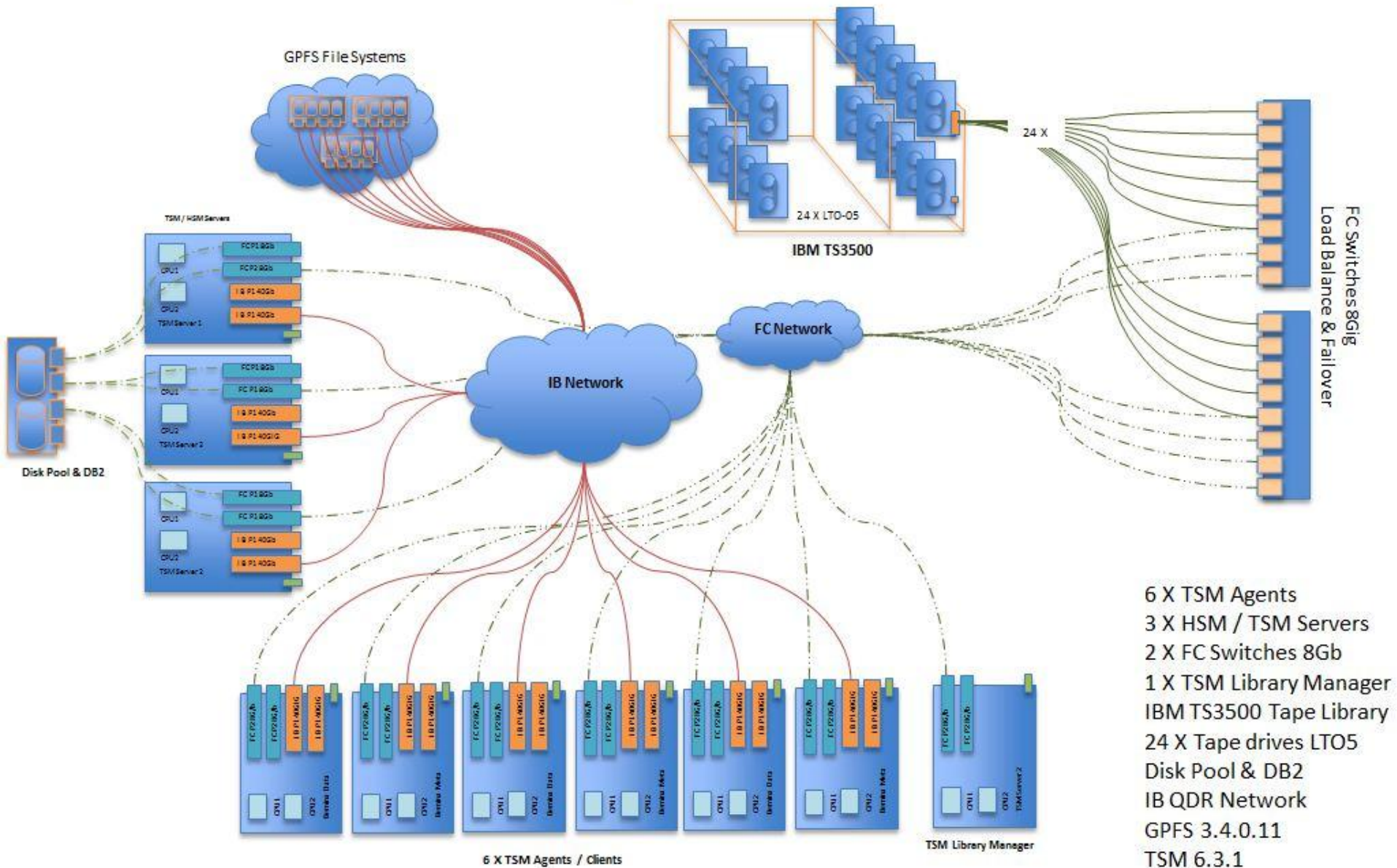
- GPFS
- 2.5PB
- ~100M files

Smaller File Systems

- Global Apps
- Global Home



CSCS Disaster recovery solution TSM / HSM



- 6 X TSM Agents
- 3 X HSM / TSM Servers
- 2 X FC Switches 8Gb
- 1 X TSM Library Manager
- IBM TS3500 Tape Library
- 24 X Tape drives LTO5
- Disk Pool & DB2
- IB QDR Network
- GPFS 3.4.0.11
- TSM 6.3.1

Archiving & Backup Solution

- General Parallel File System(GPFS)
- Tivoli Storage Manager (TSM)
- System x servers and disk storage
- LTO tape storage systems
- 3 TSM Servers + 1 Spare
- 6 TSM Storage Agents
- 24 LTO Tape Drives
- 5,719 LTO5 Slots & 5,719 Cartridges
- 8.58 PB uncompressed



Backup & Restore

- 24 Tape drives deliver ~2400MB/s
- Restoring Metadata time current GPFS
- Time between exchanging drives
- Location of data (which tapes)
- GPFS Today & Tomorrow
- 10 to 20TB of restore data per day

The challenging part

- In case of Disaster:
 - HW limitation and raiding issues
 - Rebuilding the filesystem
 - Time to restore the metadata from the DB
 - Critical projects and high priorities
 - Finalizing the restore procedure

Open Issues & Improvements

- Optimize the Space on Performance
- Recovering MetaData image from Tapes (GPFS 3.5)
- Real Sustained Bandwidth
- Colocation by Group (optimize restore)
- mmbackup and TSM shadow file
- TSM error and GPFS policy

All details will be followed by Stefano Gorini