ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CSCS
Swiss National Supercomputing Centre

HPC Storage Systems at CSCS

hpc-ch Forum on Parallel File Systems for HPC

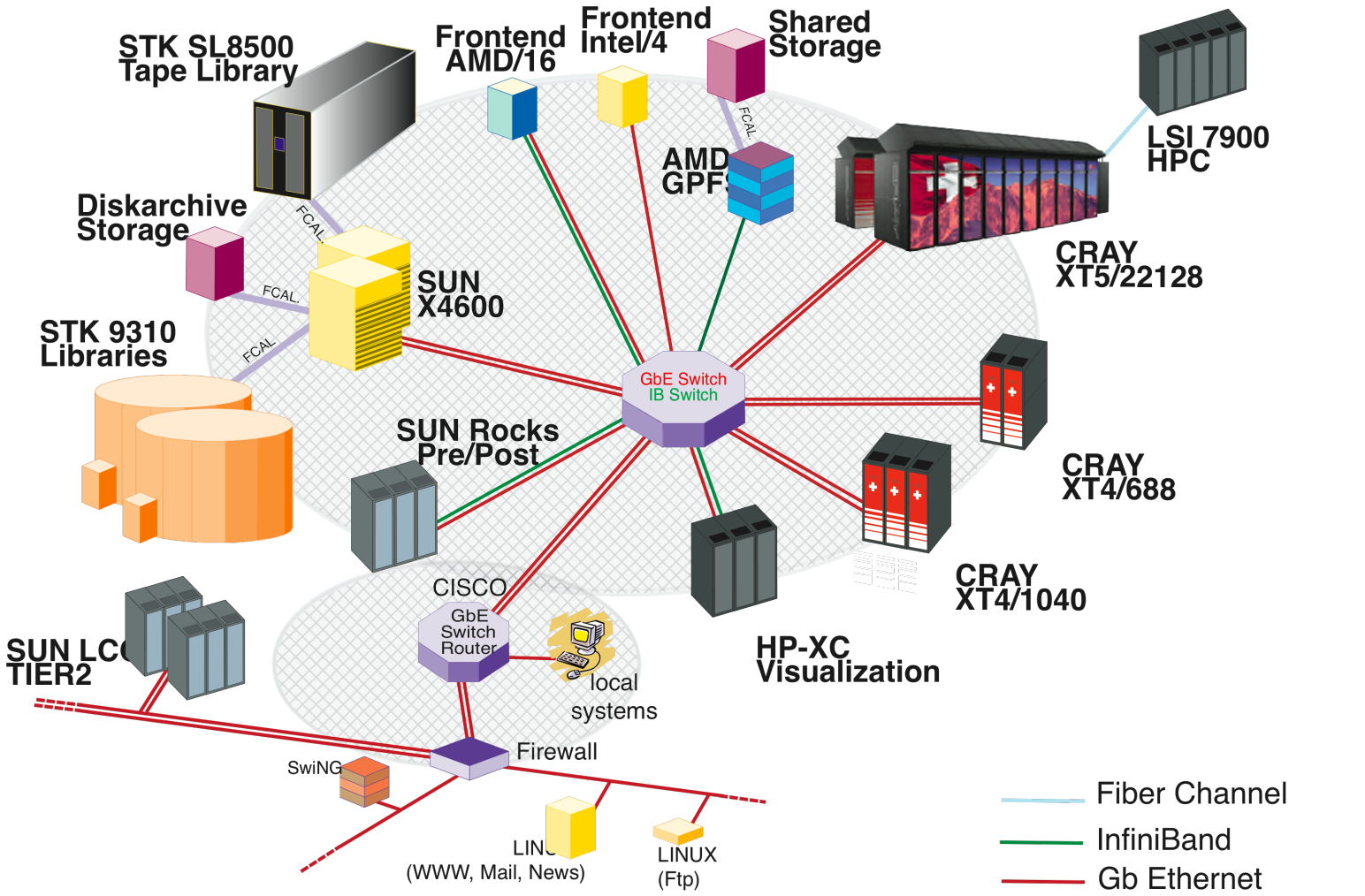
October 28th, 2010

Davide Tachella, CSCS

Points to Cover

- Overview of CSCS resources
- Storage systems
- Global file-system
- IB network
- Implementation

Non-technical Map of CSCS



Storage systems

- Cray XT5
- ✓ 5 LSI HPC 7900
- ✓ Lustre 1.6.X
- ✓ 20 IO servers
- ✓ 287 TB Used space

- Cray XE6
- ✓ 1 X DDN 9900 (SFA 10K 2010)
- ✓ Lustre 1.8.X
- ✓ 4 IO servers
- ✓ 480 TB Used space

A scratch file-system available on each XT / XE machines

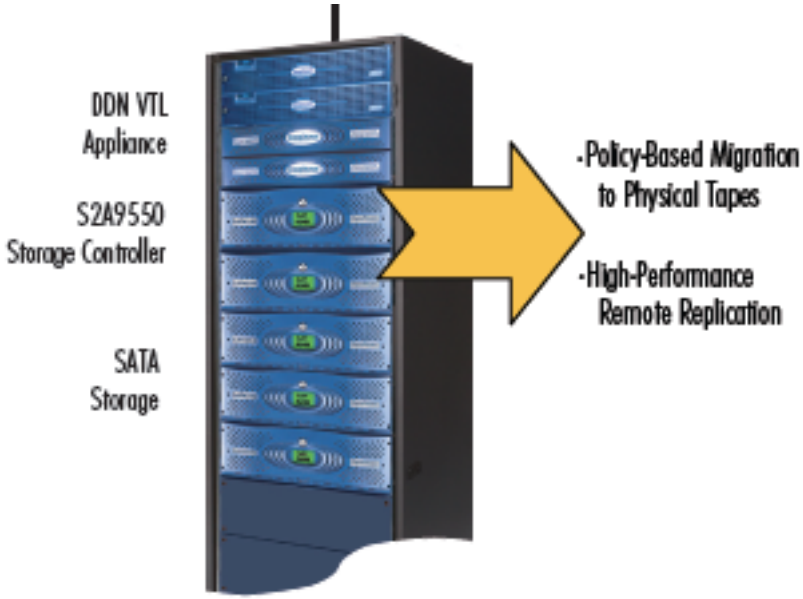
Fast access, short term use and under clean policy

Storage systems

- XT4 (Meteo Swiss)
- ✓ SUN STK 6540
- ✓ Lustre file-system 1.6.X
- ✓ 6 IO servers

IBM DS3400

- XT4 (Backup Meteo Swiss)
- ✓ IBM DS3400 & Exp3000
- ✓ Lustre file-system 1.8.X
- ✓ 4 IO servers
- ✓ 2 Inet routers



Each cluster usually has its own HW where scratch file-system is hosted

Storage systems

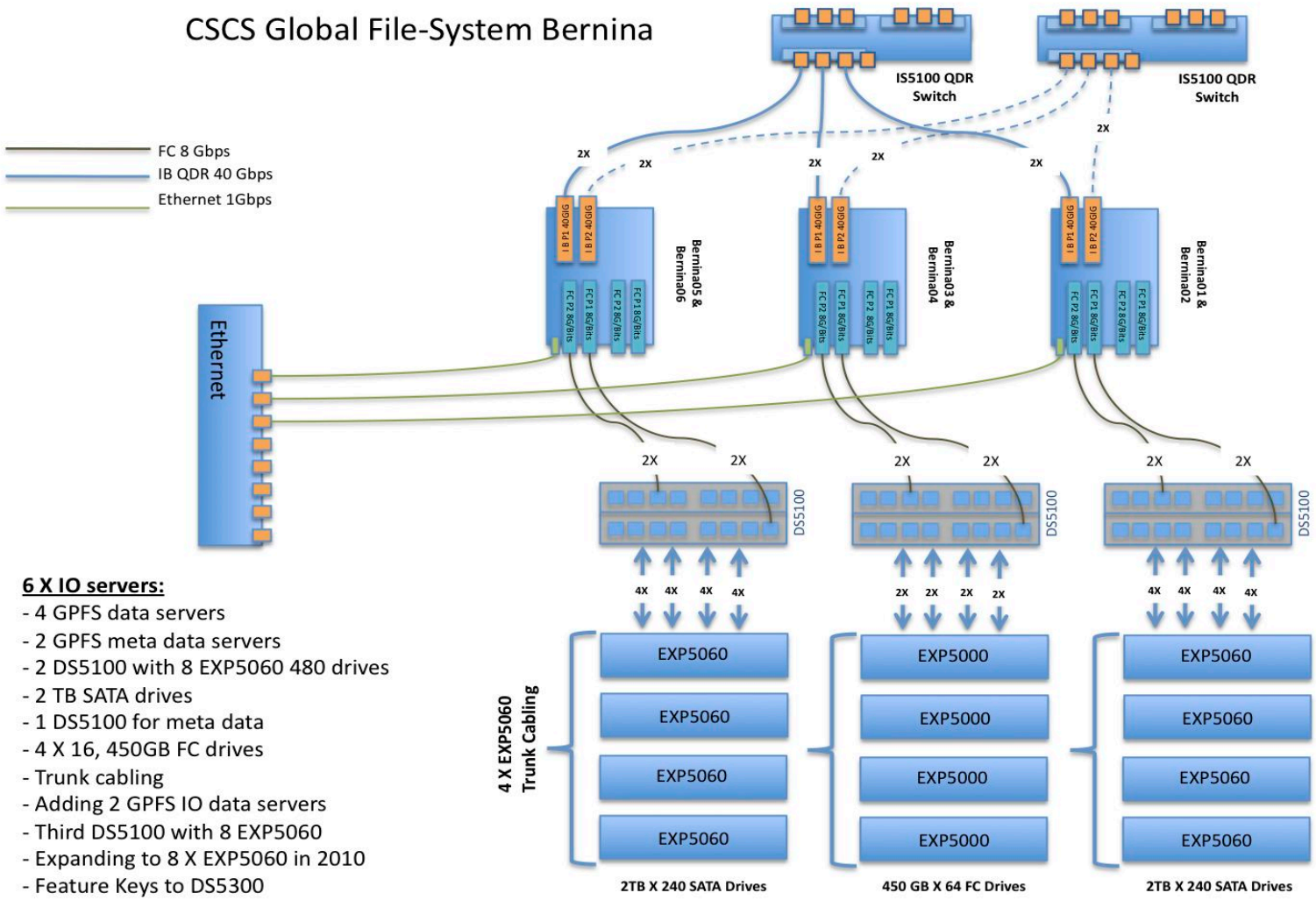
- Global File System /project (In production)

✓ GPFS
✓ 700 TB Used Space
✓ 3 X IBM DS5100 controllers
✓ 8 X EXP5060 Enclosures SATA disks
✓ 4 X EXP5000 Enclosures FC disks
✓ 4 X Data servers
✓ 2 X Meta data servers
✓ 12 X FC HCA's 8G/bits Host Ports

IBM DS5100

Reliable, Reasonable access time and Medium term use
Accessed from all CSCS HPC systems

CSCS Global File-System Bernina



6 X IO servers:

- 4 GPFS data servers
- 2 GPFS meta data servers
- 2 DS5100 with 8 EXP5060 480 drives
- 2 TB SATA drives
- 1 DS5100 for meta data
- 4 X 16, 450GB FC drives
- Trunk cabling
- Adding 2 GPFS IO data servers
- Third DS5100 with 8 EXP5060
- Expanding to 8 X EXP5060 in 2010
- Feature Keys to DS5300

CSCS, Hussein N. Harake

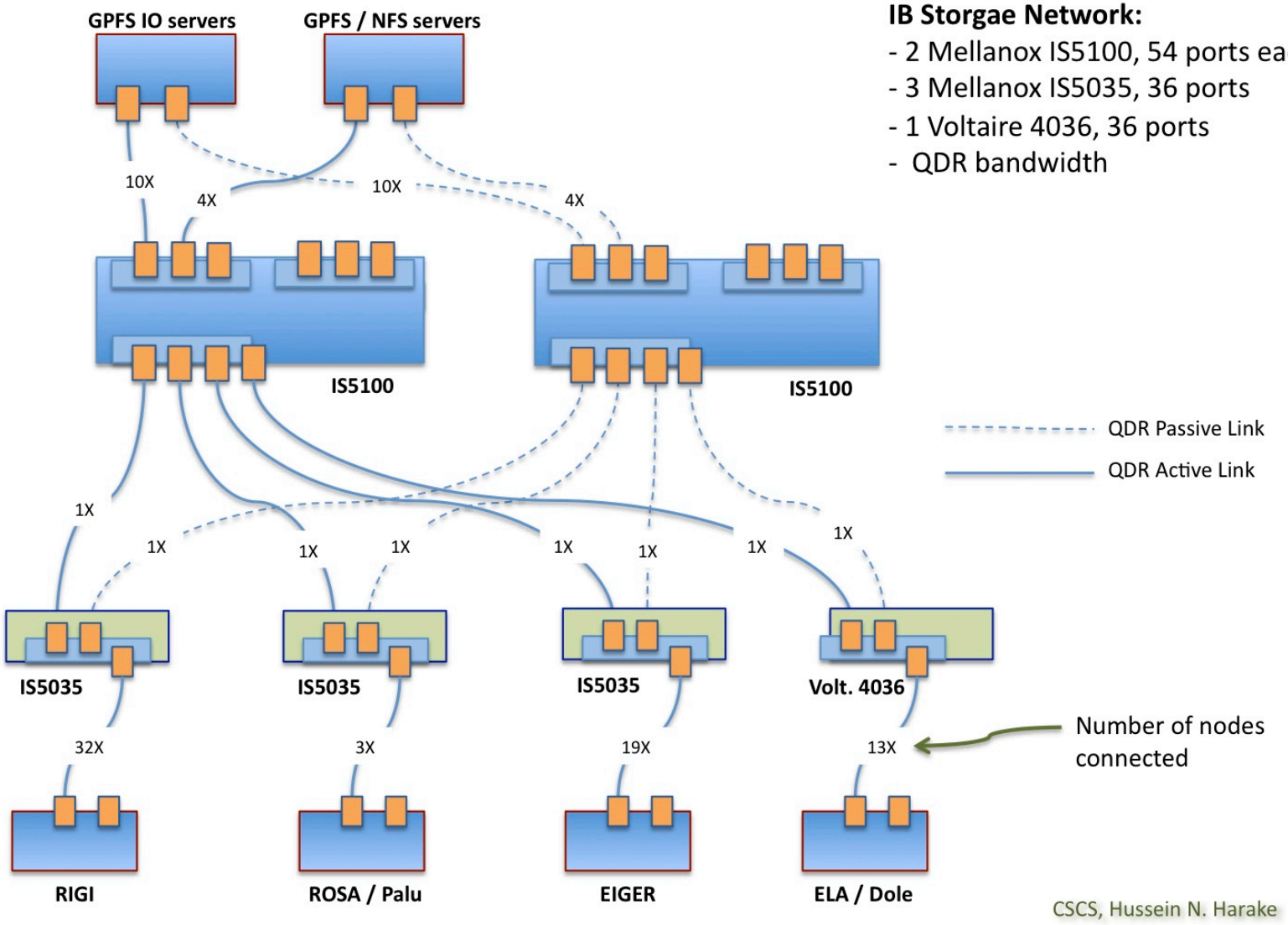
Infiniband resources

IS5100

- Production Systems

✓ 2 X IS5100 up to 108 ports each QDR
✓ 4 X IS5035 QDR 36 ports (Mellanox)
✓ 2 X 4036 Voltaire Switches 36 Ports QDR
✓ 2 X Mellanox Bridge BX4010 GW (IB – Ethernet – FC)
✓ 2 X Mellanox BridgeX BX5020 GW (IB – Ethernet – FC)
✓ SUN DS648 (M9) up to 648 Switch

Infiniband is used for storage and computing



IB Storgae Network:

- 2 Mellanox IS5100, 54 ports each
- 3 Mellanox IS5035, 36 ports
- 1 Voltaire 4036, 36 ports
- QDR bandwidth

CSCS, Hussein N. Harake

/project by the end of 2010

- ✓Adding 1 X DS5300
- ✓Converting the DS5100 to DS5300
- ✓8 X HD trays 5060 for each CTL.
- ✓4 X EXP5000 FC Disks
- ✓Increasing capacity to 2.1PB

EXP 5060

- •Performance improvement up to 11G/Byte SUSTAIN
- •2 X 8 G/bit dual FC card on each IO server
- •Raid6, snapshots and disaster recovery solution
- •Keeping a balanced system