

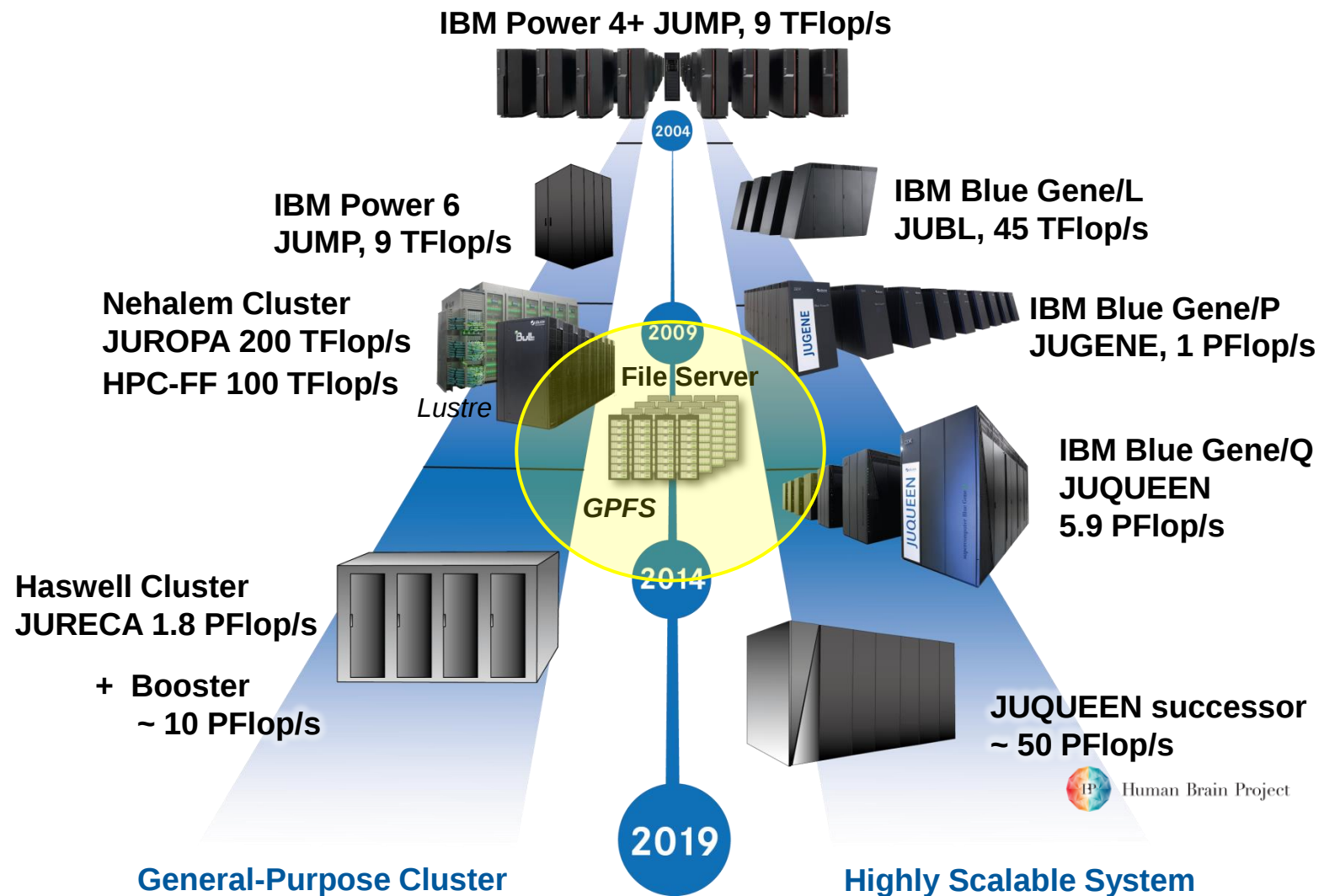


JUST JUelich STorage

St. Graf, T. Kondylis, M. Lischewski, U. Schmidt, L. Wollschläger

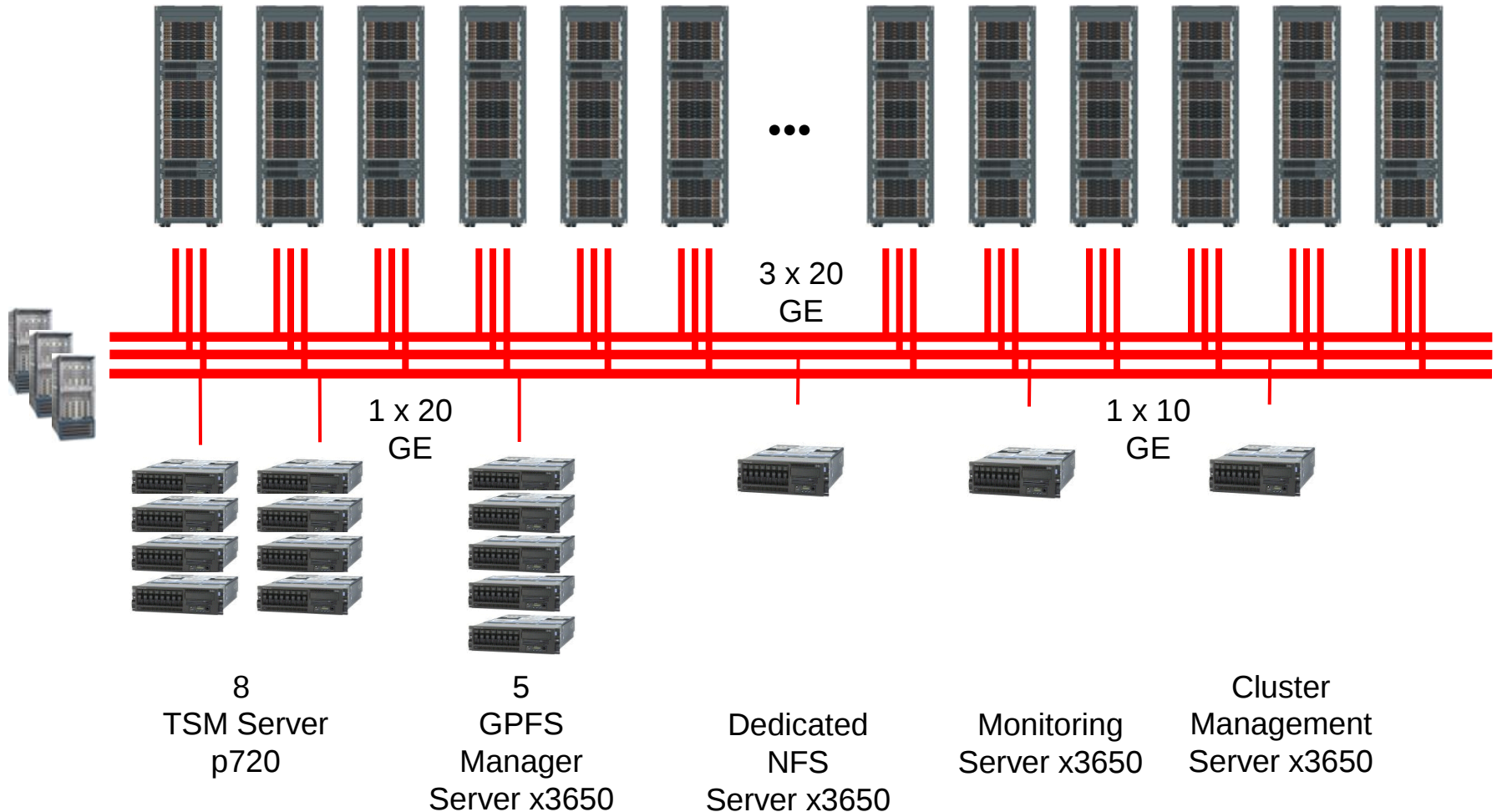
GPFS Production Cluster

HPC Systems: Dual Architecture Strategy



JUST – Physical View (Q1 2016)

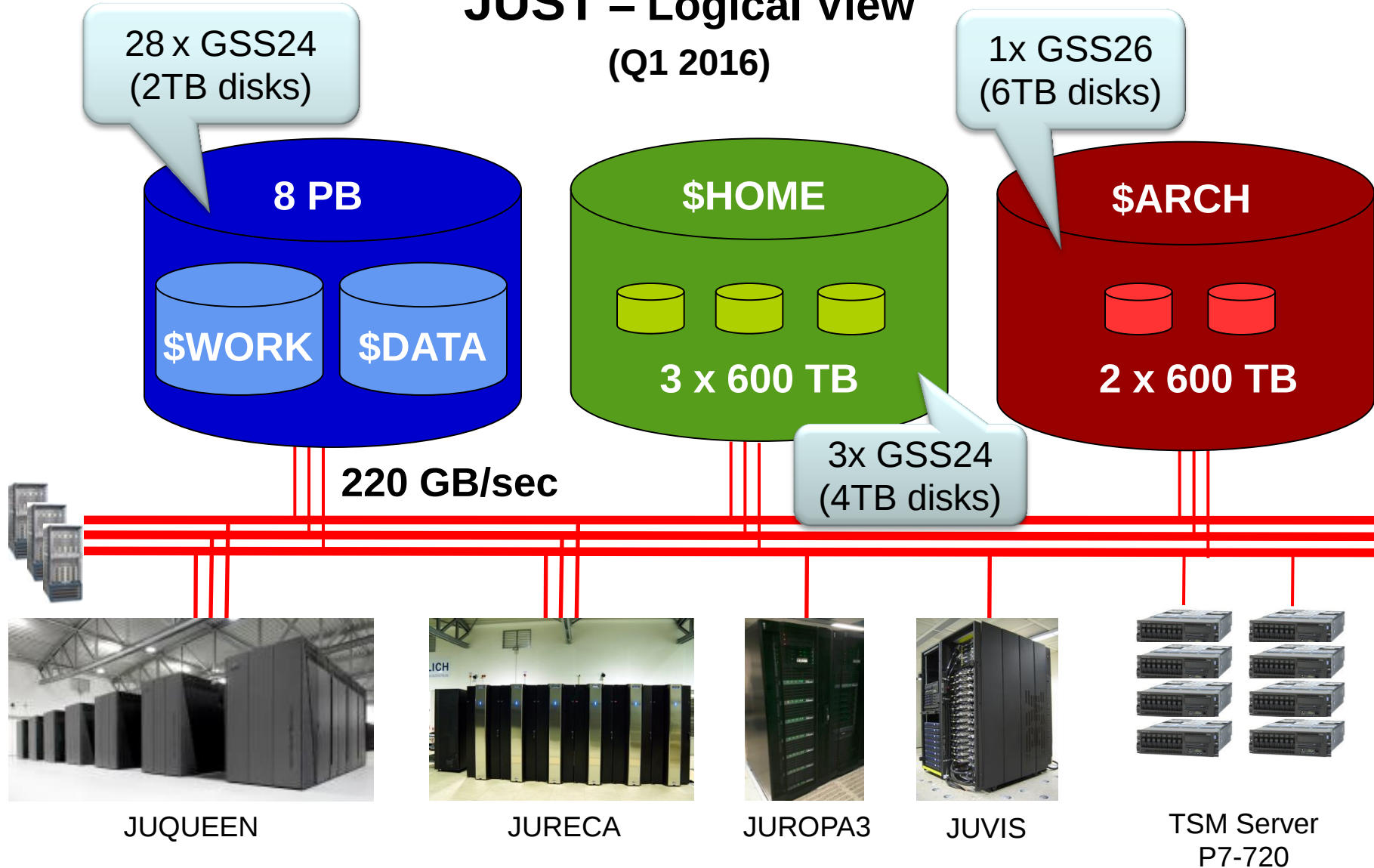
31 x GSS-24 + 1 x GSS-26 Systems → 64 x x3650 NSD Server, 136 x EXP3700 Storage



JUST4 Summary of Components

- GSS/GPFS part
 - Capacity: 19.8 PB (15.9 PB net)
 - 33x2 GSS server + 8 manager server
 - 7888 (+ 66 SSD)
 - 19 racks
- TSM part
 - Capacity: 450 TB (320 TB net)
 - 8 server + 2 manager server
- 1 Cluster Monitoring server
 - Icinga1 + check_mk
Mail notification, automatic RCMS calls for disks
 - Ganglia monitoring

JUST – Logical View (Q1 2016)



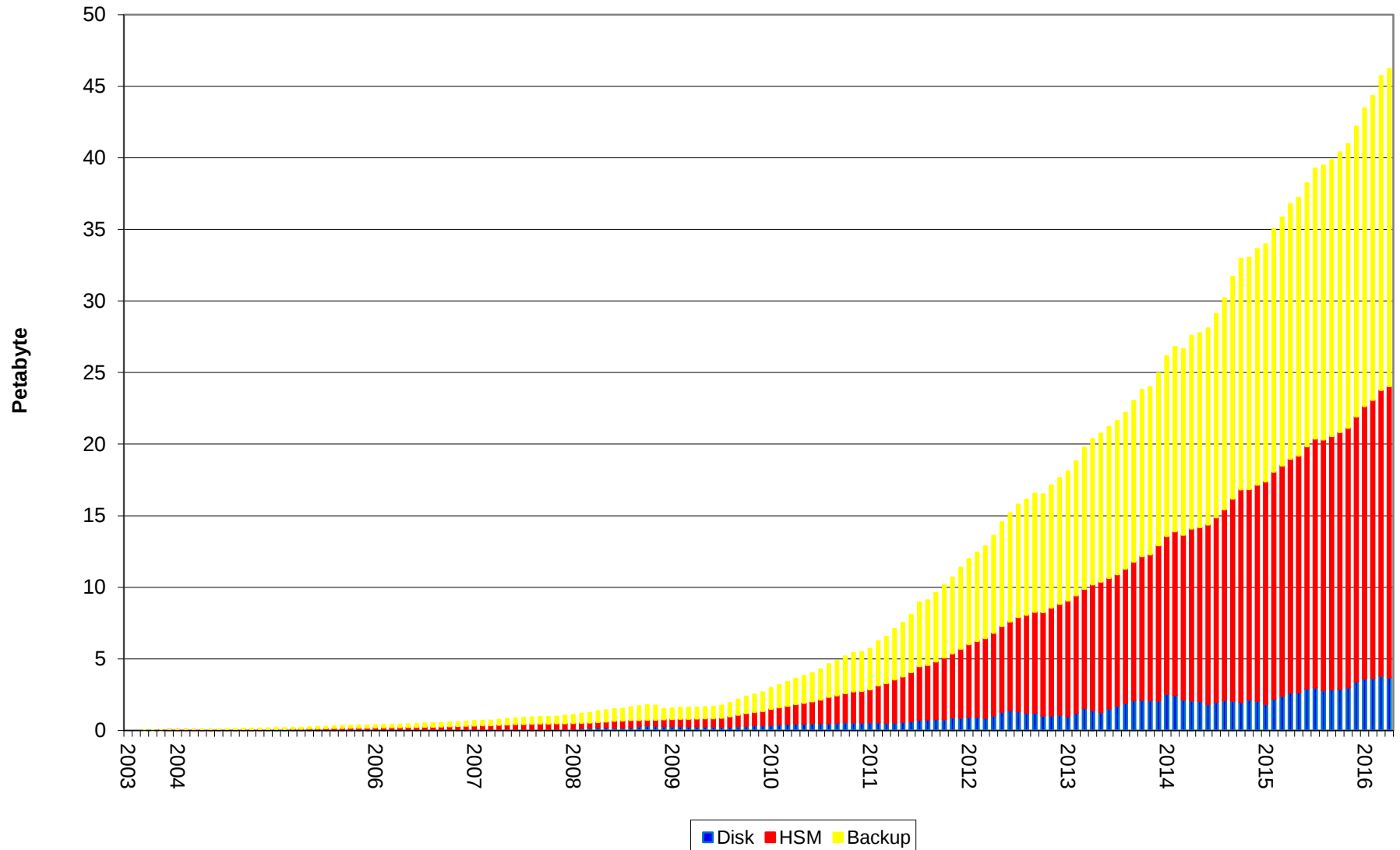
File System	Description	Characteristics
homea homeb homec	Users HOME File Systems	Group Quota: 6TB/2Mio Files Files in Backup
arch arch2	Users Archive File System	Group Quota: 2Mio Files Files in Backup Migration to Tape
work	Scratch File System	Group Quota: 20TB/4Mio Files No Backup Files deleted after 90 Days (last access) Empty directories after 3 days
data	Big data based File System For special Projects only Approval by JSC Management	Group Quota: 1GB/300 Files * Files in Backup Quota Info only for approved Project Groups

* For administrative functions only

File System Usage (April 2016)

File system	Capacity	used	Inodes
/homea	600TB	385TB	~100 Mio
/homeb	600TB	400TB	~125 Mio
/homec	600TB	389TB	~120 Mio
/work	5.3PB	3.3PB	~ 97 Mio
/data	2.7PB	2.1PB	~ 23 Mio
/arch	600TB	232TB (HSM: 10PB)	~ 52 Mio
/arch2	600TB	231TB (HSM: 8PB)	~28 Mio
6x small FS (/usr/local)	Σ 55TB	Σ 6TB	13 Mio (max)
Σ	~11PB	~7PB	> 550 Mio

Userdata GPFS



GPFS Configuration (incomplete)

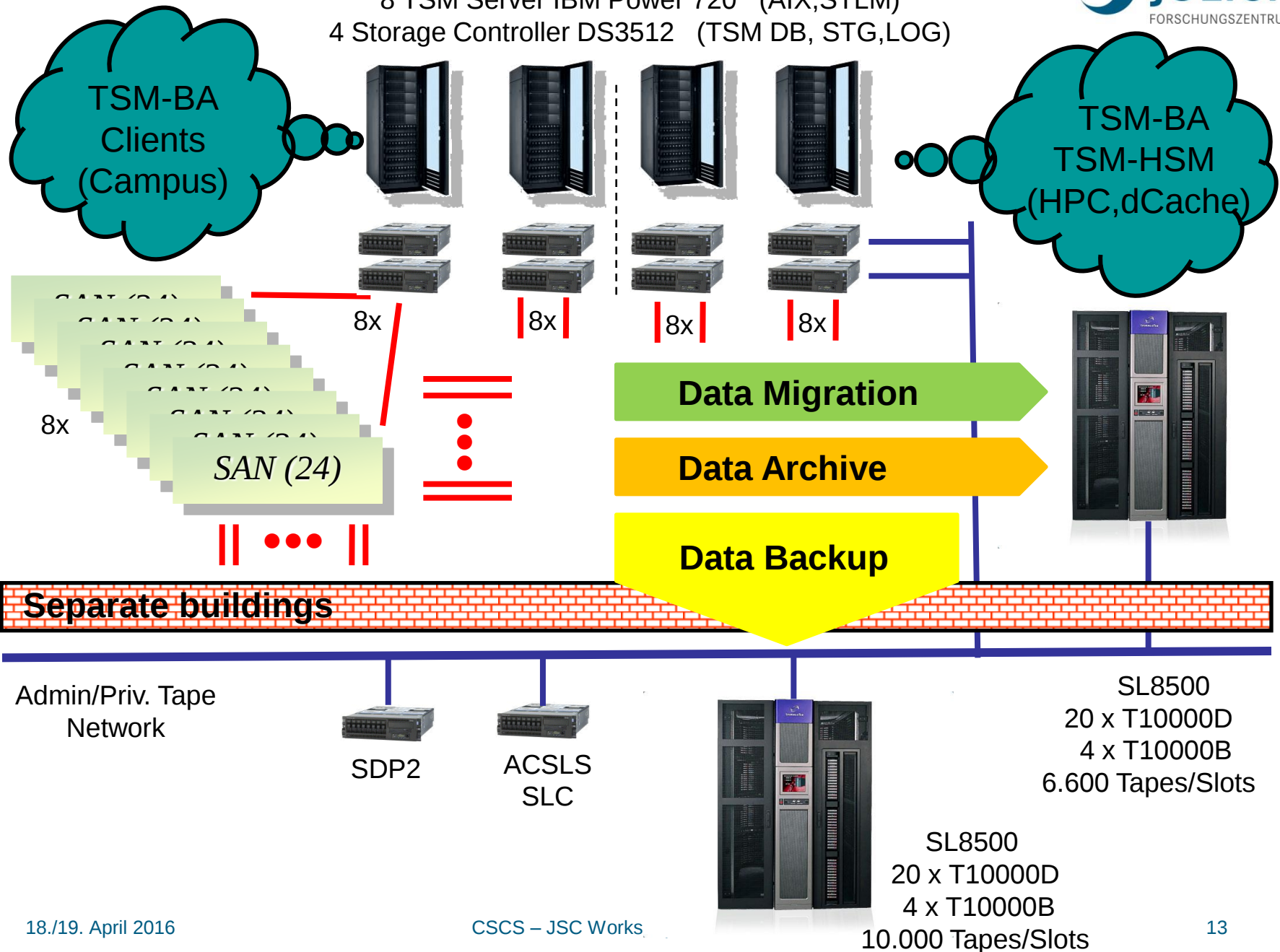
- JUST Cluster
 - GSS 2.0 (GPFS 4.1.0.8 EFIX 32)
 - 1 separate GPFS cluster for community services (GSS26)
- 10 Remote Cluster
 - GPFS 4.1.0.8 (JURECA, DEEP,)
 - GPFS 3.5.0.17 (BG/Q , waiting for version 4.1)
JUVIS, JUAMS
- Bandwidth Performance:
 - JUQUEEN: 193 GB/s write, 140 GB/s read
(target: 200 GB/s)
 - JURECA: 100 GB/s (target: 120 GB/s)

Upcoming events

- New ESS 26 integration
- JUQUEEN GPFS upgrade
- GSS 3.0 upgrade (incl. GPFS 4.2)
 - xCat Prepararion
 - RHEL 6.X -> RHEL 7.?
- Project for new usage model
 - USER centric – PROJECT centric (old)
→ Target: One account per user
 - Data projects + Research projects
 - Test environment is in place

HSM & Backup

8 TSM Server IBM Power 720 (AIX,STLM)
4 Storage Controller DS3512 (TSM DB, STG,LOG)



TSM Resources

- 2x Oracle STK SL8500 with actual 100 PB capacity (max. 10.000 + 6.600 tapes (x 8.5 TB = 141 PB))
- Library resource is only accessible by TSM instances
 - 1x TSM shared library manager
 - 20x TSM server (on 8 AIX systems)
- 4x3 TSM server instance managed by IT-S
 - Backup for FZJ desktop & server systems
- TSM server managed by JSC
 - HPC user backup data
 - HSM managed GPFS file systems
 - Special community services (e.g. dCache)

GPFS User Data Backup

- TSM native **dsmc incr** command is too slow
Backup GPFS file systems > 24h, SLA not achieved
=> IBM solution: use **mmbackup**
- mmbackup concept:
 - Get list of backed up files from TSM Server or Shadow DB
 - Create list of files from the file system using the very fast GPFS policy engine
 - Compare both lists and create
 - *backup candidate list*
 - *expired file list*
 - Execute 'dsmc' command using a file list
 - Create/update Shadow DB

Upcoming events

- Prepare migration of TSM Server from AIX to Linux (2016) and migrate all server (2016/17)
- View at TSM Operation Center
- Upgrade to TSM 7.1.X (2017)
- Cooperation with RWTH Aachen (Node replication)
- Archive concept for FZJ (IT-S)
- Project: Scientific Data Management (FZJ+external)