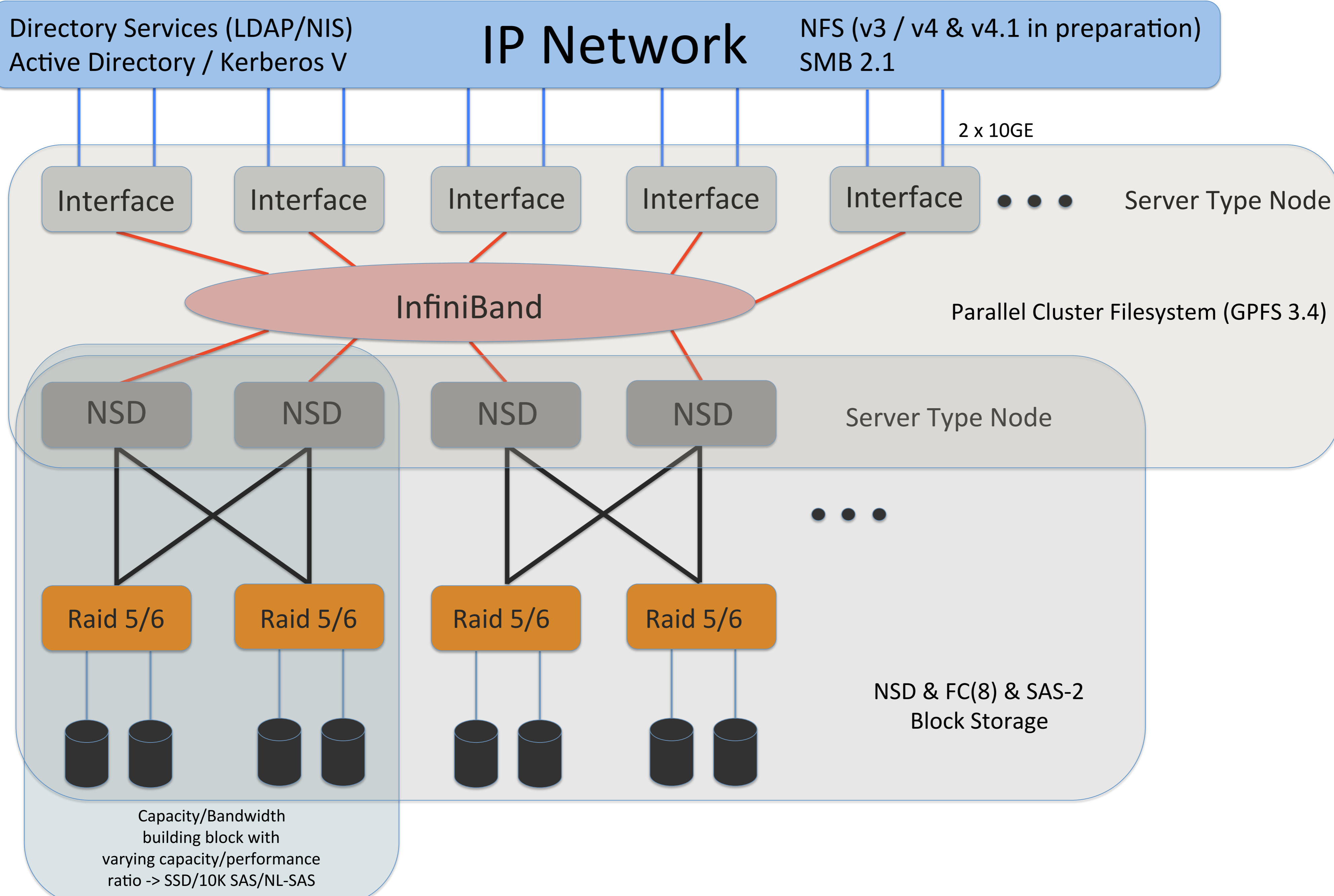


# Applicability of modern, Scale-Out file services in dedicated LHC data analysis environments.

M. Gasthuber, Y. Kemp DESY/Hamburg

## Basic structure of the system



## Specification of installed System:

- optimized for IOPS – good for ~300.000 IOPS – SPECsfs-2008 shows >400.000 IOPS
- GPFS metadata all on SSDs
- ~500TB capacity (using SSD/300GB (26), 10K SAS/600GB (480), NL-SAS/1TB (220) & 3TB(48))
- 3-Tier-Level GPFS configuration – initial placement on fast SAS
- Storwize V7000 Raid-Controller
- 12 Interface Nodes – each with 2 x 10GE & 2 x IB
- 8 Storage Nodes (NSD) each with 4 x FC8 & 2 x IB – 4 building blocks

## User Perspective Demands (Analysis Facility)

- quota (user & (sub)group & directory)
- multilevel tiering with regular policy runs
  - file within certain subdirs and atime > XZ days will be moved to slow disks/deleted
  - very hot files (and reasonable filesize) moved to SSDs
- detailed access reporting – find dead/cold data
- detailed monitoring (history & current)
- individual performance optimization

## Ongoing activities (partnership with IBM)

- GPFS cross cluster – allows native (concurrent) GPFS access with native GPFS client
- cloud storage access – WSS (Object Store) – write once/read many – SNIA's CDMI, RESTful HTTPS protocol
- NFS Ganesha tracking for early v4 and v4.1 usage – thru Cross-Cluster-Mount
- User initiated/System controlled tertiary storage interface with container based operations efficiently handling plethora of small objects/files.
- JBOD building blocks, no HW Raid – GPFS managed JBOD block storage (Perseus) tracking
- GridFTP service (using Cross-Cluster-Mount) – Grid Jobs can write data directly

## Conclusion

- Replacement for Lustre in National Analysis Facility in Hamburg as a central, shared and fast file repository with tertiary storage access.
- Usage planned for other areas requiring standard file access (VFS/Posix)
- Local development withing national LSDMA project for the next 4 years
- early bird usage ongoing – detailed benchmarking will follow