Karlsruhe Institute of Technology

Experiences with different HP SFS/Lustre systems at KIT/SCC

Steinbuch Centre for Computing (SCC)
Roland Laifer

Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft

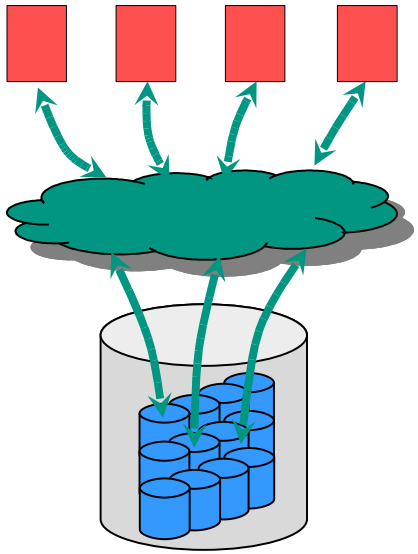
Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Overview

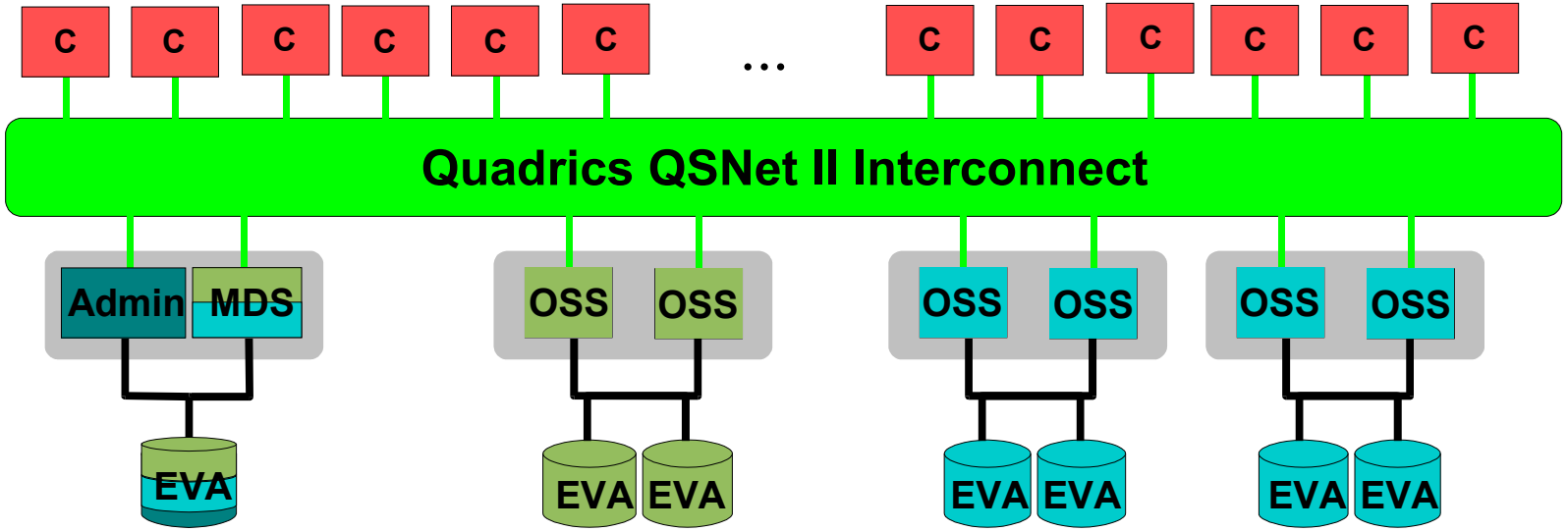
- Parallel file system basics
- HP SFS/Lustre systems at KIT/SCC
- SFS version 2 versus SFS G3
- Experiences with MSA2000
- General observations

Parallel file system basics

- What is a parallel file system?
  - Distributed file system with parallel data paths from clients to disks
    - No need for applications to be aware of these different paths
  - Applications on all clients typically see one consistent single namespace
- Why using a parallel file system?
  - Might be required if applications need any of these features:
    - Single namespace for 100‘s of clients
    - 10‘s of TB of data
    - Multiple GB/s throughput
    - 10‘s of millions of files

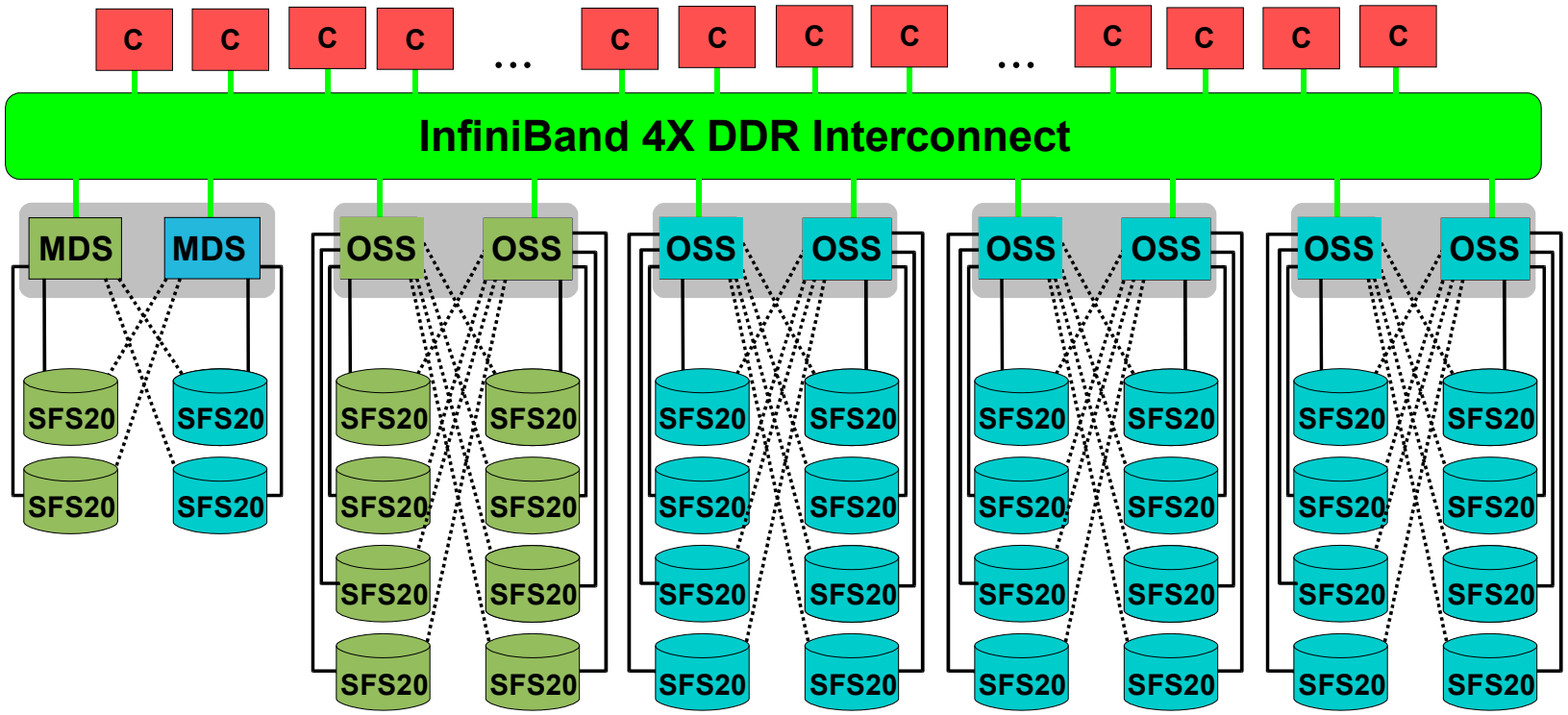


Itanium system (xc1)



- HP SFS on xc1 runs extremely stable
- Good experiences with home directories in Lustre since 2005
- System will be replaced soon

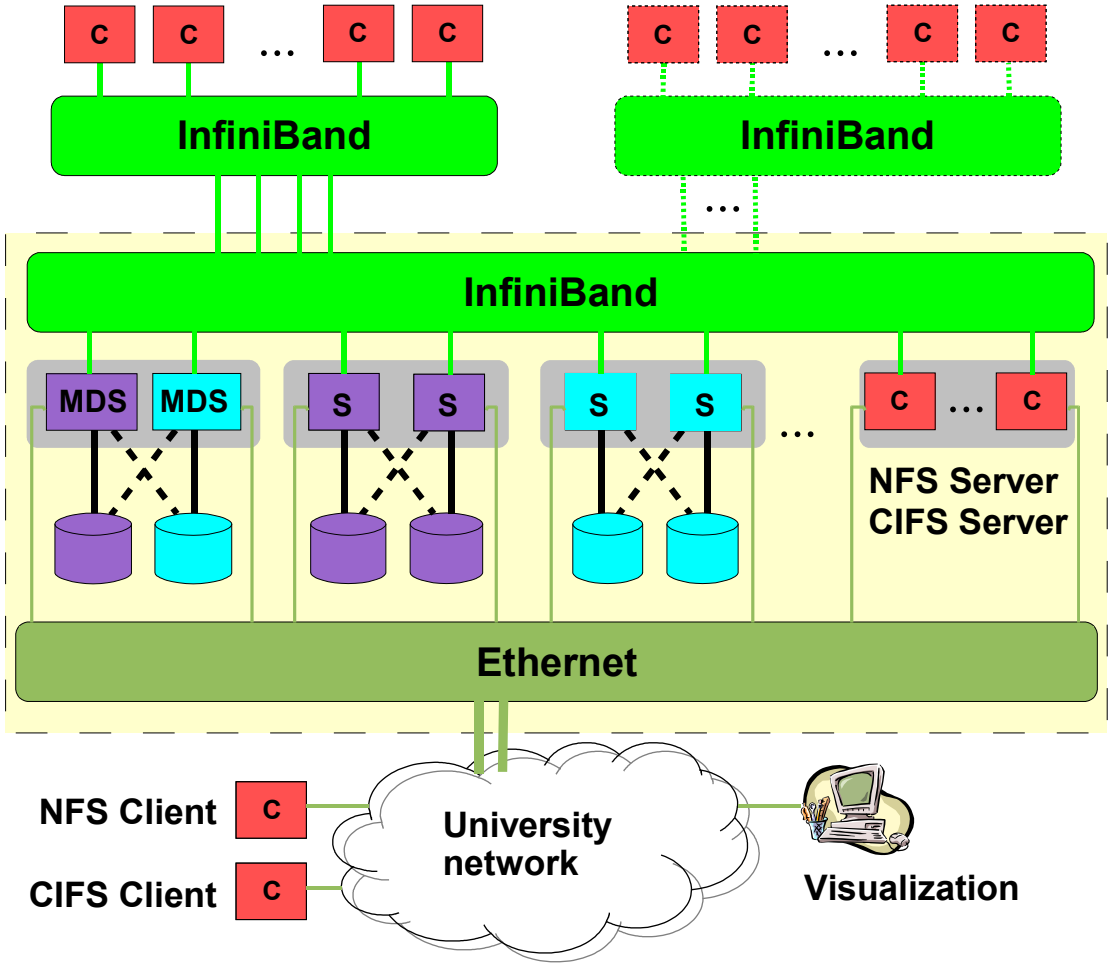
Opteron system (xc2)



- HP SFS on xc2 runs very stable
- Very high I/O requirements of different applications

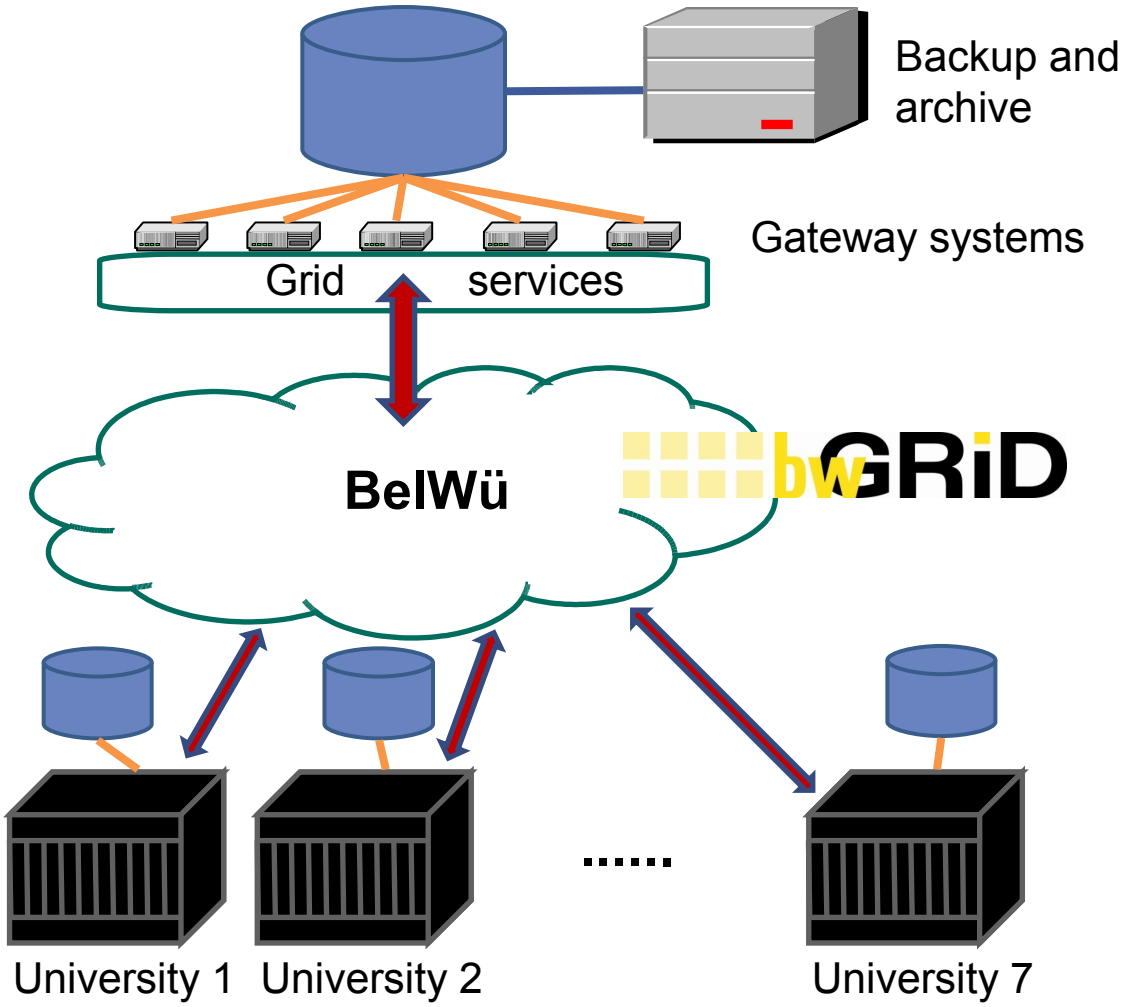
Central parallel file system (pfs)

- Additional clusters will be attached
- Virtual machines currently use NFS gateways



bwGRiD storage system (bwfs)

- Grid middleware for user access and data exchange



Lustre systems at KIT

System name	xc1	xc2	pfs	bwfs
Users	University of Karlsruhe	universities, industry	departments, multiple clusters	universities, grid communities
Lustre version	HP SFS 2.1-1	HP SFS 2.1-1	Lustre 1.6.4.3	HP SFS G3.0-0
# of clients	120	762	>200	>1400
# of file systems	2	2	2	9
Capacity (TB)	3.8 + 7.6	16 + 48	73 + 304	4*32 + 3*64 + 128 + 256
Throughput (GB/s)	0.24 + 0.48	0.7 + 2.1	1.8 + 6.0	8*1.5 + 3.5
Storage hardware	HP EVA5000	HP SFS20	transtec provigo	HP MSA2000
# of enclosures	14	36	62	138
# of disks	168	432	992	1656

SFS version 2 versus SFS G3

- Different strategy:
  - Version 2 uses old Lustre version with lots of backported patches
  - G3 uses newest Lustre version
- Advantages of SFS G3:
  - New features become available without delay
    - E.g. patchless client
  - No backporting by HP required
  - Fixing Lustre bugs can be forwarded to Sun
  - Better price (hopefully)

SFS version 2 versus SFS G3 (cont.)

- Current limitations of SFS G3 compared to version 2:
  - Stability has not improved
  - Increased management overhead:
    - No diskless installation of OSS
    - System database for easy migration no longer available
  - Currently missing features:
    - Monitoring and system check (sfsmgr syscheck)
    - Email alerts
    - Easy and enhanced file system check
    - Excellent documentation, e.g. for replacement of hardware parts
    - Preconfigured crashdump

Experiences with MSA2000

- Advantages compared to SFS20:
  - Improved performance
  - Higher capacity and hot spares
  - No single point of hardware failure and no batteries
- Current limitations:
  - Stability:
    - System might hang after a disk failure
    - Bugs required vdisk and OST recreation
  - Lots of recommended firmware upgrades
  - No good procedures for firmware upgrades:
    - Instructions not complete or changing
    - Each disk needs to be upgraded manually while system is offline

General observations

- Parallel file systems are the most critical cluster component
  - Underlying storage systems frequently cause most trouble
- I/O requirements often scale linearly with number of cores
  - Most applications are only scalable to some degree
    - Hence users are running different data sets in parallel
- Lustre provides a very good parallel file system
  - Software test including scalability by vendors is still required
  - Today most parallel file systems are Lustre installations
- Further information
  - http://www.rz.uni-karlsruhe.de/dienste/lustretalks
  - roland.laifer@kit.edu

Karlsruhe Institute of Technology

Thank you for your attention!

Steinbuch Centre for Computing (SCC)
Roland Laifer

Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825