Karlsruhe Institute of Technology

Performance of and experiences with Lustre over a long distance InfiniBand connection

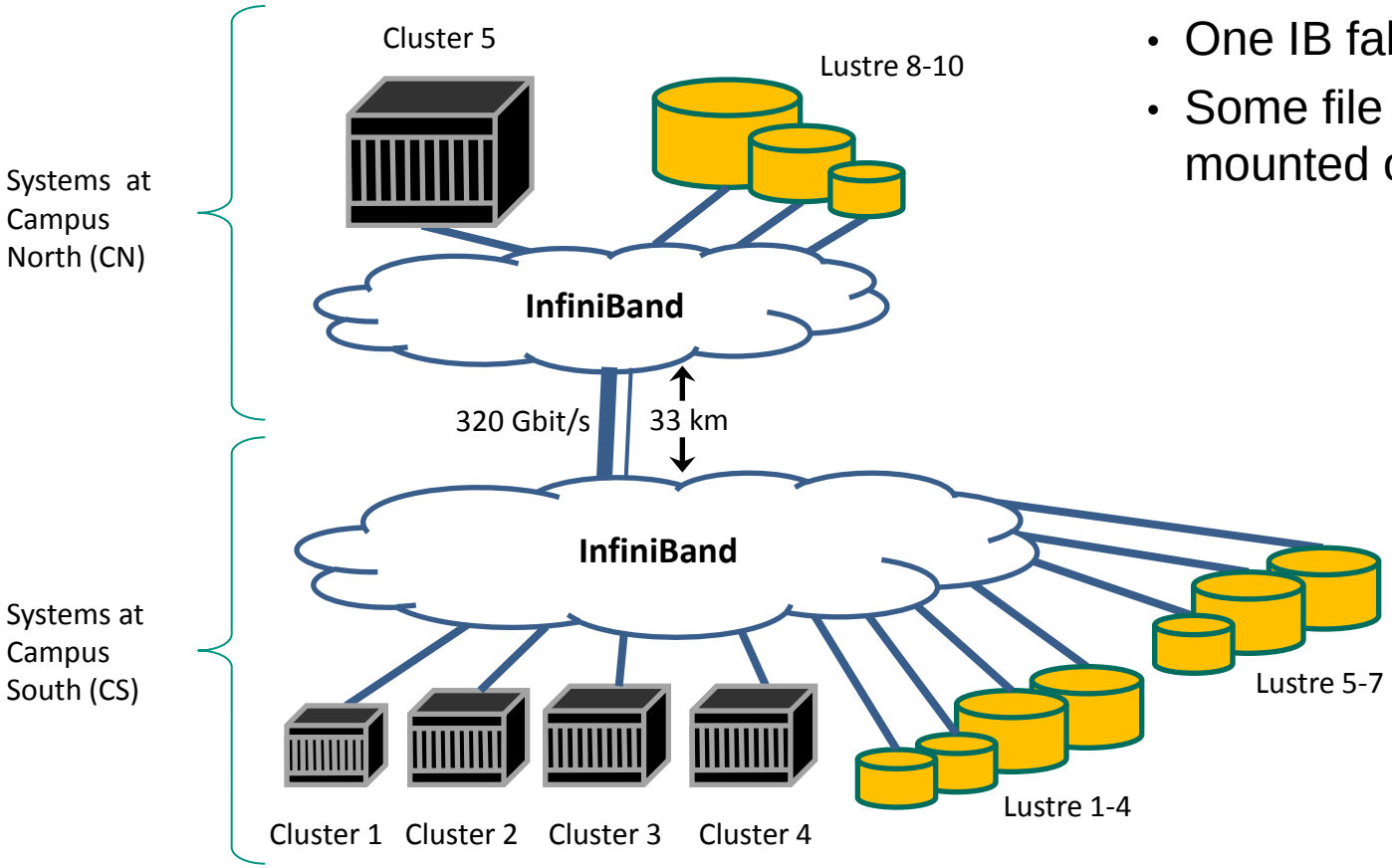
Roland Laifer

STEINBUCH CENTRE FOR COMPUTING - SCC

Overview

- Lustre systems at KIT
  - and details of our complex InfinBand fabric
- Investigation of Lustre related InfiniBand (IB) problems
  - based on two examples
- Lustre performance over long distance IB
  - for throughput and metatdata

Lustre systems at KIT - diagram



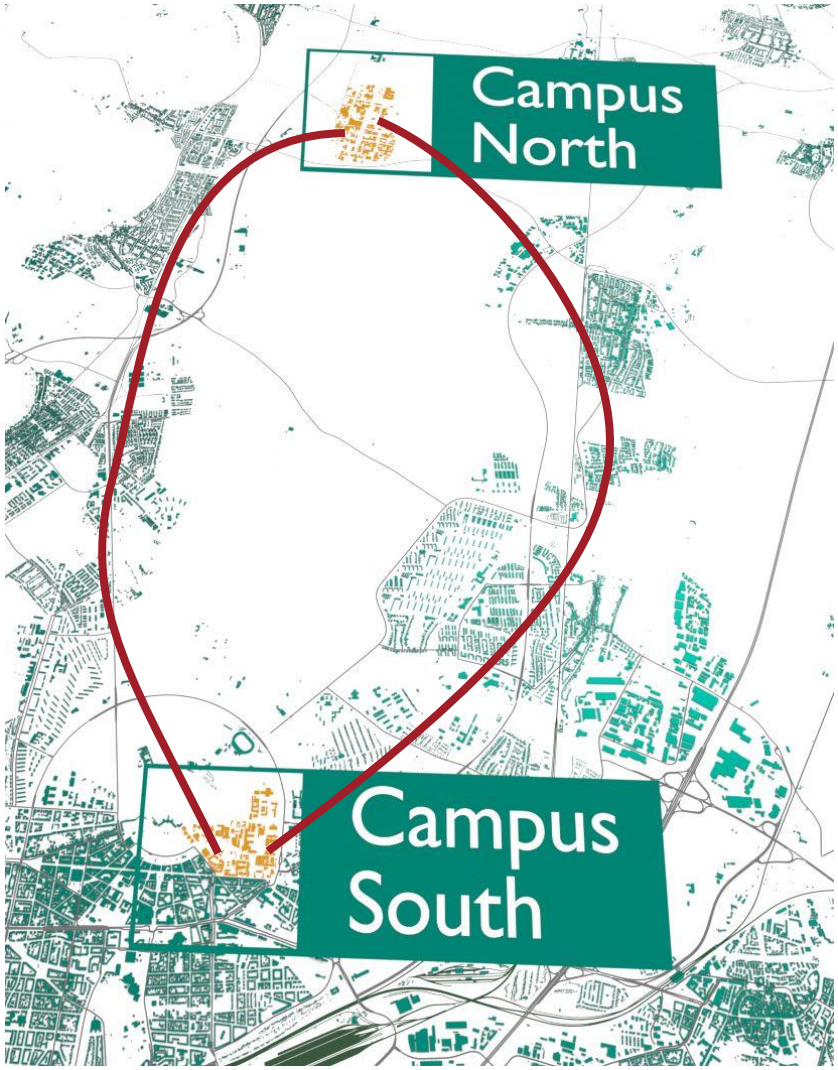
- One IB fabric
- Some file systems mounted on all clusters

Lustre systems at KIT - details

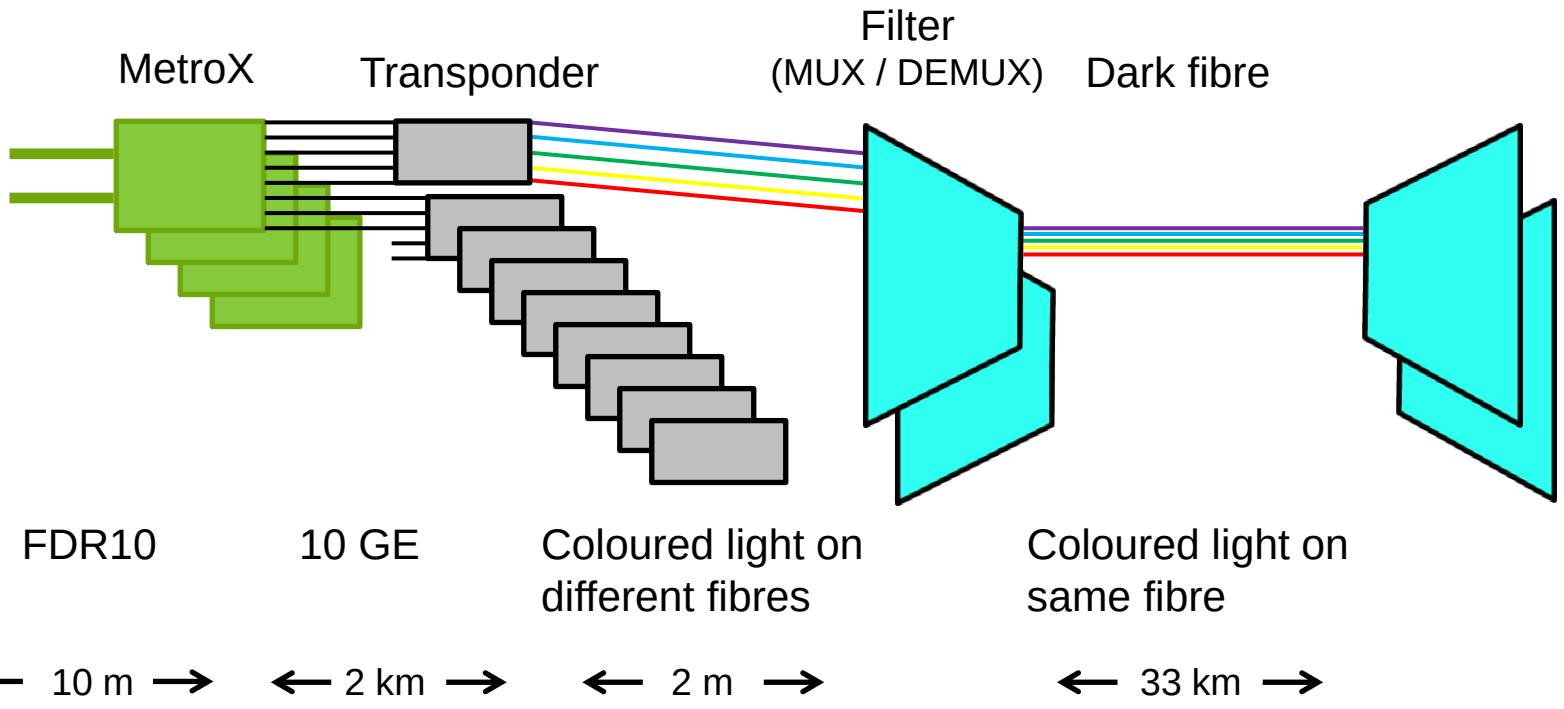
System name	pfs2	pfs3	pfs4
Users	universities, all clusters	universities, tier 2 cluster (phase 1)	universities, tier 2 cluster (phase 2)
Lustre server version	DDN Exascaler 2.3	DDN Exascaler 2.1	DDN Exascaler 2.3
# of clients	3100	540	1200
# of servers	21	17	23
# of file systems	4	3	3
# of OSTs	2*20, 2*40	1*20, 2*40	1*14, 1*28, 1*70
Capacity (TiB)	2*427, 2*853	1*427, 2*853	1*610, 1*1220, 1*3050
Throughput (GB/s)	2*8, 2*16	1*8, 2*16	1*10, 1*20, 1*50
Storage hardware	DDN SFA12K	DDN SFA12K	DDN ES7K
# of enclosures	20	20	16
# of disks	1200	1000	1120

Long distance connection - overview

- Two dark fibres for reliability
  - Allow transparent failover
  - IB failover same as with redundant switches
  - Length is 28 and 33 km
- Use the same dark fibres for different protocols
  - 10 GE, 100 GE, IB, FC
  - Underlying technology is dense wavelength division multiplexing (DWDM)
  - Spacing was reduced from 100 GHz to 50 GHz to provide additional channels for IB



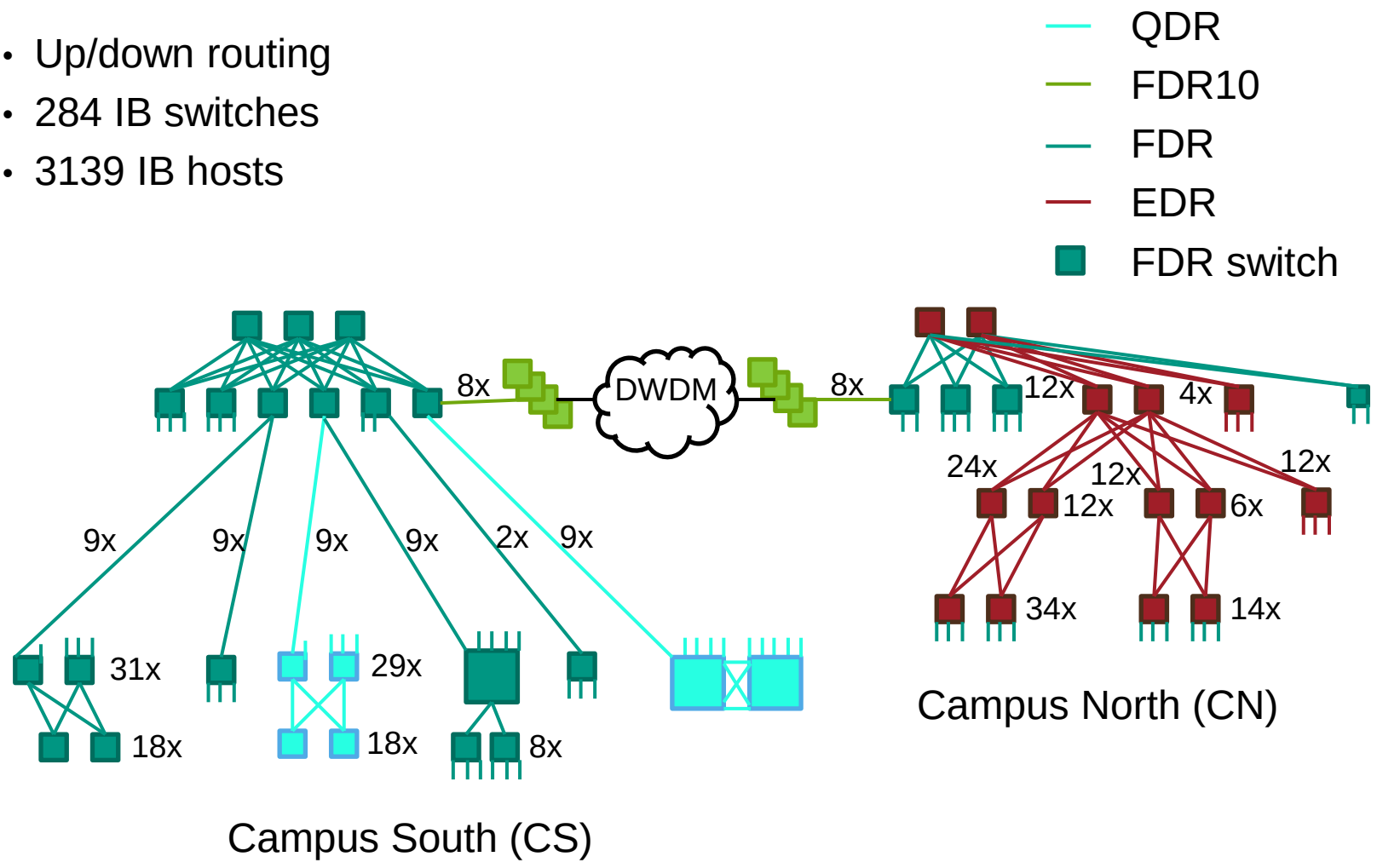
Long distance connection - details



- Mellanox MetroX IB switches
  - Special hardware with enough buffers to fill full length of dark fibre
  - Obsidian's Longbow is an alternative product

IB network details

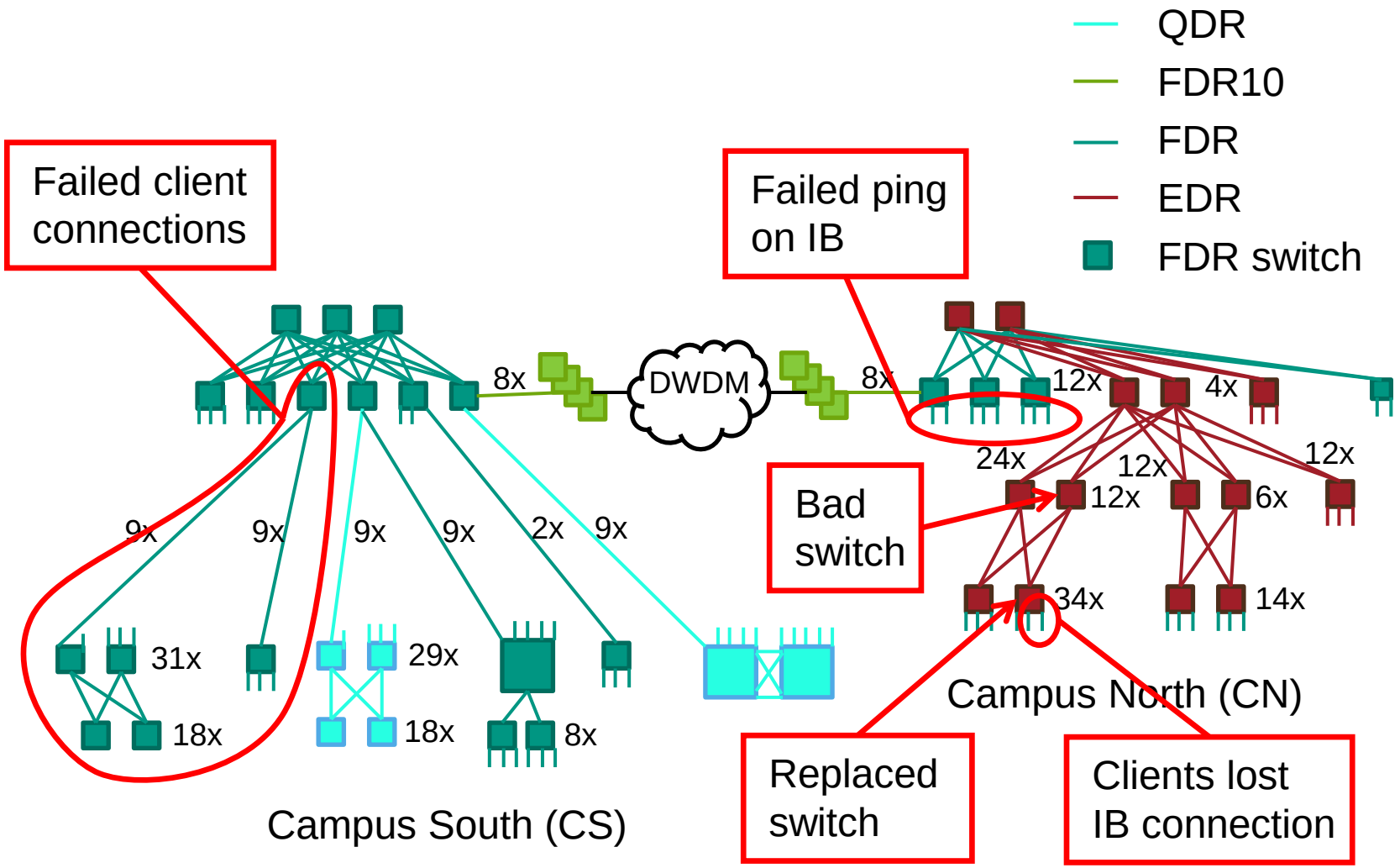
- Up/down routing
- 284 IB switches
- 3139 IB hosts



Ex. 1: Investigate Lustre connectivity problems

- Initial problem
  - Failover servers of all file systems reported new client connections
    - Obviously many clients had problems to reach active servers
  - New file system servers were failing over Lustre services
    - Reason was that ping on IB had failed (and bad configuration)
  - 2 clients on new cluster lost connection to everything on IB
    - Happened while IB throughput benchmark was running
- Further investigation
  - Problem disappeared after reboot of the switch with the bad clients
  - Problem was reproducible
    - Even with running the benchmark on few clients (including the 2 bad)
  - IB subnet manager showed healthy fabric
  - Replacing the EDR switch with the 2 bad clients did not help

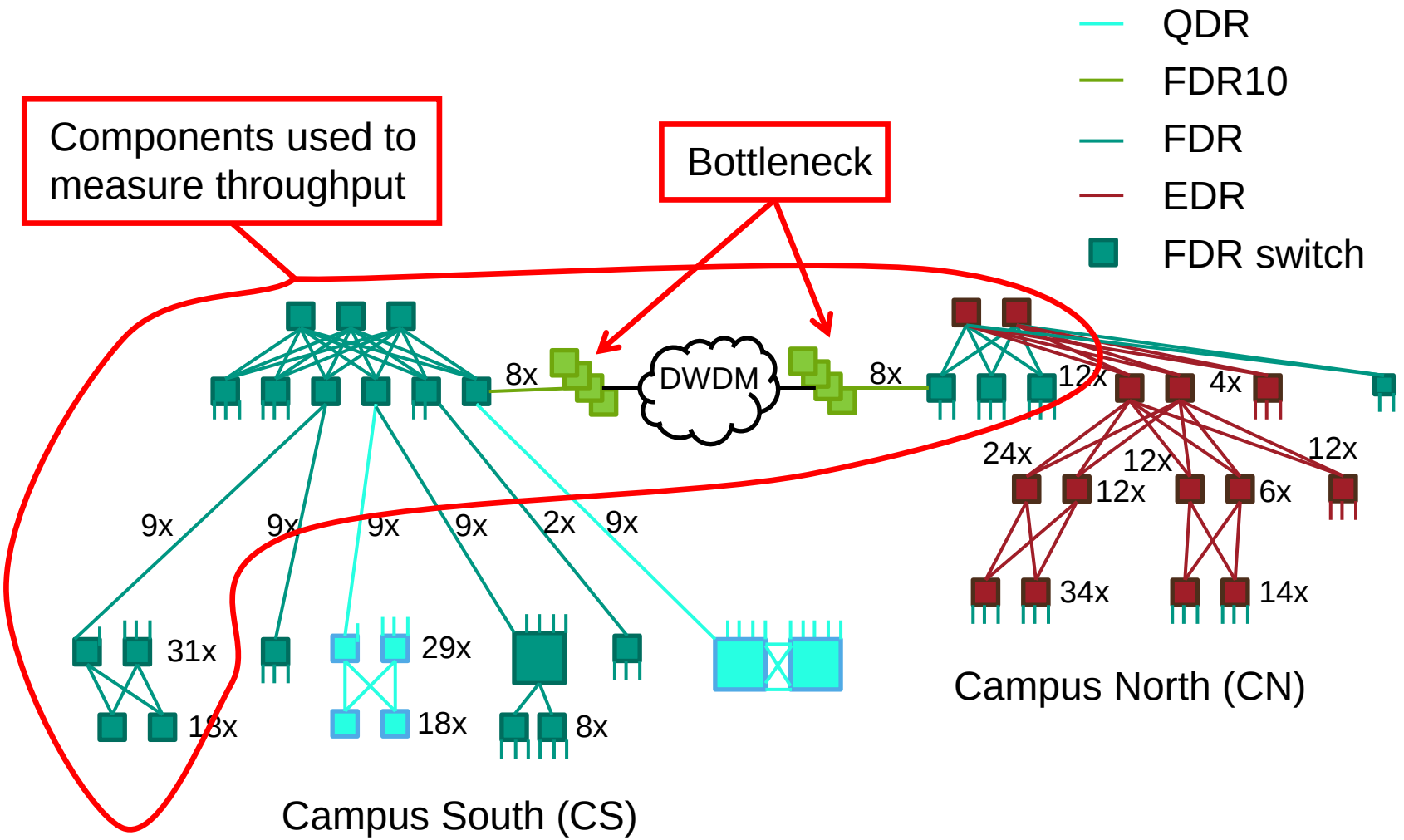
Ex. 1: Investigate Lustre connectivity problems



Ex. 1: Investigate Lustre connectivity problems

- Solution
  - Routes to 2 bad clients used the same port on core EDR switch
    - Replacing that switch fixed the problem
  - Do not ignore problems on few bad clients
  - Investigate routing to check which components are shared
- *Possible* root cause
  - Management communication still worked
    - Therefore no new route assignment by subnet manager
    - Maybe subnet manager used alternative connection
  - Data communication on bad EDR switch port was blocked
    - Possibly caused backlog and full switch buffers on other switches
  - Cascading blocked ports on complete fabric might be possible
  - We do not know how such issues could be clearly investigated

Ex. 2: Investigate Lustre performance problems



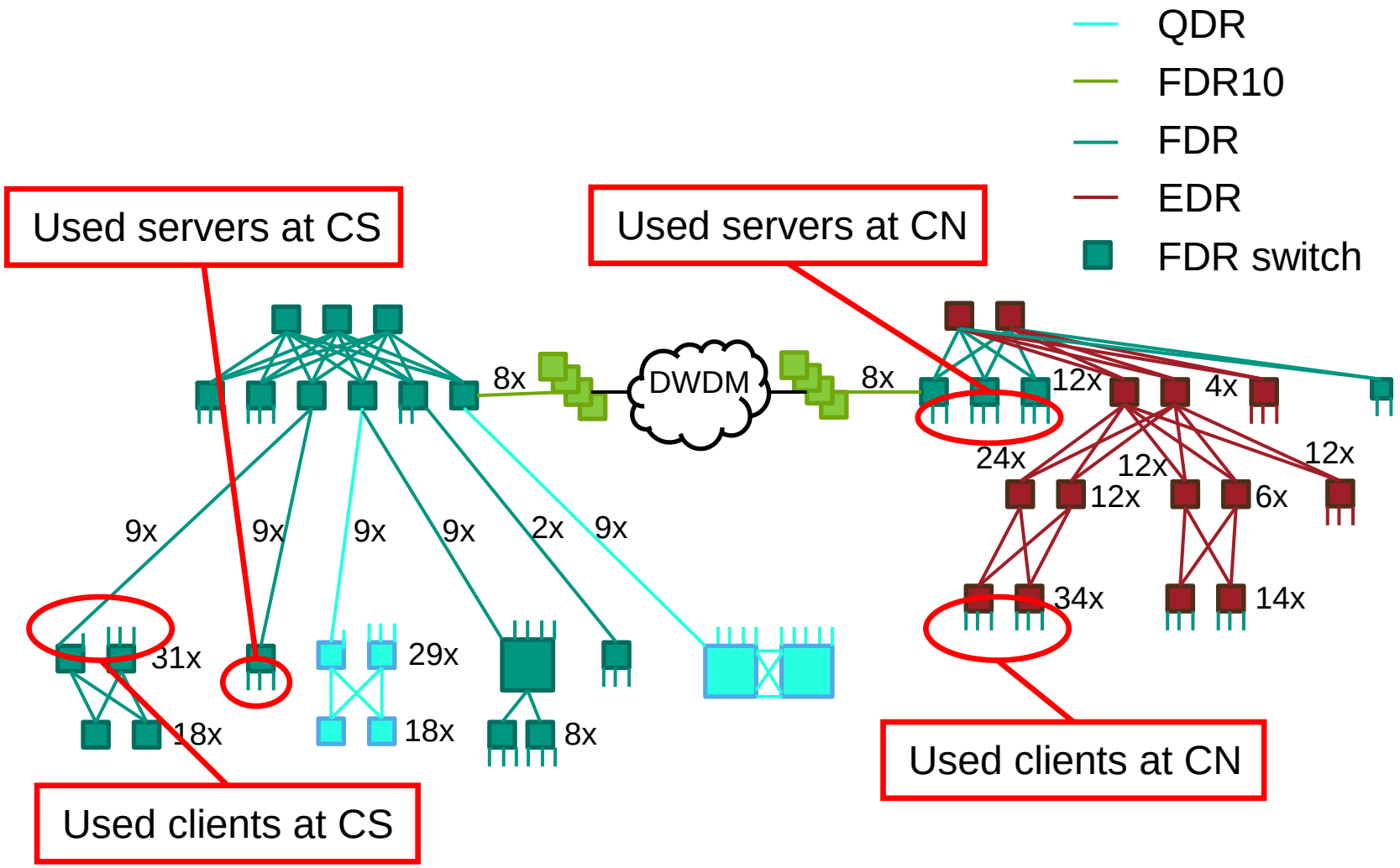
Ex. 2: Investigate Lustre performance problems

- Insufficient performance with iozone over long distance IB
  - On 50 GB/s file system only reached 22 GB/s
    - Peak IB bandwidth is 8 * FDR10 = 38 GB/s
    - With one FDR10 connection reached 4.6 GB/s which is good
- Investigate IB network topology
  - MetroX have 8 FDR10 connections but file system has 10 OSS
    - IB has static routing, i.e. some OSS share same FDR10 connection
    - Lustre evenly distributes files to OSTs, i.e. performance per OSS is half of FDR10
  - Peak with 10 OSS is FDR10 / 2 * 10 = 24 GB/s
    - Measured 22 GB/s are good

Ex. 2: Investigate Lustre performance problems

- Double check
  - With ibtracert checked which OSS used same FDR10 cable
    - MDS used one cable, i.e. only 7 connections were used by OSS
  - Created directories with different OST stripe index for iozone
    - Assigned only half number of files to OSTs with shared connections
  - Measured 29.2 GB/s with 7 connections
    - Would be 33.4 GB/s with 8 connections, i.e. 12% below peak

Performance measurement details

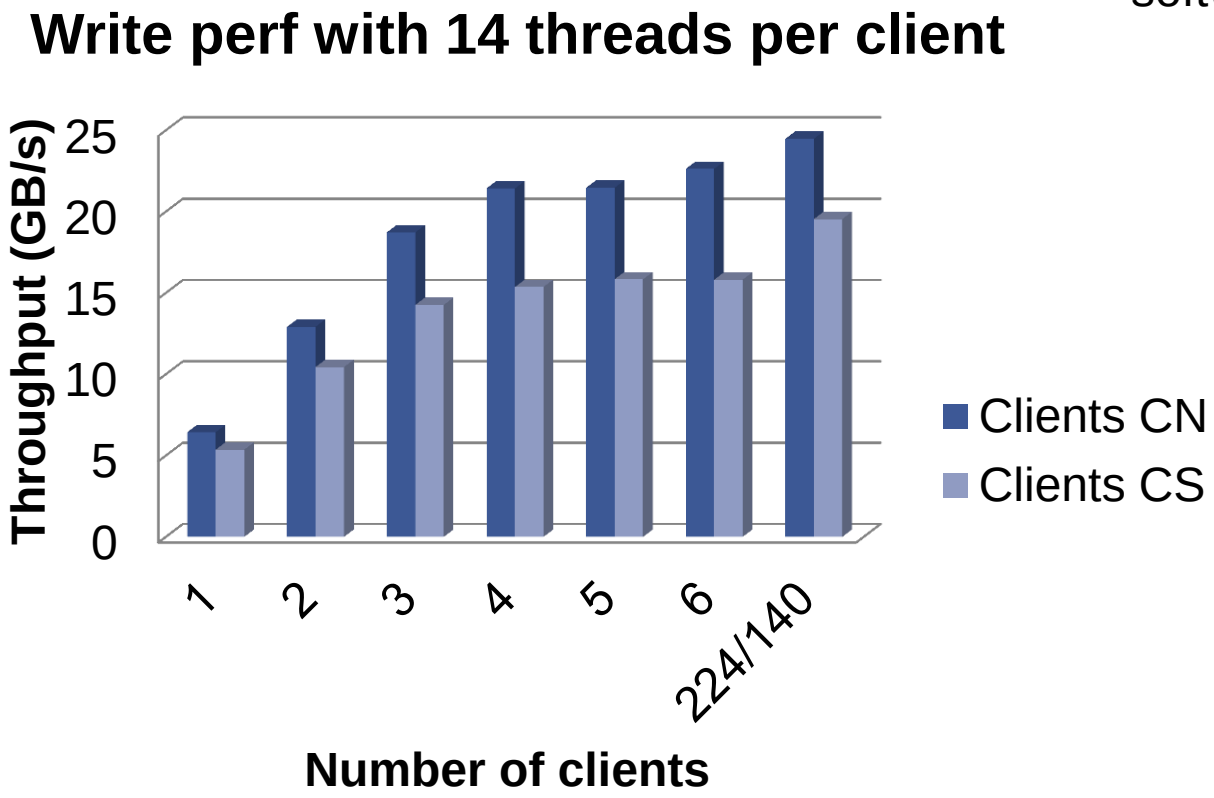


Performance measurement details

- Done while some of the systems were in production
  - Just show trends, no focus on best performance
- Write performance measured with iozone
  - Options: -+m <file_name> -i 0 -+n -r 1024k -t <thread_count> -s 8g
- Metadata performance measured with mdtest
  - Options: -u -n 10000 -i 3 -p 10 -d <lustre_dir>
- Used clients
  - CN: RH7, Mellanox OFED, FDR Connect-IB, Exascaler 2.3
  - CS: RH6, RH OFED, FDR ConnectX-3, Exascaler 2.1
- Used file systems
  - CN: EF4024 (MDT), 28 OSTs on ES7K, 6 TB disks, Exascaler 2.3
  - CS: EF3015 (MDT), 40 OSTs on SFA12K, 3 TB disks, Exasc. 2.1

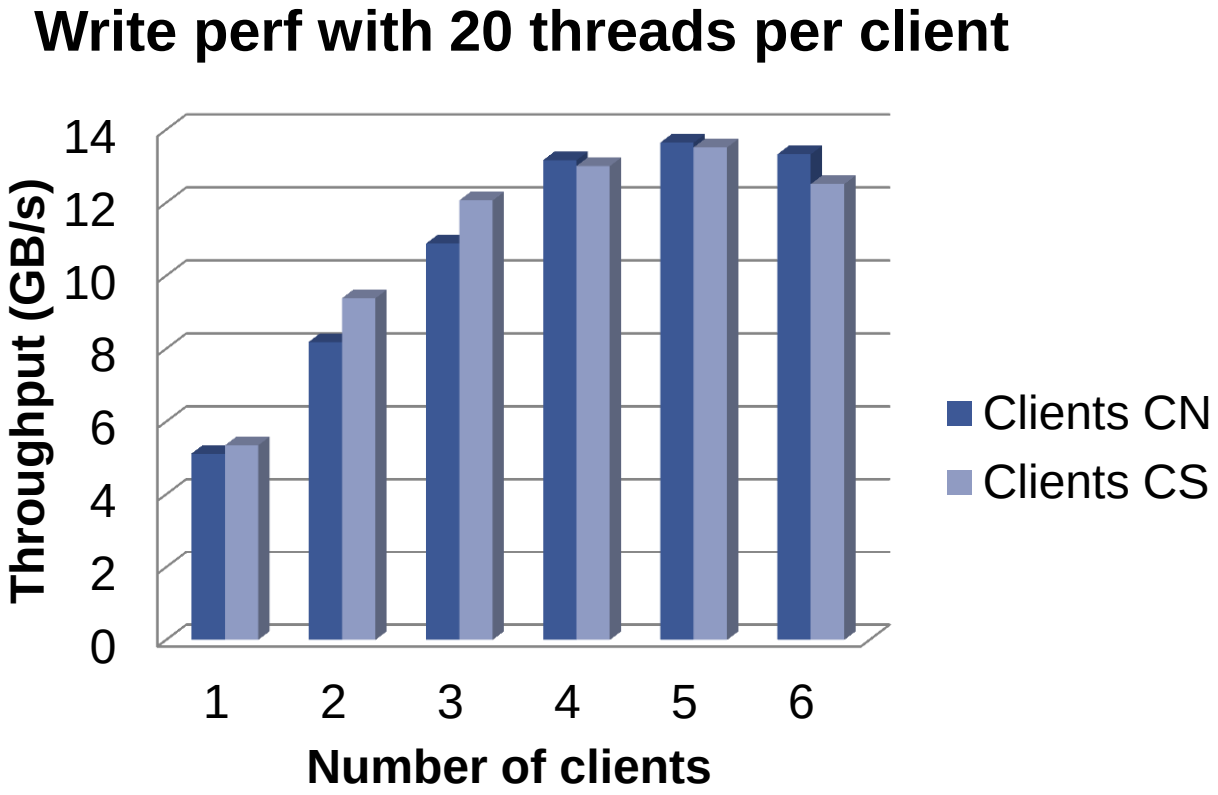
Write performance to file system at CN

- Remote CS clients on older hardware and software are slightly slower

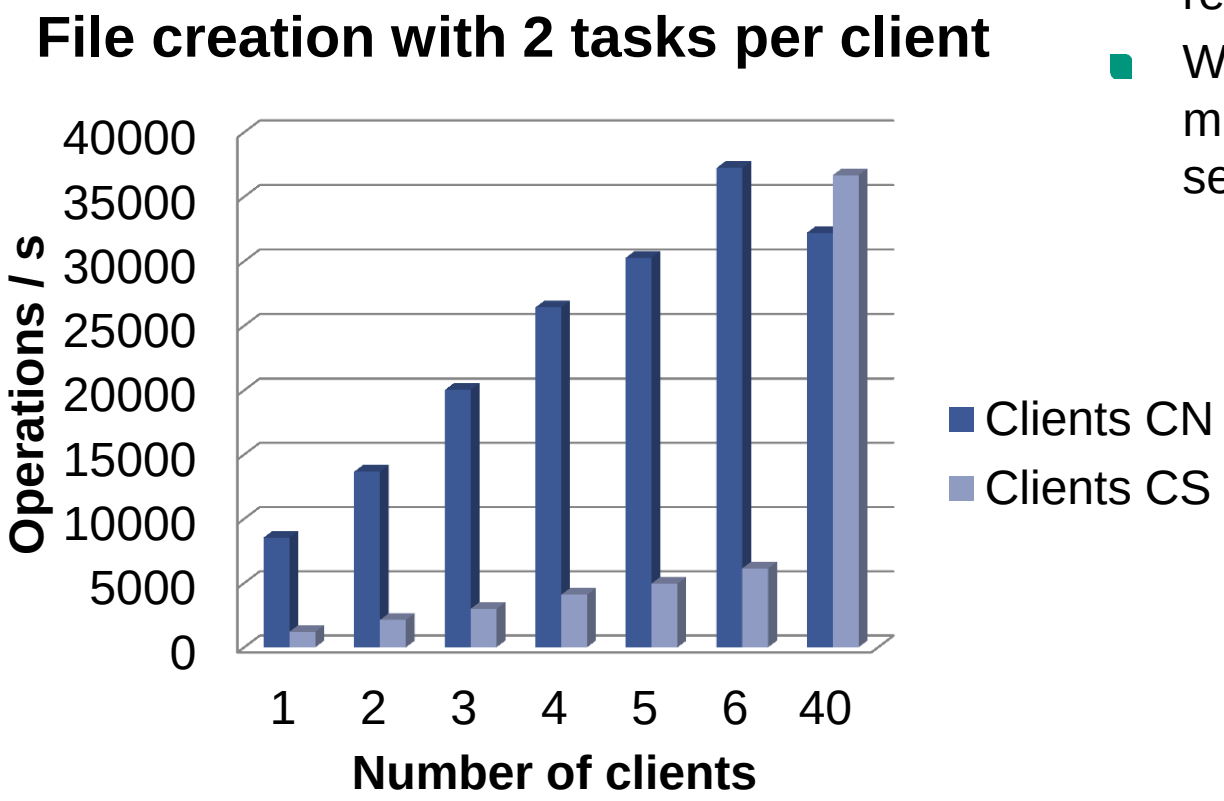


Write performance to file system at CS

- Same performance from both sites



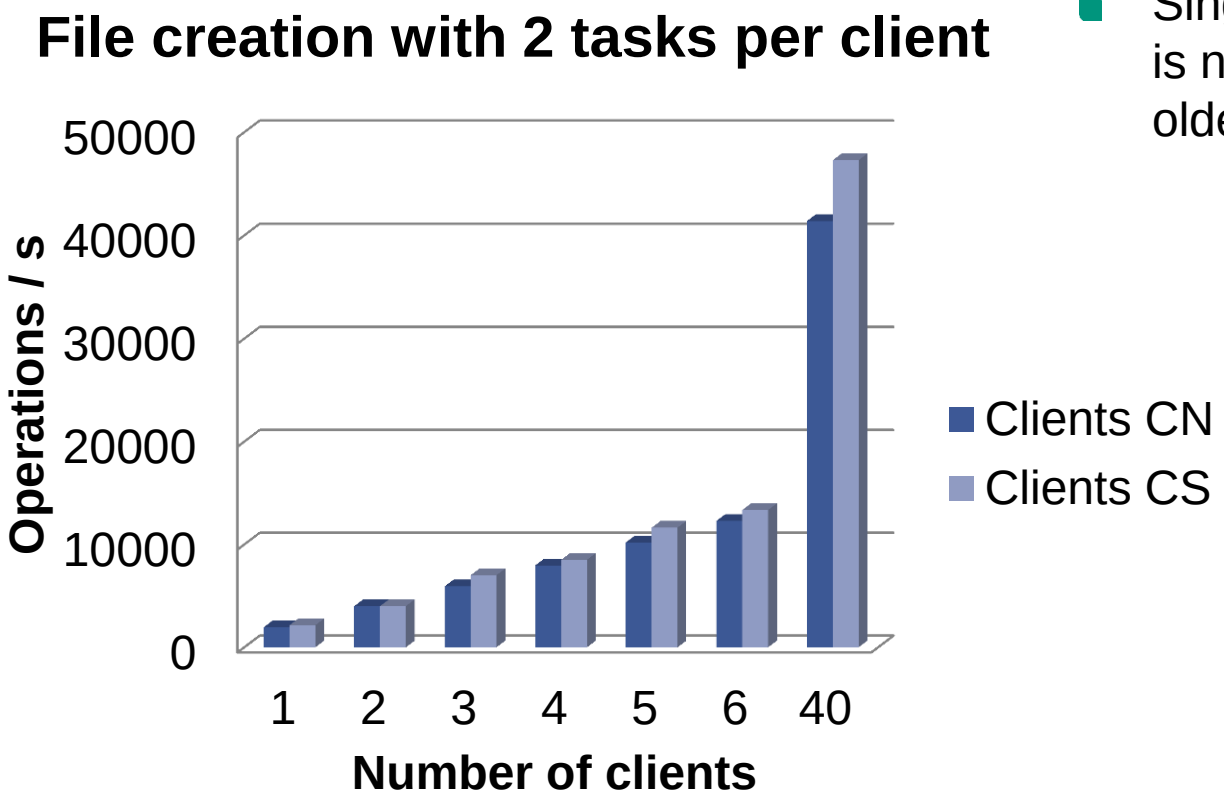
File creation performance to file system at CN



- Much slower performance from remote site
- With high load from many clients delay on server is dominating

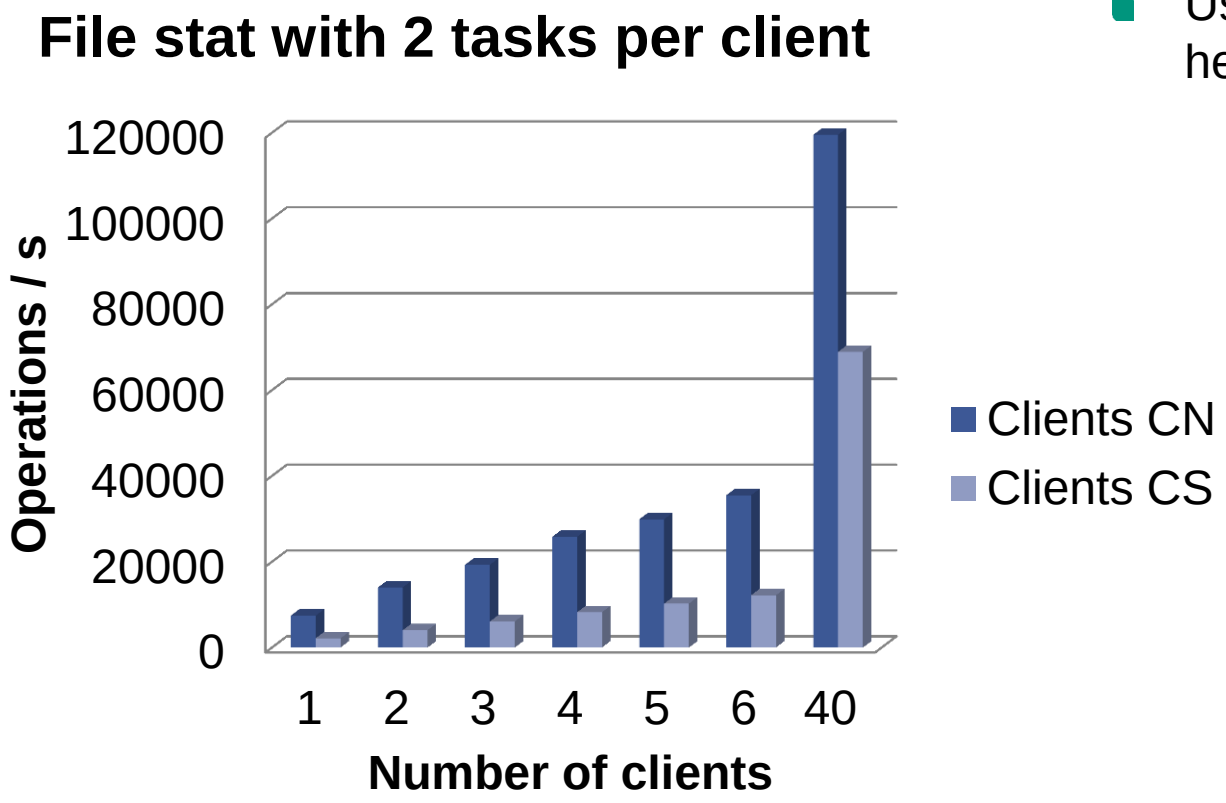
File creation performance to file system at CS

- Same performance from both sites
- Single client performance is not good, impact of older server version?



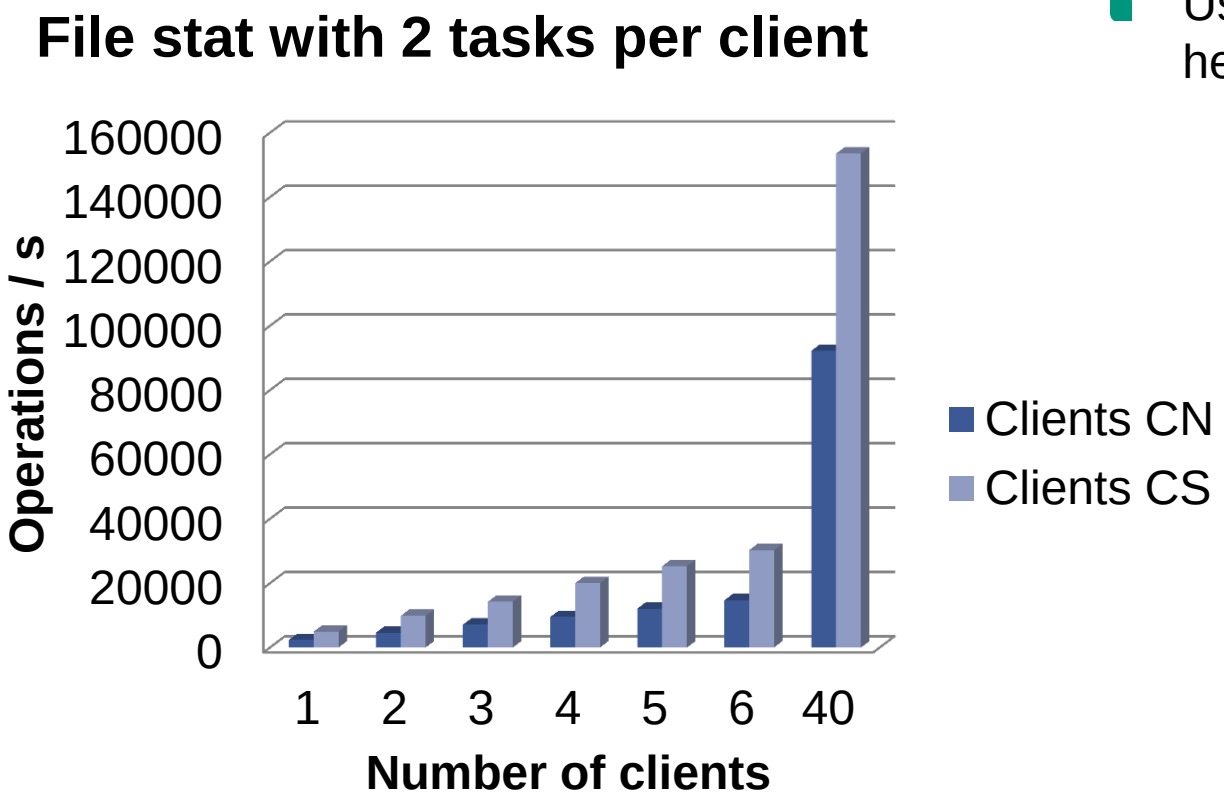
File stat performance to file system at CN

- Remote clients are slower
- Using more clients helps to hide latency



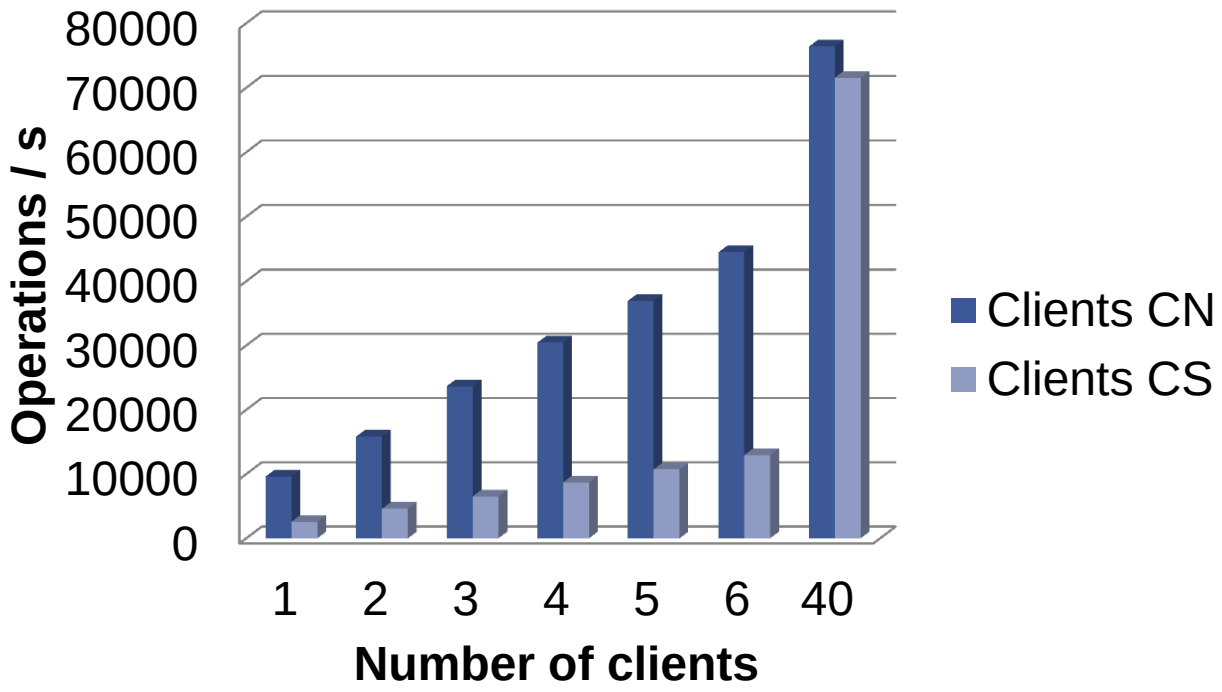
File stat performance to file system at CS

- Remote clients are slower
- Using more clients helps to hide latency



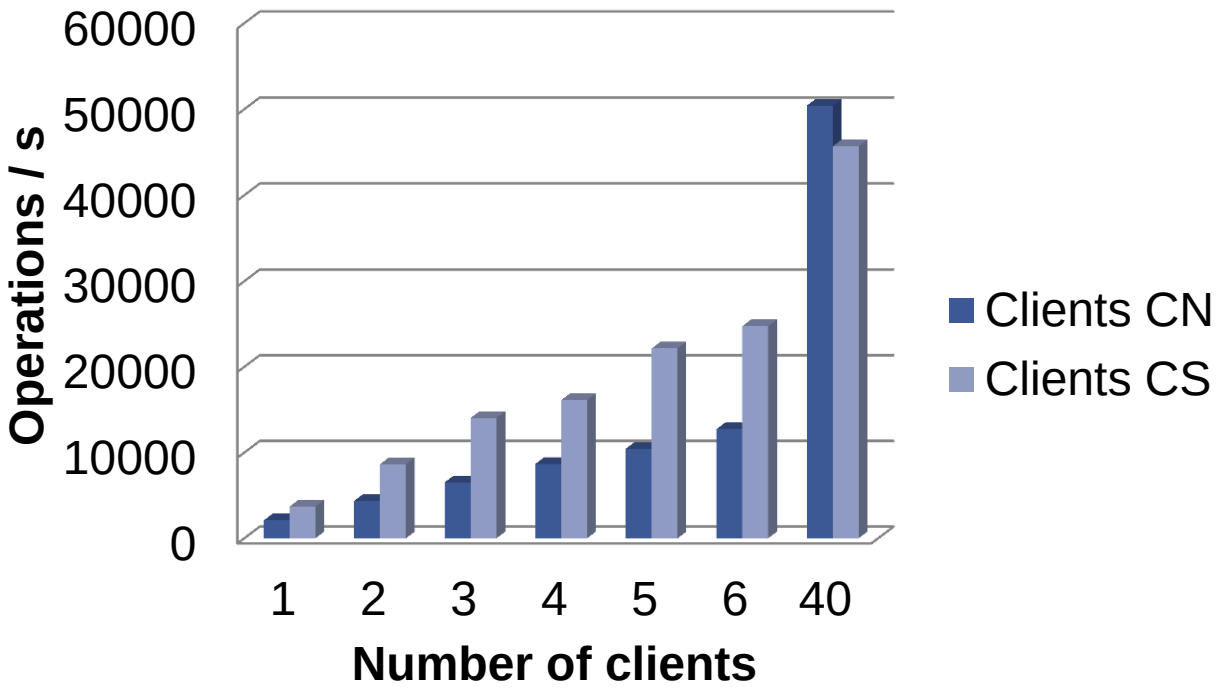
File remove performance to file system at CN

- Remote clients are slower
- Using more clients helps to hide latency



File remove performance to file system at CS

- Remote clients are slower
- Using more clients helps to hide latency



Experiences and summary

- Experiences with complex IB network at KIT
  - Analyzing network related problems is not easy
  - We saw very few critical problems
  - Long distance IB connection just worked as expected
- Performance over long distance IB
  - Throughput is similar to local usage
  - Metadata performance depends on distance
    - With 33 km drops to about one-third for some operations
    - Impact can be reduced by using more clients
- All my talks about Lustre
  - http://www.scc.kit.edu/produkte/lustre.php
- roland.laifer@kit.edu