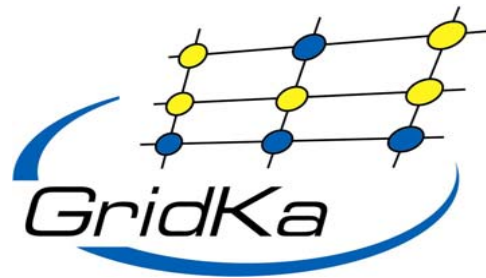


GridKa FZK



Status and plans for the
Service Challenges



Overview

- Setup
- Last challenge results (till 2rd week of Feb)
- Plans and Roadmap



People

- Bruno Hoefft – Networking, Security
- Doris Ressmann – Tape and Disk storage, dCache
- Jos van Wezel - Glue
- In fact all of GridKa staff

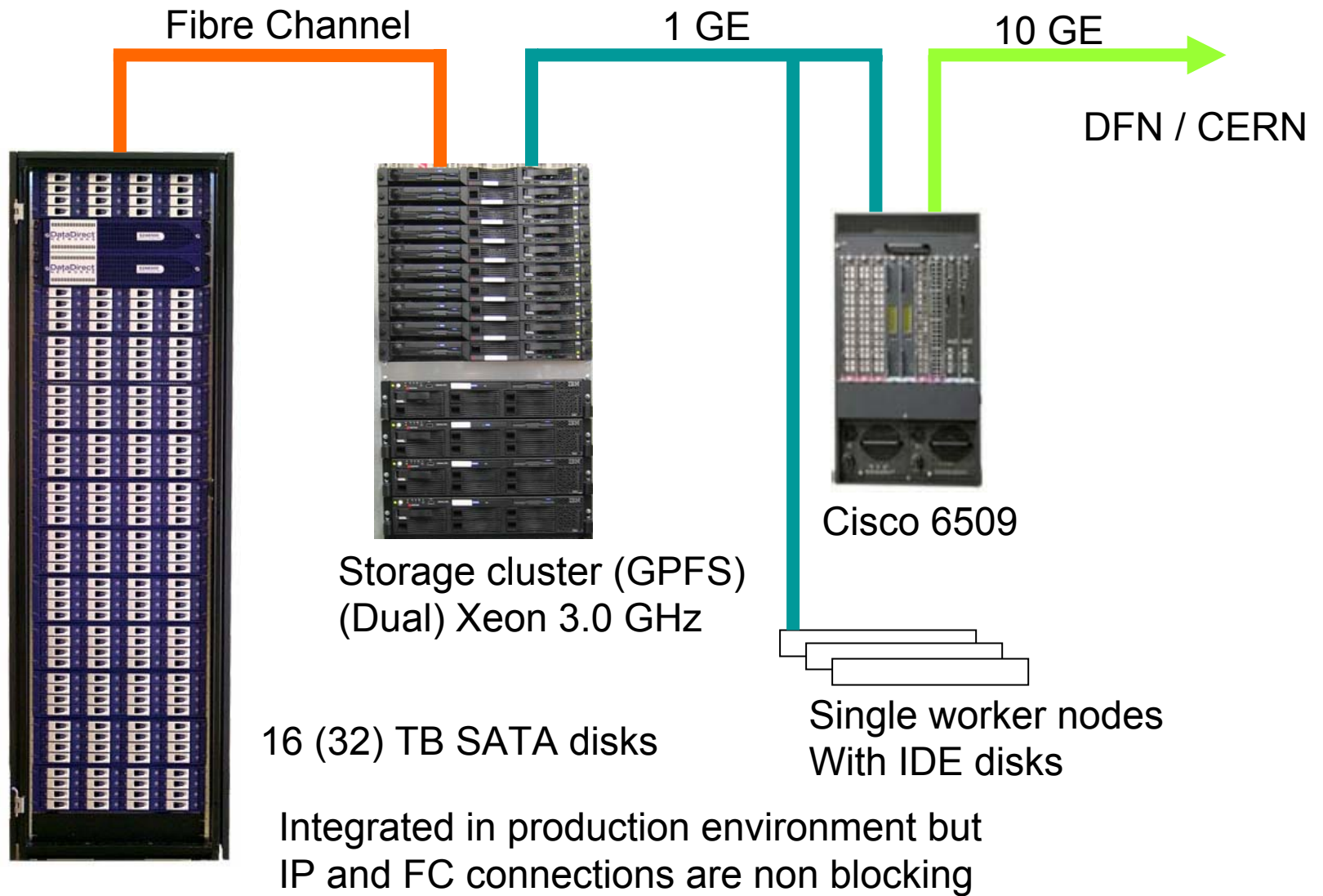
Hardware

- Various Xeon dual 2.8 and 3.0 GHz IBM x-series (Intel and Broadcom NIC)
- Recently added 3.0 GHz EM64T (800 FSB)
- Cisco 6809 with 2 10 Gb ports and lots of 1 Gb
- Storage:Datadirect 2A8500 with 16 TB
- Linux RH ES 3.0 (U2 and U3), GPFS
- 10 GE Link to GEANT via DFN (least best effort ☹)

TCP/IP stack

- 4 MB buffer
- 2 MB window size

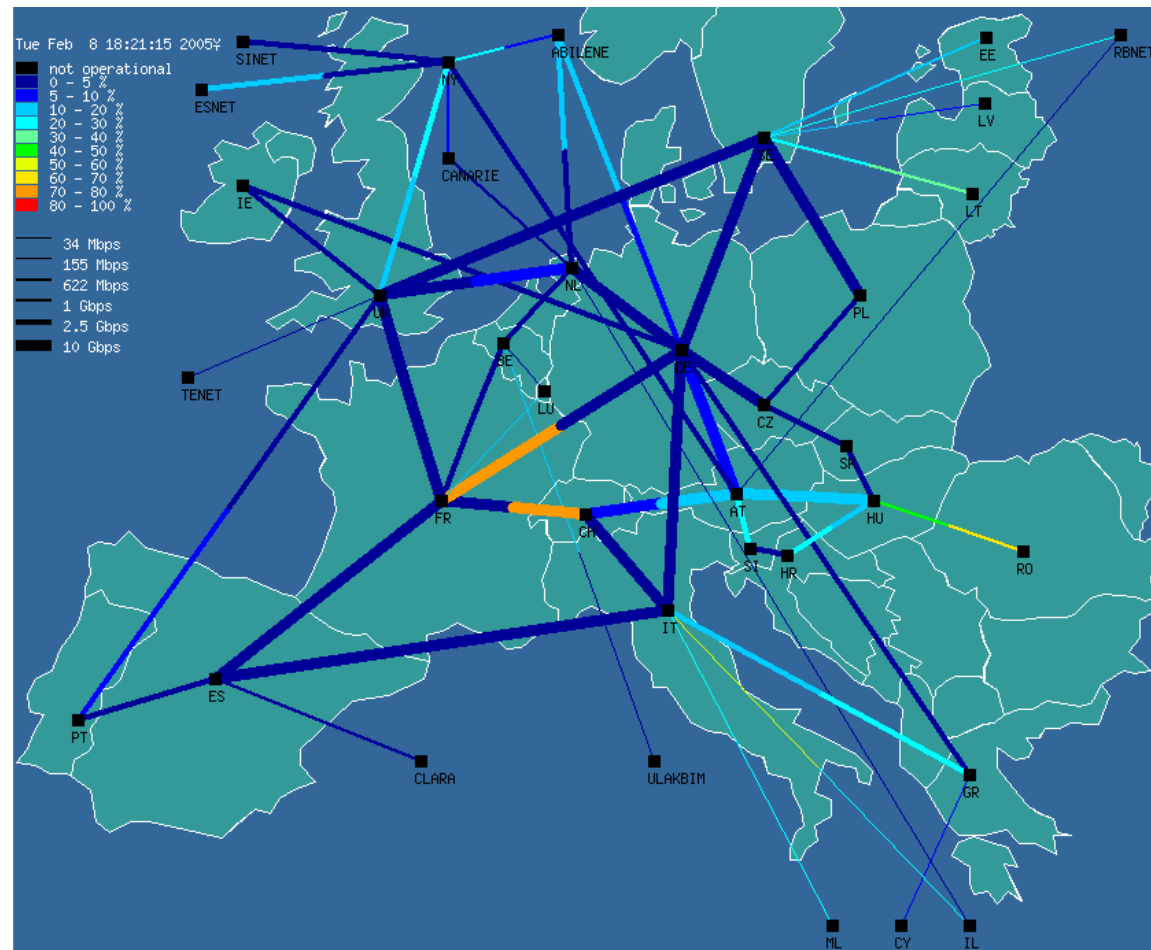
Setup



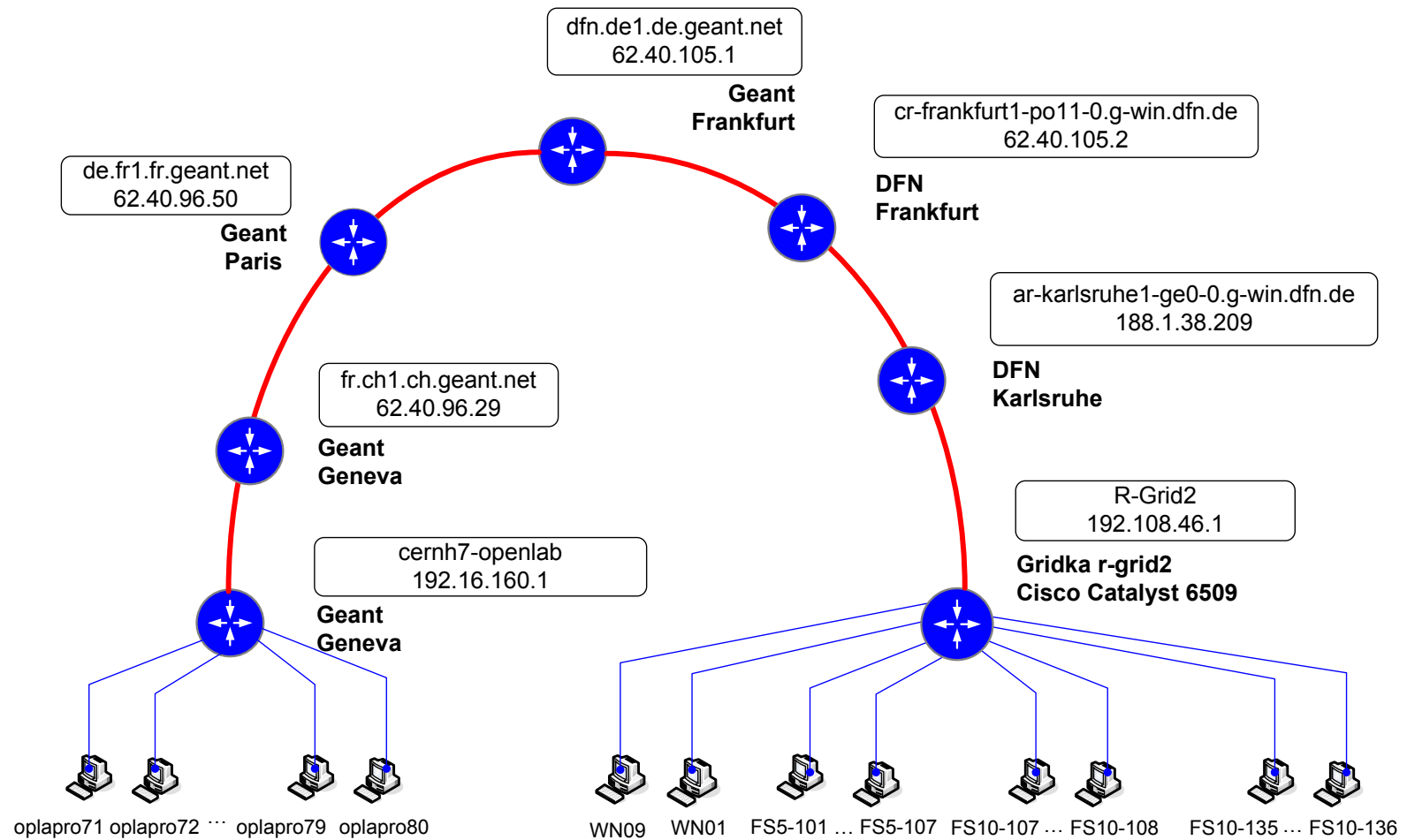
Lighting the net

Traffic during FZK SC1

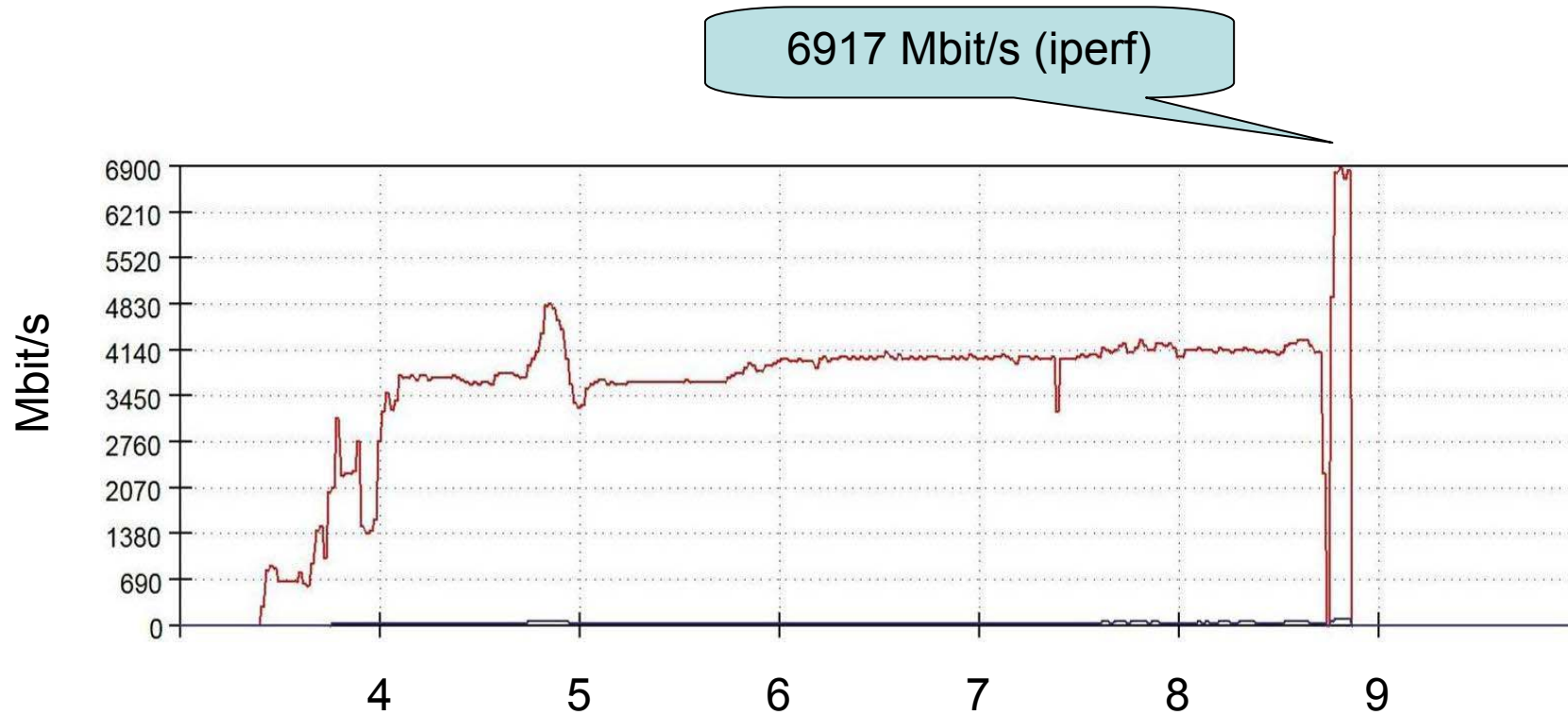
- Intermittant drops
- Least best effort



Network routes

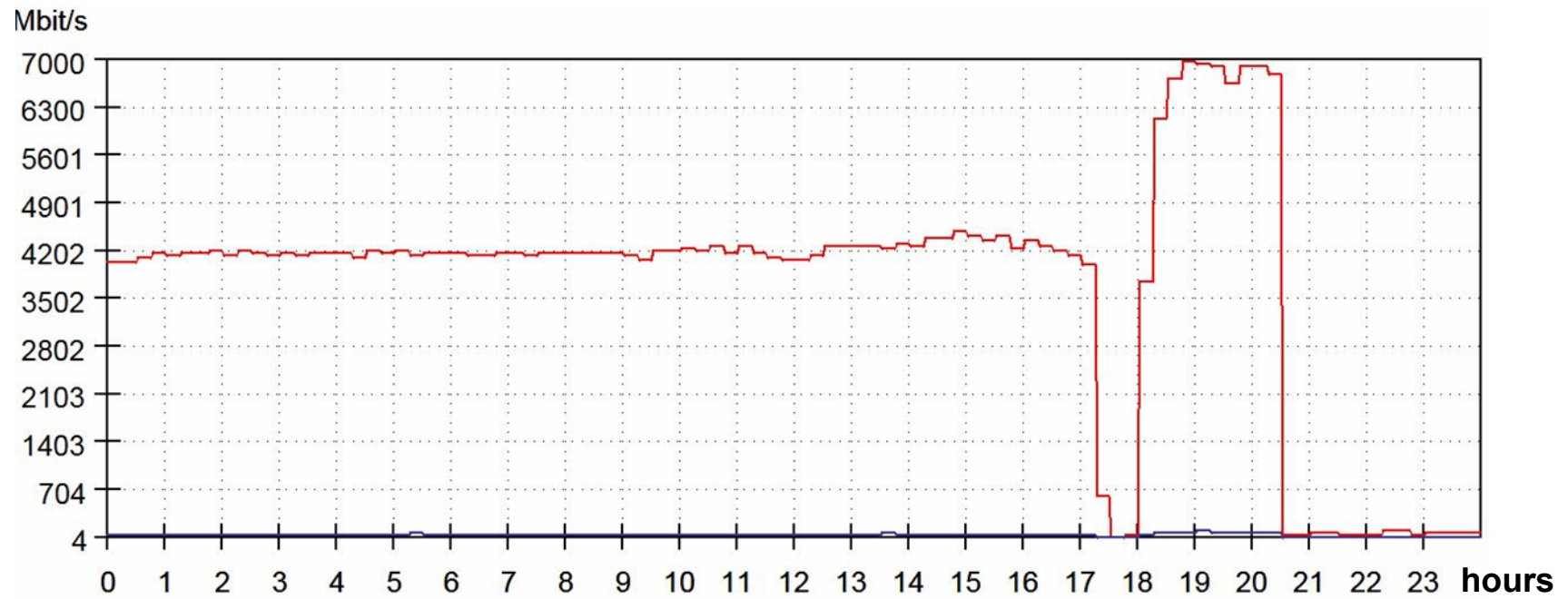


Week statistic 3th Feb 2005



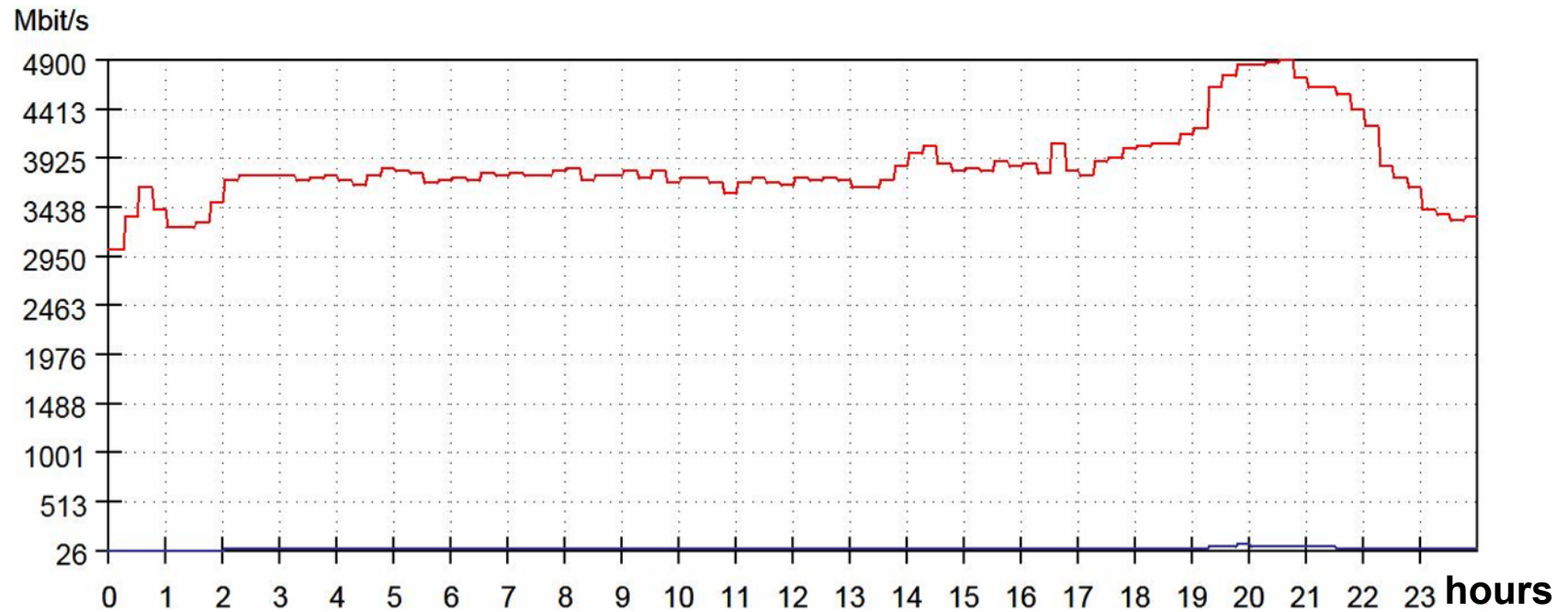
Source: DFN

Iperf reaches 7.2 Gb



9 iperf streams

Disk to disk





Results

- Local to disk: 200 + 360 MB/s
- OpenLab (disk) to FZK /dev/null: 700 MB/s
- OpenLab (disk) to FZK disk: 150 + 200 (9 nodes)

Issues

- Servers and clients needed matching to optimize (needs to be resolved)
- Less streams is more throughput (overhead vs. recovery time)
- Latencies to storage have large influence?
- Linux VM and TCP/IP tuning needed: suggestions? SC wiki?
- Did we forget this offer by Harvey Newman and should we track back:
“I'd like to suggest a review of the work by Yang Xia, Sylvain Ravot and Dan Nae on server systems and network interfaces, along with protocol stacks and associated tuning. Also- a series of standalone tests should be planned, so that the knowledge of exactly where the bottlenecks are, and what they are, is not lost in the mix.”
- What happened to Monalisa?
- SC takes a lot of time. We must also guarantee production.
- Datadirect machine does not perform as expected.

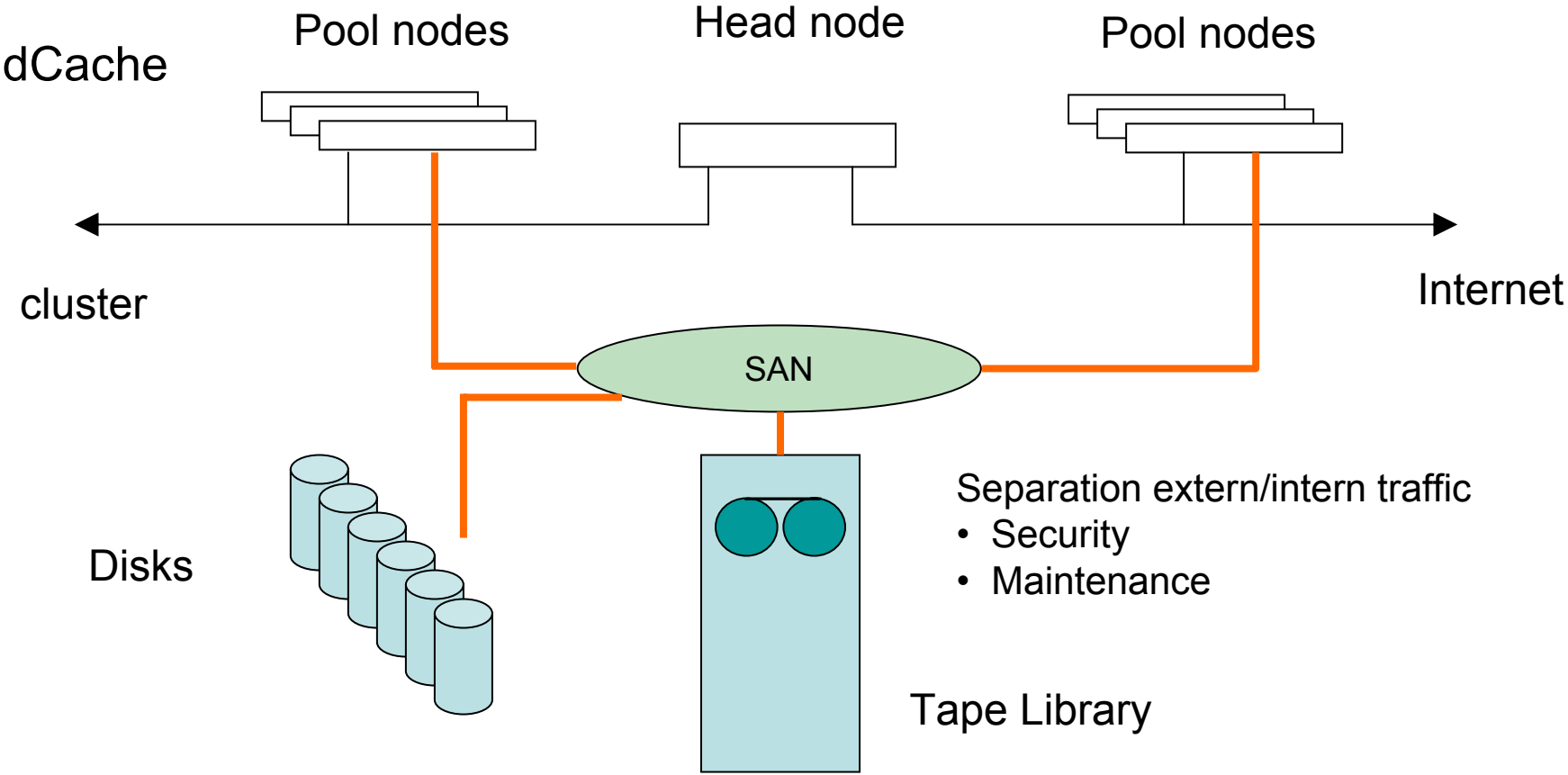
- We still do not know how many machines and disks and tape drives we need.



Plans

- Get more servers for storage farm and dCache
- Get SC dedicated tape lib (S-AIT1 or 2)
- Integrate dCache into SC setup (SC is now production too)
- Optimize storage farm (possibly add SRM)
- Start transfers via Radiant/SRM/gLite/xrootd?
- Keep 10 GE dedicated to SC till June

Storage sketch



Roadmap

- March
 - 100 MB/s disk to disk
 - Cooperation with SARA
- April/ May
 - Adding servers
 - Installation Tape Library with 4 drives (expect 100 MB/s)
 - Installation dCache on SC cluster
 - Installation (multiple) TSM connections for dCache
- June
 - 500 MB/s to disk
 - 10 GE connection shared with other traffic
- July
 - 50 MB/s to tape
- September
 - Install additional tape drives (expect 300 MB/s)
- December
 - 100 – 150 MB/s to tape