



HPC GSI/FAIR

Walter Schön, GSI

HPC at GSI: People



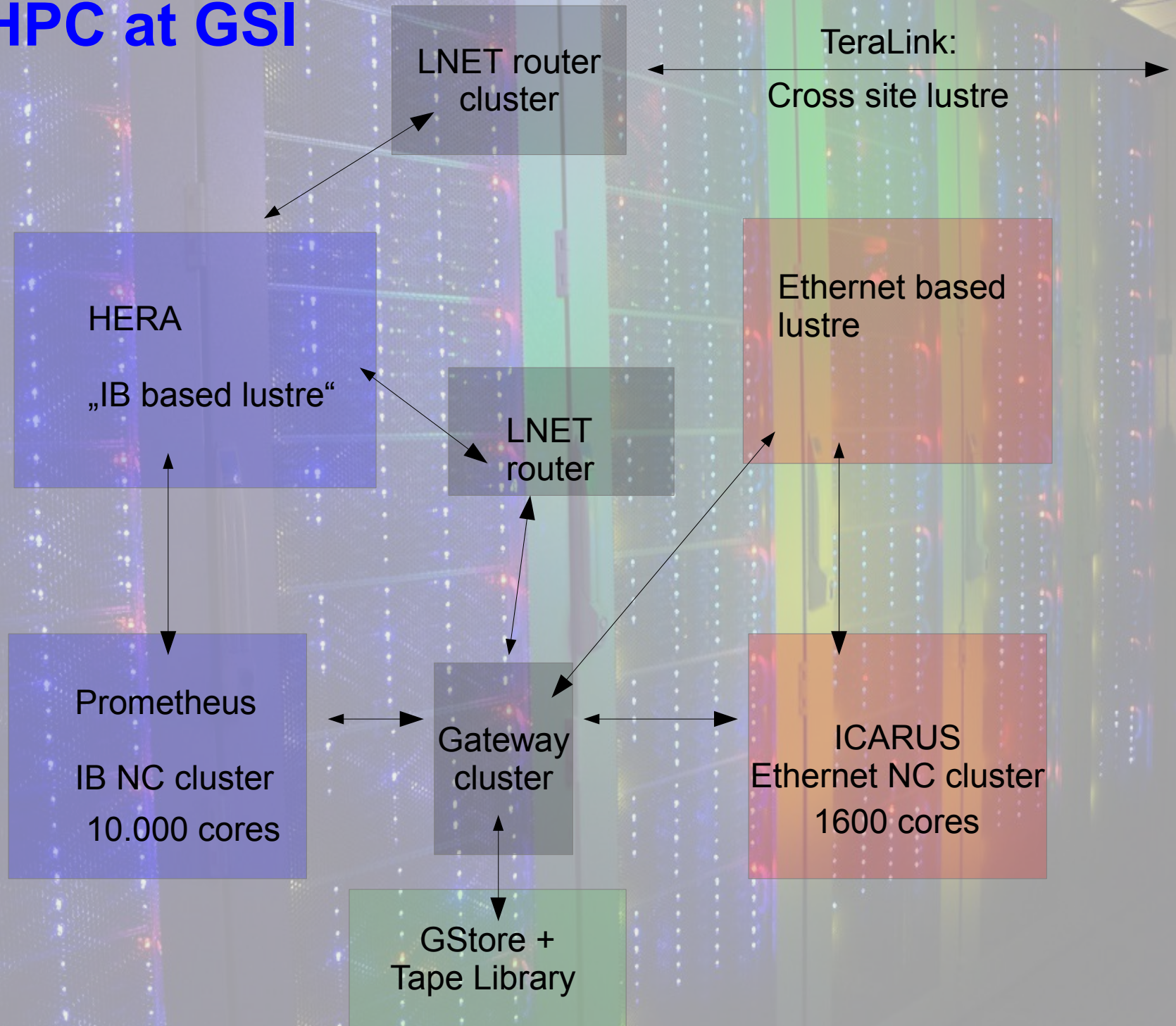
HPC – Unix based central services

- Providing OS level + HA service for central services
 - Web, Postgres DB, DHCP, DNS, Ticket System, Virtualisation Platform (PKI,....),
- Providing OS + Applications for GSI and/or „World“
 - SVN, Wiki, community web server,
 - Data Room (GF + BMBF), security platform
 - MTA's + Spam filter + Grey Listing, Virus scan, exe scan?
- Monitoring + Analysis

work to do

- Heimdal Kerberos/LDAP testing
=> production (1 Quarter 2014)
 - Integrating into PKI, 2 factor authentication
- Deployment of new hardware for virtualisation
- Migration of services on old hardware/OS to VMs ...
 - SVN: done
 - WIKI: In Progress (Dez 2013)
 - MTAs: In Progress (1 Quarter 2014)
 -
- Consolidation of configuration management
- Monitoring + Analysis of Monitoring
=> logstash project

HPC at GSI



Usage of the Farm:

HERA: Full

=> need new hardware (some boxes from old lustre can be reused)

Prometheus: usage increasing, user experience „almost full“

2013/11 Prometheus Accounting

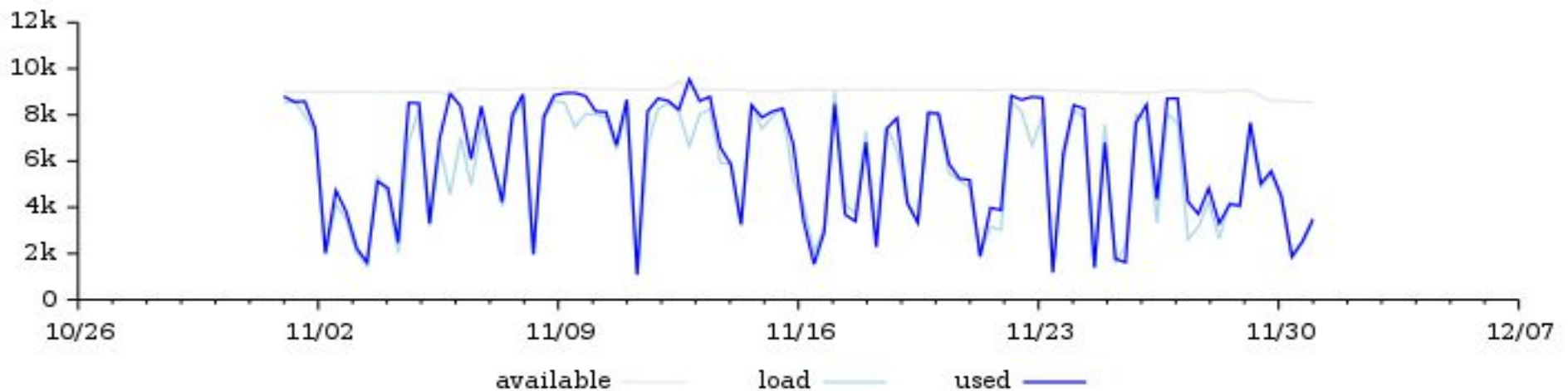
Utilization

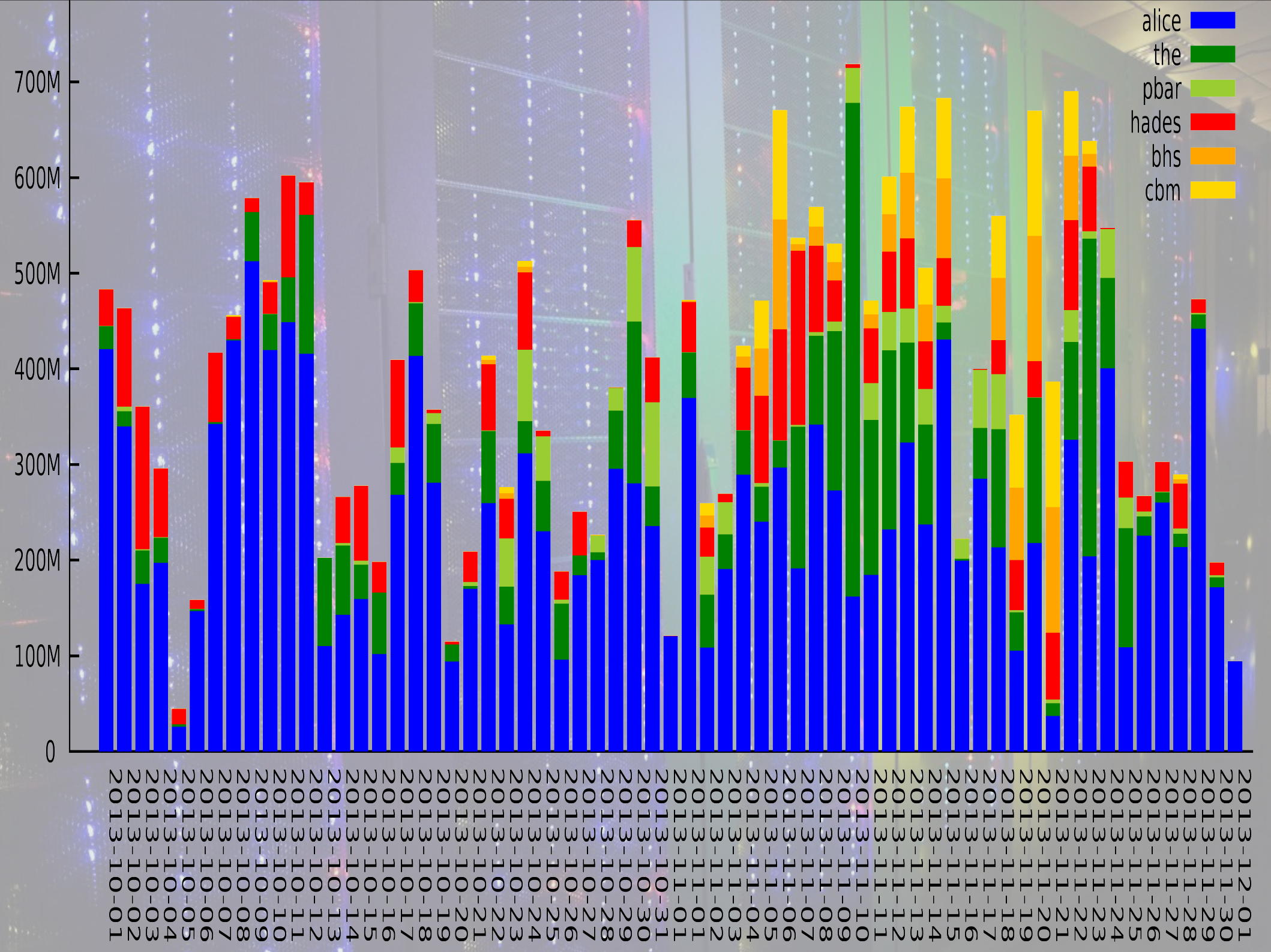
Average

Slots 9009 available

5991 used (66.50%)

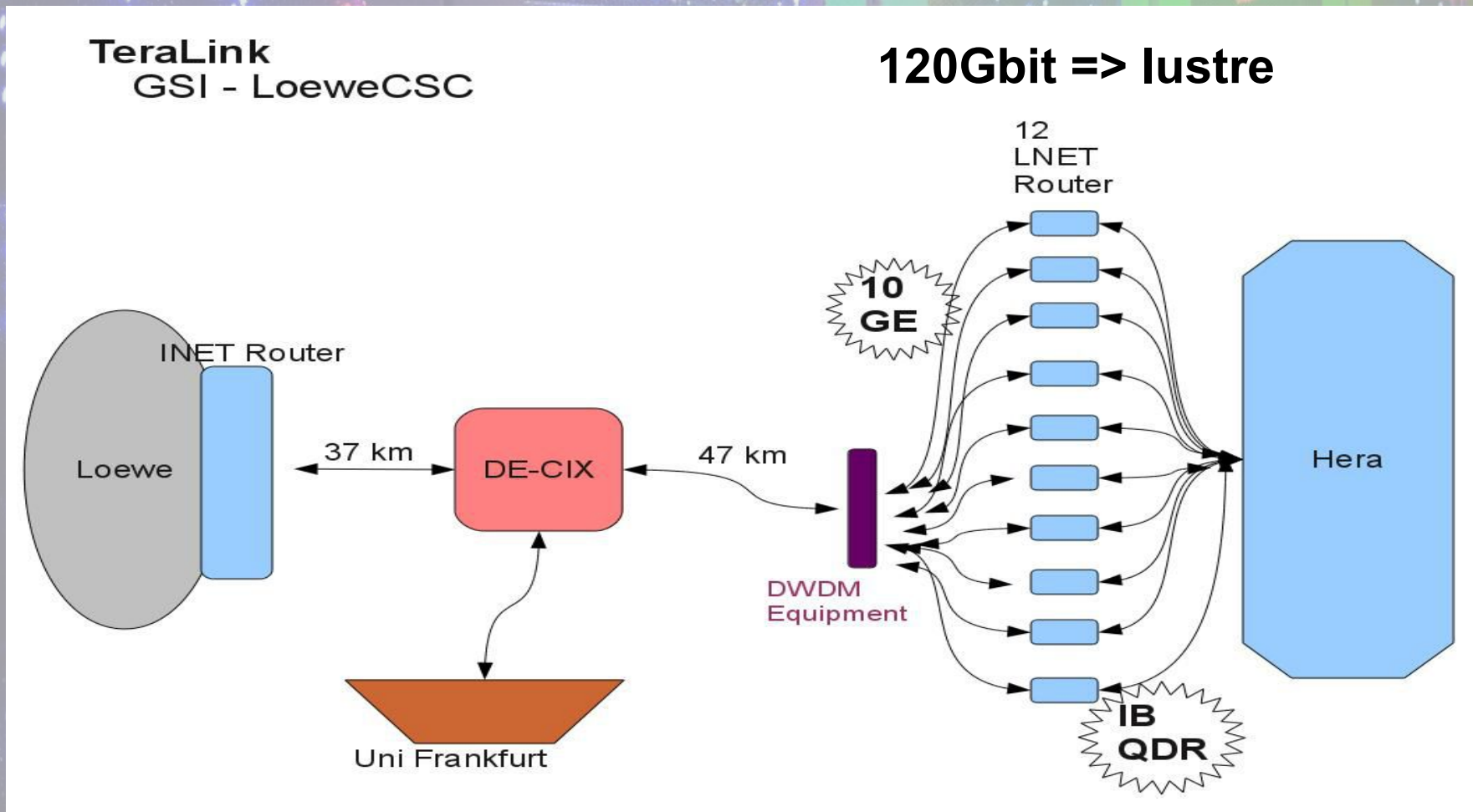
Load 5665 (relative 94.56%)





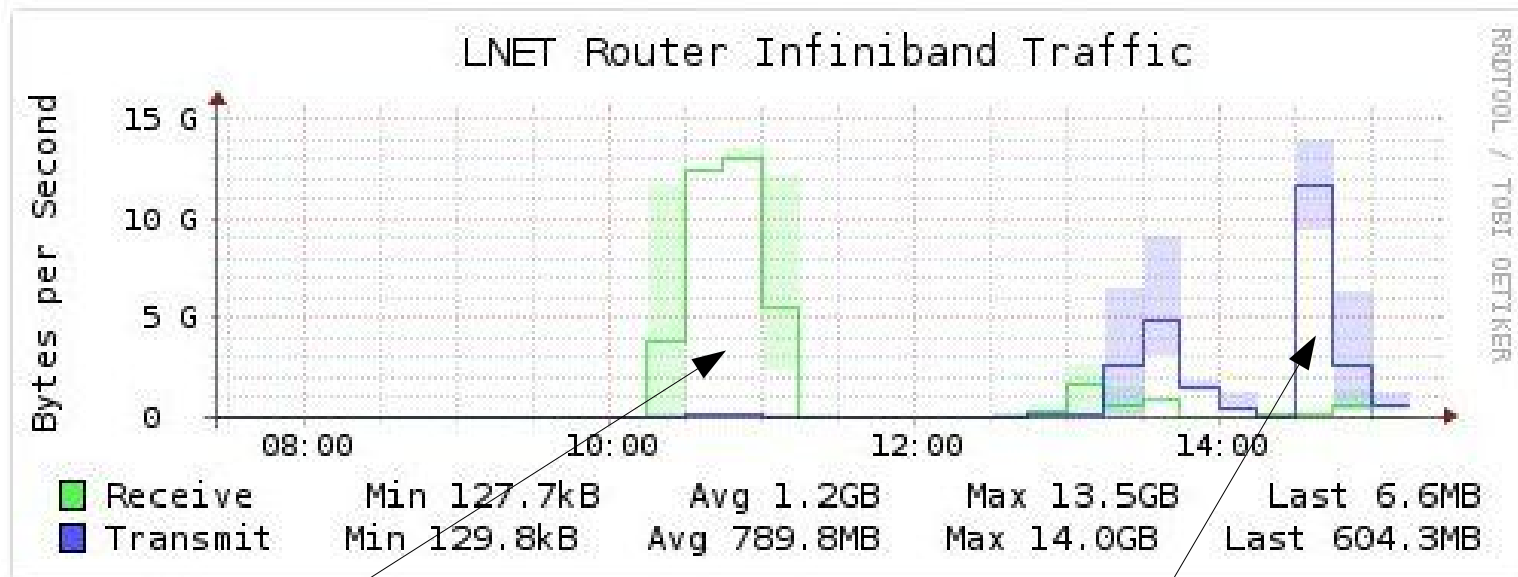
Connecting Loewe-CSC (Frankfurt) with Hera:

- Building HPC router cluster (LNET) for the TeraLink 12 x 10 Gbit Ethernet => lustre IB loadbalancing



.... and it really worked :-)

lustre clients at Gauss Supercomputer Frankfurt
Hera lustre cluster at GSI

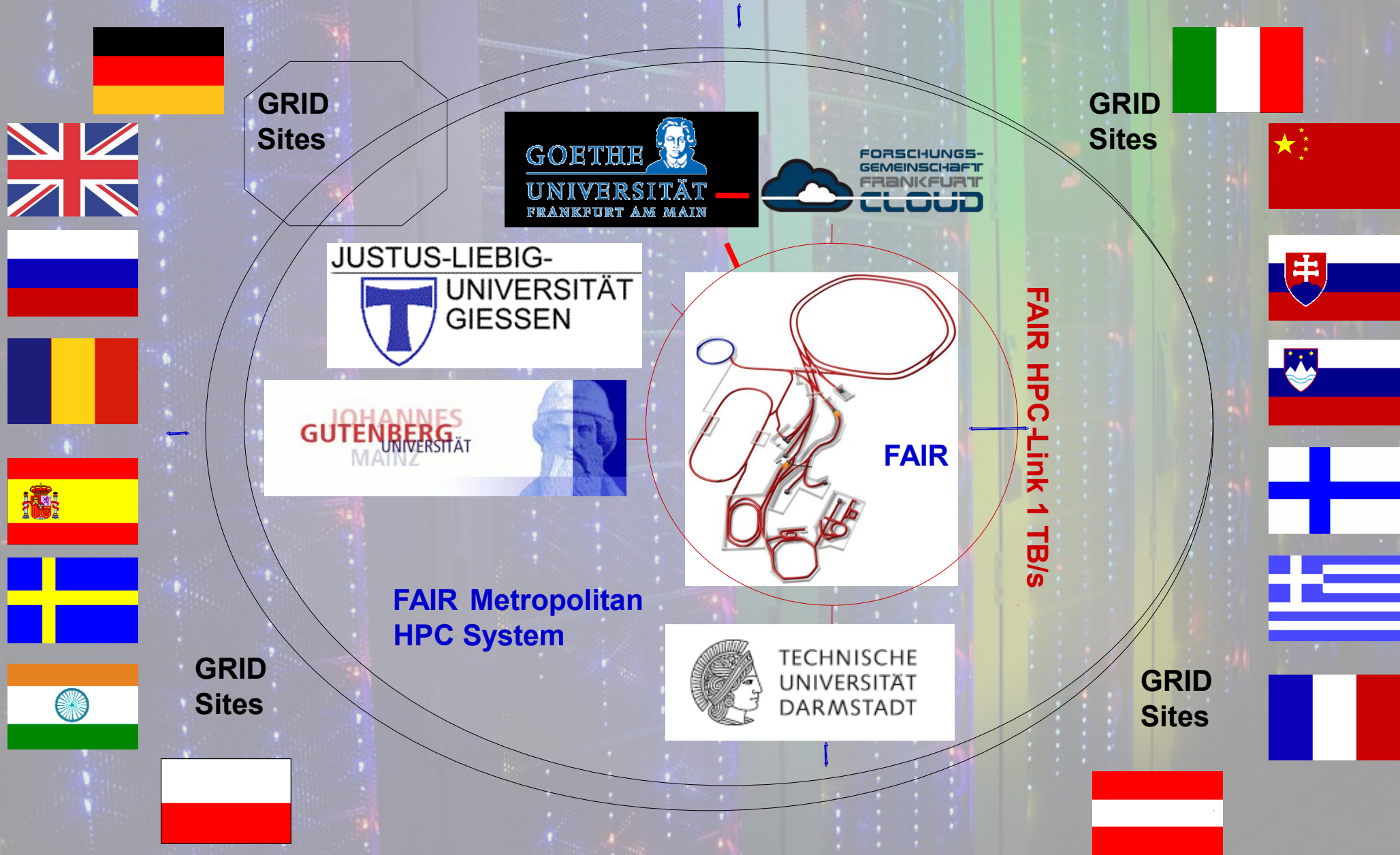


/graph?lustre/router/fin/infiniband/traffic?start=end-8h&end=now

Reading: 13.5 GB/s

Writing: 14GB/s

FAIR HPC Backbone 1TB/s



New Developments/Projects HPC

Parallel FS

- Clustered MDS for lustre => load balancing metadata
- Prototype ZFS: end to end data integrity & „self healing“
 - Performance?
 - Data integrity? Reliability?
 - Management (e.g. „auto“ disk replace functionality)
- Advanced monitoring (Robin Hood, CEA)
- More capacity for HERA, Migration to lustre 2.x

„Cloud/Distributed“ computing

- Large Scale Virtualisation Environment
 - => open Nebula: tested, not sufficient, Open Stack?
- kernel security module for cross site lustre security
- State of the art Kerberos for lustre (in development)
- CVMFS/alternatives for software deployment
- „distributed/cross site“ monitoring (new module for Robin Hood?)

• Storage

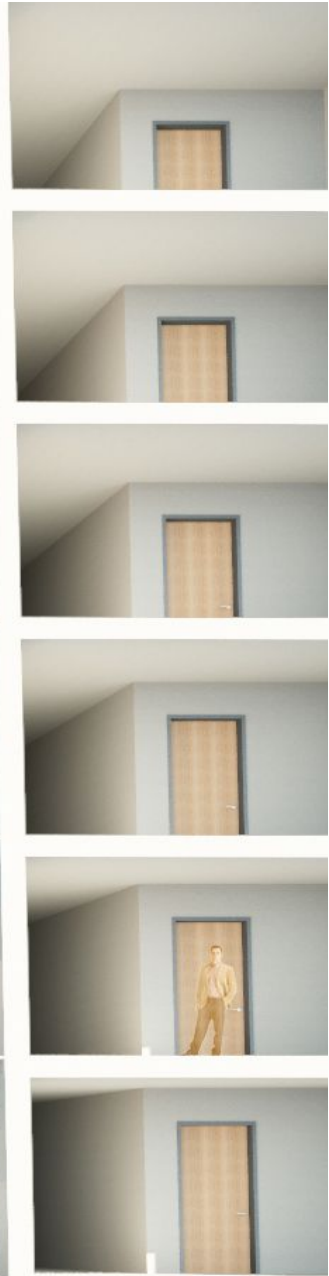
- HSM
- Long time Storage of experiment data

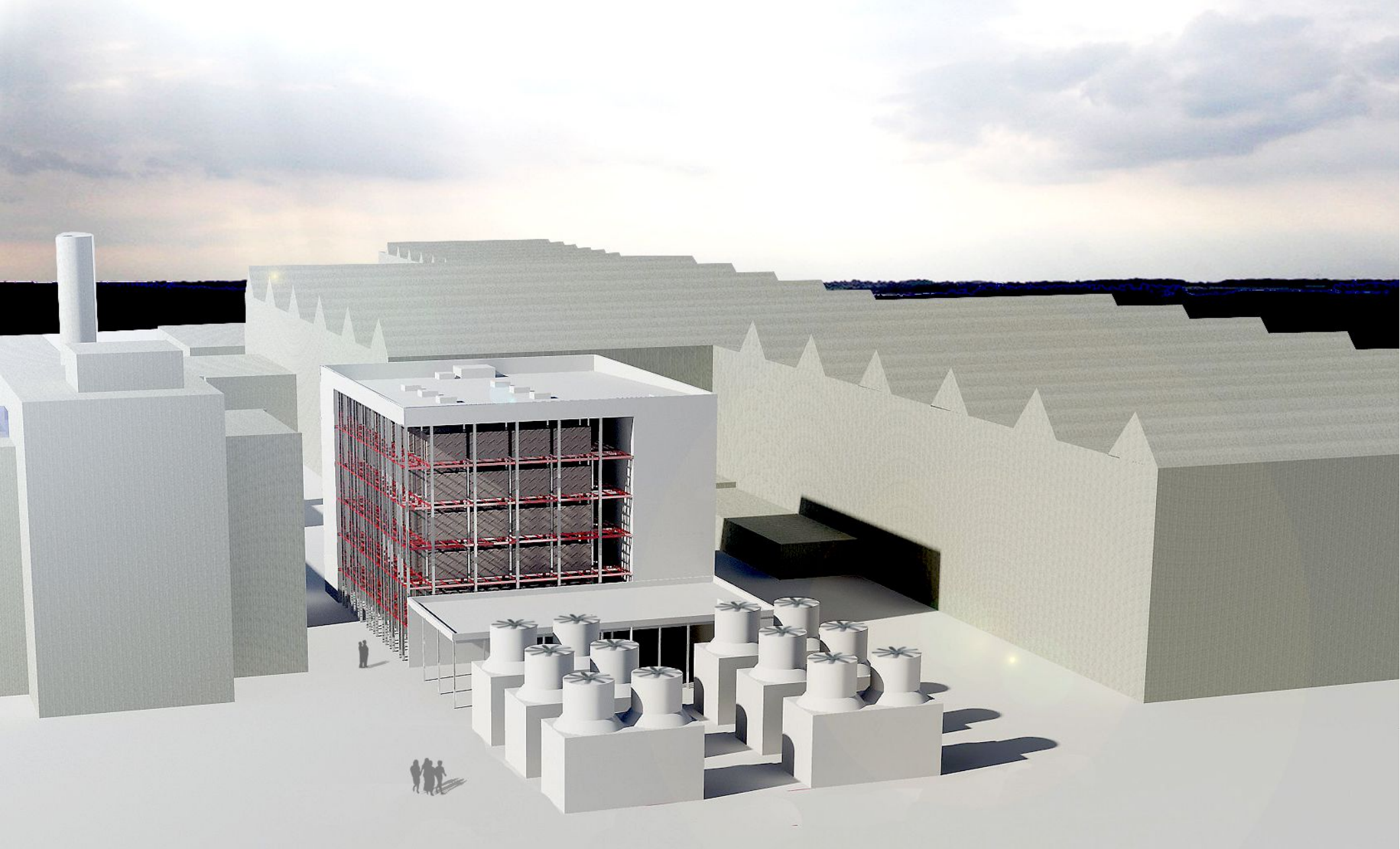
- **Farm Scheduler**
 - GridEngine needs upgrade (Improved accounting)
 - Investigation of „topologie“ sensitive scheduler
=> SLURM?
 - Scalability? Robustness?


New Data Center

„ green cube“

=> cube design (award winning) => minimising cable length, latency
=> PUE ~ 1.06, passive cooling with „water towers“







Thanks to

Bastian Neuburger, Birgit Alborea, Carsten Preuss
Christo Huhn, Dennis Klein, Jan Trautmann,
Matteo Desalvi, Matthias Pausch, Stefan Haller,
Thomas Roth, Thomas Stibor, Victor Penso