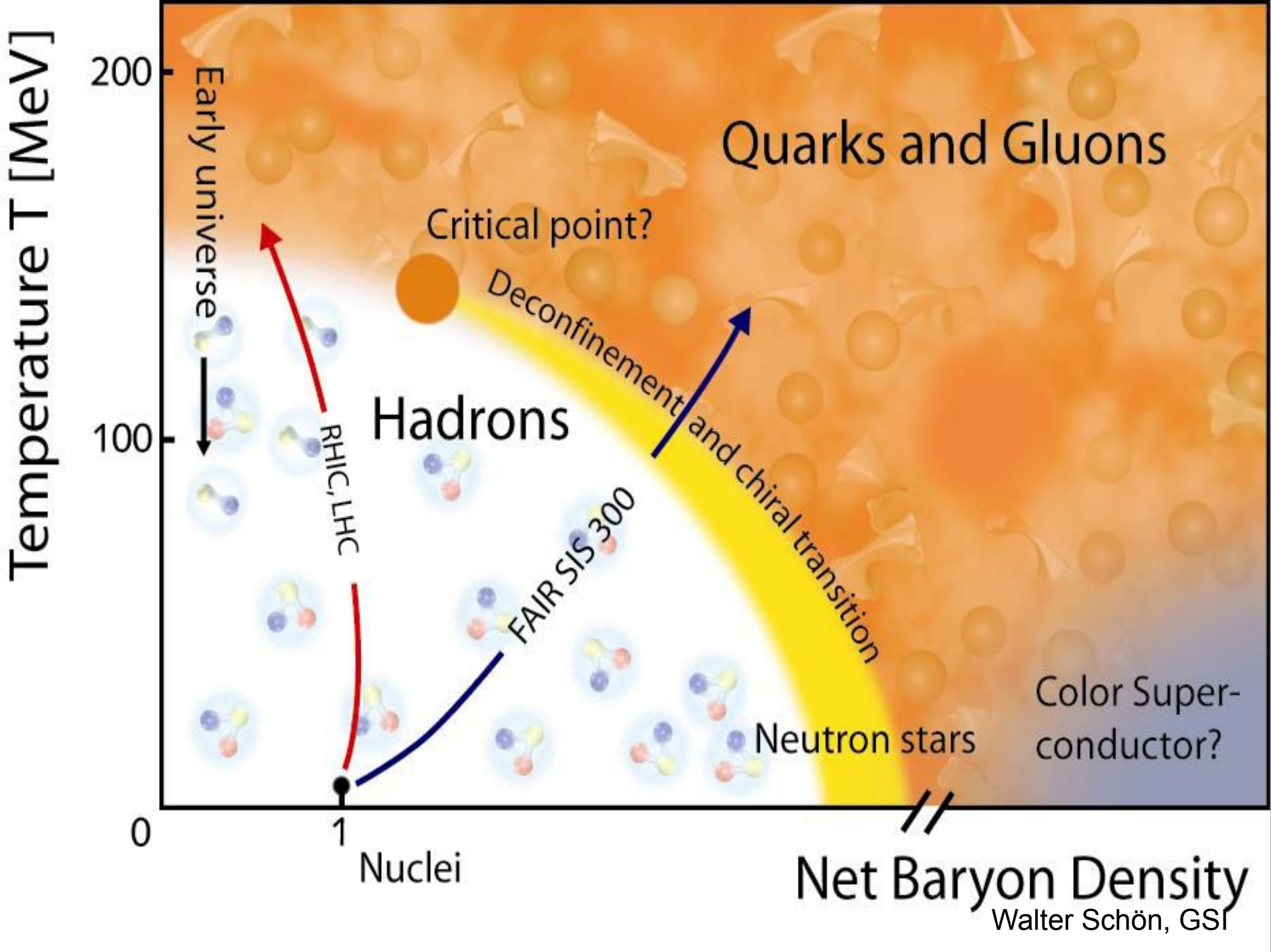
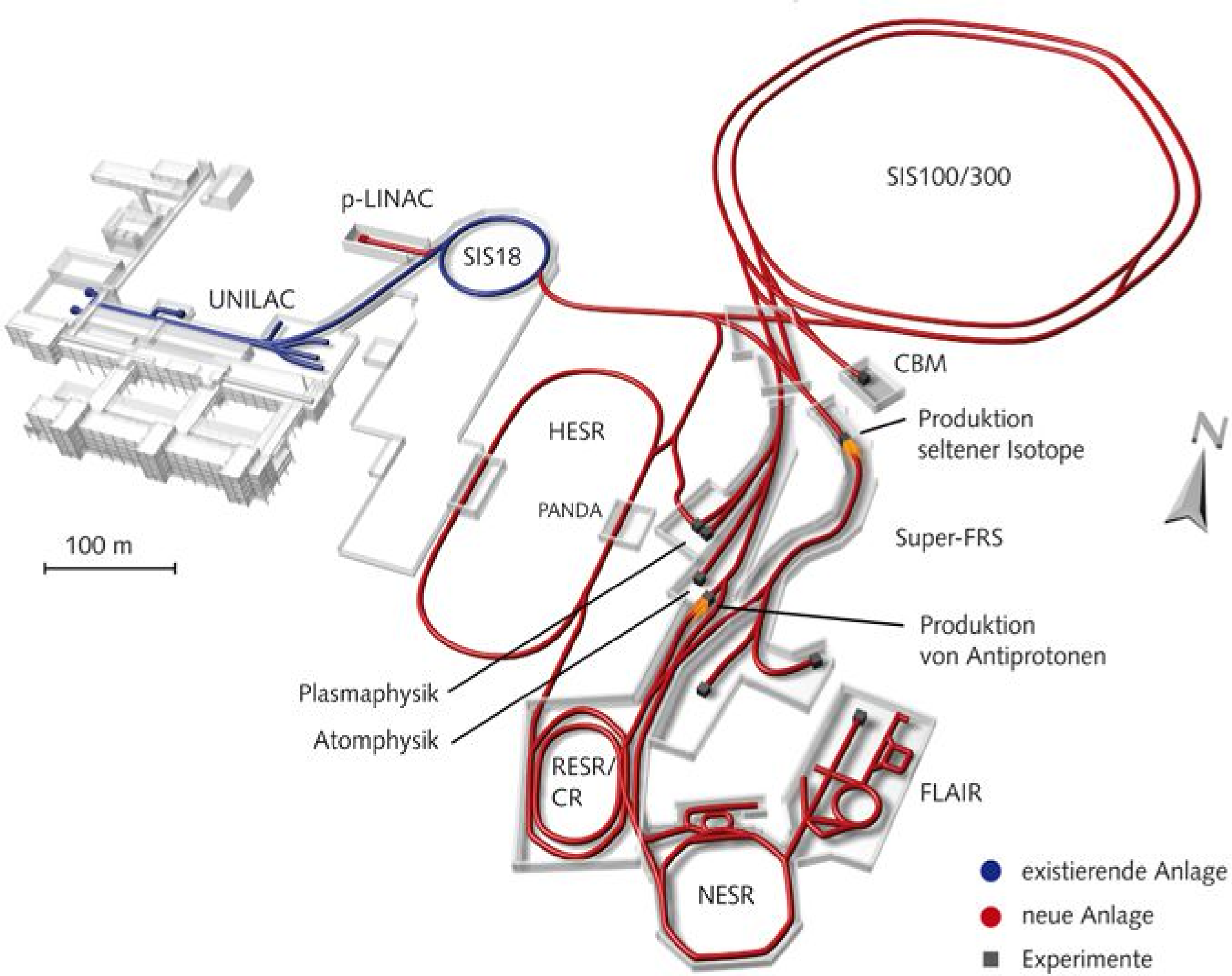
LAD 2013
Lustre at GSI/FAIR

Walter Schön, GSI


Temperature T [MeV]
200
100
0
Early universe
Quarks and Gluons
Critical point?
Deconfinement and chiral transition
Hadrons
RHIC, LHC
FAIR SIS 300
Neutron stars
Color Super-
conductor?
1
Nuclei
Net Baryon Density
Walter Schön, GSI


p-LINAC
SIS18
UNILAC
SIS100/300
CBM
Produktion
seltener Isotope
HESR
PANDA
Super-FRS
100 m
N
Produktion
von Antiprotonen
Plasmaphysik
Atomphysik
RESR/
CR
FLAIR
NESR
existierende Anlage
neue Anlage
Experimente

Walter Schön, GSI

I/O Requirements

- about 1 TB/s for CBM
- about .5 TB/s for Panda

……. + „smaller“ Experiments

=> new concept: no Hardware Trigger
=> Flexible Event Selector: compute farm calculates „Trigger“
ca 60.000 cores only for CBM FLES :-)

I/O after FLES:

1 GB/s CBM
1 GB/s Panda
+ „smaller' Experiments

=> „ Big Data“ ;-)
=> lustre :-)

Current lustre System

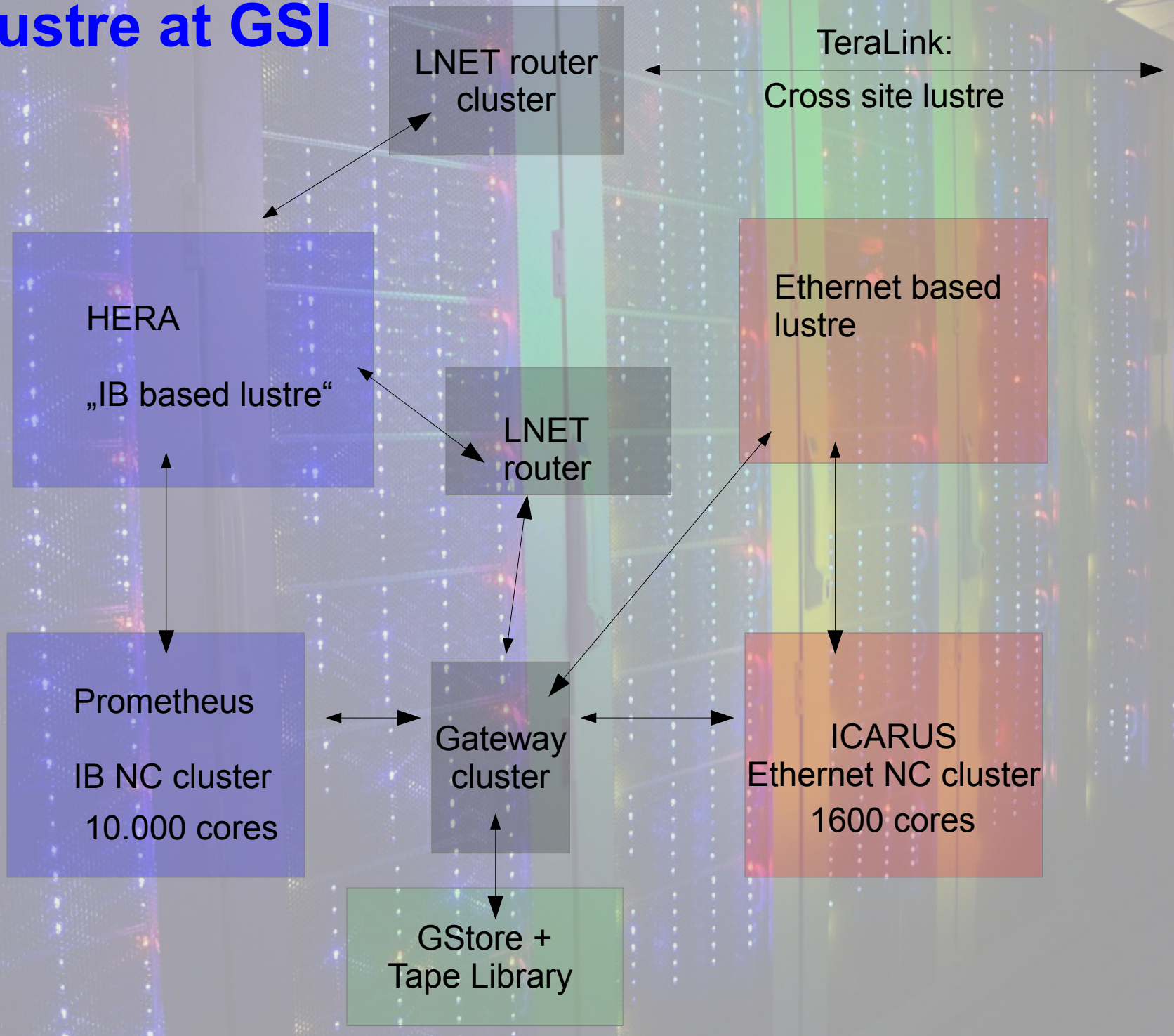
- LHC experiments (ALICE)
- GSI experiments (e.g. HADES)
- simulations of FAIR experiments
- lattice QCD for FAIR physics
- accelerator physics ……

Numbers:

7.000 disks
10 PB gross capacity (including RAID6 + global spare)
6.5 PB for users
2Tbit/s I/O capacity

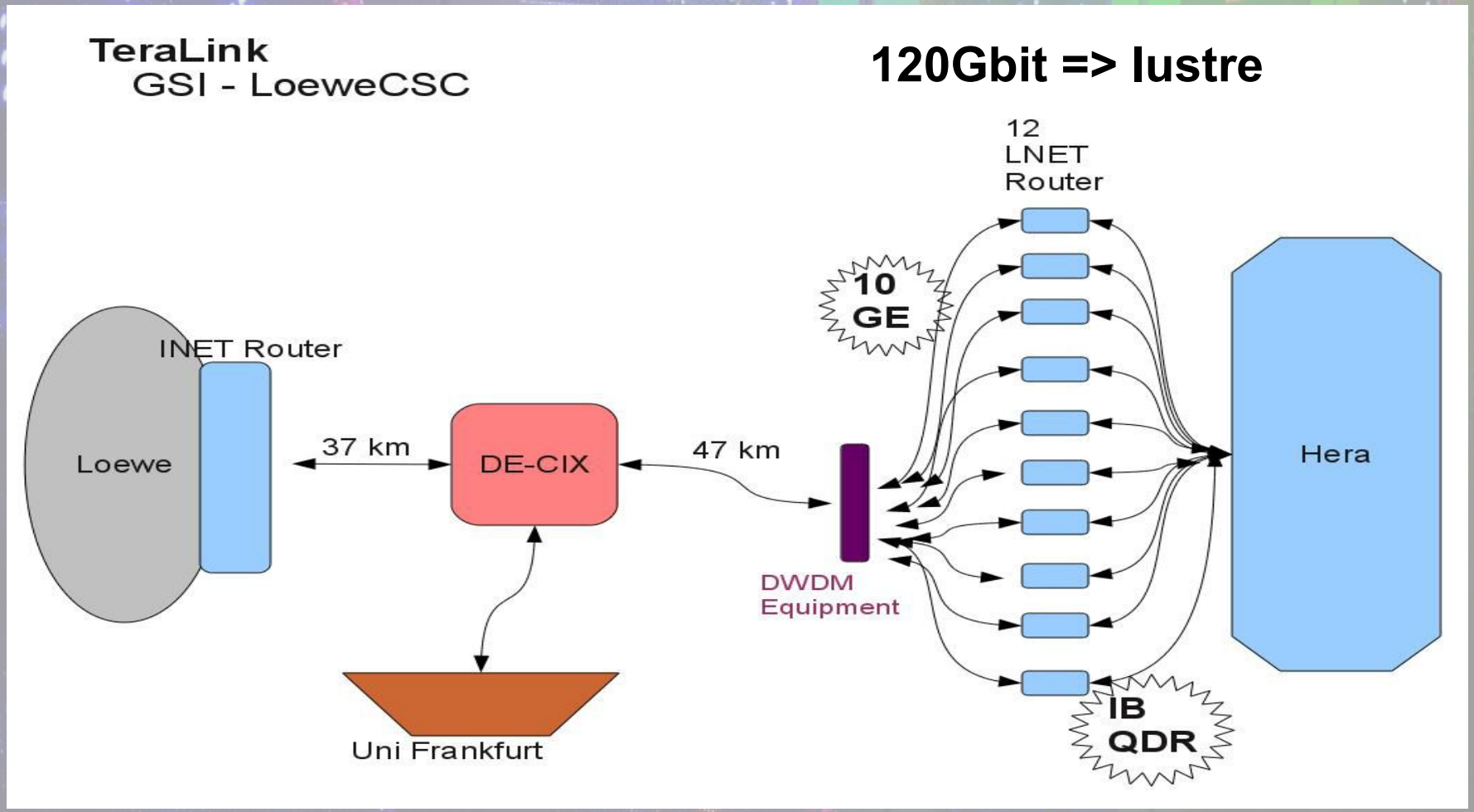
short summary:
-built by HPC Department of GSI
-commodity hardware
-very sucessfull: reliable, stable, fast
- inside the prototype „mini-CUBE“

lustre at GSI



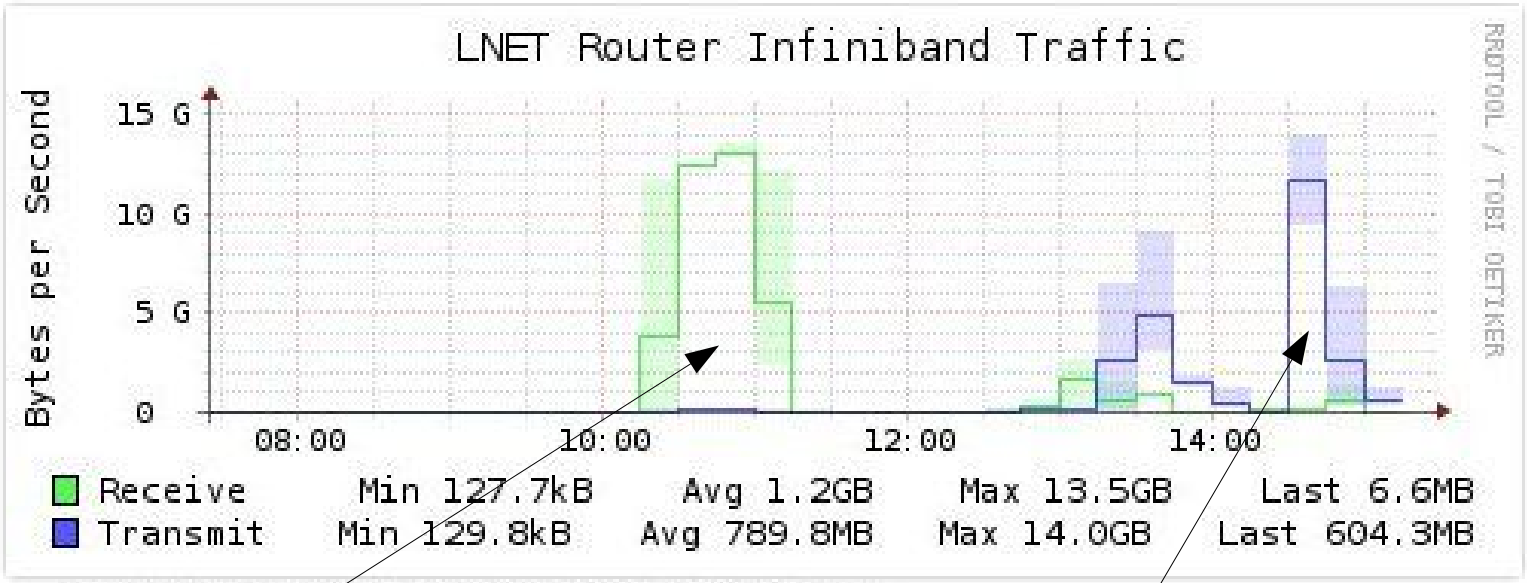
Connecting Loewe-SC (Frankfurt) with Hera:

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



…. and it really worked :-)

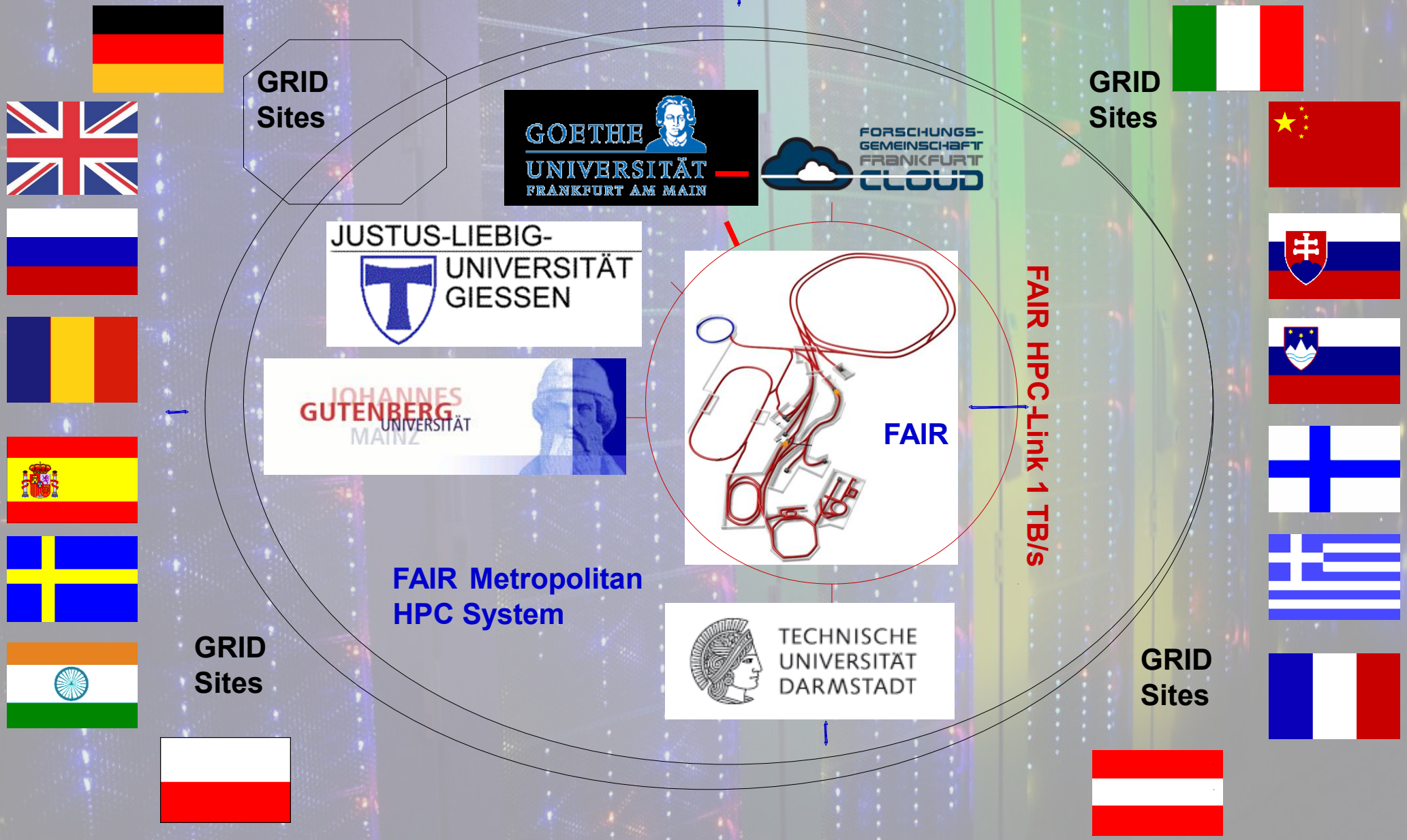
lustre clients at Gauss Supercomputer Frankfurt
Hera lustre cluster at GSI



Reading: 13.5 GB/s

Writing: 14GB/s

FAIR HPC Backbone 1TB/s



Walter Schön, GSI

New Developments

Security between remote sites

- „state of the art“ kerberos in lustre 2.x
- kernel security module for lustre

=> talk by Thomas Stibor, code is merged in 2.4 tree

Testing

- lustre 2.4, HSM, clustered MDS, ZFS

Lustre future at GSI/FAIR

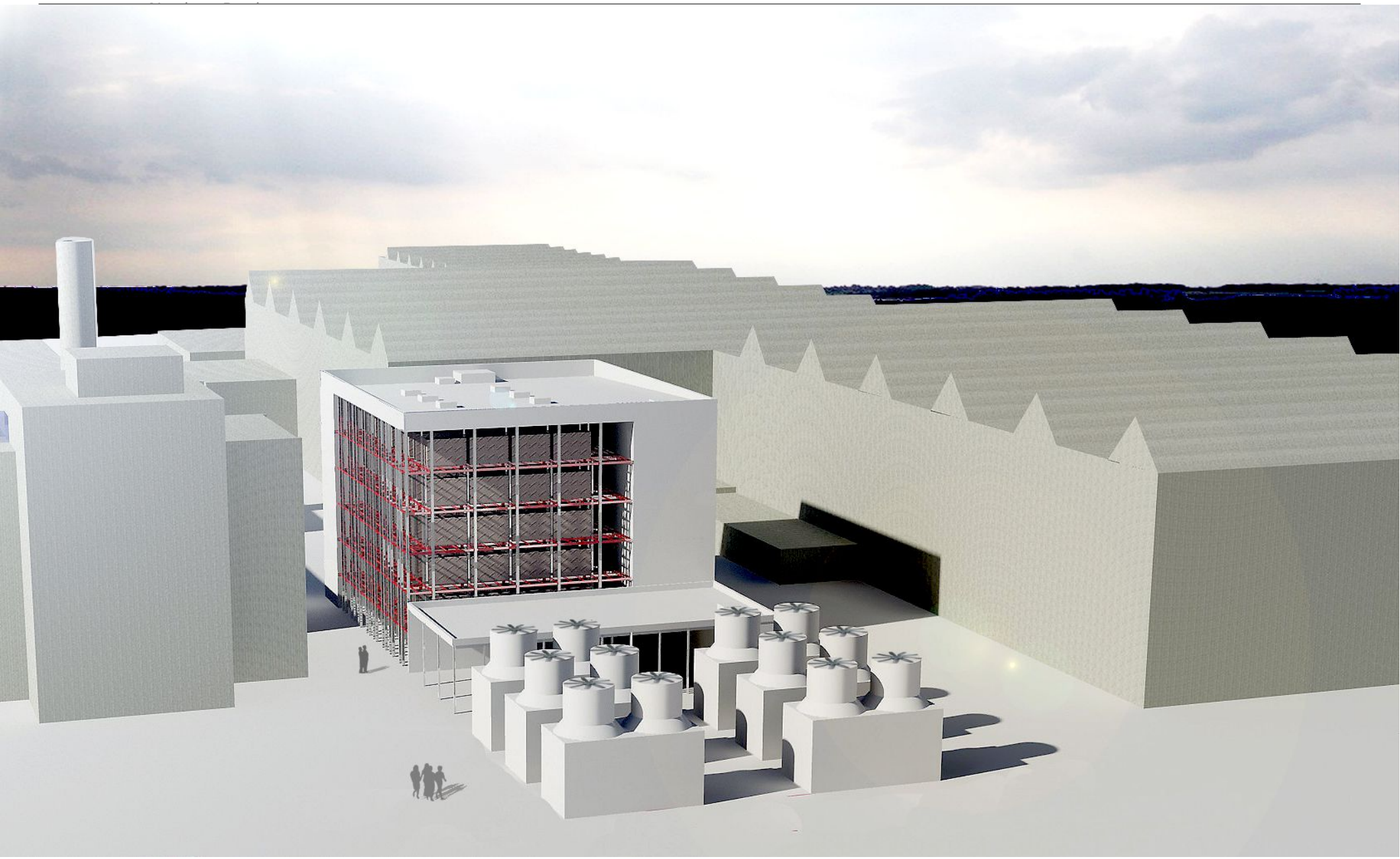
- about 300.000 cores number crunchers for FAIR
- about 50 Petabyte „online capacity“ for start of project
  => lustre
- „cache cluster“ for interfacing to long time storage
  => lustre?
- long time storage, archives Tape? Lustre with „RAID8“ :-) ??

=> new Data Center

„ green CUBE“

=> Cube design (award winning) => minimising cable length, latency

=> PUE ~ 1.06, passive cooling with „water towers“



Lustre at GSI: Hardware, Design, Kernel-Patches,Maintenance by HPC Department of GSI.....