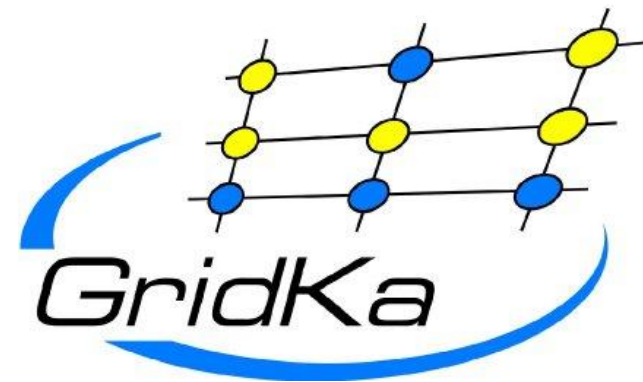


GridKa

Andreas Heiss

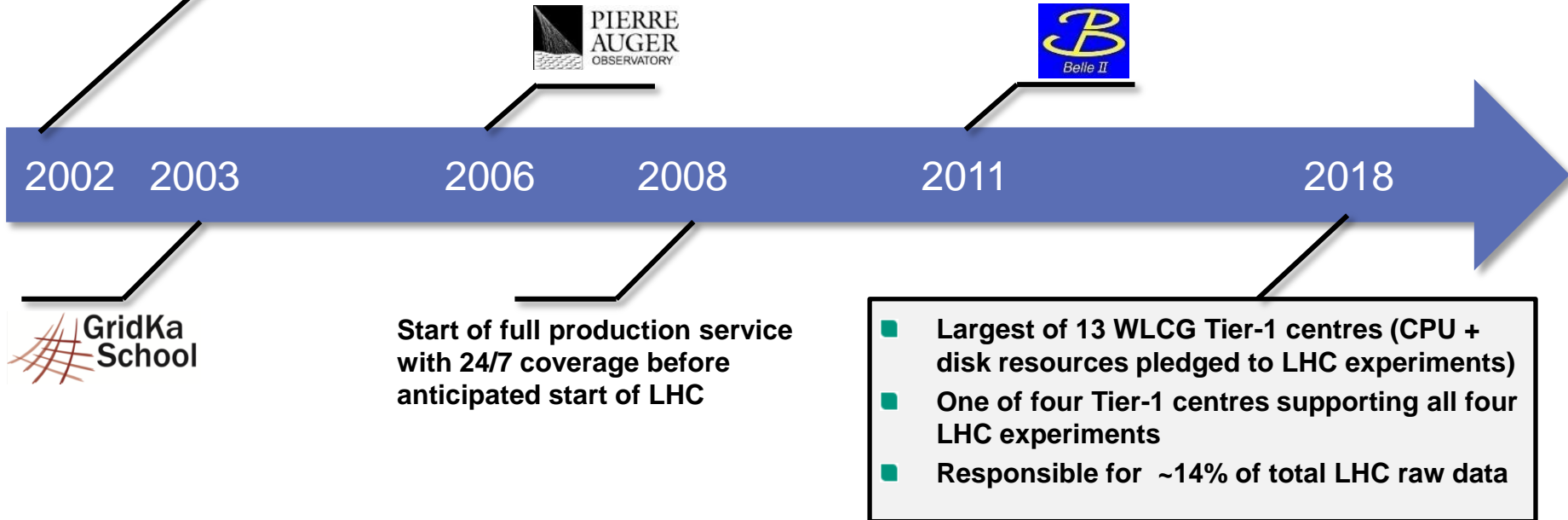


Helmholtz Research Field Matter

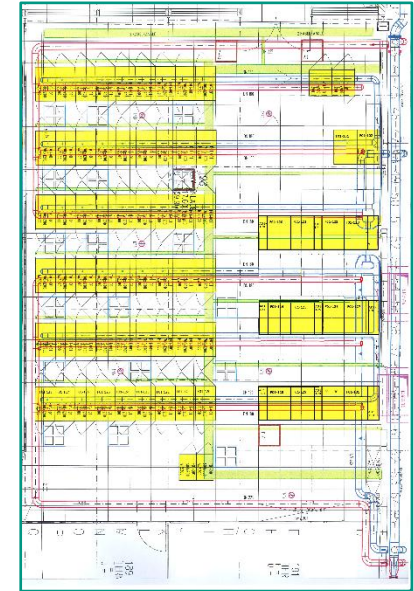
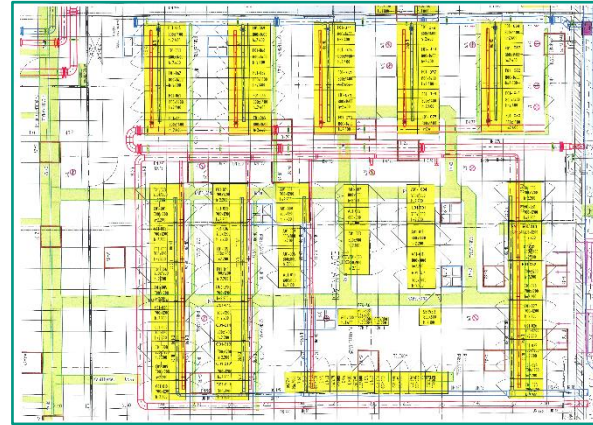


GridKa History

Initially supported experiments:



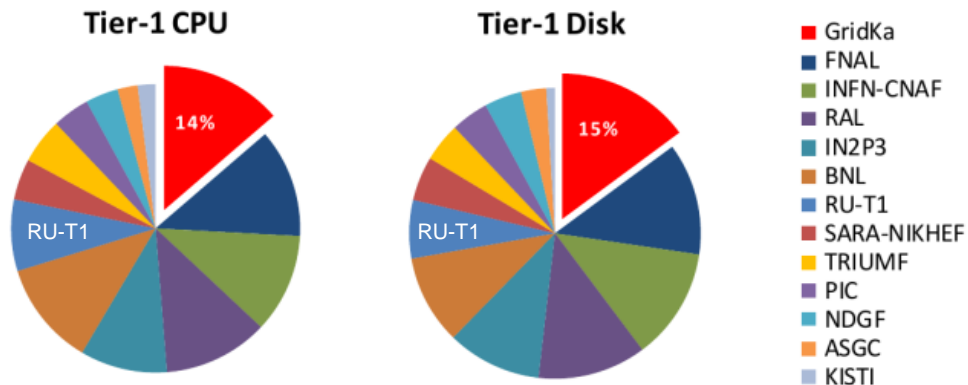
GridKa – A Helmholtz Research Infrastructure



- LK-II: Large-scale scientific user facility in Helmholtz
- Associated to the Research Program *Matter and the Universe*
- 2 server rooms + room for tape robots
- 150 racks, water cooled

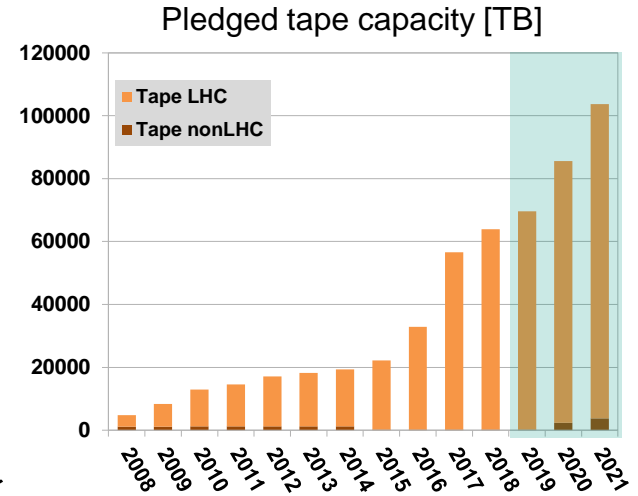
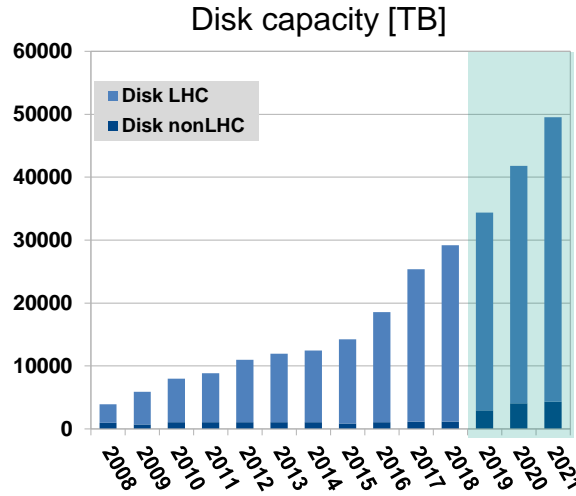
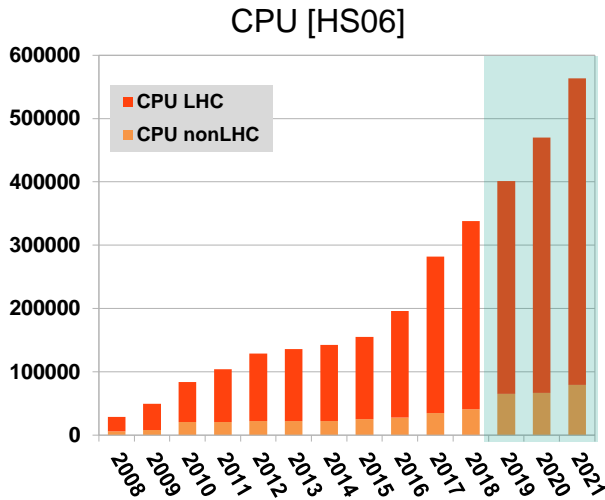
CPU and disk resource contribution to WLCG

2018 pledges to the LHC experiments



Number of cores	29,000
Number of compute jobs (2017)	23 million
Number of CPU-hours delivered (2017)	178 million
Disk space	23 PB (29 PB)
Tape space (used)	>40 PB

Resources: history and immediate future



■ ~20% per year resource increase supported until 2021

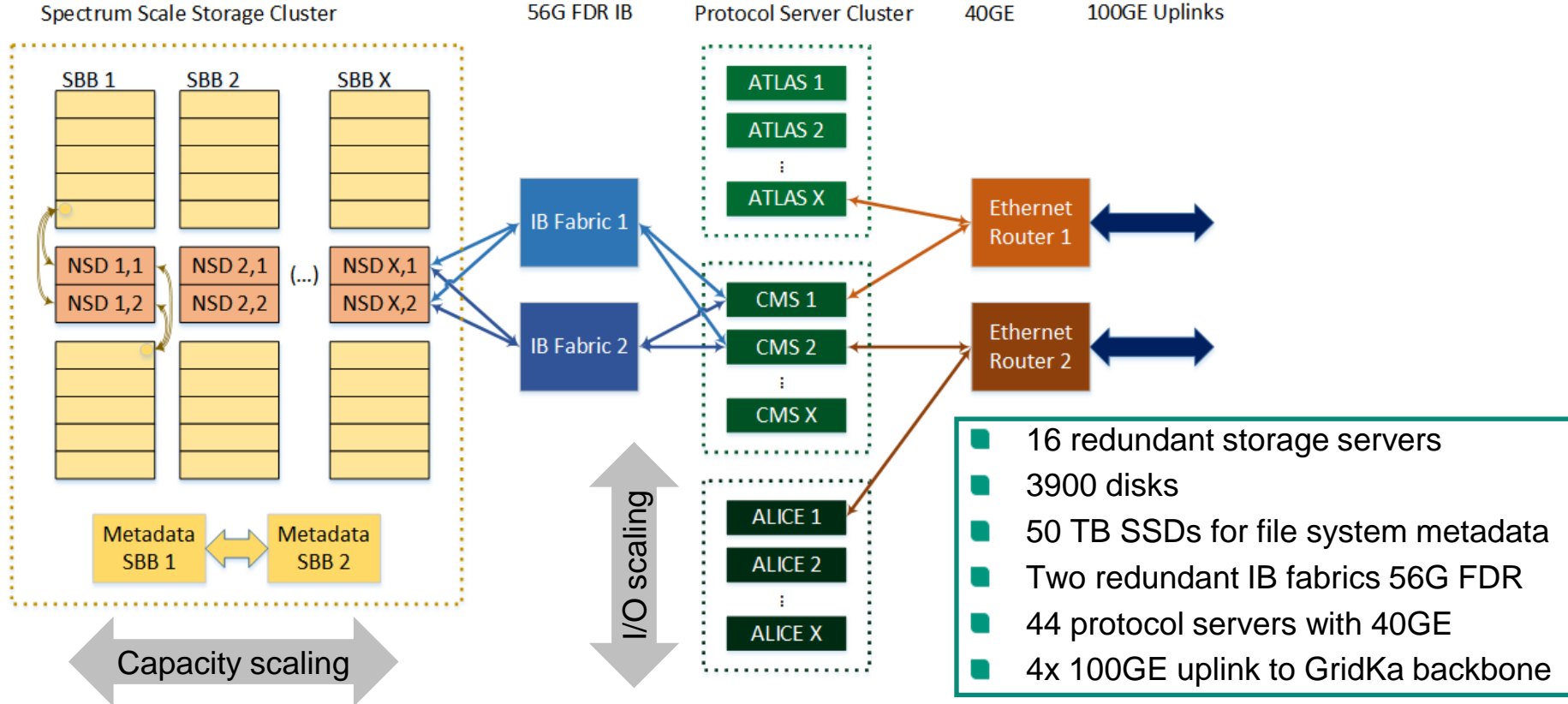
■ CPU +220 kHS06, Disk +20 PB, Tape +40 PB

Online Storage

- New storage system in production since 2017
- Storage system design allows for scalability both in size and performance
- Spectrum Scale (v4.2) software defined storage
 - Few large file systems for better scalability and manageability
 - Allows to adapt to various use cases (4 experiments, tape buffer, ...)
- Hardware: NEC GxFS
- Total capacity: 23PB (70 GB/s)
 - (extension to 29 PB and 100 GB/s installed)



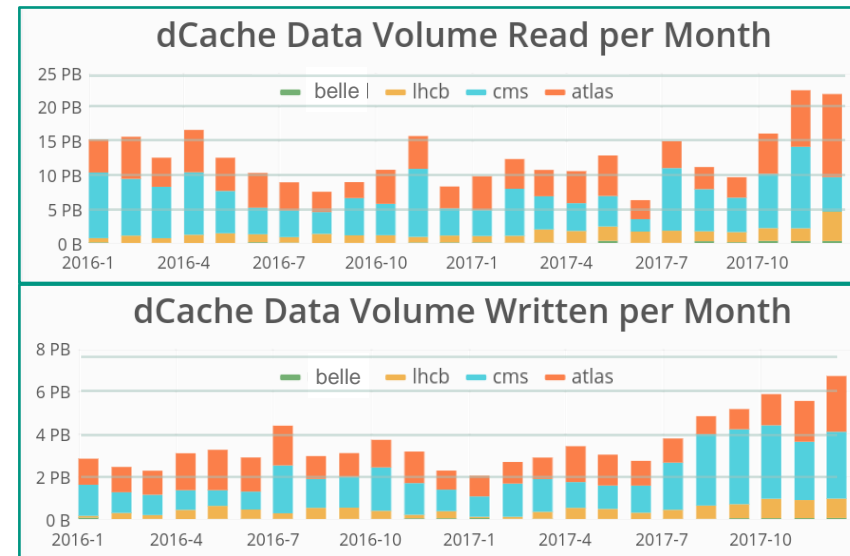
Online storage: Technical Design



- 16 redundant storage servers
- 3900 disks
- 50 TB SSDs for file system metadata
- Two redundant IB fabrics 56G FDR
- 44 protocol servers with 40GE
- 4x 100GE uplink to GridKa backbone

Online Storage – Grid enabled data management systems

- **dCache** for ATLAS, CMS, LHCb, Belle II, Auger, ...
 - GridKa is among the largest dCache installations
 - Excellent operations team @GridKa and cooperation with dCache.org (DESY)
 - Native support for protocol zoo (SRM, webdav, dcap, xrootd, gridftp, NFSv4.1)
- Native **xrootd** setup for ALICE
 - Supported by ALICE representative
 - Light-weight setup for large capacity
 - Benefits directly from parallel file system



Tape robots

- 3 locations



2 TS3500 Libraries
OKD building



1 TS3500 Library
KIT Campus South



3 SL8500 Libraries
SCC building

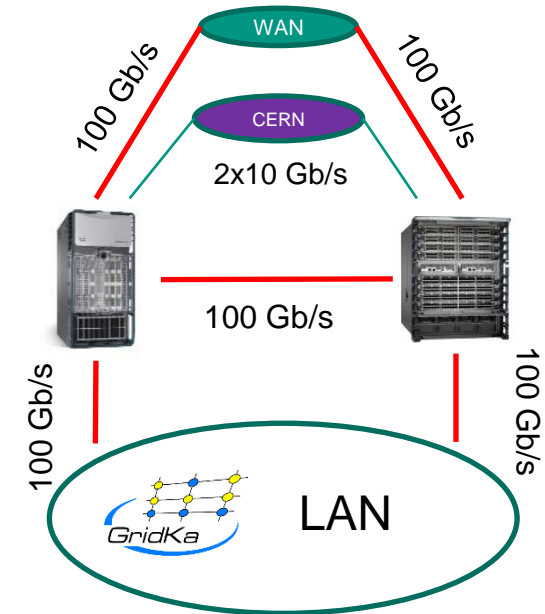
WAN & LAN Networking

■ WAN

- 2x10 Gb/s to CERN, upgrade to 100 Gb/s in 2018
- 100 Gb/s connection to DFN
 - Shared LHCOPN & LHCONE
- Switch from 100 Gb/s to 200 Gb/s once financially viable

■ LAN

- Core router setup with 4 routers
- 80-200 Gb/s backbone between routers
- Storage server connections 40 Gb/s (200 Gb/s per rack)
- WN connections 10 Gb/s (40-80 Gb/s per rack)



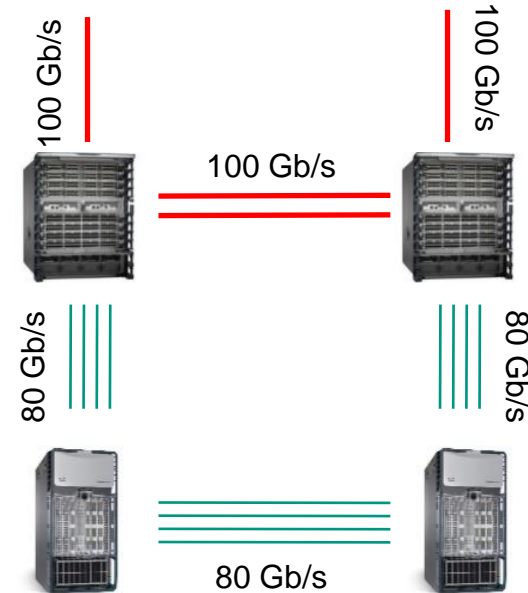
WAN & LAN Networking

■ WAN

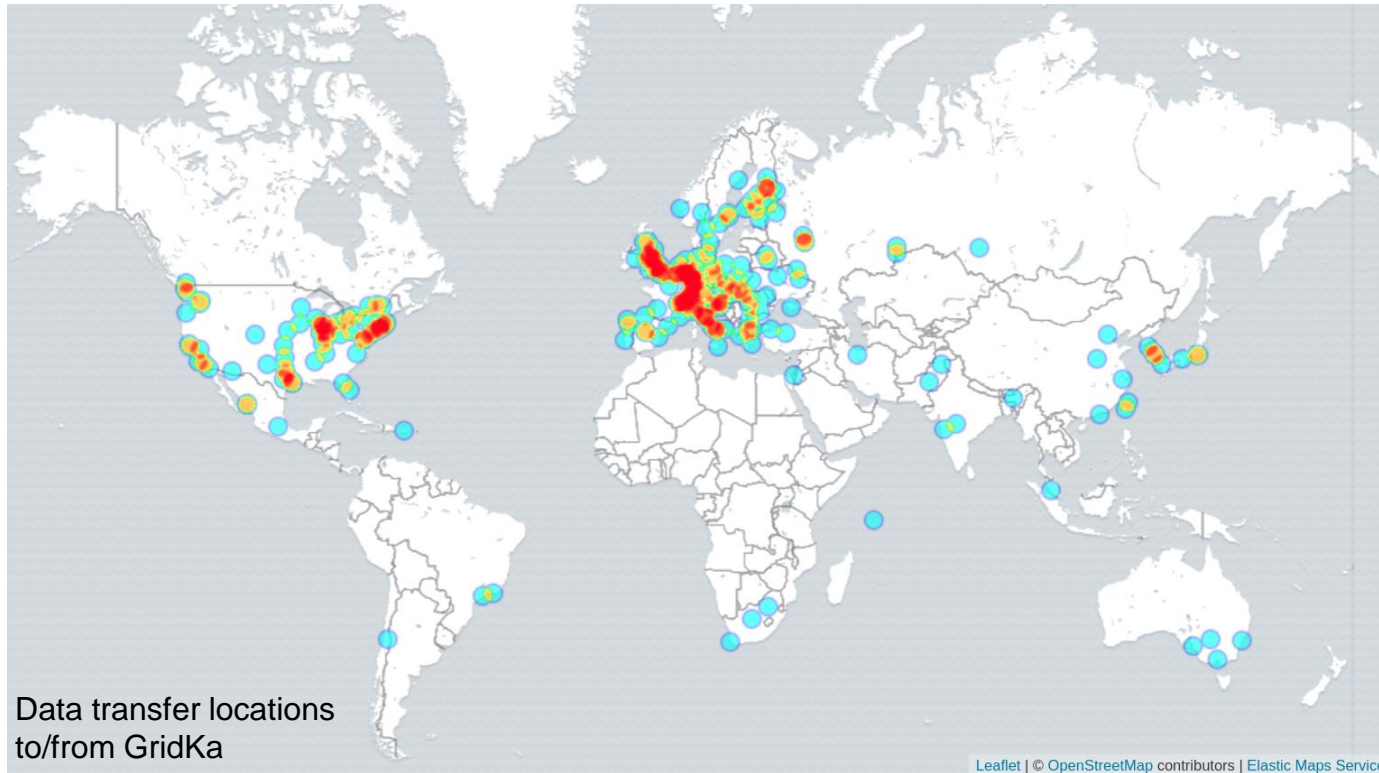
- 2x10 Gb/s to CERN, upgrade to 100 Gb/s in 2018
- 100 Gb/s connection to DFN
 - Shared LHCOPN & LHCONE
- Switch from 100 Gb/s to 200 Gb/s once financially viable

■ LAN

- Core router setup with 4 routers
- 80-200 Gb/s backbone between routers
- Storage server connections 40 Gb/s (200 Gb/s per rack)
- WN connections 10 Gb/s (40-80 Gb/s per rack)

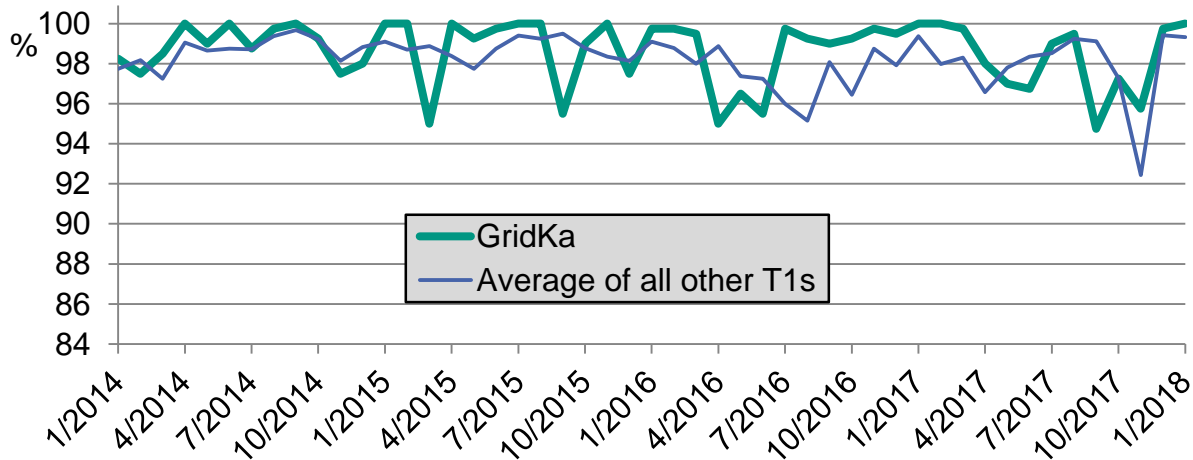


Data transfers



GridKa Service Reliability

■ WLCG measured reliability (average of the four experiments)



01/2014 – 01/2018: 98.6%

GridKa above average on
34 out of 49 months.

Interdisciplinary R&D at GridKa

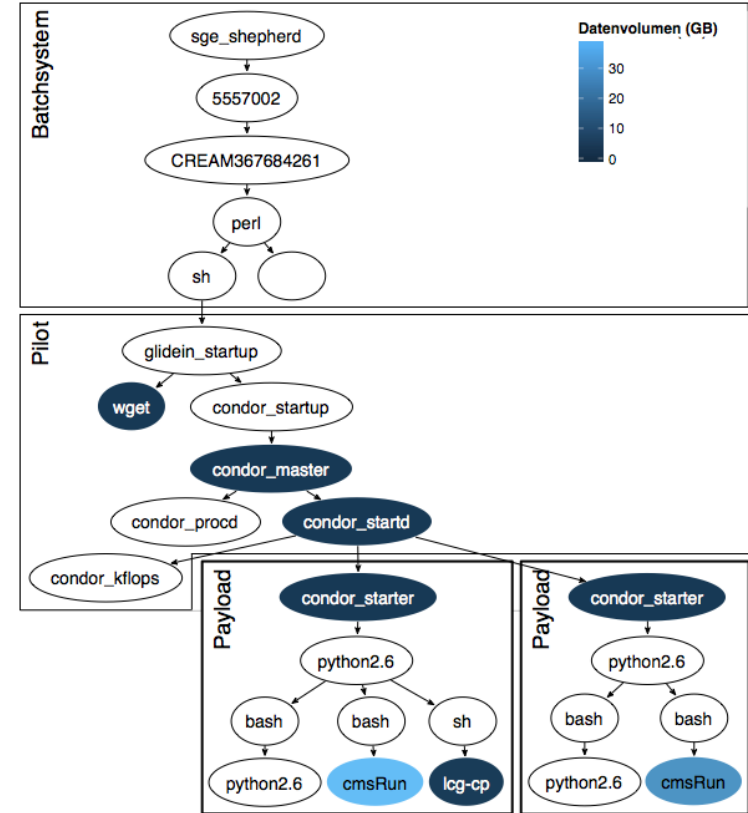
- R&D so far mainly in cooperation with
 - High Energy Physics community / WLCG
 - Particle physics CMS group of KIT
 - Helmholtz research program Supercomputing & Big Data

- Intend to strengthen R&D with astroparticle physics!

- The following slides show 3 examples of recent projects (PhD theses)

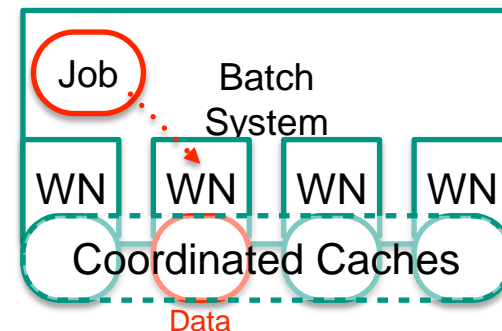
R&D 1: Online Analysis of Dynamic Streaming Data

- Developed a fine-grained monitoring of HEP payloads at GridKa
- Adopted computer science methods to classify payloads and discover anomalies
- Established a new method for online analysis of dynamic trees in data streams
- Results are an excellent foundation for further research topics

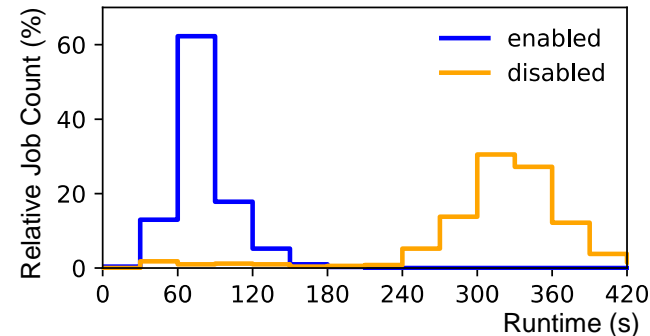


R&D 2: High Performance through Coordinated Caching

- Established new approach for user data analysis via distributed, coordinated caching on local SSDs
- Introduced data locality to HTCondor batch system
- Performance gain of factor 3-4 on typical recurrent end user data analysis payloads (prototype system)
- Third party funding for production system was recently granted (hardware)

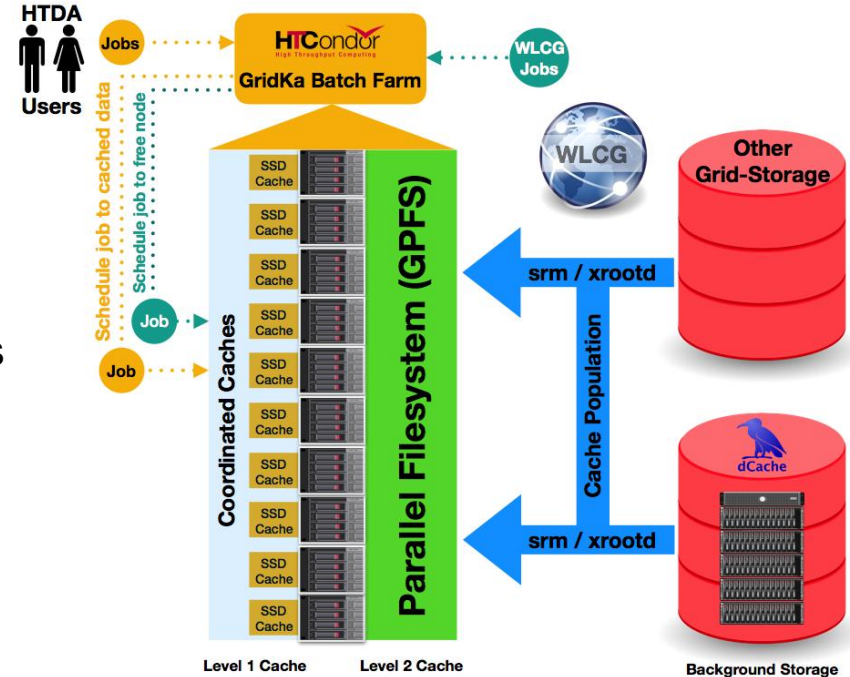


CMS Calibration Job Runtime



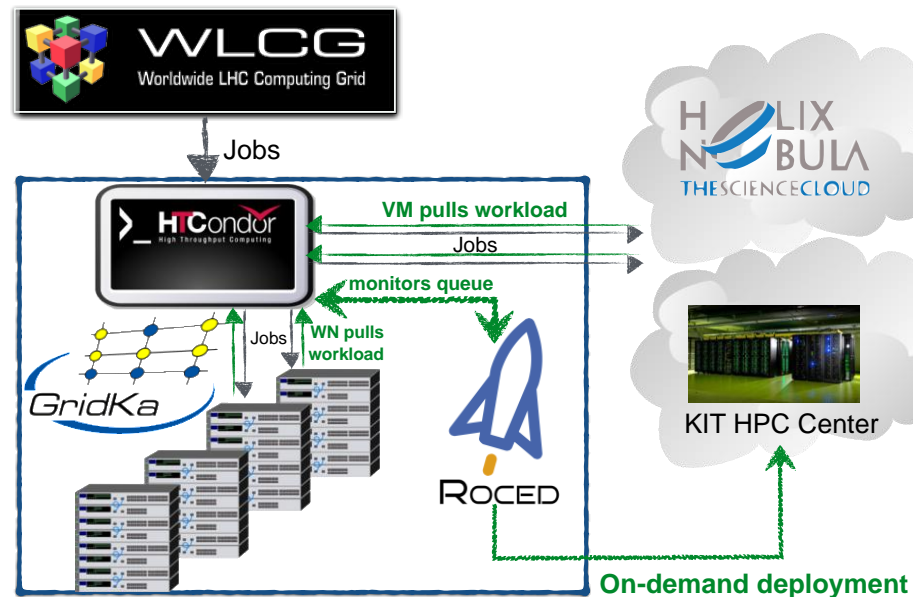
GridKa High-Throughput Analysis Extension

- Knowledge transfer from R&D to the GridKa production system
- Include high-throughput data analysis (HTDA) nodes into the GridKa batch farm
- Nodes can serve traditional WLCG jobs as well as HTDA user jobs profiting from caches
- Increase performance of typical recurrent end user data analysis payloads at GridKa



R&D 3: Dynamic Compute Expansion of GridKa

- Transparent on-demand integration of opportunistic resources using ROCED
 - Helix Nebula Science Cloud (based on traditional virtualization)
 - KIT HPC Center (FORHLR II) (based on container technology)
- Automated detection and redirection of suitable CPU-intensive workflows (relies on results by Eileen Kühn)



GridKa School

- Started in 2003 as summer school on Grid Computing
- Developed to a broadly positioned international IT summer school
- ~1700 participants (2003 - 2017)

Audience

- Graduate and PhD students, PostDocs
- Various scientific fields
- Participants from industry



GridKa School 2018: August 27 – 31

The Team

