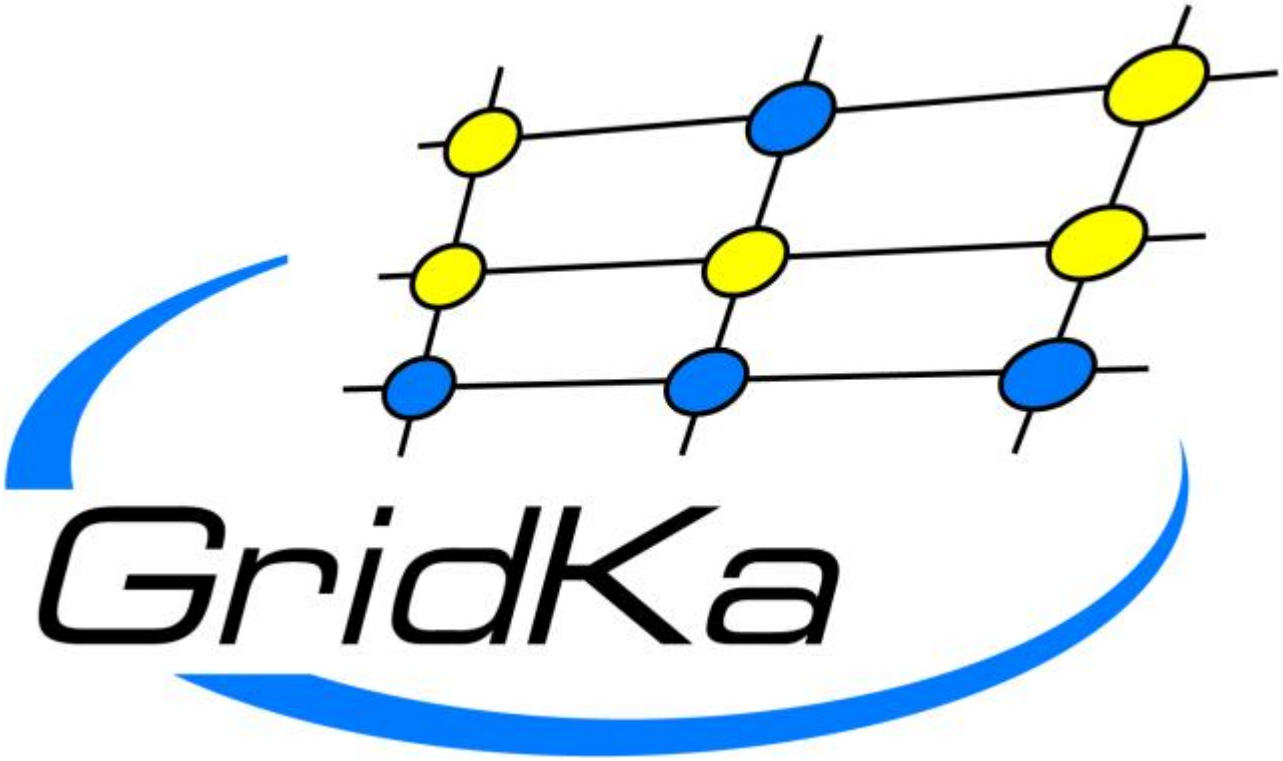
Karlsruhe Institute of Technology

KIT Site Report
HEPiX Spring 2014

Manfred Alef, Andreas Petzold, Nico Schlitter, Pavel Weber, Jos van Wezel

STEINBUCH CENTRE FOR COMPUTING – SCC

Grid Computing Centre Karlsruhe



GridKa Batch System

- GridKa cluster:
  - Dimensions:
    - **140 kHS06**
    - 610 worker nodes
    - 9,510 (physical) cores
    - 15,160 logical (hyperthreaded) cores
    - **12,800 job slots**
    - **1.5...2.5 million jobs per month**
  - Batch system:
    Univa Grid Engine (since mid 2012)
    - Robust
    - Performant
    - Very few issus (e.g. black hole nodes)

GridKa Batch System

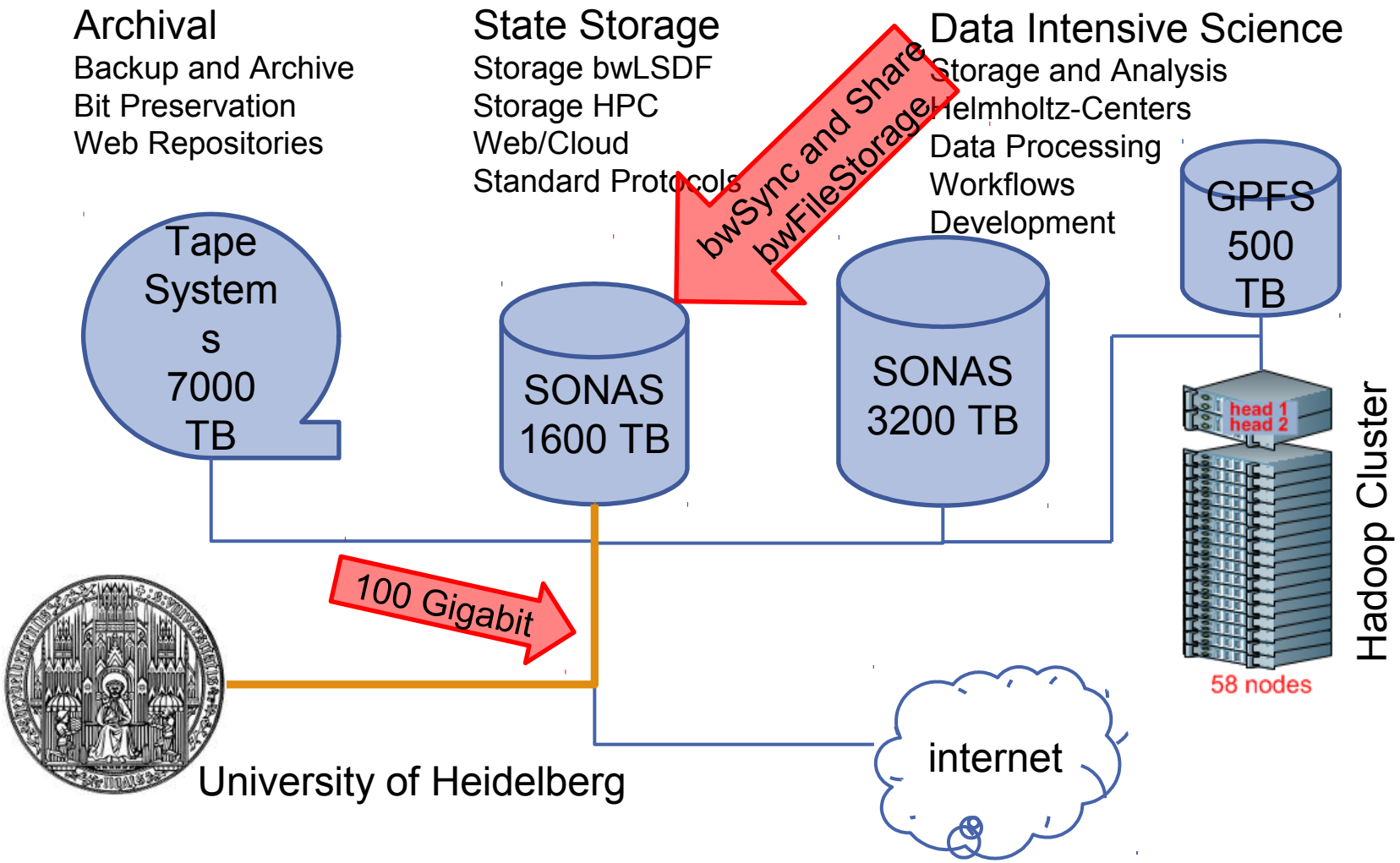
- GridKa cluster:
  - Multicore job support
    - Dynamic scheduling, no cluster partitioning
    - "Max_reservation" setup to boost pending multicore jobs
      - Limits number of multicore jobs on take-off ramp, not the total number of running multicore jobs
      - Setting: 10...20
    - Submit flag "`-R y`" required to enter reservations
      - Current Cream release doesn't add this flag to qsub call
      - Workaround: running cron job "`qalter -R y $list_of_pending_multicore_jobs`"
    - Degradation in cluster utilization of 0.5% (per 10 reservations)
      - No efficient backfilling because almost no pending jobs declaring the estimated run time
      - Wave-like job submission pattern by Atlas and CMS, number of running multicore jobs fluctuates between 0 and some 100

GridKa Firewall

- GridKa cluster:
  - Remote data access / federated storage
    - Remote data access gaining importance (Atlas FAX, CMS AAA, Alice)
    - Remote access by batch jobs can cause firewall congestions
      - Example: Alice ....
    - Monitoring of access pattern very important
    - Local monitoring of batch job network usage supplements experiment monitoring
      - → Talk by Eileen Kühn on Wednesday

Large Scale Data Facility

LSDF Storage Instances



Recent LSDF developments

- Archive infrastructure based on HPSS
  - Production start in Q3,2014
  - Various projects will use this for long time storage
    - i.e. GridKa, Federal, State, Helmholtz Association, KIT
- Extensive evaluation of CEPH
- Integration with state wide Shibboleth AAI
  - Users in Baden-Württemberg can use LSDF storage and archives using the registration of their home institute
- Added 2 PB disk storage in late 2013
- GridKa and LSDF are jointly using the tape infrastructure
  - 4 Libraries (IBM 3500 and STK 8500, LTO drives, T10K drives)
    - FW code of no longer available without vendor maintenance contract
    - This rules third party maintenance contracts: high costs

LSDF: bwSync&Share

- Dropbox-like data storage, but privacy aware
  - Data are stored at LSDF –
    hence, the service is subject to German law
- On-premise solution based on PowerFolder
- Production start: 1 January 2014
- Available to 450.000 students and scientists in the state of Baden-Württemberg
- Federated user management:
  - Users can use their home institution credentials
  - Shibboleth authentication for the webinterface and the desktop clients
  - bwIDM (→ Talk by Andreas Petzold, HEPiX Fall 2013, Ann Arbor)
- Clients: Windows, Mac OS, Linux, iOS, Android; HTML

http://de.wikipedia.org/wiki/Datei:Locator_map_Baden-W%C3%BCrttemberg_in_Germany.svg

Announcement: GridKa School

- Annual summer school for advanced computing techniques
  - September 1-5, 2014, Karlsruhe
  - http://gridka-school.scc.kit.edu/

Job Advertisement

KIT is hiring!

Karlsruhe Institute of Technology

- We are looking for physicists or computer scientists interested in
  - Large scale storage systems
  - Large scale data analysis
  - IT security (Linux, Windows, network)

  - Interested? ⟹ e-mail to: andreas.heiss@kit.edu

Questions, Comments?