

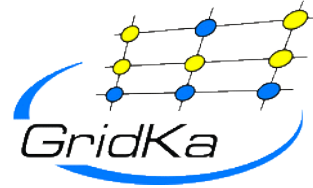
# KIT Site Report

Andreas Petzold

STEINBUCH CENTRE FOR COMPUTING - SCC



# GridKa Tier 1 - Batch Farm



## ■ Hardware

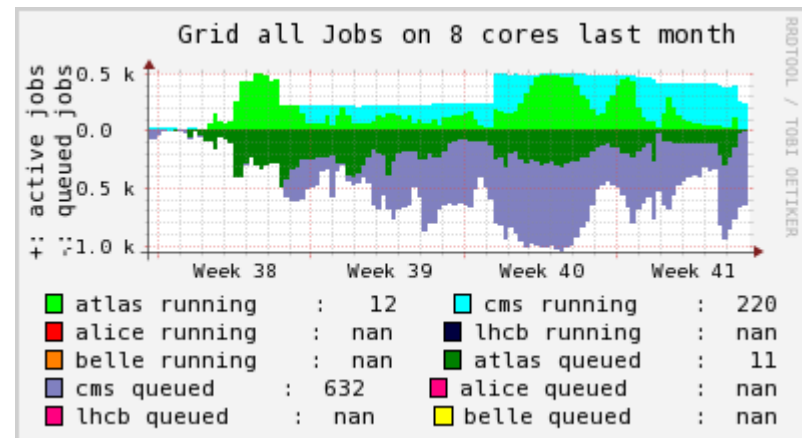
- 620 WNs, mix of Intel & AMD
- 13k job slots, 150kHS06 (2014)
- 42kHS06 will be retired, 52kHS06 new

## ■ Univa Grid Engine 8.1.7

- very smooth operation

## ■ multi-core jobs

- dynamic slot allocation,  
no static cluster partition



# GridKa Tier 1 – Disk Storage

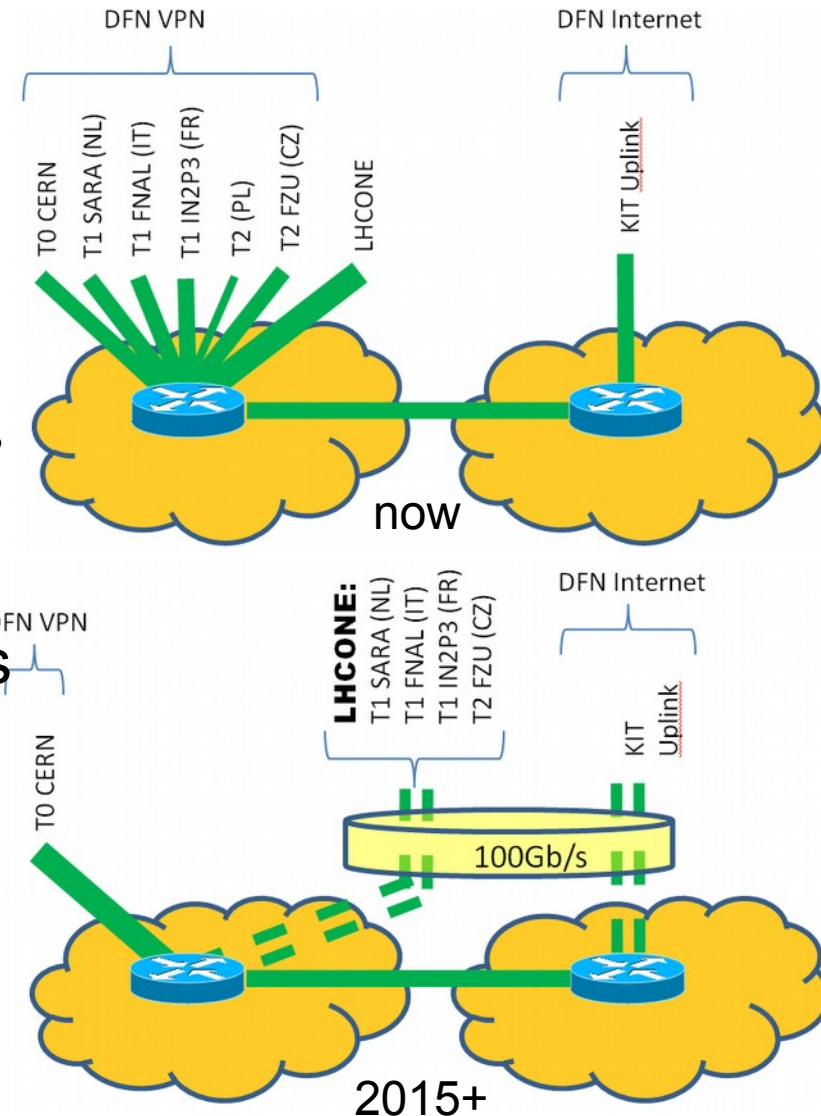
- 14PB on DDN systems S2A9900, SFA10k, SFA12k
- dCache for ATLAS, CMS, LHCb, Belle2, others
  - Version 2.6.33/34
  - plan to move to 2.10 by December
  - CMS
    - now using NFSv4.1
    - automatic file replication off → manual data set distribution
  - ATLAS & CMS instances integrated in FAX & AAA w/o xrootd proxy
- xrootd for ALICE
  - some servers already on version 4

# GridKa Tier 1 – Tape Storage

- 15.5PB used in 3 libraries (STK, IBM, Grau)
- tape technology
  - currently LTO3/4/5 with 45 drives, ~22k cartridges
  - 6 new T10kC drives, 400 cartridges in production by end of year
- tape management
  - currently TSM & ERMM
  - gradual migration to HPSS to start mid 2015
- HPSS status
  - intensive testing phase finished
  - dedicated new STK library
  - 1<sup>st</sup> user: HLRS (Stuttgart) will archive data to HPSS at KIT starting January 2015

# GridKa Tier 1 – Network

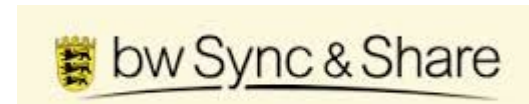
- currently 80Gb/s WAN connection based on 10Gb/s
- move to 100Gb/s based setup early in 2015
  - combine most existing connections into 100G link
  - keep 2x10Gb/s VPN to CERN
- challenge of remote data access (ALICE&FAX&AAA)
  - GridKa batch farm behind NAT
  - 15Gb/s available bandwidth often saturated



# Large Scale Data Facility



- storage and computing for non-HEP sciences
- 6.7PB disk; IBM SONAS + DDN based GPFS cluster w/ HPSS integration
- Hadoop cluster
- bwSync&Share in production since Jan 1<sup>st</sup> 2014
  - Dropbox-like service for universities and colleges in state of Baden-Württemberg
  - based on Powerfolder
  - steadily increasing number of users; currently ~8000
- bwFileStorage
  - central storage for HPC in Baden-Württemberg
- Storage for Human Brain Project
  - S3 based on DDN WOS systems
- Climatology data (ENES) replication using EUDAT B2SAFE service



# Federated Identity Management

- bwIDM
  - federated IdM for Baden-Württemberg
  - all state wide service integrate bwIDM
    - bwSync&Share, bwFileStorage, bwUniCluster, ...
  - Shibboleth and LDAP available
- extend beyond Baden-Württemberg
  - now working on extension to DFN-AAI (Germany wide)
  - integration of Umbrella

- federally funded project to enable industry and research partners to process their data on commercially available Big Data platforms
- communities: Industry 4.0, Energy, Smart Cities, Medicine
- 40 partners from industry and research
- project leader SAP

### Core Partners

Bayer Technology Services GmbH	Robert Bosch GmbH	Deutsches Forschungszentrum für Kognitive Intelligenz GmbH	EnBW Energy Baden-Württemberg AG	IBM Deutschland GmbH	Karlsruher Institut für Technologie (KIT)
Fraunhofer Institut Intelligente Analytik und Informationssysteme IAS	Forschungszentrum Jülich GmbH	SAP AG	Siemens AG	SOFTWARE AG	

### Other Partners — Research

Carl von Ossietzky Universität Oldenburg, Abt. Wirtschaftsinformatik 1 / VLBA	Deutsches Krebsforschungszentrum	Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO	FZI Forschungszentrum Informatik	RWTH Aachen, Joint Research Center for Computational Biomedicine	TU Berlin, Fachgebiet Datenbanksysteme und Informationsmanagement

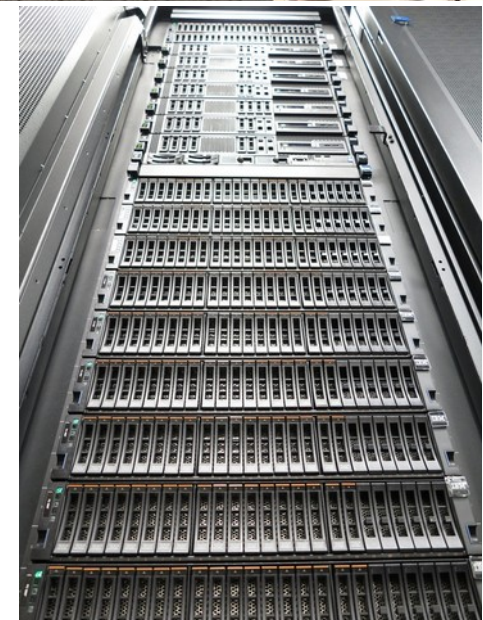
### Other Partners — Industry

ABB AG Forschungszentrum	BASF SE	Blue Yonder GmbH & Co. KG	BOGEN Electronic GmbH	CAS Software AG	Cogia GmbH
EBID Service AG	echobot Media Technologies GmbH	econda GmbH	Hitachi Data Systems GmbH	Huber Verlag für Neue Medien GmbH	Infineon Technologies AG
Leserauskunft GmbH	Microsoft Deutschland GmbH	PTV GROUP - PTV Planung Transport Verkehr AG	Stratosphere	TRUMPF Werkzeugmaschinen GmbH + Co. KG	Volkswagen AG



# Smart Data Innovation Lab

- platform operated by SCC
- SAP Hana
  - 4 machines: 2 IBM X3850, coupled with QPI interconnect, 80 (160 w/ HT) cores, 1TB RAM, 30TB HDD
- IBM Watson Foundations
  - Hadoop, SPSS Modeler, Watson Content Analytics
  - 7 Power8 SL822 machines, 140 cores, 4TB RAM total, 300TB HDD
- Software AG Terracotta & Apama
  - soon to come
- will be integrated into federated IdM



# HPC at KIT

- SCC operates HPC systems for KIT and state of Baden-Württemberg
- IC2 for Institutes of KIT (236 on Top500 2012)
  - 485 nodes, 162 TFLOPS, 32.6TB RAM, 470 TB Lustre
- bwUniCluster for general HPC work in BW
  - 520 nodes, 176 TFLOPS, 41.1TB RAM, 470TB Lustre shared with IC2, 230TB Lustre
- ForHLR1 for HPC research in BW
  - 530 nodes, 216 TFLOPS, 41.1TB RAM, 470TB Lustre shared with IC2, 230TB Lustre
- ForHLR2 for HPC research in BW
  - 45°C water cooling, free cooling
  - installation mid 2015

# New Building for ForHLR II





# We are hiring!

- We are looking for computer scientists and physicists interested in
  - IT Security
  - Large Scale Data Management
  - Large Scale Data Analysis
- Contact [andreas.heiss@kit.edu](mailto:andreas.heiss@kit.edu)