

About Belle II Computing at DESY

A Brief Overview

Andreas Gellrich

Belle II Germany (FSP) Meeting 2023

Computing Session 2023-09-26

Contents



- Collaborative Services and Tools (B2CS)
- Belle II Computing Facilities at DESY
- Belle II Data Analysis at DESY



Collaborative Services and Tools at DESY

Collaborative Services and Tools (B2CS)

Motivation



Why are collaborative services and tools needed?

- **Communication** is essential for any (virtual) collaboration
- Sharing and exchanging **information**
- **Integration** of a global virtual communities
- **Administration** and **organization**



Collaborative Services and Tools (B2CS) cont'd



Design

- Access with personal **credentials** (account/password) only
 - Centrally managed users
 - Membership management system
- As little as possible self-developed and self-maintained products
- Services embedded into a existing **IT infrastructure**
- Usage of **state-of-the-art** technologies
- Individual (web) **services** may run on virtual hosts
 - Centrally controlled updates and maintenance
 - Secured web services with certificates
 - Demilitarized zone (DMZ) in the network



Collaborative Services and Tools (B2CS)



Infrastructure

- DESY IT Infrastructure

- User registry (accounts)
- ZMS (web pages) (www.belle2.org)
- Sympa (mailing lists)
- Indico (agenda)
- GitLab (repo, issue tracking)
- Confluence, agira, bamboo, stash, gitlab



- XEN-infrastructure for Belle II

- 4 x [HV (20 cores), SSD (447GiB), SAS (237GiB)]; NetApp (800GiB) (>= 5y)
- ~35 VMs; systems (OS, cert, storage) provided and maintained by DESY (AG)
- authors, b2mms, calibration, chat, docs, evdisp, dqm, elog, mirabelle, pxd, rundb, questions, shift, software ...
- Buildbot (EL7/8/9, Ubuntu 20/22); DB squids



Collaborative Services and Tools (B2CS) cont'd

Membership Management System (B2MMS)



The B2MMS plays a central role

- Keep track of all **users** who are connected to Belle II
- Acts as an interface to the DESY **registry** for credentials (account/password) to log in
- Source all Belle II member related **issues**
 - Mailing lists
 - Author lists
 - Voting
 - Shift lists
 - M&O costs
- **Delegation** of the management of users to the institution representatives (IR)
- Based on the DESY Identity and Access Management (**IAM**) (*Apex*)
- Developed at DESY in 2017; in production since January **2018**

Collaborative Services and Tools (B2CS) cont'd

Experiences



- **Integration** into a production-grade IT infrastructure at DESY
 - Adoption of recent security **standards**
 - Implementation of **state-of-the-art** technologies
 - But: Utilization of professional software creates **dependencies** (licensing, cost models)
- B2MMS is a powerful tool
- Belle II self-made services needs sustainable maintenance models (**manpower!**)
- **User support** is known to be a non-scaling crucial issue
 - Not reading documentation (to the end)
 - Forgotten **passwords** (recovery procedure in place!)
 - Unmaintained and out-dated information in the **wiki**

A tool doesn't replace a concept.



Belle II Grid Computing at DESY

Belle II Grid Computing at DESY

Infrastructure



HEP is data-oriented.

- Grid

- **Multi-VO** site w/ **federated** resources and **opportunistic** usage
- **Home** of the VOs ICECUBE, ILC, ILDG
- Tier-2 for ATLAS, CMS (one of the largest world-wide); **Raw Data Center** for Belle II
- *HTCondor*, EL7, *apptainer*; ~20k cores == 320kHS06; 4GB/core; hardware (<=7 y)
- Mass storage **dCache**-based (SRM, XROOTD, DAVS, HTTPS)
- **tape** back-end optional
- Common name space `/pnfs/desy.de/_VO_NAME_/`

Jobs are transient,
Data are persistent.

- (National) **Analysis Facility** (NAF)

- *HTCondor*, EL7, *apptainer*; 9.5k cores
- See the mass storage via the common name space

Analysis facilities complement the Grid.

- Network

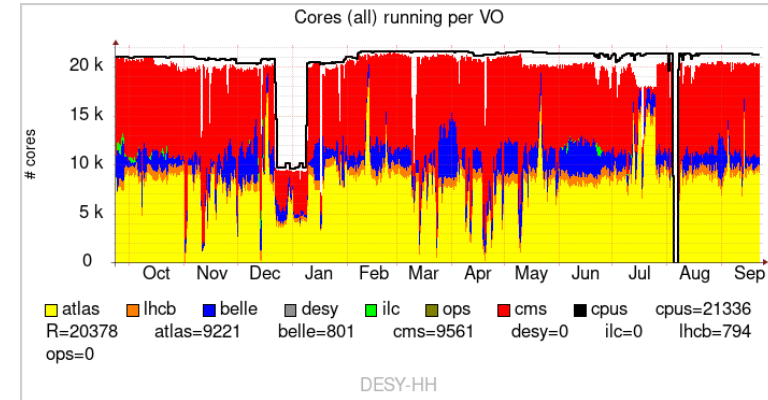
- 2 * 50Gbps WAN (X-WIN/LHCone) (since 2019)

Belle II Grid Computing at DESY cont'd

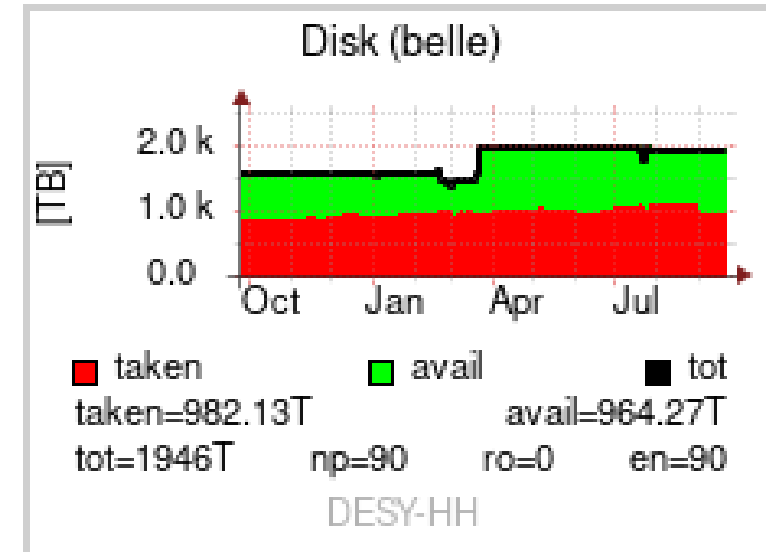


Belle II Raw Data (Grid) Center

- CPU pledges: ~30kHS06
- Usage: 29kHS06 (Apr-Dec 2023)
- Contribution: ~20% of total Belle II Grid computing



- Disk pledges: 2001TB (47% of German fraction)
- Disk usage: 982TB
- Tape usage: 107TB tape

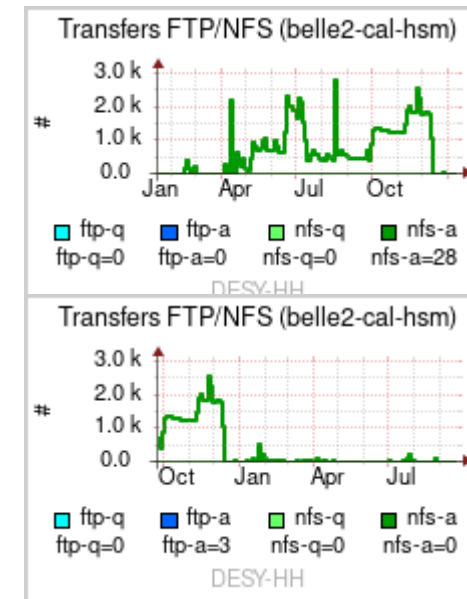
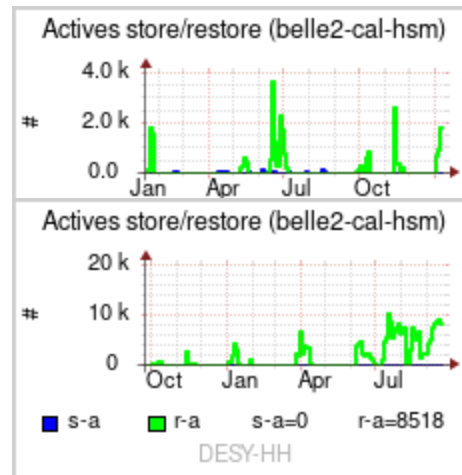
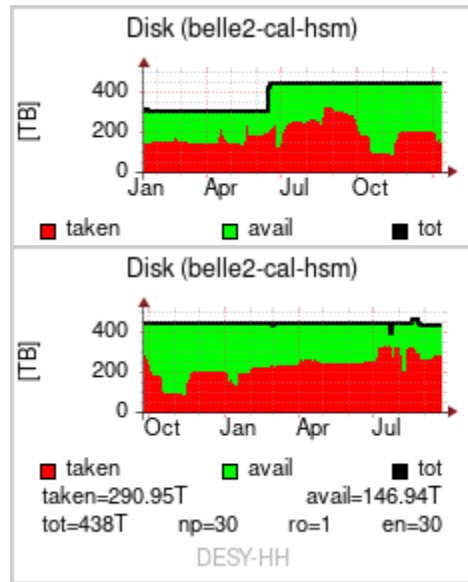


Belle II Grid Computing at DESY cont'd



Belle II Re-calibration Center (deprecated)

- Planned and funded for 2021/22
- CPU usage: 4kHS06
- Disk storage: 451TB
- Tape storage: 921TB (BTW: restoring from tape is a nightmare ...)

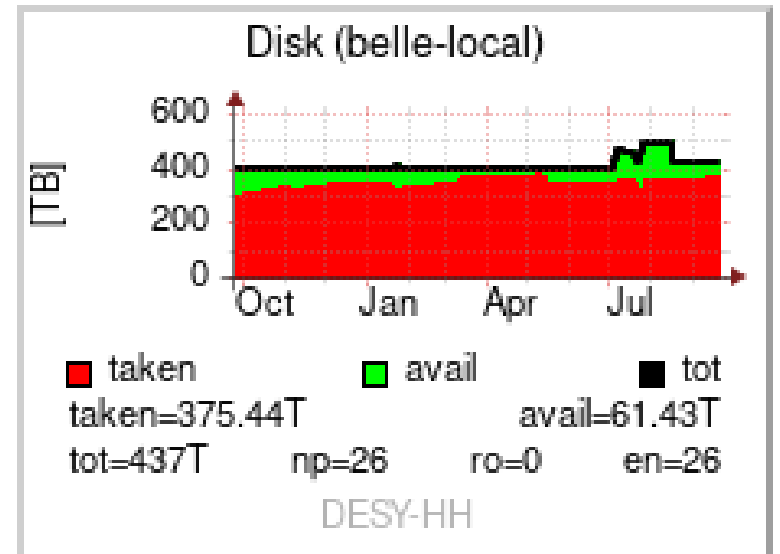
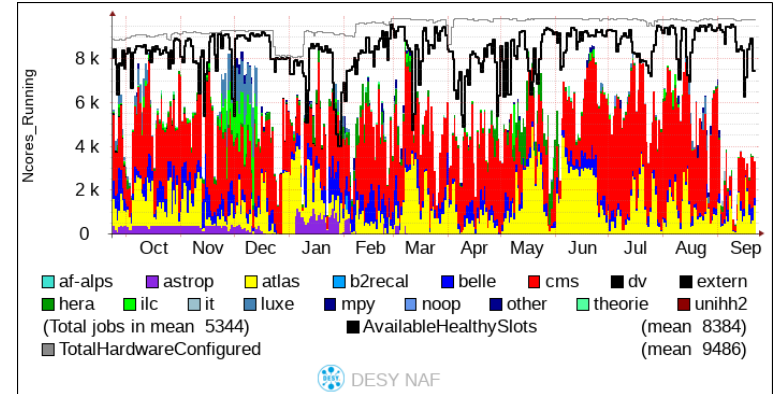


Belle II Data Analysis at DESY cont'd



Analysis Facility

- Belle II @ NAF:
 - Users:347 countries:23 institutions:75
 - Europe:210
 - Germany:~150 Italy:~30 U.S.A.:~30 France:~20 Japan:~20
 - DESY:~50 LMU:~30 Bonn:~20 HEPHY:~20 Karlsruhe:~15 CPPM:~10
 - JupyterHub: 30 Belle II users
 - Groups: WG1, WG2, napoli, tau, hephy
- Usage:
 - CPU cycles: 9kHS06
 - DUST (GPFS): 160TB
 - Disk storage: 437TB
 - ~50% data and MC
 - ~50% user files





Belle II Data Analysis at DESY

Belle II Data Analysis at DESY



- Current computing model considers interaction with MDSTs via Grid client tools
- On the Grid batch-like processing takes days
- Good for massive processing. Bad for initial stages of the analysis, prototyping.

- Analysis Facilities (AF) are not part of the Belle II computing model
- The multi-VO DESY AF (NAF) is a complement to the Grid allowing for interactive usage
- Quick and frequent iterations
- As HEP is data-oriented AFs must be close to a large Grid site which provide data
- O(TB)

- JupyterHub (with limitations though)



Belle II Data Analysis at DESY cont'd



Datasets at DESY intended for analysis (Michel Hernandez Villanueva)

- Data and MC samples copied to NAF on request of the analysis teams.
 - 91 TB of data, 76 TB of Belle II MC (out of 1 PB available on Grid) copied on demand.
 - Samples from obsolete reprocessings periodically removed.
- Analysts request samples after data-taking and prompt calibration
 - Samples from major calibration reprocessings copied ~ once per year.
 - **Major scalability issue**, as rely on manual operations.
- **Analysis facilities are not part of computing model of Belle II**
 - Grid → NAF downloads on user request, manually operated
 - **Scalability** will become a problem w/ high(er) luminosities
 - Same problem expected at Japanese, Italian sites ...
- LHC experiments are moving towards **analysis facility** concepts
 - NAF is already used for LHC experiments as well
 - It is all about automatization of data distribution

Again: **HEP is data-oriented!**

Conclusions



- Belle II **Collaborative Services and Tools**
 - Essential for today's work in Belle II
 - Compromise between IT infrastructure and self-made service

- Belle II **Grid Computing**
 - Utilization of the large multi-VO DESY Grid site
 - Federated resources and opportunistic usage → large contribution to Belle I Grid computing

- Belle II **Data Analysis**
 - Analysis Facilities (AF) are not part of the Belle II computing model
 - The multi-VO DESY AF (NAF) is a complement to the Grid allowing for interactive usage
 - As HEP is data-oriented AFs must be close to a large Grid site which provide data

Backup

Collaborative Services and Tools (B2CS) cont'd



Legacy

- Existing data and information must be migrated

Security

- Data and services must be protected; authentication and authorization for all users

Reliability, Stability, Availability

- Experiment must be operational at any time and integrity of data and information is crucial

Flexibility

- New requirements demand new features or new services and tools

Sustainability

- The collaboration will exist for >10 years

Open Issues



- Personnel (Daniel Knittel will be on parental leave Feb – Aug 2023)
- Volunteers needed (*RocketChat* 'chat.belle2.org', eLogbook)
- Plans (priorities) for services/data in case of potential (partial) shutdown due to energy market costs
- Re-calibration on the Grid as planned from 2023 on?
- Belle II analysis facilities ('BAF')
 - Belle II Data Model
 - Distribution of physics topics to other 'BAF's?