



Leonardo Sala :: AWI :: Paul Scherrer Institut

7904 - review and outlook

AWI Department update meeting 2024.01.09 / PSI

It's a long way to Tipperary

3 years lived dangerously

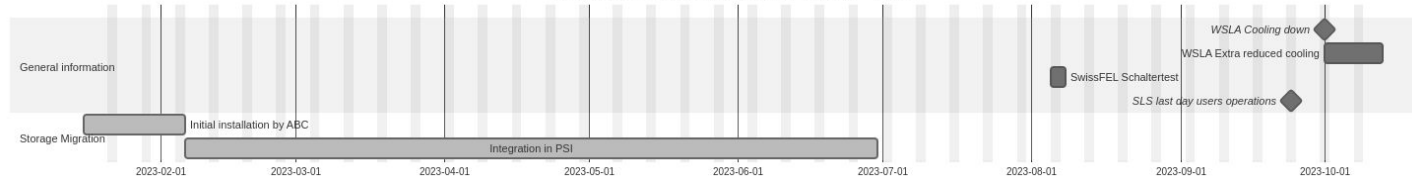


WSLA → WHGA migration overview

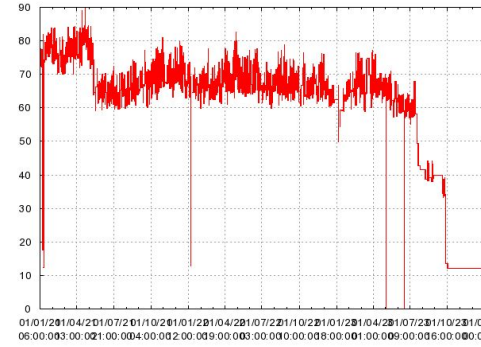
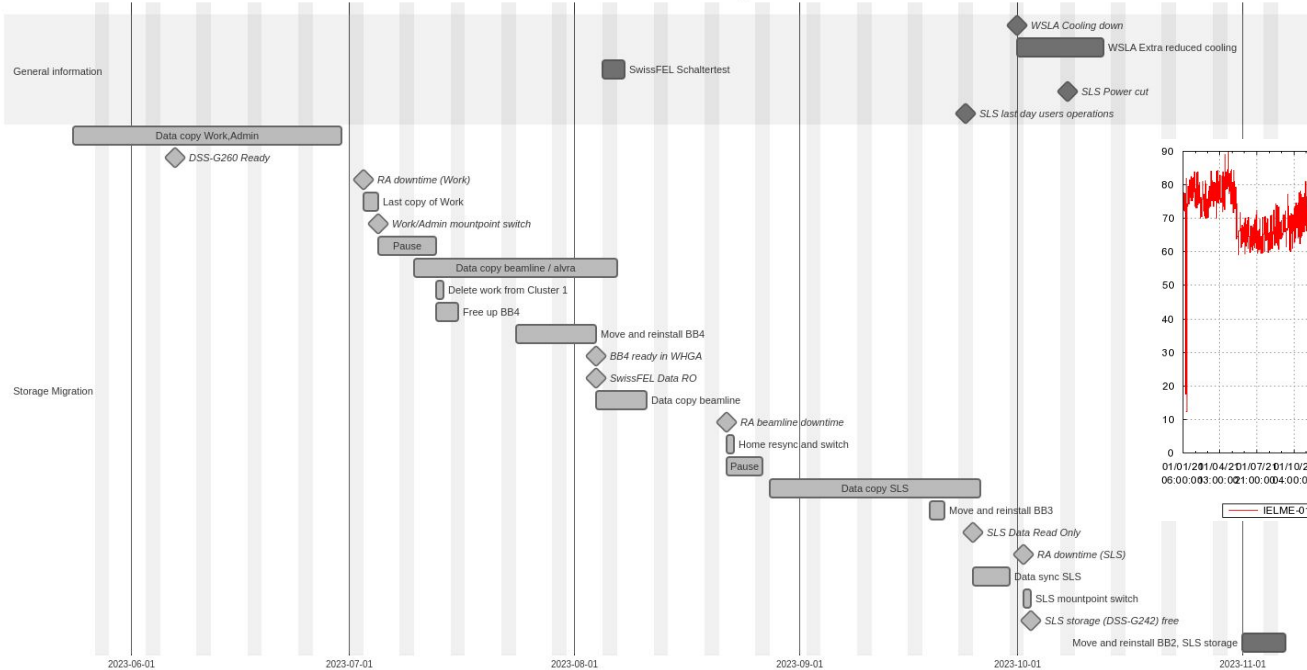
- Preparations started in **2020** ("Long range Infiniband")
- **Mid-2021**: discussions and decisions about new Cold Isle in WHGA. AIT coordinated all the efforts
- **2022**:
 - June: Racks installed and usable, start installation network
 - procurement of new Ra Storage
 - Ra in stretched mode cluster
 - August-September: more Infiniband tests
 - September: Ra in stretched cluster mode
 - requested 72 fibers from WSLA to WHGA
- **2023**: the fun begins
 - including acquisition tests from Gigafrost in June (kudos to L. Hax, Controls)
 - 15th December: WSLA empty

Busy times

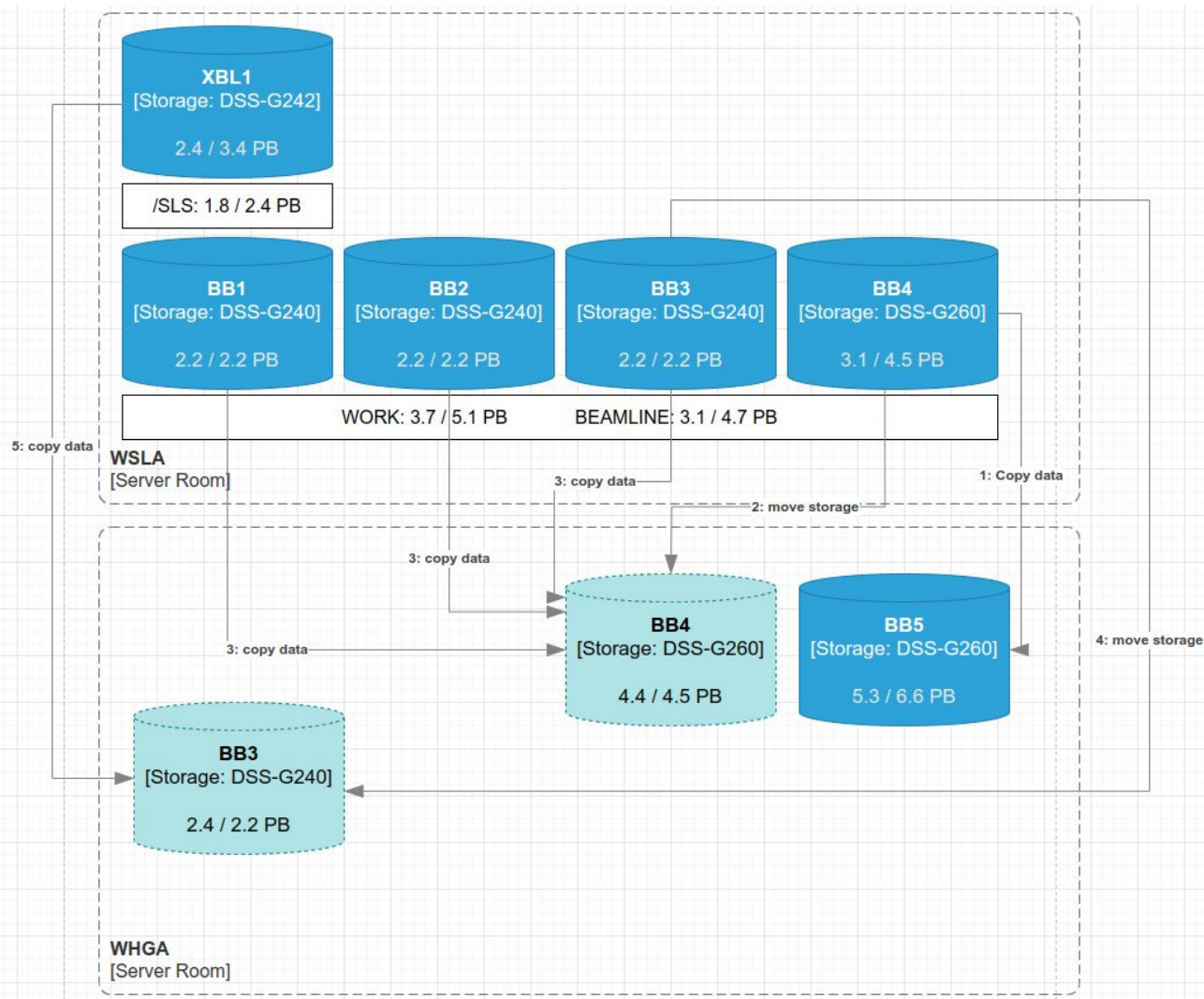
WSLA to WHGA server room migration - Part I



WSLA to WHGA server room migration - Part II



IELME-0101-EP001 WSLA-Serverraum Wirkleistung [kW]



Not included in the schema:

- Move of XBL 1 to SwissFEL
- Reconfiguration of AFM-based data copy at SwissFEL
- Borys single-hand fight with an enclosure
- All the horrible details of the Infiniband migration

What just happened?

Can you power it off and on again?



Schaltestest 2024 / I

Multiple long-prepared operations on **Ra and SwissFEL**:

- Slurm upgrade: v21.08 → v23.02
 - including RHEL8 reinstallation of service nodes
- Migration of Slurm DB server to RHEL8
- JupyterHub upgrade: v1 → v4 (and RHEL8)
- Various recabling, last modifications to Infiniband fabric
- Integration of old supported SLS nodes on Ra (as RHEL8 nodes)

Schaltestest 2024 / II

On Ra:

- First step to full RHEL8
 - default **partitions** are now RHEL8 (day, week, ...)
 - RHEL7 queues will exist until March / April (day-rhel7, ...)
 - new queue `jupyter`: minimal "guaranteed" resource for every user to run a notebook (1 CPU, 5 GB RAM)
 - default **login nodes** are RHEL8
 - RHEL7 login node can be selected as: `ra-1-003`
- Introduced flatter Slurm accounting to improve compute usage
 - users are not assigned a beamline anymore
 - removal of beamline-specific limits

Other stuff happened.



What happened in 2023

Storage

- first SSD-only storage system delivered at SwissFEL
 - 30 GB/s, ~100 TB
 - can be doubled by adding more drives
 - possible usage (among the others):
 - detector buffer
 - fast analysis cache
- discussions with IBM on purchase systems through ETHZ Framework contract
 - first two HDD system should be in production by summer 2024
- a lot of work and discussions on data retention policies
 - including writing a new python library from scratch

What happened in 2023 - cont'd

Compute

- lifecycle, lifecycle!
 - No, we won't keep 10 years old hardware unless necessary
- RHEL8 transition - at the moment ~45%
 - testing Icinga2

Services

- Tape -> Ra retrieve
- SwissFEL data deletion service
- New documentation on gitpages:
<https://hpc-sysadmins.gitpages.psi.ch/index.html>
- Update of our Data Center inventory

And now what?

Outlook



Outlook / I

- **Finalize RHEL8 migration**
 - Ra before April 2024
 - SwissFEL possibly before August shutdown
 - SLS will be installed as RHEL8 (maybe even RHEL9)
- **WHGA server room**
 - add more racks (up to 4)
 - hope we don't have to remove a rack to install storage anymore
- **Resources management**
 - start implementing REST API
 - get retention policy out
 - explore DUO integration (IT as part of beamtime resources)
 - explore GPFS and Slurm REST APIs

Outlook / II

SLS 2

- start implementing storage in WHGA by Summer 2024
- finalize Ra shared compute resources infrastructure
- DAQ: support detector integration tests
- continue support for SLS beamlines (some are still taking data)

SwissFEL

- integrated old SLS storage to replace aging Aramis storage
- integrate new SSD storage
- finalize pipelines for DAQ deployment

Outlook / III

Overall, continue to automatize and remove ourselves from plain sight

- e.g. Ra is supported by less than 1 FTE (*everything*)
 - this includes also user support from Alain, Ivan and Sven - THANKS!
 - or e.g. SLS was supported without dedicated FTEs
- we need to **delegate** as much as possible to the users or the systems
 - automatic Slurm reservations
 - automatic data deletions
 - self-service resources requests
 - special software deployment, e.g. DAQ
- a lot of good work already done
 - streamlined hardware installation
 - automatic tape archive on SwissFEL
 - automatic warranty expiration checks
 - ...



Or remove services :)

Questions?



