bw|HPC – S5

Remote Computing Services: bwCloud, bwHPC, and Beyond

Social Science Data Lab

Hendrik Winkhardt (bwHPC-S5, University of Mannheim)
winkhardt@uni-mannheim.de

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

UNI FREIBURG

UNIVERSITÄT HOHENHEIM

Hochschule Esslingen
University of Applied Sciences

Universität Konstanz

UNIVERSITÄT MANNHEIM

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

KIT Karlsruher Institut für Technologie

ulm university universität ulm

Motivation

- At some point, the personal computer is no longer feasible
  - Excessive computation time
  - Lack of hardware
  - Lack of software

    → Buy a better laptop

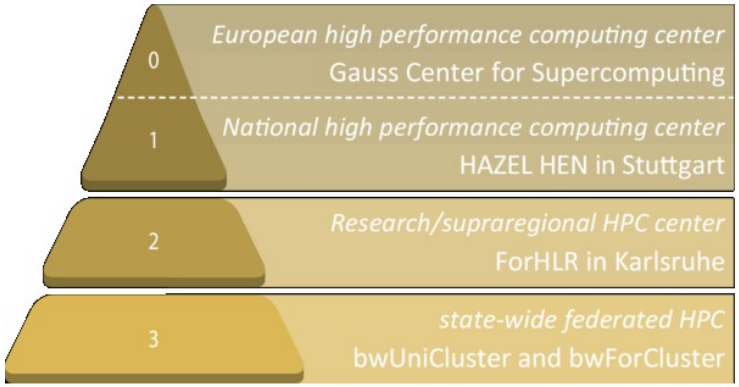
    → Improve your code

    → Get your resources elsewhere
  - An HPC-system: Access to a supercomputer
  - „The Cloud“: Server-based solutions: Virtual Machines etc.

- Data-intensive vs. Computationally expensive

bwHPC – Tier 3 Clusters

- **B**aden-**W**ürttemberg **H**igh **P**erformance **C**omputing
- **bwUniCluster** (KA): General Purpose, Undergraduates...
- **bwForClusters**: Research area-specific
  - **BinAC** (TÜ): Bioinformatics, Medical Informatics, Pharmacy, Geosciences and Astrophysics
  - **NEMO** (FR): Neuroscience, Elementary Particle Physics, Microsystems Engineering and Material Science
  - **JUSTUS** (UL): Computational Chemistry and Quantum Sciences
  - **MLS&WISO Production** (HD+MA):
    - Structural and Systems Biology
    - Medical Science
    - Soft Matter
    - Computational Humanities
      - Economics and Social Sciences



bwHPC-S5

- „Scientific Simulation and Storage Support Services“
- Focus on simulation and storage, but also support by research area
- HPC Competence Centers
  - **Here:** Molecular Life Sciences, Economics and Social Sciences

    mls-wiso@bwhpc.de

    hpc-support@mailman.uni-mannheim.de
  - **Wiki**:

    https://wiki.bwhpc.de

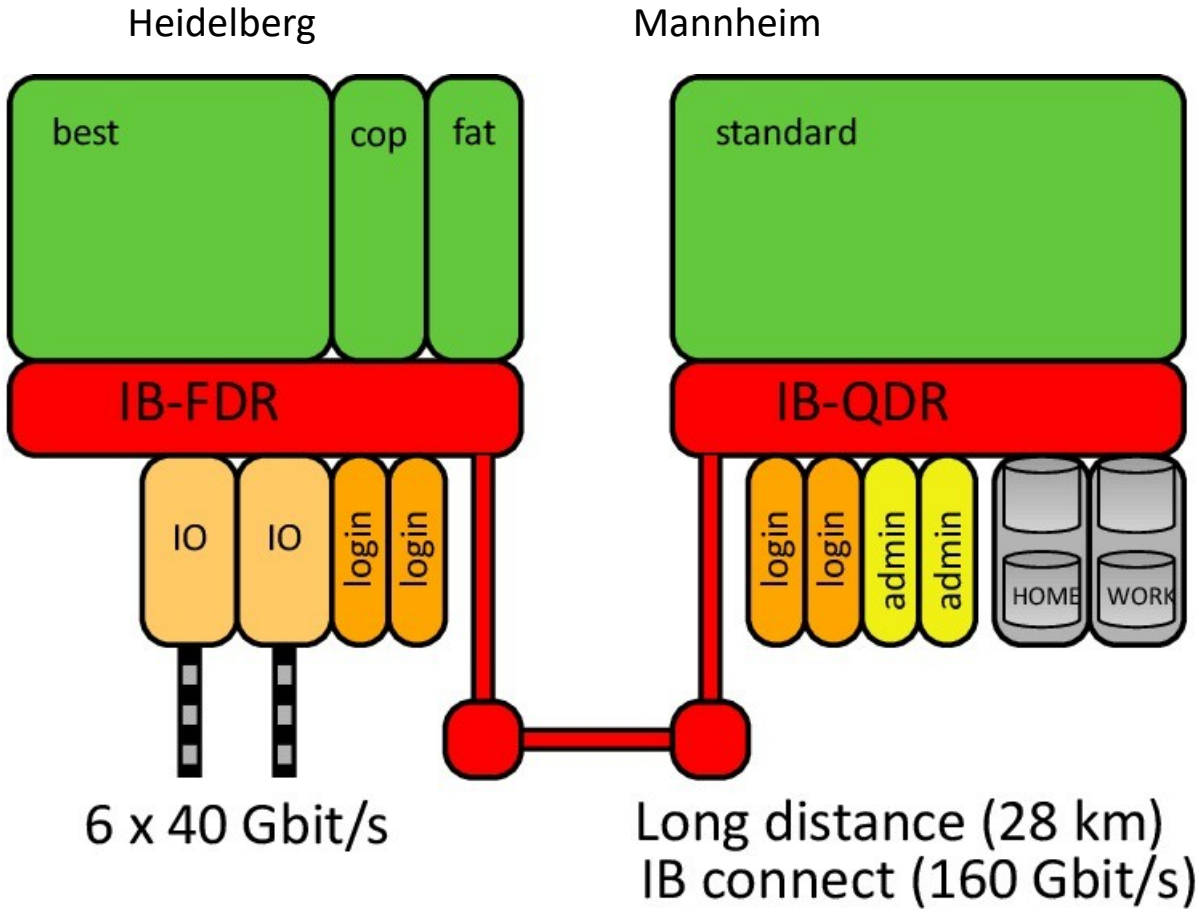
- Day-to-day user support
- Support for larger projects („Tiger Teams“)

MLS&WISO Production

- Hardware
  - **5 Login** nodes (not for computation!)
  - **476 Standard** nodes (16 Cores, 64 GB RAM, 120h walltime)
  - **184 Best** nodes (16 Cores, 128 GB RAM, 120h walltime)
  - **8 Fat** nodes (40 Cores, 1536 GB RAM, 48h walltime)
  - **18 GPU** nodes (Standard + 2 nvidia Tesla K80, 48h walltime)
  - Various different nodes in a Skylake architecture (also newer GPUs)

MLS&WISO Production

- Architecture:



Software

- CentOS 7.5 on all nodes
  - Yes, that‘s Linux
  - No, you don‘t get root :)

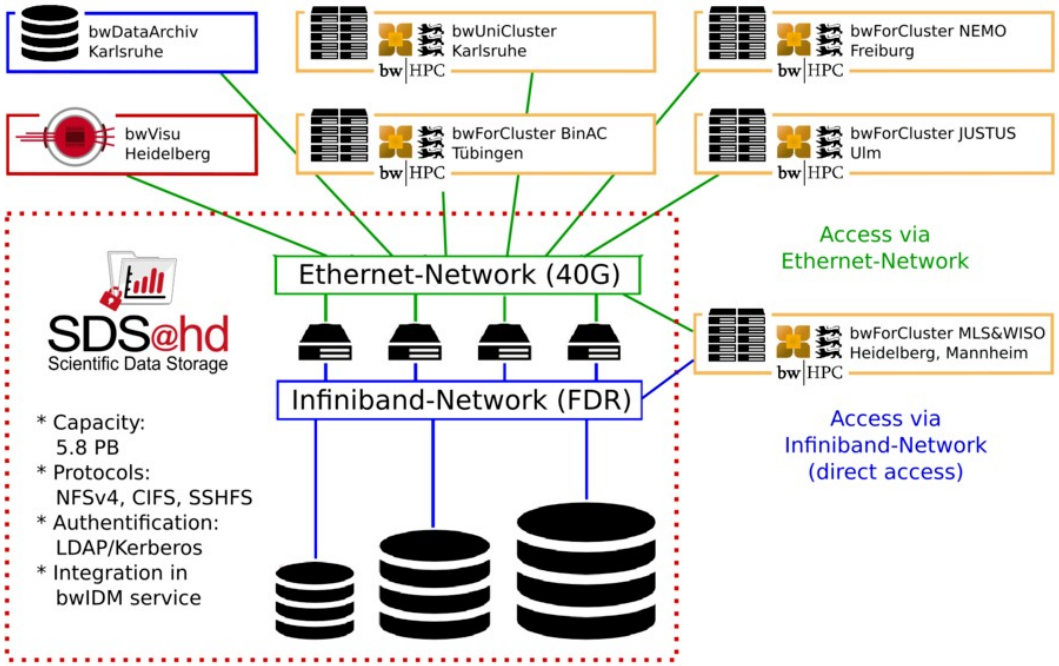
- Specific Software is loaded through modules
  - Intel Python
  - R, Matlab, Mathematica, Stata-MP...
  - Various Compilers
  - Various MPI-versions, Cuda...
  - And many more!
  - Users can import software that doesn‘t require root, e.g. installing locally or through Conda, Singularity...

Filesystem

- Filesystem
  - BeeGFS 7.1.2
- HOME
  - 100 GB total quota, permanent
  - For organization and minor preprocessing
- Workspaces
  - 10 TB total quota, time-limited
  - Any number of workspaces
  - For productive work
- TMP
  - 128 GB per node, temporary
- SDS@HD may be mounted

SDS@HD

- Scientific Data Storage @ Heidelberg
  - For frequently accessed (‚hot') Data
  - 5.8 PB total



https://wiki.bwhpc.de/e/Sds-hd

Access

- Application for a *Rechenvorhaben (RV, Computation Project)*
  - Joint project for multiple users, also for institutes
  - Registered at state-level by one project-leader

    https://zas.bwhpc.de/
- Application for an Entitlement
  - For a single user
  - Granted by their home university, but administered by bwIDM

    https://sp-grid-webregistration.uni-mannheim.de/
- Register to MLS&WISO production
  - Requires an approved *RV* and an entitlement
  - Also used to set the service password for the Cluster

    https://bwservices.uni-heidelberg.de/

First steps – Login and Administration

- Needs a terminal and an ssh-Client, either natively or emulated
- On Linux/Mac:
  - `ssh ma_xxx@bwfor.cluster.uni-mannheim.de` (Login MA)
  - `ssh ma_xxx@bwforcluster.bwservices.uni-heidelberg.de` (Login HD)
- On Windows:
  - MobaXTerm
  - Putty / WinSCP
  - Windows Subsystem for Linux

Workspaces

- Create a workspace called *myws* for 90 days

```
ws_allocate myws 90
```

- Extend the lifetime of your workspace

```
ws_extend myws
```

- Release your workspace

```
ws_release myws
```

- List all current workspaces

```
ws_list
```

The Module System

- Software is organized in modules
  - For one session / job only!
- All available modules:

```
module avail
```

- Loading a module:

```
module load math/R
module load math/R/3.6.1
```

- List of currently loaded modules:

```
module list
```

The Batch System (MOAB)

- All computations have to be sent to the Cluster-Batch System (MOAB/Torque)
- No calculating on the Login-nodes!
- Jobs will be started once the requested resources are free, and no higher-priority jobs are in the same queue
- Two kinds of jobs:
  - Batch Jobs: a Bash-Script is remotely executed
  - Interactive Jobs: a node with terminal access is given to you for manual use

Job-Scripts (MOAB)

- Job-Script (*example.sh*)

```
#!/bin/bash
#MSUB -l nodes=1:ppn=16:gpu
#MSUB -l walltime=5:00:00
#MSUB -l pmem=1gb
#MSUB -j oe

module load math/R/3.6.1
cd myworkspace

R CMD BATCH --no-save --no-restore keras.R
```

```
msub example.sh
```

- Interactive Jobs:
  - `msub -I -l nodes=1:ppn=16, walltime=1:00:00, pmem=1000mb`

man msub

- `-l` (resources)
  - nodes=1:ppn=16 – node number and specifications
  - walltime=1:00:00 – maximum „real“ time, either dd:hh:mm:ss or s
  - mem=16gb – total memory
  - pmem=1gb – memory per process (total: ppn x pmem)
- `-v` (variables)
  - OMP_NUM_THREADS=8
- `-N` (Name)
  - ThisWillGetMeTheNobelPrice
- `-j` (join output)
  - [oe | eo] – output and errors in one document

Monitoring

- `showq` – shows all of your queued or running jobs
  - `-c` – your completed jobs in the last 24h
- `showbf` – shows free resources
  - `-f <type>` – shows free resources of one specific type (e.g. best)
- `checkjob <jobID>` – shows details for a job
  - `-v` – verbose
  - `-v -v` – very verbose
  - `-v -v -v` – very, very verbose
- Direct monitoring through ssh onto an allocated node

So what if bwForCluster MLS&WISO Production just isn‘t enough?

- **If you aren‘t sure your code is optimized…**
  - ask us for support! We can help improving it.
- **If you are sure your code is optimized...**
  - try other hardware
  - Try cloud providers (federal or commercial) for VMs without walltime limits
  - apply for a different or higher tier cluster (bwUniCluster, ForHLR, ...)

„The Cloud“

- Commercial example: Amazon Web Services (AWS)
  - Commercial vendor: Costs money, data security might be tricky
  - The free package (t2-micro): 1 VCPU (10% load), 1GB RAM
  - The equivalent to a decent laptop: at 200-250€/month with high load

- But there is also… the bwCloud!

bwCloud

- Currently free of charge for MA-users
- One permanent (medium-sized) instance per user as default
- No walltime-limits, almost complete control over the instance
- Completely hosted at four of Baden-Württemberg‘s universities within their own research network, **BelWü**
- Sensible research data stays within the BelWü

BwCloud – Hard- and Software

Location	No. of Nodes	Storage (brutto)	Cores
Karlsruhe	27	504 TB	432
Ulm	36	1200 TB	576
Freiburg	27	504 TB	432
Mannheim	27	504 TB	432
Total	**117**	**2712 TB**	**1872**

(as of 11/2019)

Powered by

openstack. ceph

bwCloud - Access

- Access: Similar bwHPC
  - Entitlement from your home-university - bwCloud-Basic
  - Registration and instance-management:

    https://www.bw-cloud.org/

- The Basic-package:
  - 1 Instance with 1 vCore
  - 1 GB RAM
  - 50 GB total storage
  - 1 fixed IP-adress

HPC vs. Cloud

HPC	Cloud
(Scientific) Computing only	Any application (e.g. webhosting)
Batch-System and Queue	Bought resources are usually available right away
Way more powerful	Way less powerful
Less applicable for development	More applicable for development
Organized software-stack	Little to no pre-installed software

MLS&WISO Online Course

- Part 1... opened yesterday!

  https://training.bwhpc.de

- Main course on April 15th + 16th
  - Lessons during the day
  - Interactive session in the afternoon

Acknowledgement

- If you use the Cluster, acknowledge us!

The authors acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG.

- Notify us of any reports, conference papers, journal articles, theses, posters, talks which contain results obtained on any bwHPC resource by sending an email to publications@bwhpc.de stating.

Questions

- Thank you for your attention.
- **Questions?**