Community Forum Data-Intensive Computing

Support Teams for bwForCluster MLS&WISO, SDS@hd, and bwVisu
Common User Meeting

bwForTreff
- for users of bwHPC systems
- for students and scientists interested in HPC

SDS@hd user meeting
- for users of Scientific Data Storage SDS@hd
- for students and scientists interested in data storage

Community Forum Data-Intensive Computing (DIC)
- Get status and news of HPC and storage systems
- Learn how to use both systems for different workflows
- Networking with support team and other users
Community Forum DIC

Agenda

➤ Scientific Data Storage SDS@hd
➤ bwForCluster MLS&WISO
➤ News
  ➤ Job Efficiency
  ➤ Container Environment
➤ Announcements
  ➤ Archive Service SDA@hd
➤ Demos
  ➤ Singularity Containers
  ➤ Remote Visualization Service bwVisu
➤ Discussion
bwHPC and bwHPC-S5
Baden-Württemberg Framework Program 2018-2024

- Concept of the Universities of Baden-Württemberg for High Performance Computing (HPC), Data Intensive Computing (DIC), and Large Scale Scientific Data Management (LS2DM) ⇒ provides Compute Resources

- bwHPC-S5: Scientific Simulation and Storage Support Services ⇒ provides User Support

Information on Website
http://www.bwhpc.de
Best Practices Guide
http://www.bwhpc.de/wiki
bwHPC-S5 Project

**AP 1.1**
Förderative Wissenschaftsunterstützung

**AP 1.2**
Schulungen

**AP 1.3**
Öffentlichkeitsarbeit

**AP 2.1**
Förderative HPC-Infrastruktur

**AP 2.2**
Förderative Dateninfrastruktur

**AP 2.3**
Basisdienste

**AP 2.4**
Querschnittsthemen

**AP 3**
Innovation

- Nutzerbezogene Aktivitäten und Öffentlichkeitsarbeit
- Föderationsbetrieb und systembezogene Aktivitäten
- Innovations- und Evaluationsaktivitäten
bwHPC-S5 Support

- Information seminars, hands-on training, HPC and storage workshops
  http://www.bwhpc.de/courses_a_tutorials.html
- Documentation & best practices repository http://www.bwhpc.de/wiki
- Providing/maintaining/developing:
  - simplified access to all bwHPC resources
  - cluster independent & unified user environment
  - software portfolio
  - tools for data transfer and management
  - interface to archive and repositories services
- Migration support:
  - code adaptation, e.g. MPI or OpenMP parallelization
  - code porting (from desktop or old HPC clusters)
  - migration to higher HPC levels
SDS@hd – Scientific Data Storage

https://sds-hd.urz.uni-heidelberg.de
Scientific Data Storage SDS@hd

Service Features and Storage Hardware

- High performance data access for large scientific “hot data”
- HW parts: LSDF2 & LSDF2b
  - High availability, no single point of failure
  - Spatially divided hardware in different fire areas
- IBM Spectrum Scale Advanced (GPFS)
- Current usable capacity: 5.8 PB netto
- Data security, encryption
- Protocols: SMB, NFSv4, SSHFS, SFTP
- Flexible group and access management
  - Collaborations among universities are possible
- 'Landesdienst’ for scientists of BW universities
  - in HD: valid Uni-ID necessary (if needed: ’unentgeltliche Tätigkeit’)
- Connection to other services: bwHPC, bwVisu, heiCLOUD
Extension of the LSDF 2 hardware:

- DFG proposal for LSDF 2 Phase 2 (End 2017) was granted!

- Extension of storage capacity to 20 PB!

- The extension of the LSDF2 storage systems with the granted funding is currently prepared and will lead to a public procurement process in 2019

- Future additional proposal, if more storage capacity is needed
  ⇒ Reporting
Scientific Data Storage SDS@hd

Reporting

The use of SDS@hd must be acknowledged in publications!
▶ Highly important for DFG reviewing and future fundings of LSDF2
▶ Attest of scientific necessity and usefulness with publications

List of scientific publications:
▶ Send info about new publications to sds-hd-reporting@urz.uni-heidelberg.de
▶ Submit list at application for renewal of Speichervorhaben (SV)

Text for acknowledgement in publications:

The authors gratefully acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) and the German Research Foundation (DFG) through grant INST 35/1314-1 FUGG.

See: https://sds-hd.urz.uni-heidelberg.de/reporting
Scientific Data Storage SDS@hd

SDS@hd Registration

Entitlement
▶ Authorization to use the service
▶ Delivered by home organization

SDS@hd-Management
▶ Register Speichervorhaben (SV) or participate on existing SV
▶ Signatures needed on contracts
▶ https://sds-hd.urz.uni-heidelberg.de/management

Registration Server
▶ Register and manage service account, e.g. set personal service password
▶ https://bwservices.uni-heidelberg.de
Scientific Data Storage SDS@hd

Costs

- Investment was funded by MWK (50%) and DFG (50%) (Art. 91b GG)
  - only operation costs have to be paid

- Price:
  - 0.25 ct/GB/month (0.0025 EUR/GB/month, = 30 EUR/TB/year)
  - Heidelberg User:
    - 0.1 ct/GB/month (12 EUR/TB/year, subsidized by central funds)

- Measurement: average of used quota
- Backup is already included
Billing

- Costs have to be charged
- Measurement: average of used quota per month
- Beginning in 2019
- Some "Kostenstellen" on existing contracts are incorrect! Will be contacted for correction in the next weeks.

Medical Faculties:

- Service registration only for scientists of universities
- Not for usage as member of university hospital
- Announcement as scientific member of the medical faculty
- If you have a Uni-ID, everything is OK!
| Operating system      | Linux                                | Windows (since Vista)  
 OS X Mavericks (10.9) | all |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Installation effort</td>
<td>one-time installation effort (Kerberos, ID mapping)</td>
<td>minimal effort: Mapping of a network drive</td>
<td>installation of SFTP client, e.g. sshfs, WinSCP, FileZilla</td>
</tr>
<tr>
<td>Possible bandwidth</td>
<td>up to 40Gbit/sec</td>
<td>up to 40Gbit/sec</td>
<td>up to 40Gbit/sec</td>
</tr>
<tr>
<td>Fault tolerance</td>
<td>highly available</td>
<td>highly available</td>
<td>–</td>
</tr>
<tr>
<td>Encryption (transfer)</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
</tr>
</tbody>
</table>
| Firewall ports       | 2049 (nfs)                           | 139 (netbios), 135 (rpc)  
 445 (smb)             | 22 (ssh) |
| Recommendation       | For workstations, servers, microscopes ... | For workstations, servers, microscopes, ... | For Linux PCs, machines in restrictive or external networks |
bwForCluster MLS&WISO
bwForCluster MLS&WISO

System Layout

Development (IWR)  Production (URZ/RUM)

@Heidelberg  @Heidelberg  @Mannheim

standard  best  cop  fat  standard

IB-QDR  IB-FDR  IB-QDR

login  login  login  login  login

admin  admin  admin  admin

4xFDR  40 GE

Long distance (28 km)

IB connect (160 Gbit/s)

to LSDF2
bwForCluster MLS&WISO

Production

Access Prerequisites

▶ bwForCluster Entitlement for LocalUserID from home university
▶ Member of RV for bwForCluster MLS&WISO (Production) at ZAS

Registration

▶ Register at https://bwservices.uni-heidelberg.de
▶ Choose bwForCluster MLS&WISO - Production
▶ Set service password

Login

▶ Login server:
  bwfor.cluster.uni-mannheim.de
  bwforcluster.bwservices.uni-heidelberg.de
▶ User ID: SitePrefix_LocalUserID (= hd.UniID for Heidelberg users)
▶ Password: Service password
### File Systems

<table>
<thead>
<tr>
<th></th>
<th>$$TMPDIR$</th>
<th>$$HOME$</th>
<th>Workspaces</th>
<th>SDS@hd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>node local</td>
<td>global</td>
<td>global</td>
<td>global¹</td>
</tr>
<tr>
<td>Lifetime</td>
<td>batch job walltime</td>
<td>permanent</td>
<td>workspace lifetime</td>
<td>permanent</td>
</tr>
<tr>
<td>Disk Space</td>
<td>128 GB</td>
<td>36 TB</td>
<td>384 TB</td>
<td>5.7 PB</td>
</tr>
<tr>
<td>Quotas</td>
<td>no</td>
<td>100 GB</td>
<td>no</td>
<td>yes²</td>
</tr>
<tr>
<td>Backup</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes³</td>
</tr>
<tr>
<td>File System</td>
<td>xfs</td>
<td>BeeGFS</td>
<td>BeeGFS</td>
<td>GPFS</td>
</tr>
</tbody>
</table>

¹ on all compute nodes except standard nodes  
² disk quota set by SDS@hd for SV  
³ done by SDS@hd
## SDS@hd on bwHPC clusters

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Location</th>
<th>Connection Mode</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>bwUniCluster</td>
<td>Karlsruhe</td>
<td>via data mover rdata</td>
<td>in production</td>
</tr>
<tr>
<td>bwForCluster</td>
<td>Heidelberg/ Mannheim</td>
<td>via data mover &amp; direct access on compute nodes</td>
<td>in production</td>
</tr>
<tr>
<td>bwForCluster</td>
<td>Tübingen</td>
<td>via data mover</td>
<td>in production</td>
</tr>
<tr>
<td>bwForCluster</td>
<td>Ulm</td>
<td>not needed</td>
<td>in evaluation</td>
</tr>
<tr>
<td>bwForCluster</td>
<td>Freiburg</td>
<td>not needed</td>
<td>in evaluation</td>
</tr>
<tr>
<td>ForHLR</td>
<td>Karlsruhe</td>
<td>via data mover rdata</td>
<td>in production</td>
</tr>
<tr>
<td>HazelHen</td>
<td>Stuttgart</td>
<td>in evaluation</td>
<td>in preparation</td>
</tr>
</tbody>
</table>
Workspaces

What are workspaces?
Workspaces are allocated folders with a lifetime.

Workspace tools

- `ws_allocate foo 10` Allocate a workspace named foo for 10 days.
- `ws_list -a` List all your workspaces.
- `ws_find foo` Get absolute path of workspace foo.
- `ws_extend foo 5` Extend lifetime of workspace foo by 5 days from now.
- `ws_release foo` Manually erase your workspace foo.

Limits for bwForCluster MLS&WISO - Production

- Max. time for one allocation: 90 days
- Number of extensions: 10 times
- Max. lifetime of a workspace: 990 days
bwHPC Training

HPC Courses

- Duration: 1/2 day up to 1 week
- Courses from all bwHPC partners in Baden-Württemberg
  http://www.bwhpc.de/courses_a_tutorials.html
- Planned courses in 2019:
  - Hands-On Workshop bwHPC (german) 15. May 2019 in Mannheim
  - OpenMP and/or MPI Workshop in 2019
  - Nvidia OpenACC Workshop in 2019
News
### Job Efficiency

#### Job Feedback Script

- generated after Job completion
- appended to job output file
- includes basic job statistics
- additional Hints and advises
- available in 11.2018

---

<table>
<thead>
<tr>
<th>Job ID</th>
<th>589742</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job name</td>
<td>STAR-testjob</td>
</tr>
<tr>
<td>Job state</td>
<td>Completed</td>
</tr>
<tr>
<td>TimeQueued</td>
<td>00:00:11 [elig. 00:00:00]</td>
</tr>
<tr>
<td>WallTime</td>
<td>00:01:07 of 1:00:00</td>
</tr>
<tr>
<td>Queue</td>
<td>single</td>
</tr>
<tr>
<td>Nodes:ppn:feature</td>
<td>1:16:standard</td>
</tr>
<tr>
<td>Tasks</td>
<td>16</td>
</tr>
<tr>
<td>Core hours</td>
<td>0.30</td>
</tr>
<tr>
<td>CPU load job avg, (0.00-1.00)*tasks</td>
<td>0</td>
</tr>
<tr>
<td>CPU efficiency, 0-100%</td>
<td>0.051885</td>
</tr>
<tr>
<td>Memory task, MB</td>
<td>0 [req. 1024]</td>
</tr>
<tr>
<td>Memory job, MB</td>
<td>0 [req. 16384]</td>
</tr>
<tr>
<td>Scratch task-job, GB</td>
<td>0 [job: 0]</td>
</tr>
<tr>
<td>Host list</td>
<td>m19s0304:16</td>
</tr>
</tbody>
</table>

---

# JOB ANALYSIS: notes, problems and solutions

* NOTE: some of the parameters are equal "0", see "Memory" and "CPU load".

Advice: because of the small job's running time please ignore it.
Container Environment
Singularity on bwForCluster MLS&WISO

▶ reproducible runtime environment on HPC System
▶ comparable to Docker, Docker Images can be re-used
▶ Modules, Workspaces & SDS@hd available
▶ available via Module System in 12.2018
Announcements
Scientific Data Archive SDA@hd

- Heidelberg Service for Archiving on Tape (in development)
- all type of research data
- typically: 10 years retention obligation of DFG
- base set of metadata
- SDS@hd connection, including migration paths
- Costs:
  - 0,05 ct/GB/month (= 6 EUR/TB/year)
- Contact: sda-hd-support@ urz.uni-heidelberg.de
Singularity
Singularity

- Open source container runtime focused on HPC
- Focus on mobility of applications – encapsulate applications with all dependencies
- Easy to use
- Containers are started (and behave) like regular applications
- Integrated support for MPI and GPUs
Singularity & Docker

▶ Easy conversion Docker → Singularity
  ▶ singularity pull tensorflow.sif
docker://tensorflow/tensorflow:latest
  ▶ Docker layers are squashed

▶ User rights
  ▶ Docker daemon needs to run as root → who
    is allowed to start containers?
  ▶ In Singularity containers are started with
    user rights
Definition File

- Similar to Dockerfiles
  - Similar to Dockerfiles
  - No multiple “RUN”-instructions

```bash
Bootstrap: library
From: ubuntu:18.04

%setup
  touch /file1
  touch ${SINGULARITY_ROOTFS}/file2

%files
  /file1
  /file1 /opt

%environment
  export LISTEN_PORT=12345
  export LC_ALL=C

%post
  apt-get update && apt-get install -y netcat
  NOW= date
  echo "export NOW="${NOW}"" >> $SINGULARITY_ENVIRONMENT

%runscript
  echo "container was created $NOW"
  echo "Arguments received: "$"
  exec echo "$"

%startscript
  nc -lp $LISTEN_PORT

%test
```
Usefull Commands

▶ `singularity build paraview-v5.6.0.sif paraview.def`
  ▶ Uses definiton files to build an immutable image
▶ `singularity run paraview-v5.6.0.sif`
  ▶ Runs through the entrypoints and commands specified in the definition file
▶ `singularity shell -B /mnt/... paraview-v5.6.0.sif`
  ▶ Open a shell inside the container
Funded by

State project for remote visualization of large scientific data
Main Focus: user-friendliness

Our Project Partners

bwVisu

Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

KIT

Karlsruhe Institute of Technology

HLRS
Powerful hardware

- Can serve up to 288 concurrent users with GPUs
- Both NVIDIA and AMD GPUs available – users can choose optimal hardware for their use case
- Powerful CPUs
- High-bandwidth connection to SDS@hd and bwForCluster
Application Images

- Containerized Apps
  - Docker and Singularity
  - Reproducibility of scientific works
  - High control of software environment
  - Simple deployment
Easy to use software interface

- Webfronted for managing jobs
- Directly connect to visualization jobs with browser
- Image registry allows users to easily develop and deploy new applications to bwVisu
- Containerized application images of most common visualization applications readily available (e.g. Paraview, VMD, VisIt, ...)

ParaView  
blender  
VMD

Visual Molecular Dynamics
Other key features

- Mount file storage services like SDS@hd
- Use your custom visualization applications on bwVisu
- Restrict visibility and access of your applications
- Share visualization sessions with other users simultaneously
bwVisu Web

- Intuitive GUI
- Uses remote desktop tool Xpra
  - Only requirement: HTML5 enabled browser
  - Platform independent
  - Local client for low latency
Discussion

▶ Questions?
▶ Your use cases?
▶ Change requests?
▶ Feature requests?
Thank you!

Support Units

▶ HPC:
  hpc-support@urz.uni-heidelberg.de

▶ SDS@hd:
  sds-hd-support@urz.uni-heidelberg.de