



bwHPC – C5

# Introductory Course: HPC in Baden-Württemberg

T. König, R. Walter, R. Barthel, S. Raffener, A. Fuchs



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



**KIT**  
Karlsruher Institut für Technologie



ulm university universität  
**uulm**

Funding:



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

[www.bwhpc-c5.de](http://www.bwhpc-c5.de)

# Agenda

Time	Title
09:30	<b>bwHPC: Concept and User Support</b>
09:45	Architecture: bwUniCluster + ForHLR
10:00	Cluster: Access, Modulefiles, Filesystem
10:30	Cluster: Batch System
11:00	Break
11:10	Tutorials: Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting
Max. 13:10	End

# High Performance Computing in Baden-Württemberg: Concept and User Support

Tobias König



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



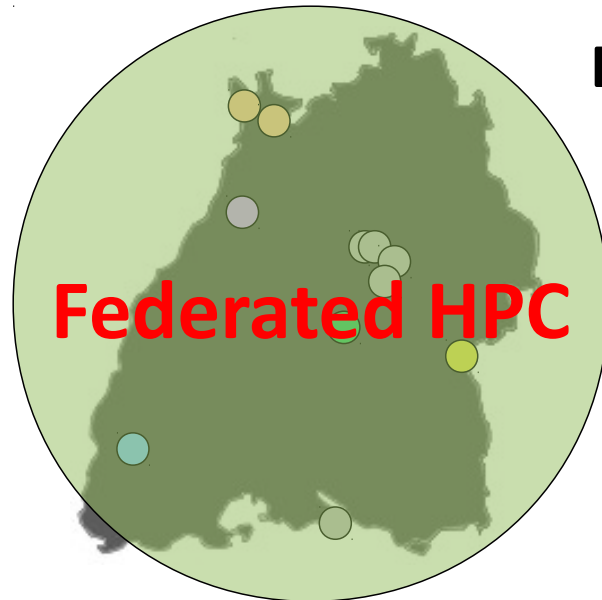
**KIT**  
Karlsruher Institut für Technologie



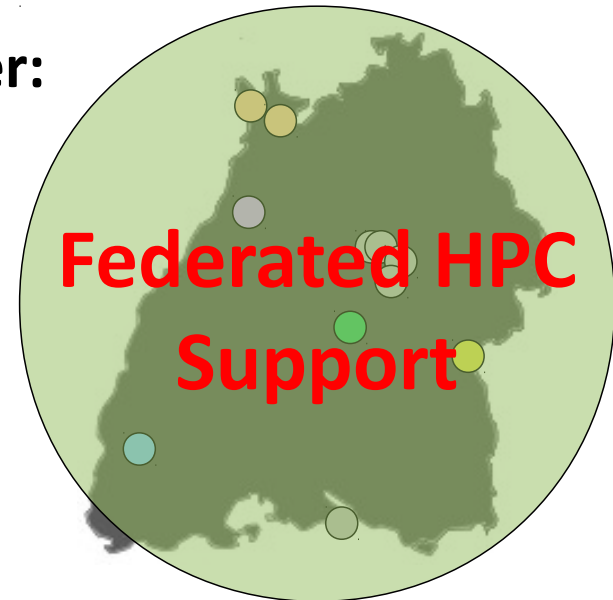
ulm university universität  
**uulm**

## To be answered:

- This presentation:
  - What is bwHPC?
  - What is bwHPC-C5?
  - Where do I get support / What kind of support?



## Brief answer:



- Following presentations
  - Where can I run my calculations/simulations and how?



# More detailed answer



## Website:

- General info: [www.bwhpc-c5.de](http://www.bwhpc-c5.de)

in english and german

- Best-practices-guide:

[www.bwhpc-c5.de/wiki](http://www.bwhpc-c5.de/wiki)

in english

The image shows two screenshots of the bwHPC-C5 website. The left screenshot displays the main navigation menu with options like 'Aktuelles', 'Projektpartner', 'Projektaufgaben', 'Mitarbeiterverzeichnis', 'Pressemitteilungen', 'Wiki / Dokumentation', 'bwHPC-Konzept', 'Leistungsleistungen', 'Kontakt', and 'Zurück'. The right screenshot shows the 'MediaWiki' interface with a search bar, navigation tabs (page, discussion, view source, history), and a 'Main Page' section. The 'Main Page' includes a 'Knowledge Base Wiki of Baden-Württemberg's IT services' and a 'Welcome' message. It also features sections for 'HPC Services' and 'Data Storage Services' with links to user guides and best practices.

## Publications:

- SCC News 02/2013, page 6:

<http://www.scc.kit.edu/publikationen/scc-news.php>

- Planned: SCC NEWS 2014



100 Gbit/s zwischen KIT und  
Universität Heidelberg  
100 Gigabit/s between KIT and Heidelberg

Landesprojekt „bwHPC-C5“ gestartet  
State project “bwHPC-C5” started

# 1. bwHPC concept

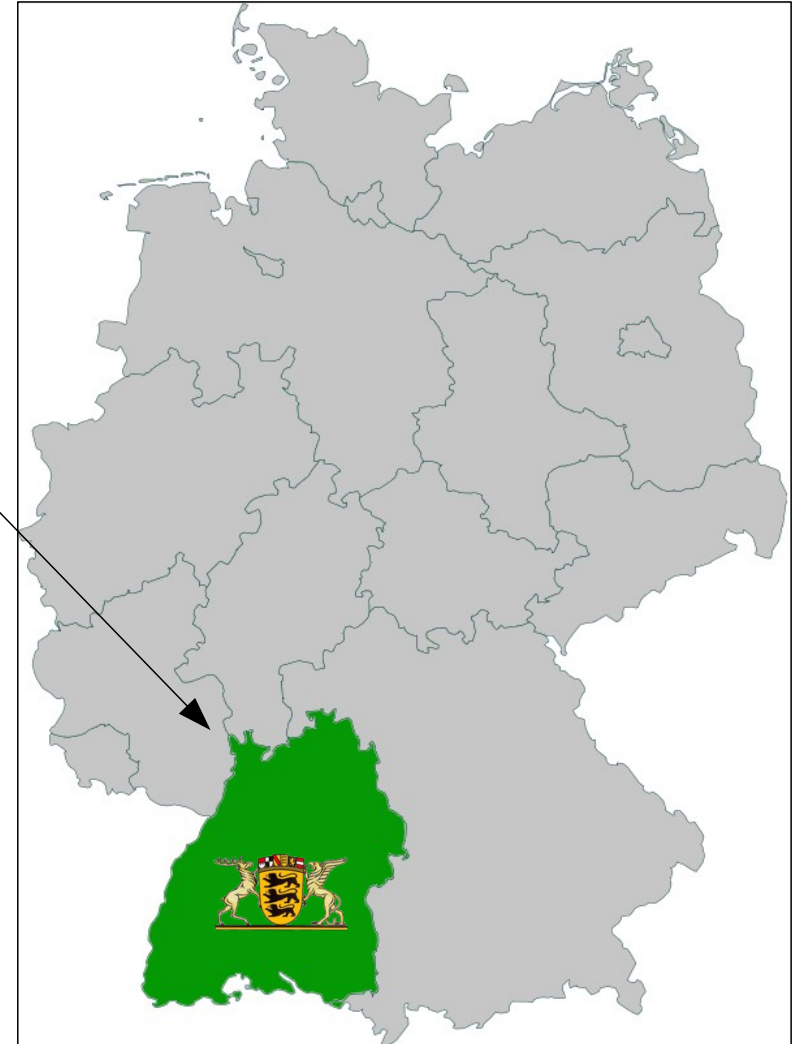
# What is bwHPC/bwHPC-C5?



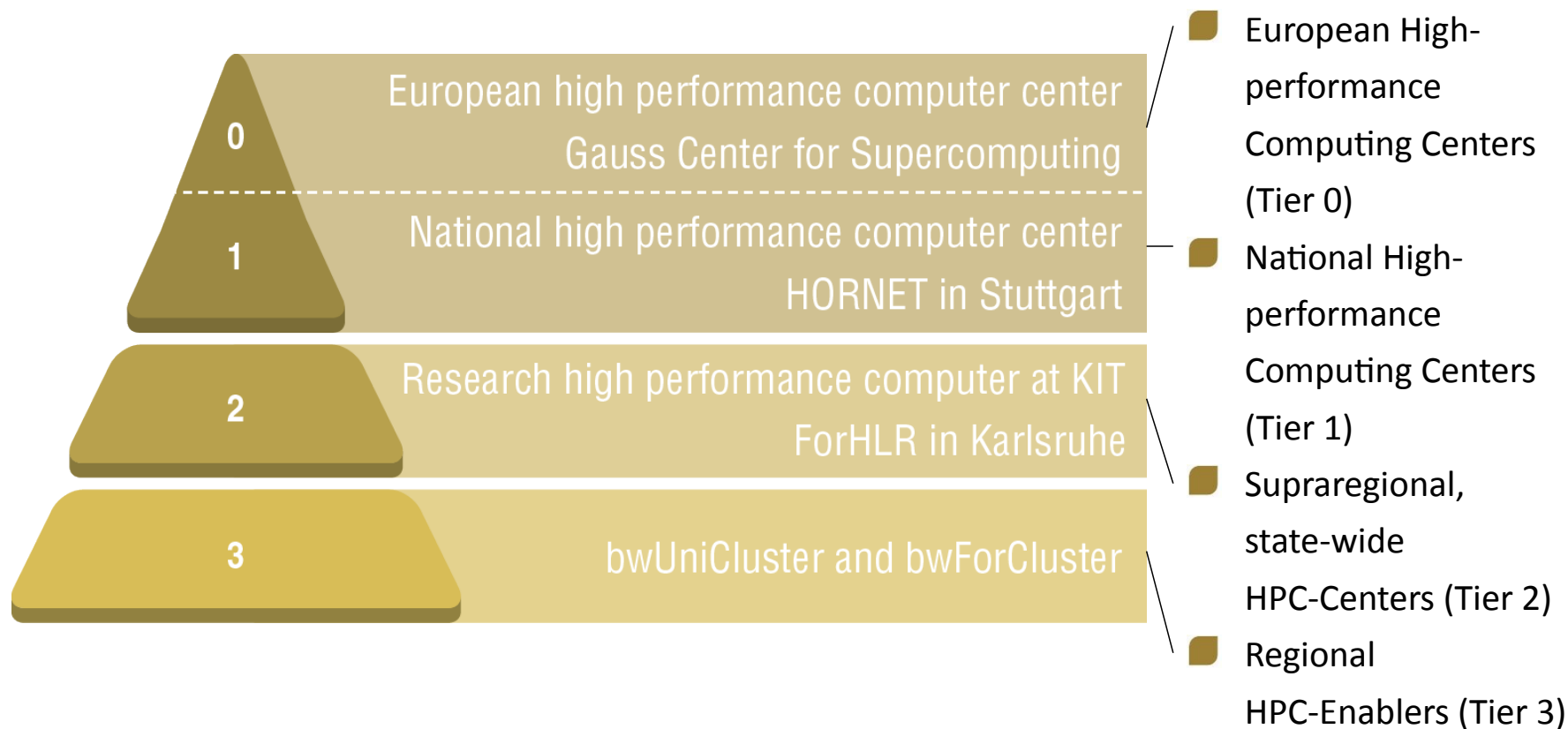
- bw = Baden-Württemberg
- bwHPC = Strategy for high performance computing in BW from 2013 to 2018, in particular for *Tier 3*
- bwHPC-C5 = Federated user and IT support activities for bwHPC

||

Cooperated activities

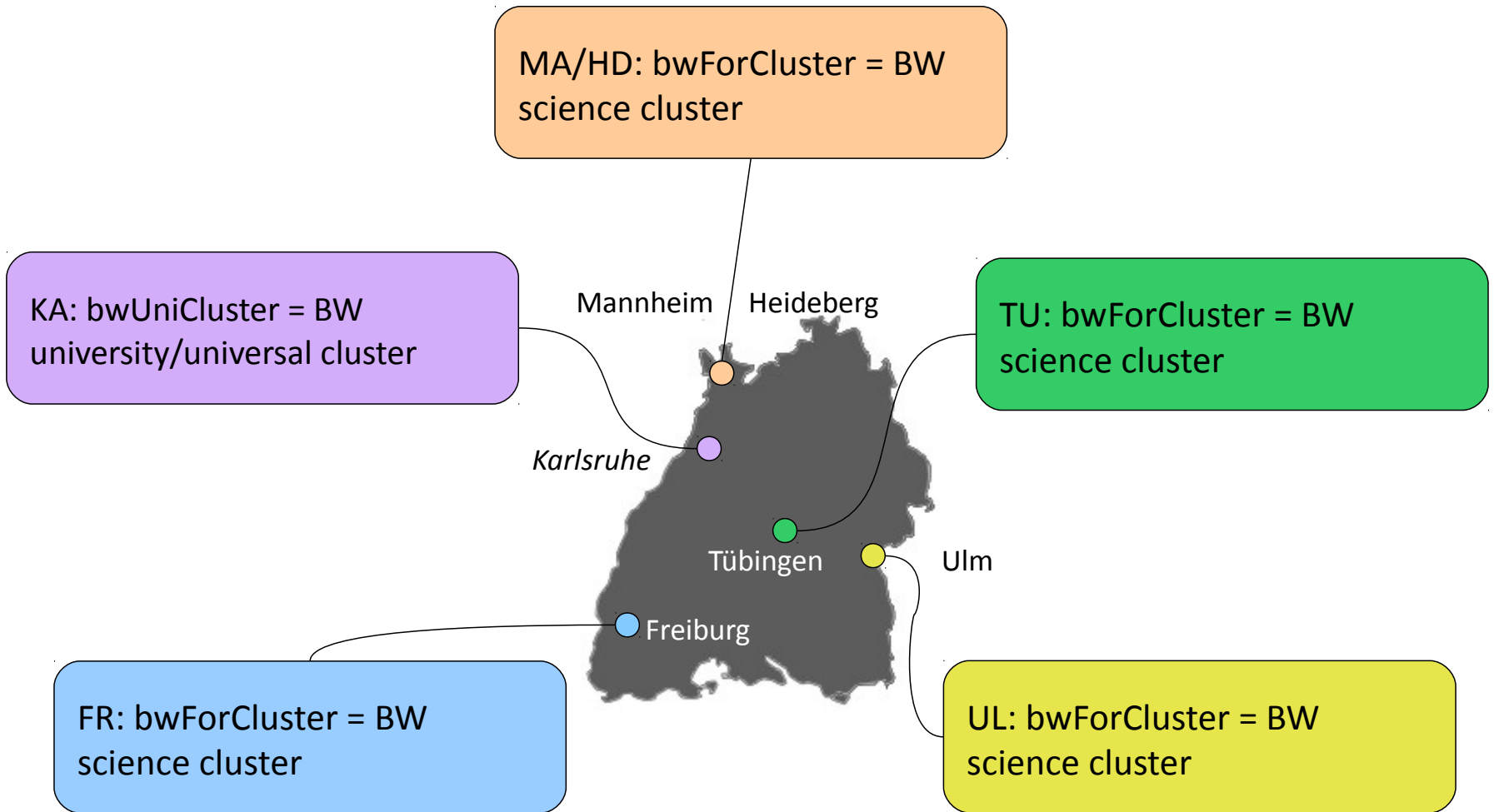


# HPC Tier classification / bwHPC

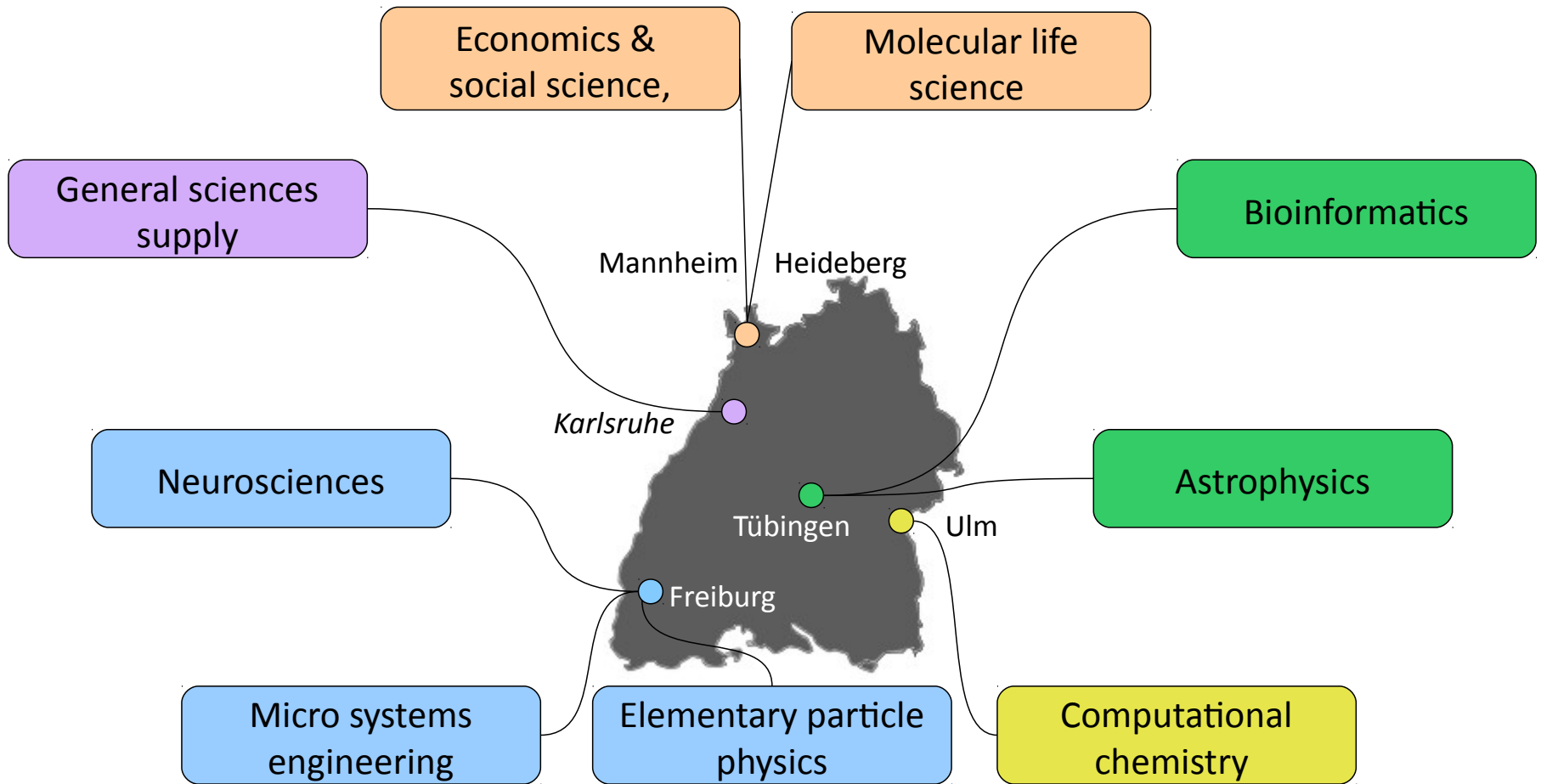


→ **bwHPC: Differentiation of Tier 3**

# bwHPC: Tier-3 (2013-2018)

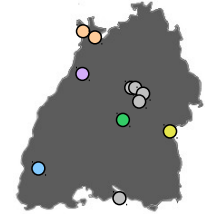


# bwHPC: currently covered research areas



→ **Independent of location, users of research area X use science cluster X**

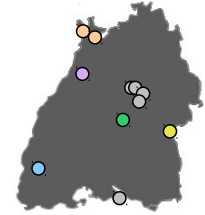
# Federated HPC @ tier 3 (1)



## ■ bwUniCluster:

- Co-financed by Baden-Württemberg's ministry of science, research and arts and the shareholders:
  - Freiburg, Tübingen, KIT, Heidelberg, Ulm, Hohenheim, Konstanz, Mannheim, Stuttgart
- Usage:
  - Free of charge
  - General purpose, teaching & education
  - Technical computing (sequential & weak parallel) & parallel computing
- Access / limitations:
  - For all members of shareholder's university
  - Quota and computation share based on own university's share

# Federated HPC @ tier 3 (2)



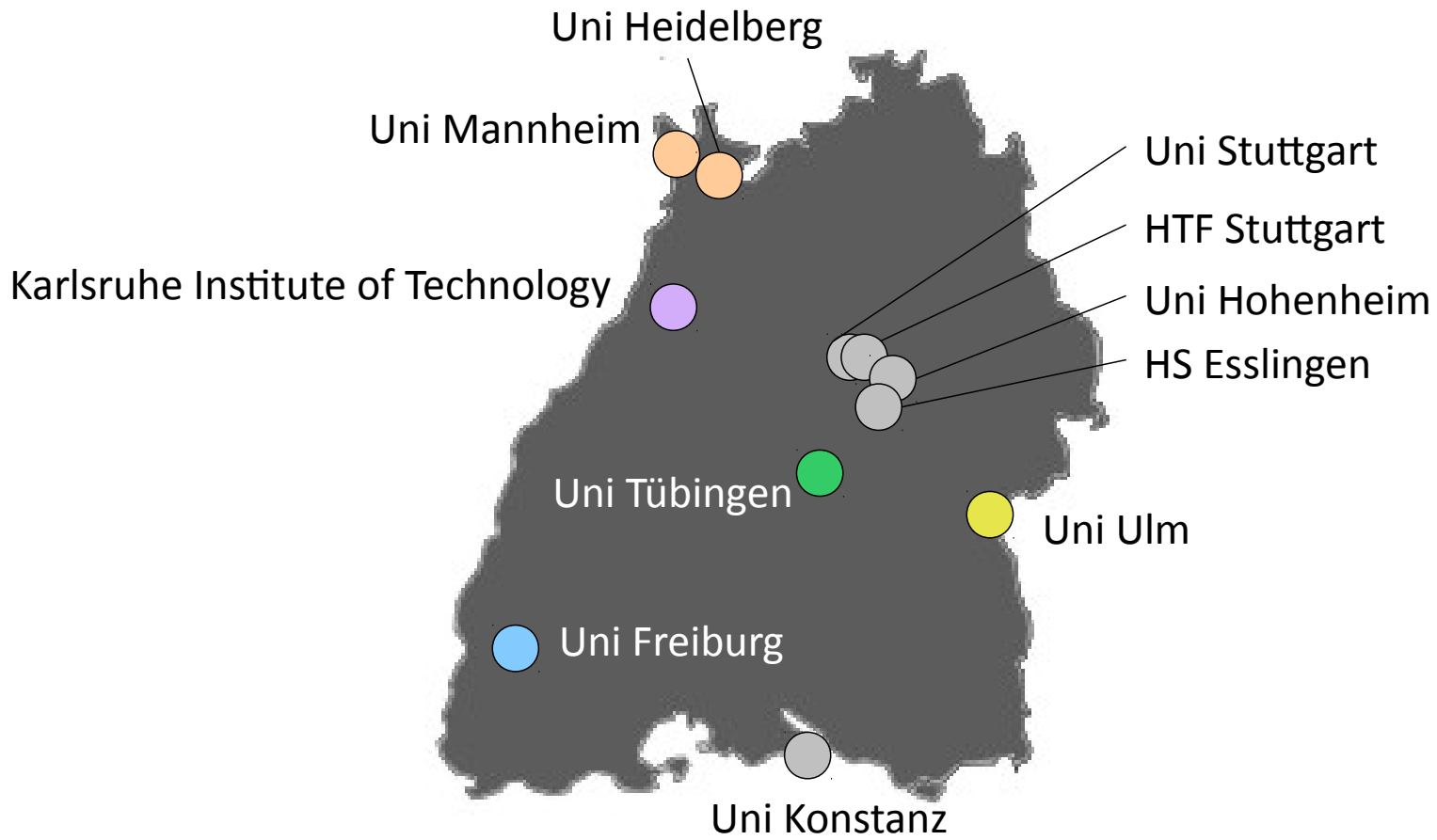
## ■ bwForCluster:

- For science communities according the DFG proposal
- Financed by German Research Society (DFG)
  
- Access:
  - All university members in Baden-Württemberg
  
- Usage, limitations:
  - Free of charge
  - Access only to science cluster matching own science community
  - Access requires approval of compute proposal



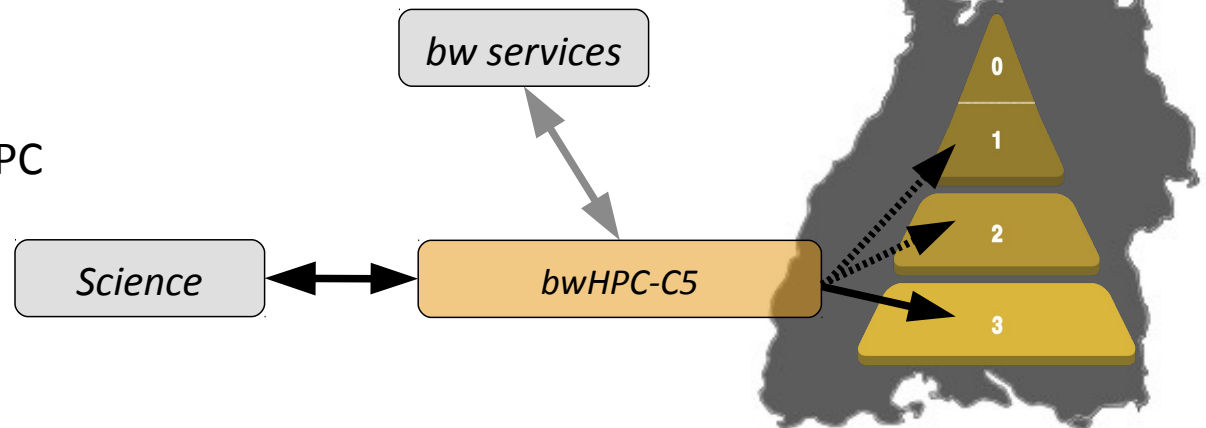
## 2. bwHPC-C5 / Federated user support

# Location of bwHPC-C5 project partners



# What is bwHPC-C5?

- C5 = Coordinated compute cluster competence centers
- CC = Competence Center
- Federated user support and IT services activities for bwHPC (2013 - 2015)
  
- For Users:
  - Support requests are **not** sent to cluster operators but to competence centers via ticket system <https://bw-support.scc.kit.edu/>
    - CC are not local, but distributed over whole BW
    - CC are community specific
  
- For BW:
  - Bridging Science & HPC
  - Bridging HPC tiers
  - Embedding services



# bwHPC-C5: federated science support

## ■ HPC competence centers

- Establish of a state wide pool of experts (community specific, parallel software development)
- Coordination of tiger teams activities (high level support teams)
- Build up best practices repository
- Identify of user key topics

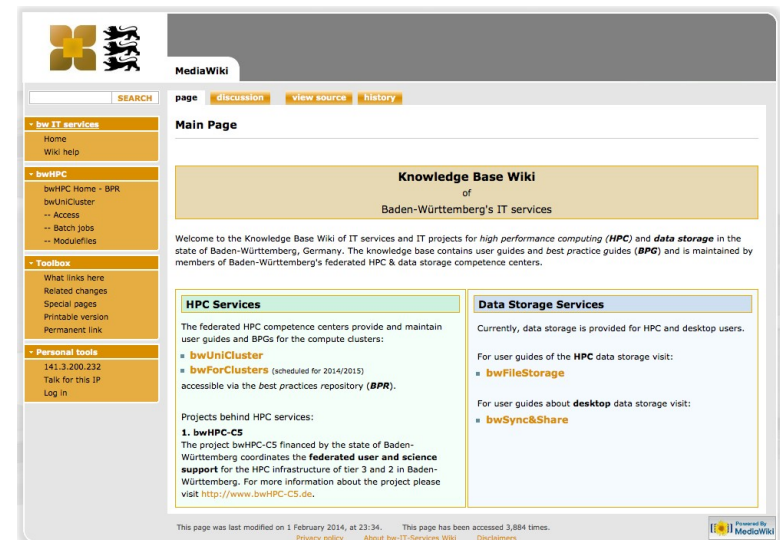
Org. Team@Mannheim/Heidelberg:  
Economics & social science,  
Molecular life science

Org. Team@Ulm:  
Computational chemistry

Org. Team@Karlsruhe:  
General sciences supply

Org. Team@Freiburg:  
Neuroscience,  
Micro systems engineering,  
Elementary particle physics

Org. Team@Tübingen:  
Bioinformatics,  
Astrophysics



## ■ Coordination of teaching activities

- Evaluate courses
- Generate offline and online material (eLearning, MOOC)

# bwHPC-C5: What kind of support?

- Information seminars, hands-on, HPC specific workshops
- Documentation + best practices repository [www.bwhpc-c5.de/wiki](http://www.bwhpc-c5.de/wiki)
- Providing/maintaining/developing:
  - simplified access to all bwHPC resources
  - software portfolio
  - cluster independent & unified user environment
  - tools for data management
  - trouble ticket system
  - „user“ information system
- Migration support:
  - code adaptation, e.g. MPI or OpenMP parallelisation
  - code porting (from desktop / old HPC cluster)
  - to tier 2 and 1

# bwHPC-C5: Service integration & development

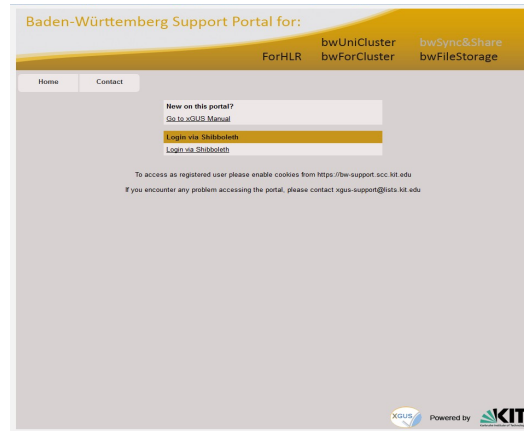
- bwHPC access + security  
e.g. bwIDM integr.



**bwUniCluster**

- 📄 Service Description
- 🔑 Registry Info
- 🔒 Set Service Password

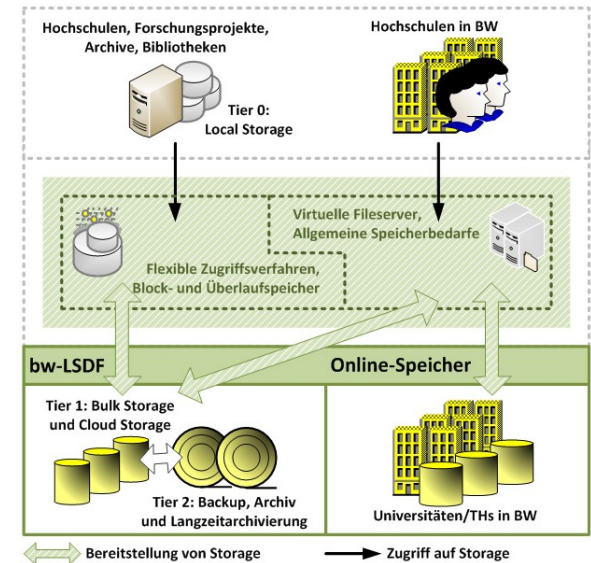
- Support Portal (trouble ticket system)



## Cluster information system (CIS)

Bibliotheken (B)		
boost	1.55.0	common
magicklib	1.3.1	common
gccfort	3.8.3-gnu-4.8	common
gccfort	3.8.3-musl-5.1.1	common
protdiff	1.4.1	common
Biologie		
blonark	0.10.1	common
biofile	1.0.1	common
biofile2	2.1.0	common
curfraks	2.2.0	common
glite	1.0.0	common
scarcads	0.1.19	common
tophat	2.0.11	common
Intermatic	0.32	common
C.A.F. (aw)		
adna	9.0.0aw	kit
amys	15.0.3aw	common
openblas	1.0-aw	common
openblas	0.7.1	common
openblas	2.2.2	common
openblas	2.1.0	common
openblas	1.0-aw	common

- Large Data integration: bwFileStorage



# bwFileStorage = central storage for HPC users

- Central online storage for HPC users
  - **bwFileStorage** (general transfer protocols: FTP, SFTP, SCP, HTTPS)
    - For HPC users: rdata tool
    - User documentation (registration, user guide, ...)  
<http://www.bwhpc-c5.de/wiki/index.php/bwFileStorage>
    - HW Installation at KIT
    - Federated user authentication → bwIDM “mechanisms” (Shibboleth)

# LNA-BW: Scientific Steering Committee

- LNA-BW = Landesnutzerausschuss-BW: Scientific steering of bwHPC and bwDATA
- Website: <http://www.bwhpc-c5.de/LNA-BW.php>
- Committee:
  - One representative for each Baden-Wuerttemberg's university
    - Representative of KIT: Prof. Dr. Marc Weber, Institut für Prozessdatenverarbeitung und Elektronik (IPE)
- Tasks:
  - Define and approve bwHPC access policies
  - Assessment of bwHPC workload
  - Regulation of bwHPC cluster expansion
  - Assignment of science communities to science clusters
  - Representation of HPC users interests concerning:
    - resource demands, HPC technologies and software licences, adjustment of resource quota



# Summary

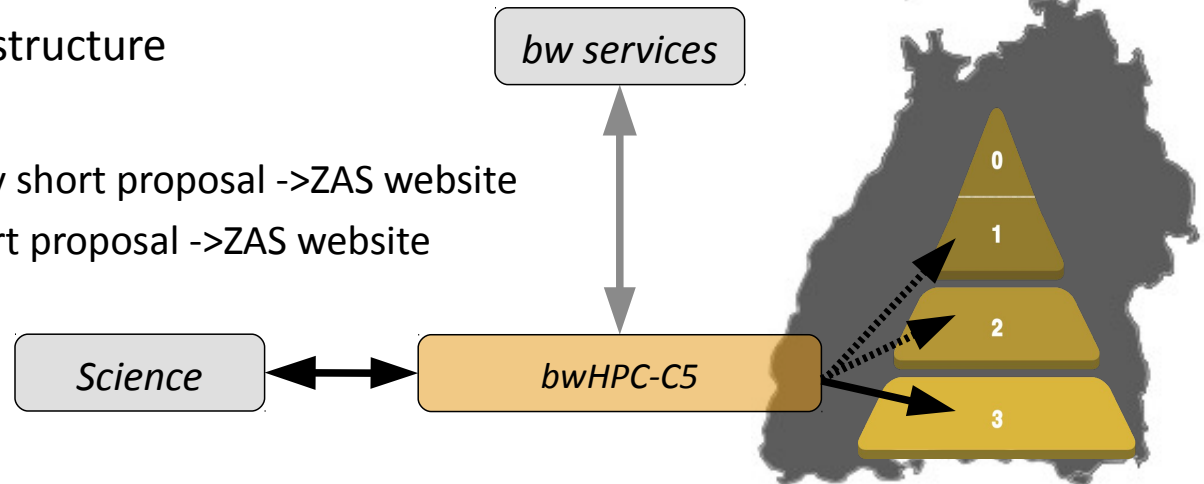
# bwHPC / bwHPC-C5: Summary

## ■ bwHPC:

- tier 3 = HPC enabler, science specific infrastructure

## ■ To get access:

- bwUniCluster: very short proposal ->ZAS website
- bwForCluster: short proposal ->ZAS website



## ■ bwHPC-C5

- your support team for HPC : **HPC competence centers**, which coordinate:


- federated HPC support (help: code migration, parallelisation etc.)
- contact via Support Portal <https://bw-support.scc.kit.edu/>

## ■ your developer team

- Implementation & development of bw services
- Evaluation & Integration of new HPC technology

## ■ LNA-BW: representatives of HPC user interests

# Acknowledgment

-  Ministry of Science, Research and Arts Baden-Württemberg

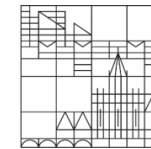
- All contributors and co-authors of bwHPC-C5:

- Uni Ulm
- Uni Konstanz
- Uni Freiburg
- HFT Stuttgart
- Uni Stuttgart
- HS Esslingen
- Uni Hohenheim
- Uni Tübingen
- Uni Mannheim
- Uni Heidelberg
- KIT



ulm university universität  
**uulm**

Universität  
Konstanz



Hochschule  
für Technik  
Stuttgart



Universität Stuttgart

**Hochschule Esslingen**  
University of Applied Sciences



EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



UNIVERSITÄT  
MANNHEIM



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

**KIT**  
Karlsruher Institut für Technologie

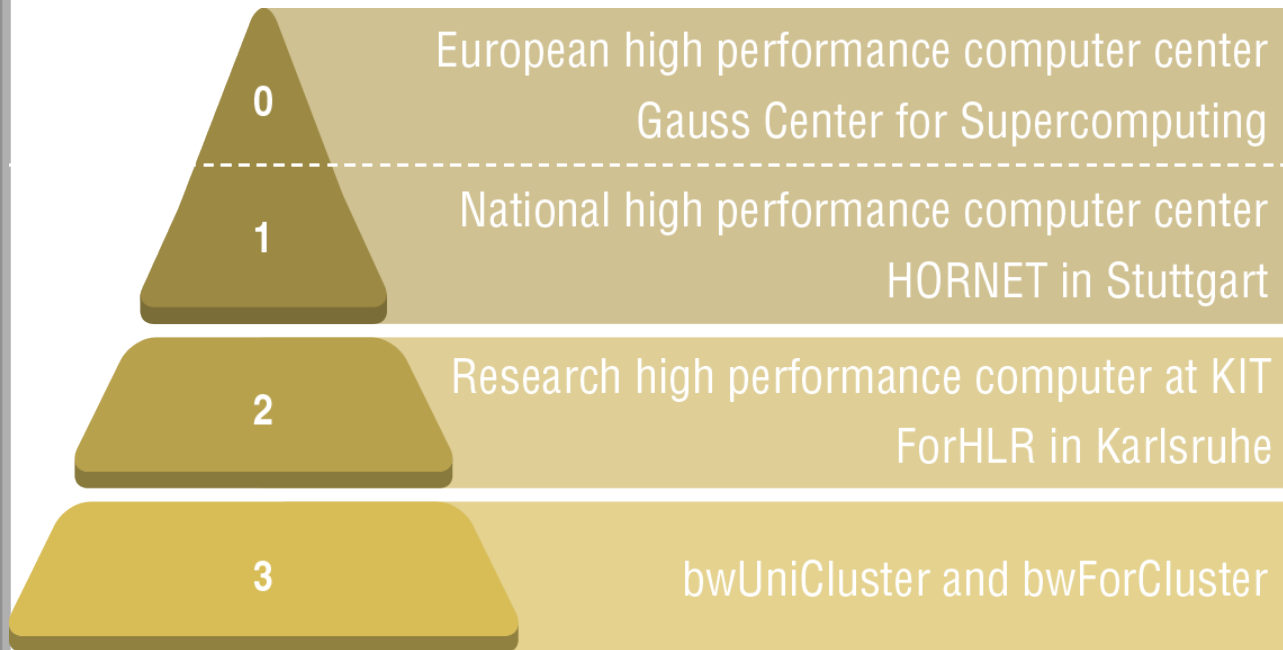
**Thank you for your attention!**

**Questions?**

# Agenda

Time	Title
09:30	bwHPC: Concept and User Support
09:45	Architecture: bwUniCluster + ForHLR
10:00	Cluster: Access, Modulefiles, Filesystem
10:30	Cluster: Batch System
11:00	Break
11:10	Tutorials: Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting
Max. 13:10	End

# HPC Tier classification / bwHPC



- European High-performance Computing Centers (Tier 0)
- National High-performance Computing Centers (Tier 1)
- Supraregional, state-wide HPC-Centers (Tier 2)
- Regional HPC-Enablers (Tier 3)

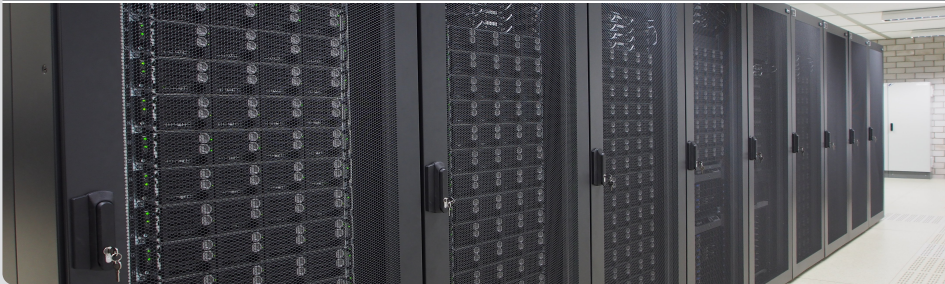
**Next presentation**

**→ Introduction bwHPC-Cluster: architecture bwUniCluster and ForHLR**

# bwUniCluster und ForHLR Hardware

(Richard Walter | 10. September 2014)

STEINBUCH CENTRE FOR COMPUTING



# Übersicht

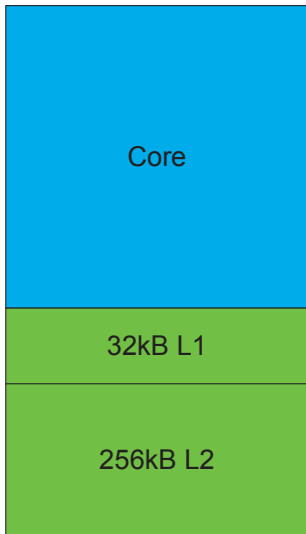
- CPUs
- Knoten
- Netzwerk



## Eigenschaften aller CPUs:

- Intel Sandy Bridge Mikroarchitektur
- ForHLR: Ivy Bridge
- Turbo Boost
  - dynamisches Übertakten
  - nicht abschaltbar
- Hyperthreading
  - derzeit nicht nutzbar
- AVX-Befehlssatz
  - Floating Point nur über Vektoreinheiten
  - 8 float oder 4 double je Vektor

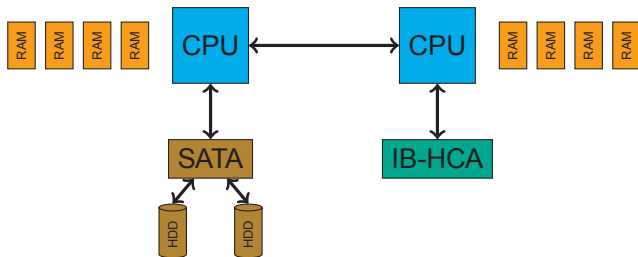
# CPUs (Aufbau)



# CPUs (Aufbau)

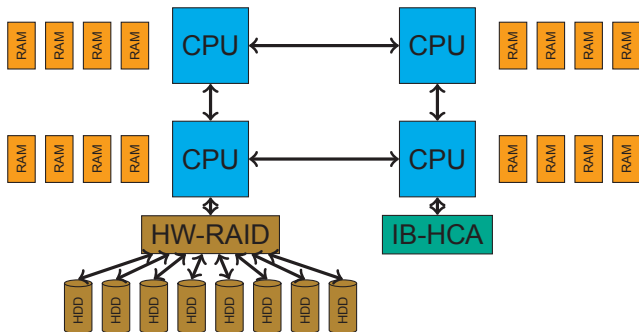
Core 1	Core 2	...	Core N
32kB L1	32kB L1	...	32kB L1
256kB L2	256kB L2	...	256kB L2
XX MB L3			

- plus 4 Speichercontroller DDR3
- plus 2 QPI 8GT/s



- 64 GiB RAM
- Festplatten je 1 TB, 7200 rpm, Software-RAID0
- InfiniBand: ConnectX-3 FDR (56 GBit/s Brutto, Netto > 6 GB/s)

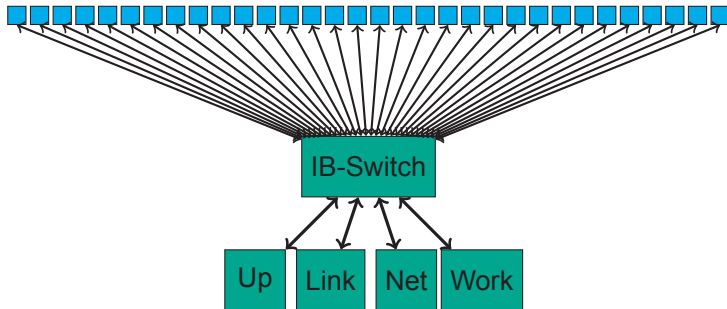
	bwUniCluster	ForHLR
Cores/Socket	8	10
CPU-Takt	2,6 GHz	2,5 GHz
L3-Cache/Socket	20 MiB	25 MiB
Speichertakt	1600 MHz	1866 MHz



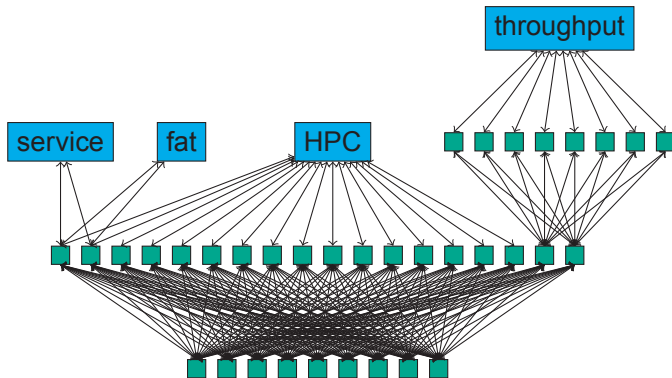
- CPUs mit 8 Cores je Socket, 20 MiB L3-Cache
- Festplatten je 1 TB, 7200 rpm, Hardware-RAID5
- InfiniBand: ConnectX-3 FDR (56 GBit/s Brutto, Netto > 6 GB/s)

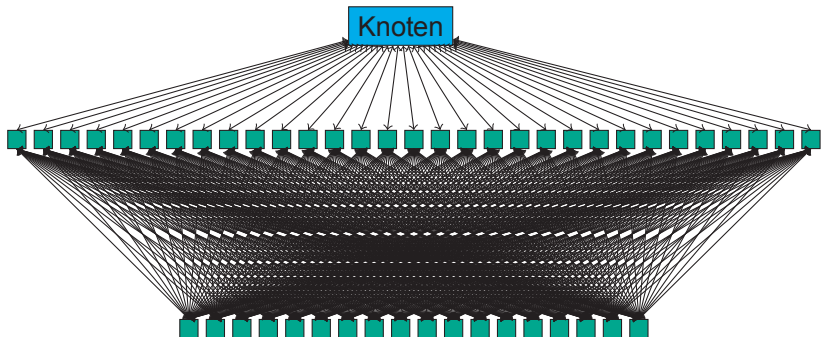
	bwUniCluster	ForHLR
CPU-Takt	2,4 GHz	2,6 GHz
Speicher	1 TiB @1333MHz	512 GiB @1600 MHz

- bwUniCluster zweigeteilt
  - Fat Tree-Bereich für HPC
  - Blocking-Bereich für Throughput Computing
    - Besteht aus 8 „Inseln“
    - Je 32 Knoten und 4 Uplinks
- ForHLR nur Fat Tree



# InfiniBandnetz (2, bwUniCluster)





Orientierung mit

- `$_SLURM_TOPOLOGY_ADDR`
- `$_SLURM_TOPOLOGY_ADDR_PATTERN`



# Agenda

Time	Title
09:30	bwHPC: Concept and User Support
09:45	Architecture: bwUniCluster + ForHLR
10:00	Cluster: Access, Modulefiles, Filesystem
10:30	Cluster: Batch System
11:00	Break
11:10	Tutorials: Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting
Max. 13:10	End





bw|HPC – C5

# Cluster: Access, Modulefiles, File System

Robert Barthel



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



**KIT**  
Karlsruher Institut für Technologie



ulm university universität  
**uulm**



# Outline

- Introduction:
  - Glossary, references
- Access
  - Registration
  - Login (Linux, Window, OS X)
- Usage
  - Command line
    - Basic commands
  - File transfer
  - Environment variables
  - File System
    - \$HOME, \$WORK, \$TMP, \$PROJECT
  - (Software) Modulefiles
    - Load, dependencies, unload

# 1. Intro

# Glossary/Remarks

## How to read this slides:

<code>\$ command</code> <code># command</code>	<code>user@machine:path\$ command</code> <code>user@machine:path# command</code>  → Command to be excuted at console
<code>\$var, \${var}</code>	Variable named <b>var</b> as used for <i>(ba)sh, ksh</i>
<code>&lt;string,int&gt;</code>	Placeholder for string, integer, etc.
<code>\$DIR</code>	Directory path
<code>\$PATH</code>	Environment variable specifying all directories with exectables
<code>foo, bar</code>	Metasyntactic variables
<b><i>ADV</i></b>	Advanced topic

## Introduction to Unix/Linux commands

■ <http://freeengineer.org/learnUNIXin10minutes.html>




# References/Literature

- Most important: [www.bwhpc-c5.de/wiki](http://www.bwhpc-c5.de/wiki)
- This slides:
  - [http://indico.scc.kit.edu/indico/event/Info-Veranst\\_2014-09\\_bwHPC](http://indico.scc.kit.edu/indico/event/Info-Veranst_2014-09_bwHPC)
  - [ab1234@bwunicluster:/opt/bwhpc/kit/workshop/2014-09-10](mailto:ab1234@bwunicluster:/opt/bwhpc/kit/workshop/2014-09-10)
- Manpages: command line applications when logged in:  
`$ man command` e.g. `$ man ls`
- Bash scripting e.g. for Batch jobs
  - <http://tldp.org/HOWTO/Bash-Prog-Intro-HOWTO.html> (intro)
  - <http://tldp.org/LDP/abs/html> (advanced)
- Environment modulefiles:
  - [http://www.bwhpc-c5.de/wiki/index.php/Environment\\_Modules](http://www.bwhpc-c5.de/wiki/index.php/Environment_Modules)
- MOAB queueing system:
  - [http://www.bwhpc-c5.de/wiki/index.php/Batch\\_Jobs](http://www.bwhpc-c5.de/wiki/index.php/Batch_Jobs)

## 2. Access

# bwHPC cluster – Registration

- HowTo: <http://www.bwhpc-c5.de/wiki/index.php/Category:Access>
- Example: **bwUniCluster**
  - Cluster entitlement application?
    - Details: [http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster\\_User\\_Access](http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster_User_Access)



**Karlsruhe Institute of Technology (KIT)**  
Steinbuch Centre for Computing (SCC)  
Leitung: Prof. Dr. Hans-Joachim Henkel  
Prof. Dr. Richard W. Vasek  
Prof. Dr. Adam Strat

ServiceDesk (BIT 8000)  
Zielfeld 2 (Geb. 20.21)  
76131 Karlsruhe

E-Mail: [service@scs.kit.edu](mailto:service@scs.kit.edu)  
Web: <http://www.scc.kit.edu>  
Tel: 0721 608 3000  
Fax: 0721 608 3006  
FaxCN: 0721 608 99308

**Antrag auf Zugangsberechtigung für den bwUniCluster**  
Nur für Mitarbeiter und Partner des KIT

Datum | by

<b>1. Angaben zum Institut</b>			
Institut	by		
Institutskostenstelle	by		
1.1. Leiter/in des Instituts bzw. der Einrichtung			
Name	by	Vorname	by
Telefon-Nr.	by	Fax-Nr.	by
E-Mail	by	KIT-Konto	by

<b>2. Angaben zum Projekt</b>	
Titel des Projekts	by
Gewünschter permanenter Plattenplatz (in GigaByte)	by
Fachliche Ausrichtung	by
Kurzbeschreibung des Projekts	by

<b>3. Angaben zu Benutzer/in</b>			
Anrede	<input type="checkbox"/> Herr <input type="checkbox"/> Frau		
Name	by	Vorname	by
Telefon-Nr.	by	Fax-Nr.	by
E-Mail	by	KIT-Konto	by

1/2 Weiter auf folgender Seite, bitte wenden!

<b>4. Hinweise</b>	
Zur Freischaltung Ihrer Zugangsberechtigung müssen Sie sich auf der folgenden Webseite mit Ihrem KIT-Konto authentifizieren: <a href="https://bwidm.scc.kit.edu">https://bwidm.scc.kit.edu</a>	
Zur Ressourcennutzung sind das Handbuch bzw. die Web-Seiten zur entsprechenden Rechenanlage zu beachten. Das SCC speichert aus betrieblichen Notwendigkeiten (Abrechnung, Fehleranalyse, Ressourcenplanung und zur Betriebssicherheit) auch personenbezogene Daten. Werden Verstöße gegen die Bestimmungen des Außenwirtschaftsgesetzes (AWG) und der Außenwirtschafts-Verordnung (AWV) eruiert, können sich daraus für die betroffenen BenutzerInnen strafrechtliche Konsequenzen ergeben.	
<b>5. Erklärung</b>	
Die Einhaltung der IoK-Ordnung des KIT wird zugesichert. Siehe <a href="http://www.kit.edu/medien/inf/amtliche/Rechenanwendungen/2013_Ab_036.pdf">http://www.kit.edu/medien/inf/amtliche/Rechenanwendungen/2013_Ab_036.pdf</a> Mir ist bekannt, dass für die Nutzung des bwUniCluster die Bestimmungen des Außenwirtschaftsgesetzes (AWG) und der Außenwirtschafts-Verordnung (AWV) einzuhalten sind. Siehe <a href="http://www.ausfuhrkontrolle.info">http://www.ausfuhrkontrolle.info</a>	
<b>6. Unterschrift Leiter/in des Instituts bzw. der Einrichtung</b>	
Institutsstempel	Name in Klerschrift   by
Datum, Unterschrift	
<b>7. Unterschrift Benutzer/in</b>	
Ich habe die obigen Hinweise und Erklärungen gelesen und verpflichte mich, diese zu beachten!	
Datum, Unterschrift	
<b>8. Nur von SCC auszufüllen</b>	
Datum	Bearbeitet von
2/2	

- Each university has its own entitlement granting policies!

left: KIT forms



# bwUniCluster – Registration (1)

■ Once entitlement is issued:

→ web registration: <https://bwidm.scc.kit.edu/>

**KIT**  
Karlsruher Institut für Technologie

**Landesdienste am KIT**

**SCC**

Browsersprache ▼

**Welcome**

In order to use bwServices at KIT you need a valid user account with one of the following organisations. Please choose your home organisation from the list and click on "Continue".

Föderation: bwIdm (id 1001) ▼

Search filter:

Heimatorganisation:

- Albert-Ludwigs-Universität Freiburg
- Hochschule Albstadt-Sigmaringen
- Hochschule Esslingen
- Hochschule für Technik Stuttgart
- Hochschule Furtwangen University
- Hochschule Karlsruhe - Technik und Wirtschaft
- Karlsruher Institut für Technologie (KIT)
- Universität Heidelberg
- Universität Hohenheim
- Universität Konstanz
- Universität Mannheim
- Universität Stuttgart
- Universität Tübingen
- Universität Ulm

Proceed

bwreg-webapp-kit-prod-1.2.0

KIT – Universität des Landes Baden-Württemberg und nationales Forschungszentrum in der Helmholtz-Gemeinschaft

# bwUniCluster – Registration (2)

- Once entitlement is issued:

→ web registration: <https://bwidm.scc.kit.edu/>

**KIT**  
Karlsruher Institut für Technologie

**Shibboleth Identity Provider**

Sie wurden von dem Serviceprovider **BW-federal services at KIT** hierher weitergeleitet und befinden sich nun auf einem Server des KIT. Bitte melden Sie sich mit Ihrem KIT-Account (z.B. ab1234 als Mitarbeiter oder uxxxx als Student) und Ihrem Passwort an.

**Anmeldung**

Benutzerkennung:  (z.B. ab1234 als Mitarbeiter oder uxxxx als Student)

Passwort:

Die oben bezeichnete Webseite des Serviceanbieters bittet Sie, sich bei Ihrer Heimateinrichtung anzumelden.

Sie bekommen auf der Folgeseite die Daten angezeigt, um deren Übermittlung der Serviceprovider bittet. Sie können dies bestätigen und damit den Vorgang fortsetzen oder durch Schließen des Fensters abbrechen. Haben Sie denselben Service bereits einmal genutzt, werden Sie nur dann erneut nach einer Bestätigung gefragt, wenn sich der Datenumfang oder der Name des Serviceanbieters geändert hat. Wenn Sie auf jeden Fall nochmal sehen möchten, welche Daten zur Übermittlung vorgesehen sind, aktivieren Sie bitte nachstehende Option.

Bitte zeige mir für diesen Serviceprovider erneut an, welche Daten gesendet werden sollen.

The service provider's website, shown on top, requests that you log in to your home site.

The following page displays the details that the service provider is requesting. You can confirm this, thereby continuing the process, or cancel it by closing the window. If you have used this service before, you will only be asked to reconfirm if the scope of data or the service provider's name has changed.

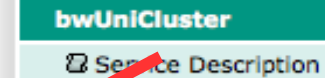
Check the option below if you would like to review the details that will be sent.

Please display once again the details that will be sent for this service provider.

- left: authentication for KIT members

# bwUniCluster – Registration (3)

## After login: select bwUniCluster Service Description



bwUniCluster  
Service Description

### Service Description: bwUniCluster

Der am Steinbuch Centre for Computing (SCC) des Karlsruher Institut für Technologie (KIT) betriebene **bwUniCluster** ist eines von mehreren zentralen Systemen für eine flächendeckende Grundversorgung der baden-württembergischen Universitäten und Hochschulen mit Hochleistungsrechnerkapazität. Der Cluster wird von den Landesuniversitäten und dem MWK getragen und ist Teil des baden-württembergischen Landeskonzepts für das Hoch- und Höchstleistungsrechnen „bwHPC“ in Forschung und Lehre ([http://www.bwhpc-c5.de/bwhpc\\_konzept.php](http://www.bwhpc-c5.de/bwhpc_konzept.php)).

Der **bwUniCluster** steht zunächst den unten genannten Einrichtungen entsprechend ihrer Finanzierungsanteile zur Verfügung und dient allein als Einstiegs- und Grundversorgungssystem für Hochleistungsrechnen in Forschung und Lehre. Eine darüber hinausgehende Nutzung, gegebenenfalls auch durch weitere Landeshochschulen oder Einrichtungen, muss über den Landesnutzerausschuss bzw. den Arbeitskreis der Leiter wissenschaftlicher Rechenzentren in Baden-Württemberg, ALWR, explizit beantragt werden.

#### Verfügbare Ressourcen:

Das bwUniCluster besteht aus 512 „dünnen“ Knoten mit jeweils 16 Kernen, 64 GB RAM Arbeitsspeicher und 2 TB lokalem Plattenplatz. Zusätzlich stehen 8 „dicke“ Knoten mit jeweils 32 Kernen, 1 TB RAM Arbeitsspeicher und 7 TB lokalem Plattenplatz zur Verfügung. Weitere Details entnehmen Sie <http://www.scc.kit.edu/dienste/9237.php>.

#### Ansprechpartner bezüglich Registrierung:

Bei Fragen bezüglich Registrierung, Nutzungsberechtigung und Nutzungsordnung für den bwUniCluster, wenden Sie sich bitte an Ihren Standortbetreuer:

- Albert-Ludwigs-Universität Freiburg: [dgrid-support@bfg.uni-freiburg.de](mailto:dgrid-support@bfg.uni-freiburg.de)
- Eberhard Karls Universität Tübingen: [hpcmaster@uni-tuebingen.de](mailto:hpcmaster@uni-tuebingen.de)
- KIT: [bwunicluster-hotline@lists.kit.edu](mailto:bwunicluster-hotline@lists.kit.edu)
- Ruprecht-Karls-Universität Heidelberg: [hpc-support@listserv.uni-heidelberg.de](mailto:hpc-support@listserv.uni-heidelberg.de)
- Universität Hohenheim: [kjm-bw-projekt@uni-hohenheim.de](mailto:kjm-bw-projekt@uni-hohenheim.de)
- Universität Konstanz: [support@uni-konstanz.de](mailto:support@uni-konstanz.de)
- Universität Mannheim: [hpc-support@mailman.uni-mannheim.de](mailto:hpc-support@mailman.uni-mannheim.de)
- Universität Stuttgart: [bwunicluster@hirs.de](mailto:bwunicluster@hirs.de)
- Universität Ulm: [helpdesk@uni-ulm.de](mailto:helpdesk@uni-ulm.de)

#### Rechtliche Hinweise zur Nutzung des bwUniClusters:

Die Ordnung für die digitale Informationsverarbeitung und Kommunikation (IuK) am Karlsruher Institut für Technologie (KIT) sowie die Bestimmungen des Außenwirtschaftsgesetzes (AWG) und der Außenwirtschafts-Verordnung (AWV) sind einzuhalten, siehe [http://www.kit.edu/downloads/AmtlicheBekanntmachungen/2013\\_AB\\_036.pdf](http://www.kit.edu/downloads/AmtlicheBekanntmachungen/2013_AB_036.pdf) und <http://www.ausfuhrkontrolle.info>. Des Weiteren sind die Benutzerordnungen der entsprechend beteiligten Einrichtungen zu beachten.

- Albert-Ludwigs-Universität Freiburg: <https://www.bwhpc-c5.uni-freiburg.de/bwunicluster/freischaltung-uni-account>
- Eberhard Karls Universität Tübingen: <http://www.zdv.uni-tuebingen.de/kontakt-antraege-beratung/antraege/e-antraege.html>
- Ruprecht-Karls-Universität Heidelberg: <http://www.urz.uni-heidelberg.de/md/urz/orginfo/ordnungen/vbo.pdf>
- Universität Hohenheim: [url]
- Universität Konstanz: <http://www.rz.uni-konstanz.de/dienste/computing-hpc-bwunicluster-bwgrid/>
- Universität Mannheim: [http://www.uni-mannheim.de/rum/ivs/benutzerverwaltung/benutzerkennung/verwaltungs\\_und\\_benutzungsordnung.pdf](http://www.uni-mannheim.de/rum/ivs/benutzerverwaltung/benutzerkennung/verwaltungs_und_benutzungsordnung.pdf)
- Universität Stuttgart: [http://www.uni-stuttgart.de/zv/bekanntmachungen/bekanntm\\_179.html#3](http://www.uni-stuttgart.de/zv/bekanntmachungen/bekanntm_179.html#3)
- Universität Ulm: [http://www.uni-ulm.de/fileadmin/website\\_uni\\_ulm/kiz/org/kiz-bo.pdf](http://www.uni-ulm.de/fileadmin/website_uni_ulm/kiz/org/kiz-bo.pdf)

Werden Verstöße gegen die Benutzerrichtlinien oder gegen gesetzliche Bestimmungen festgestellt, können sich daraus für die betroffenen Benutzer/innen strafrechtliche Konsequenzen ergeben. Bei unsachgemäßer oder missbräuchlicher Benutzung behält sich der Betreiber des bwUniClusters auch zivilrechtliche Schritte vor.

#### Hinweise zur Danksagung:

Bei der Veröffentlichung von Ergebnissen, die unter Zuhilfenahme des bwUniClusters erhalten wurden, ist eine schriftliche Erwähnung des bwUniClusters in der Danksagung notwendig:

*"This work was performed on the computational resource bwUniCluster funded by the Ministry of Science, Research and Arts and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC."*

Weiterhin ist die Öffentlichkeitsabteilung des Projekts bwHPC-C5 ([bwhpc-c5\\_dissemination@lists.kit.edu](mailto:bwhpc-c5_dissemination@lists.kit.edu)) formlos mit Angaben über Autoren, Titel, Erscheinungsform (Bandnummer etc.) und Veröffentlichungsdatum zu informieren.

[Register](#)

# bwUniCluster – Registration (3)

## Read terms of usage and accept:

### Register: bwUniCluster

**Requirements**

All requirements are met for this service.

**Register**

**Um sich für den Dienst zu registrieren, müssen Sie den folgenden Nutzungsbedingungen zustimmen.**

Nutzungsbedingungen bwUniCluster - Version 1 (Januar 2014)

**Nutzungsbedingungen und -richtlinien:**

Der am Steinbuch Centre for Computing (SCC) des Karlsruher Institut für Technologie (KIT) betriebene **bwUniCluster** ist eines von mehreren zentralen Systemen für eine flächendeckende Grundversorgung der baden-württembergischen Universitäten und Hochschulen mit Hochleistungsrechnerkapazität. Der Cluster wird von den Landesuniversitäten und dem MWK getragen und ist Teil des baden-württembergischen Landeskonzepts für das Hoch- und Höchstleistungsrechnen "bwHPC" in Forschung und Lehre ([http://www.bwhpc-c5.de/bwhpc\\_konzept.php](http://www.bwhpc-c5.de/bwhpc_konzept.php)).

Der **bwUniCluster** steht zunächst den unten genannten Einrichtungen entsprechend ihrer Finanzierungsanteile zur Verfügung und dient allein als Einstiegs- und Grundversorgungssystem für Hochleistungsrechnen in Forschung und Lehre. Eine darüber hinausgehende Nutzung, gegebenenfalls auch durch weitere Landeshochschulen oder Einrichtungen, muss über den Landesnutzerausschuss bzw. den Arbeitskreis der Leiter wissenschaftlicher Rechenzentren in Baden-Württemberg, ALWR, explizit beantragt werden.

Zur Ressourcennutzung sind das Handbuch bzw. die Webseiten zur entsprechenden Rechanlage zu beachten ([http://www.bwhpc-c5.de/wiki/index.php/bwUniCluster\\_User\\_Guide](http://www.bwhpc-c5.de/wiki/index.php/bwUniCluster_User_Guide)). Das KIT speichert aus betrieblichen Notwendigkeiten (Abrechnung, Fehleranalyse, Ressourcenplanung und zur Betriebssicherheit) auch personenbezogene Daten. Werden Verstöße gegen die Benutzerrichtlinien oder gegen gesetzliche Bestimmungen festgestellt, können sich daraus für die betroffenen Benutzer/innen strafrechtliche Konsequenzen ergeben. Bei unsachgemäßer oder missbräuchlicher Benutzung behält sich der Betreiber des bwUniClusters auch zivilrechtliche Schritte vor.

Bei der Veröffentlichung von Ergebnissen, die unter Zuhilfenahme des bwUniClusters erhalten wurden, ist eine schriftliche Erwähnung des bwUniClusters in der Danksagung notwendig: "This work was performed on the computational resource bwUniCluster funded by the Ministry of Science, Research and Arts and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC."

Weiterhin ist die Öffentlichkeitsabteilung des Projekts bwHPC-C5 ([bwhpc-c5\\_dissemination@lists.kit.edu](mailto:bwhpc-c5_dissemination@lists.kit.edu)) formlos mit Angaben über Autoren, Titel, Erscheinungsform (Bandnummer etc.) und Veröffentlichungsdatum zu informieren.

**Datenschutz:**  
Bei der Registrierung für den bwUniCluster werden Nutzerinformationen von der Heimateinrichtung an den Dienstbetreiber KIT verschlüsselt übermittelt und dort gespeichert.

**Erklärung:**  
Die Einhaltung der oben genannten Nutzungsbedingungen und -richtlinien sowie der IuK-Ordnung des KIT wird zugesichert, siehe [http://www.kit.edu/downloads/AmtlicheBekanntmachungen/2013\\_AB\\_036.pdf](http://www.kit.edu/downloads/AmtlicheBekanntmachungen/2013_AB_036.pdf) Mir ist bekannt, dass für die Nutzung des bwUniCluster die Bestimmungen des Außenwirtschaftsgesetzes (AWG) und der Außenwirtschafts-Verordnung (AWV) einzuhalten sind, siehe <http://www.ausfuhrkontrolle.info>.

I have read and accepted the terms of use.

bwUniCluster

Service Description

With acceptance you also agree to the acknowledgement policy, for details see:

[http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster\\_Acknowledgement](http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster_Acknowledgement)

# bwUniCluster – Registration (4)

## If successful:

You have already registered with the following services:

bwUniCluster	IC2	HC3
<p>Der am Steinbuch Centre for Computing (SCC) des Karlsruher Institut für Technologie (KIT) betriebene bwUniCluster ist eines von mehreren zentralen Systemen für eine flächendeckende Grundversorgung der baden-württembergischen Universitäten und Hochschulen mit Hochleistungsrechnerkapazität.</p> <ul style="list-style-type: none"> <li>Service Description</li> <li>Registry Info</li> <li>Set Service Password</li> </ul>	<p>Das SCC betreibt ein von der DFG gefördertes und mit mehreren Instituten des KIT gemeinsam beschafftes Computersystem mit dem Namen InstitutsCluster II. Er steht den an dem Förderantrag beteiligten Instituten gemäß ihrer dort genannten Anteile zur Verfügung</p> <ul style="list-style-type: none"> <li>Service Description</li> <li>Registry Info</li> <li>Set Service Password</li> </ul>	<p>Die HC3 ist ein paralleles Rechnersystem HP XC3000, das aus vielen SMP-Knoten mit 64-bit Xeon Prozessoren von Intel besteht. Der Rechner kann und soll sowohl die Aufgaben eines parallelen Hochleistungsrechners als auch die Aufgaben eines seriellen, durchsatzorientierten Rechners erfüllen.</p> <ul style="list-style-type: none"> <li>Service Description</li> <li>Registry Info</li> <li>Set Service Password</li> </ul>

In order to see details from the services you registered with, click on the 'Registry Info' under the service.

The following services are available:

bwFileStorage	bwSync&Share
<p>Der Dienst bwFileStorage bietet den Nutzern der Rechenzentren der Universitäten und Hochschulen im Land Baden-Württemberg filesystembasierten Zugriff auf den Datenspeicher der Large Scale Data Facility (LSDF) am Karlsruher Institut für Technologie (KIT).</p> <ul style="list-style-type: none"> <li>Service Description</li> <li>Register</li> </ul>	<p>bwSync&amp;Share ist ein Online Speicherdienst, der es ermöglicht, Ihre Daten zwischen verschiedenen Computern, mobilen Endgeräten und Benutzern zu synchronisieren bzw. auszutauschen und gleichzeitig in der Large Scale Data Facility (LSDF) am Karlsruher Institut für Technologie (KIT) zu sichern.</p> <ul style="list-style-type: none"> <li>Service Description</li> <li>Register</li> </ul>

To register with a service, click on the 'Register' link below the service, you want to register with.

**bwUniCluster**

- Service Description

**bwUniCluster**

- Service Description
- Registry Info
- Set Service Password






# bwUniCluster – Registration (5)

## Read registry information:

Registry Info: bwUniCluster	
Registered at:	14.08.2014 17:25
Status:	ACTIVE
Name	Value
mail	robert.barthel@kit.edu
sn	Unknown
homeDir	/home/kit/scc/scc-adm
groupName	scc
cn	scc-adm @kit.edu
sambaEnabled	false
description	9421132
localUid	scc-adm
uidNumber	Unknown
gidNumber	Unknown
givenName	Unknown
groups	ka_scc-entitlement-bwgrid;ka_scc-entitlement-bwunicluster;ka_scc-entitlement-ic2;ka_scc-entitlement-bwlsdf-fs;ka_scc-entitlement-hc3

**bwUniCluster**


-  Service Description
-  Registry Info
-  Set Service Password

More Properties: bwUniCluster	
Fetching Account from 5 Server(s): Success	
Name	Message
Server 1	Fetching Account success
Server 2	Fetching Account success
Server 3	Fetching Account success
Server 4	Fetching Account success
Server 5	Fetching Account success

Deregister: bwUniCluster
You no longer want to use this service. <b>Deregister</b>

For deregistration

## ForHLR – Registration

- [http://www.bwhpc-c5.de/wiki/index.php/ForHLR\\_User\\_Access](http://www.bwhpc-c5.de/wiki/index.php/ForHLR_User_Access)
  
  - Submit Project Proposal
    - After approval by referees
  
  - Web Registration
    - <https://bwidm.scc.kit.edu>
- 

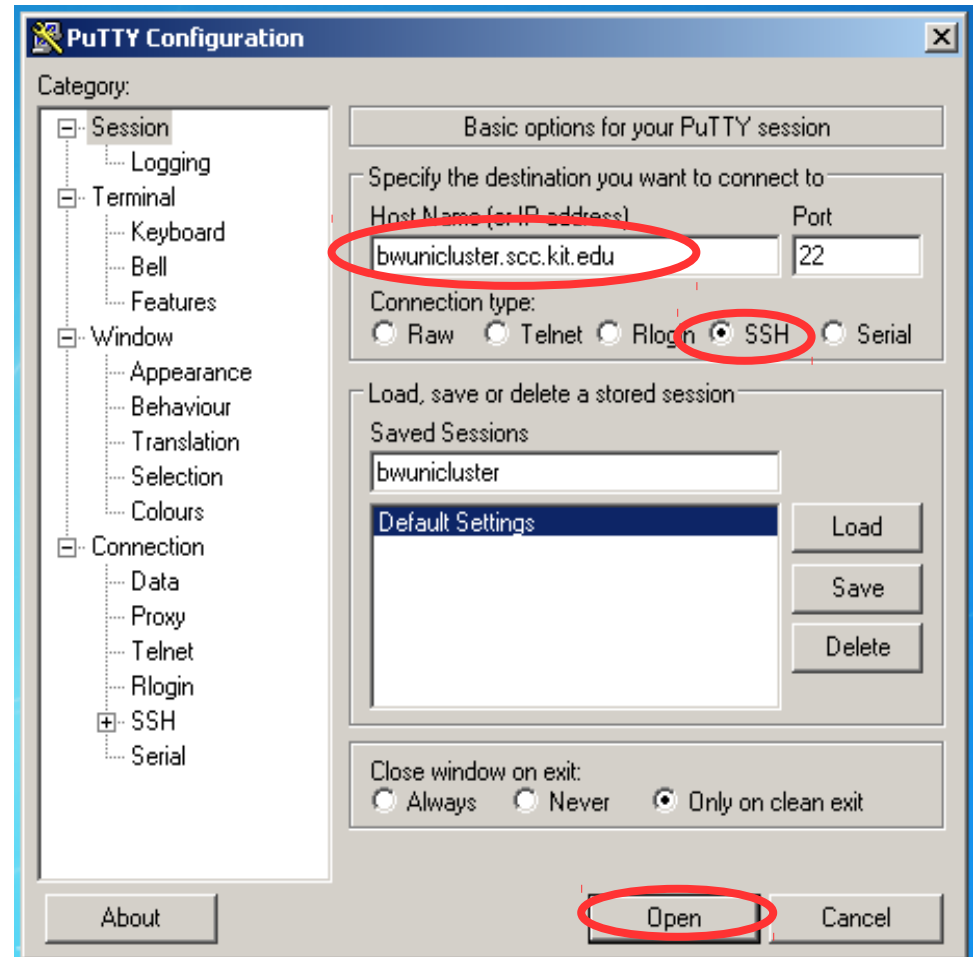




# Login (2)

## Windows

- Download SSH client, e.g. PuTTY (<http://www.putty.org/>)
  - Run putty.exe
    - Host Name:  
**bwunicluster.scc.kit.edu**
    - Connection type:  
**SSH**
    - Username:
      - a) For KIT:  
**ab1234**
      - b) For all other universities:  
**<prefix>\_<username>**



# Login (2)

## Windows

### ■ Add host key



# Login (2) Windows

- Enter your account name + password

→ If done:

```

bwunicluster.scc.kit.edu - PuTTY
login as: █

login as: yc8563
yc8563@bwunicluster.scc.kit.edu's password:
Last login: Sun Feb 16 10:10:29 2014 from openvpn-cl-200-232.scc.kit.edu
*****
*
*      Universal HPC cluster of Baden-Wuerttemberg's universities:
*
*
*      [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / /
*      [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / /
*      [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / / [ _ ] \ \ \ \ / /
*
*
*      (KITE 2.0/RHEL6.4/Lustre 2.4.1)
*
*      https://www.bwhpc-c5.de/wiki/index.php/bwUniCluster_User_Guide
*
*
*      hotline: bwunicluster-hotline@lists.kit.edu
*
*
*
*
*
*
*      KIT News:
*      2014-02-06:
*      - seminar about bwHPC/bwUniCluster (+ hands-on) on February 19th 2014
*      http://indico.scc.kit.edu/indico/event/Info-Veranst_2014-02_bwUniCluster
*
*
*
*****
[Feb-16 10:12] yc8563@uc1n996:~$ █
  
```

## 3. Usage

# Command line interface/interpreter

- Default: **BASH** = bourne again shell

- Introduction to Unix/Linux commands

- <http://freeengineer.org/learnUNIXin10minutes.html>

- Directories: „/“ to separate in path

- Program execution:

- a) global program:

```
$ program
```

- b) from local directory:

```
$ ./program
```

- Global Unix/Linux commands:

- Moving around:

*cd, pwd*

- Listing directory content:

*ls*

- Changing file permissions and attributes:

*chmod, chown*

- Moving, renaming, and copying files:

*mv, cp*

- Viewing and editing files:

*cat, less, vim, nano*

- Filename completion:

*TAB key*

## File transfer – SCP via Linux/OS X

- Copy *testfile* from your computer's local directory to your bwUniCluster directory \$HOME/transfer:

```
$ scp testfile ab1234@bwunicluster.scc.kit.edu:~/transfer
```

- Copy folder *testfolder* to your bwUniCluster \$HOME

```
$ scp -r testfolder ab1234@bwunicluster.scc.kit.edu:~
```

- Copy your bwUniCluster folder \$HOME/transfer to your local computer:

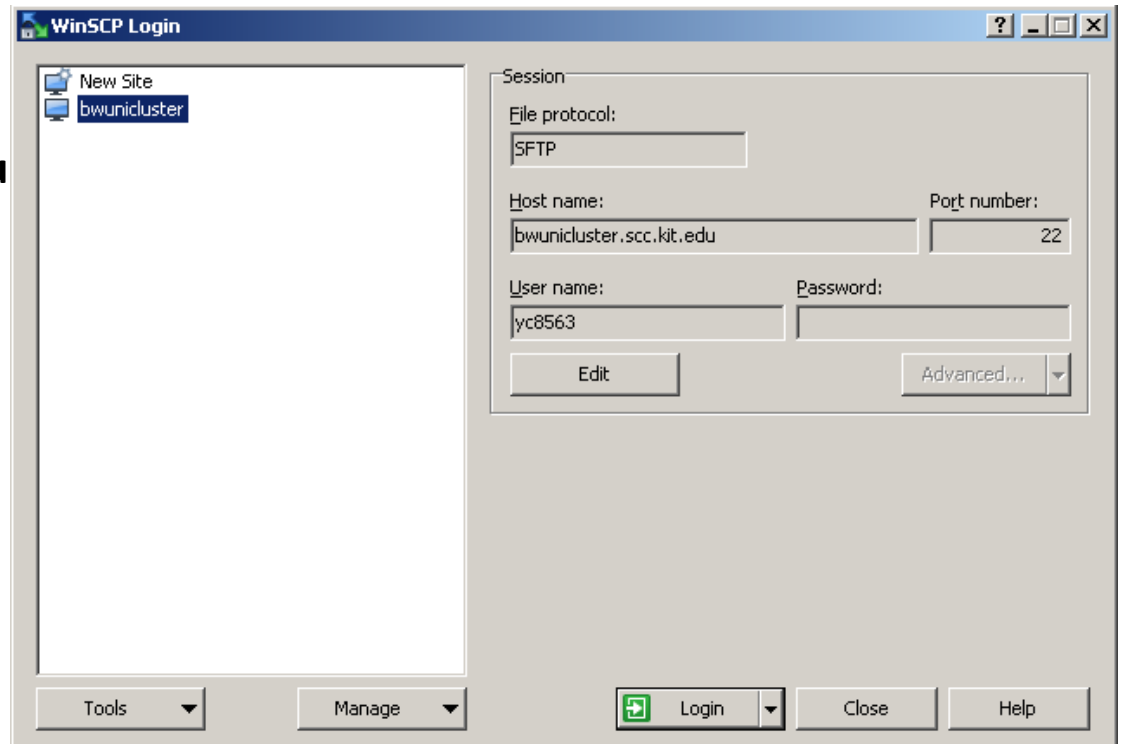
```
$ scp -r ab1234@bwunicluster.scc.kit.edu:~/transfer .
```

- further information:

```
$ man scp
```

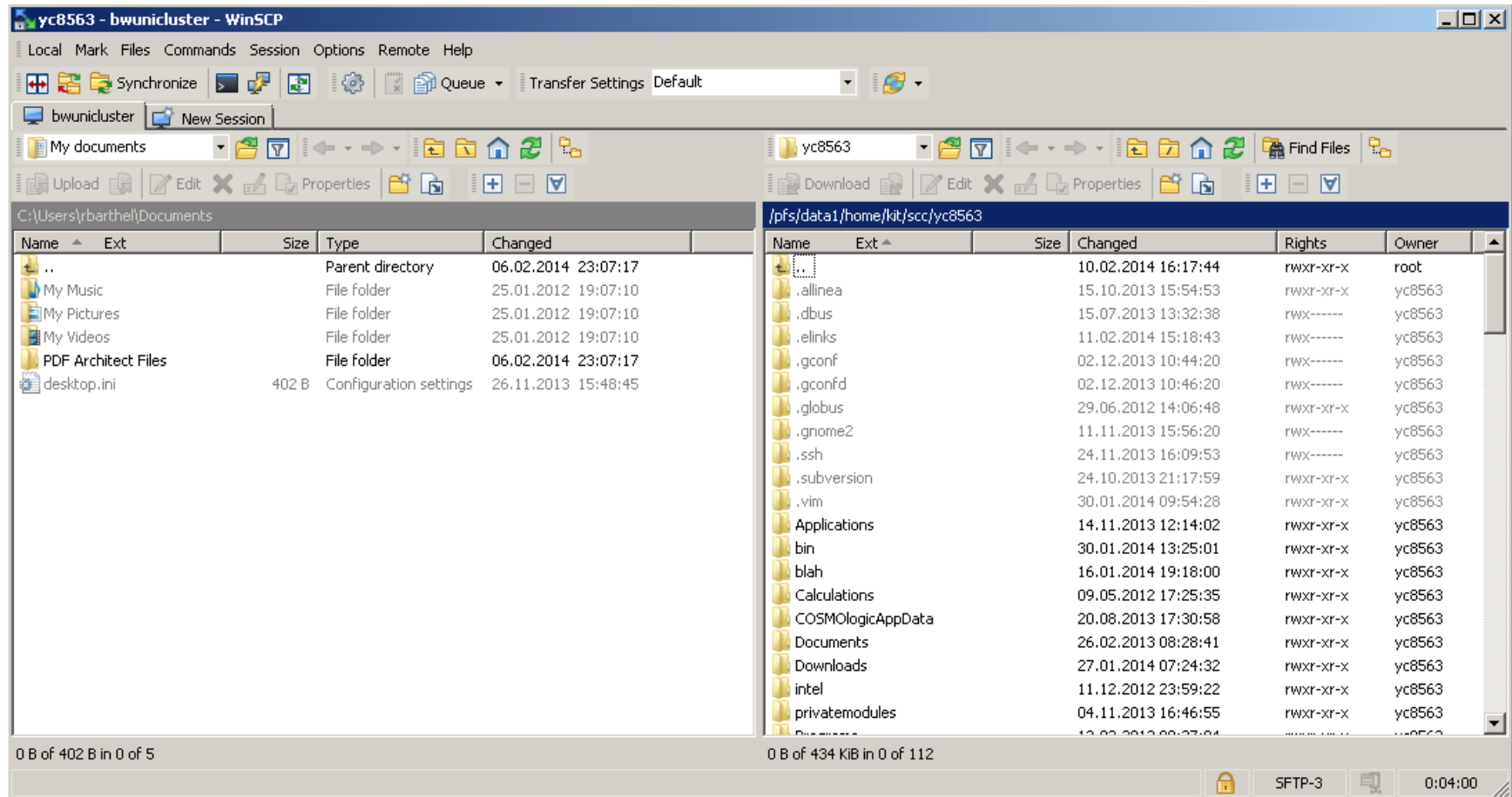
## File transfer – SCP (2)

- for Windows user → install WinSCP: <http://winscp.net>
- run WinSCP
  - Host Name:  
**bwunicluster.scc.kit.edu**
  - File protocol:  
**SFTP**
  - Username:  
**ab1234**



# File transfer – SCP (2)

- for Windows user → install WinSCP: <http://winscp.net>
- run WinSCP





# Environment variables

- you can set variables, e.g. seminar=2014-09-10, remembered by your shell:

```
$ seminar=2014-09-10
```

→ to return value of seminar:

```
$ echo $seminar
$ 2014-09-10
```

- Some important Linux bwUniCluster/ForHLR global variables:

\$HOME or ~/	Path to your home directory, parallel file space under backup
\$PATH	Set of directories (separated by :) searched through by system when calling a program
\$USER, \$LOGNAME	Login user name
\$WORK	Additional file space, parallel file space <b>not under</b> backup
\$TMP	Local file space on login or compute nodes, no backup
\$PROJECT	ONLY ForHLR: Path to your project directory

# File System

## ■ Characteristics of e.g. bwUniCluster

Property	\$TMP	\$HOME	\$WORK	workspace
Visibility	local	global	global	global
Lifetime	batch job walltime	permanent	> 7 days	max. 240 days
Disk space	2 TB @ thin nodes 7 TB @ fat nodes 4 TB @ login nodes	469 TB	938 TB	938 TB
Quotas	no	if required	if required	if required
Backup	no	yes (default)	no	no
Read perf./node	280 MB/s @ thin node 593 MB/s @ fat node 416 MB/s @ login node	1 GB/s	1 GB/s	1 GB/s
Write perf./node	270 MB/s @ thin node 733 MB/s @ fat node 615 MB/s @ login node	1 GB/s	1 GB/s	1 GB/s
Total read perf.	n*280 593 MB/s	8 GB/s	16 GB/s	16 GB/s
Total write perf.	n*270 733 MB/s	8 GB/s	16 GB/s	16 GB/s

## ■ \$HOME, \$WORK and workspaces: on the parallel file system Lustre

→ **BUT:** only \$HOME under backup

# \$HOME = Home directory

## ■ \$HOME:

- Quotas based on each university's share

- Current quota: `$ lfs quota -u $(whoami) $HOME`

- \$HOME directories of bwUniCluster, HC3 and IC2 are the same

- But: different operational systems, hardware, libraries, queueing etc.  
→ bwUniCluster (OS = REHL) vs. IC2 & HC3 (OS = SLES)

→ Implications: *you need different binaries and scripts*

e.g.: bin\_ic2.x, bin\_uc1.x, script\_ic2.sh, script\_uc1.sh

**ADV:** generalise your scripts to work on all systems using **\$CLUSTER**

```
if [ ${CLUSTER} == "ic2" ]; then
  <operations>
elif [ ${CLUSTER} == "uc1" ]; then
  <operations>
fi
```

# \$PROJECT = Project directory of ForHLR

- ONLY ForHLR:
  - All features of \$HOME
  - Access granted based on approved projects
    - assigned „name/acronym“
    - \$PROJECT\_GROUP
  - Access project home directory: `$ cd $PROJECT`
  - **Do not use: \$HOME** → since it has very low quota for the project group
  - Quota of Project: `$ lfs quota -g ${PROJECT_GROUP} /project`

# \$WORK = Working directory

- bwUniCluster/ForHLR → **additional parallel file system** with **limited lifetime, no redundancy, quotas**
  - especially designed for parallel access and for a high throughput to large files
- 2 concepts of access via:
  - (A) → \$WORK
  - (B) → *workspaces*
- (A) **\$WORK**:
  - Change to it via: `$ cd $WORK`
  - Quota: `$ lfs quota -u $(whoami) $WORK`
  - But: files no longer needed should be removed
    - any file inside your \$WORK older than 28 days will be deleted

## Workspaces = Working directory

### ■ (B) Workspaces: lifetime on allocated folder

#### ■ HowTo:

→ [http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster\\_File\\_System#Workspaces](http://www.bwhpc-c5.de/wiki/index.php/BwUniCluster_File_System#Workspaces)

\$ ws_allocate foo 10	Allocate a workspace named <i>foo</i> for 10 days
\$ ws_list -a	List all your workspaces
\$ ws_find foo	Get absolute path of workspace <i>foo</i>
\$ ws_extend foo 5	Extend lifetime of your workspace <i>foo</i> by 5 days from now. You can extend 3 times → <b>max.</b> lifetime of <i>foo</i> = <b>240</b> days
\$ ws_release foo	Manually erase your workspace <i>foo</i>

Example:

```
$ ws_allocate scratch
$ SDIR=$(ws_find scratch)
$ echo $SDIR
/work/workspace/scratch/ab1234-scratch-0
```

# \$HOME/\$WORK: Improving Performance

## ■ Improving Throughput Performance

- Sequential I/O in large buffers is good
- More clients/processes doing I/O improves overall throughput
- Multiple files are automatically distributed to storage systems
  - Change nothing if many (>10) large files are used concurrently
- If many clients are doing I/O to one large file
  - Force distribution of newly created files to all storage systems:  

```
$ lfs setstripe -c-1 <parent directory>
```
  - Try to omit concurrent write over 1 MB (stripe) boundaries

## ■ Improving Metadata Performance

- Try to reduce number of files
- Avoid competitive file or directory access
  - e.g. appending or writing to the same area from different clients
- If lots of small files are created on \$WORK use stripe count 1
  - Default stripe count for \$WORK is 2

# \$TMP = local scratch during batch job

- Local file space on login and compute nodes
  - not for multinode jobs writing to disk
    - Fat nodes: 4 TB (bwUniCluster), 8 TB (ForHLR)
    - Thin nodes: 2 TB
- No backup
- Lifetime = batch job walltime
- Usage:
  - On login nodes: use for pre- and postprocessing (e.g. compilation)
  - Jobs on compute nodes:
    - Copy your data to \$TMP and results from \$TMP to \$HOME/\$WORK



# TSM Backup + Archive

- Automatically TSM Backup of \$HOME
- Archive-Pool for HPC-Cluster (KIT users only)
- Commands:
  - Backup: `tsm_q_backup`, `tsm_restore`
  - Archive: `tsm_archiv`, `tsm_d_archiv`, `tsm_q_archiv`, `tsm_retrieve`
  - Common: `tsm_q_fi`, `tsm_q_fi.ba`, `tsm_q_fi.ar`
  - Help: `tsm_<command> -h`

## Example:

```
$ tsm_q_backup
IBM Tivoli Storage Manager
...
Size          Backup Date      Mgmt Class   A/I File
-----
7 B  24.09.2013 04:20:00  DEFAULT     A  /pfs/data1/home/kit/scc/ab1234/blubb
```

# rdata

- Access to other HPC-Filesystems
- Executes the filesystem operations on "data mover" nodes
- Simplifies data transfer and distributes load
- Available Filesystems:  $\$HOME$ ,  $\$WORK$  of HPC-Clusters, bwFileStorage, ForHLR I Project
- Supported commands: ls, cp, rsync, ...
- More Information: man rdata

Example:

```
$ rdata ls $BWFILESTORAGE/blubb  
/bwfilestorage/ka/ka_scc/ab1234/blubb
```

# Environment modules

- Default usage of compilers, libraries and software packages  
→ set up manually their session environment
- Environment modules software → dynamic modification of the session environment
  - instructions stored in *modulefiles*
- Why?
  - multiple versions of the same software can be installed and be used in a controlled manner, i.e., by *load* and *unload* modulefiles
- How to use modulefiles in general?  

```
$ module help
```
- More information:
  - [http://www.bwhpc-c5.de/wiki/index.php/Environment\\_Modules](http://www.bwhpc-c5.de/wiki/index.php/Environment_Modules)

# modulefiles: available / search

## ■ Display all modulefiles

```
$ module avail
```

```
----- /opt/bwhpc/kit/modulefiles -----
cae/abaqus/6.13-5 cae/ansys/15.0 cae/comsol/4.4 system/d-default
cae/adina/9.0 cae/ansys/15.0.7 cae/starccm+/9.4

----- /opt/bwhpc/common/modulefiles -----
bio/bismark/0.10.1 lib/boost/1.55.0
bio/bowtie/1.0.1 lib/matplotlib/1.3.1
bio/bowtie2/2.1.0 lib/netcdf/3.6.3-gnu-4.8
bio/bowtie2/2.2.3 lib/netcdf/3.6.3-intel-13.1
bio/cufflinks/2.2.0 lib/pnetcdf/1.4.1
bio/qiime/1.8.0 math/R/3.0.2
bio/samtools/0.1.19 math/matlab/R2013a
bio/tophat/2.0.11 math/matlab/R2013b
bio/trimmomatic/0.32 math/matlab/R2014a
cae/ansys/15.0.7_bw mpi/impi/4.1.0-gnu-4.4
cae/ansys/15.0_bw mpi/impi/4.1.0-gnu-4.5
cae/openfoam/1.6-ext mpi/impi/4.1.0-intel-12.1
```

## ■ Display all modulefiles with prefix „compiler“

```
$ module avail compiler
```

```
----- /opt/bwhpc/common/modulefiles -----
compiler/gnu/4.5 compiler/gnu/4.8 compiler/intel/12.1
compiler/gnu/4.7(default) compiler/gnu/4.9 compiler/intel/13.1(default)
```

# modulefiles: help / whatis

- Show help of modulefiles, e.g. `$ module help compiler/intel`

```
----- Module Specific Help for 'compiler/intel/13.1' -----  
  
This module provides the Intel(R) compiler suite version 13.1.3 via  
commands 'icc', 'icpc' and 'ifort', the debugger 'idb' as well as the Intel(R)  
Threading Building Blocks TBB and the Integrated Performance Primitives IPP  
libraries (for details see also 'http://software.intel.com/en-us/intel-compilers/').  
  
The related Math Kernel Library MKL module is 'numlib/mkl/11.0.5'.  
The related Intel MPI module is 'mpi/impi/4.1.1-intel-13.1'.  
The Intel icpc should work well with GNU compiler 4.7.  
  
Commands:  
icc          # Intel(R) C compiler  
icpc         # Intel(R) C++ compiler  
ifort        # Intel(R) Fortran compiler  
idb          # Intel(R) debugger in GUI mode  
idbc         # Intel(R) debugger in console mode  
  
Local documentation:  
Man pages: man icc; man icpc; man ifort  
firefox $INTEL_DOC_DIR/documentation_c.htm  
firefox $INTEL_DOC_DIR/documentation_f.htm
```

- Show short info modulefile

```
$ module whatis compiler/intel
```

```
compiler/intel      : Intel(R) compiler suite (icc, icpc, ifort), debugger (idb), IPP and TBB ver 13.1.3
```

# modulefiles: show

- Show all instructions of modulefile `$ module show compiler/intel`

```
/opt/bwhpc/common/modulefiles/compiler/gnu/4.7:
```

```
module-whatis  GNU compiler suite version 4.7.3 (gcc, g++, gfortran)
setenv        GNU_VERSION 4.7.3
setenv        GNU_HOME /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64
setenv        GNU_BIN_DIR /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/bin
setenv        GNU_MAN_DIR /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/share/man
setenv        GNU_LIB_DIR /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/lib64
prepend-path  PATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/bin
prepend-path  MANPATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/share/man
prepend-path  LD_RUN_PATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/lib
prepend-path  LD_LIBRARY_PATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/lib
prepend-path  LD_RUN_PATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/lib64
prepend-path  LD_LIBRARY_PATH /opt/bwhpc/common/compiler/gnu/4.7.3/x86_64/lib64
setenv        CC gcc
setenv        CXX g++
setenv        F77 gfortran
setenv        FC gfortran
setenv        F90 gfortran
setenv        TEST_MODULE_SCRIPT /opt/bwhpc/common/compiler/gnu/4.7.3/install-doc/test-compiler-gnu.sh
setenv        TEST_MODULE_NAME compiler/gnu/4.7
conflict      compiler/gnu
```

-----

## Load modulefiles (3)

- Modulefiles are sorted in categories, software name and versions:

```
$ module load <category>/<software_name>/<version>
```



- Load a default software:

```
$ module load <category>/<software_name>
```

- e.g. Intel compiler

```
$ module load compiler/intel mpi/impi
```


→ loads currently Intel compiler suite 13.1

→ loads currently Intel-MPI 4.1.1 for Intel compiler 13.1

```
$ module list
```

- Display all loaded modules

```
Currently Loaded Modulefiles:  
 1) compiler/intel/13.1(default) 2) mpi/impi/4.1.1-intel-13.1(default)
```



# modulefiles: categories & dependencies

- Module names already implicate dependencies:

→ **Category/softwarename/version\_attributes-dependencies**

e.g. **numlib/fftw/3.3.3-impi-4.1.1-intel-13.1**

→ fftw package version 3.3.3, compiled with Intel 13.1 and Intel-MPI 4.1.1

→ attributes could be: single, double precision etc.

- Categories:

compiler/	for compiler, e.g. intel, gnu, pgi, open64
devel/	for debugger, e.g. ddt, and development tools, e.g. cmake, itrac
mpi/	for MPI libraries, e.g. impi, openmpi, mvapich(2)
numlib/	for numerical libraries, e.g. Intel MKL, ACML, nag, gsl, fftw
lib/	for other libraries, e.g. netcdf, global array
bio/	for biology software, e.g. bowtie, abyss, mrbayes
cae/	for CAE software, e.g. ansys, abaqus, fluent
chem/	for chemistry software, e.g. gromacs, dacapo, turbomole
math/	for mathematics software, e.g. matlab, R
phys/	for physics software, e.g. geant4
vis/	for visualisation software, e.g. vmd, tigervnc





# modulefiles: conflicts

## ■ Conflicts:

- a) load different software version in the same session, e.g. Intel:

```
$ module load compiler/intel/12.1  
$ module load compiler/intel/13.1
```

```
compiler/intel/13.1(394):ERROR:150: Module 'compiler/intel/13.1' conflicts  
with the currently loaded module(s) 'compiler/intel/12.1'
```

- b) load module with dependencies on other modules

```
$ module load mpi/openmpi/1.6.5-intel-13.1
```

```
Loading module dependency 'compiler/intel/13.1'.  
compiler/intel/13.1(394):ERROR:150: Module 'compiler/intel/13.1' conflicts  
with the currently loaded module(s) 'compiler/intel/12.1'
```

# modulefiles: unload/swap

- To remove module *foo*:

```
$ module unload foo
```

```
$ module remove foo
```

be aware that you might create **inconsistencies**,

e.g. you can remove

*compiler/intel/13.1* while *mpi/openmpi/1.6.5-intel-13.1* is still loaded

- Swap = remove + load

e.g.:

```
$ module swap compiler/intel/12.1 compiler/intel/13.1
```

## ADV: Private modulefiles

- Each user can create own modulefiles:

e.g. modulefiles that adds path of own programs, `$HOME/special`, to `$PATH`

→ content of this modulefile „*mybin*“

```
#%Module1.0  
  
Append-path    PATH    "$env(HOME)/special"
```

→ place „*mybin*“ under `$HOME/privatemodules`

→ to make all own modules visible to “module avail” command, enter:

```
$ module load use.own    or    $ module use $HOME/privatemodules
```

→ former: own modules have lower priority than system ones if equally named

→ latter: own module have higher priority

- Remove own modules:

```
$ module unload use.own or $ module unuse $HOME/privatemodules
```

# Agenda

Time	Title
09:30	bwHPC: Concept and User Support
09:45	Architecture: bwUniCluster + ForHLR
10:00	Cluster: Access, Modulefiles, Filesystem
10:30	Cluster: Batch System
11:00	Break
11:10	Tutorials: Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting
Max. 13:10	End



bw|HPC – C5

# bwHPC Cluster: Batch System

Simon Raffener



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



**KIT**  
Karlsruher Institut für Technologie



ulm university universität  
**uulm**

Funding:



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

[www.bwhpc-c5.de](http://www.bwhpc-c5.de)

# Resource and workload manager

## ■ bwUniCluster:

- compute job will be **only** processed by the batch system  
→ **define resources** of your compute job (sequence of commands & programs) **in advance**

## ■ How?

- via workload manager = MOAB  
→ will be also running on all forthcoming bwForCluster
  1. Setup your compute job script
  2. Submit your compute job to workload manager

```
$ msub <resource_options> <job_script>
```

## ■ Implications:

- fairshare based queue
- waiting time depends on: **your university's share, your job demands, your demand history**

# msub options

■ [http://www.bwhpc-c5.de/wiki/index.php/Batch\\_Jobs#msub\\_Command](http://www.bwhpc-c5.de/wiki/index.php/Batch_Jobs#msub_Command)

■ msub options: command line or in your job script

msub Options		
Command line	Script	Purpose
-l <i>resources</i>	#MSUB -l <i>resources</i>	Defines the resources that are required by the job. See the description below for this important flag.
-N <i>name</i>	#MSUB -N <i>name</i>	Gives a user specified name to the job.
-o <i>filename</i>	#MSUB -o <i>filename</i>	Defines the filename to be used for the standard output stream of the batch job. By default the file with defined filename is placed under your job submit directory. To place under a different location, expand <i>filename</i> by the relative or absolute path of destination.
-q <i>queue</i>	#MSUB -q <i>queue</i>	Defines the queue class
-v <i>variable=arg</i>	#MSUB -v <i>variable=arg</i>	Expands the list of environment variables that are exported to the job
-S <i>Shell</i>	#MSUB -S <i>Shell</i>	Declares the shell (state path+name, e.g. /bin/bash) that interprets the job script

→ command line option overwrites script option

# `msub -l resource_list`

[http://www.bwhpc-c5.de/wiki/index.php/Batch\\_Jobs#msub\\_-l\\_resource\\_list](http://www.bwhpc-c5.de/wiki/index.php/Batch_Jobs#msub_-l_resource_list)

<code>msub -l resource_list</code>	
resource	Purpose
<code>-l nodes=2:ppn=8</code>	Number of nodes and number of processes per node
<code>-l walltime=600</code>	Wall-clock time. Default units are seconds.
<code>-l walltime=01:30:00</code>	HH:MM:SS format is also accepted.
<code>-l pmem=1000mb</code>	Maximum amount of physical memory used by any single process of the job. Allowed units are kb, mb, gb. Be aware that <b>processes</b> are either <i>MPI tasks</i> if running MPI parallel jobs or <i>threads</i> if running multithreaded jobs.
<code>-l mem=1000mb</code>	Maximum amount of physical memory used by the job. Allowed units are kb, mb, gb. Be aware that this memory value is the accumulated memory for all <i>MPI tasks</i> or all <i>threads</i> of the job.
<code>-l advres=res_name</code>	Specifies the reservation "res_name" required to run the job.
<code>-l naccesspolicy=policy</code>	Specifies how node resources should be accessed, e.g. <code>-l naccesspolicy=singlejob</code> reserves all requested nodes for the job exclusively. Attention, if you request <code>nodes=1:ppn=4</code> together with <code>singlejob</code> you will be charged for the maximum cores of the node.

→ for workshop: `-l advres=workshop.6`

→ resource can combined, but must be separated by comma, e.g.:

```
$ msub -l nodes=1:ppn=1,walltime=00:01:00,pmem=1gb <job_script>
```



# `msub -q queues`

■ [http://www.bwhpc-c5.de/wiki/index.php/Batch\\_Jobs\\_-\\_bwUniCluster\\_Features#msub\\_-q\\_queues](http://www.bwhpc-c5.de/wiki/index.php/Batch_Jobs_-_bwUniCluster_Features#msub_-q_queues)

<code>msub -q queue</code>				
<i>queue</i>	<i>node</i>	<i>default resources</i>	<i>minimum resources</i>	<i>maximum resources</i>
develop	thin	<i>walltime=00:10:00,procs=1, mem=4gb</i>	<i>nodes=1</i>	<i>walltime=00:30:00,nodes=1:ppn=16</i>
singlenode	thin	<i>walltime=00:30:01,procs=1, mem=4gb</i>	<i>walltime=00:30:01,nodes=1</i>	<i>walltime=3:00:00:00,nodes=1:ppn=16</i>
multinode	thin	<i>walltime=00:10:00,procs=1, mem=4gb</i>	<i>nodes=2</i>	<i>walltime=2:00:00:00,nodes=16:ppn=16</i>
verylong	thin	<i>walltime=3:00:00:01,procs=1, mem=4gb</i>	<i>walltime=3:00:00:01,nodes=1</i>	<i>walltime=6:00:00:00,nodes=1:ppn=16</i>
fat	fat	<i>walltime=00:10:00,procs=1, mem=32gb</i>	<i>nodes=1</i>	<i>walltime=3:00:00:00,nodes=1:ppn=32</i>

■ If queues not specified:

- jobs assigned to „develop“, „singlenode“ and „multinode“ based on your requested walltime, nodes and processes.

# Environment variables

- once eligible, MOAB adds the following variables to the job's environment

Environment variables	Description
MOAB_CLASS	Class name
MOAB_GROUP	Group name
MOAB_JOBID	Job ID
MOAB_JOBNAME	Job name
MOAB_NODECOUNT	Number of nodes allocated to job
MOAB_PARTITION	Partition name the job is running in
MOAB_PROCCOUNT	Number of processors allocated to job
MOAB_SUBMITDIR	Directory of job submission
MOAB_USER	User name

→ can be used to generalize you job scripts, e.g.:

```
## add suffix to job output file  
./program > $program_${MOAB_JOBID}.log
```

# Check/change status of your jobs

- after submission → msub returns <job-ID>

```
$ msub job.sh
```

```
659562
```

- commands:**

\$ showq	All your active, eligible, blocked, and/or recently completed jobs
\$ showstart <job-ID>	Get information about start time of job with <job-ID>
\$ showstart 16@12:00:00	Get information about start time of 16 procs with run time of 12 hours
\$ checkjob <job-ID>	Get detailed information of your job → explains why your job is pending
\$ showq -c	Display completed job
\$ canceljob <job-ID>	Cancel the job with <job-ID>

# Example

```
#!/bin/bash
#MSUB -l nodes=2:ppn=16
#MSUB -l walltime=01:00:00
#MSUB -l pmem=2gb
#MSUB -N serial-test

mpirun ./hello
```

→ Is equal to:

```
$ msub -l nodes=2:ppn=16,walltime=01:00:00,pmem=2gb -N serial-test
<job_script>
```

# Common problems

- Wrong „ppn“ setting:

```
$ msub -l nodes=3:ppn=38,walltime=00:01:00,pmem=1gb <job_script>
```

- „mem“ instead of „pmem“:

```
$ msub -l nodes=4:ppn=16,walltime=00:01:00,mem=1gb <job_script>
```

- Wrong queue

- Data in \$HOME instead of \$WORK

- `# MSUB` instead of `#MSUB` (note the space...)

# Agenda

Time	Title
09:30	bwHPC: Concept and User Support
09:45	Architecture: bwUniCluster + ForHLR
10:00	Cluster: Access, Modulefiles, Filesystem
10:30	Cluster: Batch System
11:00	Break
11:10	Tutorials: Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting
Max. 13:10	End



bw|HPC – C5

# bwUniCluster Tutorial

Access, Data Transfer, Compiling, Modulefiles, Batch Job Scripting

Annika Fuchs



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



**KIT**  
Karlsruher Institut für Technologie



ulm university universität  
**uulm**



# Login

- Username <username>
    - Same username as your user account at university.
    - Users from other universities than KIT have to prefix their username by the organization's token, e.g. ho\_anfuchs
  - Host <host>
    - bwUniCluster: `uc1.scc.kit.edu`
- 
- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>■ Linux / OS X<ul style="list-style-type: none"><li>■ open terminal:<br/>&gt; <code>ssh &lt;username&gt;@&lt;host&gt;</code></li></ul></li></ul> | <ul style="list-style-type: none"><li>■ Windows<ul style="list-style-type: none"><li>■ use SSH-Client, e.g. PuTTY</li><li>■ connect to &lt;host&gt;:<br/>&gt; <b>Login as:</b> <code>&lt;username&gt;</code></li></ul></li></ul> |
|--|--|



# Basic commands

<code>\$ pwd</code>	show path of working directory
<code>\$ mkdir &lt;dirname&gt;</code>	make directory
<code>\$ cp &lt;sourcefile&gt; &lt;targetfile&gt;</code>	copy file
<code>\$ mv &lt;sourcefile&gt; &lt;targetfile&gt;</code>	move file
<code>\$ rm &lt;filename&gt;</code>	remove file
<code>\$ man &lt;command&gt;</code>	show command's manual

# Data Transfer

- From localhost to cluster:
  - use `scp` (secure copy) or `sftp` (secure file transfer program)
  - Read manual for options/syntax questions (`man scp`, `man sftp`)
- Linux / OS X
  - Open terminal at your computer:
    - `$ scp <sourcefile> <username>@<host>:<targetfile>`
    - or
    - `$ sftp <username>@<host>:<targetdir>`
    - `$ put <sourcefile>`
- Windows
  - use SCP/SFTP-Client, e.g. WinSCP
  - connect to `<username>@<host>`
  - copy data by drag&drop mechanism

# Module Environment

- Users require different software in different versions.
- Software is installed and can be used by loading corresponding modules.

> <b>module avail</b>	show all installed software packages
> <b>module avail compiler</b>	show all available compilers
> <b>module load &lt;modulepath&gt;</b>	load a module in list
> <b>module unload &lt;modulepath&gt;</b>	remove a module from list
> <b>module list</b>	show all loaded modules
> <b>module show &lt;modulepath&gt;</b>	show environment variables of module
> <b>module help &lt;modulepath&gt;</b>	show usage information of module

# From \$HOME to \$WORK

- Compute nodes read&write in \$WORK very much faster than in \$HOME directory.
- **DO NOT COMPUTE IN \$HOME !!**
- \$HOME:
  - Source code
- \$WORK:
  - Program input (e.g. initial and boundary conditions)
  - Program output

If lifetime of \$WORK is too short, create a workspace.  
But **never** compute in \$HOME!

# Exercise

- Download source code from indigo
- Copy source code to bwUniCluster
- Log on bwUniCluster
- Load module file corresponding to the compiler of choice
- Compile the source code, e.g. sequential version with Intel-Compiler:  

```
$ icc hello.c -o hello
```
- Move your binary in \$WORK

Source code is written in C and Fortran90 and provided in a sequential version or with OpenMP, MPI or hybrid parallelization.

# Submitting jobs via script

- Example: requesting one CPU and 3000 MB of main memory for 5 hours to run the sequential program `hello`

```
#!/bin/bash
-----
#MSUB -l nodes=1:ppn=1
#MSUB -l walltime=5:00:00
#MSUB -l mem=3000mb
#MSUB -q singlenode
#MSUB -N serial-test
#MSUB -m abe
-----
```

Interpreter

Header with msub options

- resource requirements
- queue definition
- notification options,...

```
./hello
```

Execution part

- Submitting the script `jobuc.sh` with MOAB:  
> `msub jobuc.sh`

# Environment variables in job scripts

## ■ MOAB variables and own environment variables

	Using MOAB Variables	Defining own variables
Header	<code>#MSUB -o \$(JOBNAME).o\$(JOBID)</code>	<code>#MSUB -v EXEC=./hello</code>
Execution Part	<code>echo „Job \${MOAB_JOBNAME} is running (ID=\${MOAB_JOBID})“</code>	<code>export EXEC=./hello</code>

# Parallel Jobs (MPI)

```
#!/bin/bash
#MSUB -l nodes=2:ppn=4
#MSUB -l walltime=05:00
#MSUB -l pmem=1000mb
#MSUB -q multinode
#MSUB -v EXECUTABLE=./hello_mpi
#MSUB -N hello_mpi
#MSUB -o $(JOBNAME).o$(JOBID)

module load mpi/openmpi

echo "Executable ${EXECUTABLE}
running on ${MOAB_PROCCOUNT}
cores"

mpirun ${EXECUTABLE}
```

■ For computations on more than 1 node use `queue multinode`.

■ The corresponding MPI module has to be loaded on the compute nodes.

■ Use `mpirun` to execute the binary.



# Parallel Jobs (OpenMP)

```
#!/bin/bash
#MSUB -l nodes=1:ppn=8
#MSUB -l walltime=05:00:00
#MSUB -l pmem=1000mb
#MSUB -q singlenode
#MSUB -v EXECUTABLE=./hello_omp
#MSUB -N hello_omp
#MSUB -o $(JOBNAME).o$(JOBID)

export OMP_NUM_THREADS=${MOAB_PROCCOUNT}

echo "Executable ${EXECUTABLE} running
with ${OMP_NUM_THREADS} threads"

${EXECUTABLE}
```

■ Shared memory restricts to 1 node.

■ Do not define number of threads explicitly. Use MOAB variables.

# Hybrid Parallel Jobs (OpenMP+MPI)

```
#!/bin/bash
#MSUB -l nodes=2:ppn=4
#MSUB -l walltime=05:00
#MSUB -l pmem=1000mb
#MSUB -q multinode
#MSUB -v EXE=./hello_mpi_omp
#MSUB -v OMP_NUM_THREADS=4
#MSUB -N hello_mpi_omp
#MSUB -o $(JOBNAME).o$(JOBID)

module load mpi/openmpi

export NTASKS=$(( ${MOAB_PROCCOUNT} / ${OMP_NUM_THREADS} ))

echo "Executable ${EXE} running on ${MOAB_PROCCOUNT} cores with ${
NTASKS} tasks and ${OMP_NUM_THREADS} threads"

mpirun -n ${NTASKS} -bynode -cpus-per-proc ${OMP_NUM_THREADS} ${EXE}
```

- Explicit declaration of task and thread numbers is required

# Keep track of a job

- Submit job script

```
$ msub <jobscript>
```

- If a job (script) is accepted the <jobid> appears at screen.

```
$ checkjob <jobid>    show job details
$ showq               list all my running, idling and blocked jobs by <jobid>
$ showq -n            list all my running, idling and blocked jobs by <jobname>
$ showq -c            list my completed jobs
$ canceljob <jobid>   cancel job
```

# GUI via X-Tunnel

Compute at bwUniCluster but display GUI „at home“ (localhost)

## ■ Modified login

### ■ Linux / OS X

```
$ ssh -X <username>@<host>
```

### ■ Windows

- Start X server, e.g. Xming

- PuTTY Configuration: Category SSH > X11

- Check box „Enable X11 forwarding“

## ■ Submit interactive job (only bwUniCluster)

```
$ msub -I -V -l nodes=1:ppn=1,walltime=02:00:00,mem=4000mb
```

## ■ Start program, e.g. Matlab:

```
$ module load math/matlab
```

```
$ matlab
```

# GUI via VNC (Virtual Network Computing)

- Log on bwUniCluster via terminal/PuTTY

- Submit interactive job

```
$ msub -I -V -l nodes=1:ppn=1,walltime=02:00:00,mem=4000mb
```

- Start VNC server

```
$ module load vis/tigervnc
```

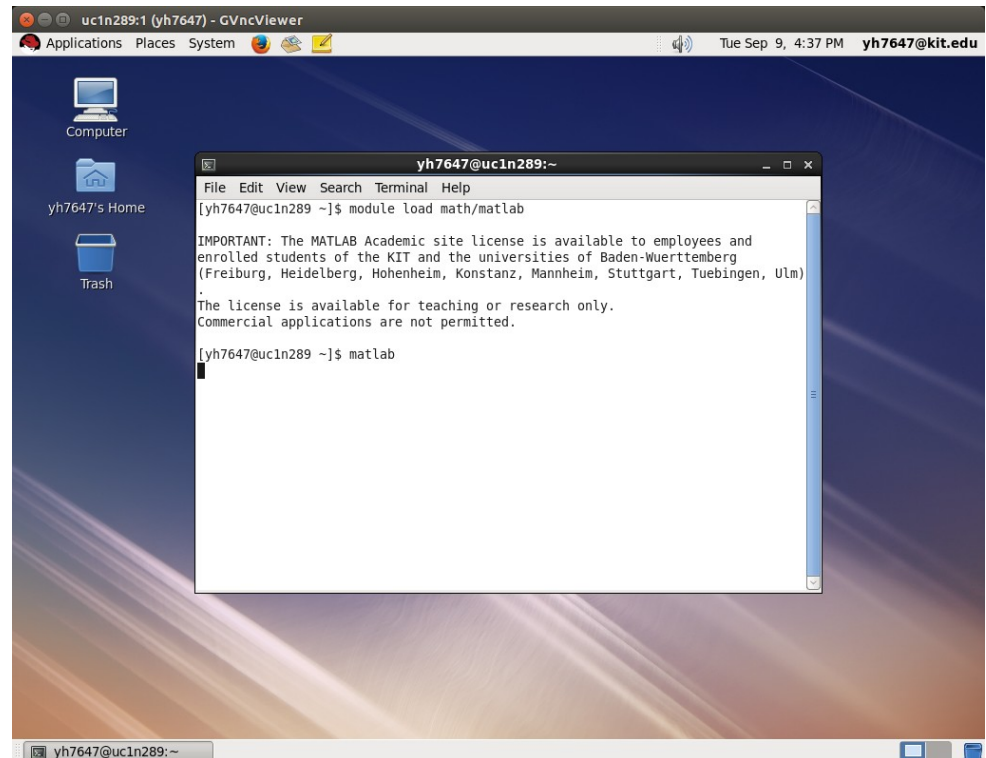
```
$ run_vncserver
```

- Set initial VNC password.

- Follow displayed instructions.

- Start VNC client at localhost

- TightVNC Java Viewer is recommended for Windows users since an SSH client is included.





bw|HPC – C5

# Tutorial: Advanced Topics

Robert Barthel



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386

Hochschule  
für Technik  
Stuttgart



**Hochschule Esslingen**  
University of Applied Sciences

Universität  
Konstanz



UNIVERSITÄT  
MANNHEIM



Universität Stuttgart

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



**KIT**  
Karlsruher Institut für Technologie



ulm university universität  
**uulm**

Funding:



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

[www.bwhpc-c5.de](http://www.bwhpc-c5.de)

# Jobs @ \$TMP (1)

- If temporary files of job > Gbyte → Run your job at \$TMP
  - but ONLY if single node jobs
- What to do:
  - Generate subdirectory under \$TMP = \$run\_DIR
  - Copy to \$run\_DIR
  - Change to \$run\_DIR & program execution
  - Copy results to start DIR
- How?
  - Check: job\_run\_under\_local\_tmp.sh with do\_files & do\_files.inp

## Jobs @ \$TMP (2)

### ■ Code snip of job\_run\_under\_local\_tmp.sh

```
. . .
## b) Define your run directory and generate it
run_DIR="${TMP}/${USER}.${MOAB_JOBID:-$$}"
mkdir -pv "${run_DIR}"

## c) check existance of run_DIR
if [ ! -d "${run_DIR}" ] ; then
    echo "ERROR: Run DIR = ${run_DIR} does not exist"; exit 1
fi

## d) copy files from submit_DIR to run_DIR
cd $MOAB_SUBMITDIR
cp -pv do_files do_files.inp "${run_DIR}"

## e) cd to run_DIR and start "binary" together with input file
cd ${run_DIR}
./do_files do_files.inp

## f) check run status
if [ $? -ne 0 ] ; then
    echo "WARNING: do_files did not run properly"
fi

## g) transfer files to submit directory
cp -pv files_*.out "${MOAB_SUBMITDIR}"
. . .
```





# Chain Jobs (1)

- Idea:
  - Split  $N$  consecutive Jobs into  $N$  MOAB Batch Jobs
- Goal:
  - Do everything in one script
  - Submit only once to MOAB
- „Pre-step“: generate script that runs interactively
  - `/opt/bwhpc/kit/workshop/2014-09-10/scripts/interactive_chain_job.sh`

## Chain Jobs (2)

```
#!/bin/bash
## Defaults
loop_max=10
cmd='sleep 2'

## Check if counter environment variable is set
if [ -z "${myloop_counter}" ] ; then
    echo " ERROR: myloop_counter is undefined, stop chain job"
    exit 1
fi
## only continue if below loop_max
if [ ${myloop_counter} -lt ${loop_max} ] ; then
    ## increase counter
    let myloop_counter+=1
    ## print current Job number
    echo " Chain job iteration = ${myloop_counter}"
    ## Define your command
    cmd='sleep 2'
    echo " -> executing ${cmd}"
    ${cmd}

    if [ $? -eq 0 ] ; then
        ## continue only if last command was successful
        export myloop_counter=${myloop_counter}
        ./${0}
    else
        ## Terminate chain
        echo " ERROR: ${cmd} of chain job no. ${myloop_counter} terminated unexpectedly"
        exit 1
    fi
fi
```

```
$ export myloop_counter=0
$ ./interactive_chain_job
```



## Chain Jobs (3) → for Moab

```
#!/bin/bash
#MSUB -l nodes=1:ppn=1,walltime=00:00:05,pmem=50mb
## Defaults
loop_max=10
cmd='sleep 2'

## Check if counter environment variable is set
if [ -z "${myloop_counter}" ] ; then
    echo " ERROR: myloop_counter is undefined, stop chain job"
    exit 1
fi
## only continue if below loop_max
if [ ${myloop_counter} -lt ${loop_max} ] ; then
    ## increase counter
    let myloop_counter+=1
    ## print current Job number
    echo " Chain job iteration = ${myloop_counter}"
    ## Define your command
    cmd='sleep 2'
    echo " -> executing ${cmd}"
    ${cmd}

    if [ $? -eq 0 ] ; then
        ## continue only if last command was successful
        msub -v myloop_counter=${myloop_counter} ./moab_chain_job.sh
    else
        ## Terminate chain
        echo " ERROR: ${cmd} of chain job no. ${myloop_counter} terminated unexpectedly"
        exit 1
    fi
fi
```

```
$ msub -v myloop_counter=0 ./moab_chain_job.sh
```



# Chain Jobs (4)

- moab\_chain\_job.sh + interactive\_chain\_job.sh =  
generalised\_chain\_job.sh

```
. . .
. . .
if [ $? -eq 0 ] ; then
  ## continue only if last command was successful
  if [ ! -z MOAB_JOBNAME ] ; then
    ## If MOAB_JOBNAME environment variable is defined
    ## -> this script is under MOAB "control"
    msub -v myloop_counter=${myloop_counter} ./generalised_chain_job.sh
  else
    export myloop_counter=${myloop_counter}
    ./${0}
  fi
else
  ## Terminate chain
  echo "  ERROR: ${cmd} of chain job no. ${myloop_counter} terminated unexpectedly"
  exit 1
fi
. . .
. . .
```

→ USE bash programming to **generalise** and **unify** your batch job scripts

# Batch jobs with input parsing

## ■ Not working:

- `msub [options] your_script -x argument`  
→ `msub` will interpret `-x` as an own option

## ■ Solution:

(A) Submit wrapper script:

```
#!/bin/bash  
your_script -x argument
```

(B) Generalise your script → see Solution B under

[http://www.bwhpc-c5.de/wiki/index.php/Batch\\_Jobs#Handling\\_job\\_script\\_options\\_and\\_arguments](http://www.bwhpc-c5.de/wiki/index.php/Batch_Jobs#Handling_job_script_options_and_arguments)

(C) Use `msub` wrapper via:

```
$ module load system/msub_addon/1.0  
$ msub <options> job.sh
```

## Common problems (2)

- Manual defining of MPI tasks for mpirun

- Wrong:

- `mpirun -machinefile=file binary`

- `mpirun -n <int> binary`

- Correct, (because the resource manager tells mpirun what to do):

- `mpirun binary`

- If you want to know about job allocated hosts in your script to:

(A) Use msub wrapper via:

```
$ module load system/msub_addon/1.0  
$ msub <options> job.sh
```

(B) Write loop into your batch job script → returns hostname of each task:

```
for tasks in $(srun hostname) ; do  
    echo $tasks  
done
```

Thank you for your attention!