bw|HPC

File Systems and Software Modules

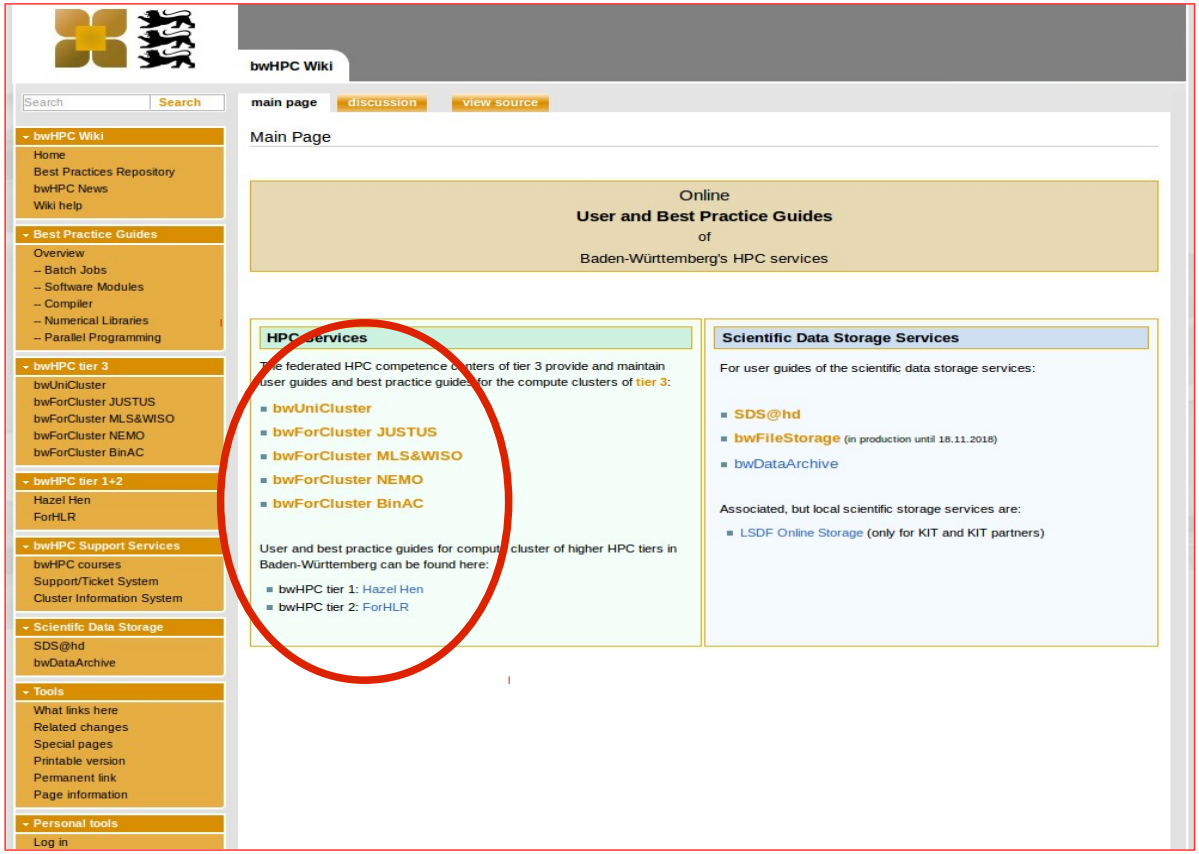
Simon Raffeiner, Scientific Computing and Simulation Dept., SCC, KIT

Universität Heidelberg Zukunft seit 1386
Hochschule für Technik Stuttgart
Uni Freiburg
Universität Hohenheim 1818
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

Funding: Baden-Württemberg Ministerium für Wissenschaft, Forschung und Kunst

Reference: bwHPC Wiki

- Most information given by this talk can be found at http://bwhpc.de/wiki, select cluster, then select
  - File Systems
  - Software Modules



Material: Slides & Scripts

- https://indico.scc.kit.edu/indico/event/487/
- bwUniCluster:

  /pfs/data1/software_uc1/bwhpc/kit/workshop/2019-04-09/

How to read the following slides

Abbreviation/Colour code	Full meaning
`$ command -option value`	$ = prompt of the interactive shell The full prompt may look like: `user@machine:path$` The command has been entered in the interactive shell session
`<integer>` `<string>`	<> = Placeholder for integer, string etc
`foo, bar`	Metasyntactic variables

File Systems

File System list and properties

- bwUniCluster, ForHLR I / ForHLR II, bwForCluster
  - Please see documentation to get complete list and details of all file systems
- Common „rules of thumb“
  - There is a small, global, permanent $HOME directory
    → Sometimes backed up, sometimes not
  - There are large, global, temporary workspace directories
    → Old files might be deleted automatically / directories might expire
    → Usually NOT backed up
  - There is local, non-global, temporary storage on the compute nodes
    → Usually named $TMP, $TMPDIR, $SCRATCH or something similar
    → Will be wiped when the job ends
  - There might be external, permanent storage
    → Usability and costs differ depending on your organization
    → Examples: bwDataArchive, SDS@hd, LSDF Online Storage

How to use each File System

- $HOME
  - → Software, configuration files, final results
  - → Omit heavy I/O
- $WORK / workspace
  - → Intermediate results, huge input/output data sets
  - → Scratch data which needs to be shared between nodes
  - → Omit small files, tiny block sizes, lots of metadata operations
- $TMP, $TMPDIR
  - → Data which is only needed during job runtime on the local node
  - → Possibly transfer data here within a batch job
  - → All sorts of I/O allowed
- External storage
  - → Archive scientific data, move data here when data sets become too large
  - → Use huge files or compressed archives, transfer with *rdata* on KIT clusters

$HOME = Home directory

- Properties of $HOME on different clusters

Cluster	Quota limits	Backup
JUSTUS	100 GB per user	Yes
MLS&WISO	100 GB per user	No
NEMO	100 GB per user	Yes
BinAC	20 GB per user	No
bwUniCluster	x TB per dep./org. group	Yes
ForHLR I / II	2 GB per project group	Yes

$HOME on bwUniCluster

- HowTo goto:

```
$ cd $HOME
```

```
$ cd
```

- User's quota usage:

```
$ lfs quota -uh $(whoami) $HOME
```

- Group's quota limit:

```
$ lfs quota -gh $(id -ng) $HOME
```

$HOME on ForHLR I/II

- ONLY ForHLR I/II:
  - $HOME is now equal to $PROJECT
    → assigned „name/acronym“, e.g. /project/fh1-project-scs/mr2515
    → /project/fh1-project-scs/mr2515/ is equal to /home/fh1-project-scs/mr2515
    → $PROJECT_GROUP, e.g. „fh1-project-scs“
  - Quota of Project:

```
$ lfs quota -gh ${PROJECT_GROUP} ${PROJECT}
Disk quotas for grp fh1-project-scs (gid 900000):
     Filesystem    used   quota   limit   grace   files   quota   limit   grace
/home/fh1-project-scs/mr2515
                  574.2G      0k  1.863T       - 2293266       0 5000000       -
```

Workspaces = Working directory with lifetime

- **Workspaces**: lifetime on allocated folder
  - Available on all clusters, maximum lifetime is slightly different
  - HowTo:

    → http://www.bwhpc.de/wiki/index.php/Workspace

`$ ws_allocate $WS 10`	Allocate workspace `$WS` for 10 days
`$ ws_list -a`	List all your workspaces
`$ ws_find $WS`	Get absolute path of workspace `$WS`
`$ ws_extend $WS 5`	Extend lifetime of your workspace `$WS` by 5 days from now. Number of extensions depends on cluster. → max. lifetime = 4*60 days (b+F), x*90 (J+M), 4*30 (B), x*100 (N)
`$ ws_release $WS`	Manually erase your workspace `$WS`

Example:

```
$ ws_allocate scratch 60
Workspace created. Duration is 1440 hours.
Further extensions available: 3
/work/workspace/scratch/mr2515-scratch-0
```

I/O statistics @ bwUniCluster/ForHLR

- Steps to get a job's I/O statistics on $HOME, $WORK or workspace:
  - Get file system name:

```
$ df <directory> | sed -ne "s|.*o2ib:/\([a-z,0-9]*\).*|\1|p"
```

  - Request I/O statistics during job submission on command line:

```
$ msub … -W lustrestats:<file system name>[,<file system name>] …
```

  - Alternatively request I/O statistics in job script:

```
#MSUB -W lustrestats:<file system name>[,<file system name>]
```

  - Commands above are for bwUniCluster,
    on ForHLR I / II replace the string lustrestats: with x=VAR=LUSTRESTATS=
  - Email example:

```
Subject: Lustre stats of your job 1141 on cluster xyz
Hello,
this is the Lustre IO statistics as requested by user john_doe
on cluster xyz for file system home.
Job 1141 has done ...
... 1 open operations.
... 1 close operations.
... 1 setattr operations.
... 10 write operations and sum of 10,485,760 byte writes (min
IO size: 1048576, average IO size: 1048576, max IO size: 1048576).
```

Exercise 1

- Allocate two workspaces

```
$ $ ws_allocate test 30
Workspace created. Duration is 720 hours.
Further extensions available: 3
/work/workspace/scratch/mr2515-test-0


$ ws_allocate scratch 50
Workspace created. Duration is 1200 hours.
Further extensions available: 3
/work/workspace/scratch/mr2515-scratch-0
```

Exercise 2

- List workspaces

```
$ ws_list
Filesystem: default
Workspace ID     Workspace location                                  Creation date    Remaining time
-------------------------------------------------------------------------------------------------
test                                                                 Apr  9 09:27:18  29 days 23 hours
                 /work/workspace/scratch/mr2515-test-0
                 acctcode:900000
                 available extensions:3
scratch                                                              Apr  9 09:27:56  49 days 23 hours
                 /work/workspace/scratch/mr2515-scratch-0
                 acctcode:900000
                 available extensions:3
```

Exercise 3

- Find workspace path and switch to it

```
$ ws_find scratch
/work/workspace/scratch/mr2515-scratch-0

$ ws_find test
/work/workspace/scratch/mr2515-test-0



$ cd $(ws_find test)
$ pwd
/work/workspace/scratch/mr2515-test-0
```

Exercise 4

- Extend the lifetime of a workspace

```
$ ws_extend test 60
Duration of workspace is successfully changed!
New duration is 1440 hours. Further extensions available: 2
/work/workspace/scratch/mr2515-test-0

$ ws_list
Filesystem: default
Workspace ID     Workspace location                          Creation date    Remaining time
-------------------------------------------------------------------------------------------
test                                                         Apr  9 09:29:08  59 days 23 hours
                 /work/workspace/scratch/mr2515-test-0
                 acctcode:900000
                 available extensions:2
scratch                                                      Apr  9 09:27:56  49 days 23 hours
                 /work/workspace/scratch/mr2515-scratch-0
                 acctcode:900000
                 available extensions:3
```

Exercise 5

- Release workspaces

```
$ ws_release test
/work/workspace/scratch/mr2515-test-0
Info: Workspace was deleted.

$ ws_release scratch
/work/workspace/scratch/mr2515-scratch-0
Info: Workspace was deleted.
```


Software System

Software (=Environment) modules

By default manual setup of $PATH, $LD_LIBRARY_PATH … for

compilers, libraries and software packages etc.

→ Getting complicated if multiple versions of same software installed

Solution:

- dynamic modification of the session environment by

  → instruction sets stored in *modulefiles*

HowTo?

- *load* and *unload* instruction sets (= modulefiles)
- How to use modulefiles in general?

```
$ module help
```

- More information:

  http://www.bwhpc.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 (gdb-ia), IPP and TBB
```

modulefiles: show

- Show all instructions of modulefile `$ module show compiler/gnu/4.7`

```
/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
-------------------------------------------------------------------
```

modulefiles: load (1)

- 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 17

    → loads currently Intel-MPI 2017.4.239 for Intel compiler 17

- Display all loaded modules

```
$ module list
```

```
Currently Loaded Modulefiles:
  1) compiler/intel/17.0(default)    2) mpi/impi/2017-intel-17.0(default)
```

modulefiles: categories & dependencies

- Module names already implicate dependencies:

  → `Category/softwarename/version_attributes-dependencies`

  e.g `numlib/fftw/3.3.5-impi-5.1.3-intel-16.0`

  → fftw package version 3.3.5, compiled with Intel 16.0 and Intel-MPI 5.1.3

- 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: load (2)

- @ bw{Uni,For}Cluster - Conflicts:
  - a) load different software version in the same session, e.g. Intel:

```
$ module load compiler/intel/14.0
$ module load compiler/intel/15.0
```

```
compiler/intel/15.0(390):ERROR:150: Module 'compiler/intel/15.0'
conflicts with the currently loaded module(s) 'compiler/intel/14.0'
```

  - b) load module with dependencies on other modules

```
$ module load mpi/openmpi/1.10-intel-16.0
```

```
Loading module dependency 'compiler/intel/16.0'.
compiler/intel/16.0(398):ERROR:150: Module 'compiler/intel/16.0'
Conflicts with the currently loaded module(s) 'compiler/intel/14.0'
```

→ NOT an issue if the cluster uses Lmod (ForHLR I/II)

→ automatically replaces conflicting modules

modulefiles: list

```
$ module list

Currently Loaded Modules:
  1) dot   2) numlib/mkl/11.3   3) mpi/openmpi/1.8   4) compiler/gnu/5
```

- Depending on the cluster, default modules might be loaded or not

modulefiles: unload/swap (1)

- To remove module *foo*:

  `$ module unload foo` or `$ module remove foo`

  @ bw{Uni,For}Cluster
  - be aware that you might create inconsistencies

    e.g. you can remove

    compiler/intel/16.0 while *mpi/openmpi/1.10-intel-16.0* is still loaded

- Swap = remove + load

  e.g.:

```
$ module swap compiler/intel/14.0 compiler/intel/16.0
```

modulefiles: unload/swap (2)

- @ ForHLR: automatic swap

```
$ module load compiler/gnu/7

Lmod is automatically replacing "compiler/intel/17.0" with "compiler/gnu/7".
```

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
```

Exercise 1

- Check loaded modules after login

```
$ module list

No Modulefiles Currently Loaded.
```

- Check default GNU Compiler version

```
$ gcc --version
gcc (GCC) 4.8.5 20150623 (Red Hat 4.8.5-36)
Copyright (C) 2015 Free Software Foundation, Inc.
This is free software; see the source for copying conditions.  There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
```

Exercise 2

- Load a different GNU compiler version

```
$ module load compiler/gnu/5.2


$ echo $PATH
/opt/bwhpc/common/compiler/gnu/5.2.0/bin:/software/all/bin:/usr/lib64/qt-3.3/
bin:/opt/moab/bin:/usr/local/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/home/kit/
scc/nz2887/.local/bin:/home/kit/scc/nz2887/bin


$ gcc --version
gcc (GCC) 5.2.0
Copyright (C) 2015 Free Software Foundation, Inc.
This is free software; see the source for copying conditions.  There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
```

Exercise 3

- Switch from GNU compiler to Intel Compiler

```
$ echo $CC
gcc


$ module load compiler/intel


$ echo $CC
icc
```

Exercise 4

- Try to compile an MPI application

```
$ mpicc
bash: mpicc: command not found…
```

- Try to load OpenMPI

```
$ module load mpi/openmpi

Loading module dependency 'compiler/intel/14.0'.
compiler/intel/14.0(401):ERROR:150: Module 'compiler/intel/14.0' conflicts
with the currently loaded module(s) 'compiler/intel/17.0'
compiler/intel/14.0(401):ERROR:102: Tcl command execution failed: conflict
compiler/intel


ERROR: Failed to load module dependency 'compiler/intel/14.0'. Please check
conflicts with already loaded modules.
```

Exercise 4

- What is happening?

```
$ module show mpi/openmpi
-------------------------------------------------------------------
/opt/bwhpc/common/modulefiles/mpi/openmpi/1.8-intel-14.0:

(..)
```

- The default OpenMPI module was compiled against Intel Compiler version 14, but we loaded version 17 before

Exercise 4

- Load a matching OpenMPI version

```
$ module load mpi/openmpi/3.0-intel-17.0
```

- Now mpicc works

```
$ mpicc --version
icc (ICC) 17.0.6 20171215
Copyright (C) 1985-2018 Intel Corporation.  All rights reserved.
```