



Course:
High Performance Computing (HPC)
in Baden-Württemberg

J. Denev, S. Shamsudeen, S. Raffener, F. Bösert, P. Weisbrod, H. Häfner, R. 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



bwHPC: Hardware and Storage Architecture

Peter Weisbrod, SCC, KIT



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



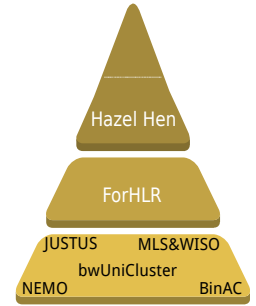
Funding:

Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

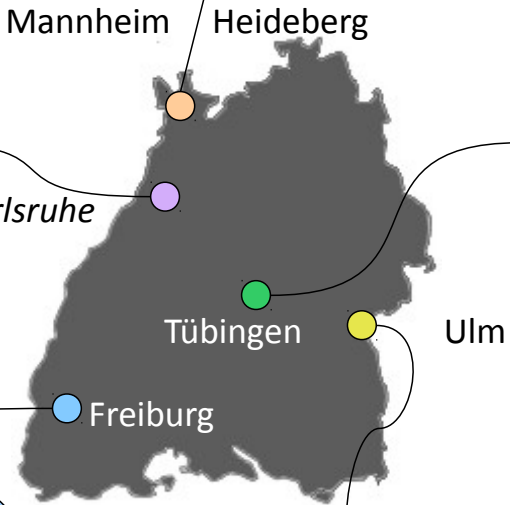
www.bwhpc-c5.de

Clusters @ Tier 2+3



bwForCluster MLS& WISO
(10/2015):
Economics & Social Science,
Molecular Life Science

bwUniCluster
(02/2014):
General purpose,
Teaching & Education
ForHLR I+II
(09/2014),(03/2016):
Research, high scalability



bwForCluster BinAC
(11/2016):
Bioinformatik,
Astrophysics

bwForCluster NEMO
(09/2016):
Neurosciences,
Micro Systems Engineering,
Elementary Particle Physics

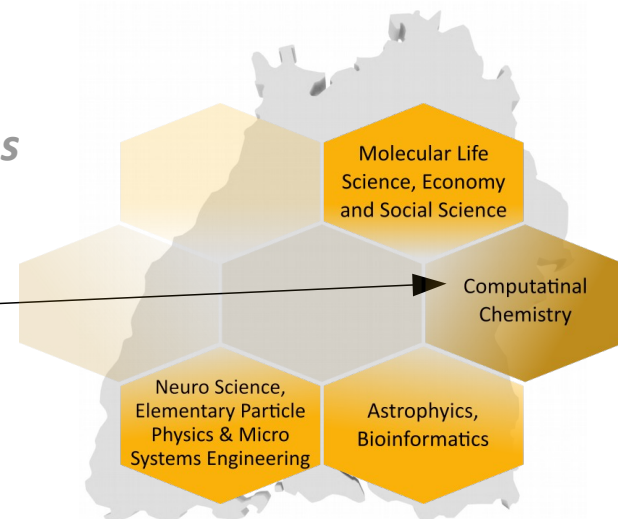
bwForCluster JUSTUS
(12/2014):
Computational Chemistry

bwForCluster JUSTUS

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **computational chemistry**
 - High I/O, large MEM jobs
- User and software support by *bwHPC competence center*



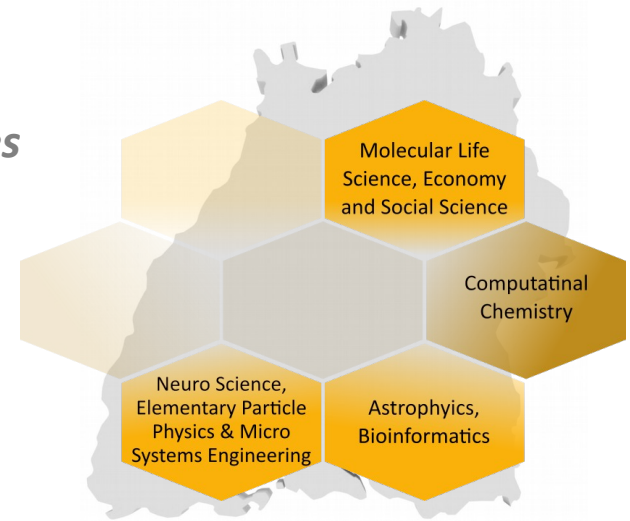
	Diskless	SSD	Big SSD	Large Mem SSD	Visual
# nodes	202	204	22	16	2
Core/node	16	16	16	16	16
Processor	2,4 GHz (Xeon E5-2630v3, Haswell)				
Main Mem	128 GiB		256 GiB	512 GiB	512 GiB
Local Storage	-	1 TB SSD	2 TB SSD		4 TB HDD
Interconnect	InfiniBand QDR				
Blocking	1:8				
HOME	192 TB NFS				
PFS – Workspaces	192 TB Lustre				
Block storage	480 TB (local mount via RDMA)				
Special feature					NVIDIA K6000

bwUniCluster

Federated HPC tier 3 resources

Selected characteristics:

- General purpose HPC entry level incl. education
- Universities are Shareholders
- Federated operations, multilevel fairsharing



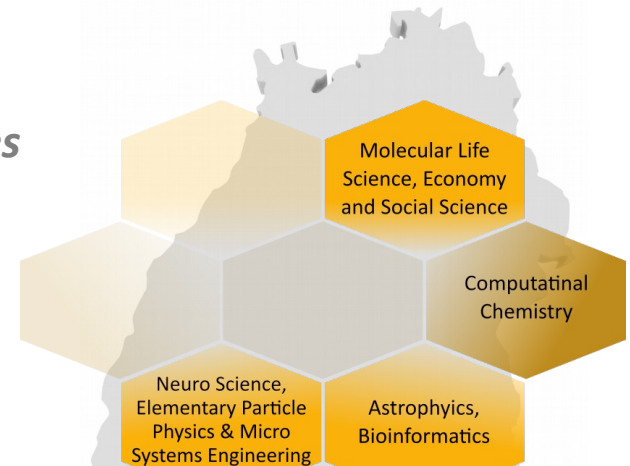
	Thin	Fat	<i>In Preparation</i>
# nodes	512	8	352
Core/node	16	32	28
Processor	2.6 GHz (Sandy Br.)	2.4 GHz (Sandy Br.)	2.0 GHz (Broadwell)
Main Mem	64 GiB	1024 GiB	128 GiB
Local Storage	2 TB HDD	7 TB HDD	480 GB SSD
Interconnect	InfiniBand 4x FDR		<i>InfiniBand FDR/EDR</i>
Blocking	1:1 (50%), 1:8 (50%)		1:1
PFS – HOME	469 TB Lustre		
PFS – Workspaces	938 TB Lustre		

bwUniCluster

Federated HPC tier 3 resources

Selected characteristics:

- General purpose HPC entry level incl. education
- Universities are Shareholders
- Federated operations, multilevel fairsharing



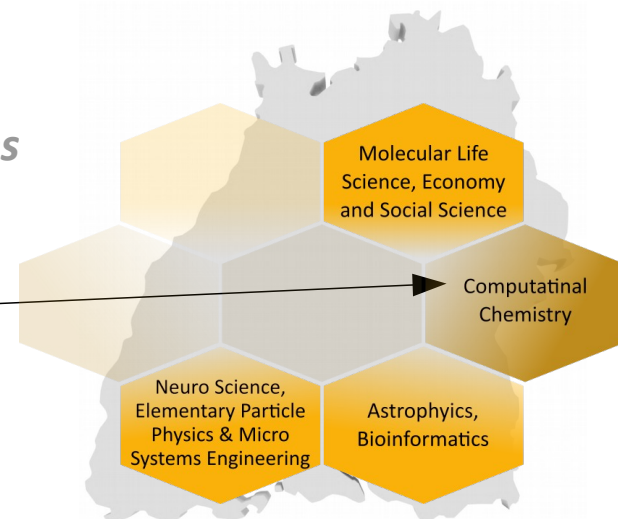
Property	\$TMP	\$HOME	\$WORK / workspace
Visibility	local	global	global
Lifetime	batch job walltime	permanent	min. 7 days / max. 240 days
Disk space	2 TB @ thin nodes 7 TB @ fat nodes 4 TB @ login nodes	427 TiB	853 TiB
Quotas	no	yes, per group	(currently) no
Backup	no	yes (default)	no
Read perf./node	280 MB/s @ thin node 593 MB/s @ fat node 416 MB/s @ login node	1 GB/s	1 GB/s
Write perf./node	270 MB/s @ thin node 733 MB/s @ fat node 615 MB/s @ login node	1 GB/s	1 GB/s
Total read perf.	n*280 593 MB/s	8 GB/s	16 GB/s
Total write perf.	n*270 733 MB/s	8 GB/s	16 GB/s

bwForCluster JUSTUS

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **computational chemistry**
 - High I/O, large MEM jobs
- User and software support by *bwHPC competence center*



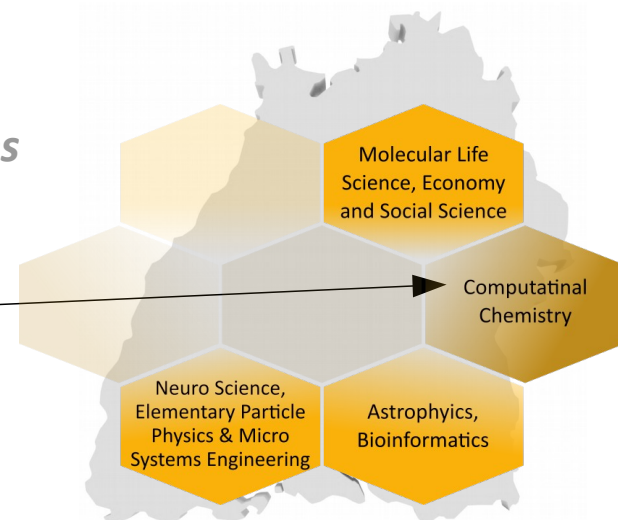
	Diskless	SSD	Big SSD	Large Mem SSD	Visual
# nodes	202	204	22	16	2
Core/node	16	16	16	16	16
Processor	2,4 GHz (Xeon E5-2630v3, Haswell)				
Main Mem	128 GiB		256 GiB	512 GiB	512 GiB
Local Storage	-	1 TB SSD	2 TB SSD		4 TB HDD
Interconnect	InfiniBand QDR				
Blocking	1:8				
HOME	192 TB NFS				
PFS – Workspaces	192 TB Lustre				
Block storage	480 TB (local mount via RDMA)				
Special feature					NVIDIA K6000

bwForCluster JUSTUS

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **computational chemistry**
 - High I/O, large MEM jobs
- User and software support by *bwHPC competence center*



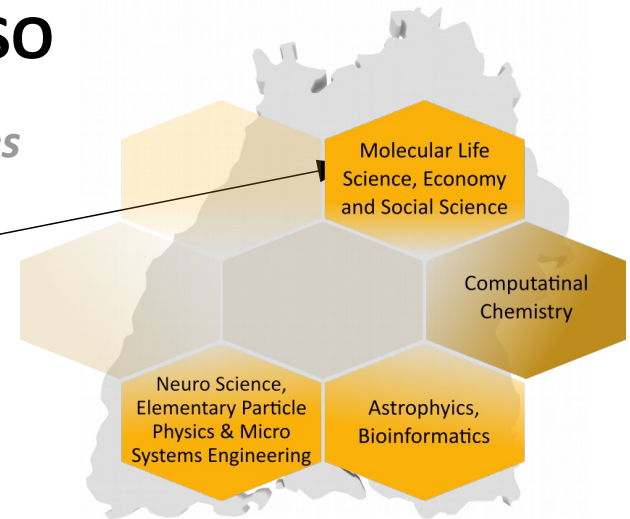
	\$TMPDIR	central block storage	workspaces	\$HOME
Visibility	local	on-demand local	global	global
Lifetime	batch job walltime	batch job walltime	< 90 days	permanent
Disk space	diskless/1TB/2TB	480 TB	200 TB	200 TB
Quotas	no	no	no	100 GB
Backup	no	no	no	yes

bwForCluster MLS&WISO

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **molecular life science, economics and social science + cluster for method development**
- User and software support by *bwHPC competence center*



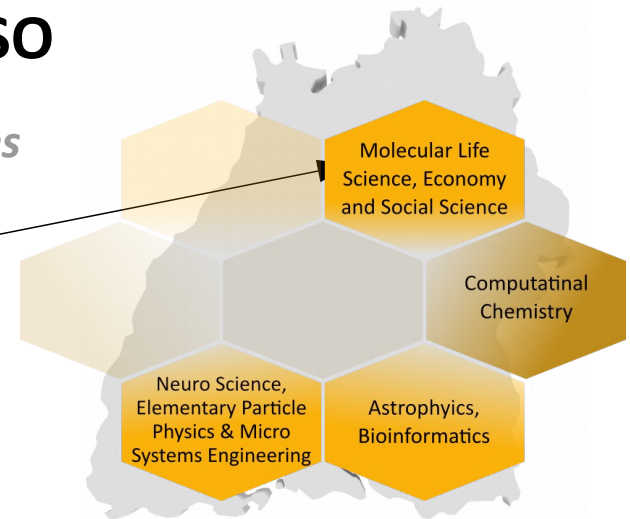
	Standard	Best	Coprocessor (GPU)	Coprocessor (MIC)	Fat	Fat (Ivy Bridge)
Node Feature	standard	best	gpu	mic	fat	fat-ivy
Quantity	476	148	18	12	8	4
Processors	2 x Intel Xeon E5-2630v3 (Haswell)	2 x Intel Xeon E5-2640v3 (Haswell)	2 x Intel Xeon E5-2630v3 (Haswell)	2 x Intel Xeon E5-2630v3 (Haswell)	4 x Intel Xeon E5-4620v3 (Haswell)	4 x Intel Xeon E4-4020v2 (Ivy Bridge)
Processor Frequency (GHz)	2.4	2.6	2.4	2.4	2.0	2.6
Number of Cores	16	16	16	16	40	32
Working Memory (GB)	64	128	64	64	1536	1024
Local Disk (GB)	128 (SSD)	128 (SSD)	128 (SSD)	128 (SSD)	9000 (SATA)	128 (SSD)
Interconnect	QDR	FDR	FDR	FDR	FDR	FDR
Coprocessors	-	-	1 x Nvidia Tesla K80	2 x Intel Xeon Phi 5110P	-	-

bwForCluster MLS&WISO

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **molecular life science, economics and social science**
+ cluster for method development
- User and software support by *bwHPC competence center*



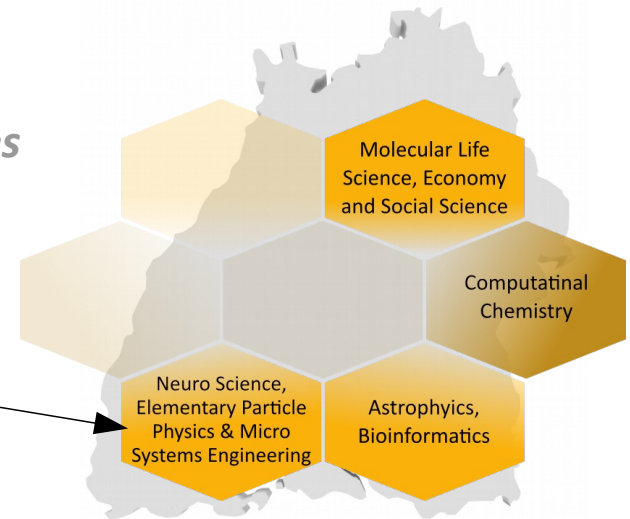
	\$HOME	Workspaces	\$TMPDIR
Visibility	global	global	node local
Lifetime	permanent	workspace lifetime	batch job walltime
Capacity	36 TB	384 TB	128 GB per node (9 TB per fat node)
Quotas	100 GB	none	none
Backup	no	no	no

bwForCluster NEMO

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **neuro science, elementary particle physics, micro systems engineering**
 - Virtual machine images deployable
- User and software support by *bwHPC competence center*



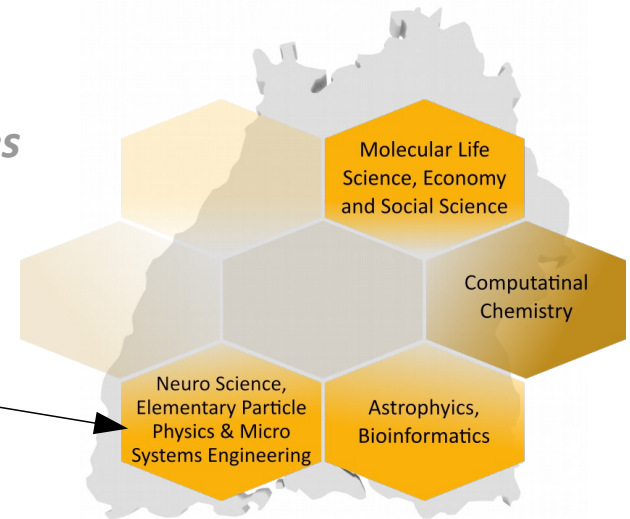
	Compute Node	Coprocessor (MIC)
Quantity	548	4
Processors	2 x Intel Xeon E5-2630v4 (Broadwell)	1 x Intel Xeon Phi 7210 Knights Landing (KNL)
Processor Frequency (GHz)	2,2	1,3
Number of Cores	20	64
Working Memory DDR4 (GB)	128	16 GB MCDRAM + 96 GB DDR4
Local Disk (GB)	240 (SSD)	240 (SSD)
Interconnect	Omni-Path 100	Omni-Path 100

bwForCluster NEMO

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **neuro science, elementary particle physics, micro systems engineering**
 - Virtual machine images deployable
- User and software support by *bwHPC competence center*



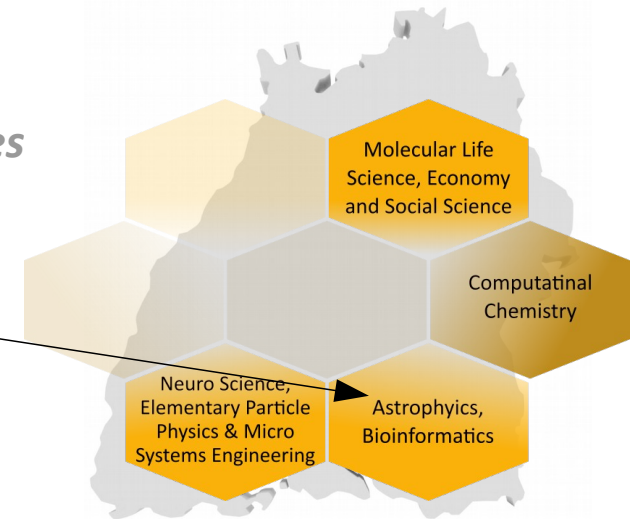
	\$HOME	Work Space	\$TMPDIR
Visibility	global (GbE)	global (Omni-Path)	node local
Lifetime	permanent	work space lifetime (max. 100 days, with extensions up to 400)	batch job walltime
Capacity	30 TB	576 TB	220 GB per node
Quotas	100 GB per user	none	none
Backup	snapshots + tape backup	no	no

bwForCluster BinAC

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **astrophysics, bioinformatics**
 - Dual GPU systems
- User and software support by *bwHPC competence center*



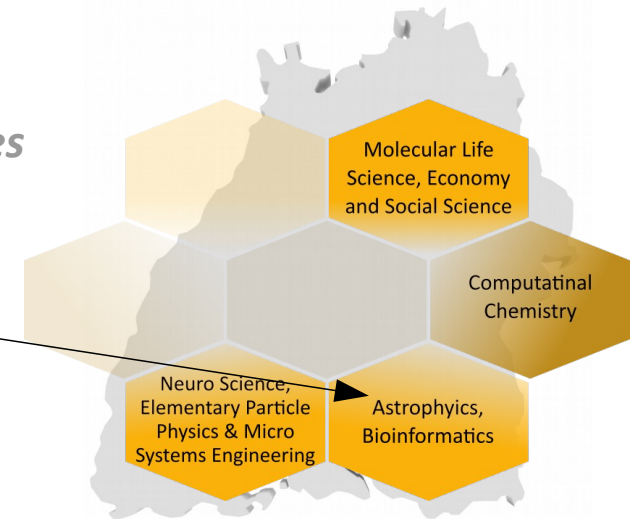
	Standard	Fat	GPU
Quantity	236	4	60
Processors	2 x Intel Xeon E5-2630v4 (Broadwell)	4 x Intel Xeon E5-4620v3 (Haswell)	2 x Intel Xeon E5-2630v4 (Broadwell)
Processor Frequency (GHz)	2.4	2.0	2.4
Number of Cores	28	40	28
Working Memory (GB)	128	128	1024
Local Disk (GB)	256 (SSD)	256 (SSD)	256 (SSD)
Interconnect	FDR	FDR	FDR
Coprocessors	-	-	2 x Nvidia Tesla K80

bwForCluster BinAC

Federated HPC tier 3 resources

Selected characteristics:

- Dedicated to **astrophysics, bioinformatics**
 - Dual GPU systems
- User and software support by *bwHPC competence center*



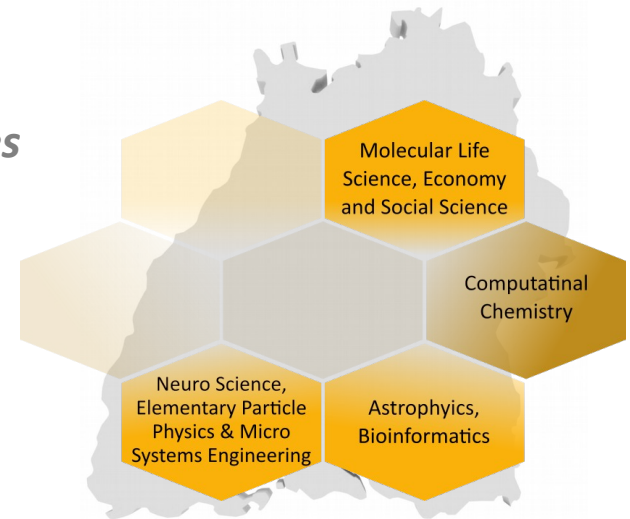
	\$HOME	Work Space	\$TMPDIR
Visibility	global	global	node local
Lifetime	permanent	work space lifetime (max. 100 days, with extensions up to 400)	batch job walltime
Capacity	unkn.	520 TB	220 GB per node
Quotas	20 GB per user	none	none
Backup	no	no	no

ForHLR I

Federated HPC tier 2 resources

Selected characteristics:

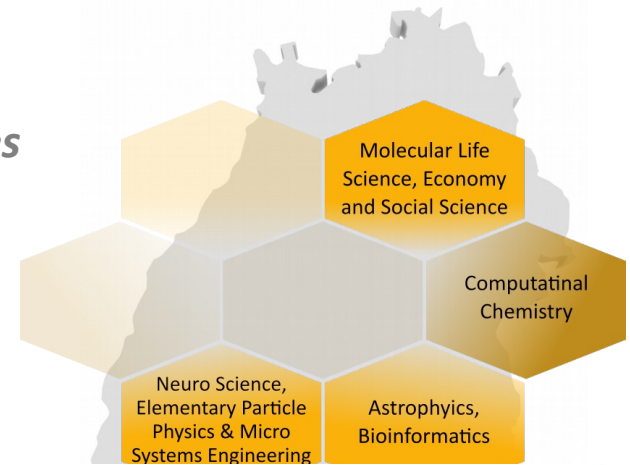
- Next level for advanced HPC users
- Research, high scalability



	Thin	Fat
# nodes	512	16
Core/node	20	32
Processor	2.5 GHz (Sandy Br.)	2.6 GHz (Sandy Br.)
Main Mem	64 GiB	512 GiB
Local Storage	2 TB HDD	8 TB HDD
Interconnect	InfiniBand 4x FDR	
Blocking	Non-blocking	
PFS – HOME	427 TB Lustre	
PFS – Workspaces	PROJECT 427 TB Lustre, WORK/workspace 853 TB Lustre	

ForHLR I

Federated HPC tier 2 resources



Selected characteristics:

- Next level for advanced HPC users
- Research, high scalability

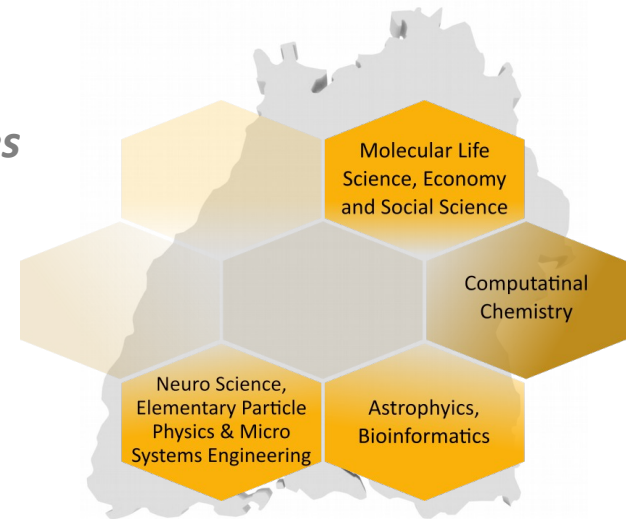
Property	\$TMP	\$PROJECT	\$WORK / workspace	\$HOME
Visibility	local	global	global	global
Lifetime	batch job walltime	permanent	usually 28 days / max. 240 days	permanent
Disk space	2 TB @ thin nodes 8 TB @ fat nodes 5 TB @ login nodes	427 TiB	853 TiB	427 TiB (limited usage)
Quotas	no	yes	no	yes
Backup	no	yes (default)	no	yes (default)
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	8 GB/s
Total write perf.	n*270 733 MB/s	8 GB/s	16 GB/s	8 GB/s

ForHLR II

Federated HPC tier 2 resources

Selected characteristics:

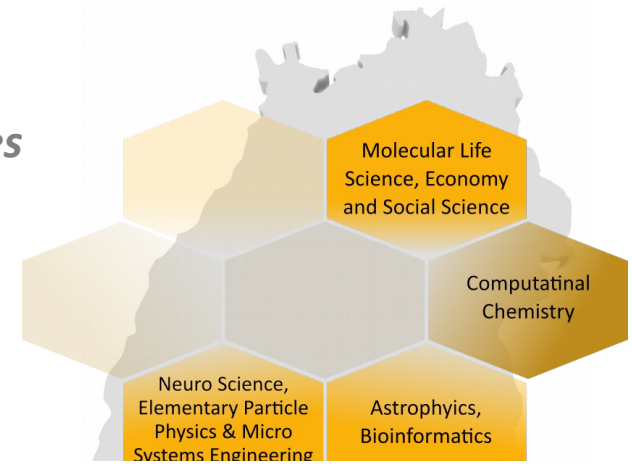
- Next level for advanced HPC users
- Research, high scalability



	Thin	Fat
# nodes	1152	21
Core/node	20	48
Processor	2.6 GHz (Haswell)	2.1 GHz (Haswell)
Main Mem	64 GiB	1024 GiB
Local Storage	480 GB SSD	3840 GB SSD
Interconnect	InfiniBand 4x FDR	
Blocking	Non-blocking	
Graphic cards		4 NVIDIA GeForce GTX980 Ti
PFS – HOME	469 TB Lustre	
PFS – Workspaces	PROJECT 610 TB Lustre, WORK 1220 TB Lustre, workspace 3050 TB Lustre	

ForHLR II

Federated HPC tier 2 resources



Selected characteristics:

- Next level for advanced HPC users
- Research, high scalability

Property	\$TMP	\$PROJECT	\$WORK	workspace	\$HOME
Visibility	local	global	global	global	global
Lifetime	batch job walltime	permanent	usually 28 days	max. 240 days	permanent
Disk space	400 GB @ compute nodes 3600 GB @ rendering nodes 400 GB @ login nodes	610 TiB	1220 TiB	3050 TiB	427 TiB (limited usage)
Quotas	no	yes	no	no	yes
Backup	no	yes (default)	no	no	yes (default)
Read perf./node	500 MB/s @ compute node ??? MB/s @ rendering node 500 MB/s @ login node	2 GB/s	2 GB/s	2 GB/s	1 GB/s
Write perf./node	500 MB/s @ compute node ??? MB/s @ rendering node 500 MB/s @ login node	2 GB/s	2 GB/s	2 GB/s	1 GB/s
Total read perf.	n*500 ??? MB/s	10 GB/s	20 GB/s	50 GB/s	8 GB/s
Total write perf.	n*500 ??? MB/s	10 GB/s	20 GB/s	50 GB/s	8 GB/s

Thank you for your attention!

Questions?