

Einführung in wiss. Software

auf dem

bwGRiD-Cluster

Dr. Heinz Kredel und Dr. Sabine Richling

Inhalt

- bwGRiD Cluster
- Java Programmierung
 - Java Threads, Java OpenMP,
 - TCP/IP Kommunikation, Java MPP, Strategien
- C und FORTRAN Programmierung
 - OpenMP und OpenMPI
- Parallelisierung (fertiger) Software
 - Matlab, Scilab, Octave, R, Stata, Mathematica

Termine (1)

9.9.: Vorbesprechung

16.9.: Einleitung bwGRiD Cluster

23.9.: Java Thread Programmierung (1)

20.9.: Java Thread Programmierung (2)

7.10: Java OpenMP Programmierung

14.10.: Java TCP/IP Programmierung

21.10: Java MPP Programmierung

28.10: Einführung in Fortran und C

Termine (2)

4.11.: Parallelisierung mit OpenMP

11.11.: Parallelisierung mit MPI – Grundlagen,
Kommunikationsarten

18.11.: Parallelisierung mit MPI – Datentypen,
virtuelle Topologien

25.11.: Parallelisierung mit MPI – Parallele Ein-
und Ausgabe, Leistungsmessung

2.12.: Statistik-Software (R, Stata)

9.12.: Mathematik-Software (Matlab & Clone)

Grid

The Grid is a system that:

- coordinates resources that are not subject to centralized control ...
- ... using standard, open, general-purpose protocols and interfaces ...
- ... to deliver nontrivial qualities of service.

Ian Foster, 2002

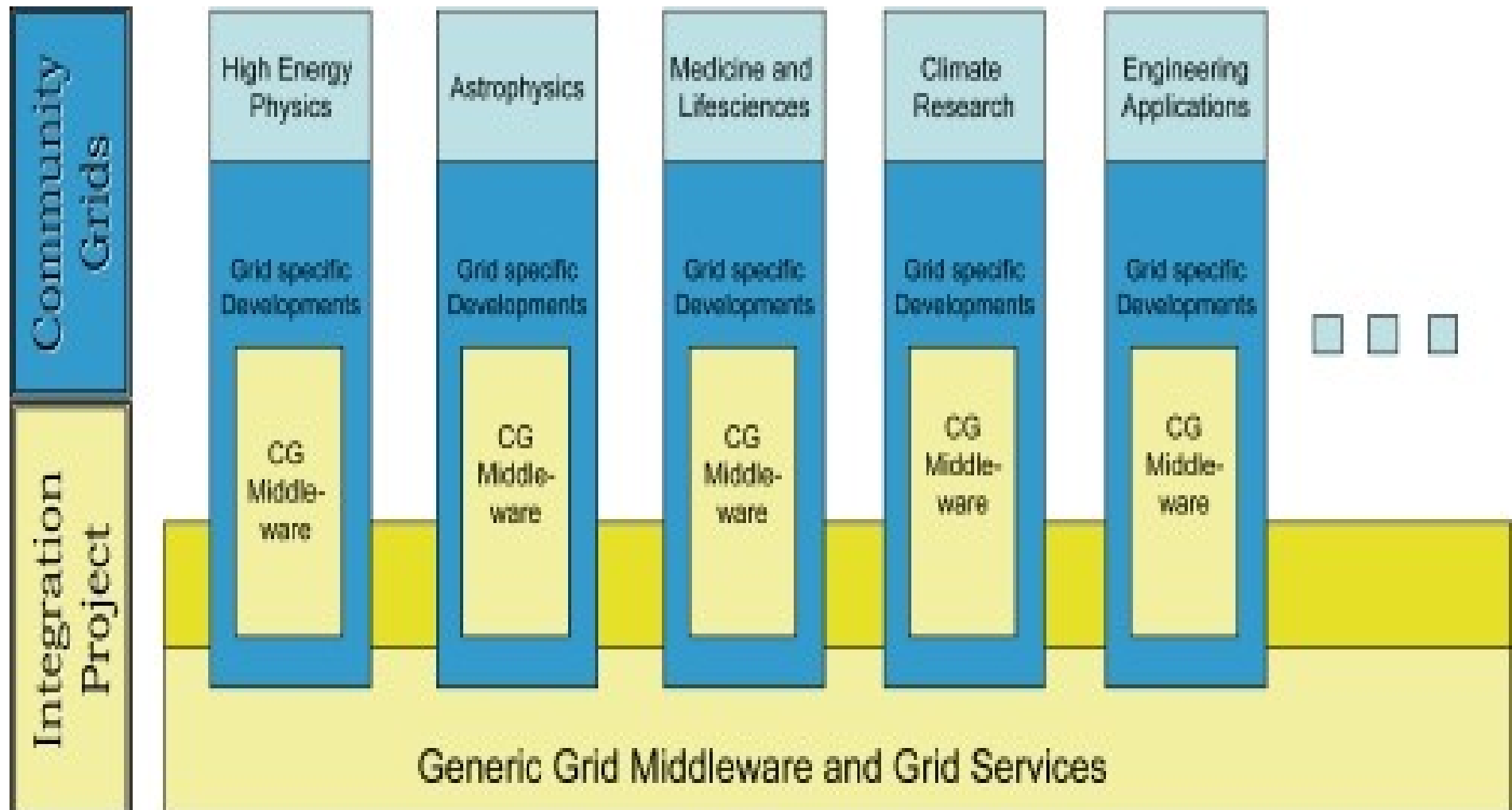
D-Grid

Das Grid

- ermöglicht den direkten Zugriff auf Ressourcen, wie Rechner, Speicher, wissenschaftliche Instrumente und Experimente, Anwendungen und Daten, Sensoren und sogenannte Grid-Middleware Dienste
- basierend auf weit verbreiteten Grid- und Web-Services-Standards.

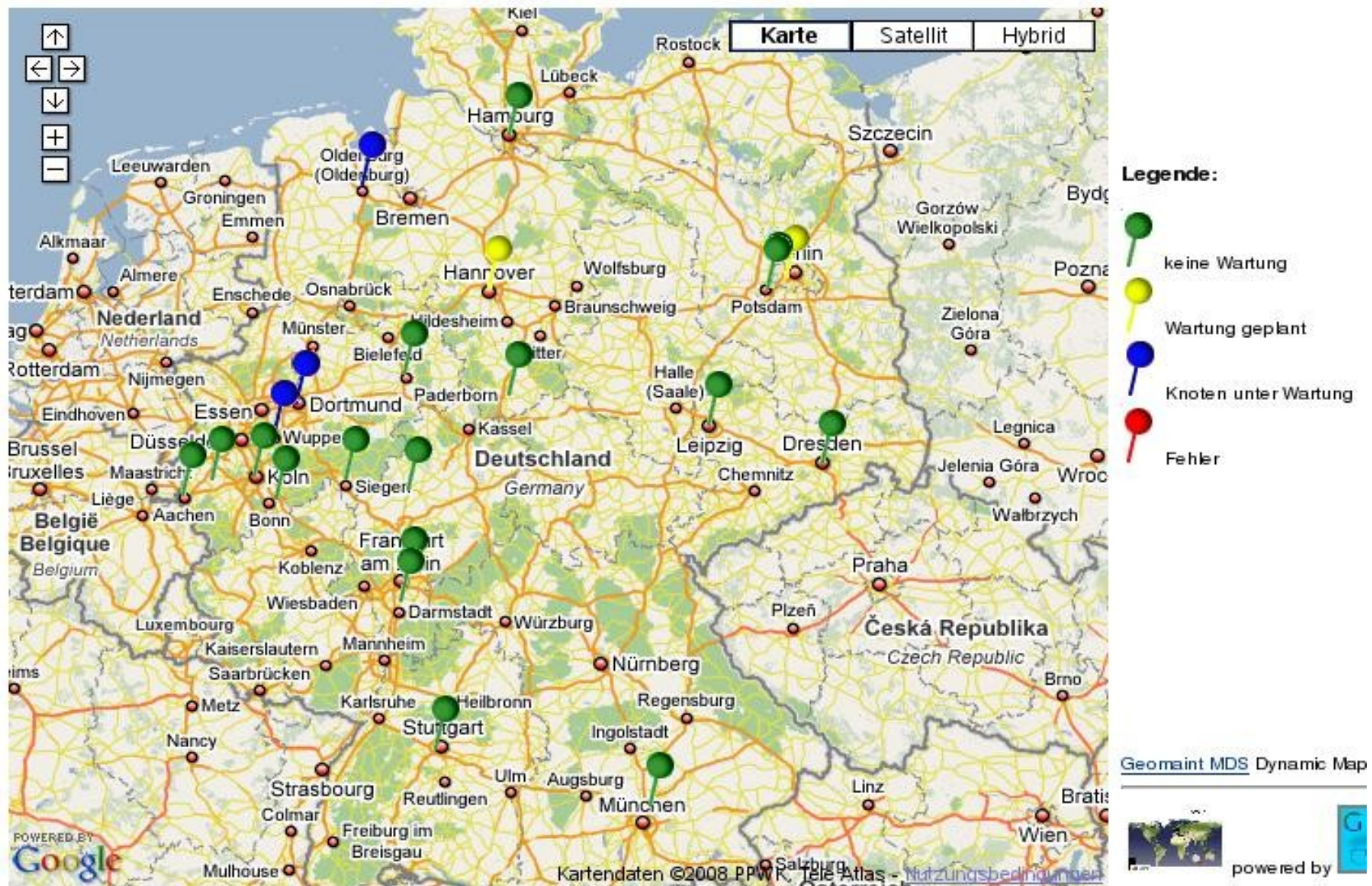
D-Grid Web-Seite, 2008

D-Grid Projekte



D-Grid Ressourcen (2008)

Dynamisch erstellte Karte von Ressourcenanbietern / Wartungszustand



bw-Grid Cluster

- Projektantrag vom HLRS an BMFT in 2007
- für D-Grid Infrastruktur an den Universitäten in Baden-Württemberg
- explizit als verteiltes System
- mit dezentraler Verwaltung
- an den Standorten
 - Stuttgart, Ulm (mit Konstanz), Freiburg, Tübingen, Karlsruhe, Heidelberg, Mannheim
- für die nächsten 5 Jahre

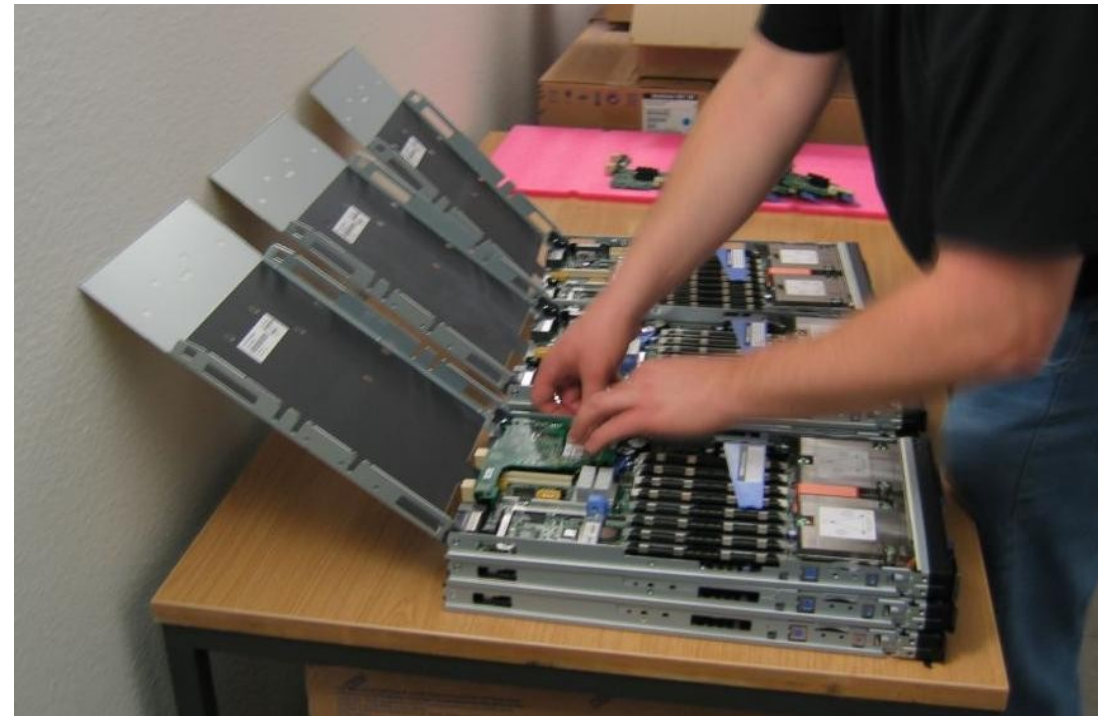
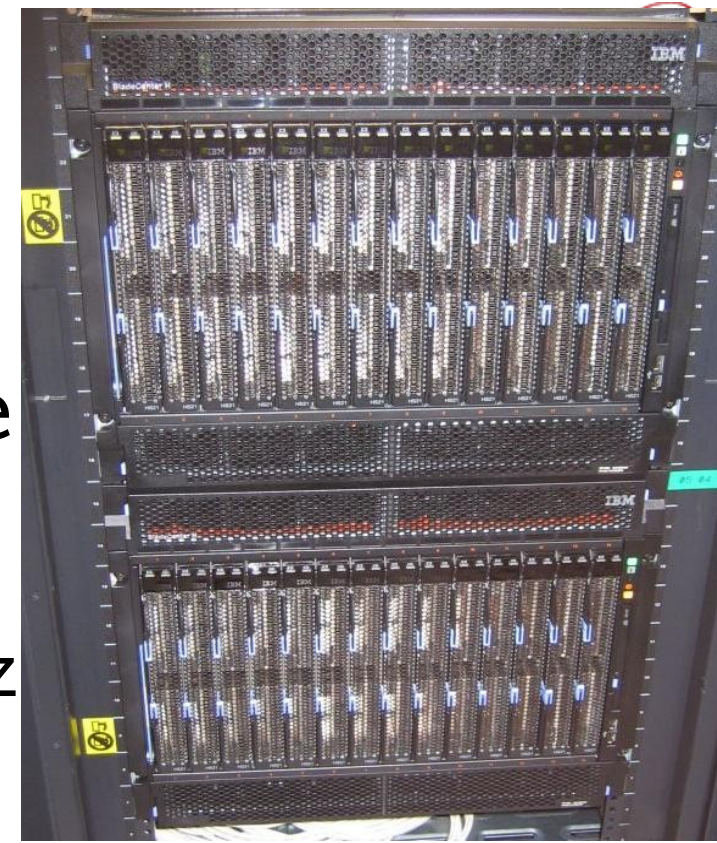
bw-Grid Ziele

- Nachweis der Funktionalität und des Nutzens von Gridkonzepten im HPC Umfeld
- Überwindung von bestehenden Organisations- und Sicherheitsproblemen
- Entwicklung von neuen Cluster- und Grid-Anwendungen
- Lösung der Lizenzproblematik
- Ermöglichung der Spezialisierung von Rechenzentren

Hardware

- 101 IBM Blade-Center Gehäuse
 - mit je 14 IBM HS21 XM Blades
 - mit je 2 Intel Xeon CPUs, 2.8 GHz
 - mit je 4 Cores
 - 16 GB Hauptspeicher
 - keine Festplatten
 - Gigabit Ethernet
 - Infiniband Netzwerk
- = 1414 Blades

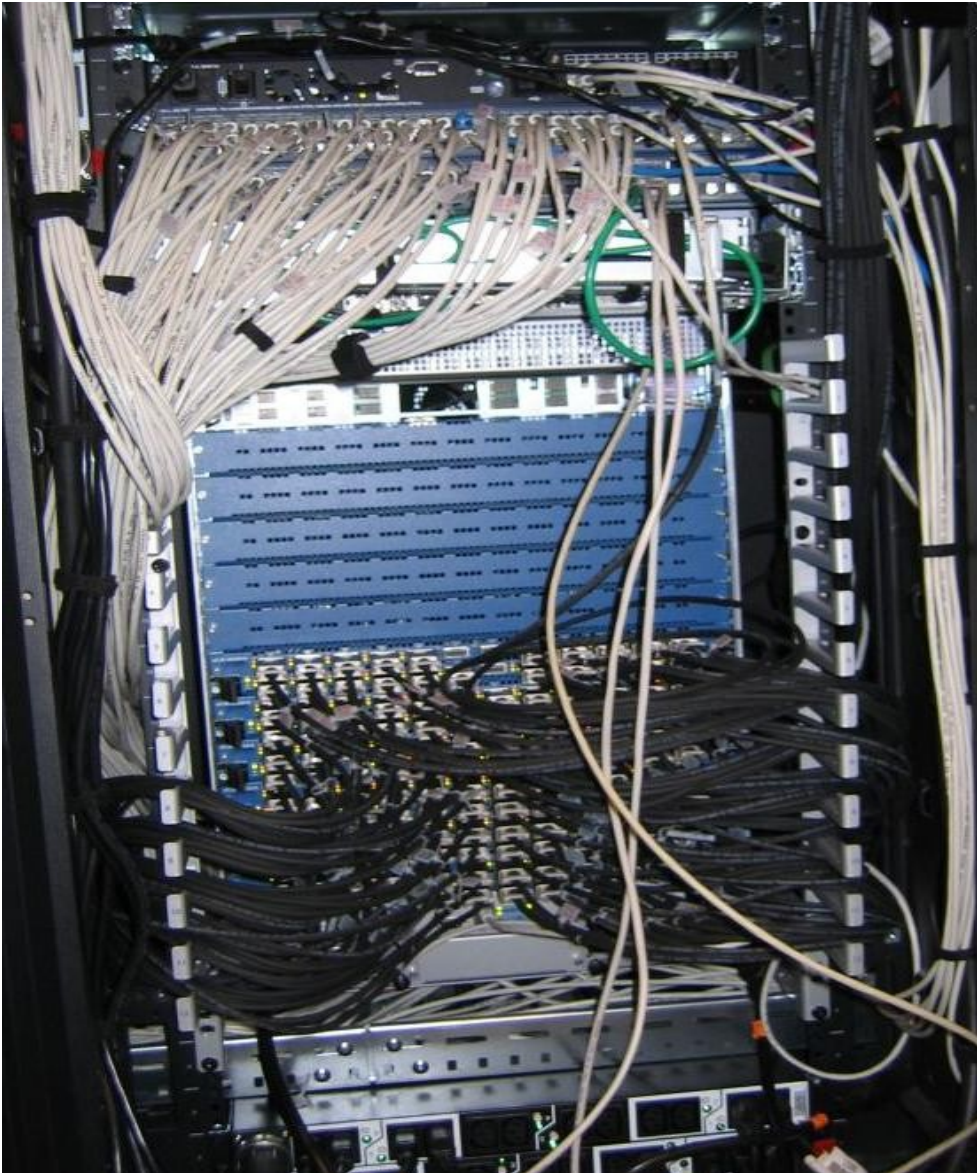
= 11312 CPU Cores,
1120 Cores in Mannheim



Racks mit Blade Center



Netz



Software

- Scientific Linux 5.0
 - basierend auf RedHat 5.0 Quellcode
 - gepflegt u.A. vom CERN
 - dort sind ca. 140.000 Rechner für LHC im Einsatz
 - ähnlich CentOS
- booten über das Netz mit DHCP und TFTP
- Festplatten über NFS, jetzt auch mit Lustre
- Swap-Bereich in lokaler Festplatte
 - am Anfang minimaler Swap-Bereich über NBD
- Kernel und System Image vom HLRS vorbereitet

Anwendungen

	Mathematik	Biowissenschaften	Ingenieurwiss.	Wirtschaftswiss.	Chemie	Physik	Informatik
Freiburg	25% CFD		20% Mikrosystemt.				
Heidelberg	25% CAS, Ug	20% Comp. NeuroSc.					
Karlsruhe	10 % LinAlg		30% CFD, CSM				
Konstanz	x SPH	x Biochemie				x Theor. Ph, QM	x DataMining
Mannheim	15% CAS			30% CAS, Simulation			
Stuttgart			35% CFD, Comp.Mech.				
Tübingen		20% BioInfo				25% Astrophysik	
Ulm		5% BioInfo, Medizin			25% Theor. Ch, MD		

Alle: ca. 55% Betrieb, Middleware, Compiler, Tools

CFD = Computational Fluid Dynamics

CAS = Computer Algebra Systems

MD = Molecular Dynamics

QM = Quantum Mechanics

Aufbau Anfang 2008

- ein tech. Admin für MA und HD zusammen
- Hardware
 - geliefert und aufgebaut, Januar bis März 2008
 - Netzwerk funktioniert, Ethernet und auch Infiniband
- Administration und Betrieb
 - Festplatten Platz für Home-Dirs über NFS-Server
 - Benutzerverwaltung ist konfiguriert, über LDAP
 - Batchsystem konfiguriert
 - in HD Nutzung der **Benutzerverwaltung und des NFS Speichers des IWR**

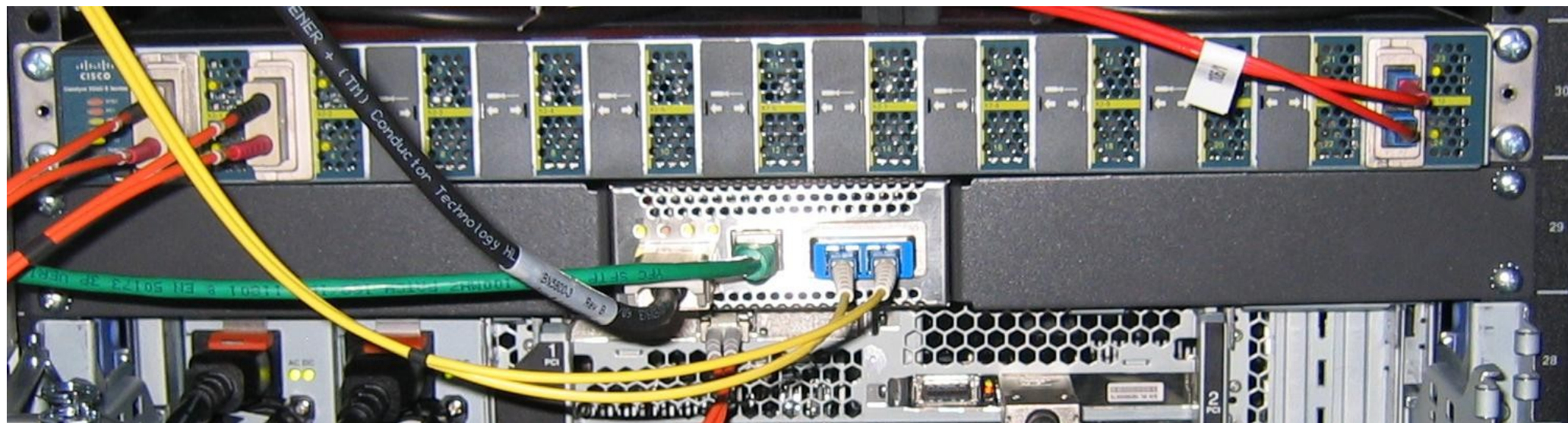
HP Lustre Speichersystem

- 32 TByte netto
- Paralleles Filesystem
- Seit Anfang 2009



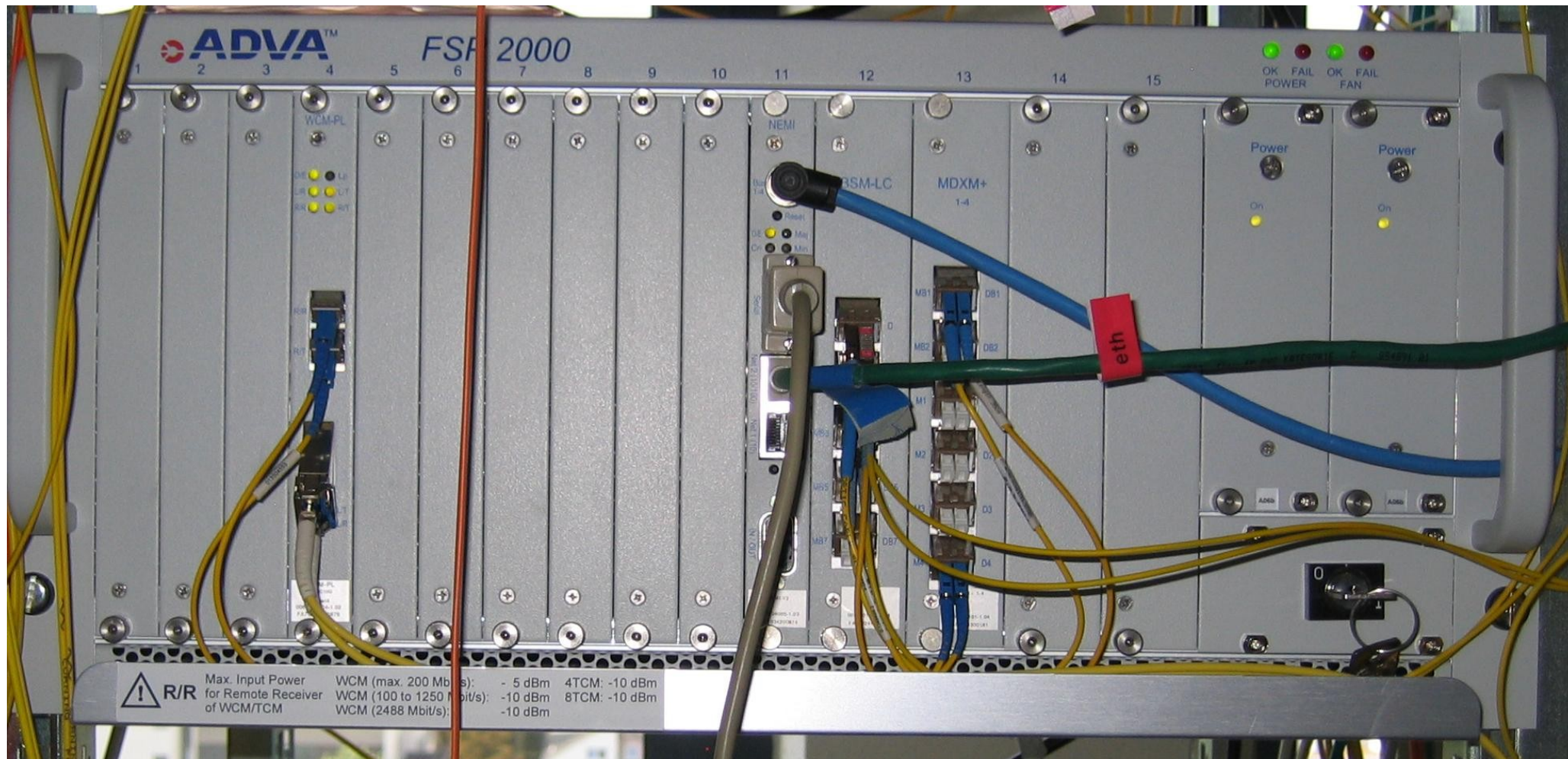
Kopplung mit Heidelberg

- Antrag in 2008
- Beschaffung und Aufbau bis Mai 2009
- in Betrieb seit Juli 2009
- InfiniBand auf Glasfaser: Obsidian Longbow

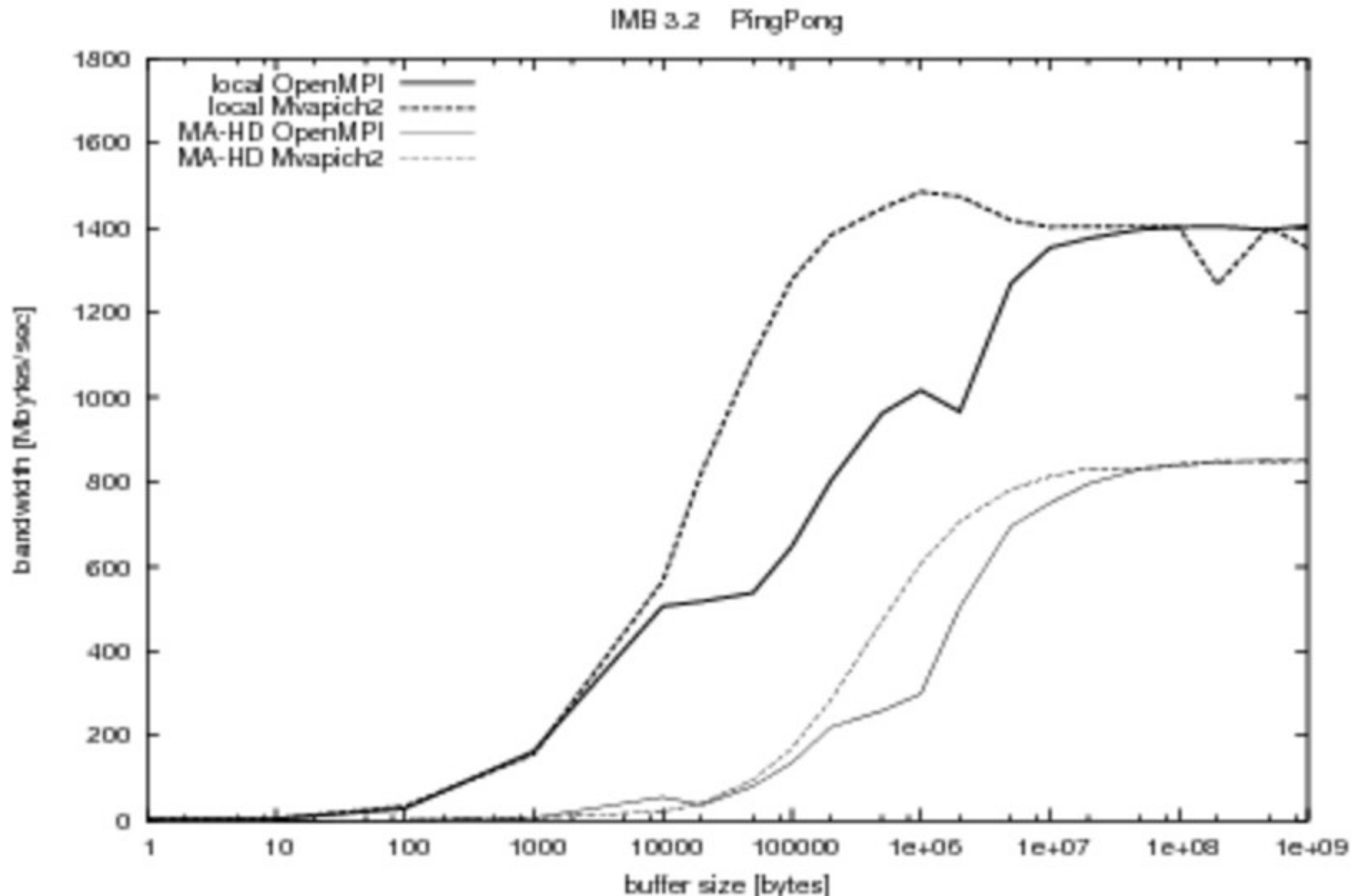


Kopplung - Hardware

- ADVA Einspeisung CWDM Strecke



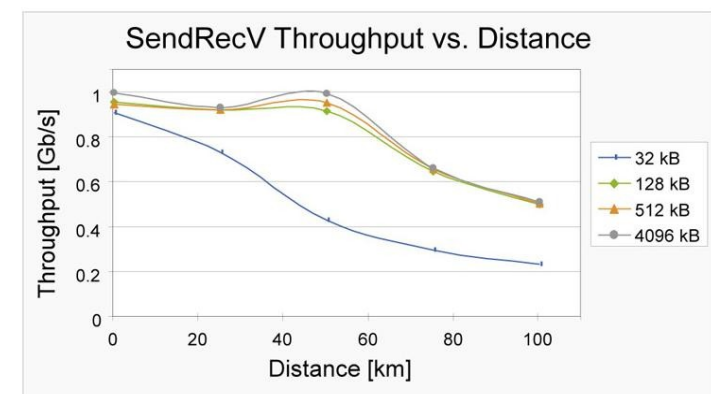
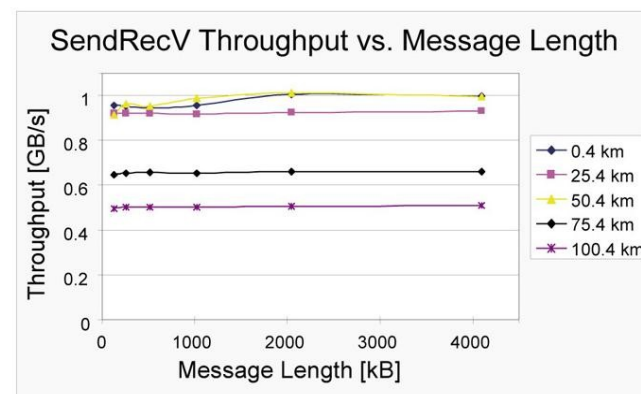
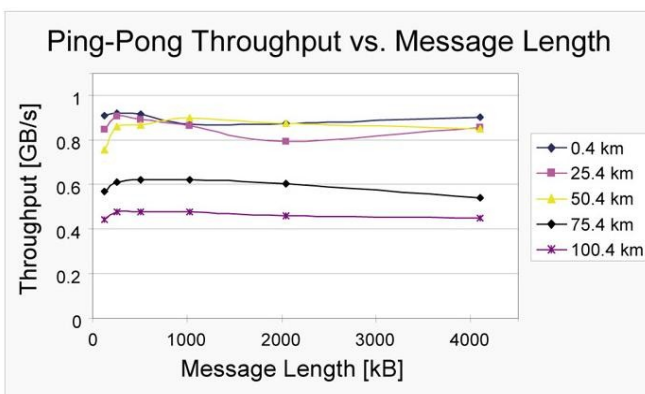
MPI Performance



HLRS MPI Performance

- Messungen für verschiedene Entfernungen
- bis 50 km akzeptabel
- Bandbreite ca. 900 – 1000 MB/sec
- keine Angaben zur gemessenen Latenz

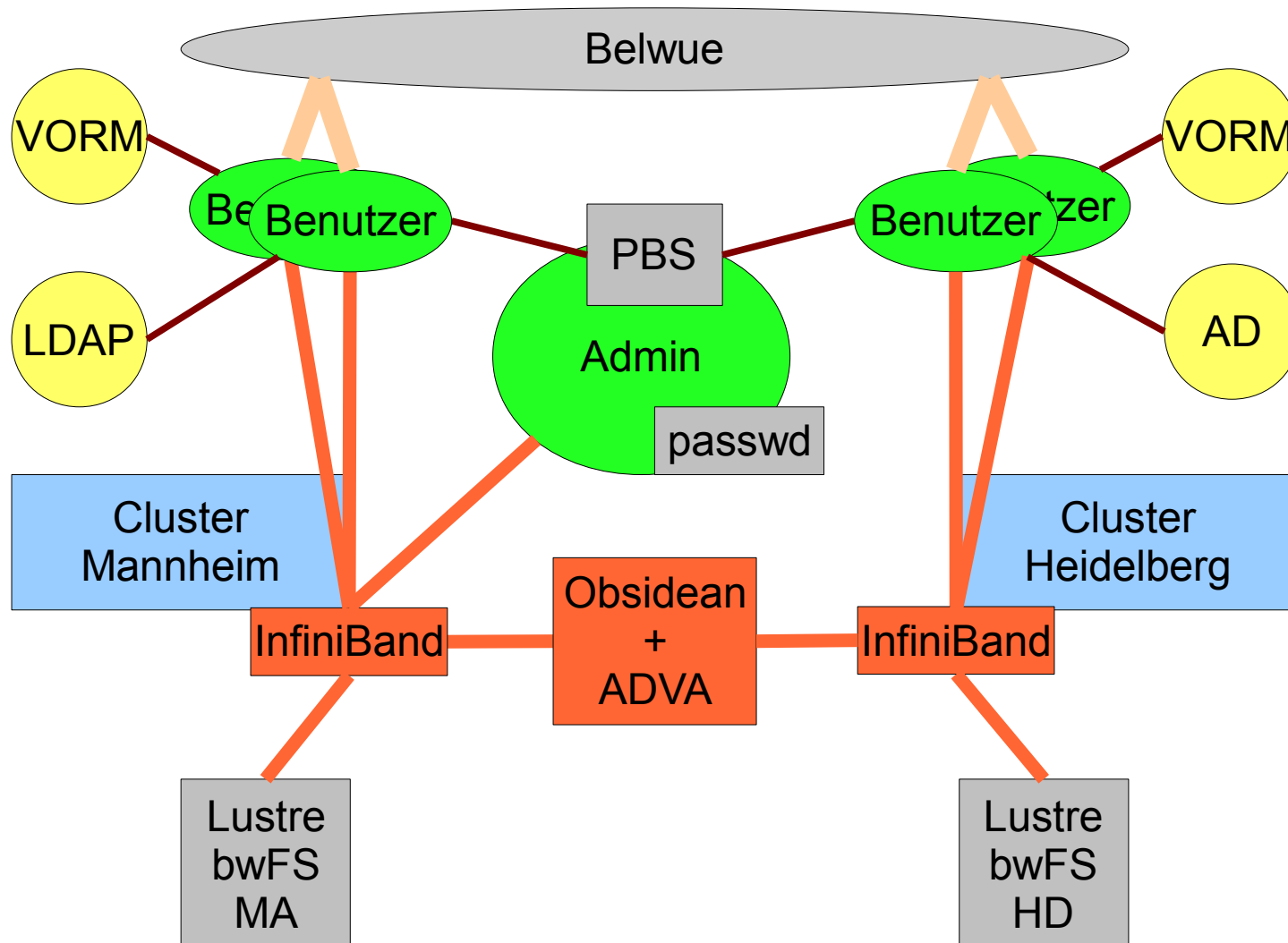
Measurement results – full InfiniBand throughput over more than 50km distance



Hardware Erfahrungen

- Entfernung MA - HD beträgt 28 km (18 Luftlinie)
 - Licht benötigt 116 us für die Entfernung
- Obsidian benötigt eine Lizenz für 40 km
 - hat Puffer für noch weitere Entfernungen
 - diese werden nur über die Lizenz aktiviert
 - wir waren der Meinung Lizenz für 10 km reicht
- Latenz mit 145 us recht hoch, Licht + 29 us
 - lokale Latenz 1.99 us P-t-P (15 us coll. comm.)
- Bandbreite mit ca. 930 MB/sec wie erwartet
 - lokale Bandbreite ca. 1200 - 1400 MB/sec

Standorte und Komponenten



neben InfiniBand noch 1 GE Verwaltungsnetz

gemeinsame Administration

- nur noch ein Admin-Server, ein PBS
- 2 Zugangsrechner für SSH, 10GE an Belwue
- 2 Zugangsrechner für GSI-SSH, ++, 10GE
- Cluster-Admin-Tools vom HLRS zur Administration der Hardware
 - MAC-Adressen, DHCP Tabelle
 - TFTP zum Booten der Kernel
 - NFS für (Admin) Software und Konfigurationen
- zusätzlich 2 extra BladeCenter für ITP (Inst. für Theor. Physik HD) in HD eingebaut

gemeinsame Benutzerverwaltung

- Nutzung RZ-Benutzerverwaltung, mit Grid-Flag
- lokale Kennungen in MA und HD (die das bwGRiD benutzen) sind alle verschieden (!)
- generieren Dateien für passwd und group
- Gruppen werden mit Prefix 'ma', 'hd', bzw. 'mh' für d-grid Nutzer versehen
- uidNumber +100.000 für MA, +200.000 für HD und +2.000.000 für d-grid
- Authentifizierung beim Zugangsrechner direkt am jeweiligen LDAP, AD oder Grid-Zertifikat

getrennte Lustre Systeme

- BWFS Lustre mit 32 TB an jedem Standort
- Homes der Benutzer bleiben am Standort
- Lustre über Kreuz gemounted, über InfiniBand
 - /bwfs/ma/home
 - /bwfs/hd/home
 - /bwfs/ma/gridhome
- InfiniBand Performance scheint auszureichen
- Backup mit TSM etwas langsam

gemeinsames Batch-System

- Jobs bleiben auf einen Cluster beschränkt
 - wegen Latenz
- also keine Jobs mit mehr als ~140 Knoten
- Performance von MPI mit InfiniBand bei Kommunikation über 28 km nicht ausreichend
 - aber keine Tests mit realen Workloads
- PBS mittlerweile mit Moab Scheduler
- 5 Queues: single, normal, itp, testkoppl, itptest

Benutzung 2008

- Batchsystem
 - qsub shell-script
 - qstat
 - nstat
- PBS Optionen
 - #PBS -l nodes=30
 - #PBS -q normal
- Ergebnis in home
 - shell-script.o113
 - shell-script.e113

```
[kredel@zug-ma-2 ~]$ nstat
```

node/job status on IBM bw-GRID v1.4						
ID	user	name	nodes	state	start [h]	remain [h]
113	kredel	versuch	10	R	-0.03	11.97
114	kredel	versuch	30	R	-0.01	11.99

waiting jobs: 0, requested nodes waiting for running: 0

Node	+0	+1	+2	+3	+4	+5	+6
n010301	-	-	-	-	-	-	-
n010308	-	-	-	-	-	-	-
n010401	-	-	-	-	-	-	-
n010408	-	-	-	-	-	-	-
n020301	-	-	-	-	-	-	-
n020308	-	-	-	-	-	F	F
n020401	F	F	F	F	F	F	F
n020408	F	F	F	F	F	F	F
n030101	F	F	F	F	F	F	F
n030108	F	F	F	F	F	F	F
n030201	F	F	F	F	F	F	F
n030208	F	F	F	F	F	F	F
n040301	F	F	F	F	F	F	F
n040308	F	F	F	F	F	F	F
n040401	F	F	F	F	F	F	F
n040408	F	F	F	F	F	F	F
n050301	F	F	F	F	F	F	F
n050308	F	F	F	F	F	F	F
n050401	F	F	F	F	F	F	F
n050408	F	F	F	F	F	F	F

F = free R = reserved . = empty - = offline
 H = Head I = I/O node + = denied

0/140 nodes reserved (0.0%)
 0/140 nodes used (0.0%)
 0/140 nodes offline (0.0%)
 100/140 nodes free (71.4%)

Benutzung nach der Kopplung

- $2 \times 140 + 28 = 308$
Knoten
- 28 reserviert für ITP
(Th. Physik)

Node	+0	+1	+2	+3	+4	+5	+6
n010301	1098	1098	1099	1099	1100	1100	1100
n010308	1101	1101	1101	1102	1102	1102	1102
n010401	1103	1103	1103	1103	1104	1104	1104
n010408	1153	1153	1104	1105	1105	1105	1105
n020301	1106	1106	1106	1106	1159	1159	1164
n020308	1164	1108	1108	1108	1108	1109	1109
n020401	1109	1109	1110	1153	1100	1110	1110
n020408	1110	1111	1111	1111	1111	1153	1112
n030101	1112	1112	1112	1113	1113	1113	1113
n030108	1048	1048	1055	1154	1154	1154	1154
n030201	F	F	F	1142	1055	1055	1055
n030208	1055	1055	1055	1055	1101	1065	1065
n040301	1115	1115	1115	1115	1116	1116	1116
n040308	1116	1117	1117	1117	1117	1118	1118
n040401	1118	1118	1119	1119	1119	1119	F
n040408	F	1059	1059	1142	1059	1059	1059
n050301	1059	1059	1059	1114	1114	1114	1114
n050308	1121	1121	1121	1121	1121	1121	1121
n050401	1121	1121	1121	1121	1121	1121	1142
n050408	1121	1121	1142	F	F	F	F
n060101	F	F	F	F	1000	1000	1000
n060108	1000	1000	1000	1000	1000	1097	1097
n060201	1097	1097	1097	1097	1097	1097	1097
n060208	-	1097	1000	1000	1000	1155	1155
n060301	1156	1156	1156	1120	1120	1120	1120
n060308	1120	1120	1120	1125	1125	1125	1125
n060401	1125	1156	1125	1125	1125	1125	1063
n060408	1063	1063	1063	1063	1063	1063	1063
n070101	1064	1064	1064	1064	1064	1064	1064
n070108	1064	1064	1064	1064	1064	1066	1066
n070201	1066	1066	1066	1066	1120	1066	1066
n070208	1066	1066	1066	1066	1125	1068	1068
n080101	1068	1068	1068	1068	1068	1068	1068
n080108	1068	1068	1068	1069	1069	1069	1069
n080201	1069	1069	1069	1069	1069	1069	1069
n080208	1069	1070	1070	1071	1071	1125	1071
n080301	1071	1146	1146	1146	1146	1146	1157
n080308	1157	1157	1157	1125	1157	1146	1146
n080401	1138	1146	1138	F	F	1095	1095
n080408	1096	1096	1096	1096	1098	1099	F
n090101	+	+	+	+	+	+	+
n090108	+	+	+	+	+	+	+
n090201	+	+	+	+	+	+	+
n090208	+	+	+	+	+	+	+

F = free R = reserved . = empty - = offline
H = Head (R) = Future reservation I = I/O node + = denied

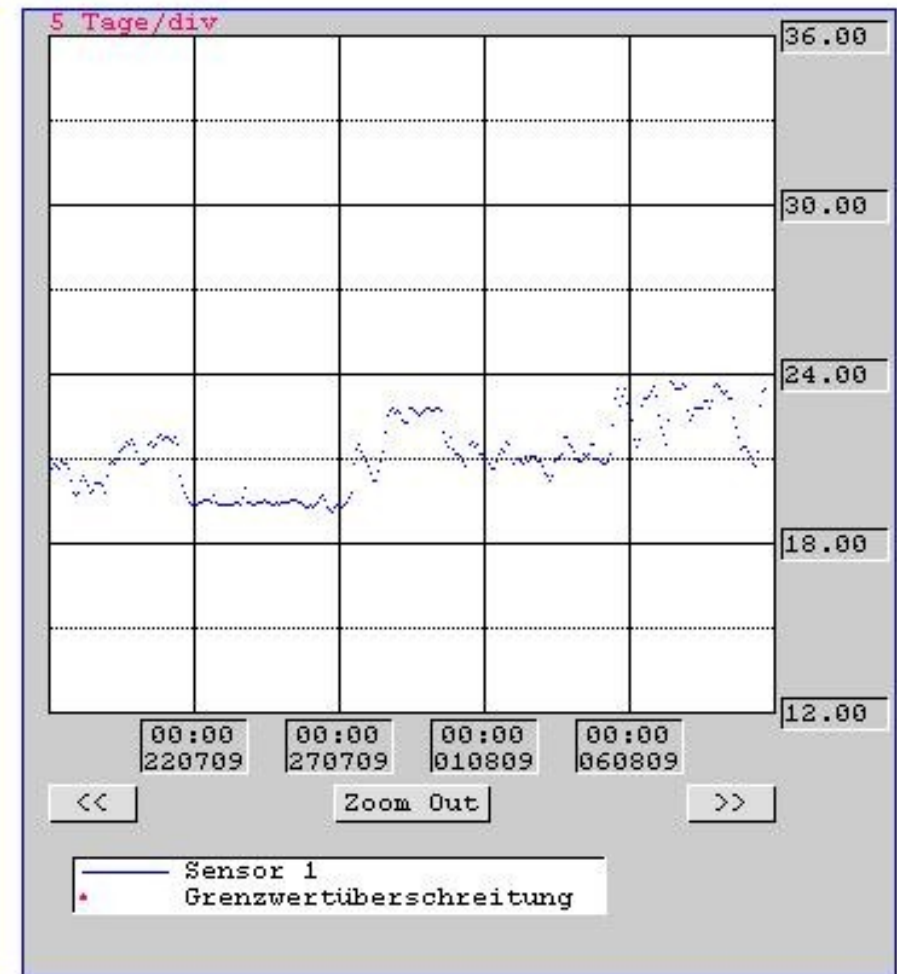
PBS queue 2009

- fast alle Knoten aktiv
- aber noch keine Wartezeiten

ID	user	name	nodes	procs	state	start [h]	remain [h]
1000	qy005	BondFluc	20	88	R	-5.43	18.57
1048	mbohn	job_scri	2	16	R	-5.43	42.57
1055	dspielma	thi-10x1	8	64	R	-5.43	42.57
1059	dspielma	80-3-axg	8	64	R	-5.43	42.57
1063	dspielma	80-3-axg	8	64	R	-5.43	42.57
1064	mbohn	job_bfm	12	96	R	-5.43	42.57
1065	mbohn	job_scri	2	16	R	-5.43	42.57
1066	mbohn	job_bfm7	12	96	R	-5.43	42.57
1068	mbohn	job_bfm1	12	96	R	-5.43	42.57
1069	mbohn	job_bfm4	12	96	R	-5.43	42.57
1070	dspielma	thi-12x1	2	16	R	-5.43	42.57
1071	dspielma	thi-12x1	4	32	R	-5.43	42.57
1095	dspielma	thi-8x12	2	16	R	-5.43	42.57
1096	dspielma	thi-8x12	4	32	R	-5.43	42.57
1097	qy005	BondFluc	20	80	R	-5.54	18.46
1098	mhbw0144	22Fe.Gru	3	24	R	-5.43	6.55
1099	mhbw0144	22Fe.Gru	3	24	R	-5.43	6.55
1100	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1101	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1102	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1103	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1105	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1106	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1108	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1109	mhbw0144	25Fe.Gru	4	32	R	-5.43	6.55
1110	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1111	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1112	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1113	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1114	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1115	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1117	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1118	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1119	mhbw0144	26Fe.Gru	4	32	R	-5.43	6.55
1120	dspielma	80-3-axg	8	64	R	-5.32	42.68
1121	joschaef	sratele8	15	120	R	-5.27	34.73
1125	rr7	STDIN	12	96	R	-5.15	2.85
1138	dspielma	thi-16x1	2	16	R	-4.78	43.22
1142	dspielma	thi-16x1	4	32	R	-1.85	46.15
1146	joschaef	600K_d15	8	64	R	-3.04	36.96
1153	dspielma	thi-14x1	4	32	R	-0.90	47.10
1154	dspielma	thi-16x1	4	32	R	-0.89	47.11
1155	dspielma	thi-12x1	2	16	R	-0.28	47.72
1156	dspielma	thi-12x1	4	32	R	-0.27	47.73
1159	dspielma	thi-16x1	2	16	R	-0.24	47.76
1164	dspielma	thi-14x1	2	16	R	-0.18	47.82
1165	joschaef	bmodul_d	5	40	R	-0.03	19.97

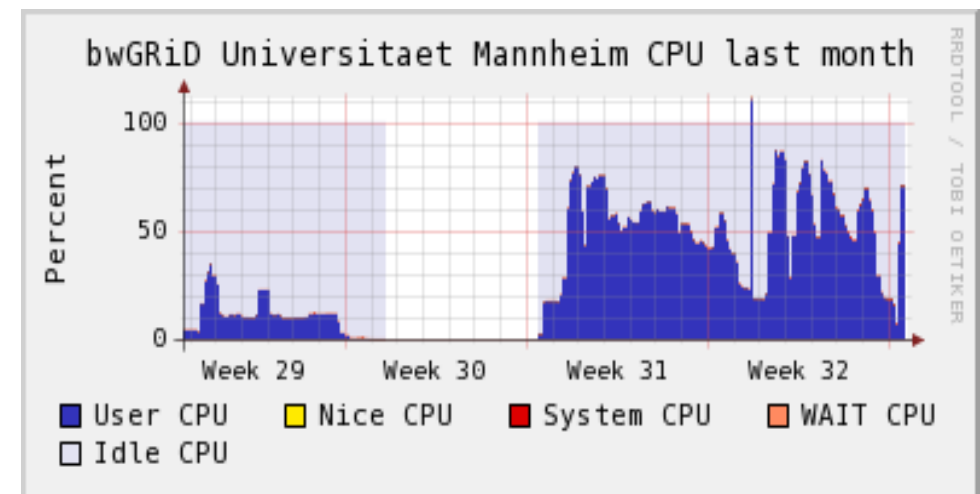
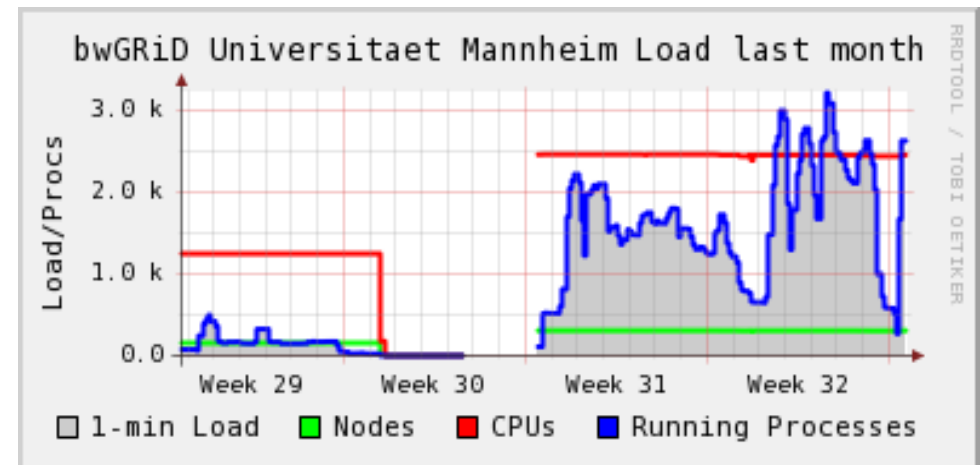
Temperatur 2009

- verlauf im letzten Monat in Mannheim

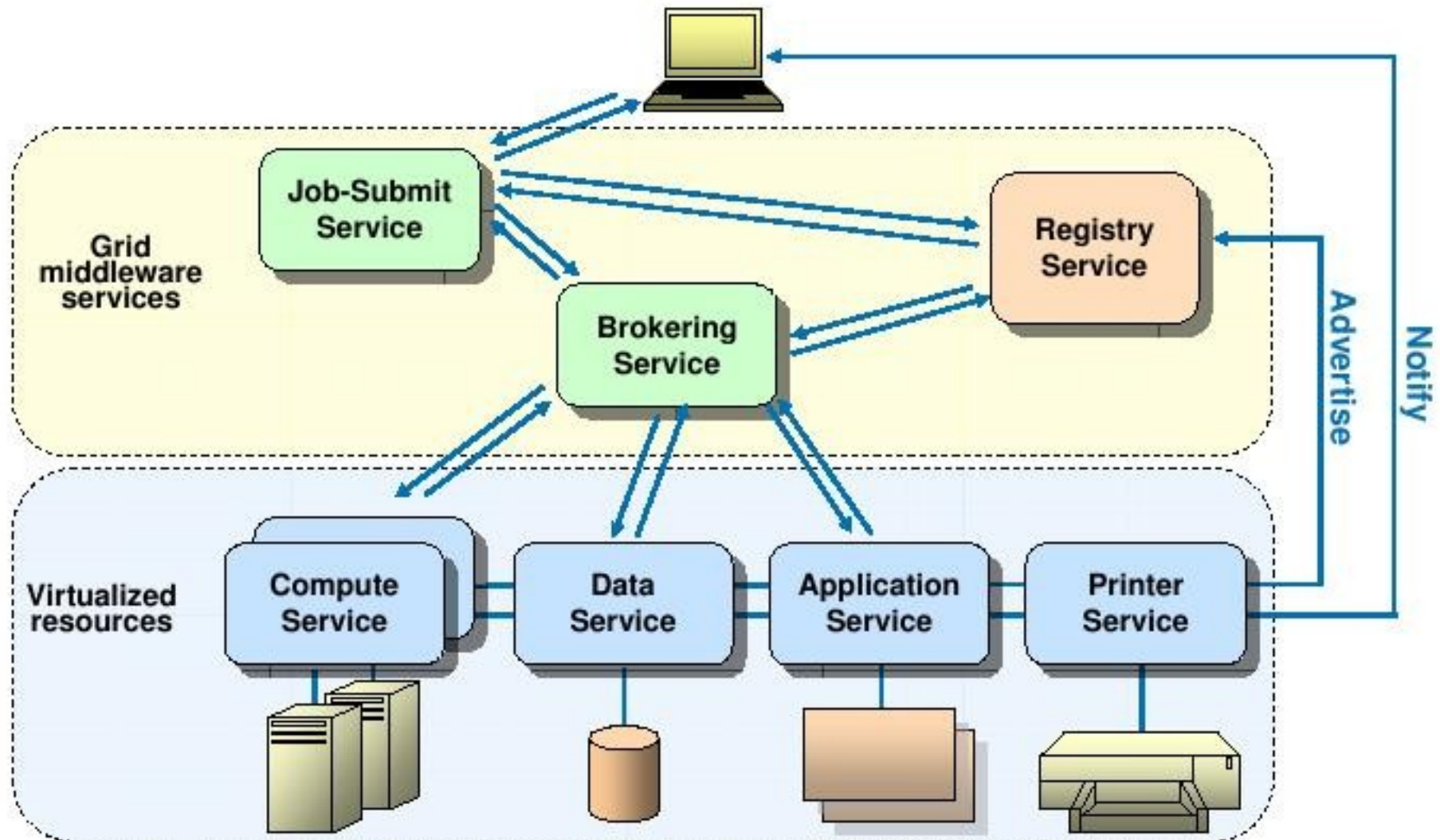


Ganglia Report

- CPU load 1-min
- Anzahl Prozesse
- User CPU Usage



Middleware und Ressourcen



Batch-System und Middleware

- Konfiguration des Batchsystems
 - mehrere Queues mit verschiedenen Qualitäten
- Installation der Grid-Middleware
 - Globus-Toolkit (DoE-, NSF-Projekt): betriebsbereit
 - UNICORE (EU-Projekt): in Vorbereitung
 - LCG/gLite (LHC CERN Projekt): in Vorbereitung
- Benutzer-Zugangsrechner
 - `bwfs.grid.uni-mannheim.de` (SSH)
 - `gtbw.grid.uni-mannheim.de` (GSI-SSH)
- Schulung und Information der Benutzer

http://www.uni-mannheim.de/rum/ag/zs/bw_grid_cluster

Virtuelle Organisationen

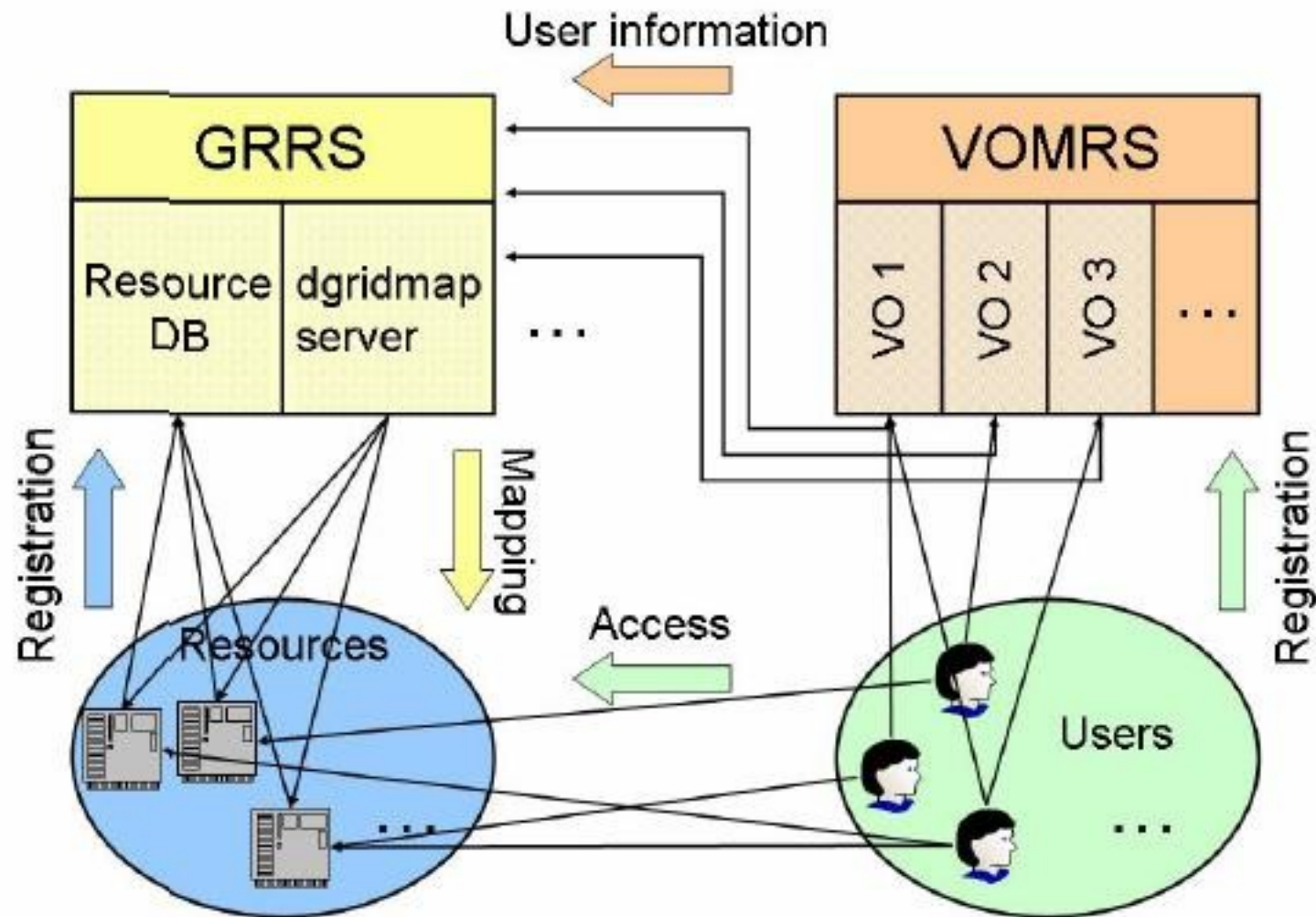


Abbildung 2 Architekturbild Betrieb der VO- und Nutzerdienste und der Ressourcen

d-grid Integration

- D-Grid Virtuelle Organisation (VO) aufsetzen
- Mitglied in VO werden, Grid-Zertifikate
- D-Grid Ressourcen aufsetzen und registrieren
- D-Grid Benutzerverwaltung
- Anbindung an zentralen Storage von bw-Grid
 - Projekt in 2009 HBFG Antrag genehmigt



D-Grid

[AstroGrid](#) [BW-Grid](#) [C3-Grid](#) [DGI](#) [FinGrid](#) [HEP](#) [InGrid](#) [MediGrid](#) [PartnerGrid](#) [TextGrid](#) [WISNET](#)

GridCSM

euqC2W

D-Grid
gesamt

[Drill-Down](#)

[Computing
Elemente](#)

[Cluster
Workernodes](#)

[Siteinfo und
Wartung](#)

[Topologie](#)

[Jawari
Benchmarking](#)

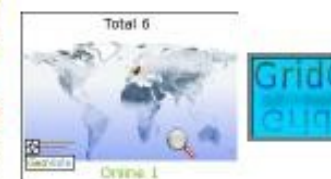
[Details
XML-Dump](#)

Dynamisch erstellte Ressourcenanbieter-Karte



- normaler Betrieb
- geplante Wartung
- Wartung
- Fehler

[Geomaint MDS Dynamic Ma](#)





D-Grid-Resources

D-Grid Compute Resources							
Site	Name	UNICORE5	UNICORE6	GLOBUS	GLITE	Interaktiv	joined VOs Administrator
desy	grid	N	N	N	grid-ce3.desy.de		➡ Birgit Lewendel
dlr	aerogrid	N	Y	N	N	aerogrid.dlr.de	➡ Gert Ohme
fzj	juggle	Y	Y	juggle-glob.fz-juelich.de	juggle-glite.fz-juelich.de	juggle-inter.fz-juelich.de	➡ Christa Dohmen Mathilde Romberg Otto Buechner
fzj	softcomp	Y		n	n		➡ Christa Dohmen
fzk	dgiref	Y	N	dgiref-globus.fzk.de	dgiref-glite.fzk.de	dgiref-login.fzk.de	➡ Dimitri Nilsen Foued Jrad
fzk	gridka	Y	N	gt4-fzk.gridka.de	ce-fzk.gridka.de	dgrid-fzk.gridka.de	➡ Foued Jrad Ingrid Schaeffner
fzk	opus	Y	N	iwrqt4.fzk.de	dgrid-ce.fzk.de	iwrtdgrid.fzk.de	➡ Olaf Schneider Foued Jrad
fzk	sx-8	Y	N	N	N		➡ Foued Jrad
gns	paportal.gns-systems.de	N	N	paportal.gns-systems.de	N	paportal.gns-systems.de	➡ Jan Niemann
gsi	gsi01	N	N	lxglobus.gsi.de	grid13.gsi.de		➡ Kilian Schwarz
gwdg	mediagrid-srv	Y	N	mediagrid-srv.gwdg.de	ce-goegrid.gwdg.de		➡ Dr. Ulrich Schwarzmann
hlrs	bwgrid-hlrs	N	N	gt4bw.dgrid.hlrs.de	N		➡ Michael Schliephake Florian Niebling
hlrs	vec-sx9	Y	N	N	N		➡ Michael Schliephake
iao	stuttgart.iao.fraunhofer.de	N	N	stuttgart-globus.iao.fraunhofer.de	N		➡ Oliver Strauss
izbi	aprilia	N	N	aprilia.izbi.uni-leipzig.de	N		➡ Michael Hartung
izbi	portal1	N	N	portal1.mediagrid.izbi.uni-leipzig.de	N		➡ Michael Hartung
izbi	portal2	N	N	portal2.mediagrid.izbi.uni-leipzig.de	N		➡ Michael Hartung
kit	bwgrid-cluster	Y	N	bwgrid-globus.scc.kit.edu	N	sccbwgrid1.fzk.de	➡ Dimitri Nilsen
lrz	cluster64bit	Y	N	lxgt2.lrz-muenchen.de	N	lxgt2.lrz-muenchen.de	➡ Helmut Heller
lrz	lrz-lcg	N	N	N	lcg-lrz-ce2.grid.lrz-muenchen.de		➡ Helmut Heller
mpe	gavosrv1	N	N	gavosrv1.mpe.mpg.de	N		➡ Arthur Carlson
mppmu	grid-lcgce	N	N	N	grid-lcgce.rzg.mpg.de		➡ Stefan Kluth
offis	wisent	Y	N	srvgrid01.offis.uni-oldenburg.de	srvgrid02.offis.uni-oldenburg.de	srvgrid01.offis.uni-oldenburg.de	➡ Guido Scherp

D-Grid Storage Resources

Site	Name	Kind	joined VOs	Administrator
aip	astrodata01.gac-grid.org	OGSA-DAI	→	Harry Enke
aip	astrodata07.gac-grid.org	DCache	→	Harry Enke
awi	c3rc2	DCache	→	Joerg Matthes
desy	dcache-se-atlas	DCache	→	Birgit Lewendel
desy	dcache-se-cms	DCache	→	Birgit Lewendel
dkrz	c3rc	DCache	→	Stephan Kindermann
dwd	c3rc.dwd	DCache	→	Sven Richter
fzj	dcache	Dcache	→	Otto Buechner
fzj	ogsa-dai	OGSA-DAI	→	Otto Buechner
fzj	testse	generic SE	→	Christina Dohmen
fzk	dgiref-dcache	DCache	→	Xavier Mol
fzk	dgiref-ogsadai	OGSA-DAI	→	Foued Jrad
fzk	gridka-dcache	DCache	→	Doris Ressmann
gwdg	textgrid-srv	OGSA-DAI	→	Andreas Aschenbrenner Dr. Ulrich Schwardmann
kit	bwgridse	generic SE	→	Artem Trunov
lrz	dcache	DCache	→	Helmut Heller
mppmu	grid-srm	DCache	→	Stefan Kluth
rrzn	dcached1	DCache	→	Harald Schwier
udo	dcache.grid.uni-dortmund.de	DCache	→	Stefan Freitag
udo	dgrzr-dcache	DCache	→	Stefan Freitag
udo	dgrzr-ogsadai	OGSA-DAI	→	Stefan Freitag
udo	ogsa-dai.grid.uni-dortmund.de	OGSA-DAI	→	Stefan Freitag
zah	alnitak	OGSA-DAI	→	Peter Schwekendiek
zaik	c3rc.zaik	DCache	→	Volker Winkelmann
zib	casim.zib.de	DCache	→	Stefan Wollny



D-Grid

[AstroGrid](#) [BW-Grid](#) [C3-Grid](#) [DGI](#) [FinGrid](#) [HEP](#) [InGrid](#) [MediGrid](#) [PartnerGrid](#) [TextGrid](#) [WISENT](#)

GridCSM
GRIFFIN

**D-Grid
gesamt**

[Drill-Down](#)

[Computing
Elemente](#)

[Cluster
Workernodes](#)

[Siteinfo und
Wartung](#)

[Topologie](#)

[Jawari
Benchmarking](#)

[Details
XML-Dump](#)

Verfügbarkeit der Computingelemente

Diese Liste zeigt die Verfügbarkeit aller [Computingelemente](#) (CE) / Job Queues und die Anzahl der vom jeweiligen Batchsystem gemeldeten Jobs.

Legende

	CE hat Jobs in der Queue
	CE ist bereit
	CE ist abgeschaltet oder in Wartung

CE			Status			CPUs		Policy	
Endpunkt	Name/Queue	Typ	Laufende Jobs	Wartende Jobs	Status	Freie CPUs	CPUs gesamt	Max. Jobs	Max. WallClockTime
srvgrid01.offis.uni-oldenburg.de	dgipar_ib	PBS	0	0	enabled	5	45	10	3 Tage
srvgrid01.offis.uni-oldenburg.de	default	PBS	0	0	enabled	5	45	10	
srvgrid01.offis.uni-oldenburg.de	dgipar	PBS	0	0	enabled	5	45	18	3 Tage
srvgrid01.offis.uni-oldenburg.de	dgiseq	PBS	160	0	enabled	5	45	18	3 Tage
srvgrid01.offis.uni-oldenburg.de	ib	PBS	0	0	enabled	5	45	10	
srvgrid01.offis.uni-oldenburg.de	test	PBS	1	0	enabled	5	45	-1	3 Tage
srvgrid01.offis.uni-oldenburg.de	verylong	PBS	1	0	enabled	5	45	18	
gtbw.grid.uni-mannheim.de	default	Fork	0	0	enabled	0	8	-1	
gtbw.grid.uni-mannheim.de	single	PBS	1	0	enabled	57	2464	-1	5 Tage
gtbw.grid.uni-mannheim.de	itptest	PBS	0	0	enabled	57	2464	-1	24 Std
gtbw.grid.uni-mannheim.de	batch	PBS	0	0	enabled	57	2464	-1	
gtbw.grid.uni-mannheim.de	itp	PBS	5	0	enabled	57	2464	-1	2 Tage
gtbw.grid.uni-mannheim.de	testkoppl	PBS	0	0	enabled	57	2464	-1	24 Std
gtbw.grid.uni-mannheim.de	normal	PBS	39	11	enabled	57	2464	-1	2 Tage

GSI SSHTerm

- SSH Fenster
- in Java implementiert
- stand-alone oder Applet
- mit Zertifikaten
- Zuordnung zu lokaler Kennung per Grid-Map

The screenshot shows the GSI SSHTerm application window. The title bar includes 'File', 'Edit', 'View', 'Tools', 'Proxy', and 'Help'. The main area displays a grid map with nodes listed on the left and their status in a grid. The status is color-coded: green for free, red for reserved, blue for empty, and grey for offline. Some nodes are marked with a '+' for denied.

Node	1	2	3	4	5	6	7	8
n070108	F	F	F	F	F	F	F	F
n070201	F	F	F	F	F	F	F	F
n070208	F	F	F	F	1201	1236	1236	1236
n080101	1236	1236	1236	1236	1236	1236	1236	1236
n080108	1236	1236	1236	1236	1236	1236	1236	1236
n080201	1236	1236	1236	1236	F	F	F	F
n080208	1258	1249	1256	1249	1249	1201	1249	1249
n080301	1255	1236	F	1236	1236	F	1201	1201
n080308	1201	1201	1255	1236	F	F	F	F
n080401	1138	1255	1138	F	1255	F	F	F
n080408	F	F	F	1236	F	F	F	F
n090101	1205	1205	1225	1225	+	+	+	+
n090108	+	+	+	+	+	+	+	+
n090201	+	+	+	+	+	+	+	+
n090208	+	+	+	+	+	+	+	+

Legend:

- F = free
- R = reserved
- . = empty
- = offline
- H = Head
- (R) = Future reservation
- I = I/O node
- +

Summary:

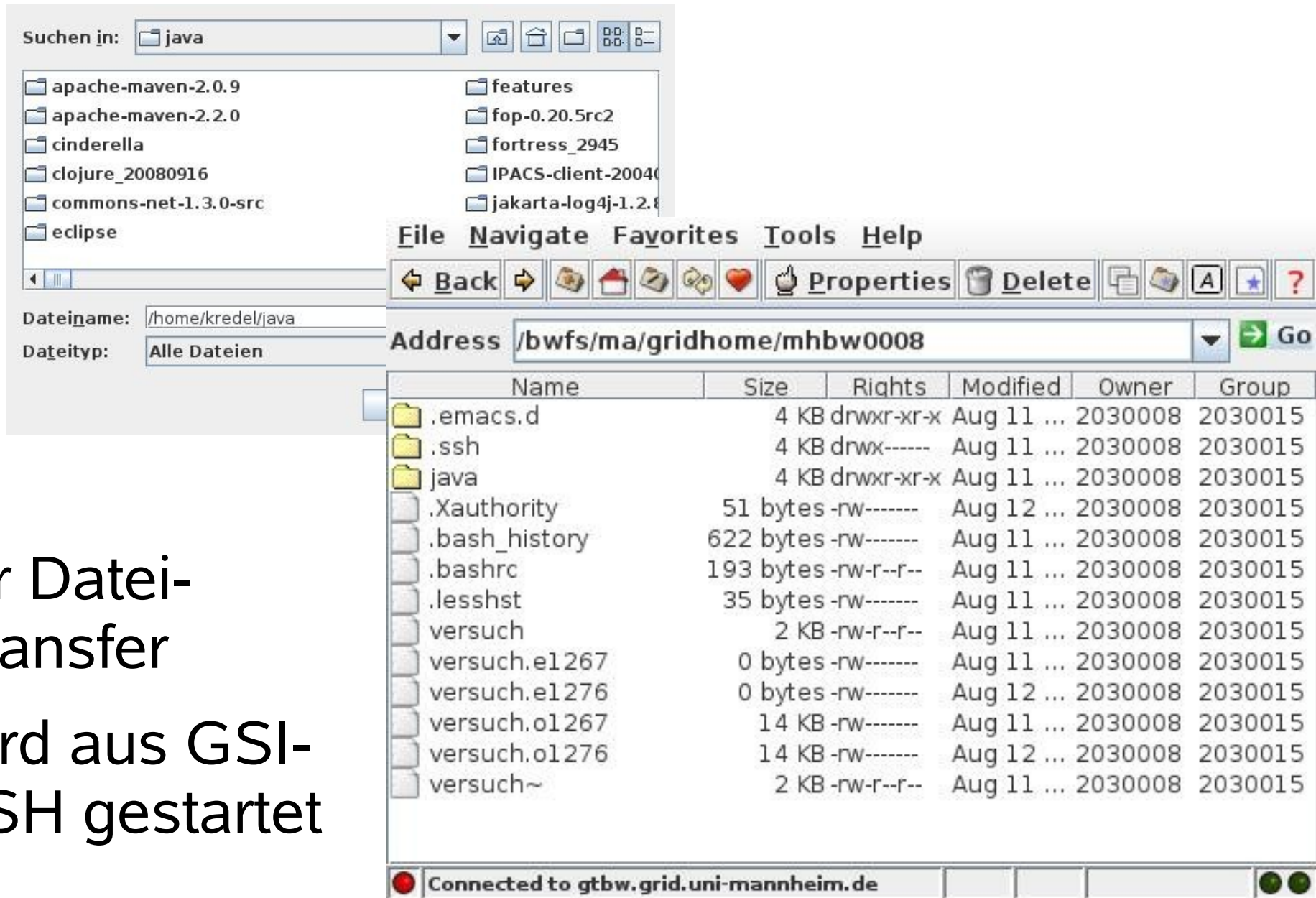
- 0/308 nodes reserved (0.0%)
- 32/308 nodes denied (10.4%)
- 151/308 nodes used (49.0%)
- 1/308 nodes offline (0.3%)
- 124/308 nodes free (40.3%)

Terminal output:

```
[mhbw0008@frbw2 ~]$
[mhbw0008@frbw2 ~]$
```

Status bar: gtbw.grid.uni-mannheim.de:2222 | Unknown | Connected

GSI SFTP



Suchen in:

File Navigate Favorites Tools Help

Back Forward Home Up Down Refresh Properties Delete

Dateiname: Address Go

Dateityp:

Name	Size	Rights	Modified	Owner	Group
.emacs.d	4 KB	drwxr-xr-x	Aug 11 ...	2030008	2030015
.ssh	4 KB	drwx-----	Aug 11 ...	2030008	2030015
java	4 KB	drwxr-xr-x	Aug 11 ...	2030008	2030015
.Xauthority	51 bytes	-rw-----	Aug 12 ...	2030008	2030015
.bash_history	622 bytes	-rw-----	Aug 11 ...	2030008	2030015
.bashrc	193 bytes	-rw-r--r--	Aug 11 ...	2030008	2030015
.lessht	35 bytes	-rw-----	Aug 11 ...	2030008	2030015
versuch	2 KB	-rw-r--r--	Aug 11 ...	2030008	2030015
versuch.e1267	0 bytes	-rw-----	Aug 11 ...	2030008	2030015
versuch.e1276	0 bytes	-rw-----	Aug 12 ...	2030008	2030015
versuch.o1267	14 KB	-rw-----	Aug 11 ...	2030008	2030015
versuch.o1276	14 KB	-rw-----	Aug 12 ...	2030008	2030015
versuch~	2 KB	-rw-r--r--	Aug 11 ...	2030008	2030015

Connected to gtbw.grid.uni-mannheim.de

- für Datei-Transfer
- wird aus GSI-SSH gestartet

GSI SSH ins bw-GRiD

The screenshot displays a terminal window with a menu bar (File, Edit, View, Tools, Proxy, Help) and a toolbar. The terminal content is as follows:

```

ne
H = Head    I = I/O node    + = denie

0/520 nodes reserved (0.0%)
2633/520 nodes used (506.3%)
21/520 nodes offline (4.0%)
151/520 nodes free (29.0%)

stbw0008@gt4bw:~> java -version
java version "1.5.0_11"
Java(TM) 2 Runtime Environment, Standard Edition (build 1.5.0_11-b03)
Java HotSpot(TM) Client VM (build 1.5.0_11-b03, mixed mode, sharing)
stbw0008@gt4bw:~>

0/136 nodes reserved
0/136 nodes used
29/136 nodes offline
101/136 nodes free

tubw0008@gt4:~>
tubw0008@gt4:~>
tubw0008@gt4:~> java -version
java version "1.6.0_13"
Java(TM) SE Runtime Environment (build 1.6.0_13-b03)
Java HotSpot(TM) 64-Bit Server VM (build 11.3-b02, mixed mode)
tubw0008@gt4:~>

[ulbw0008@koios ~]$ java -version
java version "1.6.0_13"
Java(TM) SE Runtime Environment (build 1.6.0_13-b03)
Java HotSpot(TM) 64-Bit Server VM (build 11.3-b02, mixed mode)
[ulbw0008@koios ~]$
  
```

At the bottom, a status bar shows three connection entries:

Host	Port	Status	Connected
gt4bw.dgrid.hlrs.de	2222	Unknown	Connected
gt4.uni-tuebingen.de	2222	Unknown	Connected
koios.rz.uni-ulm.de	2222	Unknown	Connected

Software

- Cluster Programmierungs Tools
 - MPI
 - OpenMP
 - Intel Compiler
- Anwendersoftware
 - Matlab, Octave, Scilab
 - Stata, R
 - geplant: Maple, Mathematica (Grid)

module System

wichtige Optionen:

- module avail
- module show <path>
- module load <path>
- module unload <path>
- module list

```
----- /opt/bwgrid/modulefiles -----
bio/biopython/1.49-python-2.4
compiler/cofort/0.91
compiler/gnu/4.3
compiler/intel/10.1
compiler/sun/12
compiler/upc/gcc4.1.1/mvapich1.0.1
compiler/upc/gcc4.1.1/openmpi1.2.7
compiler/upc/icc10.1/mvapich1.0.1
compiler/upc/icc10.1/openmpi1.2.7
debugger/ddt/2.3.1
math/R/2.7.2
math/R/2.8.1
math/R/2.9.0
math/matlab/R2008b
math/octave/3.0.3
math/scilab/4.1.2
math/singular/3.1.0
math/stata/10
mpi/mpiexec/0.83
mpi/mvapich/1.1-gnu-4.1
mpi/mvapich/1.1-intel-10.1
mpi/mvapich2/1.2p1-gnu-4.1
mpi/mvapich2/1.2p1-intel-10.1
mpi/openmpi/1.2.7-gnu-4.1
```


Vor- und Nachteile von Cluster-Systemen

- ✗ Lernaufwand für Entwicklung paralleler Programme und Batch-System
- ✗ Lernaufwand für die Grid-Middleware und Einrichtung der Grid-PKI Anforderungen
- ✓ State-of-the-art Plattform für Mannheim
- ✓ Zugang zu (fast) allen D-Grid-Ressourcen
- ✓ Auslastung unseres Clusters wird höher
 - seine Programme muss jede(r) immer noch selbst parallelisieren
 - Lizenzen für kommerzielle Software können recht teuer werden

Vielen Dank für die Aufmerksamkeit!

- Fragen? Bemerkungen?
- Links
 - <http://www.bw-grid.de/>
 - <http://web.urz.uni-heidelberg.de/server/grid/>
 - http://www.uni-mannheim.de/rum/ag/zs/bw_grid_cluster/
 - <http://www.hlrs.de/>
 - <http://www.d-grid.de/>