

Das bwGRiD – „High Performance Compute Cluster“ als flexible, verteilte Wissenschaftsinfrastruktur

Marek Dynowski^a, Michael Janczyk^a, Janne Schulz^a, Dirk von Suchodoletz^a
Sven Hermann^b

^a Technische Fakultät / Rechenzentrum
Albert-Ludwigs Universität Freiburg

^b Steinbuch Centre for Computing
Karlsruher Institut für Technologie (KIT)

marek.dynowski@rz.uni-freiburg.de
michael.janczyk@rz.uni-freiburg.de
janne.schulz@rz.uni-freiburg.de
dirk.von.suchodoletz@rz.uni-freiburg.de
sven.hermann@kit.edu

Abstract: Das bwGRiD-Projekt startet 2008 an acht Universitäten in Baden-Württemberg, um Wissenschaftlern aller Fachrichtungen Ressourcen im Bereich des High Performance Computings effizient und hochverfügbar zur Verfügung zu stellen. Im Vordergrund steht der Aufbau einer dezentralen Grid-Struktur, bei der homogene Parallelrechner-Cluster transparent zu einem Grid-Verbund gekoppelt werden. Das Projekt soll die Machbarkeit und den Nutzen von Grid-Konzepten für die Wissenschaft nachweisen und bisherige Organisations- und Sicherheitsproblematiken überwinden. Die Grid-Struktur ermöglicht eine Spezialisierung der einzelnen Rechenzentren im Anwendungs- und Hardwarebereich, sowie die Entwicklung neuer Cluster- und Softwarewerkzeuge. Die durch diese Struktur entstehende Lizenzproblematik für proprietäre Software soll im Rahmen dieses Projektes gelöst werden. Durch den kontinuierlichen Ausbau der lokalen bwGRiD-Cluster und die Integration neuer Standorte kommt es zu einer wachsenden Heterogenität, welche durch die stetige Weiterentwicklung von Software und Konzepten überwunden werden muss, um eine maximale Kompatibilität zwischen den Standorten zu gewährleisten. Die Hardware des Projektes wurde vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der D-Grid-Initiative und die Personalstellen vom Ministerium für Wissenschaft, Forschung und Kunst (MWK) Baden-Württemberg finanziert. In diesem Artikel werden Konzepte, Erfahrungen und Resultate des bisherigen bwGRiD-Projektes vorgestellt.

1 Das bwGRiD – ein Community-Grid im Südwesten

Die Konzepte der Compute- und Daten-Cloud werden im Zusammenhang mit wissenschaftlichem Rechnen zunehmend diskutiert und umgesetzt. Die hierbei formulierten Ziele wie Resource-On-Demand, schnelles Deployment, flexible Anpassung an bestimmte Nutzerwünsche, Virtualisierung von Ressourcen und geographische Verteilung sind oft

gar nicht so neu und vielfach bereits Realität [BK10]. Das bwGRiD ist Teil der D-Grid-Initiative [Gen06] der Bundesregierung und wurde mithilfe einer Infrastrukturförderung des BMBF und mit zusätzlicher finanzieller Unterstützung des Landes Baden-Württemberg auf Basis des Landeshochschulnetzes BelWue realisiert. Der gewählte Ansatz ist eine dezentrale Struktur, bei der die über das Land verteilten Cluster mit Hochgeschwindigkeitsdatenleitungen vernetzt werden.

Die Gründungsmitglieder des bwGRiD-Projektes sind, zunächst unter Federführung des Stuttgarter Höchstleistungsrechenzentrums (HLRS), die Universitäten Freiburg, Heidelberg, Karlsruhe, Konstanz, Mannheim, Tübingen und Ulm. Später schließt sich Esslingen als erste Hochschule dem Projekt an [RV10]. Das Ziel des Projektes ist die Untersuchung von „High Performance Computing“-Grids (HPC) im Hochschulumfeld hinsichtlich ihres Nutzens, sowie die Identifikation der verschiedenen Anforderungen an solche Umgebungen [Mic06]. Zu diesem Zweck werden die lokalen HPC-Cluster für die Nutzer transparent zu einem Verbund zusammengeschlossen. Diese Struktur soll sicherstellen, dass verfügbare Computing-Ressourcen für die akademische Forschung standortunabhängig und hochverfügbar bereitstehen. Geprägt wird diese Infrastruktur durch gemeinsame Standards beim Zugriff auf die Ressourcen, Authentifizierung und der Verteilung der Jobs. Dabei wird keine durchgehend einheitliche Hardware gefordert, jedoch wurden zu Projektbeginn identische Komponenten gemeinsam beschafft, um die Vorteile eines solchen Vorgehens zu evaluieren. Dieser Ansatz löst das klassische, zentralisierte HPC ab. Nicht einzelne Forschungseinrichtungen oder Arbeitsgruppen mit Bedarf an Rechenleistung beschaffen und betreiben ihre eigenen Cluster. Stattdessen werden die Aufwendungen in einen gemeinsamen Pool überführt, um diese in Summe effektiver einsetzen zu können. Das Konzept erlaubt neben einem verbesserten Support durch eine größere Anzahl von Administratoren und Entwicklern auch die HPC-Versorgung mit unterschiedlicher Hard- und Software, was auch die Bildung von untereinander vernetzten lokalen Kompetenzzentren für wissenschaftliche Fachbereiche ermöglicht. Dadurch kann gezielter auf die Erfordernisse der verschiedenen Wissenschaftsdisziplinen und Forschergruppen eingegangen werden und die daraus resultierenden Aufgaben lassen sich zwischen den Projektpartnern aufteilen. Ferner wird eine höhere Effizienz der Ressourcennutzung durch optimale Auslastung über den gesamten Lebenszyklus der Cluster-Hardware erzielt.

Eine entscheidende Rolle spielt das Betriebskonzept, das auf der einen Seite den Benutzern eine stabile und performante Umgebung sicherstellt, auf der anderen Seite jedoch flexibel auf neue Anforderungen, wie Erweiterungen der Hardwarebasis oder Modifikationen der Softwareausstattung, reagieren muss. Die Projektpartner im bwGRiD entwickeln und prüfen entsprechende Konzepte hinsichtlich einer gemeinsamen Sicherheitsstruktur zwischen den Rechenzentren zur Lösung von Organisations- und Sicherheitsproblematiken. Weitere zentrale Herausforderungen sind zudem der Aufbau einer performanten, verteilten Datenhaltung sowie die Integration unterschiedlicher Nutzerverwaltungen der jeweiligen Standorte. Die Koordination gemeinsamer Beschaffungen der im Projekt vorgesehenen Soft- und Hardware soll die Anschaffungskosten für alle Beteiligten senken, Wartungsverträge optimieren und den gegenseitigen Support zwischen den Rechenzentren erleichtern. Ein weiteres erklärtes Ziel des Projektes liegt in der Unterstützung vielfältiger Forschungsfelder, weshalb dezidiert Mittel für Lizenzen kostenpflichtiger Software bereitgestellt wur-

den. Hierbei wird auch die Lizenzproblematik für den Grid-Verbund analysiert und evaluiert. Des Weiteren werden die zur Verfügung stehenden Ressourcen für eine Optimierung von Grid-Konzepten für Forschergruppen die bereits im HPC-Bereich aktiv sind genutzt. Auch hat in den letzten Jahren das Hochleistungsrechnen in vielen nicht-klassischen HPC-Disziplinen erheblich an Bedeutung gewonnen. Daher kommt dem Heranführen solcher Forschergruppen an „High Performance“- und Grid-Computing, durch innovative und speziell für diese Gruppen entwickelte Werkzeuge, Hilfestellungen und Dokumentationen, eine außerordentliche Bedeutung in dem Projekt zu.

2 Verteilte Hardware – verteilte Systeme

Das bwGRiD, zunächst als ein dezentrales Grid homogener Cluster geplant und beschafft, wurde im Laufe der Zeit aufgrund von Nutzeranforderungen und Investitionen von Arbeitsgruppen an einigen Standorten heterogen ausgebaut (siehe Abbildung 1). Die Hauptinvestition bestand zunächst aus 101 Bladecentern des Typs H von IBM, die jeweils 14 Blades des Typs HS21 XM¹ fassen. Im Primärausbau des bwGRiDs entspricht das einer Gesamtzahl von 11312 Kernen und 22624 GByte RAM. Um die Rechenleistung des Clusters vollständig nutzen zu können, sind die Rechenknoten an den Standorten mit einem 20 Gbit/s-InfiniBand-Netzwerk verbunden.² Untereinander sind die Cluster über das BelWue-Netz mit 10 Gbit/s angeschlossen. Dies ermöglicht einen schnellen Datentransfer und gewährleistet dem Nutzer eine hohe Flexibilität bei der Auswahl des jeweiligen Standortes.

Die 101 Bladecenter wurden, mit Ausnahme von Stuttgart, gleichmäßig an alle teilnehmenden Standorte verteilt (Stuttgart erhielt 31 Bladecenter). Die restlichen Standorte wurden mit je 10 Bladecentern ausgestattet, wobei der Cluster in Konstanz mit dem Cluster am Standort Ulm vereint wurde. Die Heidelberger und Mannheimer Cluster wurden 2009 über eine InfiniBand-Glasfaser-Kopplung miteinander verbunden [RHKK10]. Esslingen investierte 2009 in 180 Appro-Bladeserver gB222X³ von NEC. Damit wuchs das bwGRiD auf eine Gesamtzahl von 1594 Rechenknoten mit 12752 Kernen und 26944 GByte RAM an.

Im HPC-Bereich ist eine effiziente Nutzung der Ressourcen unerlässlich. Neue Hardware muss zeitnah eingegliedert werden, da mehrere Faktoren wie Garantie, Abschreibung und Auftraggeber auf eine möglichst schnelle Nutzung der Ressource hin drängen. Hierbei hat sich eine zentrale Installation, die über das Netzwerk verteilt wird, bewährt [SvW⁺11]: Änderungen am Betriebssystem können zentral eingepflegt und in kürzester Zeit an alle Maschinen verteilt werden. Außerdem können alternative Linux-Betriebssysteme angeboten werden, sodass schnell zwischen den Umgebungen gewechselt werden kann. Als Betriebssystem lief zunächst ein Scientific Linux Version 5.0, das im Laufe der Zeit auf die Version 5.5 aktualisiert wurde.⁴ Ein Upgrade des Betriebssystems auf Version 6.x wird

¹Zwei Vierkern-CPU's Intel Harpertown Xeon E5440 (2.83 GHz) und 16 GByte RAM.

²Die Knoten sind mit einer Mezzanine-Karte ConnectX von Mellanox ausgestattet. Verbunden sind sie über einen InfiniBand-Switch Voltaire Grid Director ISR 2012.

³Zwei Intel CPU's Gainestown Xeon E5520 bzw. X5560 und 24 GByte RAM.

⁴Ein RHEL-Derivat. Projektseite: <http://www.scientificlinux.org/>

derzeit evaluiert.

Zur Verbesserung der Performance wurden an den meisten Standorten Festplatten für temporäre Daten nachgerüstet. Für große Datenmengen steht ein paralleles Netzwerkdateisystem zur Verfügung. Die Ausschreibung gewann eine HP-Speicherlösung, die auf Lustre 1.8.3 aufsetzt⁵ und auf hohe Redundanz ausgelegt ist. Stuttgart, Ulm und Tübingen wurden mit je 64 TByte Speicherplatz, die anderen Standorte mit je 32 TByte Speicherplatz ausgestattet. In Esslingen wurde 36 TByte LustreFS-Speicher der Firma NEC beschafft. Der parallele Lustre-Speicher ist über InfiniBand mit den Knoten verbunden. Untereinander sind die Server, welche die Festplatten ansprechen, über 10 Gbit/s FibreChannel vernetzt.

Der Zugang zum Grid und die Jobsubmission erfolgen über die Globus-Middleware.⁶ Jedoch können Jobs auch lokal am Cluster abgeschickt werden. Der Login an einem Cluster erfolgt für Grid-Nutzer über `GSISSH`⁷ und für lokale Nutzer via `SSH`.

3 Geschäftsmodell und Organisation

Das bwGRiD soll ein möglichst breites Spektrum an Anwendern und Anwendergruppen ansprechen. Die Nachteile, die durch die Beschaffung und den Betrieb von HPC-Ressourcen durch einzelne Arbeitsgruppen oder im Zuge einzelner kleiner Projekte entstehen, sollen dadurch vermieden werden. Synergieeffekte durch die gemeinsame Beschaffung von identischer Hardware führen zu einer effizienteren Nutzung monetärer Mittel. So war es unter anderem möglich, günstigere Großkundenkonditionen zu erhalten, was sich neben der Beschaffung auch auf die Konditionen von Wartungsverträgen positiv auswirkt. Weiterhin erleichtert eine homogene Basisausstattung die Administration der Systeme, da im Verbund Probleme gelöst und Erfahrungen ausgetauscht werden können. Die Wartung und Konfiguration der Systeme, wie auch das Scheduling werden unproblematischer, da häufig Hilfestellung durch andere Standorte angeboten werden kann. Außerdem wird eine gemeinsame Beschaffung Hardware-spezifischer Software (z.B. Compiler) ermöglicht.

Im Rahmen des Projektes wurden fachspezifische Kompetenzzentren an den jeweiligen Standorten gebildet. Durch die starke Vernetzung der Wissens-Cluster kann eine optimale Nutzerunterstützung koordiniert und gewährleistet werden. Mitarbeiter eines Standortes können Gruppenprofile präziser abschätzen und bieten den Nutzern über lokale und damit kurze Kommunikationswege einen idealen Support. Verschiedene Aufgaben und Verantwortungsbereiche wurden an einzelne Standorte übertragen: So übernahm Karlsruhe im Jahr 2010 die Projektleitung von Stuttgart und Konstanz ist für den Betrieb und die Gestaltung der Projektseite www.bw-grid.de zuständig, die Informationen zum Projekt, die an den jeweiligen Standorten installierte Software, die Zugangsinformationen zu den Clustern und Beschreibungen der auf dem Grid gerechneten Projekte, anbietet. Die interne

⁵Derzeitiger Stand: Lustre Server Version 1.8.4, Lustre Client Version 1.8.5 (www.lustre.org/)

⁶Derzeitige Version: 4.0.8. (<http://www.globus.org/>)

⁷Ein auf GSI (Grid Security Infrastructure) basierender SSH-Client (<http://grid.ncsa.illinois.edu/ssh/>)

Abstimmung der Projektpartner und der Unterprojekte erfolgt über eine E-Mail-Liste, die von Ulm betreut wird, und durch eine zweiwöchentlich stattfindende Videokonferenz mit allen Projektpartnern. Gemeinsame Dokumente werden über die BSCW-Groupware⁸ zur Teamarbeit, die in Stuttgart verwaltet wird, organisiert.

3.1 Einheitliche Softwareausstattung

Um eine möglichst homogene Softwareausstattung im bwGRiD zu gewährleisten, ist eine klare Softwarestruktur zwingend erforderlich. Das Softwarepaket `modules`⁹ wird hierbei genutzt, um eine temporäre Benutzerumgebung für ein bestimmtes Programm durch setzen der entsprechenden Umgebungsvariablen der Linux-Shell, zu erzeugen. Dadurch können mehrere Versionen eines Programms parallel installiert und bei Bedarf von den Anwendern geladen und genutzt werden. Eine einheitliche Benennung der Module ist unabdingbar, um Nutzern das Abschicken ihrer Jobs auf jedem Cluster des bwGRiDs ohne vorherige Anpassung ihrer Skripte zu ermöglichen. Daher wurde ein Standard für die Namensgebung der Modulnamen entwickelt. Weiterhin wurde die Verantwortlichkeit für Entwicklung, Nutzersupport und Pflege der Module auf die einzelnen Standorte ihren Schwerpunkten entsprechend aufgeteilt. Um die Organisation und Nutzung der Module zu vereinfachen, wurden diese in Klassen eingeteilt.¹⁰ Die mit `mandatory` gekennzeichneten Pakete müssen an allen Standorten installiert sein. Bei Software aus dieser Klasse können sich die Nutzer ohne vorherige Überprüfung darauf verlassen, dass diese auf jedem Cluster installiert ist. Andernfalls obliegt es den Nutzern zu prüfen, ob die erforderlichen Module am gewünschten Standort verfügbar sind.¹¹ Angaben zu den Paketen wie Modulname, Softwareversion, verantwortlicher Standort und Klasse werden zurzeit zentral in einer Tabelle gespeichert und die Module zum Download über einen Repository-Server in Freiburg angeboten.

Da sich die Grid-weite Bereitstellung standardisierter Softwaremodule als kritischer Punkt erwiesen hat, wurde im Rahmen der *Ergänzenden Maßnahmen* zum bwGRiD-Projekt die Stelle des Softwarekoordinators geschaffen. Dieser entwickelt Konzepte für eine Qualitätskontrolle der von den Standorten bereitgestellten Softwaremodule.

3.2 Lizenzproblematik

Der Erfolg der bwGRiD-weiten Beschaffung von Softwarelizenzen ist stark von den Anbietern abhängig. Das AMBER-Projekt erkannte das bwGRiD als eine Institution an, wo-

⁸Web-Präsenz: <http://www.bscw.de/>

⁹Projektseite: <http://modules.sourceforge.net/>

¹⁰`mandatory` – garantierte Verfügbarkeit an allen Standorten, `optional` – Installation optional, `local` – Lizenzpflichtige Programme, die nur an bestimmten Standorten verfügbar sind, `private` – kann von interessierten Standorten installiert werden und `on.request` – werden auf Nutzeranfrage zur Verfügung gestellt.

¹¹Web-Maske mit Suchfunktion: <http://www.bw-grid.de/benutzerinformation/software/software-suchen/>

durch eine Lizenz für das gesamte Grid erworben werden konnte. Somit kann die Software ohne Einschränkung an jedem Standort des Grids installiert und genutzt werden. Auch die Anbieter der „Schrödinger Molecular Modeling Suite“ und der „ANSYS Computer Aided Engineering und Multiphysik“-Software erlauben die Nutzung im bwGRiD durch dynamisch abrufbare Lizenztokens.¹² Entsprechende Lizenzserver werden an den Standorten Tübingen (Schrödinger) und Karlsruhe (ANSYS) betrieben. Insbesondere für Arbeitsgruppen, die eine bestimmte Software nicht permanent nutzen, ist dies ein erheblicher Vorteil. Administrativ ergibt sich allerdings ein signifikanter Mehraufwand durch das Betreiben der Lizenzserver, da diese nicht nur eingerichtet werden, sondern auch permanent mit hoher Verfügbarkeit erreichbar sein müssen.

Probleme mit diesem Lizenzmodell ergeben sich auch in Verbindung mit dem verwendeten Batchsystem.¹³ Es ist nicht auszuschließen, dass ein Job startet, obwohl nicht genügend freie Lizenzen vorhanden sind, was zu einem Jobabbruch führt. In diesem Bereich besteht noch erheblicher Handlungsbedarf, da die Lizenzproblematik vom Scheduler derzeit nicht berücksichtigt wird. Für die Intel Compiler Suite wurde ein Preis für alle Standorte ausgehandelt, jedoch handelt es sich um lokale Standortlizenzen, sodass auch die Lizenzserver an den Standorten betrieben werden müssen. Weiterhin gibt es Lizenzvereinbarungen, die auf einen Standort beschränkt sind, aber von allen Nutzern des bwGRiDs genutzt werden können. Mit dem Anbieter der Computerchemie-Software Gaussian konnte trotz intensiver Verhandlungen keine angemessene Grid-weite Lizenz ausgehandelt werden. Letztendlich bleibt zu vermerken, dass lediglich bwGRiD-Lizenzen, wie sie von AMBER bereitgestellt werden, eine Alternative zu Open Source Software darstellen.

3.3 Entwicklung des bwGRiDs

Das Bestreben eines Anbieters von HPC-Diensten ist die permanente Bereitstellung leistungsstarker Hardware für Anwender. Daher wurden die bwGRiD-Cluster an verschiedenen Standorten stetig ausgebaut (Abbildung 1). Gerade kleineren Standorten bietet das bwGRiD eine potente Basis für eine stete Erweiterung ihrer HPC-Ressourcen. Ferner können dadurch auch Kompetenzzentren im Hardwarebereich aufgebaut werden, was die Attraktivität des jeweiligen Standorts und des gesamten bwGRiDs für Anwender erhöht. Insbesondere in Freiburg und Tübingen wurden die bwGRiD-Cluster durch Betreiber- und Anwender-finanzierte Hardware kontinuierlich erweitert. Teilweise ist es gelungen, bestimmte Hardware dem gesamten Grid zur gemeinsamen Nutzung zur Verfügung zu stellen. Im Gegenzug profitieren die Arbeitsgruppen von der fachkundigen Administration ihrer Systeme durch die bwGRiD-Betreiber und dem Know-How der anderen Grid-Standorte sowie von der teilweisen Übernahme der Kosten für den Betrieb. Aktuell können Anwender von Arbeitsgruppen mit eigener Hardware bei Bedarf mit höherer Priorität versehen werden, sodass der Zugriff auf die eigenen Ressourcen jederzeit gewährleistet ist. Durch dieses Modell können auch andere Grid-Nutzer von der zusätzlichen Hardware pro-

¹²Web-Präsenz AMBER <http://ambermd.org/>, Schrödinger <http://www.schrodinger.com/>, ANSYS <http://www.ansys.com/>, Gaussian <http://www.gaussian.com/>

¹³TORQUE und Moab (<http://www.adaptivecomputing.com/>)

Integration neuer Ressourcen			
Standorte	# Knoten	CPU/GPU	RAM (GByte)
Freiburg	16	Intel Xeon X5550, 2.67GHz (2x4 Kerne)	24
	8	Intel Xeon X5650X5550, 2.66GHz (2x6 Kerne) / Nvidia Tesla M2090** (1x512 Kerne)	
	4	Intel Xeon E5520, 2.27GHz (2x4 Kerne) / Nvidia Tesla C1060* (2x240 Kerne)	
Stuttgart	1	AMD Opteron 8360 SE, 2.44GHz (8x4 Kerne)	512
		AMD Opteron 8384, 2.64GHz (8x4 Kerne)	256
	8	Intel Xeon 5472, 3.00GHz (2x4 Kerne) / Nvidia Quadro FX 5800 (240 Kerne)	8
Tübingen	24	Intel Xeon L5530, 2.4GHz (2x4 Kerne)	72
	18	AMD Opteron 6172, 2.1GHz (2x12 Kerne)	32
	16	Intel Xeon 5150, 2.66GHz (2x2 Kerne)	
	8	Intel Xeon 5355, 2.66GHz (2x4 Kerne)	16
		Intel Xeon 5150, 2.66GHz (2x2 Kerne)	
1	Intel Xeon E7-4830, 2.13GHz (4x8 Kerne) / 6,5 TByte Storage	512	

* 2x Tesla S1070 mit je 4x C1060 (2x4 GByte RAM) | ** intern verbaut (6 GByte RAM)

Abbildung 1: **Hardware des bwGRiDs** (Stand: 02.04.2012)

fitieren, wenn die Auslastung durch die priorisierten Nutzer gering ist. In diesem Zusammenhang wird aktuell an weiteren Betriebs-Modellen intensiv gearbeitet.

3.4 Nutzungsrichtlinien

Der Zugang zu den bwGRiD-Clustern erfolgt grundsätzlich über die im D-Grid verwendete Globus-Middleware. Globus unterstützt die Authentifizierung mithilfe von X.509-Zertifikaten. Daher haben alle Standorte des bwGRiDs eine Registration Authority für DFN-Gridzertifikate eingerichtet. Außerdem wurde innerhalb des Projektes beschlossen, dass neben den üblichen Grid-CAs¹⁴ auch eigene CAs der beteiligten Einrichtungen gegenseitig akzeptiert werden. Grundsätzlich können alle Mitglieder der Hochschulen in Baden-Württemberg und deren Projektpartner Mitglied der VO bwGRiD werden. Zusätzlich steht das bwGRiD auch Mitgliedern anderer D-Grid-VOs zur Verfügung. Für die Nutzung der Ressourcen ist lediglich eine Zustimmung zur D-Grid-Einverständniserklärung¹⁵ obligatorisch. In diesem Dokument sind die Rechte und Pflichten der Nutzer von D-Grid-Ressourcen geregelt, insbesondere die Voraussetzungen für das Erlangen eines Grid-Nutzerzertifikats. Zusätzlich kann jeder Standort seinen lokalen Nutzern den Zugang zum jeweiligen Cluster erlauben. Jedoch stehen diesen dann nur die Ressourcen des Standortes zur Verfügung. Die Administratoren des bwGRiDs sind angehalten den Betrieb auf den HPC-Clustern zu überwachen und bei Verstößen gegen die Einverständniserklärung entsprechende Maßnahmen gegen die Nutzer zu ergreifen.

¹⁴Certificate Authorities: Grid-CA des DFN-Vereins, GridKa-CA des Forschungszentrums Karlsruhe

¹⁵Einverständniserklärung: <http://www.fz-juelich.de/dgrid/AUF/D-Grid-User-AUP.pdf>

3.5 Hilfswerkzeuge zur Unterstützung der Nutzer

Die effiziente Nutzung verteilter Rechenumgebungen wie die des bwGRiDs erfordert, dass Anwendern und insbesondere Einsteigern adäquate Lehrmaterialien und Softwarewerkzeuge zur Verfügung gestellt werden. Zunächst muss eine umfassende Dokumentation der jeweiligen HPC-Cluster erfolgen. Dabei sollte nicht nur der unterschiedliche Kenntnisstand der Benutzer berücksichtigt werden, sondern auch ausführlich auf die Unterschiede bei der Benutzung der einzelnen Standorte, die es trotz aller internen Abstimmungen immer noch gibt, eingegangen werden. Deshalb werden im bwGRiD entsprechende Materialien derzeit noch dezentral auf den Webseiten der Standorte bereitgestellt. Zusätzlich wird gegenwärtig ein Benutzerhandbuch für das bwGRiD entwickelt, das sowohl in elektronischer als auch in gedruckter Form verfügbar sein wird. Es enthält neben standortspezifischen Besonderheiten auch Informationen und Beispiele für verschiedenste Benutzergruppen. Weitere Softwarewerkzeuge für die Nutzung des bwGRiDs wurden von den Projektteilnehmern entwickelt und verteilt. Dazu gehört eine vorkonfigurierte virtuelle Maschine (bwGRiD-VM) auf Basis von Ubuntu 10.04LTS. Diese stellt bereits eine vollständige Globus-Installation, sowie zahlreiche Skripte für die Konfiguration und Nutzung des bwGRiDs zur Verfügung. Außerdem wurden mittlerweile zwei Informationsflyer mit generellen und technischen Informationen über das Grid erstellt und veröffentlicht.

Im April 2010 startete das bwGRiD-Portal-Projekt.¹⁶ Es richtet sich vor allem an Forschungsgruppen, deren Mitglieder wenig Erfahrung im Umgang mit HPC-Ressourcen haben und soll ihnen einen einfachen Zugang zum Grid-Computing¹⁷ ermöglichen. Insbesondere der teils komplizierte Umgang mit der Kommandozeile soll durch das Webportal ersetzt und die Verwaltung und Überwachung von eigenen Jobs und Daten vereinfacht werden. Anwendungsspezifische Komponenten (Portlets) bieten hierbei die Möglichkeit zum Erstellen beziehungsweise Hochladen von Eingabedaten und zum Abschicken der Jobs. Die Portlets werden an verschiedenen Standorten in enger Zusammenarbeit mit den jeweiligen Anwendern entwickelt. Zurzeit ist nur die Grid-Anbindung via Globus (4.0.8 WS-GRAM) realisiert. Der Zugriff auf das Portal erfolgt über Grid-Nutzerzertifikate (siehe Abschnitt 3.4). Im Rahmen des Projektes wurde das Firefox-Addon „Grid Proxy Manager“ entwickelt, welches das Erstellen und Hochladen von MyProxy-Zertifikaten [NTW01] für die Nutzeridentifikation innerhalb des Grids erheblich erleichtert.¹⁸

4 Resultate

Die erfolgreiche Entwicklung des bwGRiD-Projektes lässt sich an der kontinuierlichen Zunahme der Publikationen ablesen. Wurden im Jahr 2008 sieben Publikationen, die das

¹⁶Technisch basiert das bwGRiD-Portal auf dem Portlet-Framework GridSphere und dem Servlet-Container Apache Tomcat

¹⁷Die Kommunikation mit den Grid-Komponenten erfolgt über das Grid Application Toolkit (GAT) und Gatlet.

¹⁸Projekt-Webseiten: <http://www.gridisphere.org/gridsphere/gridsphere>,
<https://gforge.cs.vu.nl/gf/project/javagat>, <http://gatlet.scc.kit.edu/>,
<https://addons.mozilla.org/de/firefox/addon/grid-proxy-manager/>

bwGRiD referenzieren, veröffentlicht, waren es im Jahr 2011 schon 101 Veröffentlichungen (Abbildung 2a). Insgesamt sind bislang 220 Publikationen oder Konferenzbeiträge eingereicht und akzeptiert worden.¹⁹

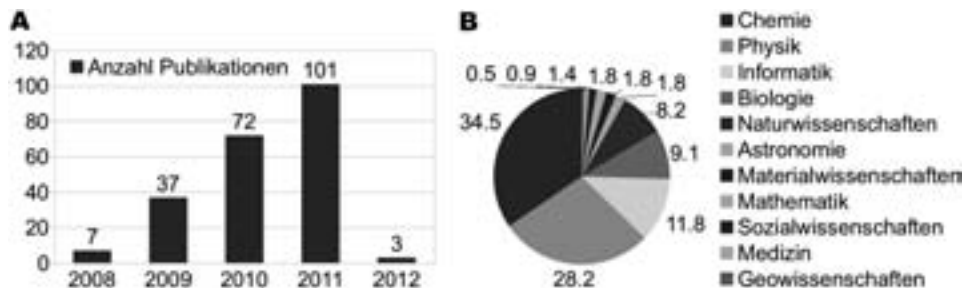


Abbildung 2: **Publikationen im bwGRiD** (Stand: 28.03.2011)

(A) Anzahl der Publikationen in denen Berechnungen auf dem bwGRiD durchgeführt wurden.

(B) Prozentualer Anteil der Schlagworte an Gesamtzahl der Publikationen.

Die Aufschlüsselung der mithilfe des bwGRiDs erzeugten Publikationen nach Fachrichtungen zeigt, dass der Großteil der derzeitigen Anwender aus dem naturwissenschaftlichen Bereich kommt. So machen Veröffentlichungen in den Bereichen Chemie, Physik, (Bio-) Informatik, Biologie und weiteren naturwissenschaftlichen Feldern zusammen einen Anteil von über 90 Prozent an allen Veröffentlichungen aus. Auf die restlichen rund acht Prozent verteilen sich Veröffentlichungen in den Gebieten Astronomie (1,8%), Materialwissenschaften (1,8%), Mathematik (1,8%), Sozialwissenschaften (1,4%), Medizin (0,9%) und Geowissenschaften (0,5%). Aus der Analyse der Publikationen lassen sich zwei Schwerpunkte für die weitere Entwicklung des bwGRiDs ableiten. Zum einen die Stärkung des Anteils von nicht-klassischen Disziplinen durch verstärkte Werbung und den Abbau der Einstiegshürden für das Grid-Computing. Zum anderen die Optimierung des bestehenden Angebotes, um die Fachbereiche, die bereits die Dienstleistungen des bwGRiDs nutzen, noch besser zu unterstützen.

5 Fazit und Ausblick

Schnelle Netze erlauben neuartige Betriebsmodelle für HPC-Cluster. Die überregionale Vernetzung der Cluster über Hochgeschwindigkeits-Weitverkehrsnetze erhöht die Redundanz erheblich und führt zu einer Hochverfügbarkeit der HPC-Ressourcen für Anwender. Ein Ausfall oder die Wartung eines Standortes können durch andere Standorte kompensiert werden. Single-Point-of-Failures werden somit minimiert. Dadurch nähert man sich dem aktuellen Cloud-Paradigma, das Community-Grids und Nachhaltigkeit fordert, weiter an. Das heißt, die für die jeweilige Berechnung am besten geeigneten Maschinen und der optimale Ort für die konkrete Ausführung verschiedener Jobs mit unterschiedlichen Anforderungen lassen sich dynamisch über große geografische Entfernungen zusammenfügen.

¹⁹Publikationsliste bwGRiD (Stand: 28.03.2011): <http://www.bw-grid.de/publikationen/>

Dabei erleichtert eine geeignete Job-Klassifikation den Grid-Nutzern sinnvolle Vorschläge zur optimalen Rechenumgebung, beziehungsweise zum Ausführungsort zu unterbreiten. Grundlage dafür sind neben einer ausreichend hohen Netzwerkbandbreite auch das Setzen gemeinsamer Standards sowie das Treffen und Einhalten gemeinsamer Absprachen. Ein bedeutender Vorteil des bwGRiD-Modells ist die Möglichkeit einer Spezialisierung innerhalb der Grid-Community, in dem sich verschiedene, verteilte Gruppen um unterschiedliche Aspekte des Cluster-Betriebs kümmern.

Durch den Verzicht einzelner Projektpartner auf ein gewisses Maß an Souveränität zugunsten eines gemeinsam abgestimmten Vorgehens bei der Beschaffung und dem Betrieb des bwGRiDs konnte der akademischen Forschung in Baden-Württemberg, und über das Bundesland hinaus, eine flexible und hochverfügbare HPC-Umgebung zur Verfügung gestellt werden. Das bwGRiD bietet HPC-erfahrenen Arbeitsgruppen und Neueinsteigern eine wertvolle Plattform für die Spitzenforschung im Land. Dies lässt sich auch an der konstant hohen Auslastung der bwGRiD-Ressourcen erkennen, welche ein auf dem bwGRiD basierendes Folgeprojekt notwendig macht.

Zuvor werden jedoch noch folgende Fragestellungen in den nächsten zwei Jahren behandelt. Zum einen wird eine Möglichkeit eines einfachen Cluster-übergreifenden Meta-Schedulings im Rahmen der *Ergänzenden Maßnahmen* untersucht. Dabei liegt die eigentliche Schwierigkeit in der Bereitstellung der Daten, einer Abschätzung der Übermittlungszeit dieser und der Wartezeit in der Schlange, sowie der Bereitstellung der eventuell angefragten Lizenztokens. Des Weiteren wird eine einfache Authentifizierung und Autorisierung gegenüber dem bwGRiD-Projekt mithilfe von Shibboleth untersucht. Es finden zudem Bestrebungen zur Entwicklung einer zentralen Datenbank und eines Software-Repositories statt, die für eine schnellere Verteilung der Software auf alle Standorte sorgen und damit die Administration erleichtern, sowie den Nutzern möglichst homogene Softwareumgebungen bieten sollen. Das bwGRiD-Portal wird gerade vom Portlet-Framework GridSphere auf das leistungsstärkere Liferay²⁰ portiert und eine Migration von Globus 4 auf Globus 5 ist ebenfalls in Planung.

Das bwGRiD hat sich im Laufe der Zeit zu einem erfolgreichen und unverzichtbaren Werkzeug für eine Vielzahl von Wissenschaftlern aus den verschiedensten Disziplinen entwickelt und wird auch in Zukunft eine bedeutende Rolle für die Forschung spielen.

Literatur

- [BK10] C. Baun und M. Kunze. Aufbau einer Computing Cloud am KIT – Betrachtung von Leistungsaspekten. *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, pages 26–35, 2010.
- [Gen06] W. Gentsch. Das Verbundprojekt D-Grid. *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, pages 132–139, 2006.

²⁰Projektseite Liferay: <http://www.liferay.com/>

- [Mic06] K.-P. Mickel. Erfahrungen mit Produktionsgrids am Beispiel des LHC-Computing-Grid (LCG). *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, pages 140 – 145, 2006.
- [NTW01] J. Novotny, S. Tuecke und V. Welch. An Online Credential Repository for the Grid: My-Proxy. In *Proceedings of the 10th IEEE International Symposium on High Performance Distributed Computing*, pages 104–, Washington, DC, USA, 2001. IEEE Computer Society.
- [RHKK10] S. Richling, S. Hau, H. Kredel und H.-G. Kruse. Operating Two InfiniBand Grid Clusters over 28 km Distance. *P2P, Parallel, Grid, Cloud, and Internet Computing, International Conference on*, 0:16–23, 2010.
- [RV10] A. Reber und P. Väterlein. Computer aus der Steckdose Hochschule Esslingen wird Teil des bwGRiD. *Horizonte*, 36(ISSN: 1432-9174):70, 2010.
- [SvW⁺11] S. Schmelzer, D. von Suchodoletz, D. Weingaertner, L. C. De Bona, G. Schneider und C. Carvalho. Universal Remote Boot and Administration Service. In *7th Latin American Network Operations and Management Symposium*, number ISBN 978-1-4577-1791-8, 2011.