



Arbeitspaket 2: Blaupausen und Beratung

Blaupause für die organisatorische Struktur eines neuen Community Grid¹

Deliverable	2.1.3 Organisatorische Struktur eines CG
Autoren	Arbeitspaket 2: Blaupausen und Beratung
Editoren	A. Rapp, O. Schmid et al.
Datum	23-05-2011
Dokument Version	2.0.0

A: Status des Dokuments

Deliverable 2.1.3, Version 2.0.0, 2011-Iteration vom Projekt akzeptiert.

B: Bezug zum Projektplan

In diesem Dokument werden die Erfahrungen der CGs bezüglich organisatorischer und struktureller Voraussetzungen und Maßnahmen wiedergegeben.

C: Abstract

Das Dokument beleuchtet die organisatorischen Problemstellungen bei der Planung und Umsetzung des Einstiegs in die Grid-Technologie. Es werden die Erfahrungen der einzelnen Communities beschrieben und mögliche Lösungsansätze erörtert. Dabei kann dieser Leitfaden nicht *die* ideale

¹ This work is created by the WissGrid project. The project is funded by the German Federal Ministry of Education and Research (BMBF).

Vorgehensweise aufzeigen, sondern nur die Suche nach der bestmöglichen Lösung erleichtern.

D: Änderungen

Version	Date	Name	Brief summary
0.1.0	10.11.2009	H. Enke	Erstellung des Ausgangsdokuments
0.1.1	23.11.2009	A. Rapp / O. Schmid	Entwicklung des Basisdokuments
0.1.2	25.11.2009	A. Rapp / O. Schmid	Strukturierung des Dokuments
0.1.3	15.12.2009	O. Schmid	Überarbeitung des Dokuments
0.2.0	17.01.2010	O. Schmid	Integration des Feedbacks anderer Communities
0.2.1	21.01.2010	O. Schmid	Ausformulierung einzelner Abschnitte
0.2.2	24.01.2010	O. Schmid	Ausformulierung weiterer Abschnitte
0.2.3	29.01.2010	A. Rapp / O. Schmid	Überarbeitung der Abschnitte
0.2.4	01.02.2010	O. Schmid	Umlaute ...
0.2.5	02.02.2010	O. Schmid	Anmerkung zu Nachteilen vermerkt
0.2.6	08.02.2010	O. Schmid	Integration des Feedbacks (Phase II)
0.3.0	10.02.2010	O. Schmid	Ausformulierung verbleibender Abschnitte
0.3.1	11.02.2010	O. Schmid	Vervollständigung des Literaturverzeichnisses
0.3.2	12.02.2010	H. Enke	Typos, Expressions
0.9.0	12.02.2010	O. Schmid	Fertigstellung des ersten Draft
1.0.0	26.02.2010	H. Enke	Veröffentlichung des ersten Draft
1.1.0	11.02.2011	O. Schmid	Überarbeitung des ersten Draft
1.1.1	14.02.2011	A. Rapp / O. Schmid	Notizen zu neuen Abschnitten
1.2.0	15.02.2011	O. Schmid	Ausformulierung Heterogenität
1.2.1	17.02.2011	O. Schmid	Verweise und Literatur
1.3.0	21.02.2011	O. Schmid	Ausformulierung TextGrid-Projekte
1.3.1	23.02.2011	B. Fritzsch	Nachnutzung, Bsp. Klimafolgenforschung
1.3.1	28.02.2011	O. Schmid	Integration Feedback T. Rathmann & G. Stöckle
1.3.2	07.03.2011	A. Rapp / O. Schmid	Weitere TextGrid-Beispiele
1.3.3	09.03.2011	O. Schmid	Erfahrungen Projektplanung
1.3.4	11.03.2011	O. Schmid	Überlegungen zu Kooperationen
1.4.0	15.03.2011	O. Schmid	Ergänzungen zu Göttinger Projekten
1.5.0	16.05.2011	O. Schmid	Integration Feedback H. Enke
2.0.0	23.05.2011	H. Enke	Release

E:

Inhaltsverzeichnis

1 Summary	4
2 Einleitung	4
3 Einführung in das Thema	4
3.1 Grundlageninformationen	5
3.2 Vorbildhafte Umsetzung in den Communities	5
3.3 Voraussetzung für die Bildung einer Community	6
4 Umsetzung / Strukturen	7
4.1 Benötigte Ressourcen / Mengengerüste	7
4.2 Langfristige Entwicklung	9
4.3 Organisationsformen	9
5 Grid-Einsatz im Rahmen eines Einzelprojekts (Join the Community)	10
6 Support	11
7 First Steps in Grid Technology - „BIG FOR EU“	12

1 Summary

The following sections provide an overview of organizational questions to consider prior to the decision in favor of or against the development of or the participation in a Grid Community. Basically the target audience for this paper are decision makers as there are few technical details described, but primarily financial and structural considerations. Thus the text does not require any technical knowledge as it has not been composed to inform members of technical staff, nevertheless it might still be an interesting supplement for them.

2 Einleitung

Auf den folgenden Seiten sollen organisatorische Fragestellungen erörtert werden, die sich bei der Diskussion eines Einstiegs in die Grid-Technologie ergeben, also beispielsweise zu finanziellen Veränderungen und politischen Konsequenzen. Im Mittelpunkt stehen somit die Überlegungen, die weniger für Techniker, sondern vielmehr für Fachwissenschaftler interessant sind, sowohl für Anwender als auch insbesondere für Entscheidungsträger. Technische Gegebenheiten werden also nur am Rande angesprochen, sofern sie von organisatorischer Bedeutung sind. Auf eine detaillierte Beschreibung technischer Rahmenbedingungen wird an dieser Stelle jedoch gänzlich verzichtet, diese können den anderen Deliverables entnommen werden. Zum Verständnis dieser Lektüre sind daher nur geringe IT-Kenntnisse notwendig, vielmehr ist es von Bedeutung, über Erfahrung in den Bereichen (Projekt-)Planung und Management zu verfügen.

Bei der Suche nach einer Entscheidung für oder gegen den Einsatz von Grid-Technologie ist zu beachten, dass es keine einheitliche „Patentlösung“ für alle Grid-Interessierten gibt, vielmehr soll dieses Kapitel als organisatorische Entscheidungshilfe für die beste Lösung unter Einbeziehung der jeweiligen Rahmenbedingungen dienen. Generell werden in den kommenden Abschnitten Erfahrungen geschildert, bei denen die Leser selbst entscheiden müssen, welche auf ihre eigene Situation zutreffen (können). Während es bei bestimmten technischen Entscheidungen nur wenige vernünftige Alternativen geben kann, welche den technischen Anforderungen einer Community gerecht werden, gibt es bei den Fragestellungen zu organisatorischen Strukturen oder zur Durchsetzbarkeit technischer Innovationen in einer Community (in Abhängigkeit von ihrer IT-Affinität) viele Einflussfaktoren und somit viele mögliche Lösungsansätze.

3 Einführung in das Thema

Aus dem Einsatz der Grid-Technologie können sich zahlreiche vorteilhafte Entwicklungen auf verschiedenen Ebenen ergeben, denen nur sehr wenige Nachteile gegenübergestellt werden können. Aus technisch-organisatorischer Sicht kann als Nachteil gewertet werden, dass im Daten-Grid gespeicherte Daten per se nicht lokal gespeichert werden würden und somit bei einem Ausfall der Grid-Infrastruktur vorübergehend nicht verfügbar wären – diesem Problem kann jedoch auf einfache Weise vorgebeugt werden. Auch erweist sich der Betreuungsaufwand als höher als bei lokaler Arbeit, diesem Kritikpunkt kann jedoch der Mehrwert des Grid entgegengestellt werden. Das Gleiche gilt in Bezug auf die Sicherheit der IT-Infrastruktur, die nach dem Einstieg in das Grid mit

mehr Problemen konfrontiert sein wird als unter „normalen“ Arbeitsbedingungen.²

Die rechtlichen Rahmenbedingungen gilt es insbesondere dann genauer zu erörtern, wenn der internationale Einsatz des Grids geplant ist, da dadurch für die Wissenschaftler bis dato nicht relevante Gesetze zu berücksichtigen sind. Doch auch auf nationaler Ebene gilt es, im Grid-Kontext Regelungen zu beachten, die unter anderen Umständen für die tägliche Arbeit nicht in dieser Form relevant sind, beispielsweise den Datenschutzrichtlinien und lizenzrechtliche Vorschriften. Auch ergibt sich aus der Teilnahme im Grid die Notwendigkeit von Zusicherungen und Vertragsbeziehungen wie sogenannte „Service Level Agreements“ (SLAs). In den kommenden Abschnitten sollen die Veränderungen dargestellt werden, die sich technisch und organisatorisch, aber auch finanziell und politisch ergeben, sowie das Innovationspotential für die Forschung.

3.1 Grundlageninformationen

Einer der grundlegenden Vorzüge der Grid-Technologie ist, dass im Prinzip alle Disziplinen von ihr profitieren, sowohl zum gegenwärtigen Zeitpunkt als auch auf lange Sicht gesehen. Die Möglichkeiten, welche die Grid-Technologie eröffnet, sind hierfür von elementarer Bedeutung, da sie vielseitige Problemlösungen bietet. So ermöglicht sie eine bessere Auslastung von Ressourcen, insbesondere von Kapazitäten im Bereich der Rechenleistung, auch erlaubt das Grid den On-Demand-Zugriff auf Ressourcen und Dienste sowie einen hohen Datendurchsatz. Vor allem bietet das Grid durch die langfristige Verfügbarkeit von Forschungsdaten einen Vorteil, auf den keine wissenschaftliche Community verzichten kann. Letzten Endes kann mit Hilfe der Grid-Technologie eine Form der Forschung realisiert werden, welche ohne Grid teilweise überhaupt nicht möglich wäre. Von Seiten der Förderinstitutionen wird das Grid vor allem vom Bundesministerium für Bildung und Forschung (BMBF) unterstützt, von der Deutschen Forschungsgemeinschaft (DFG) wird die Technologie bei Bewilligungsentscheidungen zunehmend berücksichtigt. Das Deliverable 1.4 befasst sich ausführlicher mit dieser Thematik [DS10].

3.2 Vorbildhafte Umsetzung in den Communities

In verschiedenen Communities hat die Grid-Technologie bereits erfolgreich Einzug gehalten, dabei handelt es sich nicht nur um IT-nahe Forschungsbereiche wie die Hochenergie-Physik, sondern beispielsweise auch um die Textwissenschaften als einen Vertreter der Geistes- und Kulturwissenschaften. Die neuen Möglichkeiten, welche durch den Einsatz der Grid-Technologie geboten werden, können hier exemplarisch am Beispiel des MediGRID aufgezeigt werden.

Im Rahmen des Projekts MediGRID wurde eine Grid-Infrastruktur für die biomedizinische Forschung geschaffen. So wurde mit dem MediGRID-Portal [MG10b] ein Portal-System eingerichtet und ein Workflow-Management-System eingebunden, gleichzeitig wurde ein Leitfaden zur Anwendungsintegration in das Grid etabliert. In diesem Portal wurden bereits etliche Anwendungen integriert [MG10c], um einen intuitiven zentralen Zugriff zu ermöglichen. Durch den Einsatz von Grid-Technologie konnte als nennenswerter Erfolg das Risiko-Gen für Gallensteine entdeckt werden [A08]. Im täglichen Einsatz kann das Grid-Computing die OP-Simulation beschleunigen [A06].

²Die veränderten Anforderungen werden jedoch nicht durch die Grid-Technologie als solche, sondern durch die mehr kollaborativen Anforderungen der wissenschaftlichen Arbeit erzeugt, für die Grid-Technologie ein Lösungsansatz ist. Veränderte, und höhere Anforderungen an die IT-Infrastruktur zieht auch Veränderungen in der Arbeit der IT-Abteilungen nach sich.

Der Einsatz von Grid-Technologie muss nicht immer bedeuten, dass sich der Community neue Quellen erschließen, oft gestaltet sich lediglich der Zugriff auf bereits genutzte Quellen einfacher und effizienter. So wurden im C3-Grid durch den Aufbau eines Daten-Grids keine neuen Archive geschaffen, jedoch kann über ein komfortables Web-Portal auf die verteilten Archive zugegriffen werden und somit das gesamte Daten-Grid durchsucht werden. Die Suche in diesen Archiven war zuvor schon möglich gewesen, durch das manuelle Abfragen der einzelnen Archive jedoch recht mühsam und zeitaufwändig. Im Fall von TextGrid ist es beispielsweise möglich, den Schutz von Daten durch eine genaue Regelung der Zugriffsrechte zu garantieren und auch bei weniger IT-affinen Communities Vertrauen in die neue Technologie zu schaffen. Dem Nutzer bleibt damit die Möglichkeit erhalten, stets den Zugang zu seinen Daten im Grid kontrollieren zu können. Dies steigert somit die Bereitschaft, an der technischen Entwicklung teilzuhaben und Forschungsergebnisse zur Verfügung zu stellen.

Durch den Zugriff auf verfügbare, nicht genutzte Rechenleistung erschließen sich im Grid hingegen wichtige Ressourcen, welche die Durchführung komplexer Berechnungen in angemessener Zeit erlauben, wie im bereits erwähnten Fall der Entdeckung des Risiko-Gens für Gallensteine. Des Weiteren können zeitkritische Anwendungen wie OP-Simulationen durch verteiltes Rechnen erheblich beschleunigt werden. Auch gilt es in der Medizin oft, mit hochdimensionalen Daten und der korrelativen Verknüpfung von vielfältigen, genotypischen und phänotypischen Daten unterschiedlichster Art zu arbeiten. Dabei erfordert das Einbinden externer Rechenkapazitäten durch die Schaffung einer Grid-Infrastruktur für den normalen Anwender oft keine speziellen IT-Kenntnisse.

Neben der besseren Nutzung und Auslastung von Ressourcen, insbesondere von Rechenleistung, können sich durch die Entstehung eines Community-spezifischen Grids auch weitgehend standortunabhängige Kooperationen zwischen Institutionen ergeben und im Lauf der Zeit auch zunächst außenstehende Communities in eine Grid Community integriert werden, wie auch kommerzielle Partner. So gelang durch die Zusammenarbeit mit Bayer Technology Services GmbH (BTS) als Vertreter der Pharma-Community die Einbeziehung eines technologischen Dienstleisters einer völlig anderen Community in das MediGRID. Dies zeigt, wie die Verwendung von Grid-Technologie der Förderung und Verbesserung der interdisziplinären Zusammenarbeit dienen kann.

3.3 Voraussetzung für die Bildung einer Community

Bevor eine neue Grid Community ins Leben gerufen werden kann, ist es wichtig, sich über ihre Zusammensetzung und die daraus resultierenden Konsequenzen Gedanken zu machen. Bei einer homogenen Community gelten für alle Mitglieder ungefähr die gleichen Voraussetzungen. Je heterogener sich jedoch eine Community zusammensetzt, desto schwieriger wird es, die Anforderungen ihrer einzelnen Mitglieder abzustimmen und ihre Wünsche unter einen Hut zu bekommen. So hat sich beispielsweise aus dem breiten Spektrum der Physik nicht *eine* Grid Community entwickelt, sondern mehrere Communities, deren Anforderungen ebenso breit gefächert sind wie das Fach selbst. Einer der Gründe für diese Aufteilung in eigenständige Communities trotz auf den ersten Blick ähnlicher Anforderungen an die eigentliche Rechen- sowie Speicherleistung von Computern ist die Tatsache, dass auf grundlegend verschiedene technische Geräte zugegriffen werden muss. Darüber hinaus sind Astrophysiker für ihre Arbeit auf verschiedene Arten von Teleskopen angewiesen und können dabei auf zahlreiche Datenquellen zugreifen, während Hochenergiephysiker für ihre Experimente Zugang zu einem Teilchenbeschleuniger und somit einer sehr exklusiven Datenquelle benötigen. Auch unterscheiden sich die Arbeitsweisen der beiden Communities im Bereich der Simulation und den daraus resultierenden Anforderungen an die informationstechnischen Gegebenheiten.

Um bei der Suche von Partnern eine größere Akzeptanz innerhalb einer wissenschaftlichen Community zu erlangen, sollte die zu bildende Grid Community bei ihrer Gründung möglichst homogen gehalten werden, dies gilt vor allem für weniger technik-affine Fächer. Der Kern dieser Community sollte jedoch einer Erweiterbarkeit offen gegenüberstehen, falls eine ähnlich ausgerichtete Fachrichtung sich der Grid Community anschließen möchte. So zählte das Projekt TextGrid in seiner ersten Projektphase nur die Literatur- und die Sprachwissenschaft zum Kreis seiner Mitglieder, dieser wurde zu Beginn der zweiten Phase um die Kunst- sowie die Musikwissenschaft erweitert und steht weiteren Vertretern der Geistes- und Kulturwissenschaften offen. Als Schnittmenge wurde bei diesen Teil-Communities die Arbeit des Edierens identifiziert, beispielsweise von Notentexten in der Musik oder von Wörterbüchern, die gemeinsame Basis bildet dabei die Auszeichnungssprache XML. Auf diese Weise konnten durch sorgfältige Analyse der Arbeitsweisen inhaltlich so unterschiedliche Forschungszweige wie die Systemlinguistik und die Kunstgeschichte in einer Grid Community vereinigt werden. Hinzu kommt, dass sich die Erweiterbarkeit der Community selbst in der Modularität ihrer Software widerspiegelt, so wird das TextGridLab für neue Community-Mitglieder durch spezielle Tools wie einen Noten-Editor für die Annotation von Musiktexten erweitert.

Die Homogenität eines Interessentenkreises vor sowie seine Flexibilität nach der Gründung einer Grid Community ist eine gute Voraussetzung für den Erfolg dieses Unterfangens. Hat eine Grid Community bei ihren Nutzern erst einmal einen gewissen Grad an Akzeptanz erreicht, sollte sie die zunehmende Heterogenisierung durch ihre Erweiterung vor keine größeren Probleme mehr stellen.

4 Umsetzung / Strukturen

Plant eine Institution die Umstellung auf den Grid-Einsatz, so stellt sich für sie die Frage, für welche Gruppe(n) bzw. Abteilung(en) diese Maßnahme sinnvoll ist, oder ob sie am Besten für die gesamte Einrichtung durchgeführt werden soll.

4.1 Benötigte Ressourcen / Mengengerüste

Der Einsatz von Grid-Technologie zieht eine Reihe von vielschichtigen Umstrukturierungsmaßnahmen nach sich. Er bedeutet sowohl eine finanzielle Umorientierung als auch neue Anforderungen an bestimmte Personalgruppen. Dabei ist in der Planung der Finanzierung zu berücksichtigen, dass Mittel aufgrund von Einsparungen umgeschichtet werden können und teilweise auch müssen, wenn Investitionen an anderer Stelle notwendig werden. Da einige Fragestellungen zur organisatorischen Umstrukturierung eng mit dem Kostenaspekt verknüpft sind, wird in den folgenden Abschnitten auf die finanziellen Gesichtspunkte nur so weit eingegangen, wie es für die Betrachtung der organisatorischen Strukturen nötig ist. Primär finanzielle Aspekte werden in Deliverable 2.1.4 im Rahmen einer Kosten-Nutzen-Abschätzung behandelt [SS10a].

Bei der näheren Betrachtung der finanziellen Gegebenheiten zeigt sich, dass auf der einen Seite Kosten eingespart werden können, auf der anderen Seite aber auch neue Kosten entstehen können. Beide möglichen Entwicklungen hängen dabei von den individuellen Rahmenbedingungen der jeweiligen Community ab. Werden weniger eigene lokale Ressourcen benötigt, weil Ressourcen gemeinsam genutzt und dadurch besser ausgelastet werden können, entsteht ein teilweise erhebliches Einsparungspotential. Durch die Senkung der Kosten für Beschaffung und Betrieb der Hardware bei gleichbleibender oder steigender Verfügbarkeit von Rechenleistung oder anderen IT-Leistungen

erhöht sich damit die Rentabilität der Investitionen. Wird hingegen die bestehende Infrastruktur in unveränderter Form weitergenutzt und in einem Daten-Grid nur durch die Vernetzung von Einzelarchiven und deren Zugänge aufgewertet, können hier keine Mittel eingespart werden. Im Gegenteil entstehen neue Kosten durch den zusätzlichen Betrieb der Grid-Infrastruktur und der Grid-Dienste, wobei aber ein höherer Nutzen gewonnen wird.

Aus der Umstellung auf Grid-Technologie ergeben sich anfangs erhöhte Personalkosten, sowohl für zusätzliche Administratoren als auch durch einen erhöhten Aufwand bei der Anwendungsentwicklung und der Anwendungsintegration in das Grid. Auch werden Schulungsmaßnahmen für Entwickler notwendig, da die Nutzung der sich aus dem Grid ergebenden Möglichkeiten nicht trivial ist. Eine Senkung der Personalkosten ist durch die Vereinfachung der Suche im Grid und des Zugriffs auf die verfügbaren Daten lediglich im Rahmen eingesparter Arbeitszeit möglich.

Bei der technischen Ausstattung kann im Rahmen von Grid-Projekten auf die übliche Hardware, aus eigenen Budgets finanziert, zurückgegriffen werden, sowie unter Umständen auch auf die D-Grid-Sonderinvestitionen. Die Anforderungen in Bezug auf Software und Sicherheitsmaßnahmen können den Ausführungen in Abschnitt 4 des Deliverable 2.1.5 entnommen werden [SA10].

Für das Personal entstehen durch die Nutzung des Grids zwar teilweise neue Herausforderungen, es ergeben sich jedoch auch neue Möglichkeiten. So müssen sich die Nutzer von Portalen keine speziellen Kenntnisse aneignen, um beispielsweise Daten im Grid zu suchen und herunterzuladen. Wollen Grid-Nutzer hingegen eigene Workflows entwickeln, brauchen sie ähnliche Grid-spezifische Kenntnisse wie die Entwickler der Grid-Infrastruktur. Diese IT-Spezialisten sollten den fortgeschrittenen Anwendern für solche Aufgaben mit Rat und Tat zur Seite stehen. In WissGrid gibt es für diese Unterstützung mit den Fachberatern, auf die in Deliverable 2.2.1 genauer eingegangen wird [SS10b], Grid-erfahrenes Personal aus den unterschiedlichsten Fachrichtungen. Im optimalen Fall sollten Forscher durch die Verwendung von Portalsystemen aber keine Grid-Kenntnisse benötigen, auch macht die Einbettung von Grid-Operationen in zur Verfügung gestellte Anwendungen eine spezielle Qualifikation nicht erforderlich. Es sollte den Anwendern aber ein grundsätzliches Verständnis für das Arbeiten im Grid vermittelt werden. Bei der Förderung des wissenschaftlichen Nachwuchses kann dieses Wissen auch im Rahmen von Anwendungs-Seminaren vermittelt werden. So binden mehrere Communities die Grid-Technologie auf diese Weise in die universitäre Lehre ein, auch werden inzwischen bereits Abschlussarbeiten mit Grid-Bezug vergeben.

Das Job-Profil eines Anwendungsentwicklers im Grid-Umfeld unterscheidet sich von den üblichen Anforderungen durch die zusätzlichen Kenntnisse über die notwendigen Maßnahmen zur Integration von Anwendungen in die Grid-Middleware und die Einbindung in das Grid-Workflow-Management. Ein Grid-Techniker sollte über umfangreiche Kenntnisse über die verwendete Middleware wie das Globus Toolkit verfügen, auch ist ein vertieftes Verständnis weiterer Grid-Dienste erforderlich. Soll die Anbindung von Datenbanken an das Grid selbst entwickelt werden, werden hierfür ebenfalls entsprechende Kenntnisse benötigt.

Wie bei den Kosten können auch beim Faktor Zeit die Entwicklungen in sehr unterschiedliche Richtungen gehen, so kann die Ausführungszeit einer Anwendung abhängig von Art und Auslastung der Systeme kürzer oder länger dauern als ohne den Einsatz von Grid-Technologie. Mehr zeitlichen Aufwand erfordert im Grid-Kontext normalerweise sowohl die Entwicklung als auch die Integration von Anwendungen in das Grid. Dafür ist nach erfolgter Umsetzung einer Grid-Lösung mit einem höheren Mehrwert jener Anwendungen zu rechnen.

4.2 Langfristige Entwicklung

An weiteren Kosten sind durch den Umstieg auf das Grid verschiedene Folgekosten zu beachten, die je nach Community-Umfeld und den technischen Anforderungen an das Grid jedoch auch Einsparungen mit sich bringen können. Unzweifelhaft entstehen neue Personalkosten, insbesondere für Administrationskräfte mit vorhandenen oder zu entwickelnden Kenntnissen der Grid-Technologie. Auch können Fortbildungsmaßnahmen für diese, aber auch für wissenschaftliche Mitarbeiter notwendig sein. Soll das Grid auch oder vor allem als Computing-Grid eingesetzt werden, so dürfte sich für die Betriebskosten in diesem Bereich ein ähnliches Einsparungspotential ergeben wie im Bereich der Investitionen. Dagegen ist für den Einsatz als Daten-Grid durchaus denkbar, dass neue Kosten durch den Betrieb zusätzlicher IT-Infrastruktur entstehen, wie beispielsweise für Strom und Wartung. Darüber hinaus wären in diesem Fall auf lange Sicht auch Folgeinvestitionen für die Reparatur oder Erneuerung von Geräten zu erwarten.

Ein wichtiger Aspekt bei der organisatorischen und finanziellen Planung von Grid-Projekten ist die Überlegung, dass man nicht nur die vom Projekt selbst benötigten personellen Ressourcen berücksichtigen sollte. Wenn zunächst nicht am Grid-Projekt beteiligten Institutionen und Projekten die Möglichkeit gegeben werden soll, beispielsweise auf die Entwicklung einer Virtuellen Forschungsumgebung Einfluss zu nehmen, müssen die hierfür potentiell erforderlichen Personenmonate ebenfalls in Betracht gezogen werden. So ergaben sich bei TextGrid in der zweiten Projektphase Kapazitätsanforderungen, die bei der Formulierung des Antrags in dieser Form noch nicht abzusehen waren. Beispielsweise wurde das Metadaten-Konzept aus der ersten Phase entsprechend der sich ergebenden Nutzeranforderungen in der zweiten Projektphase vollständig überarbeitet. Auch der Text-Bild-Link-Editor, ein philologisches Tool zur Verknüpfung von Bildausschnitten mit XML-Daten, erfuhr eine erhebliche Erweiterung seines Funktionsumfangs auf Basis von neuen Nutzerwünschen.

Bei der langfristigen Planung mag sich die Frage stellen, was passieren wird, wenn das Grid von einer neueren Technologie abgelöst wird. Eine solche Entwicklung wird von den zurzeit existierenden Communities jedoch nur eingeschränkt als wahrscheinlich angesehen. So ist nach einhelliger Meinung für die Eigenschaften des Daten-Grids keine ähnliche Technologie in Sicht, welche das Grid nach derzeitigem Verständnis ablösen könnte. Die Langzeitarchivierung von Daten ist eine wichtige Herausforderung für die Forschung, das Erreichen dieses Ziels spielt im Grid-Kontext eine große Rolle und wird in Deliverable 3.1 ausführlich behandelt [WA11]. Im Bereich des Computing wird derzeit über das Cloud-Computing diskutiert und die Nutzung von Cloud-Ressourcen in Betracht gezogen. Jedoch ist nach derzeitigem Stand Cloud-Computing für viele Anwendungen aus dem akademischen Umfeld nicht geeignet, insbesondere wäre die Portierung von Daten und Anwendungen in die Cloud-Umgebung, neben den Kosten für die Nutzung von Cloud-Ressourcen, mit zusätzlichen Kosten verbunden.

4.3 Organisationsformen

Auch für die Organisationsform bieten sich zwei Alternativen an. Es kann einerseits in einer zentralen Einrichtung eine bestimmte Anzahl an Personen für die zu betreuenden Institutionen zuständig sein, andererseits kann bei einer dezentralen Lösung jede Abteilung einen eigenen „Grid-Beauftragten“ haben. Die Entscheidung für eine der beiden Alternativen hängt von den jeweiligen Organisationsgrößen und der Anzahl der zu betreuenden Organisationseinheiten ab. So empfiehlt sich bei vielen kleinen Abteilungen eine zentrale Lösung, um eine gleichmäßige Auslastung der Grid-Spezialisten zu gewährleisten, während für wenige und insbesondere für größere Einrichtungen der dezentra-

le Ansatz bevorzugt werden dürfte. Auch ist in diesem Zusammenhang zu beachten, in welchem Umfang die Grid-Umstellung vollzogen werden soll.

Bei der technischen Umsetzung muss sich eine Grid Community nicht auf die eigenen (informations)technischen Kenntnisse und Fähigkeiten verlassen, hier stehen verschiedene Formen der Unterstützung zur Verfügung. So können kommerzielle Kooperationspartner technische Aufgaben übernehmen und Dienste anbieten, wie es beispielsweise die Firma DAASI International im Projekt TextGrid macht. Auch hat DAASI International z. B. im Auftrag des Rechenzentrums der Universität Tübingen mit dem TUSTEP-Wrapper ein Softwaremodul für das Projekt TextGrid programmiert. Das Einrichten von Grid-Servern an Einrichtungen wie die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) kann ebenfalls vergeben werden. Hier ergeben sich die Kostenvorteile beim Outsourcing durch die Professionalität, mit der die ausgelagerten Aufgaben vom darauf spezialisierten Unternehmen erledigt werden.

5 Grid-Einsatz im Rahmen eines Einzelprojekts (Join the Community)

Soll ein einzelnes Projekt an einem bestehenden Community Grid partizipieren, so muss dieses Vorhaben von Anfang an gründlich durchdacht werden. Bereits in der Planungsphase gilt es verschiedene Aspekte zu berücksichtigen, seien sie organisatorischer, finanzieller, personeller oder sonstiger Natur. Auch kommen auf ein bereits genehmigtes oder laufendes Projekt zahlreiche Veränderungen zu, wenn dieses den Ein- bzw. Umstieg auf Grid-Technologie in Erwägung zieht.

Beispielsweise hat sich das Projekt „Blumenbach-Online“ der Universität Göttingen schon bei der Formulierung des Antrags dafür entschieden, die Infrastruktur sowie die Software des Grid-Projekts TextGrid zu nutzen. So soll unter anderem das TextGridRepository der Langzeitarchivierung der gewonnenen Erkenntnisse dienen, zudem entspricht die für geistes- und kulturwissenschaftliche Belange entwickelte Software den Anforderungen des Projekts. Gegen Ende seiner Laufzeit hat sich das ebenfalls in Göttingen angesiedelte Projekt „Archaeo 18“ dazu entschlossen, ebenfalls die Ressourcen von TextGrid zu nutzen, um in den Genuss der Vorzüge der Langzeitarchivierung zu gelangen. Über ihre bis dahin in der Kooperation mit dem Grid-Projekt gemachten Erfahrungen werden die Projekte „Blumenbach-Online“ und „Archaeo 18“ im Rahmen der TextGrid-Tage am 12. und 13. Juli 2011 berichten.

Weitere Projekte, die sich derzeit in der Antragsphase befinden, haben die zu erwartenden Aufwendungen für die Anpassung an eine existierende Grid-Umgebung in ihrer Antragsformulierung bereits berücksichtigt. So wurde in einem der Projekte im Umfeld von TextGrid eine Mitarbeiterstelle für Programmierung zur Weiterentwicklung von bestimmten Komponenten der vorhandenen Grid-Infrastruktur beantragt.

Gemäß Deliverable 1.4 gibt es für Projekte, die auf bereits vorhandene Grid-Infrastrukturen aufbauen wollen, bessere Förderungsmöglichkeiten als für den Aufbau neuer Community Grids, da die für den Aufbau von Community Grids zur Verfügung gestellten Sondermittel auslaufen [DS10]. Insofern bietet es sich für ein neues Projekt oder eine neue Community an, bei bereits existierenden Grid Communities mit ähnlichen Anforderungen anzufragen, ob sie Kapazitäten zur Verfügung stellen können. Auf diese Weise kann beispielweise durch Sharing von Rechenkapazität, also die Mitnutzung nicht voll genutzter Rechenleistung die Auslastung der beschafften Hardware erhöht werden.

Ein weiteres Beispiel für die Nachnutzung bestehender Infrastrukturen durch neue Communities ist die Klimafolgenforschung. Eine ausführliche Analyse der Anforderungen hat ergeben, dass die Komponenten von C3Grid auch von dieser Community für ihre Anwendungen genutzt werden kann. Das hat den Vorteil, dass die Klimafolgenforschung auf die gleiche Datenbasis zugreifen kann, ohne dass die Datenprovider zusätzlichen Aufwand haben. Die speziellen Analysen der Klimafolgenforschung werden dann in das entsprechende Portal integriert. Daher ist das Climate Service Center als ein wesentlicher Vertreter der neuen Community in dem vom BMBF-Fachreferat „Globaler Wandel“ geförderten Projekt „C3-INAD“ beteiligt, in dem die C3Grid-Infrastruktur an weitere Anforderungen angepasst werden soll.

In diesem Zusammenhang ist eine Neuausrichtung des Leistungsangebots von universitären Rechenzentren zu erwarten. Diese werden zukünftig voraussichtlich über die Grenzen des eigenen Campus hinweg Dienstleistungen anbieten und sich nicht mehr auf die Anforderungen der eigenen Hochschule beschränken. Dadurch können sich die Rechenzentren auf bestimmte Aspekte wie die Langzeitverfügbarkeit von Forschungsdaten spezialisieren und sind nicht mehr gezwungen, das komplette Leistungsspektrum in abgespeckter Version anbieten zu müssen.

Die hier beschriebenen Formen der gemeinsamen Nutzung sind ein wichtiger Gesichtspunkt für die effiziente Nutzung von Grid-Ressourcen. Die Entwicklung ist derzeit noch in vollem Gange, zum Teil muss die Standortbestimmung erst noch erfolgen. Aus diesem Grund wird diese Möglichkeit des Grid-Einstiegs für einzelne Projekte weiter verfolgt und in der kommenden Iteration dieser Blaupause ausführlicher beschrieben.

6 Support

Neben der vorliegenden Blaupause als erster Hilfe zum Einstieg in die Grid-Technologie gibt es noch viele weiterführende Quellen. Die umfangreichsten Sammlungen von Informationen dürften dabei die Webpräsenzen jener Grid-Projekte sein, die sich am Projekt WissGrid beteiligen [AG10], [C3G10a], [HG10], [MG10a], [TG10a]. Dort finden sich beispielsweise die Projektanträge und die Kontaktdaten der jeweiligen Ansprechpartner der Projekte sowie zahlreiche Veröffentlichungen über verschiedene Themen aus dem Bereich Grid [KLA07], [KLA09], [NJ+09]. Einzelne Communities bieten auf ihren Internetseiten auch Tutorials für unterschiedliche Zielgruppen an, so gibt es von C3Grid Anleitungen für Ressourcen-Provider [W08], [P07], von TextGrid ein Tutorial für Entwickler [LV09] und sowohl von TextGrid [TG10b] als auch von C3Grid [C3G10b], [B06] Tutorials für Benutzer.

Des Weiteren hat WissGrid ein Team von Fachberatern zusammengestellt, das neben der Erfahrung mit der Arbeit im Grid mit seinem Spektrum an Fachrichtungen auch die verschiedenen Bedürfnisse und den unterschiedlichen Zugang zur Informationstechnologie abdeckt. Darüber hinaus bietet das D-Grid User Support Portal (DGUS) Unterstützung für alle aktuellen und potentiellen Nutzer des D-Grid an. Weitere Informationen finden sich auch in Deliverable 2.2.1 [SS10b].

7 First Steps in Grid Technology - „BIG FOR EU“

Die ersten Schritte zum Einstieg in das Grid lassen sich in acht Punkten zusammenfassen.

Bedarfsanalyse durchführen: Ist der Umstieg auf das Grid sinnvoll? Können die Betroffenen von seiner Notwendigkeit überzeugt werden?

Informationen sammeln: Welche Erfahrungen haben andere, vergleichbare Communities gemacht? Welche Veränderungen kämen auf die eigene Community zu?

Grid Community definieren: Welche Forschungsbereiche eignen sich bzw. eignen sich nicht für den Zusammenschluss zu einer Community? Welche der geeigneten Kandidaten bieten sich dafür besonders an?

Finanzierung klären: Ist der Umstieg auf das Grid finanziell realisierbar? Stehen bei zusätzlich benötigten Investitionen Förderinstitutionen zur Verfügung?

Organisatorische Änderungen berücksichtigen: Welche Organisationsform ist am besten geeignet? Welche Organisationsstrukturen sind in welchem Ausmaß betroffen und unterliegen welchen Veränderungen?

Rechtliche Rahmenbedingungen prüfen: Welche lizenz- und datenschutzrechtlichen Bestimmungen gilt es zu beachten? Ergeben sich aus einer internationalen Ausrichtung des Grids besondere Probleme?

Entscheidung treffen: Sind die Argumente für das Grid überzeugend genug? Sprechen Argumente gegen den Umstieg?

Umsetzung der Entscheidung: Gegebenenfalls die notwendigen Anträge stellen, Mitarbeiter mit Grid-Kenntnissen einstellen, Maßnahmen zur technischen Umsetzung treffen (z. B. Beschaffung von Hardware), Anwender auf die Umstellung vorbereiten.

Literatur

- [AG10] *AstroGrid-D – German Astronomy Community Grid (GACG)*. www.astrogrid-d.org/ (Stand: Februar 2011)
- [A08] Beate Achilles: *Forschung: Mit GRID-Technologie schneller zum Ergebnis*. Pressemitteilung vom 09.07.2008 <http://www.idw-online.de/pages/de/news269773>
- [A06] ÄrzteZeitung.de (o.V.): *Computer im Verbund schaffen Op-Simulation in Minuten*. Artikel vom 05.07.2006 <http://www.aerztezeitung.de/panorama/default.aspx?sid=411042>
- [B06] Benny Bräuer: *Grid-Architektur und Dienste: Portal*. (Oktober 2006) <http://www.c3grid.de/fileadmin/c3outreach/generation-0/Dokumentation-Portal-G0.pdf>
- [C3G10a] *C3-Grid – Collaborative Climate Community Data and Processing Grid*. www.c3grid.de/ (Stand: Februar 2011)
- [C3G10b] *C3Grid: Deliverables*. <http://www.c3grid.de/index.php?id=52&L=10%20order%20by%209999999--> (Stand: Februar 2011)
- [DS10] Frank Dickmann, Ulrich Sax *Konzept eines Betriebsmodells für Virtuelle Forschungsumgebungen* WissGrid Deliverable 1.4 (Dezember 2010) http://www.wissgrid.de/publikationen/deliverables/wp1/WissGrid_AP1_D1-4_final_v2.pdf
- [HG10] *D-Grid Initiative: HEP Grid*. www.hepcg.org/ (Stand: Februar 2011)
- [KLA09] Marc Wilhelm Küster, Christoph Ludwig, Andreas Aschenbrenner: *TextGrid: eScholarship und vernetzte Angebote*. In: it - Information Technology, Heft 4 (August) 2009, Themenheft Informatik in den Philologien (Gastherausgeber: Claudine Moulin, Thomas Burch, Andrea Rapp). <http://www.textgrid.de/fileadmin/TextGrid/veroeffentlichungen/kuester-ludwig-aschenbrenner.pdf>
- [KLA07] Marc Wilhelm Küster, Christoph Ludwig, Andreas Aschenbrenner: *TextGrid as a Digital Ecosystem*. IEEE DEST 2007, 21.-23. Februar, Cairns, Australien, SPECIAL SESSION 3: e-Humanities for Digital Eco-systems: A Social, Cultural, Economic and Political Agenda. http://www.textgrid.de/fileadmin/TextGrid/veroeffentlichungen/DigitalEcosystem07_Camera.pdf
- [LV09] Christoph Ludwig, Thorsten Vitt: *An Introduction to Development with TextGrid*. (August 2009) http://www.textgrid.de/fileadmin/TextGrid/reports/developers_tutorial_090803.pdf
- [MG10a] *MediGRID – Home*. www.medigrid.de/ (Stand: Februar 2011)
- [MG10b] *MediGRID: Grid Portal der MediGRID Community*. <https://portal.medigrid.de/> (Stand: Februar 2011)
- [MG10c] *MediGRID: Anwendungen im Applikationsportal*. <http://medigrid.de/anwendungen.html> (Stand: Februar 2011)
- [NJ+09] Heike Neuroth, Fotis Jannidis, Andrea Rapp, Felix Lohmeier: *Virtuelle Forschungsumgebungen für e-Humanities. Maßnahmen zur optimalen Unterstützung von Forschungsprozessen in den Geisteswissenschaften*. In: Bibliothek. Forschung und Praxis, 2/2009. http://www.bibliothek-saur.de/preprint/2009/ar2581_neuroth.pdf

- [P07] Stefan Petri: *How to set up a C3-Grid Data Provider site?* (Februar 2007)
<http://www.c3grid.de/fileadmin/c3outreach/generation-0/howto-setup-dataprovider.pdf>
- [SA10] Frank Schlünzen, Ilya Agapov: *Evaluation und Dokumentation existierender Architekturkonzepte* WissGrid Deliverable 2.1.5 (Februar 2010)
<http://www.wissgrid.de/publikationen/deliverables/wp2/Gridarchitekturen-im-Ueberblick.pdf>
- [SS10a] Joachim Schultes, Stephan Sandvoss: *Kosten-Nutzen-Abschätzung des Aufbaus von neuen Communities* WissGrid Deliverable 2.1.4 (Mai 2010)
<http://www.wissgrid.de/publikationen/deliverables/wp2/Kosten-Nutzen-CG-Aufbau.pdf>
- [SS10b] Joachim Schultes, Stephan Sandvoss: *Aktivitäten von WissGrid-Fachberater-Teams* WissGrid Deliverable 2.2.1 (Oktober 2010)
http://www.wissgrid.de/publikationen/deliverables/wp2/Initial-Report-FB-WissGrid_2010-1.pdf
- [TG10a] *TextGrid – Vernetzte Forschungsumgebung in den eHumanities*. www.textgrid.de/
(Stand: Februar 2011)
- [TG10b] TextGrid: *TextGridLab Beta – Tutorials*. <http://www.textgrid.de/beta/tutorials.html>
(Stand: Februar 2011)
- [W08] Volker Winkelmann: *Computerressourcenanbindung*. (Juli 2008)
<http://www.c3grid.de/fileadmin/c3outreach/generation-1/Computerressourcenanbindung.pdf>
- [WA11] WissGrid Arbeitspaket 3 (verschiedene Autoren): *Generische Langzeitar-chivierungsarchitektur für D-Grid* WissGrid Deliverable 3.1 (Februar 2011)
<http://www.wissgrid.de/publikationen/deliverables/wp3/WissGrid-D3.1-LZA-Architektur-v1.pdf>