



Fachgebiet 3-4 – Aufbau einer AA-Infrastruktur
für das D-Grid

Ergebnisse der Interviews mit den D-Grid-Communities

Koordination

Christian Grimm (grimm@rvs.uni-hannover.de)

Marcus Pattloch (pattloch@dfn.de)

D-Grid Integrationsprojekt (DGI)

Autoren

Stefan Piger (RRZN, Leibniz Universität Hannover)

Christian Grimm (RRZN, Leibniz Universität Hannover)

Joachim Götze (ICSY, TU Kaiserslautern)

Markus Hillenbrand (ICSY, TU Kaiserslautern)

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01AK800B gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

1.	Einführung	4
2.	Allgemeine Angaben.....	4
2.1.	Anzahl der beteiligten Einrichtungen	4
2.2.	Anzahl der Nutzer innerhalb der Community	4
3.	Middleware und verwendete Software	5
3.1.	Grid-Middleware	5
3.2.	Grid-Portalsoftware	5
3.3.	Batch-System	5
3.4.	Verwendete Betriebssysteme.....	6
4.	VO-Konzept.....	7
5.	Authentifizierung	8
5.1.	Einsatz von PKI	8
5.2.	Unterstützte Certificate Authorities	8
5.3.	Vergabe von Accounts auf Ressourcen	9
5.3.1.	Planungen für die Laufzeit des Projektes	9
6.	Autorisierung	10
6.1.	Vergabe von Berechtigungen auf Ressourcen	10
6.2.	Use-Cases – Compute Services	10
6.2.1.	Limitierung der maximalen Ressourcennutzung je Nutzer.....	10
6.2.2.	Priorisierung von Nutzern auf Compute-Ressourcen	11
6.2.3.	Autorisierung von Lizenznutzung	11
6.2.4.	Autorisierung spezieller Ressourcen.....	12
6.2.5.	Exklusive Nutzung von Ressourcen.....	12
6.3.	Use-Cases – Data Services	13
6.3.1.	Autorisierung des Zugriffs auf temporärem und permanenten Speicher.....	13
6.3.2.	Quotierung von Speicher-Ressourcen	13
6.3.3.	Setzen von Rechten auf Speicherressourcen durch den Administrator	14
6.3.4.	Setzen von Rechten auf Speicherressourcen durch den Nutzer	14
6.3.5.	Replikation von Dateien	14
6.3.6.	Use-Cases – Freigabe von Daten in Informationssystemen	15
7.	Zusammenfassung und Ausblick.....	15
7.1.	Ausblick D-Grid AAI-Prototyp	16
7.1.1.	Phase 1	16
7.1.2.	Phase 2 (Skizze).....	16

1. Einführung

Die in diesem Bericht zusammengefassten Interviews fanden im Zeitraum von September bis November 2006 statt. Geführt wurden die Interviews per Telefon- und Videokonferenz mit dem Ziel, den Bedarf der D-Grid-Communities in Bezug auf Authentifizierungs- und Autorisierungsfunktionen für eine deutsche Grid-Infrastruktur aufzunehmen. Auf Basis der Ergebnisse der Interviews soll ein Prototyp für eine AA-Infrastruktur im D-Grid konzipiert werden, der im Idealfall die im D-Grid unterstützten Grid-Middlewares Globus Toolkit 4, gLite und UNICORE vollständig einbezieht.

An den Interviews nahmen neben den Mitgliedern des FG3-4 (AAI) des D-Grid Integrationsprojekts (DGI) Partner aus den folgenden Communities teil:

- TextGrid
- AstroGrid
- C3Grid
- MediGrid
- InGrid

Die HEP-Community stand bis zum Erstellungszeitpunkt des Berichts nicht für ein Interview zur Verfügung.

Erfragt wurden die Anforderungen der Communities bezüglich Themen aus dem VO-Management, der Authentifizierung von Nutzern und Diensten, sowie der Autorisierung von Nutzern für den Zugriff auf Grid-Ressourcen. Die Diskussionen zwischen DGI FG1.10 und den Communities über ein VO-Rahmenkonzept erfolgte separat, führte in überlappenden Bereichen (hier u. a. Kap. 4 VO-Konzept) zu vergleichbaren Ergebnissen.

2. Allgemeine Angaben

2.1. Anzahl der beteiligten Einrichtungen

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Projektpartner	k.A.	k.A.	15	gesamt 19	gesamt 9
assoziierte Partner	k.A.	k.A.	3	s.o	s.o

Tabelle 1: Anzahl der an den Communities beteiligten Einrichtungen

2.2. Anzahl der Nutzer innerhalb der Community

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Nutzer (momentan)	200	k.A.	k.A.	k.A.	25
Nutzer (mittelfristig)	10% der Germanistik-Studenten	k.A.	k.A.	k.A.	k.A.

Tabelle 2: Anzahl der Nutzer in den Communities

3. Middleware und verwendete Software

3.1. Grid-Middleware

Bei den D-Grid-Communities herrscht eine starke Präferenz für die Middleware Globus Toolkit in der Version 4 vor. Einige Communities planen insbesondere in Hinblick auf die im Aufbau befindlichen Rechen-Cluster aus den Sonderinvestitionen des BMBF auch die Nutzung anderer Grid-Middleware wie gLite oder UNICORE.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Middleware (momentan)	Globus Toolkit 4	Globus Toolkit 4	Globus Toolkit 4	Globus Toolkit 4	Globus Toolkit 4
Middleware (mittelfristig, zusätzlich)	gLite	gLite, UNICORE	k.A.	gLite, UNICORE	k.A.

Tabelle 3: Eingesetzte Grid-Middleware

3.2. Grid-Portalsoftware

Grid-Infrastrukturen sind momentan durch die CLI¹-basierte Nutzung nur mit großen Hürden für Erstnutzer zugänglich, daher planen nahezu alle Communities den Einsatz von so genannten Grid-Portalen. Dieser Typ Nutzerinterface ist im Grid-Umfeld noch relativ unerforscht, die Zahl der Lösungen in diesem Bereich folglich relativ gering. Die am weitesten fortgeschrittene Lösung ist die Software Gridsphere, die auf Web-Technologien aufbaut. Aus diesem Grund streben bis auf Text-Grid alle befragten D-Grid-Communities konkret die Nutzung dieser Software an, ihre Unterstützung sollte somit auch im Rahmen der AA-Infrastruktur gewährleistet werden.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Portal	noch offen	Gridsphere	Gridsphere	Gridsphere	Gridsphere

Tabelle 4: Portalsoftware

3.3. Batch-System

Die Anbindung von Compute-Ressourcen, z.B. von Rechner-Clustern an Grid-Umgebungen erfolgt in der Regel mittels so genannter Batch-Systeme. Verbreitet sind hier neben herstellereinspezifischen (NQS, LoadLeveler) und der Sun Grid Engine (SGE) PBS basierte Systeme, wie Torque.

In den D-Grid-Communities existiert eine große Bandbreite an Batch-Systemen, bedingt zum Teil durch die Tatsache, dass bei Projektbeginn eigene, bereits bestehende Systeme als Eigenanteile eingebracht wurden. Vorherrschend sind jedoch die PBS basierten Systeme sowie die Sun Grid Engine. Diese sollten daher vorrangig für eine D-Grid-AAI berücksichtigt werden.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Batch-System	noch offen	Torque, OpenPBS, Rocks	NQS, PBS, PBSPro, LoadLeveler, SGE	Torque, OpenPBS, LFS	Torque, Moab, Maui, SGE

Tabelle 5: In den Communities verwendete Batch-Systeme

¹ CLI = Command Line Interface (Bedienung über eine Kommandozeile)

3.4. Verwendete Betriebssysteme

Im Bereich des Grid-Computing ist das vorherrschende Betriebssystem das freie UNIX-Derivat Linux. Diese Dominanz wird auch in den Angaben der D-Grid-Communities deutlich. Alle Communities verwenden Linux als primäres Betriebssystem, Nennungen anderer Systeme stammen aus der Notwendigkeit, bestehende Compute-Ressourcen an das Grid anzubinden. Hier sind insbesondere AIX (IBM), SuperUX (NEC) und Irix (SGI) vertreten.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Verwendete Betriebssystem	allgemein Linux	SuSe, Scientific Linux 4.x, AIX	SuSE, RedHat, Solaris, AIX, SuperUX	SuSe, Redhat, CentOS, AIX, Irix	Scientific Linux, SuSE, Debian, Redhat, Fedora, AIX

Tabelle 6: In den Communities verwendete Betriebssysteme

4. VO-Konzept

Eine Virtuelle Organisation (VO) ist ein Zusammenschluss von Benutzern und Ressourcen verschiedener Institutionen zum Zweck der Kollaboration in einem gemeinsamen Projekt. Dabei können sich die teilnehmenden Institutionen jeweils mit Personen und Ressourcen an der VO beteiligen, es ist aber auch möglich, sich nur mit Personen oder auch ausschließlich mit Ressourcen zu beteiligen.

Die D-Grid-Communities sehen Bedarf für eine VO-Struktur in ihren jeweiligen Projekten, die Planungen hierzu sind unterschiedlich weit fortgeschritten. Für den ersten Projektabschnitt planen die meisten Communities mit einer oder zwei VOs, über die zur Verwaltung benötigten Werkzeuge herrscht bisher weithin noch keine Klarheit. In der AstroGrid Community wurde bereits eine Entscheidung zugunsten von VOMRS gefällt.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
VO-Struktur	zwei VOs in Community: Editionsphilologischen und Wörterbuch	VO „AstroGrid“ ist angelegt. Einsatz von VOMRS zur Verwaltung.	VO-Konzept wird noch diskutiert.	In der Entwicklungsphase ist eine VO ausreichend.	Bedarf für VOs vorhanden.

Tabelle 7: VO-Struktur – Status

Die Planungen der D-Grid-Communities sehen eine Ausweitung der Anwendung des VO-Konzeptes vor. Weiterhin ist die Verwendung von Shibboleth zur Verwaltung von VO-Mitgliedschaften im Rahmen von InGrid geplant.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
VO-Struktur	k.A.	Zusätzliche instituts-spezifische VOs für nicht VO- (lokale) Ressourcen in Planung.	k.A.	Vorgesehen sind 12 VOs für die Anwendungsprojekte und vier Kernmodule.	Evtl. Verwendung von Shibboleth.

Tabelle 8: VO-Struktur – Planungen für die Projektlaufzeit

5. Authentifizierung

Authentifizierung ist das Beweisen einer Identität. Die sichere Identifizierung von Nutzern und Diensten ist eine unabdingbare Grundvoraussetzung für den Betrieb einer sicheren Grid-Infrastruktur. Im Grid-Umfeld hat sich die Verwendung von X.509-Zertifikaten, die von akkreditierten Certificate Authorities (CA) ausgestellt werden, als Standard durchgesetzt. In Zukunft erscheint jedoch auch eine weitere Verbreitung von Shibboleth als Ergänzung bzw. Substitution von Public Key Infrastructures (PKI) denkbar. Hierbei ist allgemein zu beachten, dass Shibboleth lediglich zur Authentifizierung von Nutzern und zur Vergabe von Attributen eingesetzt werden kann.

5.1. Einsatz von PKI

Die D-Grid-Communities setzen zum Zeitpunkt der Interviews ausschließlich auf X.509-Zertifikate zur Authentifizierung von Nutzern und Diensten. Text-Grid und C3Grid gehen jedoch davon aus, dass sie zu einem späteren Zeitpunkt auf Shibboleth zur Authentifizierung umsteigen können werden.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Verwendung von X.509 Zertifikaten für Authentifizierung von Nutzern und Diensten	Ja, in Phase 1	Ja.	In Generation 1 von C3Grid ja	Ja.	Ja.

Tabelle 9: Verwendung von X.509 Zertifikaten – Status

Die meisten D-Grid-Communities haben die AAI ihrer ersten Grid-Prototypen auf Basis von PKI aufgebaut. Zukünftig tendieren die Communities Text-Grid und C3Grid jedoch dazu, zusätzlich oder als Ersatz von PKI Shibboleth zu verwenden. Auch in MediGrid und InGrid wird eine AAI auf Basis von Shibboleth untersucht werden. Unter Berücksichtigung zukünftiger rechtlicher Rahmenbedingungen wird in MediGrid der Einsatz ergänzender Verfahren (u. a. Heilberufeausweis) untersucht.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Planung bzgl. Authentifizierung von Nutzern und Diensten für die Zukunft	Unsicher, ob in Phase 2 weiterhin mit Zertifikaten gearbeitet wird. Untersuchung von Shibboleth.	k.A.	Lösung zur Authentifizierung auf Basis von Shibboleth angestrebt.	Der Medizinbereich bekommt in Zukunft Heilberufeausweis. Dieser soll auch im Grid eingesetzt werden.	Untersuchung von Shibboleth.

Tabelle 10: Verwendung von X.509 Zertifikaten – Planungen für die Projektlaufzeit

5.2. Unterstützte Certificate Authorities

Bei der Verwendung von X.509 Zertifikaten zur Authentifizierung ist es entscheidend, dass beide an dem Vorgang beteiligten Partner den gleichen Certificate Authorities vertrauen. Im europäischen Grid-Umfeld haben sich CAs durchgesetzt, die von der European Grid Policy Management Authority (EUGridPMA) akzeptiert werden. Die beiden deutschen EUGridPMA-konformen Zertifizierungsstellen werden vom Forschungszentrum Karlsruhe (GridKa-CA) und vom DFN-Verein (DFN-Grid-CA) betrieben.

Die D-Grid-Communities haben sich alle dafür entschieden, Zertifikate sowohl vom Forschungszentrum Karlsruhe als auch vom DFN-Verein anzuerkennen und zu nutzen. Zusätzlich akzeptieren zwei Communities (Text-Grid und InGrid) auch Zertifikate von weiteren CAs, um externe Partner zu unterstützen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Unterstützte CAs	Sowohl Zertifikaten von GridKa als auch vom DFN-Verein wird vertraut. Evtl. auch weitere EuGridPMA kompatible CAs.	Sowohl Zertifikaten von GridKa als auch vom DFN-Verein wird vertraut.	Sowohl Zertifikaten von GridKa als auch vom DFN-Verein wird vertraut.	Sowohl Zertifikaten von GridKa als auch vom DFN-Verein wird vertraut.	Sowohl Zertifikaten von GridKa als auch vom DFN-Verein wird vertraut. Eigene CAs von den Partnern werden u.U unterstützt.

Tabelle 11: Durch die Communities unterstützte Certificate Authorities

5.3. Vergabe von Accounts auf Ressourcen

Eine wichtige Funktionalität von Grid-Middlewares ist deren Fähigkeit, User, die durch den Distinguished Name (DN) im Zertifikat oder durch Attribute aus dem VO-Management gekennzeichnet sind, auf lokale Accounts auf den Ziel-Ressourcen abzubilden. Die Abbildung von Nutzer auf Account wird dabei in Globus Toolkit basierten Grid-Umgebungen in vielen Fällen mittels des so genannten grid-mapfile durchgeführt. Hierbei bekommt jeder Nutzer einen eigenen Account (1:1 Abbildung) auf der jeweiligen Ressource zugewiesen. Als Erweiterung sieht die Grid-Middleware gLite eine Abbildung von Gruppen von Nutzern (typischerweise ganze VOs oder Untergruppen von diesen) auf eine Gruppe von Accounts vor. Mit diesen so genannten nicht personalisierten Pool-Accounts wird eine m:n Abbildung ermöglicht.

Die D-Grid-Communities verwenden grid-mapfiles, um Nutzer 1:1 auf Accounts abzubilden. Dabei werden Accounts, die nicht nach dem Nutzer benannt sind, als Pool-Accounts bezeichnet.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Vergabe von Accounts	Noch offen, Abbildung von Nutzergruppen oder auch einzelnen Nutzern auf Accounts wird erwartet	1:1 Mapping von Nutzer (DN) auf Account. Account wird bei Bedarf erstellt.	(personalisierte) Pool-Accounts für Nutzer.	Dynamisches Mapping von Nutzer auf Pool-Account	(personalisierte) Pool-Accounts akzeptabel

Tabelle 12: Vergabe von Accounts – Status

5.3.1. Planungen für die Laufzeit des Projektes

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Vergabe von Accounts	s.o.	s.o.	Wegen Forderungen von Ressourcenprovidern evtl. auch Einzelaccounts	k.A.	k.A.

Tabelle 13: Vergabe von Accounts – Planungen für die Projektlaufzeit

6. Autorisierung

Das Ziel von Zugriffskontrolle bzw. Autorisierung ist es, Aktionen und Operationen von Benutzern gemäß den geltenden Sicherheitsrichtlinien zu beschränken. Dazu gehört auch, Programme einzuschränken, die von einem Benutzer ausgeführt werden oder einem Benutzer Rechte zu verschiedenen Bereichen eines Systems zu gewähren oder zu verwehren.

6.1. Vergabe von Berechtigungen auf Ressourcen

Die Vergabe von Berechtigungen soll in allen befragten D-Grid-Communities mittels lokaler UNIX-Berechtigungen erfolgen. Zu diesem Zweck werden Accounts eingerichtet, auf die die Grid-Nutzer abgebildet werden. Die Abbildung erfolgt über lokale grid-mapfiles, die aus zentralen Listen der Mitglieder einer VO gebildet wird.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Berechtigungsvergabe	lokale grid-mapfiles	lokale grid-mapfiles	lokale Nutzerlisten	zentrale Nutzerliste, Abbildung von Nutzer auf Rolle in VO.	lokale grid-mapfiles

Tabelle 14: Berechtigungsvergabe – Status

Der Prozess der Pflege der grid-mapfiles wird in den Communities unterschiedlich gehandhabt. Es gibt zentrale Verfahren in AstroGrid, C3Grid und MediGrid, lediglich InGrid verwaltet die Listen lokal auf den Ressourcen. AstroGrid verfügt bereits über Mechanismen, die eine automatische Aktualisierung der grid-mapfiles auf den Ressourcen vornimmt.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Pflege von Userlisten auf Ressourcen.	k.A.	Aus VOMRS werden lokale grid-mapfiles gebildet.	Verteilung per E-Mail	Administratoren erhalten Nutzerliste. Lokale Umsetzung der Liste.	momentan lokal

Tabelle 15: Pflege von Userlisten auf Ressourcen der Communities

Zwei der Communities, Text-Grid und InGrid, planen, eine Evaluierung der Autorisierungsfunktionalitäten von Shibboleth im Grid-Umfeld durchzuführen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Berechtigungsvergabe	zentrale Verwaltung, Autorisierung über Shib-Attribute + VO-Management	k.A.	k.A.	k.A.	Evaluation, ggf. Verwendung von Shibboleth

Tabelle 16: Berechtigungsvergabe – Planungen für die Projektlaufzeit

6.2. Use-Cases – Compute Services

6.2.1. Limitierung der maximalen Ressourcennutzung je Nutzer

Der Zugriff auf Rechner-Ressourcen erfolgt – wenn nicht anders konfiguriert – nach dem First-Come-First-Serve-Prinzip, d. h. der Job, der zuerst an der Ressource eintrifft, wird auch zuerst abgearbeitet. Weiterhin gibt es derzeit keine Beschränkung von Nutzern in Bezug auf ihre maximale Ressourcennutzung.

Für den produktiven Betrieb innerhalb der D-Grid-Communities kann dies unter Umständen inakzeptabel werden. Die befragten Communities haben das Problem erkannt, sehen es aber als nicht vordergründig an. Zukünftig kann es für einige der Communities relevant werden, wenn die Zahl der Nutzer innerhalb der Communities steigt oder Erfordernisse aus dem regulären Betrieb die Einführung von QoS-Mechanismen unumgänglich machen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Quotierung von Compute-Ressourcen	Noch unklar. Im Fall von Ressourcenknappheit evtl. erforderlich.	In der Anfangszeit nicht erforderlich	Noch unklar. Im Fall von Ressourcenknappheit evtl. erforderlich.	QoS ist vordergründig. Gäste bekommen nur beschränkten Zugang.	Fairness zwischen Nutzern, Priorisierung von interaktiven gegenüber Batch-Jobs oder umgekehrt.

Tabelle 17: Quotierung von Compute-Ressourcen – Status

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Planungen für die Projektlaufzeit.	k.A.	Zukünftig müssen in D-Grid eingebrachte Anteile der Cluster berücksichtigt werden.	k.A.	k.A.	k.A.

Tabelle 18: Quotierung von Compute-Ressourcen – Planungen für die Projektlaufzeit

6.2.2. Priorisierung von Nutzern auf Compute-Ressourcen

Die Klassifizierung von Nutzern nach ihrer Bedeutung für die jeweilige Community und eine daraus abgeleitete Priorisierung ihrer jeweiligen Jobs kann bei Ressourcenknappheit zu einer effizienteren Allokation von Ressourcen führen.

Die befragten D-Grid-Communities haben zu einer solchen Hierarchisierung von Nutzern keine einheitliche Haltung. Text-Grid, C3Grid und InGrid lehnen sie ab und streben eher eine freiwillige Priorisierung von Jobs an, während AstroGrid und MediGrid mittelfristig einen solchen Mechanismus anstreben.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Bevorzugung bestimmter Nutzer	Nein, evtl. aber von bestimmten Jobs	Im Prinzip ja, aber in den nächsten ein bis zwei Jahren nicht relevant.	Der Nutzer soll zwischen eiligen und weniger wichtigen Jobs unterscheiden können.	Aus QoS-Sicht soll es verschiedene Rollen mit verschiedenen Prioritäten geben.	Für D-Grid Ressourcen sind alle Nutzer gleichwertig.

Tabelle 19: Priorisierung von Nutzern

6.2.3. Autorisierung von Lizenznutzung

Die Verwendung von lizenzpflichtiger Software ist auch in vielen Bereichen des wissenschaftlichen Umfelds unumgänglich. Daraus ergeben sich für Grid-Umgebungen sowohl wirtschaftliche als auch administrative und operative Konsequenzen.

Lizenzpflichtige Software wird momentan von den Communities AstroGrid und InGrid im Grid eingesetzt und pragmatisch berechtigten Grid-Nutzern zur Verfügung gestellt. Dies kann z. B. über eine host-basierte Autorisierung erfolgen, nach der Compute-Nodes im Grid pauschal für die Nutzung dieser Lizenzen berechtigt werden.

Für die anderen Communities spielt die Nutzung von lizenzpflichtiger Software entweder keine Rolle (Text-Grid) oder wird erst in späteren Projektphasen vorgesehen (C3Grid, MediGrid).

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Autorisierung der Nutzung lizenzpflichtiger Software	Verwendung von Open Source Software. Content ist evtl. DRM geschützt.	Momentan werden verfügbare Lizenzen den Nutzern unkompliziert zur Verfügung gestellt.	Momentan kein Problem.	Momentan kein Problem.	Noch zu diskutieren. Momentan werden Lizenzen host-basiert vergeben.

Tabelle 20: Autorisierung der Nutzung lizenzpflichtiger Software – Status

Mittelfristig werden in allen Communities außer Text-Grid Autorisierungsmechanismen für die Nutzung dieser Ressourcen benötigt. AstroGrid sieht darüber hinaus einen Bedarf an Accounting- und Billing-Verfahren.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Autorisierung der Nutzung lizenzpflichtiger Software	k.A.	In Zukunft wird es evtl. explizite Berechtigungen und Gebührenerhebung geben müssen.	Zukünftig muss es evtl. Berechtigungen für die Nutzung lizenzpflichtiger Software geben.	Zukünftig wird es Applikationen geben, die nur von bestimmten Rollen (Administrator, Nutzer, Entwickler) genutzt werden dürfen.	Zukünftig Berechtigungsvergabe erwünscht.

Tabelle 21: Autorisierung der Nutzung lizenzpflichtiger Software – Planungen für die Projektlaufzeit

6.2.4. Autorisierung spezieller Ressourcen

Im Grid-Umfeld ist die Nutzung generischer Speicher- und Rechner-Ressourcen das Hauptanwendungsgebiet. Für einige Communities kann jedoch darüber hinaus auch die Nutzung spezieller Ressourcen über Grid-Interfaces sinnvoll sein. Diese speziellen Ressourcen können z.B. Beschleunigerboards für bestimmte Rechenoperationen sein oder auch robotische Teleskope in der AstroGrid-Community.

Diese speziellen Ressourcen sind in der Regel sowohl in der Anschaffung als auch im Betrieb sehr kostenintensiv und sollten daher nur von autorisierten Nutzern in Anspruch genommen werden können. Eine solche Funktionalität ist daher in einem Autorisierungskonzept für das D-Grid mit zu berücksichtigen.

Für drei Communities (Text-Grid, AstroGrid und C3Grid) ist eine solche Funktionalität mittelfristig von Bedeutung.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Vergabe von Zugriffsrechten auf speziellen Ressourcen	Zukünftig vermutlich ja.	Zukünftig ja, z.B. für robotische Teleskope.	Zukünftig vermutlich ja.	Keine diesbezügliche Planung.	Keine diesbezügliche Planung.

Tabelle 22: Vergabe von Zugriffsrechten auf speziellen Ressourcen

6.2.5. Exklusive Nutzung von Ressourcen

Teile des wissenschaftlichen Umfelds sind heute – ähnlich wie die Industrie – von starkem Wettbewerb geprägt. In diesen Bereichen kann es von entscheidender Bedeutung sein, dass die Wettbewerber keine Informationen über die eigene Arbeit erlangen. Aus diesem Grund ist es für eine

gute Akzeptanz von Grid-Infrastrukturen von Bedeutung, dass Funktionalität bereitgestellt wird, die es ermöglicht, einzelne Rechnerknoten exklusiv zu nutzen, so dass andere Nutzer keine Einsicht in die eigenen Jobs erhalten.

Derartige Mechanismen sind für die InGrid Community zwingend erforderlich, für MediGrid im Fall einer Industriebeteiligung, z.B. während der Durchführung von Arzneimittelstudien.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Exklusive Nutzung von Ressourcen	Wird nicht benötigt.	Wird nicht benötigt.	Im Allgemeinen nicht benötigt.	Nur im Falle der Beteiligung von Industrie wichtig.	Zwingend erforderlich.

Tabelle 23: Autorisierung der exklusiven Nutzung von Ressourcen

6.3. Use-Cases – Data Services

6.3.1. Autorisierung des Zugriffs auf temporärem und permanenten Speicher

Die Unterscheidung von temporärem Speicher, der üblicherweise zum Ablegen von Zwischenergebnissen genutzt wird, und permanentem Speicher, der für Rohdaten oder Endergebnissen verwendet wird, findet auch im Grid-Umfeld Anwendung.

Einige Communities sehen vor, explizite Berechtigungen für die Nutzung so genannten Archiv-Speichers zu vergeben. Dies trifft insbesondere auf die Communities Text-Grid und C3Grid zu. Auch AstroGrid und InGrid benötigen explizite Berechtigungen für die Nutzung von Archiv-Speicher, während MediGrid an das Grid angebundene Speicher-Ressourcen nur als Zwischenspeicher ansieht und plant, zu archivierende Daten auf externen Systemen abzulegen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Unterscheidung zwischen temporärem und permanentem Speicher.	Es werden Berechtigungen für die Nutzung permanenten Speichers benötigt.	Vermutlich ja.	Nicht jeder Nutzer darf permanenten Speicher nutzen.	Grid-Ressourcen werden eher als Zwischenspeicher gesehen.	Ja.

Tabelle 24: Unterscheidung zwischen temporärem und permanentem Speicher

6.3.2. Quotierung von Speicher-Ressourcen

Die Einhaltung von Grenzen zur maximalen Nutzung von Speicherplatz, so genannte Quotas, ist bei beschränkten Ressourcen eine entscheidende Funktionalität von Grid-Umgebungen. Diese Quotas können für einzelne Nutzer bzw. ganze Gruppen oder VOs vergeben werden.

Kurzfristig erscheint diese Funktionalität den meisten Communities nicht wichtig zu sein, da die meisten erst im Aufbau ihrer Grid-Umgebungen begriffen sind und noch nicht über große Nutzergruppen verfügen. Mittelfristig sehen jedoch alle befragten Communities derartige Mechanismen vor, deren genaue technische Ausgestaltung den meisten jedoch noch unklar ist. In den meisten Fällen wird sie auf den Basismechanismen des auf den Speicher-Ressourcen installierten Betriebssystems beruhen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Beschränkung von Speicherplatz für Nutzer	Vorstellungen noch vage, vermutlich festes Quota für jeden Nutzer.	Quotierung wird benötigt, Mechanismen sind jedoch noch unklar.	Im Moment nicht, perspektivisch benötigt.	Es ist die Vergabe von Prioritäten für Nutzer oder Rollen vorgesehen.	Ja.

Tabelle 25: Beschränkung von Speicherplatz für Nutzer

6.3.3. Setzen von Rechten auf Speicherressourcen durch den Administrator

Ein sicherer Betrieb von Grid-Umgebungen setzt eine sichere Grundkonfiguration derselben voraus. So ist es z.B. essentiell, dass Speicher-Ressourcen so konfiguriert werden, dass Benutzergruppen nur Zugriff – insbesondere schreibend, aber auch lesend – auf die Daten der eigenen Gruppe erhalten und Daten anderer für sie nicht sicht- oder zugreifbar sind.

Kurzfristig genügen den D-Grid-Communities die Rechte, die ein UNIX-Betriebssystem für Dateien bietet. Langfristig erfordert der Datenschutz für personenbezogene Daten von MediGrid höhere Schutzlevel, die Zugriffsschutz für Teile von Dateien erfordert. Diese können evtl. durch den Einsatz von Multi-Level-Verschlüsselung befriedigt werden. Im Detail ist dies im Laufe des Projektes noch zu klären.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Setzen von Rechten auf Speicherressourcen durch den Administrator	Verzeichnisse bzw. Dateien abzusichern steht im Vordergrund.	VO-basierte Vergabe von Rechten denkbar. UNIX-Rechte bilden unteren Level.	Funktionalität von UNIX-Systemen ist ausreichend.	Kurzfristig ist eine Autorisierung auf Dateiebene vorgesehen. Langfristig Vergabe von Zugriffsrechten für Teile von XML-Dateien.	UNIX-basierte Rechte ausreichend.

Tabelle 26: Setzen von Rechten auf Speicherressourcen durch den Administrator

6.3.4. Setzen von Rechten auf Speicherressourcen durch den Nutzer

Neben einer sicheren Grundkonfiguration von Zugriffsrechten auf Speicher-Ressourcen im Grid ist es für den Nutzer in vielen Fällen wichtig, die Default-Rechte weiter einschränken zu können. Außer MediGrid reichen für alle befragten Communities hier die UNIX-Rechte aus. MediGrid benötigt langfristig die Funktionalität, Zugriffsrechte aus den Rollen des Nutzers bilden zu können. Diese müssen bei Bedarf auch kombinierbar sein.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Setzen von Rechten auf Speicherressourcen durch den Nutzer	Benutzer sollen die Rechte für ihre Verzeichnisse und Dateien selbst vergeben können.	k.A.	UNIX-Rechte ausreichend.	Kurzfristig: Unix-Rechte Langfristig: Kombinierte Zugriffsrechte (nach Rolle, Einrichtung, VO, Beruf, Rolle und Beruf zusammen)	UNIX-basierte Rechte ausreichend.

Tabelle 27: Setzen von Rechten auf Speicherressourcen durch den Nutzer

6.3.5. Replikation von Dateien

Grid-Umgebungen basieren auf stark verteilten Ressourcen. Diese sind durch Weitverkehrsnetze verbunden, an deren Übergängen zu den lokalen Netzen zusätzlich Firewalls für eine Begrenzung des Datendurchsatzes sorgen. Somit ist es in vielen Fällen sinnvoll, dass Compute-Ressourcen und Ressourcen zur – ggf. temporären – Speicherung von Input-Daten in einem gemeinsamen lokalen Netz angeordnet sind. Dies erfordert in vielen Fällen ein Kopieren der Dateien vor Beginn der Berechnung. Über diese Replikas muss bzgl. Anzahl und Position im Grid buchgeführt werden; weiterhin ist sicherzustellen, dass sie mit denselben Zugriffsrechte versehen sind wie das Original.

Die Notwendigkeit für die Replikation von Daten im Grid ist mit Ausnahme von InGrid in allen befragten D-Grid-Communities vorhanden. Eine spezielle Autorisierung hierfür durch den Nutzer ist

dafür von Text-Grid, AstroGrid und MediGrid erwünscht. C3Grid geht davon aus, dass Replikation ein automatischer Prozess ist, auf den der Nutzer keinen Einfluss haben sollte.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Autorisierung des Anlegens von Replikas	Ja, sowohl prinzipiell als auch durch Einschränkungen der Ressourcen.	Ja, bei Simulationen notwendig	Nein, Anlegen von Replikas ist ein automatischer Prozess in C3Grid.	Ja.	Nein.

Tabelle 28: Autorisierung des Anlegens von Replikas

6.3.6. Use-Cases – Freigabe von Daten in Informationssystemen

In Grid-Umgebungen werden von verschiedenen Diensten Informationen über Aktionen von Nutzern gespeichert, so z.B. Informationen über deren Jobs. Derartige Metadaten können unter Umständen sensibel sein, da sie Dritten Rückschlüsse über die Aktionen der Nutzer ermöglichen. In Communities mit Beteiligung aus der Industrie kann dies aus Wettbewerbsgründen problematisch sein. Auch in Communities, die starken datenschutzrechtlichen Anforderungen unterliegen, werden diese Daten zu schützen sein. Auf der anderen Seite ist es in kollaborativen Umgebungen von Vorteil, wenn Mitglieder einer Gruppe, die an einem Problem arbeiten, Einblick nehmen können in die Arbeit anderer Gruppenmitglieder.

Die befragten D-Grid-Communities sehen die Freigabe von Metadaten generell als problematisch an. Drei Communities (Text-Grid, C3Grid und InGrid) lehnen diese ab und planen nicht, diese Funktionalität zu implementieren. Für MediGrid wird diese Funktionalität einzuschränken sein, wenn in einer späteren Projektphase Partner aus der Industrie beteiligt werden. AstroGrid sieht eine solche Funktionalität vor, wünscht jedoch seinen Mitgliedern die Möglichkeit zur Einschränkung der Freigabe zur Verfügung zu stellen.

	Text-Grid	AstroGrid	C3Grid	MediGrid	InGrid
Freigabe von Daten in Informationssystemen	Nicht vorgesehen.	Der Benutzer soll die Möglichkeit haben, den Zugriff auf Metadaten von Jobs zu beschränken.	Nein.	Im Falle der Beteiligung von Industrie problematisch.	Nein.

Tabelle 29: Freigabe von Daten in Informationssystemen

7. Zusammenfassung und Ausblick

In den Communities wird derzeit hauptsächlich Globus Toolkit 4 verwendet, in den späteren Phasen der Projekte wird aber auch auf gLite und UNICORE zurückgegriffen. Neben der Kommandozeile wird auch auf GridSphere als Zugangsweg zum D-Grid gesetzt.

Der Bedarf der Communities an VO Management Tools ist noch nicht klar umrissen, da die Planungen in den meisten Communities noch nicht weit fortgeschritten sind. Bisher existieren wenige übergeordnete VOs ohne weitere Struktur; zukünftig wird jedoch eine feinere Struktur angestrebt.

Die Authentifizierung von Nutzern und Diensten wird derzeit von allen Communities mit X.509-Zertifikaten von GridKa und DFN durchgeführt. Eine Unterstützung von Shibboleth/GridShib ist zukünftig notwendig.

Für die Communities ist beim aktuellen Stand der Projekte eine Autorisierung mittels grid-mapfile ausreichend. Zukünftig muss jedoch nach Attributen autorisiert werden können, insbesondere im Hinblick auf den Zugriff auf spezielle Ressourcen, auf lizenzpflichtige Software und zur Quotierung des Ressourcenverbrauchs.

7.1. Ausblick D-Grid AAI-Prototyp

Die Ergebnisse aus den Interviews der Communities führen zu einer zweiphasigen Auslegung des AAI-Prototyps für das D-Grid. In der ersten Phase werden die aktuellen Gegebenheiten berücksichtigt, während die zweite Phase den zukünftigen Anforderungen an Authentifizierung und Autorisierung gerecht wird.

7.1.1. Phase 1

In Phase 1 des AAI-Prototyps werden Globus Toolkit 4, gLite und UNICORE unterstützt. Die Verwaltung der virtuellen Organisationen geschieht mittels VOMRS und VOMS. Die Erstellung der Nutzerlisten geschieht auf Basis der VO-Verwaltung zentral und automatisiert. Für Globus Toolkit 4 und gLite werden grid-mapfiles erzeugt und für UNICORE entsprechende UUDB-Datensätze. Diese werden auf den einzelnen Ressourcen in regelmäßigen Abständen per Cronjob geladen und integriert. Damit ist die Authentifizierung auf Basis von X.509-Zertifikaten im gesamten D-Grid möglich.

Für die Autorisierung wird in Phase 1 nur eine reine Ja/Nein-Entscheidung möglich; nur anhand des DN aus dem Nutzerzertifikat wird entschieden, ob ein Nutzer auf eine Ressource zugreifen darf. Die Abbildung der Nutzer auf UNIX-Accounts geschieht mit einer 1:1-Abbildung, die jedem Eintrag im grid-mapfile bzw. der UUDB einen Account zuweist.

7.1.2. Phase 2 (Skizze)

In Phase 2 des AAI-Prototyps werden die technischen Lösungen aus Phase 1 durch komplexere, aber im Zuge der Vergrößerung der Nutzerzahl notwendige Techniken ergänzt. Als Middlewares werden ebenfalls Globus Toolkit 4, gLite, UNICORE unterstützt. Die Verwaltung der virtuellen Organisationen mit VOMS/VOMRS wird ergänzt durch eine Unterstützung für Shibboleth. Damit einher geht eine Anbindung an die DFN-AAI². Die Nutzer-Authentifizierung mit X.509-Zertifikaten wird im Zuge dessen durch weitere noch festzulegende Mechanismen ergänzt.

Eine Autorisierung der D-Grid-Nutzer geschieht nicht mehr auf Basis einer einfachen Ja/Nein-Entscheidung, sondern durch eine feingranulare Auswertung von VO-Attributen, die jedem Nutzer zugeordnet werden. VO-Attribute sind u.a. Rollenzuweisungen in VOMS, die ggf. für die in Phase 1 erzeugten Zuweisungen der Nutzer zu VOs ergänzt werden müssen. Dies ist u.a. bedingt durch die Anforderungen einzelner Communities nach Priorisierung (siehe 6.2.2), Lizenzierung (siehe 6.2.3) und der Einbeziehung nicht allgemein zugänglicher Ressourcen (siehe 6.2.4). Durch die Auswertung von Attributen wird ein Mechanismus verwendet, der einen geringen Wartungsaufwand hat und weniger fehleranfällig ist als eine grid-mapfile basierte Einzellösung pro Ressource.

Die Abbildung von Nutzern auf Accounts ist dann nicht mehr ausschließlich 1:1 möglich, sondern muss um Pool-Accounts erweitert werden. Auch das dynamische Anlegen von Accounts oder virtuellen Workspaces beim Ressourcenzugriff aufgrund der Attribute eines Nutzers wird in Phase 2 unterstützt.

Phase 2 bleibt damit kompatibel zu Phase 1, ermöglicht aber Ressourcen-Betreibern, eine wesentlich feingranuläre Autorisierung zu verfolgen.

² <http://www.dfn.de/content/dienstleistungen/dfnaai/>