

SLA4D²GRID

Statusbericht und Gap-Analysis: D-Grid Infrastruktur, Middleware und Service Level Agreements

D 1.1

Arbeitspaket AP 1: Anbindung an die D-Grid Infrastruktur

Veröffentlichungsdatum: 20/11/2009

Verantwortlicher Editor: Benjamin Henne

Version: 1.0

Status: Final

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Autoren

Benjamin Henne, Regionales Rechenzentrum für Niedersachsen (RRZN)

Philipp Wieder, Technische Universität Dortmund (TuDo)

Alexander Willner, Universität Bonn (UBO)

Internes Review

Dominic Battré, Technische Universität Berlin (TUB)

Versionshistorie

Version	Datum	Kommentar	Autoren
0.1	11.09.2009		Benjamin Henne
0.2	07.10.2009		Benjamin Henne
0.3	10.10.2009		Philipp Wieder
0.4	30.10.2009		Benjamin Henne
0.5	01.11.2009	Internes Review	Dominic Battré
0.6	09.11.2009		Benjamin Henne
0.7	10.11.2009		Benjamin Henne
0.8	17.11.2009	Kleines Review	Alexander Willner
0.9	18.11.2009	Ergebnisse Projekttreffen	Benjamin Henne
1.0	20.11.2009		Benjamin Henne

Kurzzusammenfassung

Dieser Bericht beschreibt die momentane D-Grid Infrastruktur mit allen existierenden zentralen Diensten. Der Fokus liegt dabei auf denjenigen Diensten, die im Rahmen des Projektes SLA4D-Grid von Interesse sein werden. Die Dienste werden in Übersicht verschaffender Tiefe vorgestellt und Quellen für detaillierte Informationen aufgezeigt. Die vorgestellten Dienste werden im Weiteren bezüglich ihrer Verwendbarkeit im Rahmen von SLA4D-Grid beurteilt und existierende Lücken aufgezeigt.

Die Betrachtung im Rahmen dieses Berichtes ist partiell nur mäßig detailliert, da zum Zeitpunkt seiner Erstellung, im Projektmonat 4, detaillierte Informationen und Anforderungen aus den anderen Arbeitspaketes des Projektes noch fehlen oder noch nicht klar definiert sind.

Inhaltsverzeichnis

1	Einleitung.....	5
2	Infrastruktur und zentrale Dienste des D-Grids	5
2.1	Überblick.....	5
2.2	Zentrale Dienste	6
2.2.1	Management von Nutzern und Virtuellen Organisationen	6
2.2.2	Management von Ressourcen	7
2.2.3	Monitoring	9
2.2.4	Accounting	11
2.3	Infrastruktur.....	13
2.3.1	Authentifizierungs- und Autorisierungsinfrastruktur	13
2.3.2	Netzwerk-Infrastruktur.....	14
2.4	Referenzinstallation und Betriebskonzept	14
2.4.1	Referenzinstallation und TAB.....	14
2.4.2	Betriebskonzept	15
3	Verwertbarkeit existierender D-Grid Dienste und Gaps	15
3.1	Management von Benutzern und Virtuellen Organisationen	15
3.2	Management von Ressourcen.....	16
3.3	Monitoring.....	17
3.4	Accounting.....	19
3.5	Referenzinstallation und Betriebskonzept	19
4	D-Grid-Middleware	20
4.1	Eingesetzte Middleware.....	20
4.2	Fokus von SLA4D-Grid.....	20
4.3	SLA-Unterstützung der Middleware.....	20
5	Offene Punkte	21
6	Referenzen.....	24

1 Einleitung

Dieser Bericht beschreibt die aktuelle Infrastruktur des D-Grid und fokussiert dabei die Komponenten, die für die Anbindung der SLA-Schicht an die D-Grid Infrastruktur von Interesse sind. Die Komponenten der D-Grid Infrastruktur werden im Folgenden beschrieben und es wird betrachtet, welche der von ihnen gelieferten Funktionen und Informationen im Rahmen von SLA4D-Grid verwendet werden können.

2 Infrastruktur und zentrale Dienste des D-Grids

Dieses Kapitel gibt einen Überblick über die Infrastruktur und die zentralen Dienste des D-Grid. Diejenigen Komponenten, die für SLA4D-Grid von Interesse sind, werden im Folgenden genauer betrachtet, um zu zeigen welche hiervon genutzt werden können, welche Schnittstellen existieren und welche eventuell noch geschaffen werden müssen.

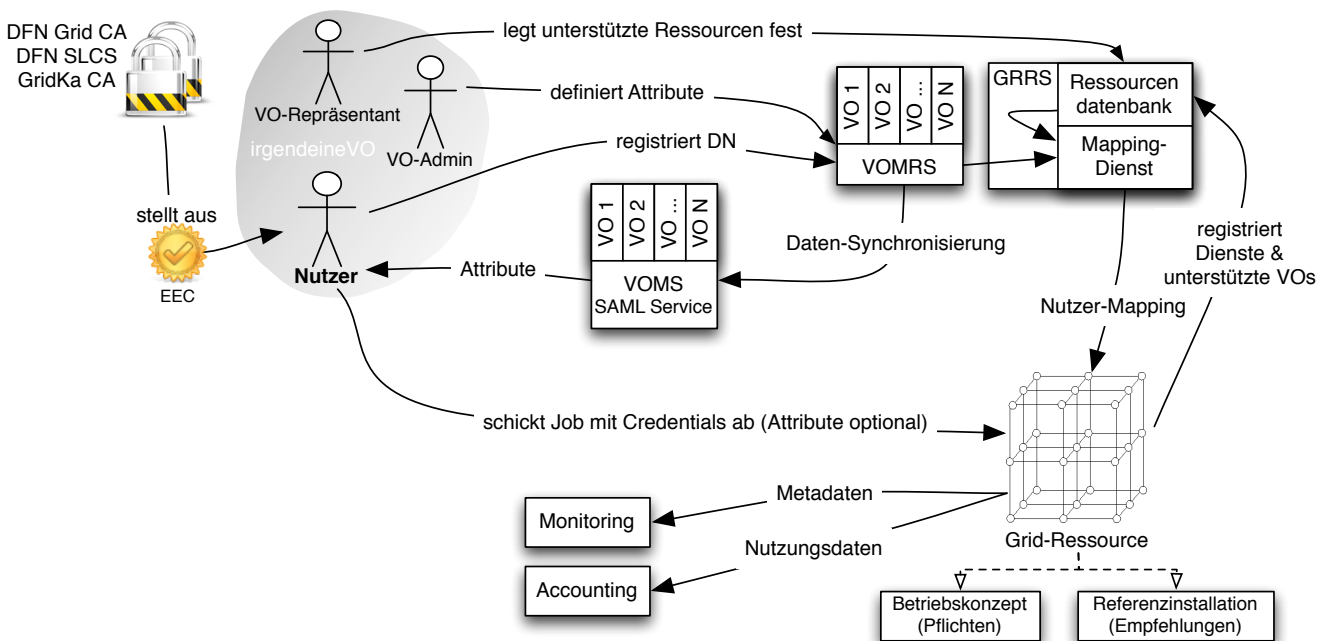


Abbildung 1: D-Grid AAI und Ressourcenverwaltung

2.1 Überblick

Abbildung 1 zeigt die aktuelle Authentifizierungs- und Autorisierungsinfrastruktur (AAI) des D-Grid, welche eng mit der Ressourcen- und Dienste-Verwaltung verknüpft ist.

Benutzer authentifizieren sich mittels X.509-Nutzer-Zertifikaten, die von einer EUGridPMA-/IGTF-akkreditierten Zertifizierungsstelle (in Deutschland: GridKa CA, DFN-Grid-CA und DFN-SLCS¹) ausgestellt werden. Um Grid-Ressourcen nutzen zu können, muss sich ein Nutzer bei mindestens einer Virtuellen Organisation (VO) über den jeweiligen *Virtual Organization Membership Registration Service* (VOMRS) registrieren. Der VOMRS speichert die Zuordnung von Benutzern (individuell identifiziert durch den Distinguished Name des verwendeten Zertifikats) zu VOs.

¹ SLCS: Short Lived Credential Service – CA für die Ausstellung kurzlebiger Zertifikate

Ressourcen werden beim Grid Resource Registry Service (GRRS) registriert, um diese innerhalb des D-Grid nutzbar zu machen. Neben Informationen über die installierte Grid-Middleware werden im GRRS für jede Ressource Informationen gespeichert, die angeben, welche VOs Zugriff auf die Ressource haben. Außerdem legt der Repräsentant einer VO im GRRS fest, welche Ressourcen von seiner VO genutzt werden sollen. Ressourcen authentifizieren sich durch X.509-Host-Zertifikate, die von einer der zwei deutschen akkreditierten Grid CAs (nicht SLCS) ausgestellt werden.

Aus den Informationen des VOMRS (Zuordnung VO und Nutzer) und des GRRS (bidirektionale „Absichtserklärung“ zwischen Ressourcen und VO) kann im Folgenden eine Liste von zugriffsberechtigten Nutzern für jede Ressource erstellt werden. Der Zugriff auf eine Ressource wird immer für komplette VOs und nicht für einzelne Nutzer festgelegt. Über den Mapping-Dienst *dgridmap* können Ressourcen (nur) die für sie gültige Zugriffskontrollliste (grid-mapfile/UUDB-Daten) abrufen.

Die Software-Installation der Grid-Ressourcen wird durch das D-Grid Betriebskonzept und die Referenzinstallation gesteuert: In beiden ist festgelegt, wie Merkmale umgesetzt werden müssen, beziehungsweise empfohlen, wie sie umgesetzt werden können.

Für die Überwachung der Ressourcen melden diese Informationen über ihren aktuellen Systemzustand an die Monitoring-Dienste der Grid-Middleware. Führt ein Benutzer einen Grid-Job aus, so meldet die genutzte Ressource einen entsprechenden Usage-Record an den Accounting-Dienst.

Für besondere Berechtigungen auf den Grid-Ressourcen können Nutzern VO-spezifische Attribute im VO-Managementsystem (im VOMRS) zugewiesen werden. Nutzer-Attribute können vom VOMS-Server, der regelmäßig via Push-Verfahren durch den VOMRS synchronisiert wird, als Attribut-Zertifikat oder als SAML Attribute Assertion von den Nutzern abgefragt werden. Eine Attribut-basierte Autorisierung auf den Ressourcen ist momentan noch vielen Einschränkungen unterworfen, nur für wenige Sonderrollen geplant und noch in der Umsetzung.

2.2 Zentrale Dienste

2.2.1 Management von Nutzern und Virtuellen Organisationen

Zur Authentifizierung werden im D-Grid X.509-Zertifikate verwendet. Um im D-Grid einsetzbar zu sein, muss eine Zertifizierungsstelle (CA) von der EUGridPMA/IGTF² akkreditiert sein. Nutzer-Zertifikate müssen von einer deutschen akkreditierten CA ausgestellt werden. Diese sind momentan die GridKa CA, die DFN-PKI Grid-CA und der DFN-SLCS. Host-Zertifikate müssen von einer der deutschen Grid CA (nicht SLCS) ausgestellt sein. Dies sind momentan die GridKa CA und die DFN-PKI Grid-CA [6][1]. Service-Zertifikate werden im D-Grid nicht eingesetzt. Nutzer werden im D-Grid durch den *Distinguished Name* (DN) ihres Zertifikates eindeutig beschrieben.

Für das Management von Nutzern einer Virtuellen Organisation steht jeder VO eine VOMRS-Instanz am Forschungszentrum Jülich (FZJ) zur Verfügung. Der *Virtual Organization Membership Registration Service*³ (VOMRS) bietet Nutzern ein Web-Interface für die Registrierung und die Verwaltung ihrer Daten. VO-Administratoren/-Repräsentanten können weitere Aufgaben, wie beispielsweise die Zuordnung von Attributen, über die Web-Schnittstelle ausführen. Die Daten des VOMRS werden täglich

² <https://www.eugridpma.org/>, <http://www.igtf.net/>

³ URL der Dienste siehe [1] bzw. [6]

an den *Virtual Organization Membership Service* (VOMS) am Forschungszentrum Karlsruhe übermittelt.

Für die Abfrage von VO-Attributen (VO-Zugehörigkeit, Gruppen, Rollen) bietet die Attribut-Autorität (AA) VOMS zwei Dienste: Der VOMS-Server ermöglicht die Abfrage der Attribute als Attribut-Zertifikat. Der VOMS SAML-Service ermöglicht die Abfrage dieser Attribute als SAML2 Attribute Assertion. Die Abfrage der Attribute eines Nutzers ist nur durch diesen selbst möglich. Ein Nutzer authentifiziert sich hierzu durch sein Nutzer-Zertifikat oder ein davon abgeleitetes Proxy-Zertifikat. Da nur Nutzer selbst ihre Informationen abfragen können, müssen sie diese selbst bei der AA abfragen und in ihre Credentials/Webservice-Aufrufe integrieren, um sie mit einem Service im Grid nutzen zu können (Attribute-Push).

Für das Monitoring existiert in der Datenbank des GRRS (genauer: in der von außerhalb lesend zugreifbaren Teilkopie in der DMZ des FZJ) die Tabelle `dgrid_dn_vo` (vgl. Abbildung 2), die – aggregiert aus den Datenbanken des GRRS und des VOMRS – Nutzernamen, Distinguished Names und die jeweilige VO-Zugehörigkeit eines Nutzers enthält. Diese Datenbank kann für die Feststellung der VO-Zugehörigkeit eines Nutzers verwendet werden.

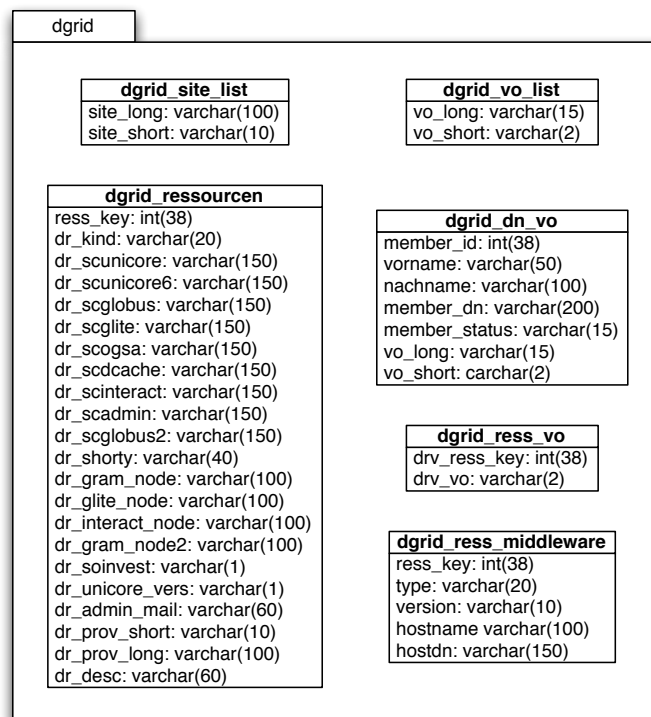


Abbildung 2: Datenbankschema D-Grid GRRS MySQL-Datenbank für Außenzugriff (Stand August 2009)

2.2.2 Management von Ressourcen

Für das Management der Ressourcen des D-Grid wird am Forschungszentrum Jülich der *Grid Resource Registry Service* (GRRS) betrieben. Der GRRS speichert Informationen über die Virtuellen Organisationen des D-Grid und über vorhandene Grid-Ressourcen. Die momentane Implementierung des GRRS baut auf einer Oracle Datenbank auf. Das Schema der Datenbank (Stand Juli 2008) ist in [7] beschrieben und lässt sich wie in Tabelle 1 dargestellt zusammenfassen.

Die ersten sieben Datenbanktabellen in Tabelle 1 beschreiben die Ressourcen, dessen Administratoren, die VOs, dessen Administratoren und die (gegenseitige) Absichtserklärungen der Nutzung bzw. Unterstützung beider. Die letzte Tabelle `dgrid_dn_vo` wird aus den vorherigen Tabellen sowie den Daten des VOMRS erzeugt. Sie wurde ursprünglich für das Monitoring erstellt und wird, wie im Folgenden erläutert, für den externen Zugriff bereitgestellt.

Aus der Tabelle `dgrid_ress_vo` (siehe auch Abbildung 2) und den Daten des VO-/Nutzermanagements VOMRS wird das Nutzer-Mapping für die Zugriffskontrolllisten der Ressourcen (`grid-mapfile` bzw. `UADB`) erzeugt, welches diese via `dgridmap` abfragen.

Tabelle	Daten	Details
<code>dgrid_ressourcen</code>	Ressourcen-Informationen	Site, Name, DN, Art, Middleware, URLs, Beschreibung
<code>dgrid_ress_admin</code>	Ress.-Administratoren	Ressource, Name, DN, Adresse, Telefon, Email
<code>dgrid_vo</code>	VO-Informationen	Name, URL, verschiedene Flags
<code>dgrid_vo_admin</code>	VO-Administratoren	VO, Name, DN, Email
<code>dgrid_vo_matrix</code>	VO <i>x</i> unterstützt Ressource <i>y</i>	Ressource, VO
<code>dgrid_ress_matrix</code>	Ressource <i>y</i> unterstützt VO <i>x</i>	Ressource, VO
<code>dgrid_ress_vo</code>	Schnitt der <code>_matrix</code> -Tabellen	Ressource, VO
<code>dgrid_dn_vo</code>	Nutzer & VO-Zugehörigkeit	Name, DN, ID, VOMRS-Status, VO

Tabelle 1: Datenbanktabellen D-Grid GRRS-interne Oracle-Datenbank (Stand Juli 2008)

Für den externen Zugriff auf Informationen des GRRS existiert eine Teilkopie der GRRS-Daten in einer MySQL-Datenbank. Diese Datenbank befindet sich in der DMZ des FZJ und ist von außerhalb des Forschungszentrums via MySQL-Client lesend zugreifbar. Projekten, die dortige Informationen nutzen möchten, wird ein entsprechender Zugang eingerichtet. Ein Zugang für SLA4D-Grid wurde eingerichtet.

Welche Informationen für die externe Nutzung bereitstehen, zeigt das Datenbankschema in Abbildung 2. Neben einigen Tabellen der GRRS-internen Datenbank enthält die externe Datenbank weitere Tabellen, wie in Tabelle 2 zusammengefasst.

Tabelle	Daten
<code>dgrid_ressourcen</code>	(Kopie) Ressourcen-Informationen
<code>dgrid_dn_vo</code>	(Kopie) Nutzer & VO-Zugehörigkeit
<code>dgrid_ress_vo</code>	(Kopie) Schnitt der <code>_matrix</code> -Tabellen
<code>dgrid_ress_middleware</code>	Liste aller Ressourcen-Frontends
<code>dgrid_site_list</code>	Liste der Sites
<code>dgrid_vo_list</code>	Liste der VOs

Tabelle 2: Datenbanktabellen D-Grid externe GRRS MySQL-Datenbank (Stand August 2009)

Weiterentwicklung

Im Rahmen des Projektes DGI-2 wird der GRRS vom DGI-2 Fachgebiet 2 „Betrieb der D-Grid-Infrastruktur“ weiterentwickelt. Das Arbeitspaket 2.3 sieht den Einsatz der *D-Grid Resource Description Language* (D-GRDL) für die Beschreibung von Ressourcen vor, um „Abläufe bei der Integration neuer Ressourcen zu automatisieren“ [5]. Durch die Umstellung der GRRS-Datenbank auf eine XML-Datenbank auf Basis des D-GRDL-

Schemas können die Informationen des GRRS maschineninterpretierbar gemacht werden und automatisch verarbeitet werden. In einem ersten Meilenstein [15] wurde die Objektbezeichnung der GRRS-Daten in der D-GRDL vorgeschlagen. Die Umsetzung der XML-Datenbank basierend auf eXist soll vorerst parallel zum etablierten GRRS geschehen und bis Ende 2009 umgesetzt werden.

2.2.3 Monitoring

Im D-Grid existieren verschiedene Monitoring-Dienste: middleware-spezifische, middleware-übergreifende und middleware-unabhängige Dienste mit historisch gewachsenen nicht einheitlichen Schnittstellen. Der Großteil der Monitoring-Lösungen ist für die Überwachung der Grid-Ressourcen konzipiert. Es gibt jedoch auch Ansätze für das Monitoring von Grid-Jobs.

Jedes im D-Grid eingesetzte Middleware-System bietet und nutzt sein eigenes Monitoring. Dementsprechend ist innerhalb des D-Grid eine Vielzahl verschiedener Monitoring-Systeme installiert:

Das *Monitoring and Discovery System* (MDS) des Globus Toolkit ist Teil einer jeden Globus Toolkit Installation auf den Ressourcen. Gemäß der Referenzinstallation (siehe Abschnitt 2.4.1) enthält der MDS für eine Rechenressource Informationen über die Knoten eines Clusters (gesammelt via Ganglia), Informationen über das Batchsystem (PBS) sowie Informationen über die Site (via Geomaint) an der die Ressource installiert ist. Die lokalen MDS-Dienste schicken die Monitoring-Daten an einen zentralen MDS (Push-Verfahren) am Leibniz-Rechenzentrum (LRZ). Seit Ende August 2009 existiert außerdem ein redundanter zentraler MDS am Forschungszentrum Karlsruhe (FZK), der ebenfalls von den lokalen MDS gespeist wird. D-Grid Sites können einen eigenen Site-Level-MDS betreiben und besitzen diesen zum Teil auch schon. Dieser ist zwischen lokalem und zentralem MDS in der Hierarchie platziert. Als Schnittstelle zum MDS steht eine WSRF-Schnittstelle bereit.

UNICORE 6 bietet mit dem *Common Information Service* (CIS) einen middleware-spezifischen Monitoring-Dienst. Ein zentraler CIS wird vom Forschungszentrum Jülich (FZJ) betrieben. Der CIS verfügt sowohl über eine Web-Schnittstelle als auch eine Service-Schnittstelle.

Die im D-Grid eingesetzte Monitoring-Komponente der gLite-Middleware ist der *Berkeley Database Information Index* (BDII). Der BDII setzt auf ein LDAP-Verzeichnis auf und bietet dadurch eine entsprechende Schnittstelle. Jede Site im D-Grid, die gLite einsetzt, besitzt einen Site-Level-BDII, der Informationen über die Site-lokalen Ressourcen speichert. Regelmäßig werden die Monitoring-Informationen zum zentralen Top-Level-BDII am FZK transferiert. Für den Top-Level-BDII soll eine redundante Installation im Rahmen von DGI-2 Arbeitspaket FG 2.5 am FZJ eingerichtet werden. Die *Relational Grid Monitoring Architecture* (R-GMA) wird im D-Grid nicht mehr eingesetzt und wurde sowohl aus der Referenzinstallation als auch aus dem Betriebskonzept entfernt⁴.

Weitere detaillierte Informationen zu den middleware-spezifischen Monitoring-Diensten, sowie eine Gegenüberstellung dieser, können entsprechenden Vorarbeiten des D-Mon-Projektes [3] entnommen werden.

Im Rahmen des D-Grid Gap-Projektes *D-Mon*⁵ (07/2007 bis 06/2009) wurde ein middleware-übergreifendes Monitoring-System entwickelt. Die entwickelte

⁴ siehe Protokoll der 3. D-Grid TAB-Telefonkonferenz am 13.07.2009 unter <http://www.d-grid.de/tab>

⁵ D-Mon Projektbericht unter <http://www.d-grid.de/index.php?id=401>, Kontakt via d-mon@d-grid.de

Softwarearchitektur [2] baut auf den vorhandenen middleware-spezifischen Monitoring-Systemen auf und aggregiert die Informationen dieser via Pull-Verfahren in einer eigenen SQL-Datenbank. Die Informationen werden dabei von den verschiedenen Darstellungen auf das GLUE 2.0 Schema abgebildet. Dem Nutzer bietet D-Mon ein Gridsphere-basiertes Webfrontend, über das er die gesammelten Informationen einsehen und auf VO-Basis (vgl.: Zugriff je Ressource per VO) filtern kann. Für den Zugriff durch andere D-Grid Services ist eine OGSA-DAI⁶-Schnittstelle vorgesehen. Bisher existiert nur die Installation des D-Mon Testbeds. Ein Übergang in den D-Grid Betrieb durch DGI-2 Fachgebiet 2 (Meilenstein 5-2) ist für Juni 2010 geplant. Voraussichtlich wird dies jedoch schon bis Ende 2009 geschehen. Die Installation für den Produktivbetrieb wird am LRZ erfolgen. Eine redundante Installation wird am FZK aufgebaut. Details zur Kopplung von D-Mon an das existierende D-Grid Monitoring (Anbindung an Site-Level- vs. Top-Level-Dienste) wurden noch nicht festgelegt.

Im Rahmen des D-Grid Integrationsprojektes DGI-2 wird in Fachgebiet 5.5 aufbauend auf dem AMon⁷-System [10] (für LCG/gLite) ein nutzer- und jobzentrisches Monitoring für das Globus Toolkit 4 entwickelt. Der Prototyp [9] sieht folgende Architektur vor: Ein Job-Monitoring-Client speichert während der Job-Laufzeit Job-Informationen in einen Kurzzeit-Datenspeicher. Nach der Beendigung des Jobs werden die Daten in einen (VO-)zentralen Langzeit-Datenspeicher transferiert. Durch eine erweiterte Version der AMon-Nutzerschnittstelle können Nutzer auf Daten im Langzeit-Datenspeicher zugreifen und diese auswerten. Die Datenspeicher werden auf Basis des Globus MDS und der GSI aufgebaut.

Am Forschungszentrum Karlsruhe wird momentan ein Nagios⁸-basiertes Monitoring für D-Grid Dienste entwickelt. Der Fokus der Überwachung liegt dabei auf der Prüfung der Erreichbarkeit bzw. Funktionsfähigkeit der Dienste an den D-Grid Sites. Die Prüfung geschieht durch die Ausführung von Testjobs und Serviceanfragen mit normalen Benutzerrechten durch Benutzer der dafür etablierten VO *dgops*. Über die Nagios-Webschnittstelle können die Ergebnisse der Tests eingesehen werden. Nagios ermöglicht zudem das Ausführen von Aktionen, wie das Versenden von Emails, wenn definierte Fehlerzustände eintreten. Momentan können nur gLite-Dienste überwacht werden, die im D-Grid Top-Level-BDII verzeichnet sind, da eine Kopplung an den GRRS noch nicht realisiert worden ist. GLite-Dienste können geprüft werden, für den Test von Diensten der Globus Toolkit und UNICORE Middleware müssen entsprechende Test noch erstellt werden.

JAWARI⁹ ist ein Benchmarking-Werkzeug für Grid-Ressourcen, entwickelt vom Fraunhofer ITWM. Die Benchmark-Suite besteht momentan aus 19 verschiedenen Tests¹⁰ für verschiedene Komponenten einer Grid-Infrastruktur. Die Tests werden auf den Grid-Ressourcen als Nutzer-Jobs ausgeführt. Die Ergebnisse werden in einer Datenbank aggregiert und durch Scoring-Algorithmen bewertet. Für den Zugriff durch Grid-Nutzer stehen sowohl ein CLI, als auch eine Webservice-Schnittstelle und ein Webfrontend bereit.

Zugriff auf die etablierten zentralen Monitoring-Dienste via Web-Schnittstelle:

- GT 4.0 WebMDS (LRZ) <http://webmds.lrz-muenchen.de:8080/webmds/>
- GT 4.0 WebMDS (FZK) <http://dgrid-mds.scc.kit.edu:8080/webmds/>

⁶ OGSA-DAI-Webseite unter <http://www.ogsadai.org.uk/>

⁷ siehe http://tu-dresden.de/die_tu_dresden/zentrale_einrichtungen/zih/forschung/grid_computing/amon/

⁸ siehe <http://www.nagios.org/>

⁹ siehe <http://jawari.net/monitoring/resource>

¹⁰ siehe <http://jawari.itwm.fraunhofer.de/web/showHelp.do?topicId=5>

- UNICORE CIS (FZJ) <https://dgrid-unic.fz-juelich.de:9113/>
- D-Mon Demo (LRZ) <http://vo-mon.lrz-muenchen.de:8080/gridsphere/gridsphere>

Zugriff auf die etablierten zentralen Monitoring-Dienste via Service-Schnittstelle:

- GT 4.0 MDS (LRZ)
wsrf-query -s https://mds-dgrid.lrz-muenchen.de:8443 \\
/wsrf/services/DefaultIndexService
- GT 4.0 MDS (FZK)
wsrf-query -s https://dgrid-mds.scc.kit.edu:8443 \\
/wsrf/services/DefaultIndexService
- UNICORE CIS (FZJ)¹¹
ucc cis-showallinfo -u "https://dgrid-unic.fz-juelich.de:9113 \\
/services/CISRegistry?res=cis_registry"
- gLite Top-Level BDII (FZK)
ldapsearch -x -H ldap://bdii-dgrid.gridka.de:2170 -b "o=grid"

2.2.4 Accounting

Das momentan im D-Grid durchgeführte Accounting basiert auf dem Einsatz des *Distributed Grid Accounting System* (DGAS). Momentan existiert nur eine Accounting-Lösung für Compute-Jobs im D-Grid und keine für Lizenzen oder Daten.

Zur Erstellung von Accounting-Datensätzen wird auf einer Grid-Ressource der DGAS-Client ausgeführt. Dieser verknüpft Informationen aus den Log-Dateien des eingesetzten Batch-Systems mit Nutzerinformationen und schickt diese an den DGAS-Server. Der zentrale DGAS-Server im D-Grid wird vom Regionalen Rechenzentrum für Niedersachsen (RRZN) betrieben. Virtuelle Organisationen haben die Möglichkeit, ein eigenes DGAS-Accounting-System zu betreiben, um ein VO-internes Accounting durchzuführen. Für das

JOB_TYPE=local	start=1214921353
GlueHostBenchmarkSI00=17.3	end=1214921403
GlueHostBenchmarkSF00=17.1	Exit_status=0
USER_DN=/C=DE/O=GridGermany/OU=Universitaet Hannover/CN=Max Mustermann <gdigrid>	exec_host=gcwn51.d-grid.uni-hannover.de/7
USER_FQAN=/gdigrid/Role=NULL/Capability=NULL	Resource_List.cput=48:00:00
ACCTLOG:07/01/2008 16:10:03;	Resource_List.ncpus=1
E;	Resource_List.neednodes=1
158568.torqued1.d-grid.uni-hannover.de;	Resource_List.nodect=1
user=gd0012	Resource_List.nodes=1
group=gdigrid	Resource_List.walltime=48:00:00
jobname=STDIN	session=16097
queue=dgiseq	resources_used.cput=00:00:27
ctime=1214921350	resources_used.mem=186732kb
qtime=1214921350	resources_used.vmem=1022648kb
etime=1214921350	resources_used.walltime=00:00:50

Tabelle 3: DGAS UsageRecord Beispiel

D-Grid-weite Accounting müssen jedoch alle Accounting-Datensätze im zentralen Accounting gesammelt werden. Hierzu lassen sich die DGAS-Server hierarchisch

¹¹ momentan noch nicht möglich, da bisher nicht via dgridmap an D-Grid AAI angeschlossen

aufbauen. VOs mit einem eigenen DGAS-Server müssen Datensätze an den zentralen Server weiterleiten, haben jedoch die Möglichkeit, die Datensätze zu anonymisieren (siehe Accounting-Konzept [13]).

Da der DGAS-Client die Accounting-Informationen aus den Log-Dateien des Batchsystems gewinnt, ist das Accounting unabhängig von der eingesetzten Grid Middleware, jedoch abhängig vom verwendeten Batchsystem. Momentan werden die Batchsysteme PBS/Torque und LSF unterstützt. Die Unterstützung für SGE wird derzeit am INFN in Italien implementiert.

Ein DGAS-Accounting-Datensatz enthält Informationen über einen Job, wie exemplarisch in Tabelle 3 dargestellt. DGAS verwendet das eigene DGAS Usage Record (UR) Format. Es unterstützt momentan nicht das XML-basierte Usage Record Format [11], welches im Rahmen der OGF Usage Record Working Group (UR-WG) entwickelt wird.

Im Rahmen von DGI-2 Fachgebiet 5.2 wird vom RRZN in Kooperation mit dem INFN eine teilweise Reimplementierung des DGAS auf Basis von Java Enterprise geplant. Bis Ende 2009 soll ein erster Prototyp der Basisimplementierung umgesetzt werden. DGAS-EE wird das Usage Record Format der UR-WG (Beispiel siehe Abbildung 3) einsetzen. Entsprechend werden auch die Client-Applikationen angepasst. Statt proprietären Schnittstellen wird DGAS-EE Webservice-Schnittstellen einsetzen. Die Implementierung soll auf Basis des JBoss Application Servers, Hibernates und Apache ActiveMQs umgesetzt werden. Offen ist, welche Schnittstellen neben denen für die Client-Server-Kommunikation zusätzlich implementiert werden.

DGAS-EE soll anfangs parallel zum bisher eingesetzten legacy DGAS betrieben werden, so dass Nutzer nicht gezwungen sind, sofort zum neuen System zu wechseln. Es soll eine Schnittstelle zwischen legacy DGAS und DGAS-EE erstellt werden, die – fungierend als DGAS-EE-Client – Accounting-Informationen in DGAS-EE einspeist.

Neben dem Accounting von Compute-Jobs wird von DGI-2 Fachgebiet 5.2 am Accounting lizenzpflichtiger Software, kostenpflichtiger Daten sowie von Speicherressourcen gearbeitet. Dieses soll im Rahmen der Entwicklung von DGAS-EE geschehen.

Mit der Überarbeitung des D-Grid Betriebskonzeptes (siehe Abschnitt 2.4.2) soll das Accounting verpflichtend für alle D-Grid Ressourcen durchgesetzt werden. Bisher nimmt nur ein Teil der D-Grid Ressourcen am D-Grid-weiten Accounting teil.

In Bezug auf den Datenschutz der während des Accounting erhobenen und gespeicherten Daten wurde von den DGI-2 Fachgebieten 5.2 und 5.5 ein Vorschlag für die Erweiterung der Acceptable Use Policy (AUP) für Nutzer gemacht [14], der bisher jedoch nicht umgesetzt worden ist.

In Bezug auf die Verfügbarkeit des Accounting-Dienstes und der erhobenen Daten arbeitet das Fachgebiet 5.2 an einem Template für ein SLA zwischen der MediGrid-Community und dem RRZN als Betreiber des Dienstes. Ein solches SLA kann momentan jedoch nur als reine Willenserklärung angesehen werden und ist keine rechtlich bindende Vereinbarung.

```

<?xml version="1.0" encoding="UTF-8"?>
<JobUsageRecord xmlns="http://schema.ogf.org/urf/2003/09/urf"
  xmlns:urf="http://schema.ogf.org/urf/2003/09/urf"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://schema.ogf.org/urf/2003/09/urf
    file:/Users/bekah/Documents/GGF/URWG/urfwg-schema.09.xsd">
  <RecordIdentity urf:recordId=http://www.emsl.pnl.gov/mscf/colony/PBS.1234.0
    urf:createTime="2003-08-13T18:56:56Z" />
  <JobIdentity>
    <LocalJobId>PBS.1234.0</LocalJobId>
  </JobIdentity>
  <UserIdentity>
    <LocalUserId>scottmo</LocalUserId>
  </UserIdentity>
  <Charge>2870</Charge>
  <Status>completed</Status>
  <Memory urf:storageUnit="MB">1234</Memory>
  <ServiceLevel urf:type="QOS">BottomFeeder</ServiceLevel>
  <Processors>4</Processors> <ProjectName>mscfops</ProjectName>
  <MachineName>Colony</MachineName>
  <WallDuration>PT1S</WallDuration>
  <StartTime>2003-08-13T17:34:50Z</StartTime>
  <EndTime>2003-08-13T18:37:38Z</EndTime>
  <NodeCount>2</NodeCount>
  <Queue>batch</Queue>
  <Resource urf:description="quoteId">1435</Resource>
  <Resource urf:description="application">NWChem</Resource>
  <Resource urf:description="executable">nwchem_linux</Resource>
</JobUsageRecord>

```

Abbildung 3: OGF UR-WG Usage Record Beispiel aus [11]

2.3 Infrastruktur

2.3.1 Authentifizierungs- und Autorisierungsinfrastruktur

Sowohl Nutzer als auch Ressourcen werden im D-Grid ausschließlich durch X.509-Zertifikate authentifiziert. Notwendige langlebige Zertifikate (Gültigkeit: 13 Monate) werden von der GridKa CA und der DFN-PKI Grid-CA für alle Grid-Nutzer und Ressourcen ausgestellt. Nutzer haben zudem die Möglichkeit kurzlebige Zertifikate (Gültigkeit ≤ 1 Mio. Sekunden $\approx 13,5$ Tage) ausgestellt durch den DFN-PKI SLCS im D-Grid zu nutzen. Dieser Dienst ermöglicht die spontane Ausstellung von X.509-Zertifikaten nach einer Authentifizierung basierend auf Authentifizierungsmerkmalen, die bei der Heimatorganisation eines Nutzers verwaltet werden. Der organisationsübergreifende Zugriff auf die organisationsintern verwalteten Authentifizierungsmerkmale geschieht dabei über Shibboleth¹². Voraussetzung für die Nutzung des Dienstes ist die Mitgliedschaft einer entsprechenden Organisation in der Shibboleth-Föderation *DFN-AAI*.

Die verschiedenen Ressourcen autorisieren die Nutzer durch die jeweiligen Mechanismen der Grid-Middleware. Das Globus Toolkit und die gLite Middleware nutzen das grid-mapfile, UNICORE die (X)UUDB und dCache die kpwd-Datei.

Für die Autorisierung beziehen die Ressourcen (mindestens einmal am Tag) die notwendigen Informationen über autorisierte Nutzer über das Tool *dgridmap* vom zentralen GRRS. *Dgridmap* authentifiziert sich dazu via Client-Authentifizierung mit dem jeweiligen Host-Zertifikat der Ressource beim GRRS und bezieht automatisch die für die

¹² Shibboleth unter <http://shibboleth.internet2.edu/>

jeweilige Ressource angepasste Autorisierungsliste. Die bezogenen Listen werden darauf in das lokale System importiert.

Die Weiterentwicklung der D-Grid AAI durch DGI-2 FG-3.2 sieht die Einführung einer Attribut-basierten Autorisierung vor. Auf Basis aktueller Implementierungen für die drei verschiedenen Middleware-Systeme beschränkt sich die Umsetzung momentan auf die Autorisierung weniger besonderer Rollen auf spezielle lokale Nutzer-Accounts.

2.3.2 Netzwerk-Infrastruktur

Während des D-Grid Integrationsprojektes 1 (DGI-1) fanden Reviews der Netzwerk-anforderungen der damaligen Communities statt. Es zeigte sich, dass alle Anforderungen an das dem D-Grid zugrunde liegende Kommunikationsnetz mit den Diensten des Wissenschaftsnetzes X-WiN¹³ gut abgedeckt werden können. Speziallösungen wurden von D-Grid Communities nicht gefordert, wurden aber beispielsweise im Rahmen des LHC-Projektes erstellt und vom *DFN Verein* bereitgestellt.

Entstehen erweiterte Anforderungen von D-Grid Communities oder Projekten, so ist der *DFN Verein* interessiert, Lösungsvorschläge für Spezialanforderungen zu erarbeiten.

2.4 Referenzinstallation und Betriebskonzept

Das Betriebskonzept und die Referenzinstallation des D-Grid beschreiben Anforderungen und Pflichten sowie Empfehlungen für den Betrieb und die Installation der zentralen Dienste und der Ressourcen des D-Grid.

2.4.1 Referenzinstallation und TAB

Der Begriff *D-Grid Referenzinstallation* bezeichnet zwei zu unterscheidende Dinge: Einerseits ist die Referenzinstallation der Prototyp einer D-Grid Standard-Site installiert auf Ressourcen am Forschungszentrum Karlsruhe, auf den alle D-Grid-Nutzer wie auf jede andere D-Grid Site zugreifen können. Andererseits – und dies ist meist gemeint – ist sie die Dokumentation wie genau jenes System installiert ist: Die Dokumentation, das DGI-Ref-Wiki¹⁴, ist eine exemplarische Realisierung und Empfehlung, wie ein Ressourcen-Anbieter seine Ressourcen für den Einsatz im D-Grid installieren und konfigurieren kann und teilweise auch muss. Spezielle Hardware (Beispiel Infiniband-Interconnect) wird dabei von der Referenzinstallation nicht berücksichtigt. Auch berücksichtigt die Referenzinstallation an einigen Stellen nur eine von mehreren Alternativen.

Die Referenzinstallation wird im Rahmen des D-Grid Integrationsprojektes DGI-2 im Fachgebiet 1.3 gepflegt und weiterentwickelt. Für die Aktualisierung der Installation definiert die Referenzinstallation einen Release-Zyklus mit je zwei Releases pro Jahr. Die Referenzinstallation an sich hat keine bindende Wirkung und auch keine Möglichkeiten Ressourcen-Anbieter zur Erfüllung von Vorgaben oder zum Wechsel von Software-Versionen zu forcieren.

Größere Veränderungen der Referenzinstallation (Neuaufnahme von Software, Major-Release-Wechsel, End-of-Life von Software-Versionen et cetera) bedürfen der Beratung und dem Beschluss des D-Grid *Technical Advisory Boards*¹⁵ (TAB). Das TAB besteht aus Anwender- und Ressourcen-Anbieter-Vertreter der drei Middleware-Systeme des D-Grid sowie einem Vertreter des D-Grid Integrationsprojektes DGI-2. Es tagt regelmäßig in nicht öffentlichen Treffen.

¹³ Wissenschaftsnetz X-WiN unter <http://www.dfn.de/xwin/>

¹⁴ D-Grid Referenzinstallation Wiki unter <http://dgiref.d-grid.de/>

¹⁵ Technical Advisory Board unter <http://www.d-grid.de/tab/>

2.4.2 Betriebskonzept

Das „Betriebskonzept für die D-Grid Infrastruktur“ legt Regeln und Verfahren zum Betrieb von Ressourcen im Rahmen des D-Grid fest. Es beschreibt Regeln für den Zugang zur D-Grid Infrastruktur für die beteiligten Nutzer, Virtuellen Organisationen und Ressourcen-Anbieter. Es legt Aspekte der Ressourceninstallation fest, die für den Betrieb notwendig sind, sowie die Softwareinstallation der Ressourcen angeschafft aus den Mitteln der BMBF Sondermaßnahmen (früher: Sonderinvestitionen). Außerdem werden die zentralen Dienste des D-Grid im Betriebskonzept beschrieben.

Öffentlich verfügbar ist momentan die Version 1.1c draft des Betriebskonzeptes [6]. Innerhalb des D-Grid Integrationsprojektes wird momentan an einer überarbeiteten Version des Dokumentes [1] gearbeitet. Das Betriebskonzept wird im Rahmen des Projektes DGI-2 unter der Leitung des Fachgebietes 2 in Zusammenarbeit mit den anderen Fachgebieten gepflegt und erweitert.

Rechtliche Regelungen zwischen den verschiedenen Parteien werden durch die *Acceptable Use Policies* (AUPs) festgelegt. DGI-2 bietet hierzu Muster-AUPs. Bisher existiert nur eine AUP für Nutzer¹⁶, die einer Virtuellen Organisation beitreten.

Informationen zum Inhalt des Betriebskonzeptes bietet auch die SuGI¹⁷-Video-Aufzeichnung des Vortrags „Betriebskonzept für D-Grid-Ressourcen“ von T. Fieseler und C. Dohmen vom 04.03.2009:

- http://sugi.d-grid.de/fileadmin/user_upload/sugi_storage//0000/431/02_fieseler.html

3 Verwertbarkeit existierender D-Grid Dienste und Gaps

In diesem Abschnitt werden die existierenden Dienste des D-Grid betrachtet und ihre Verwertbarkeit im Rahmen der SLA-Implementierung beurteilt. Als Grundlage für die Betrachtung dient der in Abbildung 4 dargestellte SLA-Lebenszyklus.

3.1 Management von Benutzern und Virtuellen Organisationen

Das momentane Nutzer- und VO-Management des D-Grids – genauer der GRRS – bietet autorisierten D-Grid Diensten den Zugriff auf Informationen über Nutzer, VOs und der Nutzer VO-Zugehörigkeit. Der auf je-VO-Basis festgelegte Zugriff auf D-Grid Rechenressourcen ist ebenfalls im GRRS gespeichert. Da sowohl VO-Zugehörigkeit als auch VO-Ressourcen-Zugriffsrechte vom GRRS von außerhalb durch autorisierte Dienste abgefragt werden können, können diese Informationen kombiniert und folglich genutzt werden, um festzustellen, welcher Nutzer im Rahmen seiner VO welche Rechenressource nutzen kann, ohne mit der Ressource selbst zu interagieren.

Diese Informationen können bei der Ressourcenvorauswahl (Capability Matching) verwendet werden, um eine erste Auswahl möglicher Ressourcen zu treffen, die im Weiteren anhand „statischer Attribute“ (z. B. Hardware-Parameter wie CPU-Zahl) dem definierten SLA entsprechend weiter gefiltert wird.

Ist ein Nutzer in mehreren VO, so muss auch die zu verwendende VO bei der Auswahl berücksichtigt werden, da es andernfalls zu Rechtsproblemen bei der Jobausführung kommen kann. Ein Nutzer kann Zugang zu einer Ressource durch die Mitgliedschaft in einer VO haben, auf welcher er im Namen einer anderen nicht agieren darf. Erschwerend kommt hinzu, dass der Nutzer nicht mit jedem Grid-Middleware-System wählen kann, im

¹⁶ siehe http://www.d-grid.de/fileadmin/user_upload/documents/Kern-D-Grid/AUP/D-Grid-User-AUP.pdf

¹⁷ Sustainable Grid Infrastructure: <http://sugi.d-grid.de/>

Namen welcher VO bzw. mit dem lokalen Nutzeraccount welcher VO er auf der Ressource agieren will¹⁸.

Für die Auswahl von Ressourcen basierend auf statischen Ressource-Attributen können momentan keine Informationen aus dem GRRS verwendet werden. Die dortigen Ressource-Beschreibungen sind Fließtext und somit nicht maschineninterpretierbar und zudem nicht für alle Ressourcen gleich detailliert ausgeführt. Solche statischen Attribute können zum aktuellen Zeitpunkt nur aus den Monitoring-Diensten des D-Grid bezogen werden, insofern sie dort von den Ressource-Betreibern eingepflegt werden.

Diese Situation kann sich mit der Migration des GRRS auf die Basis der D-GRDL ändern. Die Migration sieht bisher keine Integration von Ressourcen-Beschreibungen in maschinenlesbarer Form vor, ermöglicht dieses jedoch technisch. Die D-GRDL sieht die Beschreibung entsprechender Ressourcen (beispielsweise von Clustern und deren Knoten) vor. Hier könnten entsprechende Zuarbeiten des SLA4D-Grid-Projektes förderlich sein. Von DGI-2 FG-2 *Betrieb* allein können für solch eine Erweiterung keine Zusagen gemacht werden, da dieses über den regulären Projektplan des Fachgebietes hinausgeht.

Zuarbeiten zur Umsetzung der Ressourcen-Beschreibung im D-GRDL-GRRS werden vorerst nicht durch SLA4D-Grid geplant. Im Rahmen der Arbeiten an der Ressourcen-Vorauswahl durch den Orchestrierungsdienst (SLA4D-Grid Arbeitspaket 5) kann dieses nochmals evaluiert werden. Eventuell können alle notwendigen Informationen durch die Erweiterung des Monitoring zugreifbar gemacht werden, so dass eine Ressourcen-Beschreibung im GRRS nicht benötigt wird. Die Zahl der zu adaptierenden Systeme soll möglichst gering bleiben.

3.2 Management von Ressourcen

Für die Ressourcenvorauswahl (Capability Matching) müssen die vorhandenen Grid-Ressourcen bekannt sein. Eine Liste aller aktuellen Ressourcen kann aus dem Ressourcen-Management des D-Grids – dem GRRS – bezogen werden. Da dem zugrunde liegenden VO-Konzept nach nicht alle Nutzer garantiert alle Ressourcen nutzen können, muss für eine erste Vorauswahl auch der Nutzer sowie seine VO-Zugehörigkeit miteinbezogen werden. Wie der GRRS und die von ihm gespeicherten Daten entsprechend eingesetzt werden können, ist in Abschnitt 3.1 beschrieben.

Im Rahmen der Implementierung der SLA-Schicht im D-Grid wird der GRRS voraussichtlich wie beschrieben im Orchestrierungsdienst (SLA4D-Grid Arbeitspaket 5) für die Ressourcenvorauswahl eingebunden werden. Es soll keine direkte Anbindung an den GRRS innerhalb des WS-Agreement-Frameworks für Java (WSAG4J) implementiert werden.

Der GRRS enthält keine Informationen über den aktuellen Zustand einer Ressource, sondern speichert nur, welche Ressourcen im D-Grid angeboten werden und somit prinzipiell verfügbar sein sollten. Für die Ressourcenvorauswahl muss außerdem der aktuelle Zustand jeder Ressource (verfügbar, nicht verfügbar, in Wartung usw.) miteinbezogen werden. Diese Information ist in einigen, aber nicht allen, Monitoring-Systemen verfügbar. Ein zu wählendes Monitoring-System sollte einen Vergleich zwischen Soll- und Ist-Zustand

¹⁸ Die Wahl eines bestimmten Account ist mit einem Teil der Middleware möglich, wenn der Name des lokalen Accounts vorher bekannt ist. Dieses widerspricht jedoch dem Grid-Gedanken. Aktuelle PDP der Middleware ermöglichen nicht die Auswertung von VO-Zugehörigkeitsattributen. Die sauberste Lösung aus Nutzersicht ist momentan die Nutzung verschiedener Zertifikate bei Multi-VO-Mitgliedschaften. Ohne eine Auswahl wird ein Nutzer momentan (seit dem 21.10.2009) auf den alphabetisch ersten Account auf einer Ressource abgebildet.

des Ressourcen-Pools machen, um diese Funktionalität nicht nochmals an anderer Stelle (im SLA-Layer) zu implementieren. Einige der middleware-spezifischen Monitoring-Systeme speichern nur Informationen über aktive Ressourcen und „vergessen“ nicht aktive Ressourcen vollständig, wenn diese keine Informationen mehr an das Monitoring-System melden.

Welche statischen Informationen der GRRS für eine Ressource gespeichert ist für den RDBMS-basierten Fall in Abschnitt 2 beschrieben. Welche Informationen voraussichtlich in der D-GRDL-basierten Variante gespeichert werden, ist dem dort beschriebenen Meilenstein zu entnehmen.

Da nicht alle Ressourcenanbieter ihre Ressourcen mittels Dienstgütegarantien, die über die SLA-Schicht verhandelt werden, anbieten werden, ist das Ressourcen-Management um ein entsprechendes Ressource-Attribut *SLA-Unterstützung* zu erweitern. Dieses ist technisch für beide Varianten des GRRS möglich. Aufgrund der bevorstehenden Migration des GRRS zur D-GRDL-basierten Variante und einer Verwendung erst zu späterem Zeitpunkt des SLA4D-Grid-Projektes (nicht in 2009), ist das Attribut vor allem in die neue Variante des GRRS zu integrieren.

Für eine Rechenressource speichert der GRRS den Namen und den FQDN des entsprechenden Submission-Hosts (wie Globus GRAM oder gLite CE). Der GRRS enthält keine vollständigen Endpoint References (EPRs) für die Submission-Services. Das Einfügen vollständiger EPR wäre eine weitere Erweiterung des GRRS.

3.3 Monitoring

Das Monitoring von Ressource- und Job-Parametern ist eine grundlegende Voraussetzung für die Umsetzung von SLA. Monitoring-Daten werden im gesamten Lebenszyklus eines SLA (vgl. Abbildung 4) benötigt – von der Auswahl einer Ressource die ein SLA erfüllen können soll, über die Ausführung eines Jobs, bis zur Verifikation, ob ein SLA erfüllt worden ist.

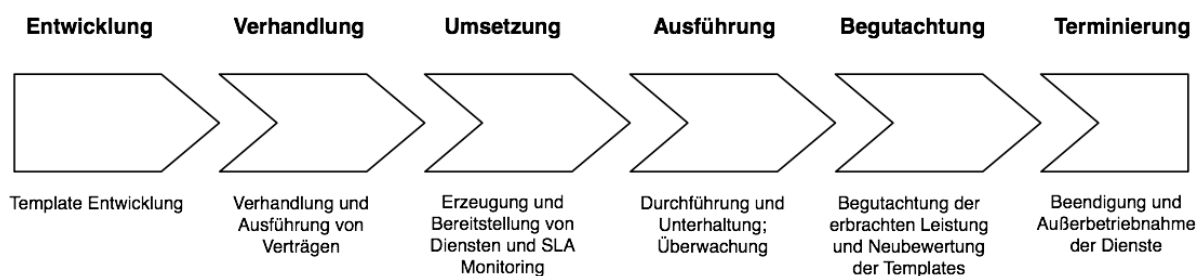


Abbildung 4: Der SLA-Lebenszyklus [12]

- i. Vor und während der Verhandlungsphase eines SLA werden Informationen über existierende Ressourcen und ihren Status sowohl aus der Ressourcenverwaltung (welche Ressourcen, statische Ressourcen-Informationen) als auch aus den Monitoring-Diensten (aktuelle Parameterwerte von Ressourcen, dynamische Informationen) benötigt. Für die Ressourcenauswahl sind statische Informationen meist ausreichend, für eine Ressourcenplanung (Scheduling) sind dynamische Informationen hingegen wertvoll.
- ii. Während der Implementierungsphase eines SLA wird das SLA-spezifische Monitoring „verdrahtet“. Alle zur Durchführung eines Jobs verwendeten Dienste werden gestartet und für das Monitoring vorbereitet.

- iii. Während der Abarbeitung eines Jobs werden permanent Monitoring-Daten über beteiligte Dienste und Ressourcen gewonnen. Sie werden benötigt um nach Abschluss eines Jobs zu beurteilen, ob im SLA definierte Parameter (wie QoS) bei der Ausführung eingehalten worden sind.

Grundsätzlich werden für die Umsetzung von SLAs Informationen zu allen in den SLAs vereinbarten bzw. zu vereinbarenden Parametern (wie Zeiten, Hardwaremerkmale, usw.) benötigt. Dies beinhaltet Informationen über Ressourcen (vor und während der Ausführung eines Jobs/eines Datenzugriffs), Informationen über Rechenjobs (während und nach ihrer Ausführung) oder Zugriffe auf andere Ressourcen. Die Überwachung dieser Informationen ist eine notwendige Anforderung an das Monitoring für die Umsetzung von SLA. Andersherum betrachtet können auch nur SLAs über die Parameter erstellt werden, die überwacht werden können.

Die zu überwachenden Parameter und die dadurch bestimmte Art des Monitorings sind stark abhängig vom jeweiligen Anwendungsfall. Im Rahmen des Projektes SLA4D-Grid geht es um den generellen Anwendungsfall, die Reservierung von Rechenressourcen sowie die kommerziellen Anwendungsfälle der Projektpartner con terra GmbH und Sun Microsystems GmbH. Es muss gewährleistet werden, dass die notwendigen Monitoring-Informationen aus der vorhandenen D-Grid Infrastruktur bzw. den anwendungsspezifischen Diensten geliefert werden. Eine Integration der existierenden Monitoring-Systeme bzw. die Schaffung und Nutzung einheitlicher Monitoring-Schnittstellen zum Andocken anwendungsspezifischer Monitoring-Dienste und Informationsquellen ist auf lange Sicht wünschenswert.

Ist die Überwachung von Parametern, die für die Erstellung von SLA in den genannten Anwendungsfällen notwendig sind, nicht gegeben, so besteht die Möglichkeit, dass SLA4D-Grid technische Lösungen für entsprechendes Monitoring mitgestaltet. Dieses sollte, wenn möglich, als Input für das D-Grid Integrationsprojekt (DGI-2) geschehen, muss eventuell aber auch durch für solche Fälle allokierte Mittel des Projektes mitgetragen werden.

Betrachtet man die Vielzahl der existierenden Monitoring-Lösungen, sollte die Anbindung an drei (oder mehr) verschiedene Informationsquellen auf keinen Fall das Ziel sein. Daher bietet sich die Nutzung der Ergebnisse des D-Mon-Projektes an. Allerdings ist es fraglich, ob der Overhead der von D-Mon gebotenen OGSA-DAI-Schnittstelle gerechtfertigt ist. Eine einfache Webservice-Schnittstelle zu diesem Dienst wäre für SLA4D-Grid höchstwahrscheinlich sinnvoller.

Generell sollten folgende Anforderungen an Monitoring-Daten erfüllt werden:

- Daten sind unabhängig von der eingesetzten Middleware
- Daten sind für jedes SLA-Template verfügbar
- Daten sind angemessen aktuell
- alle Daten liegen in einem Format vor
- Zugriff auf Daten via eine Webservice-Schnittstelle möglich

Im Weiteren sind folgende Aspekte von SLA4D-Grid zu verfolgen. Die vorhandenen Anwendungsfälle müssen analysiert werden, um festzulegen, welche Monitoring-Daten für die korrespondierenden SLA benötigt werden. Gleichzeitig muss das Angebot vorhandener Monitoring-Daten betrachtet werden, um aus beiden einen Soll-Ist-Vergleich zu erarbeiten, um zu sehen, welche Daten zusätzlich zu denen des Status Quo überwacht werden müssen. Hieraus kann ein Satz notwendiger Monitoring-Daten definiert werden, der im Folgenden D-Grid-weit umgesetzt werden muss. Für die Umsetzung von SLAs

müssen notwendige Monitoring-Informationen von allen beteiligten Grid-Ressourcen einheitlich und vollständig vorhanden sein. Der Soll-Ist-Vergleich ist in Punkto Vorhandensein, aber auch in Punkto Aktualität der Informationen zu erarbeiten. Im Rahmen der Betrachtung müssen auch die vorhandenen Ansätze des Job-Monitoring betrachtet werden, da diese für 3.3.iii. von hoher Bedeutung sind. Hier ist auch die rechtliche Frage des Zugriffs auf Job-Monitoring-Daten zu betrachten.

Die momentane D-Grid Infrastruktur lässt ein Monitoring über den ganzen Lebenszyklus eines SLAs nicht zu. Es ist an dieser Stelle anzumerken, dass es nicht möglich sein wird, dies innerhalb der Projektlaufzeit generisch zu lösen. Dafür ist die momentane D-Grid Infrastruktur nicht vorbereitet und potentielle Anwendungsfälle zu divergent. Allerdings wird SLA4D-Grid exemplarisch Lösungen schaffen und seine Anforderungen an das D-Grid weitergeben.

Im Bereich des Monitoring werden für die Umsetzung einer SLA-Schicht im D-Grid die meisten Lücken und schwerwiegendsten Probleme gesehen. Es existiert eine Vielzahl von Monitoring-Lösungen, welche aber nicht alle notwendigen Informationen bereitstellen. Zu den notwendigen Informationen wird eine einheitliche Schnittstelle benötigt, da im Rahmen der SLA-Implementierung nicht alle verschiedenen Monitoring-Systeme adaptiert werden können.

3.4 Accounting

Das Accounting stellt neben dem Monitoring die Grundlage für die Verifikation von SLAs dar. Nach Beendigung eines Jobs wird anhand der im SLA definierten Parameter und der gemessenen Parameter entschieden, ob ein SLA erfüllt worden ist.

Um die Accounting-Informationen nutzen zu können, muss ein Accounting-Datensatz nicht nur einem Verursacher (lokaler Nutzer/VO), sondern auch einem bestimmten Job zugeordnet werden können. Dieses ist momentan nicht möglich, da es keine Verbindung zwischen Middleware-spezifischer Job-ID und der Job-ID des lokalen Batchsystems gibt. Letztere wird momentan im Accounting-System gespeichert. Erstere ist dem Nutzer bei der Submission eines Jobs bekannt. Um diese Informationen zu verknüpfen bietet sich eine entsprechende Datenbank bzw. ein entsprechender Dienst auf der Grid-Ressourcen an, die die Schnittstelle zwischen Grid-Middleware und lokalem Batchsystem darstellt, da an dieser Stelle beide IDs bekannt sind. Solch ein Ansatz wurde schon in DGI-2 FG-5 vorgeschlagen [4], ist jedoch bisher nicht umgesetzt worden. Sollen die Accounting-Datensätze für die SLA-Verifikation verwendet werden, so ist solch eine ID-Zuordnung von SLA4D-Grid zwingend erforderlich und muss entsprechend kommuniziert werden.

Um auf die Accounting-Daten von DGAS bzw. DGAS-EE zugreifen zu können wird eine entsprechende Webservice-Schnittstelle benötigt.

Desweiteren muss der rechtliche Rahmen für den Zugriff auf Accounting-Daten für die SLA-Verifikation geklärt werden. Bisher können nur Verursacher eines Jobs und Accounting-Datensatzes selbst bzw. dessen VO Accounting-Informationen einsehen. Für die Verifikation muss ein entsprechender Dienst als dritte Person auf Accounting-Informationen zugreifen, die zudem eindeutig mit einem Job in Verbindung gebracht werden können.

3.5 Referenzinstallation und Betriebskonzept

Software, die im Rahmen von SLA4D-Grid entwickelt wird und später im D-Grid eingesetzt werden soll, muss sowohl in der Referenzinstallation als auch im Betriebskonzept des D-Grid eingepflegt werden. Die Referenzinstallation dient dabei der Dokumentation der

Installation der Systeme der Ressourcen-Anbieter. Das Betriebskonzept muss um neue zentrale Dienste erweitert werden.

Um SLA im D-Grid rechtlich bindend umsetzen zu können müssen Rahmenverträge zwischen den verschiedenen Beteiligten existieren. Solche Rahmenverträge könnten als Erweiterung der D-Grid Acceptable Use Policies (AUP) umgesetzt werden.

4 D-Grid-Middleware

4.1 Eingesetzte Middleware

Im D-Grid werden drei verschiedene Middleware-Systeme eingesetzt: das Globus Toolkit, gLite und UNICORE. Um Nutzer und aufbauende Dienste fortlaufend zu unterstützen und einen Migrationszeitraum zu bieten, werden zudem verschiedenen Versionen dieser Middleware Pakete parallel betrieben.

Bis dato wird im D-Grid die Version 4.0 des Globus Toolkit eingesetzt. Es ist jedoch ein Wechsel zur Version 4.2 von Seiten des DGI geplant. Die neue Version wird mit dem Release 2010-1 in die Referenzinstallation aufgenommen. Details zum parallelen Betrieb beider Versionen, zum End-of-Life der Version 4.0 und der Installation eines zentralen MDS 4.2 wurden noch nicht geklärt bzw. entschieden.

UNICORE 5 ist noch die hauptsächlich im D-Grid installierte Version der UNICORE-Middleware. Nur wenige Sites haben bisher ihre Ressourcen um ein Frontend für UNICORE 6.2 erweitert. Am zentralen Gateway der Middleware am FZJ ist momentan keine wesentliche Nutzung von UNICORE 5 festzustellen. Aus diesem Grund soll UNICORE 5 aus der D-Grid Referenzinstallation entfernt und nur noch auf UNICORE Version ≥ 6.2 gesetzt werden. Der Gateway wird jedoch noch weiterhin getrieben. Bis zu welchem Termin die Sites auf UNICORE 6 migrieren müssen und das zentrale Gateway von UNICORE 5 abgeschaltet werden soll, muss noch geklärt und beschlossen werden¹⁹.

Momentan wird im D-Grid gLite in den Versionen 3.0 und 3.1 eingesetzt, die Referenzinstallation beschreibt die Installation von gLite 3.1 auf Basis von Scientific Linux 4 und der x86-Architektur. Da inzwischen gLite 3.2 x64-Pakete für Scientific Linux 5 existieren, ist eine Aktualisierung der Cluster-Workernodes auf diese in einem zukünftigen Release der Referenzinstallation geplant. Bisher wird das LCG-CE genutzt, am FZK existiert eine Testinstallation des neuen Cream CE.

4.2 Fokus von SLA4D-Grid

Im Rahmen des SLA4D-Grid Kick-off-Meetings wurde in Projektmonat 1 festgelegt, dass Komponenten mit dem Fokus auf die neuen Versionen der Middleware-Systeme entwickelt werden sollen. In SLA4D-Grid sollten daher die Komponenten der SLA-Schicht für UNICORE 6, das Globus Toolkit 4.2 sowie für gLite entwickelt werden. Die Wahl der Version 4.2 des Globus Toolkit muss aufgrund der neusten Entwicklungen im Globus-Projekt und im D-Grid nochmals eruiert und validiert werden (siehe Abschnitt 5).

4.3 SLA-Unterstützung der Middleware

Welche Unterstützung/Komponenten für die Umsetzung von SLAs von der im D-Grid eingesetzten Middleware bisher geboten wird und welche Projekte und Software sich mit diesem Thema schon befasst haben, wird im Bericht des SLA4D-Grid Arbeitspaketes 4 „State of the Art“ betrachtet.

¹⁹ siehe Protokoll der 1. D-Grid TAB-Videokonferenz am 02.09.2009 unter <http://www.d-grid.de/tab>

5 Offene Punkte

Es ist zu beachten, ob sich durch die Nutzung der zentralen Dienste des D-Grids sicherheitsrelevante Fragen ergeben. Vor allem im Bereich der Zugriffsrechte auf Daten der zentralen Dienste – wie GRRS, Monitoring oder Accounting – sowie die Verwendung dieser Daten können weitere Hindernisse entstehen. Neben rein technischen Gründen sind vor allem Datenschutzbestimmungen als potentieller Fallstrick zu sehen. Dies gilt insbesondere im Fall personenbezogener Accounting-Daten.

Sollen durch SLA4D-Grid Anforderungen an Informationsdienste – wie den GRRS – beziehungsweise an das Monitoring von Ressourcen oder Jobs eingebracht werden, stellt sich die Frage, ob es dafür einen (optimalen) Prozess gibt. Wer sind die besten Ansprechpartner beziehungsweise mit welcher Instanz sollten solche Zuarbeiten gemeinsam initiiert werden?

Einige Schnittstellen mit der existierenden D-Grid Infrastruktur müssen noch genauer betrachtet werden. Während der GRRS im D-Grid zur Verwendung der D-GRDL migriert werden soll, soll in SLA4D-Grid das GLUE-Schema für die Beschreibung entsprechender Ressourcen eingesetzt werden.

Das Billing von Ressourcen-Nutzung ist für das von SLA4D-Grid verfolgte Ziel der kommerziellen Nutzung von Interesse. Es wurde jedoch bisher nicht im D-Grid implementiert.

Wie in Projektmonat 5 bekannt wurde, gibt es im Rahmen der Weiterentwicklung des Globus Toolkits²⁰ weitreichende Veränderungen²¹ an diesem. Mit der Entwicklung des GT 4.2 hatten die Globus-Entwickler die WSRF-Implementierung an den gültigen Standard angepasst, so dass GT 4.0 und GT 4.2 bezüglich des WSRF inkompatibel sind und nicht miteinander genutzt werden können. Die nun vor der Veröffentlichung stehende Version 5 des Toolkits lässt sämtliche Web Services entfallen und stellt eine Weiterentwicklung der Komponenten des Non-WS GT 2 dar. Zu einem späteren Zeitpunkt soll eine neue WS-Umgebung mit Namen CRUX²² entstehen. Diese ist aber erst in einem Jahr zu erwarten. Diese weitreichenden Veränderungen werden sich mit hoher Wahrscheinlichkeit auch im D-Grid niederschlagen. Eine Vielzahl der Projekte und Communities im D-Grid baut bisher auf dem GT 4.0 auf und wird keine doppelte Migration (zu GT 4.2 und zu GT 5 mit oder ohne CRUX) durchführen wollen. Aus diesem Grund ist der ursprünglich durch das DGI geplante Umstieg auf GT 4.2 fragwürdig. Zu dieser Fragestellung laufen momentan Umfragen unter den Communities und den Ressourcen-Anbietern des D-Grids. Die Ergebnisse der Umfragen und ein eindeutiges Statement des DGI Globus-Supports am LRZ sollten abgewartet werden, um die Verwendung des Globus Toolkit 4.2 im Rahmen von SLA4D-Grid zu eruieren und zu validieren. Eine Alternative wäre die fortgeführte Nutzung des GT 4.0 um später auf GT 5 mit CRUX zu migrieren.

²⁰ Globus-User-Mailingliste unter <http://lists.globus.org/pipermail/gt-user/2009-October/008578.html>

²¹ Globus-User-Mailingliste unter <http://lists.globus.org/pipermail/gt-user/2009-October/008553.html>

²² Globus-User-Mailingliste unter <http://lists.globus.org/pipermail/gt-user/2009-October/008574.html>

Abkürzungsverzeichnis und Glossar

AA	Attribute Authority – Attribut-Autorität
AUP	Acceptable Use Policy
BDII	Berkeley Database Information Index
CA	Certificate Authority – Zertifizierungsstelle
CIS	Common Information System
CLI	Command Line Interface
DFN	Deutsches Forschungsnetz (Verein)
DFN-AAI	Authentifizierungs-Autorisierungs-Infrastruktur/ Shibboleth-Föderation des DFN-Vereins
DFN-PKI	Public-Key-Infrastruktur des DFN-Verein
DFN-PKI Grid-CA	Zertifizierungsstelle der DFN-PKI für langlebige Grid-Zertifikate
DFN-PKI SLCS	Zertifizierungsstelle der DFN-PKI für kurzlebige Grid- Zertifikate
DGAS	Distributed Grid Accounting System
DGAS-EE	Distributed Grid Accounting System - Enterprise Edition
DGI	D-Grid Integrationsprojekt
DGI-2	D-Grid Integrationsprojekt 2 (Folgeprojekt zu DGI)
D-GRDL	D-Grid Grid Resource Description Language
DMZ	Demilitarisierte Zone in einem Computer-Netzwerk, nicht Teil des internen Netzes sowie Internets abgegrenzt durch Firewalls
EUGridPMA	European Policy Management Authority for Grid Authentication
eXist-DB	XML-basiertes Datenbank-Management-System
FG	Fachgebiet
FZJ	Forschungszentrum Jülich
FZK	Forschungszentrum Karlsruhe
GLUE	Grid Laboratory for a Uniform Environment
GridKa	Grid Computing Centre Karlsruhe
GridKa CA	Zertifizierungsstelle für Grid-X.509-Zertifikate der GridKa
GRRS	Grid Resource Registry Service
GSI	Grid Security Infrastructure
GT	Globus Toolkit
IGTF	International Grid Trust Foundation
INFN	Istituto Nazionale di Fisica Nucleare (Italien)
ITU-T	International Telecommunication Union - telecommunications standardization sector
ITWM	Fraunhofer ITWM – Institut für Techno- und Wirtschaftsmathematik
LCG	Large Hadron Collider Compute Grid
LDAP	Lightweight Directory Access Protocol
LRZ	Leibniz-Rechenzentrum (München)
LSF	Load Sharing Facility
MDS	Monitoring and Discovery System (Globus Toolkit)
OGF	Open Grid Forum

OGSA	Open Grid Services Architecture
OGSA-DAI	Open Grid Services Architecture - Data Access and Integration
PBS	Portable Batch System
PKI	Public Key Infrastructure – Public-Key-Infrastruktur
RRZN	Regionales Rechenzentrum für Niedersachsen (Hannover)
SAML	Security Assertion Markup Language
SAML2	Security Assertion Markup Language Version 2
SGE	Sun Grid Engine
SLA	Service Level Agreement
SLCS	Short Lived Credential Service
SQL	Structured Query Language
SuGi	Sustainable Grid Infrastructure
UNICORE	Uniform Interface to Computing Resources
UR	Usage Record
UR-WG	Usage Record - Working Group (OGF)
UUDB	UNICORE User Database
VO	Virtual Organization – Virtuelle Organisation
VOMS	Virtual Organization Membership Service
VOMRS	Virtual Organization Membership Registration Service
WSAG4J	WS-Agreement-Framework für Java
WSRF	Web Service Resource Framework
X.509	ITU-T-Standard für Public-Key-Infrastrukturen
XML	Extensible Markup Language
XUUDB	UUDB ab UNICORE Version 6

6 Referenzen

- [1] Büchner, Otto, et al. *Betriebskonzept für die D-Grid Infrastruktur*. Version 1.3.22 Draft (unveröffentlicht), Forschungszentrum Jülich: D-Grid Integrationsprojekt (DGI-2) Fachgebiet 2: Betrieb, 30.10.2009.
- [2] Baur, Timo, Anne Milbert, Mathilde Romberg, Gevorg Poghosyan, and Tobias Lindinger. *D-Mon: Softwaretechnische Analyse und Design*. AP 1.3, D-Grid Gap-Projekt D-Mon, 2008.
- [3] Baur, Timo, Tibor Kalman, Tobias Lindinger, Anne Milbert, Gevorg Pogoshyan, and Mathilde Romberg. *Middleware-übergreifendes Monitoring: Evaluierung und Auswahl von Komponenten*. AP 1.2, D-Grid Gap-Projekt D-Mon, 2008.
- [4] Brenner, Michael, and Jan Wiebelitz. *Accounting von vermittelten Grid-Jobs*. Version 1.0, D-Grid Integrationsprojekt DGI-2 Fachgebiet 5.2, 2008.
- [5] "D-Grid Integrationsprojekt (DGI-2) – Anhänge der Fachgebiete." 2007.
- [6] Dohmen, Christa, et al. *Betriebskonzept für die D-Grid Infrastruktur*. Online. Version 1.1c Draft, Forschungszentrum Jülich: D-Grid Integrationsprojekt (DGI-2) Fachgebiet 2: Betrieb, 2007.
- [7] Dohmen, Christa, Thomas Fieseler, and Wolfgang Gürich. *Tabellenstruktur des GRRS - Grid Resource Registry Service*. Forschungszentrum Jülich: D-Grid Integrationsprojekt (DGI-2) Fachgebiet 2: Betrieb, 2008.
- [8] Fachgebiet Referenzinstallation, D-Grid Integrationsprojekt. *DGI reference installation*. <http://dgiref.d-grid.de/> (letzter Zugriff am 07.09.2009).
- [9] Hilbrich, Marcus. *Erster Prototyp Job-Monitoring-Infrastruktur mit Dokumentation*. Meilenstein D5-5.1-3, D-Grid Integrationsprojekt DGI-2 Fachgebiet 5.5.
- [10] Lorenz, D., et al. "Job monitoring and steering in D-Grid's High Energy Physics Community Grid." *Future Generation Computer Systems* (Elsevier), no. 25 (2009): 308-314.
- [11] Mach, R., R. Lepro-Metz, S. Jackson, and L. McGinnis. *Usage Record – Format Recommendation*. GFD-98, OGF UR-WG, 2007.
- [12] The TeleManagement Forum. *SLA Management Handbook, Volume 2, Concepts and Principles*. Morristown, New Jersey: The Management Forum, 2005.
- [13] Wiebelitz, Jan, and Michael Brenner. *Konzept für das Accounting im D-Grid*. Version 1.0, D-Grid Integrationsprojekt DGI-2 Fachgebiet 5.2, 2008.
- [14] Wiebelitz, Jan, Ralph Müller-Pfefferkorn, and Michael Brenner. *Erweiterung der D-Grid Acceptable Use Policy um Aussagen zum Datenschutz*. D-Grid Integrationsprojekt (DGI-2) Fachgebiete 5.2 und 5.5, 2008.
- [15] Wolf, Armin, Thomas Fieseler, and Christa Dohmen. *Festlegung der Objektbezeichnung für die D-GRDL*. Meilenstein M2-1, D-Grid Integrationsprojekt (DGI-2) Fachgebiet 2.3, 2008.