

Benutzergesteuerter Datenschutz in Grids*

Wolfgang Hommel
Munich Network Management Team
Leibniz-Rechenzentrum

Michael Schiffers
Munich Network Management Team
Ludwig-Maximilians-Universität München

Abstract: Grid Computing erlaubt die aus Benutzersicht weitgehend transparente, verteilte Nutzung von Ressourcen und unterscheidet sich durch Konzepte wie Virtuelle Organisationen zum Teil stark von herkömmlichen verteilten Systemen. Die sich daraus ergebenden Grid-spezifischen Anforderungen an den Datenschutz und die Datensicherheit wurden von Lösungsansätzen, die beispielsweise im Bereich des Federated Identity Management bereits erfolgreich eingesetzt werden, bislang nur unzureichend erfüllt.

In diesem Beitrag werden zuerst die aus Datenschutzperspektive wichtigsten Unterschiede zwischen Grids und herkömmlichen föderierten Umgebungen dargelegt und sich daraus ergebende neue Anforderungen diskutiert. Anschließend wird gezeigt, wie bestehende Privacy Management Ansätze erweitert werden können, um Grid-spezifische Datenschutzerfordernungen zu erfüllen. Der Schwerpunkt liegt dabei auf den Steuerungs- und Kontrollmechanismen, die den Grid-Benutzern eingeräumt werden. Aufbauend auf einer policybasierten Privacy Management Infrastruktur, die wir in einer früheren Arbeit präsentiert haben, werden die Anwendung der vorgestellten Methodik am Beispiel der standardisierten Policysprache XACML konkret demonstriert und die Integration der notwendigen Komponenten in Grid-Architekturen diskutiert.

1 Einführung: Datenschutz in Grid-Umgebungen

Grid Computing ermöglicht die transparente, organisationsübergreifende Nutzung von Ressourcen (meist in Form von Rechen- und Speicherkapazitäten) auf der Basis entsprechender Middleware-Technologien. An aktuellen Grid-Projekten wie dem deutschen D-Grid [NKH07] sind als Ressourcenanbieter primär Höchstleistungsrechenzentren beteiligt; sie stellen Systeme für wissenschaftliche Anwendungen zur Verfügung, um sehr umfangreiche Berechnungen mit großen Datenvolumina effizient durchführen zu können.

Abbildung 1 zeigt den prinzipiellen Ablauf aus Benutzerperspektive: Eine der am Grid beteiligten Organisationen fungiert als Heimatorganisation des Benutzers; dieser erhält dort die Möglichkeit, so genannte Grid-Jobs in Auftrag zu geben, wofür entweder kommandozeilen-orientierte Werkzeuge oder Grid-Webportale zur Verfügung stehen. Komponenten der als Grid-Middleware bezeichneten Softwareinfrastruktur regeln und überwachen dann den Zugang zu Grid-Ressourcen und -Diensten.

Die beliebig verteilte Ausführung von Grid-Jobs basiert auf einer föderierten Benutzerverwaltung, deren technische Realisierung je nach im Grid-Verbund eingesetzter Middleware

*Teile der Arbeit wurden im Rahmen der D-Grid Initiative unter FKZ 01AK810B unterstützt.

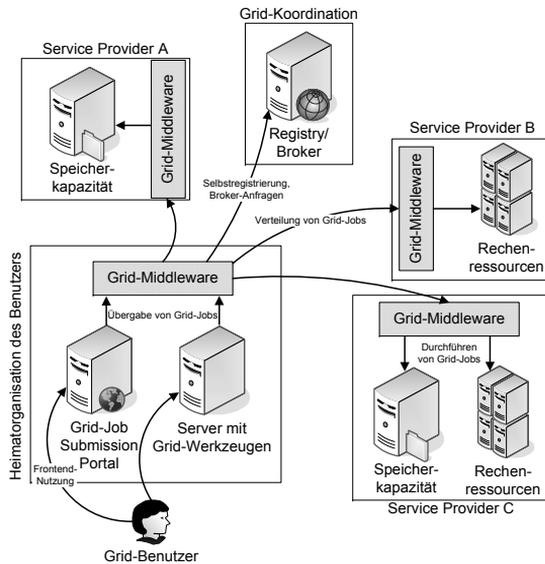


Abbildung 1: Prinzipielle Ausführung von Grid-Jobs über mehrere Service Provider

unterschiedlich sein kann. Da Grid-Dienste primär nicht anonym oder pseudonym nutzbar sind, spielt der Schutz personenbezogener Daten eine sehr wichtige Rolle. Aufgrund der historischen und zum Teil sehr kernfunktional orientierten Entwicklung der Middleware-Produkte setzt sich das Bewusstsein für Datenschutz und Datensicherheit in Grids jedoch erst nach und nach durch; bislang stehen keine technischen Lösungen zur Verfügung, die praktisch eingesetzt werden können.

Aus technischer Perspektive müssen beim Datenschutz in Grids – wie bei anderen organisationsübergreifend genutzten Diensten – zwei Phasen unterschieden werden: In einem ersten Schritt muss entschieden werden, welche personenbezogenen Daten einer am Grid beteiligten Organisation überhaupt zugänglich gemacht werden dürfen; in der sich anschließenden Betriebsphase muss sichergestellt werden, dass diese Daten nur zweckgebunden verwendet und den Vereinbarungen gemäß wieder gelöscht werden. In beiden Phasen müssen neben gesetzlichen Auflagen auch Grid- sowie organisationsweite Vereinbarungen und insbesondere auch benutzerspezifische Vorgaben berücksichtigt werden.

In den Bereichen des Enterprise Identity & Access Managements sowie des Federated Identity Managements wurden sowohl von der Forschung als auch von der Industrie Lösungen für das Privacy Management erarbeitet, die sich jedoch nur bedingt auf Grids anwenden lassen. Zum einen fokussieren sie auf den Schutz der personenbezogenen Daten des Benutzers selbst, vernachlässigen jedoch den Schutz Grid-anwendungsspezifischer Daten – beispielsweise umfangreicher Patientendatensätze, die als Eingabe von Grid-Anwendungen dienen. Zum anderen konnten Grid-Konzepte wie die Sichten Virtueller Organisationen (VOs) auf diese Daten technisch nur rudimentär abgebildet werden [Sch07].

Im nächsten Abschnitt fassen wir relevante aktuelle Lösungsansätze knapp zusammen, da die hier vorgestellte Arbeit auf ihnen aufbaut. In Abschnitt 3 diskutieren wir die Datenschutzspezifika von Grids und stellen die wesentlichen Ergebnisse einer Anforderungsanalyse für Grid-spezifisches Privacy Management vor. In Abschnitt 4 zeigen wir, wie existierende Ansätze aus dem Bereich des policybasierten Privacy Managements erweitert werden können, um die Grid-spezifischen Anforderungen benutzerzentriert zu erfüllen. Die Anwendung dieser Methodik wird am Beispiel der weit verbreiteten Policysprache XACML in Abschnitt 5 demonstriert. Die Integration einer solchen Lösung in bestehenden Grid-Architekturen ist Gegenstand von Abschnitt 6. Eine Diskussion der bisherigen Ergebnisse mit Ausblick auf weitere Forschungsschwerpunkte schließt diesen Beitrag ab.

2 Bisherige Lösungsansätze

Die in Grid-Umgebungen relevante Weitergabe personenbezogener Daten von einer als Identity Provider bezeichneten Organisation an einen Service Provider (SP) wurde im Umfeld des Federated Identity Management (FIM), das auf Technologien wie SAML, den Liberty Alliance Spezifikationen oder WS-Federation beruht, bereits eingehend untersucht.

Im FIM-Umfeld haben sich für die benutzerzentrierte Umsetzung von Datenschutzaspekten im Wesentlichen zwei orthogonale Lösungswege herausgebildet: Im einfachsten Fall wird der Benutzer jedesmal vor der Weitergabe seiner Daten interaktiv um Zustimmung gebeten [A⁺04]. Diese Variante leidet unter schlechter Usability, wenn eine Vielzahl von FIM-Transaktionen durchgeführt wird, da Benutzer der ständigen Zustimmung zur Datenfreigabe schnell überdrüssig werden und dann dazu tendieren, den Datenfreigaben ohne genaues Lesen immer zuzustimmen; für Grid-Jobs, die potentiell auf eine Vielzahl von SPs verteilt werden, scheidet dieser Ansatz deshalb in der Praxis aus.

Der zweite Lösungsweg basiert auf dem policybasierten Management; Policies enthalten maschinell verarbeitbare Regeln, die angeben, welche Benutzerdaten an welche SPs herausgegeben werden dürfen. Die Übermittlung von Daten kann einerseits an Bedingungen gebunden sein, beispielsweise dass der SP eine bestimmte Zweckbindung der Daten garantiert. Andererseits können Auflagen, die häufig als (Langzeit-)Obligationen bezeichnet werden, definiert werden, die vom Dienstleister einzuhalten sind, beispielsweise das Löschen personenbezogener Daten nach erfolgter Dienstabrechnung.

Policybasierte Ansätze unterscheiden sich primär bezüglich der gewählten Policysprache und deren mathematischen Eigenschaften wie der Abgeschlossenheit unter Schnittmengenbildung; ein Überblick kann [Hom07] entnommen werden. Die architekturelle Realisierung dieser Lösungen umfasst typischerweise Variationen der folgenden Komponenten:

- Ein Policy Decision Point (PDP) wertet alle für eine Anfrage relevanten Policies aus und entscheidet darüber, ob die Daten freigegeben werden dürfen oder nicht. Die Policies werden in einem zentralen Policy Repository gespeichert und über Policy Administration Points verwaltet.
- Alle Zugriffe auf die Daten werden über Policy Enforcement Points (PEPs) kanali-

siert; ein PEP kommuniziert die Anfrage an den PDP und setzt dessen Zugriffsentcheidung um. Es muss technisch sichergestellt werden, dass an den PEPs vorbei kein direkter Zugriff auf die Daten möglich ist.

- Ein Obligation Monitor überwacht die Einhaltung der vereinbarten Auflagen; er stößt bei Bedarf z. B. das lokale Löschen personenbezogener Daten an.

Eine weiterhin aktuelle Forschungsfragestellung betrifft die benutzergesteuerte Überprüfung, ob die vereinbarten Obligationen erfüllt worden sind. Bisherige Ansätze sehen den Einsatz von zertifizierter Software auf Trusted Computing Platforms vor [Mon04], um gefälschte Auskünfte zu unterbinden; sie haben sich aufgrund des damit verbundenen Aufwands in der Praxis jedoch noch nicht etabliert.

3 Grid-Charakteristika und resultierende Anforderungen

Die Nutzung von Grid-Infrastrukturen, die häufig mit der Metapher der Rechenleistung „aus der Steckdose“ verbunden wird, unterscheidet sich in mehrfacher Hinsicht von der Inanspruchnahme herkömmlicher Dienste einer Einrichtung über das Internet. Die folgenden beiden Charakteristika liegen im Fokus unserer Betrachtungen:

- Grids sehen wie andere High Performance Computing-Dienste im allgemeinen vor, dass ein Benutzer eigenen Programmcode auf den Servern der beteiligten Organisationen ausführen kann. Er ist somit einerseits auch selbst dafür verantwortlich, dass die von seinen Programmen durchgeführte Datenverarbeitung die vorhandenen Zweckbindungen der genutzten Eingabedaten nicht verletzt. Andererseits muss berücksichtigt werden, dass die Ein- und Ausgabedaten auf Maschinen von Dritten verarbeitet werden, die zum Zeitpunkt der Datenerfassung aufgrund der von der Grid-Middleware zur Laufzeit realisierten Ortstransparenz noch unbekannt sind. Da sich die genutzten Rechen- und Speicherkapazitäten unter vollständiger Kontrolle dieser Dritten befinden, liegt nahe, dass auch herkömmliche Verfahren wie die Verschlüsselung der Ein- und Ausgabedaten nur eine begrenzte Wirksamkeit bezüglich der potentiellen Einsichtnahme und Verwendung durch Dritte haben. Somit sind explizite Vereinbarungen zum Schutz der Daten unerlässlich, die über den Schutz personenbezogener Daten des Grid-Benutzers selbst deutlich hinausgehen.
- Die Grids zugrunde liegende virtuelle Organisationsform muss berücksichtigt werden. Beispielsweise können in VOs Datenschutzaspekte im Rahmen von Acceptable Use Policies (AUP) a priori vertraglich für alle Teilnehmer geregelt werden. Sofern solche Vereinbarungen existieren, müssen sie auf die eingesetzten technischen Systeme abgebildet werden können. Andernfalls muss bedacht werden, dass die von der Middleware geschaffene Ortstransparenz im Widerspruch zu den Entscheidungs- und Kontrollmöglichkeiten, die Benutzern bei der Weitergabe ihrer Daten eingeräumt werden müssen, stehen kann.

Den Anwendern müssen deshalb Möglichkeiten eingeräumt werden, die in bisherigen Grid-Konzepten und -Realisierungen aufgrund der fehlenden technischen Unterstützung

nur unzureichend berücksichtigt werden könnten. Dazu gehören insbesondere:

- Jeder Benutzer muss bestimmen können, welchen am Grid beteiligten Organisationen die ihn betreffenden personenbezogenen Daten in welchem Umfang und für welchen Zeitraum zur Verfügung gestellt werden. Als Entscheidungsbasis muss einerseits eine semantisch unmissverständliche und verbindliche Zweckbindung definiert werden (z. B. die Kontaktaufnahme bei technischen Problemen oder zum Zwecke der Rechnungsstellung). Andererseits müssen Auflagen an die Datenhaltung vereinbart und umgesetzt werden, so dass z. B. nach erfolgter Abrechnung und definierter Karenzzeit nicht mehr benötigte Daten vom Dienstleister gelöscht werden. Letzteres kann an den Lebenszyklus von VOs gekoppelt werden [Sch07].
- Analog dazu müssen Kontrollmechanismen für den vom Benutzer gelieferten Programmcode, die Eingabe- sowie die erzeugten Ausgabedaten angeboten werden, die über die Mechanismen heutiger Middleware-Implementierungen hinausgehen. Über die Zweckbindung kann beispielsweise vereinbart werden, dass eine Analyse des Programmcodes nur zur optimierenden Anpassung an die bereitgestellte Rechnerarchitektur oder zur Analyse auf enthaltene Schadfunktionen (z. B. Viren, trojanische Pferde) durchgeführt wird.
- Die Einhaltung der Zweckbindung und Erfüllung der vereinbarten Auflagen muss in beiden Fällen vom Benutzer überwacht werden können.

Im nächsten Abschnitt wird eine Erweiterung bestehender Privacy Management Ansätze um diese Aspekte erläutert.

4 Erweiterung existierender Privacy Management Ansätze

Wie in Abschnitt 2 dargelegt wurde, werden Datenfreigaberegeln in den bisherigen Anwendungsfällen typischerweise einzeln pro SP definiert. In Grids bietet jedoch häufig eine größere Menge von SPs zueinander äquivalente Dienste wie CPU- oder Speicherkapazität an, so dass technische Aspekte wie die aktuelle Auslastung der Ressourcen oder ökonomische Aspekte wie das beste Preis-/Leistungsverhältnis darüber entscheiden, bei welchen SPs ein Grid-Job am besten ausgeführt werden sollte. Privacy Management Ansätze müssen für den Einsatz in Grids deshalb wie folgt ausgebaut werden:

1. Die in Datenfreigaberegeln verwendeten SP Identifikatoren müssen um die Adressierbarkeit von Gruppen von Organisationen bis hin zu Virtuellen Organisationen als Ganzes erweitert werden. Neben der expliziten Benennung muss auch die Verwendung beschreibender Merkmale möglich sein, beispielsweise eine Gruppierung aller Organisationen, die vom Benutzer vorgegebene Datenschutzauflagen erfüllen.
2. Die Formulierung von Datenfreigaberegeln durch den Benutzer muss über die personenbezogenen Daten des Benutzers hinaus auf Programmcode sowie Ein- und Ausgabedaten ausgedehnt werden.
3. Die benutzergesteuerte Überwachung der Einhaltung der Datenschutzvereinbarungen muss ebenfalls um die Bestandteile von Grid-Jobs erweitert werden.

Die ersten beiden Punkte betreffen sowohl den Inhalt als auch die Organisation von Datenfreigaberegeln; der letzte Punkt impliziert eine technische Erweiterung der bereits vorhandenen Privacy Monitoring Infrastruktur.

Die Unterscheidung zwischen den personenbezogenen Daten des Benutzers selbst und dem Schutzbedarf der im Rahmen von Grid-Jobs verwendeten Programme und Daten bedingt die folgenden syntaktischen Erweiterungen von Datenfreigaberegeln:

- Der Namensraum zur Spezifikation der durch eine Freigaberegeln abgedeckten Daten muss um Grid-Job-Komponenten wie *Programmcode*, *Eingabedaten* und *Ausgabedaten* erweitert werden. Damit verbunden ist eine Erweiterung des VO-weit eingesetzten Vokabulars für die Spezifikation von Verwendungszwecken, beispielsweise um den Anwendungszweck „Optimierung des Programmcodes für providerspezifische Ressourcen“ formal festhalten zu können.
- Es bietet sich an, die Datenfreigaberegeln für personenbezogene Daten des Benutzers getrennt von denen für Grid-Jobs zu verwalten, da letztere an den jeweiligen Einzelfall angepasst werden müssen, während die Datenschutzpräferenzen des Benutzers typischerweise langlebiger und unabhängig von einzelnen Grid-Jobs sind. Bei der Realisierung ist somit zu beachten, dass einerseits die eingesetzten Policy Decision Points mit der Verteilung der für eine Anfrage relevanten Freigaberegeln auf mehrere Policies geeignet sind und dass andererseits für die Benutzer geeignete Frontends angeboten werden, die eine getrennte Konfiguration beider Arten von Freigaberegeln zulassen und unter Usability-Aspekten optimal unterstützen.

Zur nachweisbaren Umsetzung der mit den Benutzern getroffenen Datenschutzvereinbarungen müssen bei jeder am Grid beteiligten Organisation die in Abschnitt 2 vorgestellten Privacy Management Systeme und Auditierungsmechanismen implementiert werden. Aufgrund der Schutzanforderungen von Grid-Jobs sind diese wie folgt zu erweitern:

- Der Zugriff auf die vom Benutzer gelieferten Grid-Job-Komponenten *Programmcode* und *Eingabedaten* muss der Kontrolle durch das Privacy Management System unterliegen. Dies bedeutet, dass nicht nur vom SP eingesetzte Werkzeuge wie Präprozessoren und Batch Scheduling Systeme um entsprechende Schnittstellen zu erweitern sind, sondern auch Komponenten der Grid-Middleware (z. B. der Globus Gatekeeper). Außerdem muss über die Konfiguration der eingesetzten Speichersysteme sichergestellt werden, dass auf Daten nicht mehr direkt, d. h. am Privacy Management System vorbei, zugegriffen werden kann.
- Der eingesetzte Obligation Monitor muss um einen Triggermechanismus erweitert werden, über den die Umsetzung der Auflagen bezüglich des Löschens von Grid-Job-Komponenten angestoßen und überprüft werden kann.

Im nächsten Abschnitt wird die Anwendung dieser Vorgehensweise auf ein konkretes Privacy Management System demonstriert.

5 Verwendung von XACML für Grid-spezifische Datenschutz-Policies

Die eXtensible Access Control Markup Language (XACML) ist eine Policysprache, die nicht nur aufgrund ihrer formalen Fundierung für wissenschaftliche Arbeiten interessant ist, sondern sich durch ihre nahtlose Integration in Web Services Infrastrukturen auch bereits in der Praxis und für die Zugangskontrolle in Grids bewährt hat [L⁺04].

Gegenüber dem von uns in [BH06] vorgestellten Konzept für eine FIM-spezifische Privacy Policy Management Architektur auf XACML-Basis wurden folgende Erweiterungen vorgenommen, die nachfolgend in einem Beispiel veranschaulicht werden:

1. Das URI-basierte Namensraumschema für die Identifikation einzelner Dienste von SPs wurde um die Benennung von Gruppen von Anbietern bis hin zu ganzen VOs ergänzt. Im Beispiel werden über das Subject-Element zuerst alle an einem Grid beteiligten Provider-Organisationen ausgewählt und über die nachfolgend unter Punkt 3 beschriebenen Bedingungen auf eine Teilmenge reduziert.
2. Die ebenfalls URI-basierte Adressierung einzelner Benutzerattribute, die durch die Policy geschützt werden, wurde um die Komponente *Grid-Job* und die unter der jeweiligen *Job-Id* angesiedelte Spezifikation von *code*, *input* und *output* ergänzt.
3. Für die Erweiterung des Vokabulars für die Spezifikation der Zweckbindung schlagen wir vor, den weit verbreiteten Standard P3P [R⁺99] u. a. um die Werte *code-optimization* und *no-backup* zu erweitern. Sie besagen, dass vom Benutzer gelieferter Programmcode für die Anpassung an die lokale Rechnerarchitektur ausgewertet bzw. zur Sicherstellung der Vertraulichkeit vom Grid-Job des Benutzers keine Offline-Datensicherung durchgeführt werden darf.

Hingegen wurden die bereits vorhandenen Möglichkeiten zur Spezifikation von Policy-Obligationen, mit denen Auflagen z. B. bezüglich der maximalen Datenaufbewahrungszeit modelliert werden können, unverändert beibehalten, da sie den erweiterten Namensraum ohne Anpassung nutzen können. Das folgende Beispiel zeigt eine XACML-basierte Privacy Policy für den Code eines Grid-Jobs:¹

```
1 <Policy id="GridJobPolicyExample1" RuleCombiningAlgorithm="first-applicable">
2   <CombinerParameter>
3     <CombinerParameter ParameterName="PolicyPriority">
4       100 <!-- Exemplarische Prioritaet, falls mehrere Policies fuer die Anfrage relevant sind -->
5     </CombinerParameter>
6   </CombinerParameter>
7   <Description> Freigabe des Programmcodes fuer Optimierungszwecke </Description>
8   <Rule id="ExampleRule1" effect="permit">
9     <Target>
10      <Resource> <!-- Daten, auf die sich die Regel bezieht, entsprechend definiertem URI-Namensraum -->
11        https://org1.example.com/username/gridjobs/projectname/code
12      </Resource>
13      <Subject> <!-- Freigabe an potentiell alle Service Provider im Grid... -->
14        https://grid.example.com/members/VO
15      </Subject>
16      <Action> <!-- ... mit Einschränkungen anhand des Verwendungszwecks -->
17        gridjobs/code-optimization
18      </Action>
19    </Target>
20  </Rule>
21 </Policy>
```

¹Zur besseren Lesbarkeit wird eine vereinfachte XML-Schreibweise ohne XML-Namespaces verwendet.

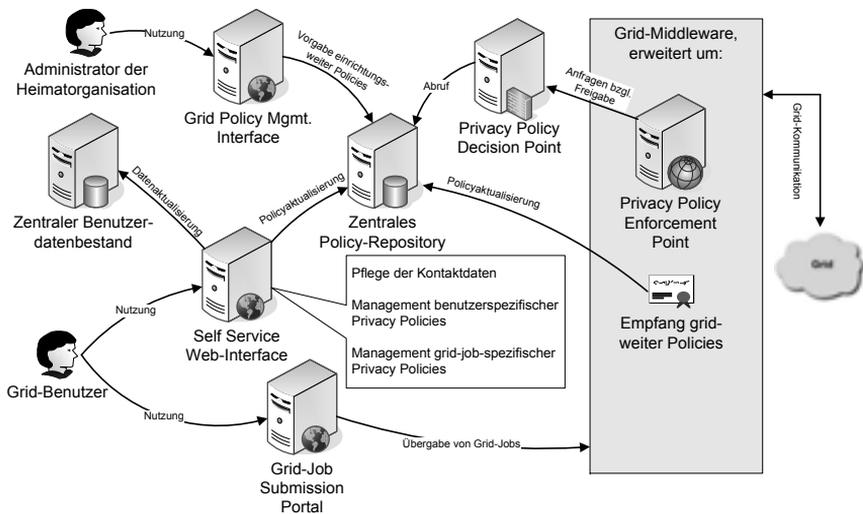


Abbildung 2: Integration der Privacy Management Komponenten bei der Heimorganisation

Analog zur gezeigten Policy können weitere Regeln für die anderen Bestandteile des Grid-Jobs sowie für die persönlichen Daten des Benutzers definiert werden; sie werden typischerweise durch eine abschließende „Catch-All“-Policy mit niedrigerer Priorität, die die Freigabe aller weiteren Daten unterbindet, komplementiert.

6 Integration der Lösung in Grid-Architekturen

Bei der Realisierung der Privacy Management Infrastruktur in Grids muss zwischen der Heimateinrichtung des Benutzers, über die dieser Zugang zu den Grid-Ressourcen erhält, und den SPs unterschieden werden. Dabei ist zu berücksichtigen, dass in aktuellen Grids die meisten Organisationen beide Rollen, ggf. in unterschiedlich starker Ausprägung, wahrnehmen. An eine skalierbare Lösung wird die nahe liegende Anforderung gestellt, dass von technischen Komponenten nur jeweils eine Instanz, und nicht beispielsweise eine Instanz pro Grid-Projekt, an dem die Organisation beteiligt ist, benötigt wird.

Abbildung 2 zeigt die Erweiterung der Heimorganisation des Benutzers um Grid-fähige Privacy Management Komponenten. Ein Privacy Policy Decision Point (PDP) wertet Grid-weite Vorgaben, administrative Voreinstellungen und benutzerspezifizierte Privacy Policies aus. Beim Management der Policies durch den Benutzer wird bereits in der graphischen Benutzeroberfläche zwischen persönlichen und Grid-Job-spezifischen Policies unterschieden. Ein Grid-Job kann nur zur Ausführung kommen, wenn von der Middleware geeignete Service Provider ermittelt werden, die die benötigten Ressourcen zur Verfügung stellen und die benutzer- und Grid-Job-spezifischen Privacy Policies erfüllen können.

Die Architektur beim SP ist hierzu komplementär. Sie besteht einerseits ebenfalls aus einem PDP, der darüber entscheidet, ob dem Benutzer u. a. aufgrund der geforderten datenschutzrelevanten Auflagen der Zugriff auf die lokalen Ressourcen gewährt werden kann. Andererseits wird ein Obligation Monitor eingesetzt, der nicht nur den lokalen Benutzerdatenbestand berücksichtigt, sondern auch die von den Grid-Jobs belegten Dateibereiche *code*, *input* und *output*. In [B⁺06] wurde gezeigt, dass sich die Polycysprache XACML auch auf der Seite des Service Providers effizient im Rahmen von Privacy Management Systemen einsetzen lässt. Fungiert eine Organisation sowohl als SP als auch als Heimateinrichtung von Benutzern, wird somit lediglich *ein* gemeinsamer XACML PDP benötigt, der aufgrund der getrennten Namensräume auch beliebig viele Grid-Projekte unterstützt.

Wie bereits in Abschnitt 4 diskutiert wurde, besteht der größte Eingriff in die vorhandene Infrastruktur darin, den Zugriff auf personenbezogene Daten und Grid-Jobs nicht mehr direkt zuzulassen, sondern die Nutzung des Privacy Management Systems zu erzwingen. Hierfür können die aus dem FIM-Umfeld bekannten Migrationskonzepte (vgl. [Hom07]) und/oder Middleware-spezifische Erweiterungen (siehe z. B. [Ami07]) eingesetzt werden. Um das Commitment zu einer raschen Grid-weiten Umsetzung zu erwirken, sollten jedoch auch Anreize auf organisatorischer Ebene geschaffen werden.

Die praktische Umsetzung dieser Konzepte ist jedoch bei Weitem noch nicht abgeschlossen und wird noch längere Zeit in Anspruch nehmen, weil sie auch eine Vielzahl von Organisationen betrifft, deren interne Benutzerverwaltungssysteme noch kein explizites Privacy Management System umfassen. Da jedoch auch die Benutzerakzeptanz und somit die Nachhaltigkeit geschaffener Grids stark von der Erfüllung von Datenschutzauflagen abhängt, zeichnen sich bereits erfreuliche Fortschritte ab (vgl. [Hr07]).

7 Zusammenfassung und Ausblick

In diesem Artikel haben wir eingangs erläutert, warum bisherige Privacy Management Ansätze für das Umfeld des Grid Computings noch nicht ausreichen. Auf Basis der Besonderheiten von Grids und den resultierenden datenschutzspezifischen Anforderungen haben wir ein Konzept vorgestellt, das als Basis zur Erweiterung von Privacy Management Systemen um Grid-Charakteristika herangezogen werden kann. Die Umsetzung dieses Konzepts wurde am Beispiel der Polycysprache XACML demonstriert; ebenso wurde eine Strategie zur Integration der resultierenden Lösung in Grid-Architekturen aufgezeigt – aus unserer Sicht eine zwingende Notwendigkeit für zukünftige Grid-Systeme, die mit spontaner Integration Grid-fremder Ressourcen in „Überlastfällen“ reagieren müssen (z. B. bei Pandemie-Berechnungen, Katastrophensimulationen oder spontanen Simulationen von Materialbelastungen in Notsituationen).

Unsere aktuellen Forschungsarbeiten konzentrieren sich auf die Zusammenführung von Grid-spezifischen Service Provisioning Mechanismen mit den aus dem Federated Identity Management bekannten, policybasierten Federated User Provisioning Methoden. Ein wesentliches Ziel dabei ist, der Ausgangssituation Rechnung zu tragen, dass Ressourcen nur zum Teil an Grid-Projekte „delegiert“ werden und parallel dazu nach wie vor

von lokalen Benutzern verwendet werden. Bei der Integration der zu unterstützenden Nutzungsvarianten stehen deshalb Aspekte wie Zugangskontrolle, Auditing und Compliance sowie der Datenschutz im Vordergrund – gepaart mit der Notwendigkeit, Middleware-unabhängig zu bleiben.

Danksagung:

Die Autoren danken den Mitgliedern des Münchner Netzwerk-Management Teams (MNM) für hilfreiche Diskussionen und wertvolle Kommentare zu früheren Versionen dieses Artikels. Das MNM-Team ist eine Forschungsgruppe der Münchener Universitäten und des Leibniz-Rechenzentrums der Bayerischen Akademie der Wissenschaften unter der Leitung von Prof. Dr. Heinz-Gerd Hegering.

Literatur

- [A⁺04] Robert Aarts et al. Liberty ID-WSF Interaction Service Specification. Liberty Alliance Specification, 2004.
- [Ami07] Hervé Amikem. Prozess-orientiertes Monitoring Virtueller Organisationen in Globus-basierten Grids. Diplomarbeit, Fakultät für Informatik der Technischen Universität München, September 2007.
- [B⁺06] Latifa Boursas et al. Policy-based Service Provisioning and Dynamic Trust Management in Identity Federations. In *Proceedings der IEEE International Conference on Communications (ICC 2006)*. IEEE Press, 2006.
- [BH06] Latifa Boursas and Wolfgang Hommel. Policy-gesteuerte Datenfreigaben und Trust Management im organisationsübergreifenden Identitäts-Management. In *Proceedings der SICHERHEIT 2006*. Springer Lecture Notes in Informatics, 2006.
- [Hom07] Wolfgang Hommel. *Architektur- und Werkzeugkonzepte für föderiertes Identitäts-Management*. Dissertation, LMU München, ISBN 978-3-89963-594-2, 2007.
- [(Hr07] Bruno Crispo (Hrsg.). First International Workshop on Security, Trust and Privacy in Grid Systems. In den Proceedings der SecureComm 2007, IEEE Press, ISBN 1-4244-0975-6, 2007.
- [L⁺04] Markus Lorch et al. An XACML-based Policy Management and Authorization Service for Globus Research Resources. Technischer Bericht im Department of Computer Science, Virginia Tech, 2004.
- [Mon04] Marco Casassa Mont. Dealing with Privacy Obligations in Enterprises. Technical Report HPL-2004-109, HP Laboratories Bristol, 2004.
- [NKH07] Heike Neuroth, Martina Kerzel, and Wolfgang Gentzsch (Hrsg.). *Die D-Grid Initiative*. Universitätsverlag Göttingen, ISBN 978-3-940344-01-4, 2007.
- [R⁺99] Joseph Reagle et al. The Platform for Privacy Preferences. In *Communications of the ACM*, volume 42, pages 48–55. ACM Press, 1999.
- [Sch07] Michael Schiffers. *Management dynamischer Virtueller Organisationen in Grids*. Dissertation, LMU München, 2007.