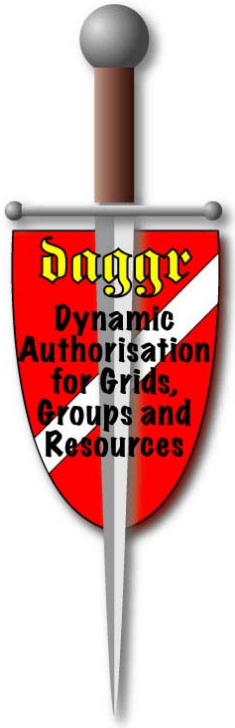
Information Society Technologies

Dynamic Authorisation for Grids, Groups and Resources

DAGGR

Part B
Version-16
October 14, 2003



The Sixth Framework Programme
(2002-2006)

Project Full Title	**Dynamic Authorisation for Grids, Groups and Resources**
Project Acronym	**DAGGR**
Proposal Number	**IST-2003-003849**
Date of Preparation	**October 14, 2003**
Type of Instrument	**Specific Targeted Research Project**

List of participants

Participant no.	**Participant name**	**Participant short name**
1 (coordinator)	Foundation for Fundamental Research on Matter, Utrecht, The Netherlands	FOM
2	University of Manchester, Manchester, United Kingdom	UoM
3	Istituto Nazionale di Fisica Nucleare, Frascati, Italy	INFN
4	Masarykova Univerzita v Brne, Brno, Czech Republic	MU BRNO
5	Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany	FZK

Coordinator

Name	K.Bos
Organisation	FOM
Email	k.bos@nikhef.nl
Telephone	+31-20-5925003
Fax	+31-20-5925054

Acronyms

ADS	Authorisation Decision Services
API	Application Programming Interface
AS	Attribute Service
ADS	Authorisation Decision Service
BGP	Border Gateway Protocol (internet routing protocol)
CA	Certification Authority
DAGGR	Dynamic Authorisation for Grids, Groups and Resources
DoA	Domain of Authority
IETF	Internet Engineering Task Force
EDG	European DataGrid Project
EDT	European DataTAG Project
EGEE	Enabling Grids for eScience in Europe
ERA	European Research Area
EU	European Union
GACL	Grid Access Control Language
GGF	Global Grid Forum
GSI	Grid Security Infrastructure
LCAS	Local Centre Authorisation Service
LCG	LHC Computing Grid
LDAP	Lightweight Directory Access Protocol
LHC	Large Hadron Collider
NGG	Next Generation Grid(s)
NIS	Network Information Service
OGSA	Open Grid Services Architecture
PAG	Project Administrative Group
PEB	Project Execution Board
PKI	Public Key Infrastructure
PM	Project Manager
PO	Project Office
PRG	Project Referee Group
RAUCS	Resource Access and Usage Control Services
RBAC	Rule Based Access Control
RFC2904	RFC on the AAA Authorisation Framework
SAML	Security Authorisation Mark-up Language
SASL	Simple Authentication and Security Layer
SoA	Source of Authorisation (generalisation of VOMS)
VC	Virtual Community
VO	Virtual Organisation
VOMS	Virtual Organisation Membership Service
WP	Work Package
WPL	Work Package Leader
XACML	eXtensible Access Control Mark-up Language
XML	eXtensible Mark-up Language

Definitions

GridSite[1]
A set of Grid extensions to the industry-standard Apache web-server, supporting GSI authentication, VOMS attributes, the Grid Access Control Language (GACL) and access policies in XML.

SlashGrid[2]
A framework for providing local and remote filesystems with access controlled by Grid credentials (including VOMS attributes) and specified by GACL access policies.

Virtual Community (VC)
Group of users and resources in the same administrative domain (Domain of Authority) sharing the same security infrastructure (e.g. PKI-X.509, Kerberos[3], Unix passwords etc.). The authorisation of a VC can be structured with roles (e.g. conforming to the RBAC model). Examples of VC's are a cluster of PCs under NIS and a Kerberos realm.

Virtual Organization (VO)
Union of autonomous entities, belonging to different administrative domains, aimed to a common goal. The VO can be structured in groups and subgroups, not necessarily corresponding to the original autonomous entities, to reflect the needs of the collaboration. In this respect, the organisation of users is centralised (the day-by-day group management can be delegated). Examples of VOs are the collaborations formed around the LHC experiments in the LCG framework.

Virtual Organisation Membership Service (VOMS)[4]
The purpose of VOMS is to grant authorisation data to users at VO level. VOMS provides support for group membership, forced groups (i.e. for negative permissions), roles and capabilities. This system, developed in the EU DataGrid and DataTAG projects, is backward compatible with user-based Grid Security, whilst providing more dynamic capabilities for hierarchies within a VO.

X.509
X.509 is a format for certified Public Key's, which are suitable for use in a Public Key Infrastructure. These certificates are useful for encrypting messages using Privacy Enhanced Mail and forming SSL network connections, such as are used in HTTPS. It is described in standard ISO/IEC 9594-8.

[1] http://www.gridpp.ac.uk/authz/

[2] A. McNab, "Grid-based access control for Unix environments, Filesystems and Web Sites", Proceedings of Computing in High Energy and Nuclear Physics (2003)

[3] http://web.mit.edu/kerberos/
[4] http://grid-auth.infn.it/

Contents

Proposal summary

Dynamic Authorisation for Grids, Groups and Resources

DAGGR

**** DISTRIBUTION IS LIMITED TO THE UvA COURSE ON GRID MIDDLEWARE, 2006 ****

Collaborative efforts, joint ventures, and the sharing of resources form the basis of both business and science in the society that we live in today. Information technology is the glue that holds these efforts together, and the Grid will bring new opportunities as content and knowledge will be shared electronically. Although current grids address some of the issues involved in setting up collaborations and joint ventures, it has been widely recognized that the technology available today is not yet ready for the dynamics of true user-centric, scalable and interoperable systems.

DAGGR – *Dynamic Authorisation for Grids, Groups and Resources* – addresses the core issues that hamper the wider adoption of Grid technology to date: the building of trust and the establishment of security across multiple administrative domains, using methods and concepts that are rigorously based on standards. This way, we can converge on one Grid, where virtual organisations can not only co-exist, but interact with each other in a secure way.

To enable a dynamic Grid, one should look at the complete spectrum of technologies that underpin security for complex problem solving: from hiding the complexities of security from the end-users, right down to the automated control of systems management for grid resources.

In this project, four key concepts for interoperation are addressed: credential management, managing authorisation attributes, taking the authorisation decisions, and the enforcement thereof. A fifth task addresses the specifics of federating multiple ‘ad-hoc’ virtual organisations in order to give a user-centric view on the entire Grid.

When all these elements are brought together a coherent picture emerges of what the next generation of grid systems will look like: transparent to the user, rigorously based on standards, pervasive and secure. Something science, industry, and individuals in the society at large will look towards as their preferred way of collaboration.

B.1 Scientific and technological objectives of the project and state of the art

B.1.1 The Problem

Collaborative efforts, joint ventures, and the sharing of resources form the basis of both business and science in the society that we live in today. Information technology is the glue that holds these efforts together and aids in new scientific discoveries and increased sales. The World Wide Web being its most prominent exponent today, the Grid will bring new opportunities as not only data, but also storage, computation and even knowledge can be shared electronically.

Although current generation grids address many of the issues involved in setting up collaborations and joint ventures, it has been widely recognized that the technology available to us today is not yet ready for the dynamics of a true user-centric, scalable and interoperable system. Although the Grid does integrate a large variety of resources, it has traditionally focussed on building rather large and long-lived "Virtual Organisations" – the collections of users and service providers that join forces to tackle a common problem. Bound by the constraints current Grid technology pushed upon them, and by the lack of adequate standards, they have to use the same 'middleware' and identical ways to identify and authorise requestors and providers.

With this project, we address one of the core issues that hampered the wide adoption of Grid technology to date: the building of trust and the establishment of security across multiple administrative domains, using methods and concepts that are rigorously based on standards. The Expert Group Report on Next Generation Grids[5] indicated clearly that, although many different Grids may exist now, they should converge into one Grid, where various virtual organisations can not only co-exist, but also interact with each other in a secure way.
We are convinced that our proposed dynamic authorisation for grids can bring together user groups and resources in all fields of science and industry. With this work we can take the initiative in setting the relevant standards for Federations of grids, and in defining the architectures for the formation and dissolution of collaborations.

To enable such dynamic joint ventures on the Grid, we need to have a look at a wide spectrum of technologies: how can the complexities of key management be hidden from the end-users, right down to the automatic control of systems management for grid resources. In this project, we propose an integrated approach to these problems. Four tasks are directed to address the fundamental concepts needed for interoperations: credential management, authorisation attributes, authorisation decisions, and the enforcement thereof. A fifth task addresses the specifics of federating multiple 'ad-hoc' Virtual Organisations in order to give a user-centric view on the entire Grid.

When all these elements are brought together a coherent picture emerges of what the next generation Grid systems will look like: transparent to the user, rigorously based on standards, pervasive and secure. Something science, industry, and individuals in the society at large will look towards as their preferred way of collaboration.

B.1.2 Objectives

DAGGR will create a new overall security architecture and new security components for a next generation of Grids that will be far more secure and more flexible. This enables the new Grids to be applicable to and cost-effective for a new range of applications not addressed by the current

[5] *Next Generation Grids, European Grid Research 2005-2010*, *Expert Group Report June 2003.*

architectures. In particular, it will enable better synergy between Grid security architectures and peer-to-peer collaboration mechanisms.

In order to enable the dynamic creation and management of this next generation of Grids, the DAGGR project will answer the following key questions:

- What additions are required to provide authentication in these dynamic environments, where users may have several identities that may be used to achieve various specific tasks?
- What new community management systems are needed to be able to provide collaboration on-demand?
- How can sites enforce appropriate access and usage control for remote users, where the chain of trust and the authorisation policies are themselves dynamically configured?
- How can access and usage control be expressed in a general way by users and administrators, and how can resources such as file servers and execution environments apply this to existing computing concepts such as files, objects and database entries?
- How can Grid federations be used to find solutions for the above questions, in case security and trust is to be managed across multiple administrative domains?

DAGGR will address design limitations of current Grids by the introduction of new concepts in security, like forwarding and federation. This will lead to more flexible and dynamic Grids and will thus greatly improve interoperability of heterogeneous systems. We will co-ordinate our efforts with the working and research groups within the Global Grid Forum, rigorously base us on standards where available, and foster the establishment of new standards based on our work. With this project we will work towards the vision of One Grid, based on the paradigm of a semantic grid that hides the complexity through new ideas like service federation and other novel security techniques. The proposed project duration is two years and the research will be done by four people at each of the five partners.

B.1.3 Expected Results

The result of the project will be a comprehensive security architecture that enables the dynamic creation, operation and decommissioning of virtual collaborations that span multiple ‘traditional’ Grids. We will create prototype environments demonstrating our ideas, and we will ensure that these prototypes allow for incremental incorporation in both existing and new Grid projects in Europe and beyond.

Specifically, we will deliver architectures and software to:

- Manage credentials in heterogeneous multi-domain environments, and mediate between domains in establishing mutually-intelligible authentication
- Create and manage virtual organisations, and provide a semantic mapping between different virtual organisations within a federation
- Decide on authorisation to use Grid services, based on locally defined, compound policies
- Enforce access rights, usage limits and other components of Service Level Agreements on managed resources
- Federate virtual organisations based on policies, leading to a user-centric view on the Grid

We will ensure that relevant standards are being set where appropriate, and undertake the necessary efforts to set new standards based on the results of this project.

B.1.4 State of the Art

The first generation Grids consisted of middleware from a single vendor or project that was used by largely unmodified applications to access remote resources. In this generation, very simple authorisation schemes were possible, largely due to their use for demonstrations of the technology rather than production, involving a small number of users, whose authorisation could be managed by hand at each resource.

The second generation consists of Grids composed of middleware from several sources, with published and stable interfaces. The EU DataGrid (EDG) provides a good example, with middleware from the Globus Alliance, Condor and systems developed by EDG. The large number of users (hundreds) and sites (tens) associated with second generation Grids means that authorisation must be managed by automated procedures that derive policy from central, but manually configured sources of authorisation. In the case of EDG, this information is either pushed to resources by the virtual organisation administrator, or supplied to users as signed attributes by attribute certificate servers. Although flexible once created, these systems involve a significant amount of start-up effort for users, the virtual organisation, and resource administrators, and are intended to be long-lived and, in that sense, static.

The next generation Grids will provide ‘Grid Services’ as an extension of Web Services, and will be rigorously based on standards, such as those endorsed by the Global Grid Forum. They will provide a dynamic, user-centric Grid, where services can be created on-demand, and higher level services be composed and provided to users by combining lower level services. The Globus Toolkit 3 and other implementations of the Open Grid Services Architecture aim to provide middleware for this generation. In these environments, it will be severely limiting if the security systems are not equally dynamic, and able, for example, to create Virtual Organisations on demand for short-term groups of users.

The DAGGR project addresses this challenge by providing the security architecture for the next-generation Grids for solving complex problems, and by providing prototype implementations demonstrating the usefulness and applicability of these architectures for real-life Grids today and tomorrow.

B.6 Detailed Implementation Plan

In the following paragraphs we will explain how the problem space has been divided into domains which quite naturally can be translated into workpackages. After some use cases from different application areas the idea of federation of services is explained by a detailed description of a general use-case. Subsequently there is a paragraph on each of the workpackages where the state of the art, the new idea and the way this new idea will improve the current situation is described. We will work towards one common demonstrator to which each of the work packages will contribute its specific development and show all the components together lead to a better security framework. This demonstrator is described in a separate paragraph. In the last paragraph of this chapter the workplan is presented showing a detailed agenda for developments within the work packages and points of adjustment and tuning where all project participants meet to make sure the work proceeds towards a common goal and along agreed standards.

B.6.0 Introduction to service federation

Below some specific use cases will give an indication of the magnitude of the problem. Subsequently a general use case is worked out in detail to illustrate some of the ideas behind federation.

The Bank-Telecom Use Case

In daily life people are members of many different Virtual Organisations and have a (often different) identity in each of them. If they can prove their identity, they can use of the services of each of these Vos. Take the simple case of a telephone company where the identity of a client is a telephone number. Another example of a more protected environment is a bank where the identity of a client is a bank account. One can access the services of the bank only with a bank card and one has to prove its identity by remembering a pin code. Another example of a less secure VO are the frequent flyer programs which airplane companies have set up where the identity is a number and the proof of identity is the combination of a member card and a passport. Many more examples can be identified and all of them are implemented as computer systems.

As a random example: suppose it is December and a bank wants its customers to put money in their saving accounts before the end of the year. They therefore want to launch a publicity campaign where customers that do so may use their telephone for free the whole day of December 31. In this case there are two Virtual Organisations, the telephone company and the bank that have to collaborate. They have to create on a short timescale and for a short time a Virtual Community of people that have put a substantial amount of money on their savings account. This Virtual Community now must be authorised to make use of the services of another Virtual Organisation: the telephone company. It goes without saying that accounting will have to be implemented because the bank will have to pay the telephone company and want to know if their campaign was profitable. There is much detail that could be added to this example but even in this simple form the problem should become clear.

There are many ad hoc ways to implement this use case and it has been done so in the past. However the problem is general and a standard approach doesn't exist. The problem is how on a short timescale can a Virtual Community be granted services in a domain and how can those rights be propagated to a different domain. How can the identity of a member of the Virtual Community be proven in one domain and how can that trust be federated to others? In the above example there were only two organisations mentioned but it is not difficult to find cases where more than two are involved. As all these organisations are implemented as computer systems all of them are problems

in the IST area. At present use-cases as the one mentioned above are solved on an ad hoc basis and such solutions cost a lot of resources each time.

The High Energy Physics Use Case

The traditional concept of Virtual Organisations is now being adopted in many scientific areas and in industry. But in the same way that hierarchical systems on the Internet made way for dynamic peer-to-peer structures, this is bound to happen for Grid security as well. The high-energy physics community may serve as an excellent example. Using today's state-of-the-art Grid authorisation methods, that scientific domain is currently setting up virtual organisations for their large experiments at the LHC accelerator at CERN. Management boards are created to delegate rights to specific working groups, and bureaucracy is put in place to manage these long-lived Virtual Organisations. With today's tools, it is the only viable way to go, and such formal structures will indeed help to do good science in the next few years.
But this hierarchical system is not resistant to the more chaotic way in which science will be done once real students start working on real data. Scientists, and especially those in Europe, are dispersed over many university groups and smaller institutes. They have good local ties with compute and network service providers, and can put their embedding in larger faculties and campuses to good use in acquiring resources locally. In such a world, many small "domains of authority" (mini-Vos) will emerge. Resource sharing and trust building amongst those will take place in many small venues, and should not be hindered by large formal structures.

The BigFilmStore Use Case

But such dynamic networks are certainly not limited to big science. Authenticated and protected peer-to-peer networks have similar needs. For example user Jane with her simple PC and her PKI certificate signed by Small-CA with *BigFilmStore* certified by Big-CA (where Jane doesn't trust Big-CA and *BigFilmStore* never heard of Small-CA). Of course both must confide in *TrustfulServer* which is the catalyser for the establishment of the Federation "Jane + *TrustfulServer* + BigFilmStore". In this way BigFilmStore could sell services to Jane in complete security, without the need for Joe to register before, and Jane having to reveal her identity to the store.

To enable such a heterogeneous environment, key questions regarding trust building, authorisation, and resource access must be addressed. This federative scenario involves the mediation of trust, the translation of authorisation data, local resource protection, and hooks for accounting and tracing. And interoperability, protocol negotiation and standards are essential to make this work.

Federation: a case description

Since federating domains of authority is a new concept, both within the Grid community as well as in other applications, it cannot readily be predicted what the model should be: it is the topic addressed in this proposal. However, an example scenario that addresses the use cases described above may help to illustrate the basic concepts that underlie Grid Federation. At the same time, the scenario highlights the core components that will be required to enable a viable federative Grid, and thus will quite naturally into the five work packages proposed.

Creating the first Domain of Authority: the Virtual Organisation

Let us assume that an organization "A" sets up a Domain of Authority (DoA) – conventionally called a single "Virtual Organisation" – to address a (scientific) issue. The organization has individuals that will work within the DoA, and brings resources (a storage system) into the domain.

The DoA is established by the instantiation of a VO Membership Grid Service. In this case, it is a VOMS service that issues X.509 attribute certificates. Organisation "A" uses a Public Key Infrastructure, outsourced to a trusted third party (a Certificate Authority or CA). Thus, the DoA "α" is established. This configuration is shown in Figure 1a. Note that both the VOMS service and the storage area provided by "A" are protected by a similar Authorisation Service (AS) that honours authorisation assertions from "α".

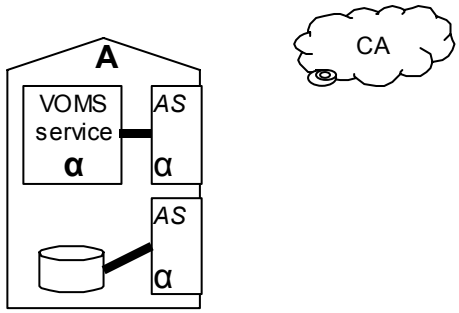


Figure 1a: a Virtual Organisation "α" set up by organisation "A".
Trust is derived from a Certification Authotity (CA).

The DoA "α" acquires computational power from organization "B". But this second organization does not use a PKI, and will not trust the CA used by organization "A" for authentication. By contract, however, it needs to interoperate with the users registered in DoA "α", so it established a credential translation service: a separate service, that will check the attributes issued by the VOMS server of "α", in a request authenticated by the CA that authenticated DoA "α" in the first place. But this translation service will produce new authentication and authorisation information, useable for access to the compute cluster at "B", derived from the authorisation structure within DoA "α". The tokens have a different format (say Kerberos tokens from realm "א") but they convey the same information as the attribute certificates issues by "α" itself.
A user at organization "A" can now go to the translation service, add the Kerberos tokens from "א" to his credential wallet, and use both the storage at "A" and the computers at "B" to do the job. This is shown in Figure 1b.

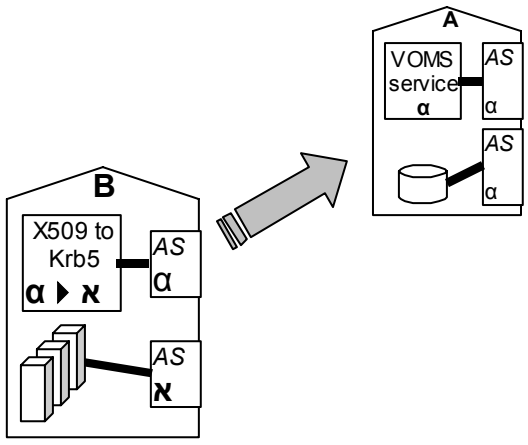


Figure 1b: A new organisation, that uses authorisation tokens of another format (א) joins VO α

Roaming users

Or course, users within a DoA need not be located at a resource provider. People are bound to travel, and roaming users can show up at many different organizations. On-line credential proxies are a reliable way of supporting such roaming users. Even if a user is in an un-trusted environment (like an airport kiosk at airport "C"), an organization that she trusts (like "A") can run an on-line credential repository, where a user get access to his or her credential wallet. The authorisation service that protects the repository can allow for many different methods of authentication, like biometrics or one-time crypto-cards, so that the user does not need direct contact with the CA or "A" or the Kerberos centre of "B". Access control will protect the individual items in the credential repository.
It is shown in Figure 1c.

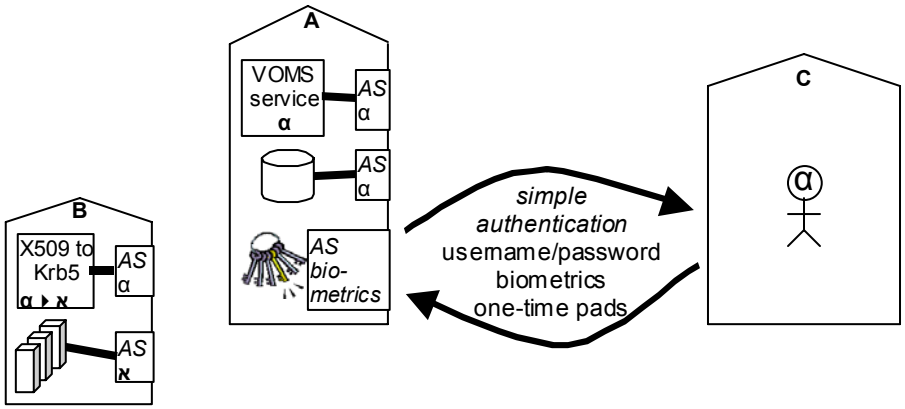


Figure 1c: a roaming user of α currently at C uses an on-line credential repository

What is described up to now can still be labelled as a "classical" Virtual Organisations. But many new concepts have been introduced: the advanced forms of credential translation and mixing of authentication methods are largely unexplored territory, as well as the credential repositories that implement restricted delegation, and the creation of VOs has become virtually instantaneous. Omnipresence of authorisation services pushes the flexibility of the VO to new heights.

Federation

It becomes more challenging when two DoAs want to join forces and work together, perhaps for a short period of time.

Figure 2a shows two DoAs, the one described previously (containing the organisations "A" and "B"), and a second DoA consisting of another organisation "D". This new organisation already has its own DoA, "β", and has given that DoA access to their storage facility. The AS of this storage system is configured to allow access to anyone in "β", and entities in "β" retain their role and group access rights that are specific to this VO.

In the simplest case the entire DoA "β" (up to now, that only is "D") will federate with "α". We will also assume for the time being that "β" uses PKI authentication with the same CA as the one used in "α", although also the more complex cases – where different kinds of credentials are used – are within the scope of this proposal.

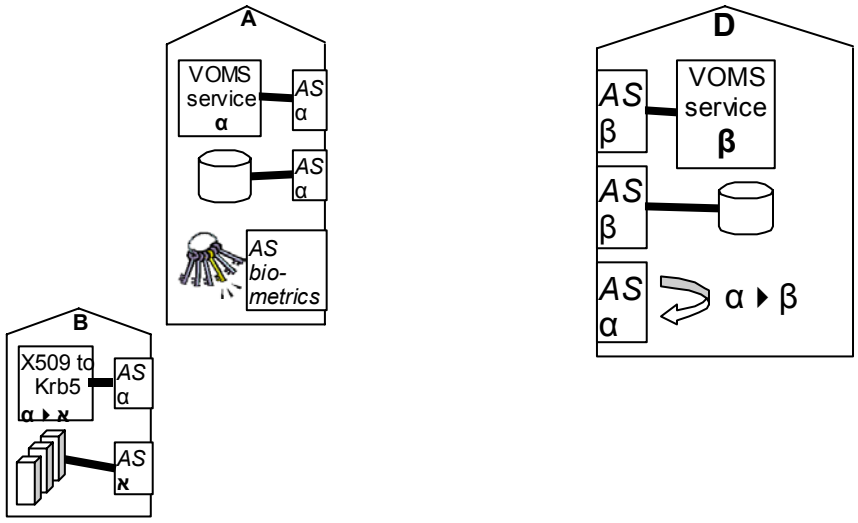


Figure 2a: A second domain of authority, β, is going to federate with α, and establishes a bridging service

There are now two options for the federation: either all users from organisations "A" and "B" will appear to the storage systems in "D" as a generic "user-of-α", or a role translation service is established at "α" or "β" that can map the appropriate groups and roles into each other. Which one of the two is best depends on the duration of the federation and the functional requirements of the VO. The Federation Activity (workpackage 5) in this project will address this question in-depth.

In any case, the AS's that control access to the resources should now negotiate a compatible authorisation method between the requestor (say, a user in "A") and their set of trusted Sources of Authority. To get the proper tokens for accessing "D", the user goes to the local VOMS server of "α", requesting tokens for use in "β". It is the VO service of "α" that will forward the request to "β", according to the rules laid out in a federation agreement. These tokens will be handed to the AS in "D" in order to use the service.

Indirect trust and trust chains

It becomes more complicated when trust is used indirectly. A resource provider in "E" may have joined forces with "D" in the "β" domain of authority, but is not willing to perform for users of "α". In the general case, this is a complex matter, and a significant amount of work in all the activities foreseen in this project, is needed: for instance preventing fraudulent actions (like "D" giving out attribute assertions of "β" without proper authorisation), and trust building through multiple SoA's. But ignoring for the time being the possibility that organisation "D" (hosting the "β" DoA) is fraudulent, part of the federation protocol can include the insertion of "derivation sources" in the attributes signed in "β" for requestors coming from "α". This is in effect a form of delegation tracing that is currently already being researched for user-initiated delegation.

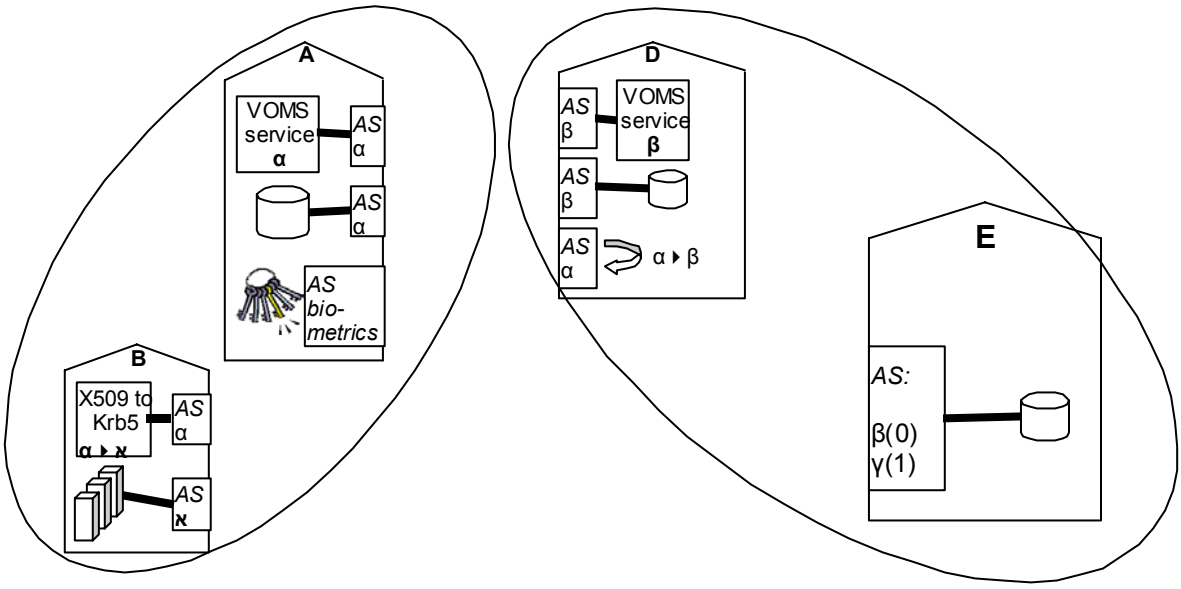


Figure 2b: organisation E in a member of Vos β and γ, and allows for one level of indirection, from γ only (indicated by "γ(1)") but does not allow indirect users from β (indicated by "β(0)")

When applied in a federative context, it allows for fine-grained access control inside a newly forged federation. Organisation "E" should be able to implement a policy in the AS protecting its resource, that considers not only the direct rights presented by the user (in this case still coming from within "α"), but can also inspect the route that leads to the ultimate source or authority: the VOMS server at organisation "D".

It is clear that the concept illustrated above for federating two DoA's can be recursively applied to more complex and more deeply nested federations.

Information and service advertisement

When complex federations are forged, the process of discovering *accessible* services becomes far more complicated than in case of a single virtual organisation. Delegation path-length constraints limit resource availability for users in the other virtual organisations, and even a VO that looks homogeneously accessible from within, may appear as a set of dispersed services to users elsewhere in the federation.
As the federation grows, it becomes increasingly more important for users to learn only about those resources that are actually willing to perform, and not merely 'visible' to the federation as a whole. Such information is to be distributed inside the federation, and its visibility will be researched as part of the Federation Service Activity.

An especially appealing case is 'route redundancy' in accessing resources in a federation: if three or more DoAs federate, there may be more then one delegation trace that leads to the same resource, as shown in Figure 3.

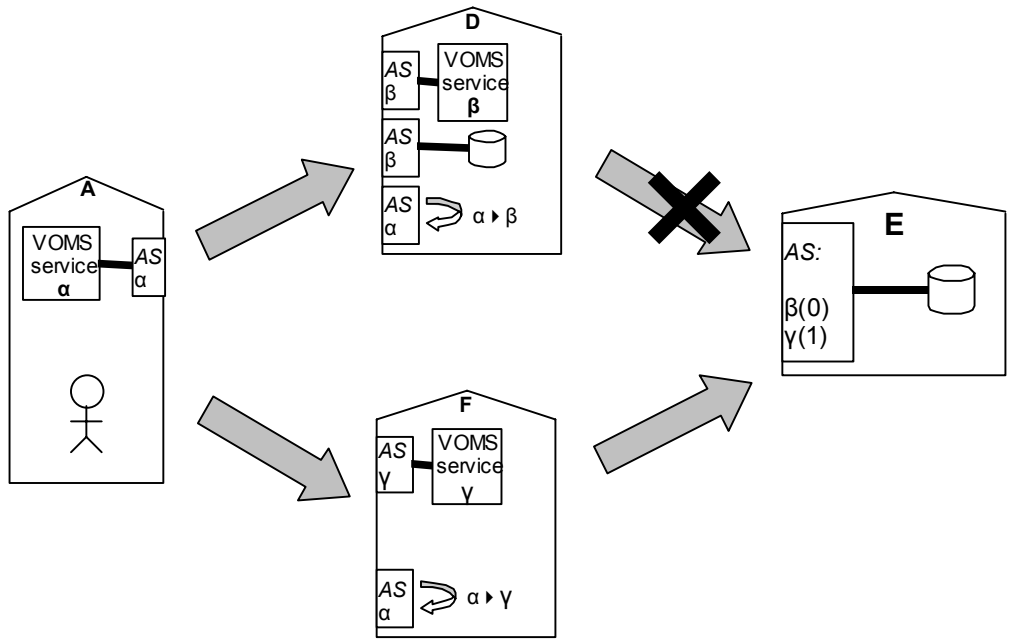


Figure 3: more than one authorisation path leads from a user in α to the resource in E, but only through γ those resources are accessible, due to delegation path-length constraints

Organisation "E" is a member of two DoA's: "β" and "γ". Part of the DoA agreement with "β" was that only direct users in "β" can use the storage in "E", but that this right is not transferable (i.e. there is a path-length constraint in effect). But "E" also has an agreement with DoA "γ", that allows for one additional level of indirection.
Suppose that a federation is forged between "α", "β" and "γ". A user in "α" cannot use the attributes asserted by "β" to access resources in "E", but if that same user asks the other member of the federation, "γ", for an attribute assertion, the resources in "E" do become accessible. Thus, depending on the delegation path, resources may or may not seem part of your federation.

Core components and the work packages

Already from the first scenario step, is becomes clear that differences in the credential mechanisms used need to be resolved. Work package 1 will be addressing these issues, thus enabling inter-domain trust establishment. Once trust is there, the `domains of authority' need to be defined; Work Package 2 will be addressing these issues, making sure that the attribute services will merge well the various Grid architectures that could be envisioned.
The interpretation of the attributes, and the protection of the Grid services involved is the domain of WP3, authorisation decision services. This component is essential in protecting underlying resources and negotiating compatible authorisation methods within a federation. The interface between the Grid and the underlying resources, and the abstraction of resources in Grid terms will be addressed by WP4. Standard ways to express resource access control are a prerequisite to enable authorisation amongst sites employing different mechanisms.
The core federation protocols that enable the complex structures described, the indirect trust and trust and authorisation chaining form the core of this work, and are addressed by WP5.

B.6.1. Workpackage 1: Credential Services

State of the art

To date, virtually all Grid infrastructures have relied on the same authentication infrastructure (GSI), that was introduced in the first release on the Globus Toolkit: a Public Key Infrastructure (PKI), modified to allow for single sign-on based on full delegation. But as the number of Grid deployments is growing, the diversity of authentication credentials and mechanisms used is increasing; the most commonly used alternative being Kerberos.

The current setup poses three main challenges. Firstly, it is becoming increasingly difficult to find a common method or trust domain for authentication. The end-user must decide on the authentication protocol to use when accessing a VO and its resources. In the authorisation area, protocols like SASL are addressing this with a negotiation phase. Work in the IETF on the PKINIT extensions to the Kerberos protocol is ongoing to allow the acquisition of Kerberos tokens based on PKI authentication.

Even when identical technology (like PKI) is applied, the agreement and establishment of a common trust infrastructure is usually a long process. Experience in the EU DataGrid project has shown that the management involved in establishing Certification Authorities and agreeing on their procedures and liabilities is a tedious exercise. And the dissemination of trust information to relying parties (resource providers and users alike) is complicated and error-prone.

A third challenge comes with the management of these credentials by end-users. This credential management is key to secure operations, but it is complicated for end-users and therefore prone to errors. A lot of sites offer their users a mechanism that makes credential management easier or more secure. Such systems include MyProxy[6] and the Virtual Smart Card project[7] that allow users to store their credentials in a dedicated repository. Others, like KCA[8], offer an on-line CA that can issue short-lived PKI credentials on-demand to users registered with an existing Kerberos realm.

Proposal

To use grids in any meaningful way, they should support the single sign-on principle across the entire breadth of resources, thus allowing users to authenticate once and further user's operation should run without requiring user to authenticate multiple times. The goal of this work package is to design and implement the core services that allow users to manage their authentication credentials in a secure way, and to acquire a compatible set of credentials for accessing services. Moreover, in the federative context, this service will support arbitration of credentials for resource access negotiation. The service is essential to trust establishment in dynamic virtual organisations that deploy a heterogeneous authentication infrastructure.

Credential services comprise two independent core components: the 'wallet' running on the user/client side, and a 'bridge' running on the end side. The 'wallet' service manages the user's credentials, and can perform transformations to various credential types – either directly using, *e.g.*, kCA or indirectly by pointing to another repository, *e.g.*, a MyProxy service. This service is also able to decide, given the authentication methods supported by the end-points, which credential type should be used to access a service. The requirements on this system will explicitly include the possibility to use the wallet service as a secure credential store, addressing the need for enhanced security by relying parties and service providers in the Grid.

[6] http://grid.ncsa.uiuc.edu/myproxy/
[7] http://slac.stanford.edu/~abh/vsc
[8] http://www.citi.umich.edu/projects/kerb_pki/

The 'bridge' service is to be called by a service end-point, when accessed using a credential not directly supported by the service. The 'bridge' service then tries to 'translate' such a credential to a compatible type of credential (*e.g.* by using PKINIT to create a Kerberos ticket from a PKI-based credential). When employed in a scenario involving a trusted third party, this service also caters for the need of anonymous but traceable authentication in the Grid. Akin to services offered by banks and clearing houses in commercial transactions, the credential translation offered by such parties hides the identity of one or both of the parties, whilst accounting information can still be exchanged.

It is considered important that a standard protocol communication of mutual authentication information is designed, akin to similar protocols defined for authorisation.

Enhancement of the state-of-the-art

The services and architecture proposed by this work package will allow more secure and convenient handling of authentication credentials in dynamic and heterogeneous Grids. It will also improve overall credential security, and address critical trust and liability issues raised by Grid service providers.
For the dynamic creation of VOs, these services will bring the ability to agree on compatible set of authentication methods without off-line negotiations of mechanisms, and allows making the single sign-on principle more general and usable in larger context.

B.6.2. Workpackage 2: Attribute Services

This workpackage addresses the task of the management of authorization attributes in the scope of an administrative domain.

State-of-the-art

One of the central concepts of the current Grid environment is the Virtual Organization (VO): an abstract entity which groups users, institutions and resources belonging to different administrative domains, sharing a common purpose.

The large number of users (hundreds) and sites (tens) associated with the current incarnation of Grids, implies that Authorization at each resource must be managed by some automated procedure, which derives local policy from one or more central, manually-managed, sources of Authorization. Local resource administrators grant rights to the VO as a whole, while VO administrators grant them to individual members of the community.

In the specific case of European DataGrid (EDG)[9] and European DataTAG (EDT)[10], this information is managed by *external, stand-alone, servers*: either published to resources by VO directory services, or supplied by users as signed attributes of their proxy certificates by VO-managed VOMS attribute certificate servers[11]. The main problem with these systems is that, although very flexible once created, involve a significant amount of start-up effort for users, VO and resource administrators. For this reason are intended to be long-lived and, in that sense, static.

For a small entity, as a Virtual Community (VC) might be, setting up a VO-based authorization infrastructure, and therefore unifying the security infrastructure, is a too heavy overhead. Moreover, within the VO model, the Authorization policies are centrally managed, even if, in principle, it is possible to delegate subgroups management. Membership in a group implies being registered as

[9] http://www.eu-datagrid.org/
[10] http://www.datatag.org/
[11] http://grid-auth.infn.it/

member of the VO. Therefore this model does not fit in a scenario where the involved entities are completely independent (e.g. peer-to-peer applications) or where anonymity is required.

Proposal

To overcome the limitations above outlined, and the strict hierarchical authorization structure, we propose to transfer the authorization machinery from the VO level and the "heavy-weight" second generation Grid servers to the VC level with "light-weight" OGSA[12] services. In this way trust establishment will be possible at the single user level, without the necessity to involve the complex procedures of VO establishment (and management). All this should be enforced, of course, in the complete respect of site security policies.

In particular, the goal of this workpackage is to build an "Attribute Service" (AS) which will be able to grant security (authorization) tokens to authenticated principals of a particular entity (e.g. Virtual Community), eventually to be a member of a Federation.

In the case of a Federation, each entity has its own AS with its specific authorization structure and mechanisms; the trust relationships between these entities allow principals to access resources in the other domains of the Federation, using credentials granted by their own domain. To make this possible, the AS must publish the information about its authorization data and structure (eventually different internal and external "views" can be assumed), in order to allow credential mapping among the systems.

To permit users and services direct access to other parties' resources, we think that the AS should support "agent", "push" and "pull" models[13], that is it should be able to:

- function as an agent between the user and the service;
- evaluate requests from the service, directly contacted by the user, returning an appropriate response;
- send authorization tokens to the user, who will forward them to the service, together with his request.

Moreover, the AS should be able to cope with the privacy and possibly anonymity requirements. A solution might be an Attribute/Pseudonym, perhaps implemented as a separated service, which would allow distributing authorized information about principals, masking their real identity (this, of course, requires that a trust domain has been established). If the Attribute/Pseudonym "service" belongs to an external (trusted) site, it should be possible to achieve anonymous access capabilities, still retaining traceability in case of need.

For the implementation of the prototype we plan to profit from the experience gained in EU DataGrid and DataTAG, where we designed, implemented and deployed the authorization infrastructure.

The use of the OGSA architecture and of recognized standards (e.g. SAML and XACML) will make it possible to interoperate with other – standards compliant – alternative solutions (e.g. the successors of CAS, Akenti and PERMIS).

Enhancement of the state-of-the-art

We think that the experience gained from this new trust establishment model could be usefully transferred to the "production grid world" like e.g. EGEE[14], where the focus is on more static relationships. Moreover, the adherence to recognized standards, the use of OGSA architecture, and the support of all the models described in RFC 2904 will make this new service (i.e. the AS) interoperable with other standard analogous services or with the legacy ones (e.g. VOMS, CAS), allowing, at the same time, a more flexible approach to the problem.

[12] http://www.globus.org/ogsa/
[13] J. Vollbrecht et al. , *AAA Authorization Framework* – RFC 2904 (http://www.ietf.org/rfc/rfc2904.txt)
[14] http://egee-ei.web.cern.ch/egee-ei/2003/index.html

B.6.3 Workpackage 3: Authorisation Decision Services

This work package provides the access control mechanisms for generic Grid services. It will provide not only the protection for the conventional Grid services, but also the access control to the attribute services and the credential translation and bridge services. It implements the access control based on the access policy language and the authentication and authorisation data presented by the clients, and performs the negotiations. It addresses site autonomy by providing hooks and call-outs for adopting hierarchical authorisation methods.

State-of-the-art

Site autonomy has been one of the key concepts of Grid computing since its conception, and is essential to convince potential participants in a Grid to collaborate. Within a Virtual Organisation, participants will share a subset of their resources under specific conditions with selected partners. Implementing such a sharing policy in the Grid thus requires an authorisation decision point and appropriate policy enforcement for all services within the VO.
Such authorisation decision functions can take many forms. The early Grid toolkits supported only lists of authorised users, maintained by out-of-band (non-automatic) communication amongst the VO members. A significant step forward was the introduction of centrally managed VO directories that are periodically retrieved by the service providers. And with the introduction of VOMS fine-grained authorisation decisions could be taken by local services, *e.g.*, using the EDG-developed Local Centre Authorisation Service (LCAS).
All presently existing VO solutions, however, are limited to static (pre-configured and 'centrally managed') Vos. This includes the central VO directories and VOMS, but also to local directories maintained by the user's home organisation, *e.g.*, in RADIUS servers used for Shibboleth. All these systems allow for relatively straightforward local Authorisation Decision Services (ADSs).

Proposal

When more flexible ways of organisation are required, the Authorisation Service controlling access to the grid services offered, needs additional functionality. In particular, it has to negotiate compatible authorisation attributes, and the format in which these are to be presented by the requesting party. Also expression of local policy and its translation into wire protocols, and the co-authorisation of services in a distributed "multi-domain" AAA request are within the mandate of the ADS.

Thus the service-local issues that need to be addressed by the ADS are:

- Interpretation of authorisation attributes presented by requesting entities.
- Hooks for negotiation of compatible security. Such a process in similar to the negotiation by, *e.g.*, SASL (simple authentication and security layer), although in the federated architecture foreseen in this proposal such negotiation can also include third-party "translator" services.
- Evaluation of local policy expression and definition, and promoting standardisation of such local policy languages.
- Delegation of authorisation decisions to other authorisation servers, and the aggregation of layered policies within a logical domain.
- Reliable co-authorisation between different services across multiple domains, supporting reliable commits in the presence of generic AAA server interaction.

The ADS is also key point in the implementation of the policies that define the federation of domains-of-authority. In order for the individual service owners to retain control, the enforcement

of the federation rules needs an implementation within the authorisation service. The federative aspects include:

- Expression and enforcement of indirect trust (models and expression of path-length constraints, transient trust and PGP models). These constrains implement the rules that result from the "meta" federation of domains.
- Publication of local federative policies. This is essential for brokering relationships within federations. An interchange format for such policies should be investigated in close interaction with both GGF and OASIS.
- Incorporation of the concept of "trust-chains". A path-length constraint in service authorisation will impose a "hop-limit" on the amount of announcement levels to be traversed (in both ways). This defined the "breadth" of the federated domain as seen by any individual entity in the federation – conceptually similar to the Internet route announcement model.

Once an entity is allowed to access a given service, the authorisation information obtained in the process needs to be conveyed to the local system. In hosted environments, this could take the form of a one-to-one translation of the attributes obtained (in a format understandable to the local domain-of-authority), but in general the resource access service (WP4) and the ADS should define an interface at with such attributes are to be exchanged and possibly mapped.
The inverse (abstracting ADS local policy expression information from the underlying resources access definitions) needs a similar interface. On these topics WP3 and WP4 will work on a common API, with potential for standardisation.

Enhancement of the state-of-the-art

When considered in the larger context of federating domains-of-authority, the model proposed opens up a complete new playing field for dynamic authorisation, and the convergence of traditional Grid Virtual Organisations and the federation model underlying modern peer-to-peer systems. Independent novel developments for the ADS include trust chaining and the inclusion of "indirect" trust, and the publication of policies for authorisation brokerage.
By providing independent implementations of standards for local policy evaluation and the negotiation of authorisation of the wire, the work will advance draft standards to a better-matured state.

B.6.4 Workpackage 4: Resource Access and Usage Control

This workpackage applies access and usage control mechanisms to fine-grained, local functionality of resources, such as file servers and filesystems, based on policies written in terms of the Grid-wide identities and authorisation credentials managed by the other workpackages.

State-of-the-art

In most Grid software, access control is currently managed by the use of lists of authorised users. For example, in systems based on Globus, the "grid-map file" list has the certificate subject names of all authorised users on a specific resource.
A few projects have begun to use more generic access policy languages, and one leading example of this is our development of the Grid Access Control Language used by the EU DataGrid for GridSite and Storage Element access control. This is a simple, XML-based language, which controls local file operations such as read or write in terms of Grid certificate identities and VO group membership.

Proposal

This workpackage of DAGGR comprises two main areas of research: the identification of how to bind Grid access control to local resources, and the management of local access policy in the dynamic, on-demand environment developed by the other work packages.

First, it will identify more general ways of applying Grid access control to low level components, such as filesystems and fileservers, but also to other local objects such as database records and batch execution queues. This will involve a survey of the requirements imposed on a general policy language by these disparate systems, the development of concrete demonstrations and the abstraction of general tools which can be applied by other research projects to their own systems.

Secondly, it will research ways of "boot strapping" local trust for externally managed policies (for example, how the owner of a resource decides whether a paying customers' policy involving third-parties is legally acceptable.) and mechanisms required for the dynamic creation of on-demand local access control environments (for example, to create a suitable sandbox environment for a user's programs, which prevents unauthorised access to the rest of the local resource.)

Enhancement of the state-of-the-art

The workpackage will be largely centred on developing a replacement for the EDG Grid Access Control List (GACL) policy language, which satisfies the additional flexibility and trust establishment requirements identified. Part of the workpackage will involve co-ordination of policy language development and prototype implementations with the work of the OGSA Authorisation Working Group in GGF.

The architecture developed will also use the Authorisation Service of Workpackage 3 as an additional source of authorisation decisions, and research the interaction between policy expression (eg GACL) within a resource, and the discovery of authorised rights from an external Authorisation Service (eg by a protocol such as SAML.)

Finally, the architecture will attempt to address requirements arising from Grid accounting research: in particular, whether these access control systems can be used for usage control, limits and quotas by the use of quantitative attributes and how they can be bound to limit mechanisms in local resources such as disks, databases and CPUs.

B.6.5 Workpackage 5: Federated Service Management

This workpackage addresses the tasks of setting up and managing a Federation of autonomous entities.

State-of-the-art

The current structure of a grid is based on the VO concept: both user authorisation management and the relationships with Resource Providers are managed in the VO framework. This VO-centric approach is probably the result of an extrapolation from the particular case of large scientific collaborations (e.g. the LHC experiments) and their relationships with the computing centres (e.g. the Tiers in LCG).

User management, even if can be distributed through a hierarchy of groups, is *de facto* centralised, since the control on each group is delegated by the VO managers only. The trust relationships between VO members are "static" and all actors must agree on a common authentication and

authorization infrastructure (e.g. PKI-X.509). This is perfectly functional to the "large scientific cooperation" model.

Moreover, all the present implementations of grid authorization mechanisms (e.g. VOMS/LCAS/LCMAPS/GACL in EDG, EDT and LCG projects) are built in this perspective, without taking into account alternative scenarios, where several autonomous entities could join, possibly dynamically, to form more complex structures, but still wanting to retain full control of their resources and users. Furthermore, these entities could use *different* authentication and authorization infrastructures (e.g. one PKIX.509, another one Kerberos, etc).

To fit in this scenario, we must *rethink* the authorisation infrastructure, introducing support for the Federation concept: a union of autonomous entities like Virtual Communities (or even Vos).

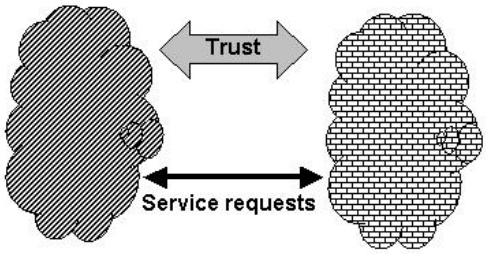
Other projects, too – developed in non-grid environments – try to address the issues arising from the Federation problem, e.g. Liberty Alliance[15], Passport[16] and Shibboleth[17]. From our point of view, the main shortcoming of all these projects is that the task of setting up the Federation is neither lightweight, nor dynamic. Moreover, in Passport, all parties need to use the same authentication infrastructure (i.e. Kerberos), and refer to a central user authentication service. Liberty Alliance, while allowing single sign-on, requires users to be "known" at every site. Shibboleth, on the other end, is oriented towards inter-institutional sharing of *web resources*, mainly fit to satisfy the requirements of roaming users.

Proposal

The goal of this workpackage is to study the mechanisms, design and implement a set of services to *dynamically* set up and easily manage a Federation, i.e. a collection of entities with trust relationships between them.

The basic trust models are the following:

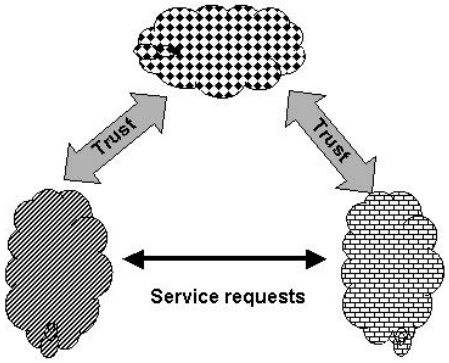
direct: the trust is a binary relationship between two entities,



Direct Trust

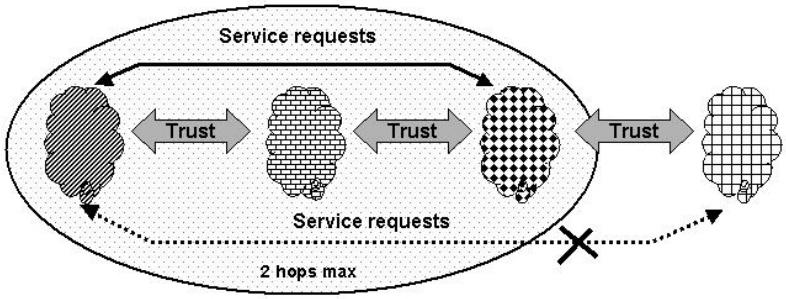
indirect: the trust relationship relies on a third-party,

[15] http://www.projectliberty.org/
[16] http://www.passport.com/
[17] http://shibboleth.internet2.edu/



Indirect Trust

delegation: where an entity acts on behalf of another one ("trust chain"), the maximum number of allowed hops is one of the basic parameters of the Federation.



Delegated Trust

These basic models can be combined to get more complex trust architectures.

The main requirements to be satisfied for the establishment of a Federation are:

- simplicity of set-up and management;
- sharing of authentication and authorization data using different mechanisms;
- no need to propagate local identity of users;
- anonymity and privacy issues.

We believe that these goals could be fulfilled by *ad hoc* service(s), active in every participating entity, with, at least, the following functionalities:

- negotiate the characteristics of the Federation with the other partner(s);
- establish the trust relationships necessary to create and modify the Federation (e.g. to change the participants or the Federation parameters);
- announce the Federation and publish its policies so that other entities can join;
- discover existing Federations to be joined;
- join and leave a Federation.

The establishment of a Federation implies also the set-up of a common policy schema that, in general, can mask the actual implementations in each domain. In this perspective, we plan to participate in the GGF group defining a (candidate standard) policy language.

Crucial points for complex Federations are the discovery of accessible services and the length limit of delegation chain to make resources available to users in other Virtual Communities.

Moreover, as the Federation grows, it becomes increasingly more important for users to learn only about those resources that are actually willing to perform, and not merely "visible" to the Federation

as a whole. Such information is to be distributed inside the Federation, and will be one of the issues addressed of this WP.

An especially appealing case is “route redundancy” in accessing resources: if three or more Virtual Communities federate, there may be more then one delegation path leading to the same resource – as shown in the example in the introduction of this section.

It is worthwhile to notice the analogies of this problem with the visibility of “Autonomous Systems” in the global Internet routing tables. The structure (with “link costs”), the dynamic character, and the information propagation (via BGP) all offer interesting sources of inspiration for constructing Grid Federations. Especially in the case of short-lived or large and dynamic federations, we can thus also expect a similar complexity level, including challenges in “Federation stability” when translation services join and leave a Federation.

Enhancement of the state-of-the-art

- The experience gained with this WP, will allow a better understanding of trust mechanisms required to set up collaborations among autonomous domains.
- Moreover the tools developed will allow a uniform and standard security interface to access different resources on the network.
- We think that also “production grid”, with “static” VO’s will benefit from this, and that these results should be valuable also for commercial applications (e.g. *authenticated p2p network*).
- We will ensure that a consistent and adequate policy language will be defined, in the context of GGF, and will use the GGF-endorsed policy language as a basis for our design.

B.6.6 Deliverables and Milestones

The deliverables of the project are in common for the five workpackages and are listed below. With the architecture deliverables some keywords are listed as to what the deliverable will contain for each workpackage. A two year project is assumed. The dates could differ slightly depending on the starting date of the project as we have to make sure to be able to attend the Global Grid Forum meetings and to have the material for the working groups for GGF. Moreover the GGF meetings will be used for additional meetings with the members of the collaboration.
The three development cycles are reflected in the list of Milestones and Deliverables. Each cycle starts with a retreat of all project members followed by a period of detailed design within each WP. The output of this period is an architectural design deliverable. Then starts a period of implementation and the result is a demonstrator deliverable. When this cycle is finished a period of software hardening and evaluation will follow but at the same time a new development cycle will start again with a retreat.

We plan to have two internal reviews. The first one after the first architecture to get feedback from other specialists in the field. The second after the final architecture is finished, in time to make adjustments to the final deliverable and to help us develop plans for future directions. We intend to invite two or three security specialists to be prepared to receive documentation and to come and spend a day of presentations, demonstrations and intensive discussions with us. We will select those reviewers among security experts in the important grid initiatives at that time.

On purpose there has been left some slag at the end of the project to allow for comprehensive dissemination of the results of the project. The time will be used also for final hardening of the code as it will be available from a public repository after the project.