

Virtual Organization Trustworthiness in the Grid World

M. Altunay, I. Gaines, D Petravick, I. Sfiligoi

Fermilab

gaines@fnal.gov

Abstract. Computing in High Energy Physics and other sciences is quickly moving toward the Grid paradigm, with resources being distributed over hundreds of independent pools scattered over the five continents. The transition from a tightly controlled, centralized computing paradigm to a shared, widely distributed model, while bringing many benefits, has also introduced new problems, a major one being the handling of trust between participating parties.

The trust problem has been recognized since the beginning of the Grid movement, and a lot of thought has been put into developing the infrastructure for handling trust between resource providers and users. In particular, recognizing the size of the problem, the trust handling has been split into two pieces;

- a) between final users and Virtual Organizations (VOs), and
- b) between VOs and resource providers.

However, the above mentioned split has only been tackling the scalability issue, and very little thought has gone into understanding the trust relationship problems that a VO itself introduces. In particular, most VOs run dozens of services, many of them handling user binaries and user credentials.

Such services are obviously critical both for the final users as well as for the security health of the whole Grid; a compromised service could easily generate a major security incident. In spite of this, there is very little, if any, formal process in place to maintain the necessary level of trust.

This presentation will give an introduction to the problem of VO trust as well as an overview of the possible solutions.

1. Introduction

One of the key characteristics of grid computing environments is that there is no hierarchical management relationship between the various entities participating in the grid systems. Traditional centrally managed computing enterprises (even widely distributed systems) are characterized by brick and mortar organizations, well defined hierarchical relationships where everyone clearly understand their role and responsibility, ample face to face interaction, and primarily vertical information flow. Grid environments, on the other hand, are characterized by virtual rather than physical organizations, flat organizations based on willing collaboration rather than following orders, limited face to face

interaction, and mainly horizontal information flow. As the grid paradigm for computing becomes increasingly important in obtaining large amounts of computing resources, particularly in global scientific environments, the problems introduced by this new style of organization must be addressed.

Traditional environments can rely on certain actions being performed because of a shared management hierarchy. The hierarchy, by defining organizationally the roles and the interactions among the roles, explicitly represents the trust relationships within the organization. It clearly communicates how one role member is expected to perform and the consequences of any deviant performance, and therefore lays the ground work for all role members to trust in the behavior of one another. Instead, in grid environments, the various parties must rely on each other to act in certain ways based on cooperative and collaborative agreements. We call such relationships "trust relationships", and the grid environments will not operate effectively without a variety of such trust relationships between parties.

Although this obvious fact has been long recognized, in most cases there are no formal mechanisms for establishing and managing such trust relationships. Instead, the parties make use of a variety of informal and ad hoc mechanisms of agreements, memoranda of understanding, joint policy documents, proper use statements, and the like, to provide the basis for the trust which is required to operate these distributed grid environments.

We will examine here how such trust relationships can be more formally defined and managed, and will look at some examples of how trust is managed today in existing operating grid systems and how this trust management can be improved, with the goal of replacing the current ad-hoc trust process that relies on a manual process, often involving individuals who must be known to one another, with automated processes that can build, grow, and monitor trust relationships through their full life cycle.

2. Formal Trust Relationships

Fundamentally, trust is a relationship between two (or more) parties where certain parties can rely on other parties to act as they have "promised", or in other words, to perform certain specific actions under certain conditions. Each such relationship is characterized by answers to the questions who, what, why and when:

- who: which entity will perform actions (trusted party), and for which other entities will they do this (trusting party)
- what: what are the specific details of the actions to be performed
- why: what are the specific conditions or precursors that will cause the actions to be performed
- when: what are the start and end dates during which the actions will be performed

For example, a VO (who) trusts that a computing resource provider (who) will run a computing job (what) upon submission of the job through the agreed upon grid interface (why) for the members of a given virtual organization (who) during a specific calendar year (when).

The essence of any trust relationship is that a trusting party does not have to check in each instance that the trusted party actually has performed the desired actions. Behaving otherwise simply shows that the trusting party in fact does not trust the other party. On the other hand, trust is a heuristic process, where the level of trust grows based on past experiences between the parties, and where how much the trusted party kept their "promise" in the past increases the likelihood that they will do that in future again, thus earning recognition as being trustworthy. There thus need to be automated mechanisms providing assessments of the performance of the desired actions, providing parties to the relationship with a level of assurance that the relationship is working properly. The details of the trust relationship need to be appropriately published so that the relevant parties are aware of the details.

Other aspects of the relationship are described below when we consider the full trust relationship life cycle..

2.1. Formalizing trust

To deal with these issues, we propose formalizing the trust process. Instead of a collaborative relationship being governed only by a detailed agreement or memorandum of understanding listing large numbers of actions the parties will perform, we suggest breaking down the relationship into a series of atomic trust relationships, each specifying a unidirectional agreement where one party agrees to perform a single specific action under appropriate circumstances. Then each of these relationships can be clearly monitored and evaluated for levels of assurance, while it is very difficult to measure the level of assurance for a complicated multi component agreement.

Thus, each separate trust relationship describes A's trust in B to perform a specific action, under specific conditions, during a specific time period. The relationship is unidirectional and non reflexive (A trusts B but B may or may not have other relationship where B trusts A). Under usual circumstances, however, the trust relationship can be transitive (A trusts B and B trusts C, then A can trusts C). This transitive feature is an important aspect to allow trust to scale to larger collaborations. Monitoring and evaluation of the performance of the actions measures a level of assurance LoA, where LoA \in {low, high, medium}.

Such formal mechanisms will enhance collaboration by making it clear in advance precisely what each party in the collaboration can expect from others. Standardization of the process will make it easier and faster to establish trust, and monitoring of the relationships will demonstrate on an ongoing basis that the trust relationship is performing as agreed.

2.2. Trust relationship life cycle

Rather than inventing procedures from scratch each time a trust relationship is established, there should be standard mechanisms for handling each of the stages in the life cycle of such a relationship. These stages include:

- Establishing the relationship (naming the parties, the actions, the conditions and the time period)
- Publishing details of the relationship
- Monitoring performance
- Measuring and reporting levels of assurance
- Terminating the relationship under normal conditions (time specified for the relationship expires)
- Renewing a relationship for a further period
- Revoking a relationship when it fails to met the necessary levels of assurance
- Restoring a broken relationship when conditions are met to reestablish trust

In particular, these mechanisms should make use of standard mechanisms for establishing identity so they can be used even by entities who have not met face to face.

3. Virtual Organizations and Transitive Trust Relationships

It would be highly inefficient if every party in a grid system must make bilateral agreements with every other party. Instead we use the idea of VOs (virtual organizations) as intermediaries to greatly reduce the number and complexity of trust relationships. The VO will form bilateral trust agreements with each of the service providers that is providing computing resources; and the VO will form a trust

relationship with each grid computing user whom it is willing to vouch for. This then obviates the necessity for each user to form a trust relationship with each service provider.

This concept can be extended by taking advantage of federations or collaborations of grid entities, such as the Open Science Grid, who can form trust relationship with particular sites and with particular VOs, thus avoiding the necessity for each VO to establish a trust relationship with each site. The established web of trust allows the grid to be scaled to worldwide collaborations and greatly broadens the resources made available to any individual user.

Further, combining the formal mechanisms suggested above with the use of the virtual organizations as the parties to the trust relationships addresses the security issues described in the abstract. Trust relationships are cycled between users and VOs and between VOs and Sites. Each separate action to be taken by a party is precisely defined by a separate trust agreement and monitored for compliance. For example, a VO can observe the Level of Assurance that a Site adhered to its trust relationship during a job submission. Likewise, a user can observe how a VO adhered to its promise to provide necessary services and respect the user's personal data collected at registration. The ongoing processes will allow the parties to detect problems in their implementation of trusted actions and correct deficiencies long before a major security incident reveals any longstanding problems.

4. Virtual Organization Trust in the Open Science Grid

The Open Science Grid is a confederation that brings together computing and storage resources from campuses and research communities into a common, shared grid infrastructure over research networks via a common set of middleware. They have been operating large scale production grids for several years.

The Open Science Grid currently manages trust through a series of:

- site agreements
- service agreements
- VO agreements
- user agreements

which are maintained on web pages which allows entities to register as a member of OSG and agree to various policy statements which describe the trust relationships. In general these policy documents and agreements each contain a large body of separate actions that the parties are "trusted" to perform, and lack mechanisms for monitoring the performance of the trusted actions.

While today these relationships are managed in a manual and ad hoc basis, as the OSG develops its security plans and procedures and scales to greater size, there is an increasing need to formalize and individually manage the specific details of the trust relationships. Procedures are under development to, for example, test out the response of grid parties to a simulated security incident to determine if they can in fact be "trusted" to carry out their appropriate roles in such as occurrence. The successful conclusion of such tests will rely on first specifying the precise details of what each party is required to do under such circumstances, leading to just the sort of formalized atomic trust relationships we have been discussing. We expect that OSG will provide a fruitful laboratory to evaluate these ideas.

5. Conclusions

The seamless and successful collaboration between a Site and a VO is the holy grail of the grid computing. We think that a collaboration is essentially a trust relationship. The requirements to collaborate, such as defining the duration, the activities involved, the resources devoted, the

permissions to act on, termination, and etc, in fact correspond to the different states of a trust life-cycle. In other words, establishing collaboration between a site and a VO is identical to establishing a trust relationship between them. Currently, this trust relationship is cultivated manually and out-of-band. However, this is not practical, prone to errors and time-consuming, not to mention the confusing disputes born out of not having formalized processes. We represent this as the missing link in the current grids due to our lack of understanding. An ideal solution is to enable Sites and VOs with processes that can handle the states of trust life-cycle. Without any third party's help, Sites and VOs would automatically invoke the processes to go through the states of the trust life-cycle. As a result, VOs and Sites can reach faster, more formal and less ambiguous collaborations. Moreover, by monitoring state, the collaborating parties can collect data for future collaborations, and complete the whole cycle by restating new collaborations with the old partners.