Proceedings of the 5th WSEAS Int. Conf. on SIMULATION, MODELING AND OPTIMIZATION, Corfu, Greece, August 17-19, 2005 (pp72-76)

Analysis of Grid Security Metadata Based on Knowledge Mining

Haidong Xiao, Jianhua Li
School of Electronic,Information and Electrical Engineering
Shanghai Jiaotong University
Shanghai
CHINA

Abstract: Grid security research is a hotspot in grid computing. Grid security application is weak in interoperability in common Single Sign-on. Interoperability with grid security includes performance, scalability, and standardization is focused first. A new Grid security architecture based on knowledge base is formed, idiographic analysis of grid security policy metadata reflexion combined with security knowledge base is done in addition, at last Cauchy sequence of knowledge space reflects to parameter space is constructed, it gives testify that knowledge of security policy knowledge base is unique. Knowledge base security architecture and grid security meaning space provide a promising approach to make grid security technologies and solutions smarter, more flexible, scaleable and adaptable.

Key Words: Single Sign-On , Grid Security, Knowledge, Cauchy sequence

1 Introduction

The GSI provides a delegation capability: an extension of the standard SSL protocol which reduces the number of times the user must enter his pass phrase. If a Grid computation requires that several Grid resources be used (each requiring mutual authentication), or if there is a need to have agents (local or remote) requesting services on behalf of a user, the need to re-enter the user's pass phrase can be avoided by creating a proxy.

A proxy consists of a new certificate and a private key. The key pair that is used for the proxy, i.e. the public key embedded in the certificate and the private key, may either be regenerated for each proxy or obtained by other means. The new certificate contains the owner's identity, modified slightly to indicate that it is a proxy. The new certificate is signed by the owner, rather than a CA. (See diagram below.) The certificate also includes a time notation after which the proxy should no longer be accepted by others. Proxies have limited lifetimes.

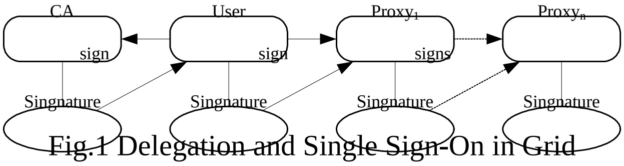

Fig.1 Delegation and Single Sign-On in Grid security application

The proxy's private key must be kept secure, but because the proxy isn't valid for very long, it doesn't have to keep quite as secure as the owner's private key. It is thus possible to store the proxy's private key in a local storage system without being encrypted, as long as the permissions on the file prevent anyone else from looking at them easily. Once a proxy is created and stored, the user can use the proxy certificate and private key for mutual authentication without entering a password.

When proxies are used, the mutual authentication process differs slightly. The remote party receives not only the proxy's certificate (signed by the owner), but also the owner's certificate. During mutual authentication, the owner's public key (obtained from her certificate) is used to validate the signature on the proxy certificate. The CA's public key is then used to validate the signature on the owner's certificate. This establishes a chain of trust from the CA to the proxy through the owner. In this model, The inherent scale, heterogeneity, dynamism, and no determinism of Grids and Grid applications have resulted in complexities that are quickly breaking current paradigms, making both the infrastructure and the applications brittle and insecure. Clearly, there is a need for a fundamental change in how Grids and Grid applications are developed and managed. This is leading researchers to consider alternative paradigms that are based on the strategies used by systems in nature to deal with complexity, dynamism, heterogeneity, and uncertainty. This emerging vision aims at realizing computing systems and applications capable of configuring, managing, interacting, optimizing, securing, and healing themselves with minimum human intervention, and has lead to a number of recent research initiatives such as Autonomic Grids, Cognitive Grids, and Semantic Grids[1].

Grid security authorized layer is a good solution defined below to apply grid security policy; the security strategy is based on security knowledge, and performs the real logic functions to ensure the chain from the CA to the proxy through the owner is available.

2 Grid security authorization police knowledge mining and Grid security model

The Grid by its nature involves access to proxies and data outside one's own company or institution. Security is therefore a major element in any Grid infrastructure, as it is necessary to ensure that only authorized access is permitted[2]. The common Grid security model is as follows:

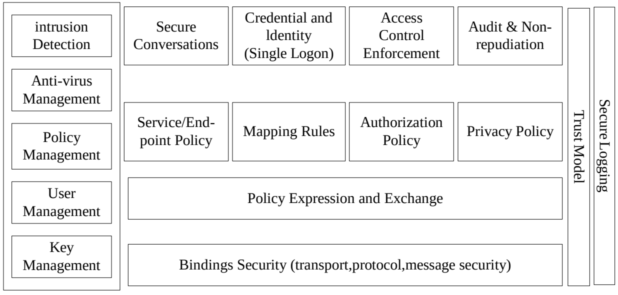

Fig. 2 Common Grid Security Model

And this common grid security model associated with the Globus toolkit is supported by a grid Security Infrastructure (GSI) based on a Public Key Infrastructure where users authenticate to the grid using X509 certificates.[3]

Here, the knowledge base is introduced to give a grid security model based on it. In order to assure the grid security, a knowledge base should be constructed to refine the data in the attacks database, as we know, when a common computing package arrived the grid node, some actors of it will be traced, from the common grid security model above, secure logging will trace every computing package, and the useful attacks information will be written to the security database. The purpose is to construct the grid security knowledge base as follows.

3 Security information metadata description of grid node

Security, encrypt is vital to grid, any grid infrastructure must consolidate the grid resource under the protection of grid security system, to creates a huge resource united database which is labeled, digital, and protected by standard security infrastructure, to supports the varies of actual grid application, and those security application is a part of the whole grid construction.

In order to construct the security knowledge base accurately, metadata is used to describe the security information on the grid node, and then comes analysis.

First, two kind of metadata should be defined: one is metadata which is defined to describe the parameter space, the other is metadata which is used to describe the security knowledge space. The reflexion relationship between these two kind metadata constructs the pattern of security knowledge mining.

When grid node is attracted by illegal computing packages from outside, the metadata of original parameter space will be changed, include property insert, delete or update actions, all these alteration is dynamically occurred in parameter space. In the gird, security knowledge information is described by metadata of security knowledge space, and metadata of knowledge space is a dynamically mapping of original parameter space. In the real grid computing circumstances, this process is the dynamically mapping between the virtual resource layer and physical resource layer. The process masks the complexity of bottom physical resource layer, and refines the security knowledge metadata indirectly through this mapping. And it is vital to construct the security knowledge base.

According to the HDM(hyper graph based data model, HDM)[4],mapping the metadata first.

n Parameters space information resource group $(x_1, x_2 ... x_n)$ and knowledge space metadata $(y_1, y_2 ... y_m)$ are known, the measurement space $C[a,b]$ is constructed with $\rho(x,y) = \max_{\rho \in [a,b]} [x(t) - y(t)]$;

The following proving process shows $x_n \in C[a,b] \quad (n = 1,2...)$ is a Cauchy sequence,

Namely: $\forall \varepsilon > 0 \quad \exists N, \forall \ m,n > N$ there is $\rho(x_n, x_m) < \varepsilon$

$$\therefore \quad \max_{t \in [a,b]} |x_n(t) - x_m(t)| < \varepsilon \qquad (1)$$

$\therefore \quad \forall t \in [a,b]$ there is $|x_n(t) - x_m(t)| < \varepsilon$

fixed $t_0 \in [a,b]$ because of $|x_n(t_0) - x_m(t_0)| < \varepsilon$

so this group sequence is Chauchy sequence

Suppose that $\lim_{n \to \infty} x_n(t_0) = x(t_0)$

let $m \to \infty$ there is $|x_n(t_0) - x_m(t_0)| < \varepsilon$, because t_0 is random picked from $[a,b]$,

so $\max\limits_{t\in[a,b]} |x_n(t) - x_m(t)| < \varepsilon$, $\rho(x_n, x_m) < \varepsilon$ 。

To expand the knowledge space metadata $(y_1, y_2 ... y_m)$ to R^m ,

we will get in the same way

$$\rho(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_m - y_m)^2}$$

$$= (\sum_{i=1}^{m} |x_i - y_i|^2)^{\frac{1}{2}} \qquad (2)$$

And security metadata knowledge space is a maturity measurement space.

Analysis to the unique attribute of security metadata knowledge space will be followed.

4 Grid security architecture based on knowledge base

The grid security knowledge with the arithmetic support that is mentioned below will be very useful in Single Sign-On Process. We can construct the grid security architecture now. To see the figure below:

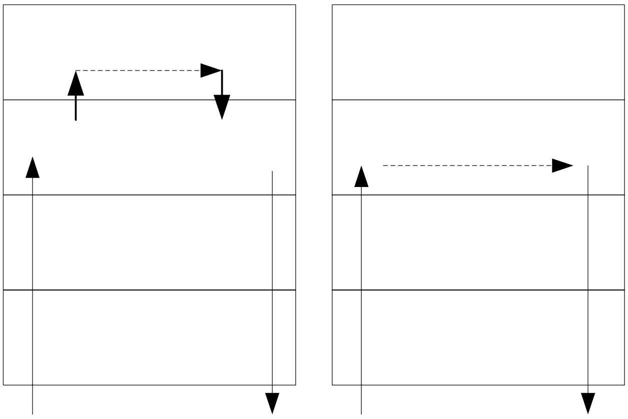

Fig.3 Grid security architecture based on knowledge base

From the figure above, we can learn that all the grid application based on the three layers under it, and Grid security authorized layer is the key layer between the grid application layer and grid security knowledge database. In this layer, all the grid security logic will be described and designed. And all the grid security logic selection based on the rules in the grid security knowledge base. In this way, we can use the knowledge grid's advantages to build a more efficient and effective intelligent application platform to solve the grid security problem we faced.

The rule of knowledge mining is often described as follow modality:

< Rule >:= (<rule ID>, <precondition1>| <precondition2> …<conclusion>, <rule intension CF>), as the legal request is sent to the grid node and ask for the grid application service, the request will pass through the authorization of grid security authorized layer because it is a legal request. From the <rule> function above, the return value is true, and we can index its rule ID from the knowledge base to confirm the requisition is the legal one.

But in the contrast way, the illegal or unauthorized requisitions will be banned by the grid security authorized layer and never brother the grid application above it.

In this architecture, the grid security authorized layer act as a grid security guard role. It filters the unauthorized grid computing requisitions or hazard computing packages and ensures the whole grid system running in efficient and effective level.

In the follow steps, we will construct this grid security authorized layer because the implementation of the policy management and evaluation are abstracted into an authorization service which may be an internal implementation or an external authorization service, as can be seen in Figure 2. The authorization service provides the following features for policy management and evaluation[5].

Step 1, a method to dynamically add a credential or list of credentials that can be trusted for a particular action.

Step 2, a method to dynamically delete a credential that is not allowed to perform an action.

Step 3, the ability to verify that the presented credential allows the user to access the requested resource.

The authorization layer need not evaluate the policy, but just needs to know which service to contact. The internal structure of the authorization layer is separated from the details of the implementation of the Authorization service. The authorization sub-layer in the framework contacts the authorization service to perform the policy check during the request-response protocol. The advantages of the model are:

First, the authorization model is a plugin module can be e can be easy put into existing applications thus allowing for the flexibility of having secure and un-secure applications

Second, the authorization service acts as a bridge between the diverse security subsystems that may exist on the Grid.

Third, the intricacies of the policy representation, management and evaluation are abstracted away from the application.

And the last, the model provides the user the flexibility to choose between any pre-existing security mechanisms or any application-specific mechanisms. For example if an application needs to restrict or grant access at the parameter level of a method, the authorization service module could be customized for the same.

5 Unique attribute of knowledge in security policy knowledge space

Semitic security layer includes policy knowledge space and parameter space, the first provides varies of security services to users according to the security policy, the latter describe the user's knowledge in a form that computer can understand. In the security policy knowledge space user and service requirement can be presented by special role, in this space, service is autonomy and self-denoted.

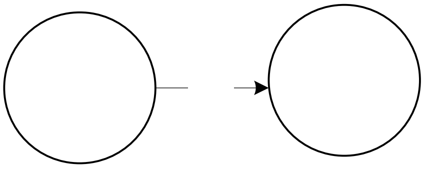

Fig. 4 the relationship of reflex from parameter space metadata to security policy knowledge space

The property and behavior of security policy knowledge space is competent for the role for it is the reflexion of parameter space metadata. R^m is constructed by security policy knowledge space metadata$(y_1, y_2...y_m)$,according to the property of maturity and Chauchy sequence, we can get $\rho(x,y) \to 0$,in the other word, $m \to \infty$ with the reflex sample metadata accumulate. Radius of high dimension security policy knowledge space goes to 0, and the knowledge of security policy knowledge space goes to unique, this is the procession of security knowledge mining while the security knowledge refined procession which from uncertain to certain. It accord with the define of knowledge itself, because the procession of knowledge abstract is a procession from uncertain to certain.

6 To improve the cooperation with grid security solutions based on security meaning space analysis

Security meaning space is a special space to support the grid security knowledge mining, grid user authorized by different CA center has his own security requirement, security role is often distributed to these users, even the same grid user has more than one different security role according to the different security requirements.

Grid security meaning space is a important part in grid security authorized layer, all the security knowledge police based on the child space of security knowledge, as figure shows below:

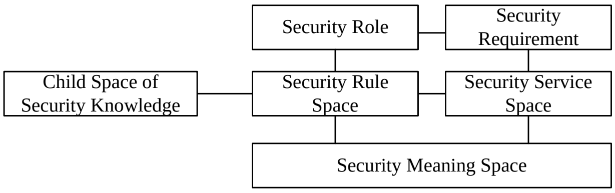

Fig.5 Construction of Grid Security Meaning Space

KGOL Grammar Specifications is available to improve the cooperation with grid security solutions. The syntax and semantics of the KGOL are designed by referring to that of the SQL. For example, the statement of getting knowledge resources from a Knowledge Grid (KG) at the Concept level and the Software category can be expressed as follows:

GET K from KG
WHERE level="Concept" and category="Security".

A major form of the KGOL statement is described as follows:

<KGOL Statement> ∷ =<Operator>[ALL | DISTINCT][Resource]
[FROM|INTO] <Knowledge Grid Name>[AT UGL]
[WHERE<Condition-Expression>]
[SORT BY <Knowledge Grid Element>][ASC | DESC];
<Operator>∷= Get | Put | Delete | Undelete;
<Condition-Expression> ∷ = Level-coordinate-expression
|Category-coordinate-expression
|Keywords-expression | Name-expression.

The Condition-Expression represents a Boolean expression specifying the level coordinate and the category coordinate of the Knowledge Grid as well as the other constraints such as keyword patterns. The KGOL also has the assign-statement, the create-statement, the browse-statement, the update-statement, the join-statement, the open-statement, and the log-statement, etc.

7 Conclusion

We have mentioned delegation and single Sign-On in grid security application in grid node. And refer to grid security model and security authorization police knowledge mining. In this architecture, we can get the security characteristics dynamically, classify them and decide the most suitable security scheme to the node resource of the grid system.

The judgment and mark of invalid computation packets offers the important basis for whole grid system to collect rubbish and detect task stop. If a grid node that is attacked, or a threaten computation packet coming, these would influence the computations security knowledge, so the knowledge of the computation packet changed. We can judge this by index it in the security knowledge base, if its character matches to the invalid computing package, we decide the computation packet is invalid, and then eliminate it at once.

With the development of new grid security technology, grid application and grid security infrastructure problems should be solved. The technology includes grid application protection and analysis of commutation among grid nodes. The grid security technology which combines the grid security and knowledge mining occurs when knowledge theory goes forward. According to this thesis, radius of high dimension security policy knowledge space goes to 0, and the knowledge of security policy knowledge space goes to unique, this is the procession of security knowledge mining while the security knowledge refined procession which from uncertain to certain. It accord with the defining of knowledge itself, because the procession of knowledge abstract is a procession from uncertain to certain according to knowledge theory.

Specifically, this grid security architecture that supports high-performance, latency tolerant and heterogeneous components to promote growth of grid security technology into grid computing environments and unity within the grid security and knowledge base science communities.

References:

[1] I. Foster, C. Kesselman, and S. Tuecke, “The nexus task-parallel runtime system,” in Proc. *1st Int. Workshop Parallel Processing*, 1994, pp. 457–462.

[2] Bajkumar Buyya. *High Performance Cluster Computations Architectures and Systems*, Volume 1 [M]. Beijing: Electronic Industry Press. 2001,pp.527-529.

[3] Mo Zeyao, Li Xiaomei. Parallel Multigrid Computations for Anisorropic Diffusion Problems. *Numerical Computation And Computer Application*, 1997.9(3):pp.218-229.

[4] P Mcbrien, A Poilovassilis. Schema evolution in heterogeneous database architectures, A schema transformation approach. In: *Proc of the 14th Int’l Conf on Advanced Information Systems Engineering*. Berlin: Springer, 2002. 484-499.

[5] Xiao Haidong, Li Jianhua Analysis and Research of Grid Security Based on Knowledge Base [M] .Guangzhou: *CIHW*, 2004. pp.177-179.