ACCESS ISOLATION MECHANISM BASED ON VIRTUAL CONNECTION MANAGEMENT IN CLOUD SYSTEMS
How to Secure Cloud System using High Perfomance Virtual Firewalls

Alexey Lukashin, Vladimir Zaborovsky and Sergey Kupreenko
Saint-Peterburg State Polytechnical University, Polytechnicheskaya street 29, Saint-Petersburg, Russia

Keywords: Security, Firewall, Cloud system, Virtualization, Virtual connection.

Abstract: The paper describes the access isolation model based on virtual connection management and proposes the mechanism of traffic filtering in transparent mode, invisible to other components. New level of complexity of information security tasks was observed in the distributed virtualized systems. The paper proposes a specialized firewall solution for implementing access isolation and information security in hypervisors and entire distributed cloud system.

1 INTRODUCTION

Virtualization of distributed computational resources and development of heterogeneous virtual machine environments are a very popular and fast growing area in information technologies. Various components belonging to this area are usually denoted by the term “cloud computing”. A lot of service providers offer such solutions, from IaaS to SaaS layers. There are not only public cloud providers such as Amazon, Google, and others. Private clouds are becoming very popular and there are a lot of solutions based on Eucalyptus, Open Nebula, VmWare, or Microsoft technologies. Therefore, ensuring information security of cloud systems is a vital problem (Cloud Security Alliance, 2010). The present paper introduces a virtual connection as an emergence essence and describes the access isolation in virtual networks based on the virtual connection management. The network traffic is described as an aggregation of virtual connections. The distributed virtual environment (cloud system) provides heterogeneous computing resources; therefore, it would be reasonable to use these resources to protect the information security of the system. Virtual connections function separately of each other and do not have any shared resources, so it is possible to establish parallel traffic filtering within the security domain. This security domain would exist in the hypervisor, would use the amount of resources (cores, memory) that is required for current information security tasks, and would be scalable and adaptable to different situations. The paper highlights the promising (perspective) approaches to information protection in the distributed computing environments. These approaches use the high-performance virtual firewalls, that operate in stealth mode in the virtualization nodes of the computing environment and provide consistency of security policies through a centralized management. Applying the methods of formalizing security to automate the generation of filtering rules in combination with hardware and software platforms based on multicore microprocessors can deliver high performance firewall. This firewall implements the filtering functions in the operating system kernel based on the application network management models in the Netgraph subsystem.

2 INFORMATION SECURITY IN CLOUD SYSTEMS

Today, many companies, including leading universities and government institutions, are transferring their computing resources to the virtual infrastructures, using both open systems (Eucalyptus, Open Nebula) and commercial solutions (VmWare, Citrix, IBM). Due to this trend, the information security of cloud systems becomes an acute problem. The major differences between

Lukashin A., Zaborovsky V. and Kupreenko S..
ACCESS ISOLATION MECHANISM BASED ON VIRTUAL CONNECTION MANAGEMENT IN CLOUD SYSTEMS - How to Secure Cloud System using High Perfomance Virtual Firewalls.
DOI: 10.5220/0003433803710375
In *Proceedings of the 13th International Conference on Enterprise Information Systems* (ICEIS-2011), pages 371-375
ISBN: 978-989-8425-55-3

the cloud systems and the distributed networks are the following:

- Information processing takes place on the virtual machines under full hypervisor's control; the hypervisor has access to all data processed by its virtual machines;
- Cloud software controls the resource planning and provision; it is a new entity in the information environment which has to be protected from the information security threats;
- Traditional information security components such as hardware firewalls cannot control the internal virtual traffic between virtual machines in one hypervisor;
- In virtualized environments, files serve as virtual storage devices; these files are located in the network storages and are more exposed to threats than hard disks;
- Transfer of instance memory occurs when migrating virtual machines between hypervisors; this memory may contain confidential information.

Therefore, due to the above-listed specifics, new information security threats appear, including:

- Attacks against the virtual machines management tools, controllers of the computing environment (cloud controller), or cluster and data storage, where the virtual machine images and user data are located;
- Unauthorized access to the virtualization node;
- Using virtual network for data transfer not allowed by the information security policy.

The major specifics of the virtual infrastructure is that an attack or an attempt of unauthorized access can come from the virtual network, where such devices as switches, hardware firewalls, and physical connections are absent. This specifics hampers applying the exiting methods and tools for ensuring information security in computer networks and GRID systems to the information security protection of cloud systems.

The distributed and virtual computing environments do not have effective methods of information security protection. One of the problems is the lack of firewalls, which can operate in virtual environment as efficiently as the existing on the market software and hardware solutions for protection of information resources and reflection of cyber attacks. For a number of cloud solutions, for example, free and open source cloud environment Eucalyptus, based on hypervisors XEN or KVM, there are no efficient solutions for the virtual machines' protection, despite of the rapidly growing popularity of this environment due to its compatibility with the interfaces of Amazon (Amazon EC2, Amazon S3) products.

3 VIRTUAL CONNECTION MANAGEMENT IN THE ACCESS CONTROL TASKS

Virtual connection (VC) is a logically ordered exchange of messages between the network nodes. (Silinenko, 2009). Computer network is a set of virtual connections. Virtual connections are classified as technological virtual connections (TVC) and information virtual connections (IVC). (Figure 1).

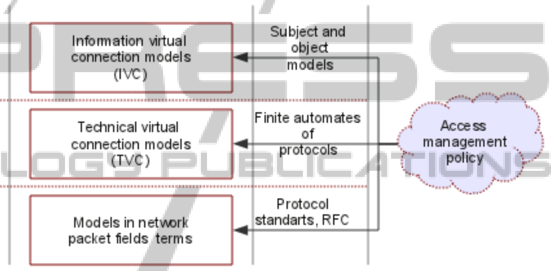
Information virtual connection models (IVC)
Subject and object models
Technical virtual connection models (TVC)
Finite automates of protocols
Access management policy
Models in network packet fields terms
Protocol standarts, RFC

Figure 1: Layers of access control policies.

To implement the policy of access control, the filtering rules are decomposed in the form of TVC and the IVC. These filtering rules can be configured for different levels of data flows description based on the network packet fields on the levels of channel, transport, and application protocols. In terms of the access control, the TVC model can be defined as a stream of packets generated by the network applications during communication. The TVC model is presented in the form of potentially countable subset of the Cartesian product set of packets P and timestamp T (1).

$$TVC=\{p_{ti}\},i=\overline{1,N},N\in[1,\infty)\subset P\times T \qquad (1)$$

This model is characterized by a finite set of parameters that describes the access subject and the access object, as well as action between them in the form of packet stream within the interconnection. The model parameters are the identifiers of the subject and the object, such as addresses, ports, and other characteristics of network protocols. For efficient traffic classification, the IVC model is used along with the TVC model. The IVC model describes the interaction between the access object and the access subject at the application services level. The IVC model is a set of technical virtual connections (TVC); the number and characteristics of these TVCs are determined by the Cartesian product of the information interaction access model (IIM), the access subject model (IMS), and the access object model(IMO) (2).

$$IVC = \{TVC_i\}, i = \overline{1, N} \subset (IIM \times IMS \times IMO) \qquad (2)$$

This formalization allows representing the access IIM as a finite subset. The size of this subset is determined based on the description of interconnection subjects permitted within the given access control policy. IMO is characterized by a finite subset of information and network resources, the access to which is This formalization allows representing the access IIM as a finite subset. The size of this subset is determined based on the description of interconnection subjects permitted within the given access control policy. IMO is characterized by a finite subset of information and network resources, the access to which is permitted in accordance with the access control policies. IMS describes the operations performed by the access subject within the bounds of IMO. In accordance with the access control policies, IMS describes the operations performed by the access subject within the bounds of IMO.

4 PARALLEL PROCESSING OF VIRTUAL CONNECTIONS

Virtual connection (VC), as some abstract, exists in parallel to and independently from other virtual connections. Virtual connections do not share any resources, which allows parallel processing of virtual connections. (Zaborovsky, Lukashin, Kupreenko, 2010). The suggested approach to the network traffic filtering is based on the concept of a virtual connection and allows extracting the connection context. The connection context can be represented as a vector Y_i, of parameters, for example, source and destination addresses, port, connection status (for TCP protocol), etc. Controlling the virtual connection is calculating the indicator function F, which requires resources such as computing processors and operating memory (3).

$$F(Y_i\} = \{1,0,*\} \qquad (3)$$

The indicator function F takes the following values: 1 – if VC is allowed, 0 – if VC is forbidden, * – if at the current moment it is impossible to clearly determine whether connection is prohibited or not, the decision is postponed and VC is temporarily allowed.

Computing problems could be divided into two groups:

1. Stream-related tasks that can be calculated with SIMD processing elements (for example, using graphic processors and CUDA technology).
2. Computational problems solved on the standard multicore computers MIMD.

Because the distributed environment is heterogeneous with respect to the available processing elements, both the streaming SIMD processors and the classic MIMD multicore processors can be used for the firewall tasks in the cloud systems. Firewalls that protect the hypervisor operate in the virtualized environment; thus, the configuration (computing cores, memory, streaming processing elements) of the protection device can be changed depending on the loading options, access policies, and the amount of available resources.

Calculation of the indicator function F can be decomposed into multiple computing processes – $\{F_i\}$. In this case, the problem of VC control can be described using the graph G(Q,X), which is called the VC control information graph (you can find the detailed description of stream tasks by graph in (Kaliaev, Levin, Semernikov and Shmoylov, 2008)). The VC control information graph consists of the set of nodes; each of these nodes is attributed with the operation F_i . If two nodes q_i and q_{i+1} are connected with an arc, then result of operation F_i is the input for the operation F_{i+1}. Each node has an arc, which corresponds to the case when $F_i = 0$. Then VC is considered as prohibited and no further analysis is performed.

The multiprocessor computing system that solves the firewall problems can be presented as a full mesh computation system graph with MIMD and stream computers as its nodes. This graph is a full mesh, because the communications between CPUs are provided by hardware and operating system, and there is no predefined path between the cores, data can pass directly from one node to another. Usually the computation system graph and the control information graph do not match each other, because of amount of computing resources is limited and is less than the amount of computational processes.

We can split the VC control graph in N non-crossing subgraphs and, thus, build a VC operating pipeline. Because the virtual connections exist separately from each other, we can process them in parallel. With the C compute nodes of MIMD type, the operating time of VC processing would be limited by (4).

$$T_{VC} \le \frac{\max(z(f_i)) \cdot \max(\tau_j)}{C} \qquad (4)$$

z(f$_i$) – number of CPU clocks, required for calculation, τ_i – real time of CPU clock.

The inequality appears in the given formula because the decision on the VC classification (allowed/forbidden) can be made before the passing all nodes of the graph.

Due to the heterogeneity and reconfigurability of the computing environments, in some cases the configuration of the firewall can be adapted to the access control tasks being solved at the current moment of time. This can be achieved by using the graph models for network traffic processing and Netgraph (Cobbs, 2003) technology. This technology allows organizing the network traffic processing in the context of the operating system kernel (Zaborovsky, Lukashin, Kupreenko, 2010).

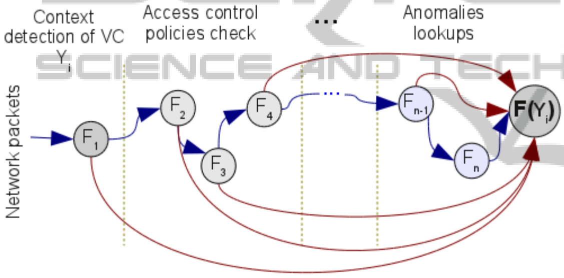

Figure 2: Information graph of the virtual connection management.

Figure 2 shows an example of the virtual connections information control graph with decomposition of the indicator control function into components. The presented approach, in the combination with using the virtualization resources technology, allows improving performance of the network traffic monitoring and using only those computing components that are required to solve current access control problems.

5 ARCHITECTURE OF A SECURE CLOUD COMPUTING ENVIRONMENT

A distributed computing environment (cloud system) consists of the following software and hardware components:

- ○ virtualization nodes;
- ○ storage of virtual machines and user data;
- ○ cluster controller; and
- ○ cloud controller.

The distributed computing environment intended for solving scientific and engineering problems is a set of various computing resources such as virtual machines, and has the following features (Lukashin and Roshupkin, 2010):

- ○ The environment is used by a wide range of users, who are solving problems of different classes;
- ○ Virtual machines of different user groups can operate within one hypervisor;
- ○ Wide range of software components (CAD/CAE applications, development tools) and operating systems is used;
- ○ Different hardware configurations are used, including virtual multicore computing machines and virtual machines, that allow performing computations using the streaming technology CUDA.

Virtualization node is powerful multicore system with hypervisor installed, on which the domain level 0 (dom0 in terms of hypervisor XEN or service console in terms of other hypervisors) and virtual computing machines (domain level U, domU) operate.

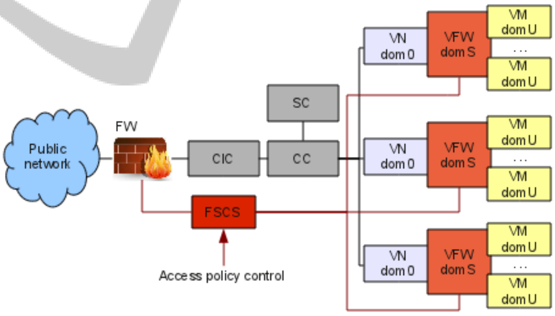

Figure 3: Secure cloud architectrure.

For information security and access control (AC) between the virtual machines that operate under a single hypervisor, the internal ("virtual") traffic and the external traffic (incoming from other hypervisors and from public networks) must be controlled. The solution of the access control problem could be achieved through the integration of a virtual firewall into the hypervisor; this firewall would functions under the hypervisor, but separately from the user virtual machines. The virtual firewall domain can be defined as "security domain" (domS). Invisible traffic filtering is an important aspect of the network monitoring; the firewall must not change the topology of the hypervisor network subsystem. This can be achieved by using "Stealth" (Zaborovsky and Titov, 2009) technology – a packet traffic control

(Figure 3) invisible to other network components**.**

Figure 3 shows the common architecture of a distributed cloud system with integrated AC components. Abbreviations: FW – hardware firewall; VFW – virtual firewall; FSCS – the central control system of all firewalls in the cloud VM – virtual machine; ClC – cloud controller; CC – cluster controller SC – storage controller.

The FSCS distributes the access control policies to all firewalls in the system. When the information security policy changes, new access rules are replicated to all components. The security domain isolates virtual machines from the hypervisor, which prevents the possibility of attack against the hypervisor inside the cloud. The hardware firewall isolates the private cloud components from the external threats.

6 CONCLUSIONS

The presented architecture is a distributed heterogeneous computing environment that provides computing resources of different configurations; this allows arranging the information protection for this system in the form of a dedicated security domain (domS). The security domain can be quickly adapted to the current situation in the cloud. The traffic filtering process in kernel shows good scalability and can be linearly scaled up to eight cores. A secure cloud prototype, based on Eucalyptus and adopted for CAD/CAE computation tasks (Lukashin and Roshupkin, 2010), was created and is currently functioning at the Telematics department of the Saint-Petersburg Polytechnical University.

REFERENCES

Cloud Security Alliance, *Top Threats to Cloud Computing, 2010. URL: http://www.cloudsecurity alliance.org/topthreats/csathreats.v1.0.pdf*

Silinenko A., 2010. *Access control in IP networks based on virtual connection control models: phd thesis 05.13.19: / Saint-Petersburg, Russia.*

Zaborovsky V., Lukashin A., Kupreenko S., 2010. Multicore platform for high performance firewalls. High performance systems // *Materials of VII International conference – Taganrog, Russia.*

Kaliaev A., Levin I., Semernikov E., Shmoylov I., 2008. Reconfigurable multipipe computation systems. – Rostov-na-Donu, Russia.

Cobbs A., 2003. All about Netgraph URL: http://www.daemonnews.org/200003/netgraph.html

Lukashin A., Roshupking I., 2010. Methods and strategies of developing distributed computation systems for CAD/CAE problems solving // *XXXIX week of science SPbSTU, Materials of Russian conference for students, Part XV, p. 13-15. – Saint-Petersburg, Russia.*

Zaborovsky V., Titov A., 2009. Specialized Solutions for Improvement of Firewall Performance and Conformity to Security Policy. *Proceedings of the 2009 International Conference on Security & Management. v. 2. p. 603-608. July 13-16, 2009.*