The Research and Design of Trusted Cloud Computing Platform based on Group theory

Ying Yang,Xuehang Shao
Department of Computer Science and Technology,
ChengDong College Of Northeast Agricultural University
Harbin, 150025,China
teatree2002@hotmail.com

***Abstract*—The cloud computing can greatly reduce the cost of computing, but is unable to ensure either the integrality or the confidentiality of data and calculation. Therefore, this paper considers of the safety of cloud computing, combining the thought of abstract algebra group theory in modern algebra, puts forward TCCPoGT (trusted cloud computing platform based on group theory) that designs in many different respects,such as general structure, public key cryptosystems and node management,etc. The analysis results show that the platform TCCPoGT can insure service security.**

Keywords- cloud computing; safety; trust cloud platform; group theory

I. INTRODUCTION

In recent years, the speedy development of computer software and hardware technology promoted the evolution of the computing model. With the advent of various calculation concept and model to be perfected, for instance, parallel computing, distributed computing, the network computing and utility computing, cloud computing model[1] is proposed in the computer science， industry and academia respectively.

Cloud computing, as a kind of calculation model which is rapidly developing and researching, is seen to be about to triggering a wave of new information age for greatly reducing the IT operating cost, improving the efficiency of management of the resources observably, and quickly meeting the needs of business as well. But, in the meantime, that also causes the uncertainty of security problems on different levels (e. g., the host, the network, and the program or data) , while makes studying on cloud computing security become a research focus recently.

It will lead to losing data and application's control directly that the individual or a company or organization uploads the information or program to the cloud, and because the cloud service provider reveals little about cloud internal information, so users have no knowledge of that the clouds service provider how to process their own data and application. The loss of Control and right-to-know will cause users worrying about that whether or not to use the solution of cloud computing, thus results in appearing a question of trust among users and cloud computing service provider, which even may impede the cloud computing development. In consequence, one of crucial issues of the research content of cloud security that how to protect the data of cloud and the integrity of calculation and confidentiality is urgent to solve. This paper bases on group theory of modern algebra theory and designs a trusted cloud computing platform to solve the trust issues between users and cloud service provider.

II. DESIGN PRINCIPLES AND FRAMEWORK OF TRUST CLOUD COMPUTING PLATFORM

Compared with the traditional computing platforms, cloud computing platform come up with higher challenges to the credibility of computer system. It is cloud computing platform that integrates the system of very large scale computer, and provides application for multiple organization users differently through the Internet. Therefore, cloud computing platform should have the following feaures:

Very large scale. Compared with traditional calculation platform, cloud computing integrates many different computer systems that gather together through the Internet. These independent computer systems may consist of some daily and unreliable components for the consideration of cost. For example, services related to cloud computing provided by Google such as web Search and Gmail usually comprise of large amounts of cheap ordinary performance PCs, storage devices and Internet devices. Therefore, the use of unreliable equipment over scale have been put forward rigorous challenges about usability, fault tolerance and maintainability of cloud computing platform overall.

Single-interface. It is cloud computing platform which generally gives a uniform interface for the user. Through that a large number of legal users access cloud computing platform. So, it is an ordeal to designed security of cloud computing platform.

Cross-organization. Cloud computing platform tend to provide services for different organizations or individuals via rental. These individuals and organizations generally lack of trust between each other, and probably become competitors. So, It requires that cloud computing platform isolates service between each user effectively, thus ensuring data integrity and security. In addition, mutual mistrust between cloud computing platform provider and users may also exist. Therefore, user should be able to make sure that the cloud computing platform provider won't malicious tamper or leak users' critical code and data or confidential commercial information and so on.

Published by Atlantis Press, Paris, France.

To serve as the core. On account of serving as the core, the cloud computing platform must hide platform details such as underlying hardware architectures or the overall number of system's resources available from the user.

It is based on these characteristics which leads to the designing and promotion of credibility of cloud computing platform. On the whole, the credibility designing of cloud computing platform mostly in the below several prominent areas:

Data security and privacy. Data security and privacy is mainly due to two hidden dangers: On the one hand, service provider of cloud computing platform, a research from famous business analytics company Garnter, said that user data which oriented cloud computing platform may be accessed and leaked uncontrollably by service providers; on the other hand ,external security attack, for instance, online pay provider PayMaxx suffer from external security attack resulting in more than 25000 users billing information leaked in 2005 for its system security hole.

Maintainability and manageability. It is also an enormous challenge that how to manage, update and maintain hardware and software of super-sized cloud computing platform. While performing such operation must ensure that it won't affect the normal operation of service provided by cloud computing platform.

Fault tolerance and usability. Cloud computing platform can to be capable of providing correct service constantly, and hence platform must be able to ensure that service won't be affected when software and hardware are in case of failure and maintenance.

So far, the design of cloud computing platform mainly in the following aspects:

A. Programming model

Google proposes a kind of data-oriented parallel programming model MapReduce [2] which renders mass data processing for abstraction, and fall into two stages: Map and Reduce. The Map stage is mainly responsible for mapping the data to a large number of sets like (Key, Values), while Reduce stage is in charge of merging and simplifying that through the key. On the MapReduce Framework, users only need to provide Map and Reduce function, without needing to pay attention to the underlying task scheduling, task parallel, and fault tolerance. At the moment, MapReduce has been widely used in development of various products at Google, including Gmail, web search, etc. Because of the convenience and usability of MapReduce, database related researchers extend that they increase a Merge stage[3] after the Map and Reduce, which is used to realize the connection among data tables, making it more suitable for the database processing. Phoenix developed by Stanford implements MapReduce in CMP, SMT and SMP platform on the back of MapReduce in multi-core systems outlook.

Microsoft puts forward the Dryad programming model [4] for parallel computing of data which basic idea is to abstract the tasks into the nodes in the graph, and then maps these nodes to compute node available in the runtime environment. At present Dryad model mostly supports multi-core and cluster system. Compared with MapReduce, Dryad model is more flexible and universal, but to some extent, it is at the cost of sacrificing simplicity.

Dryad LINQ integrates the Dryad and Microsoft.net Language Integrated Query (.Net Language Integrated Query), in which users can process massive data as long as writing programs similar to SQL Query Language. And this model can completely compatible with MapReduce as well, is a kind of more generic parallel data programming model.

Merge is a programming model and operating environment [5] which aims at that designing multi processor system about isomerism. It offers a programming interface based on library and realizes the language models such as MapReduce, in order to hide the isomerism of underlying platform. At run-time, Merge is capable of mapping operation to heterogeneous processing unit automatically according to its input and computing resources available

B. The system support

Cloud computing needs support in highly scaleable and extensible file system and operating system owing to it requires much more management resources than before, for this reason, the design of trusted cloud computing poses a series of new challenges for the design of the bottom of the operating system and file system. Meanwhile the availability and maintainability of the system also need to pay close attention to since cloud computing platform heavy employs daily low configuration computing components such as PC and so forth.

Google File System [6] (GFS) is a typical extended file system developed by Google which is based on cheap PC and storage equipment. Yet unlike traditional file, the GFS no longer processes the equipment failure as exception event, instead it is handled as general system event. Furthermore, the update operation of the GFS file is seen to be adding files and thus once the data was added, it is managed as read-only file, so as to improve reading efficiency of the file effectively.

III. GROUP THEORY FOUNDATION

Group theory is a branch of abstract algebra [9], which mainly studies all kinds of algebraic structures, that is the sets of algebraic operation, and is the foundation of modern science. With technological improvements, especially the development of computer technology, the theory and method of abstract algebra are also to be perfect step by step, and have more application fields as well, which has penetrated into all fields of science and each practical application department. This section makes a brief introduction on the foundational mathematics of group theory of trusted cloud platform.

Definition 1: Semi-group. If S is a non-empty set, in case of satisfying that there is an algebraic operation ☆ in S, this operation can meet that to arbitrary element of the set S: a, b, c ,the associative law (a ☆ b) ☆ c = a ☆ (b ☆ c) is correct, S is called a semi-group about operation ☆, recorded as (S, ☆).

Definition 2: Commutative semi-group. If the operation of semi-group S ☆ can satisfy the commutative law: a ☆ b

= b ☆ a, and a, b are the elements of the set S, then S is referred to as commutative semi-group.

Definition 3: Suppose f is a mapping from set A to set B, if satisfying that, for any elements a, b of A, when a≠b, f (A)≠f (B), then f is called an injection from A to B; If for any elements b of B and any elements a of A, exist f (a) = b, then f is called a surjection from A to B; If f is both an injective and a surjection, then f can be called a bijection.

Define 4: Suppose S and S ' are all semi-group, f is a mapping from S to S', if f can maintain operation, that is arbitrary element of S: x, y meet the operation: f (x ☆ y) = f (x) ☆ f (y), then f is called the homomorphism from S to S'.

When the homomorphism f is an injection, then says that f is a monomorphism ; When the homomorphism f is a surjection, then says that f is an epimorphism, the homomorphism of S and S ' recorded as S ~ S'; When the homomorphism f is a bijection, says that f is an isomorphism, the isomorphism of S and S ' recorded as S ≅ S'.

IV. DESIGN OF TRUSTED CLOUD COMPUTING PLATFORM BASED ON GROUP THEORY

A. *The overall design*

In this paper the concept of the trusted platform is expanded to the whole background, infrastructure as a service IaaS, puts forward trusted cloud computing platform based on group theory (TCCPoGT), which provides a completely enclosed execution environment, in order to ensure the confidentiality and integrity of users, and users can do an advance judge, decide that whether IaaS takes thest measures. As the figure 1 shows, the trusted computing base of TCCPoGT includes the following two contents: trusted public key crypto system based on group theory and trusted coordinator (TC). In the diagram, N is a trusted node, TC is trusted coordinator, CM is an untrusty cloud managers which provides a series of services to the user and ETE is an external reliable entity that is responsible for the maintenance of TC.

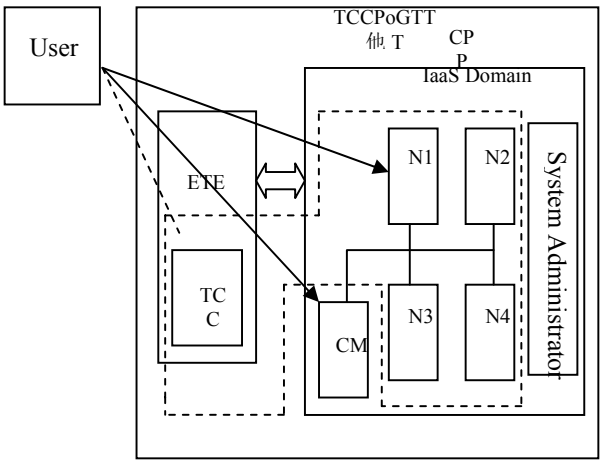

Figure 1. Trusted cloud computing platform architecture based on group theory

Each node behind the scenes runs the TMMoGT that controls user, and prevents privileged users from peeping or tampering with. TMMoGT can be able to protect its own safety and abide by the protocol TCCPoGT. Trusted platform module is embedded into node, and can load TCCPoGT through starting process safely. The essence of TCCPoGT is actually that it is a local enclosed environment which could resist malicious system administrator, TC is responsible for managing a series of client nodes that operate safely, and the nodes are called trusted nodes as well. In order to achieve the purpose of trusted, node is necessary to appear in the security field, and run TCCPoGT, which requests that TC ought to save the record when its node emerges in secure domain, and can verify node platform so as to be able to judge whether the node is running reliable TCCPoGT. TC manages the events such as adding and removing nodes in the cluster , or temporarily closing nodes due to upgrade and maintenance, the user can make the judgment about the safety of IaaS via the verification of TC. To make sure the safety of users, each TCCPoGT running in each node must be in cooperation with the TC , the aim is to ensure that the users' state for migration will not be peeped or tampered with. Here we assume that external reliable entity ETE is maintained by a third party which has no common motive with IaaS service provider.

B. *Utilizing group theory to realize public key cryptosystems*

Suppose group S is a subgroup of group G, and $T = C_G(H)$ is a centralizer of S on G, then to the integer m, n (n > 1), $M \in \{0.1\}^m$ is clear text, hash function H are defined as follows: $H : G^{\rightarrow}\{0,1\}^m$ .And the new key algorithm based on the group theory is: private key is $\beta \in S_X^{n+1}$, and public key is $\alpha \in G_X^n$. Encryption process is selected at random $\gamma \in T_X^{n+1}$, the cryptograph is $(\gamma(a), H(\gamma_1 \pi(\beta(\alpha, e))\gamma_{n+1}^{-1}) \dot{Y} M)$, if the encrypted cryptograph is (a, b) , then decrypts in order to recovering the plaintext through $H(\pi(\beta \bullet (a, e))) \dot{Y} b$.

The security analysis of key cryptosystem: according to the definition of the core of the group theory's difficulties MSRP (Multi - variant Subgroup Root Problem) and SAP (Subgroup Action Problem), it is known that anyone who want to get β from public key will have to solve the problem MSRP, and anyone who want to get β from must solve the puzzle SAP. When the hash function H is secure enough, it is impossible to obtain $\gamma_1 \pi(\beta(a, e))\gamma_{n+1}^{-1}$, as a result it is impossible to obtain β or M with the help of b.

C. *Node management*

Through the management of the nodes during the security domain and the public key of trusted security platform module TPM which is used to recognize nodes and the directory of expected measure list, TC is capable of supervising a series of controlled reliable node dynamically. ETE can guarantee the safety and publicly available of TC's part parameters , for instance, the value MLTC and MLN that the identification node N of remote operation platform or TC want to receive, etc.

Node must register according to the agreement described in figure 2 at TC, the first two steps are the verification of node N to TC, firstly, node N sends challenge information nN to TC, then TC returns encrypted information MLTC after receiving that, if the received MLTC is consistent with what N expects, it says the TC is reliable. At the same time the return information in the second time that TC sends contains a challenge information nTC to node N. After that, step 3, the node N generates key (TKPN, TKPN), and the public key and the third validation message are transmitted to TC together . If the node N is confirmed as a trusted identity by TC, then send the forth acknowledgement to affirm that the node N is credible. Platform should be able to guarantee that the trusted node is still credible after restarting, otherwise it may threaten the security of the system, moveover, the nodes will be stopped by platform after restart and must register again.

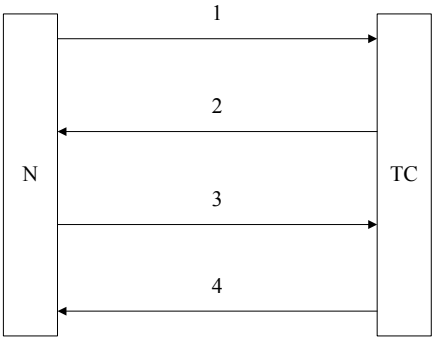

Figure 2. Information interaction of node management processing

V. CONCLUSIONS

Cloud computing platform puts forward higher challenge about system's credibility than traditional computing platform for the integration of supersized computer system, and it provide applications through the Internet for users from different organizations. This paper first have in depth analyses of the design imperative of the trusted cloud computing platform, then gives design principles and the current research situation of trusted cloud platform, and then devises a trusted cloud computing platform TCCPoGT based on group theory upon it analyses the basic concept of group theory, and introduces the design of the platform in several ways in detail: the overall architecture, public key cryptosystems based on the theory and node management , the platform is able to guarantee the security of the service.

REFERENCES

[1] Buyya R, Yeo C S, Venugopal S. Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. Keynote at the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08). IEEE, 2008 ,PP.5-13.

[2] Dean J, Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. Proceedings of Usenix Symposium on Operating System Design and Implementation.PP.137-150,2010

[3] Yang H, Dasdan A, Hsiao R, Parker D. 14Zap-reduce-merge: simplified relational data processing on large clusters.Proceedings of the 2007 ACM SIGMOD international conference on Management of data, 2007,PP.1029-1040

[4] Isard M, Budiu M, Yu Y, Birrell A, Fetterly D. Dryad: distributed data parallel programs from sequential building blocks.Proceedings of the 2007 conference on EuroSys, 2007,PP.59-72

[5] Linderman M, Collins J, Wang H, Meng T. Merge: a programming model for heterogeneous mufti-core systems.Proceedings of ACM Symposium on Architecture Support on Programming Language and Operating Systems. ACM New York, NY, USA, 2008 ,PP.287-296.

[6] Ghemawat S, Gobioff H, Leung S. The Google file system.ACM SIGOPS Operating Systems Review, 2003,PP. 29-43.

[7] Ping-tian Zhu, Bo-hong Li, Yuan-bian Zou. Modern algebra . Beijing: science press. 2001