

A BRIEF SURVEY ON FEW COMMON ISSUES OF HIGH PERFORMANCE COMPUTING : GRID AND CLOUD

Smriti Priya Medhi¹, Esha Barlaskar²

ABSTRACT

The recent trend of computing or analysing bulk amount of data needs a substantial support of a number of parallel resources. This can only be achieved via a high performance computing environment/infrastructure. Although going for a supercomputer is also a probable solution to it, but hiring/buying the same involves a large amount of money. Existing supercomputers are very exotic machines with high-end technologies inside them. They are gigantic in terms of scalability. In some of the present ones, 10,000 of thousands of processors make up a single machine. However high performance computing environment can be thought of a scenario, where a supercomputer have been divided into small clusters of nodes/machines, where the clusters may be part of the same organization or even across organizational boundaries. It is obvious that such type of a high-end processing environment will provide a fruitful contribution to the computing community. But the probable issues associated with it cannot be ignored. Secure data

transaction/distribution/sharing among all the resources are some of the major problems faced by developers and programmers. This paper primarily discusses the issues of security associated with grid and cloud computing and provides a brief review of the probable solutions to these issues.

Keywords : HPCA, grid computing, cloud computing,

I. INTRODUCTION

The field of computing have been in a continuous need of new, advanced and faster means of computing. As computing is the root of every application, and almost all applications now-a-days involves a large percentage of data to be computed. The computation of these large quantities of data cannot be realized in a single machine, if we want the results in a considerable time frame. This demand of faster processing, led researchers develop and opt for supercomputing or parallel computing (parallel supercomputers) techniques. One of the other method, that is in high demand and use today, to satisfy the requirement of high level processing, is the high performance computing. The high performance computing environment is achieved by integrating low-end resources in a cluster, so that their processing capabilities can be used to provide a performance elevation to the computation of bulk data in various applications.

Over the years, with the decline in performance of single-core architectures [7]; people have seen the concept of

¹Department of Computer Science Engineering and Information Technology, Don Bosco College of Engineering and Technology, Assam Don Bosco University, Airport Road Azara, Guwahati 781017, India
E-mail : smriti_official@yahoo.com

²Department of Computer Science Engineering and Information Technology, Don Bosco College of Engineering and Technology, Assam Don Bosco University, Airport Road Azara, Guwahati 781017, India
E-mail : eshabarlaskar@gmail.com

IV. SOME MAJOR ISSUES IN A HPC ENVIRONMENT

The HPC environment involves open sharing of computing resources, databases, important and sensitive data. The data that is to be analysed is also in bulk amount. The ultimate goal of constructing a HPC infrastructure is not only of achieving an accelerated performance level in computing but also to facilitate a secure transmission/sharing/distribution of data. There are many issues in different HPC technologies. We are discussing few common issues between Grid and Cloud computing only.

◆ Common Issues Between Cloud And Grid Computing:

Cloud and Grid Computing are quite similar in the sense that both share lots of issues in common. The most important issues includes:

1. Scalability : Data must be distributed over many computers so as to achieve good scalability [1].
2. Security: People can be afraid of sending sensitive data through a large number of computers.[1].
3. Data must be moved repeatedly to distant computers, which generates the bottleneck of the process, since the data is not always available everywhere and sometimes it is necessary to make this data available[1]
4. Data can be requested regardless of its location[1].
5. Service-level agreements (SLAs) are very important in both Cloud and Grid computing. SLAs are made between customers and providers where uptime availability is guaranteed of, say, 98 or 99 percent. The consumer will get service credit for receiving data late if the service moves down the level of

the guaranteed uptime service as mentioned in SLA. [2]

6. Both systems must be able to determine the amount of unused resources and then distribute the work from over loaded resources to the unused ones. Thus Load balancing is another important issue[2].

Although there are many common issues between Grid and Cloud Computing, the most important on the other hand, Grid Computing just like Cloud Computing has a lot of security issues that should be considered since it is a distributed system where a heterogeneous set of computers share their idle resources. The authors in [6] presented various solutions on how security of data in grid should be tackled and what are the issues, that a user has to face in doing so. The main challenges and solutions in Grid are described in Table 2 below:

From the Tables 1 and 2 we came to know some of the security issues in Cloud [14] and Grid Computing[13] and their contributed solutions. However, we should agree on the fact that the security is better in Grid than in Cloud because Grid Computing was introduced much before issue which is of great concern is Security issues which are explained in details in [5]. Every HPCA have different issues related to security, likewise, Cloud Computing too has lots of problems. The network infrastructure has to be very carefully controlled but it is not possible to guard it always thus it may so happen that if the data are not carefully handled then it may be captured or tampered by third parties. Hence, proper security solutions must be incorporated into cloud. We are presenting very few issues which were discussed by the authors in [2] and [4]. They have proposed solutions for some security problems

A Brief Survey on Few Common Issues of High Performance Computing : Grid and Cloud

which are described in Table I as follows. Cloud came into existence and as Cloud computing is less mature than Grid Computing till now in providing solutions for different security issues. Ian Foster et. al. [1] stated the same thing in one of their works where the authors said that the security model adopted by Cloud appears to be comparatively easy and less secure than the Grids security model. We can justify this comment by taking the example of any software which can be thought of as secure if after its development process it is properly tested until it becomes bugs free. Thus in this context we can say that since Cloud Computing is new compared to Grid Computing so Cloud it is less tested and as a result there is the possibility of discovering more bugs in future than in Grid which has already been tested time and again.

TABLE 1 : SECURITY ISSUES IN CLOUD COMPUTING

CLOUD SECURITY ISSUES	SOLUTION
Web application vulnerabilities: Cross scripting, SQL Injections	Develop a security oriented framework that teaches the best programming practices
Vulnerabilities inherent to the TCP/IP stack and/or the operating systems: DoS[9], and DDos(Distributed denial of service)	Deactivate unused services, update applications and control rights
Authentication problems: IP spoofing, RIP attacks, ARP poisoning	Use encrypted protocols if possible prevents IP spoofing, controlling rights to access ARP tables, etc.
Verification problems: The verification, tampering and loss of data.	Encrypted data would be a solution, but, 'since the unencrypted data must reside in the memory of the host running the computation', this must be encrypted in order to avoid memory copies
Physical access: Only authenticated users can use their data but sometimes it unauthenticated users access it	Control rights and log actions when accessing the hardware.
Privacy control of data: Data goes to wrong hands thereby causing damage to it	Use Service-level agreements.

TABLE 2 : SECURITY ISSUES IN GRID COMPUTING

GRID SECURITY ISSUES	SOLUTIONS
It is very important to check that a machine which is globally sharing its resources with users running their applications on it has not been compromised.	Schedulers that allow only authentic users to run their applications and give them enough rights to deploy the same. Authenticity A specialized scheduler that allows users with enough rights to run applications.
Mapping of the ID of the user using the grid and the ID of the grid.	Grids with centralized control can solve this issue. Some other solutions to it however include a single Grid sign-up mechanism.
The problem of having access of all logs of all users	Special libraries can be used to control and protect the issue of giving access to the logs of the users of the grid.
Accessing policies, locally or globally.	In order to access policies, the owner can digitally sign the policies[10] keeping them secure. In global scenarios, it is mandatory to have a secure connection and he must be able to authenticate himself.
Achieving data integrity and its confidentiality totally	Algorithms including MAC help in maintaining integrity and confidentiality can be achieved through encryption methods having temporary and dynamic keys.
The management of keys in encryption methods	Smart Cards can be a probable solution
A factor of trust between the users and the service providers	SSL credentials or secure DNS and IPSec can provide support in this case.
Grids are always guarded by firewalls and VPNs. This becomes a challenge.	Configuring servers to allow known ports to run on firewalls and authenticating VPNs with identity certificates such as x509 can be the probable solutions.