Approach to effective and efficient Cloud Computing environment

Alexander Bogdanov [1]
[1] *Institute for High-performance computing and integrated systems, e-mail: bogdanov@csa.ru, Saint-Petersburg, Russia*

Kyaw Zaya[2], **Pyae Sone Ko Ko**[3]
[2,3]*St.Petersburg State Marine Technical University, e-mail: kyawzaya4436@gmail.com, pyaesonekoko@gmail.com, Saint-Petersburg, Russia*

In this paper we discuss the approach to cloud computing environments and their efficiencies, evaluated by their load balancing systems, resources management system, performance of workload and data processing. The concept of cloud computing has been around for years under numerous labels: mainframe multitasking, grid computing, on-demand computing, and hosted IT. Essentially, cloud computing uses the internet to deliver remotely hosted applications, data, and infrastructure services. The advantage of cloud computing is the ability to virtualize and share resources among different applications with the objective for better server utilization [9]. The cloud computing is a large distributed system that employs distributed resources to deliver a service to end users by implementing several technologies. Hence, providing acceptable response time for end users presents major challenge for cloud computing [9][11].

Introduction

Recently, cloud computing as a new service concept has become popular to provide various services to user such as multi-media sharing, online office software, game and online storage. The cloud computing is bringing together multiple computers and servers in a single environment designed to address certain types of tasks, such as scientific problems or complex calculations. This structure builds up a lot of data, distributed computing nodes and storage. Typically, applications executed in a distributed computing environment, apply to only one data source. Cloud computing delivers infrastructure, platform, and software (application) as services, which are made available as subscription-based services in a pay-as-you-go model to consumers. These services in industry are respectively referred to as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). It provides the scalable IT resources such as applications and services, as well as the infrastructure on which they operate, over the Internet, on pay-per-use basis to adjust the capacity quickly and easily. It helps to accommodate changes in demand and helps any organization in avoiding the capital costs of software and hardware. Approach to effectiveness and efficiency Cloud computing environment, we should take the attention on many factors, the existing load balancing techniques in cloud computing and further compares them based on various parameters like performance, scalability, associated overhead [4] [11].

Approach to hybrid Cloud

Basically we can describe, that the hybrid cloud is a combination of public and private clouds bound together by either standardized or proprietary technology that enables data and application portability [5]. It could be a combination of a private cloud inside an organization with one or more public cloud providers or a private cloud. We propose to use a cloud approach based on open standards and utilizing several up-today technologies, that make it very effective for large scale problems. Some of the main factors to approach to hybrid cloud environment are depend on it distributed file system, that can process to deal with massive amount, processing of large data sets is done via shared virtual memory and the use of API [8][9][4].

Load Balancing in Cloud Computing

Load balancing in clouds is a mechanism that distributes the excess dynamic local workload evenly across all the nodes. It is used to achieve a high user satisfaction and resource utilization ratio, making sure that no single node is overwhelmed, hence improving the overall performance of the system. Proper load balancing can help in utilizing the available resources optimally, thereby minimizing the resource consumption. It also helps in implementing fail-over, enabling scalability, avoiding bottlenecks and over-provisioning, reducing response time; and its technique that facilitates networks and resources by providing a maximum throughput with minimum response time [3]. Dividing the traffic between servers, data can be sent and received without major delay. Different kinds of algorithms are available that help traffic loaded between available servers. Generally in Cloud computing use the dynamic algorithm, which allows cloud entities to advertise their existence to presence servers and also provides a means of communication between interested parties [6]. We present some fundamental load balancing techniques in Cloud environment are[2][3][5]:

1. Decentralized content aware load balancing: proposed a new content aware load balancing policy named as workload and client aware policy (WCAP). It uses a unique and special property (USP) to specify the unique and special property of the requests as well as computing nodes. USP helps the scheduler to decide the best suitable node for the processing the requests this technique improves the searching performance and, hence, overall performance of the system. It also helps in reducing the idle time of the computing nodes, hence, improving their utilization [3].
2. Server-based load balancing for Internet distributed services: proposed a new server based load balancing policy for web servers which are distributed all over the world. It helps in reducing the service response times by using a protocol that limits the redirection of requests to the closest remote servers without overloading them. A middleware is described to implement this protocol. It also uses a heuristic to help web servers to endure overloads [3].
3. Scheduling strategy on load balancing of virtual machine resources: proposed a scheduling strategy on load balancing of VM resources that uses historical data and current state of the system.

This strategy achieves the best load balancing and reduced dynamic migration by using a genetic algorithm. It helps in resolving the issue of load-imbalance and high cost of migration, thus achieving a better resource utilization [3].

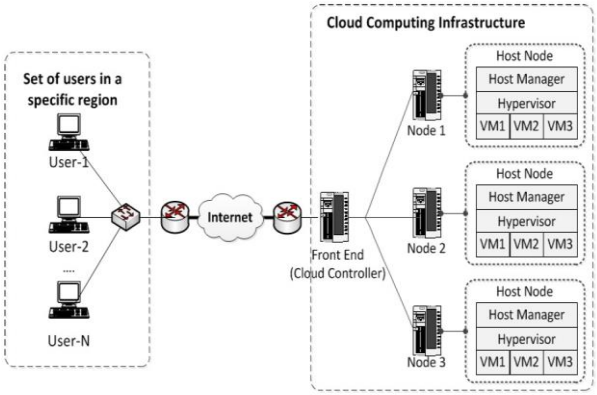

Figure1. Cloud computing infrastructure

Resources management

Services-based resources should not be different from the resources in your own environment, except that they live remotely. Ideally, you have a complete view of the cloud computing resources. In most cloud environments, the customer is able to access only the services they're entitled to use. Entire applications may be used on a cloud services basis. Development tools are sometimes cloud based [7]. In fact, testing and monitoring environments can be based on the cloud. It enables users to migrate their data and computation to a remote location with minimal impact on system performance. Clouds are designed to deliver as much computing power as any user wants. While in practice the underlying infrastructure is not infinite, the cloud resources are projected to ease the dependence on specific hardware [12].

Figure2. Cloud resources management

Data processing

Massive data processing in Cloud computing must handle effectively large amount entities such as applications, accessing to data resources. That's why, we need to realize the efficiency of massive data processing in Cloud environment with the combination of subsystems for judgment and accumulation with large number of server in order to evaluate massive event streams and need to be ensured that failure of any of the data processing server does not stop the entire services [1].

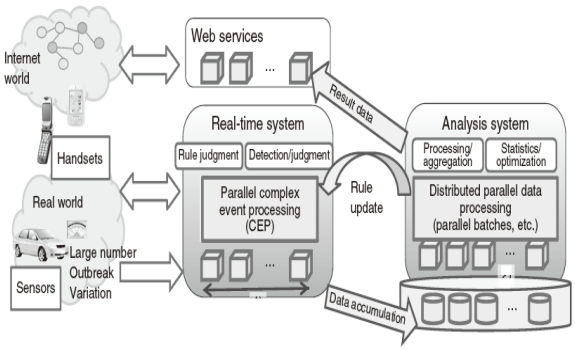

Figure3. Massive data processing in Cloud Infrastructure

Conclusion

The point to determine the abilities of Cloud computing environments is the aggregation of resources and data into data centers on the internet. Cloud services realize an improve execution efficiency by aggregating application execution environments at various levels for sharing them. Cloud computing is to provide on demand services according to the clients requirements within a stipulated time [2]. Further, it provides the users for accessing the shared pool of distributed resources. Cloud is a pay-go model where the consumers pay for the resources utilized instantly, which necessitates having highly available resources to service the requests on demand.

References

1) Satoshi Tsuchiya, Yoshinori Sakamoto, Yuichi Tsuchimoto, Vivian Lee. Big Data Processing in Cloud Environment.
2) Rashmi. K. S, Suma. Vaidehi. M. Load Balancing Approach to Avoid Deadlocks. 2012.
3) Nidhi Jain Kansal1, Inderveer Chana. Cloud Load Balancing Techniques: A Step Towards Green Computing. IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012.
4) Samuel Greengard. IDG White Paper. The Silver Lining: To Build an Effective Cloud Computing Infrastructure.
5) Rui Han, Li Guo, Moustafa M. Ghanem, Yike Guo. Lightweight Resource Scaling for Cloud Applications.2012.
6) Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, Sonia Lopez-Alarcon, Warren Carithers. Efficient Resource Management for Cloud Computing Environments.
7) Moshe Shadmon, CTO, ScaleDB Inc. Clustered Database Storage Engine: Scaling MySQL in the Cloud. November 12, 2010.
8) Preeti Agrawal1, Yogesh Rathore: Resource Management In Cloud Computing With Increasing Dataset. June, 2012.
9) Alexander Bogdanov, Thurein Kyaw Lwin, Anatoly Shuvalov, Soe Moe Lwin. Unconventional use of distributed databases from server consolidation to consolidation resources. The 4th International Conference "Distributed Computing and Grid-technologies in Science and Education" June 28 - July 3, 2010, Dubna, Russia.
10) Amazon elastic compute cloud (EC2): http://aws.amazon.com/ec2/
11) D. A. Bacigalupo, J. van Hemert, et al, "Managing dynamic enterprise and urgent workloads on clouds using layered queuing and historical performance models," Simulation Modelling Practice and Theory, vol. 19, pp. 1479-1495, 2011.
12) An Approach For Effective Resource Management in Cloud Computing (International Journal EnggResearch.net, Issue Dec 2011)