Volume 4, Issue 5, May 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Survey of Virtual Machine Scheduling in Cloud Computing

Deepti Gupta *
Research Scholar MIET Department of Computer Sc & Engineering
Kurukshetra University

Amit Chaudhary
Assistant Professor MIET
Department of Computer Sc & Engineering
Kurukshetra University

Abstract— Cloud computing is a most recent new computing paradigm where applications, records and IT services are provided over the Internet. Cloud computing is the fastest new paradigm for delivering on demand services over internet and can be described as internet centric software. Cloud computing has come out to be an interesting and beneficial way of changing the whole computing Schedulers for cloud computing determine on which processing resource jobs of a workflow should be allocated . Cloud computing is basically an Internet-based network made up of large numbers of servers - mostly based on open standards, modular and inexpensive. Clouds contain vast amounts of information and provide a variety of services to large numbers of people.

Keywords— Cloud Computing; SAAS,PAAS,V.M

I. INTRODUCTION

In recent 10 years, Internet has been developing very quickly. The cost of storage, the power consumed by computer and hardware is increasing. The storage space in data center can't meet our needs and the system and service of original internet can't solve above questions, so we need new solutions. At the same time, large enterprises have to study data source fully to support its business [1]. The collection and analysis must be built on a new platform. So we need a new computing model to utilize the vacant resources of computer, increase the
economic efficiency through improving utilization rate,
decrease the equipment energy consumption. Cloud computing is everywhere. When we open any IT magazines, websites , radios or TV channels, "cloud" will definitely catch our eye. Today's most popular social networking, e- mail, document sharing and online gaming sites, are hosted on a cloud. More than half of Microsoft developers are working on cloud products. Even the U.S government intends to initialize cloud-based solutions as the default option for federal agencies of 2012. Cloud com- puting makes software more attractive as a service, and shapes the way in which IT hardware is purchased. Predictably, it will spark a revolution in the way organizations provide or consume information and computing. Cloud computing is rising as the next generation platform for computation. May be in future Cloud computing will be the main platform to save the world; this makes people can have everything they need on it. Main advantages of the Cloud computing is used for on-demand gathering of information, technology services and products. The name Cloud has come from the Internet, based on how it is depicted in computer network diagrams, and is an abstraction for the complex infrastructure it conceals. In general we can define cloud computing is style of computing in which IT-related capabilities are provided "as a service", allowing users to access technology-enabled services from the Internet without knowledge of, expertise with, or control over the technology infrastructure that supports them. Email application was probably the first service on the "cloud" [5] [6]. As now a day many computing industry shifting toward providing Platform as a Service (PaaS), Infrastructure as a Service (IaaS) and Software as a Service (SaaS) for consumers and enterprises to access on demand regardless of time and location, this helps in avoiding the over-supplying of the resource when used with utility pricing, which meets the demands of millions of users.

II. Cloud Computing Architecture

Cloud computing architecture, just like any other system, is categorized into two main sections: Front End and Back End. Front End can be end user or client or any application (i.e. web browser etc.) which is using cloud services. Back End is the network of servers with any computer program and data storage system. It is usually assumed that cloud contains infinite storage capacity for any software available in market. Cloud has different applications that are hosted on their own dedicated server farms.

Cloud has centralized server administration system. Centralized server administers the system, balances client supply, adjusts demands, monitors traffic and avoids congestion. This server follows protocols, commonly known as middleware. Middleware controls the communication of cloud network among them.

Cloud Architecture runs on a very important assumption, which is mostly true. The assumption is that the demand for resources is not always consistent from client to cloud. Because of this reason the servers of cloud are unable to run at their full capacity. To avoid this scenario, server virtualization technique is applied. In sever virtualization, all physical servers are virtualized and they run multiple servers with either same or different application. As one physical server acts as multiple physical servers, it curtails the need for more physical machines.

As a matter of fact, data is the most important part of cloud computing; thus, data security is the top most priority in all the data operations of cloud. Here, all the data are backed up at multiple locations. This astoundingly increases the data storage to multiple times in cloud compared with a regular system. Redundancy of data is crucial, which is a must-have attribute of cloud computing.

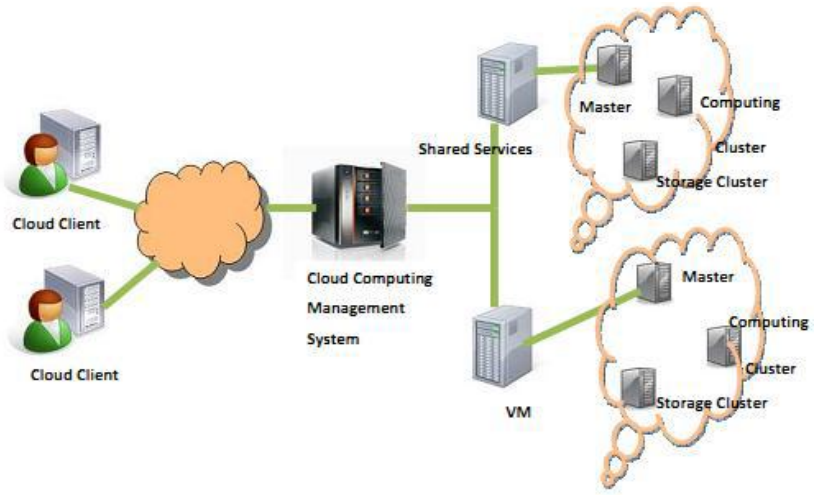


Figure3.1. Cloud Computing Architecture [2]

Cloud computing provides three different kind of services are-

- Different Services
- Platform as a Service(PaaS):

Examples: Google AppEngine, Azure services and Amazon web services.

- Infrastructure as a Service (IaaS):

Examples: VMware, Amazon EC2.

- Software as a Service (SaaS):

Examples: salesforce.com: for buying software's on demand.

Cloud computing is used in many applications are-

- Applications
- Educational – Image filtering, rendering.
- Commercial – online gaming.
- Scientific- Protein structure prediction etc.

III. Cloud Service

As an underlying delivery mechanism, cloud computing ability is provisioned as services, basically in three levels: software, platform and infrastructure.

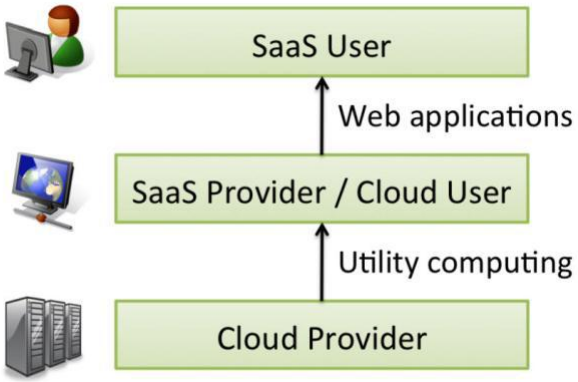
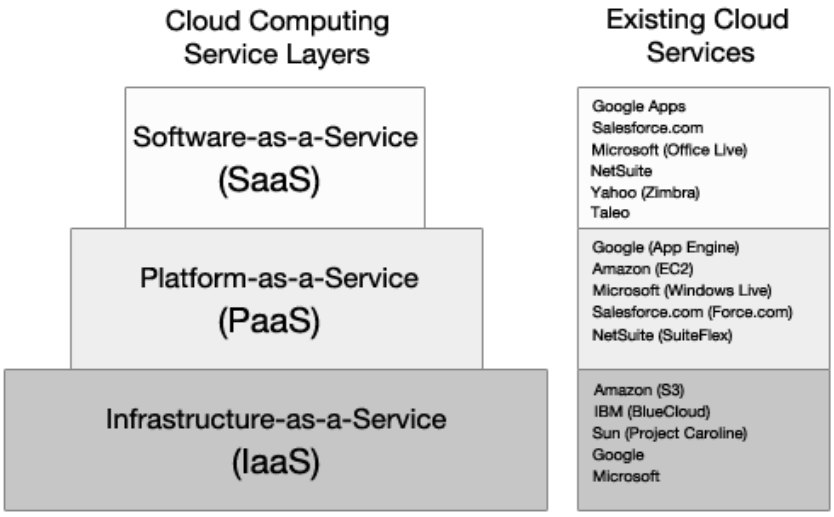


Figure 3.1. Users and Providers of Cloud Computing.

3.1.1 Software as a Service

Software as a Service (SaaS) is a software delivery model in which applications are accessed by a simple interface such as a web browser over Internet. The users are not concerned with the underlying cloud infrastructure including network, servers, operating systems, storage, platform,etc. This model also eliminates the needs to install and run the application on the local computers. The term of SaaS is popularized by Salesforce.com, which distributes business software on a subscription basis, rather than on a traditional on-premise basis. One of the best known is the solution for its Customer Relationship Management (CRM). Now SaaS has now become a common delivery model for most business applications, including accounting, collaboration and management



3.1.2 Platform as a Service

Platform as a Service (PaaS) offers a high-level integrated environment to build, test, deployand host customer-created or acquired applications. Generally, developers accept some restrictions on the type of software that can write in exchange for built-in application scalability. Customers of PaaS do not manage the underlying infrastructure as SaaS users do, but control over the deployed applications and their hosting environment configurations .PaaS offerings mainly aim at facilitating application development and related management issues. Some are intended to provide a generalized development environment, and some only provide hosting-level services such as security and on-demand scalability. Typical examples of PaaS are Google App Engine, Windows Azure, Engine Yard, Force.com, Heroku, MTurk.

3.1.3 Infrastructure as a Service

Infrastructure as a Service (IaaS) provides processing, storage, networks, and other fundamental computing resources to users. IaaS users can deploy arbitrary application, software, operating systems on the infrastructure, which is capable of scaling up and down dynamically. IaaS user sends programs and related data, while the vendor's computer does the computation processing and returns the result. The infrastructure is virtualized, flexible, scalable and manageable to meet user requirements. Examples of IaaS include Amazon EC2 , VPC, IBM Blue Cloud, Eucalyptus, FlexiScale, Joyent, Rackspace Cloud,

3.1.**4MSP (management service provider)**

This is one of the ancient applications of cloud computing.

This application mostly serves the IT industry instead of end users. It is often used in mail virus scanning and program monitoring.

3.1.5 Commercial service platform

The commercial service platform is the mixture of SAAS and MSP (Mixed signal Processor), this kind of computing provides a platform for the interaction between users and service provider. For instance, the user individual expense management system can manage user's expense according user's setting and coordinate all the services that users purchased.

IV. Cloud computing Technique

Google File System (GFS)

Google File System (GFS) is a proprietary distributed files system developed by Google Inc. for its own use. It is designed to provide efficient, reliable access to data using large clusters of commodity hardware.

GFS is optimized for Google's core data storage and usage needs (primarily the search engine), which can generate enormous amounts of data that needs to be retained; Google File System grew out of an earlier Google effort, "Big Files", developed by Larry Page and Sergey Brin in the early days of Google, while it was still located in Stanford. Files are divided into chunks of 64 megabytes, which are only extremely rarely overwritten, or shrunk; files are usually appended to or read. It is also designed and optimized to run on Google's computing clusters, the nodes of which consist of cheap, "commodity" computers, which means precautions must be taken against the high failure rate of individual nodes and the subsequent data loss. Other design decisions select for high data throughputs, even when it comes at the cost of latency. beneath the set number).

V. VIRTUAL MACHINE SCHEDULING

Virtual machine images will always change because the layers of software within them will always need to be patched, upgraded, or reconfigured. What doesn't change is the process of creating the virtual machine image, and this is what developers should focus on. A developer might build a virtual machine image by layering a Web server, application server, and MySQL database server onto an operating system image, applying patches, configuration changes, and interconnecting components at each layer. Focusing on the model, rather than the virtual machine image, allows the images themselves to be updated as needed by re-applying the model to a new set of components. With this standard

deployment unit, cloud architects can use appliances that help to speed deployment with lower costs. A developer might use an appliance that is preconfigured to run Hadoop on the OpenSolaris OS by interacting with the appliance's API. Architects can use content switches that are deployed not as physical devices, but as virtual appliances. All that needs to be done to deploy it is interact with its API or GUI.

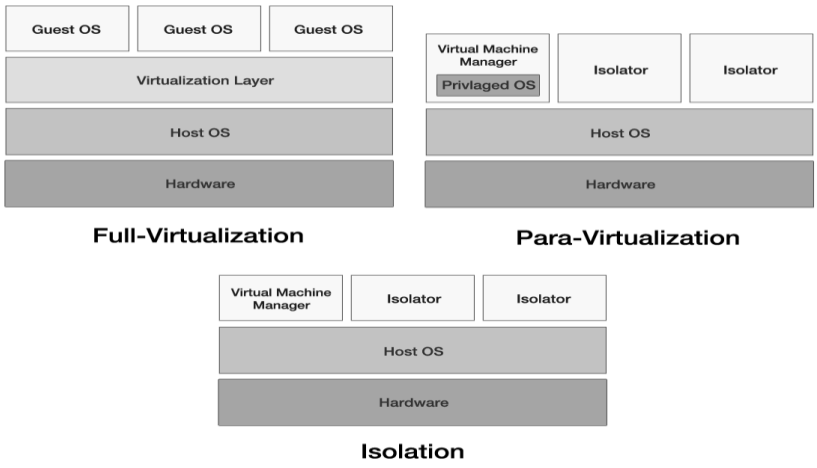


Figure 1.5 - Three forms of virtualization

VI. Related work

Sandeep Tayal et al[], Cloud computing is a latest new computing paradigm where applications, data and IT services are provided over the Internet. The Task management is the key role in cloud computing systems. Task scheduling problems are premier which relate to the efficiency of the whole cloud computing facilities. Task scheduling algorithm is an NP-completeness problem which play key role in cloud computing. Cloud computing is Internet-based development and use of computer technology. The cloud is a metaphor for the Internet and is an abstraction for the complex infrastructure it conceals. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. It define in three models Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS).

Andrew et al Cloud computing is becoming one of the most explosively expanding technologies in the computing industry today. It enables users to migrate their data and computation to a remote location with minimal impact on system performance. There are a number of underlying technologies, services, and infrastructure-level configurations that make Cloud computing possible. One of the most important technologies is the use of virtualization [10], [11]. Virtualization is a way to abstract the hardware and system resources from a operating system. This is typically performed within a Cloud environment across a large set of servers using a Hypervisor or Virtual Machine Monitor (VMM) which lies in between the hardware and the Operating System (OS). From here, one or more virtualized OSs can be started concurrently as seen in Figure 1, leading to one of the key advantages of Cloud computing. This, along with the advent of multi-core processing capabilities, allows for a consolidation of resources within any data center. It is the Cloud's job to exploit this capability to its maximum potential while still maintaining a given QoS.

Karthik Kumar et al[], In this cloud computing, a resource is a virtual machine that guarantees a certain level of performance to the user. For example, Amazon's EC2 cloud service defines virtual machines with speeds in "compute units"; this provides a hardware independent definition of speed for the virtual machine by abstracting away variations in the underlying physical hardware.. There are three different types– standard, high memory, and high CPU, including different amounts of processors, memory, and storage at different costs. Each virtual machine is rented multiple hours, and the user is charged a fixed cost irrespective of the virtual machine's utilization within the hours. This motivates the need to find a cost-efficient allocation for a given set of tasks. In the rest of this paper, we use the terms "resources", "virtual machines (VMs)", and "processors" interchangeably. Several researchers have developed efficient allocations for real-time tasks on multi-processor systems. However, previous studies schedule tasks on a fixed number of processors. For scalable computing, the virtual machines (VMs) are rented, and can be scaled up to any number. This creates some fundamental changes in the problem. First, it implies that if a task is sufficiently parallelizable, its deadline can always be met since more VMs can be allocated to complete the task before its deadline. This is different from previous studies that examine the schedule ability on a given number of processors. Second, since the number of available VMs is nearly infinite, at every time instant, there are options using different numbers and types of VMs, based on their computing speeds and costs. For example, to finish a task, a user can select a larger number of slower, cheaper VMs or a smaller number of faster, more expensive VMs, or a combination between them. Third, acquiring VMs by rent implies a fixed charge for a given rental period. Suppose a user rents a VM for an hour and the task completes before the end of the hour, the VM becomes available for running other tasks that arrive within the hour.

Gaochao Xu et al In this paper the Cloud computing is experiencing a rapid development both in academia and industry; it is promoted by the business rather than academic which determines its focus on user applications. This technology

aims to offer distributed, virtualized, and elastic resources as utilities to end users. It has the potential to support full realization of 'computing as a utility' in the near future[1]. With the support of virtualization technology[2, 3], cloud platforms enable enterprises to lease computing power in the form of virtual machines to users. Because these users may use hundreds of thousands of virtual machines (VMs)[4], it is difficult to manually assign tasks to computing resources in clouds[5,6]. So we need an effective algorithm for task scheduling in the cloud environment. A good task scheduler should adapt its scheduling strategy to the changing environment and the types of tasks. Therefore, a dynamic task scheduling algorithm, such as Ant Colony Optimization (ACO), is appropriate for clouds. ACO algorithm is a random search algorithm, like other evolutionary algorithms. It imitates the behavior of real ant colonies in nature to search for food and to connect to each other by pheromone laid on paths traveled.

VIII. Conclusion

In cloud computing, each application of users will run on a virtual operation system, the cloud systems distributed resources among these virtual operation systems. Every application is completely different and is independent and has no link between each other whatsoever, for example, some require more CPU time to compute complex task, and some others may need more memory to store data, etc. Resources are sacrificed on activities performed on each individual unit of service. In order to measure direct costs of applications, every individual use of resources (like CPU cost, memory cost, I/O cost, etc.) must be measured.

REFERENCES

[1] Sandeep Tayal "**Task Scheduling optimization for the cloud computing system**"IJAEST Volume no-5,Issue 2.

[2] Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, Sonia Lopez-Alarcon, Warren Carithers" **Efficient Resource Management for Cloud Computing Environments**"IEEE 2010.

[3] Karthik Kumar, Jing Feng, Yamini Nimmagadda, and Yung-Hsiang Lu**" Resource Allocation for Real-Time Tasks using Cloud Computing"**IEEE 2011.

[4] Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Dan Wang" **Cloud Task scheduling based on Load Balancing Ant Colony Optimization**" 2011 Sixth Annual ChinaGrid Conference978-0-7695-4472-4/11 $26.00 © 2011 IEEE

[5] Ghemawat S, Gobioff H, Leung ST. "The Google file system," In:Proc.Of the 19th ACM Symp. on operating Systems Principles. New York: ACM Press. 2003. 29-43.

[6] Dean J, Ghemawat S. "MapReduce: Simplified dataprocessing on large clusters," In: Proc. of the 6th Symp. on Operating System Design and Implementation. Berkeley: USENIX Association, 2004. 137-150.

[7] Brrows M. "The chubby lock service for loosely-coupled distributed systems," In: Proc. Of the 7th USENIX Symp. on Operating Systems Design and Implementation. Berkeley: USENIX Association, 2006. 335- 350.

[8] Amazon Elastic Compute Cloud, http://aws.amazon.com/ec2/

[9] Google App Engine, https://appengine.google.com/

[10] MicrosoftAzure, http://www.microsoft.com/azure/windowsazure.mspx

[11] Twenty Experts Define Cloud Computing, http://cloudcomputing.syscon. com/read/612375_p.htm.

[12] R.Buyya, C.S.Yeo, S.Venugopal. "Market-Oriented Cloud Computing: Vison, Hype, and Reality for Delivering IT Services as Computing Utilites," The 10th IEEE International Conference on High Performance Computing and Communications.

