Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 2, Number 1; January-March, 2015 pp. 35-38

http://www.krishisanskriti.org/acsit.html

The Distributed Computing Paradigm: Cloud Computing

Prabha Sharma

CSE UIET, PUSSGRC Hoshiarpur
E-mail: prabha_sharma1@rediffmail.com

Abstract—*The distributed computing system uses multiple computers to solve large-scale problems over the Internet. It becomes data-intensive and network-centric. In distributed computing, the main stress is on the large scale resource sharing and always goes for the best performance. In this article, we have reviewed the emerging area of distributed computing paradigm i.e. cloud computing.*

Keywords*: Distributed Computing Paradigm, cloud computing, utility computing*

1. INTRODUCTION

Distributed computing has been an essential component of scientific computing for decades. It consists of a set of processes that cooperate to achieve a common specific goal. computing.[29] The various paradigm of distributed computing has been shown below. The utility computing is basically the grid computing and the cloud computing which is the recent topic of research. This classification is well shown in the Figure 1.1.

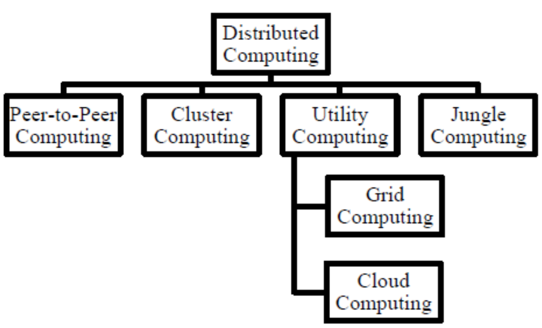

Figure 1.1: Classification of Distributed Computing

Cloud computing is becoming main computing paradigm. The use of cloud computing has been increased exponentially. It has various characteristics and advantages of on-demand computing, shared infrastructure, pay-per-use model, scalability and elasticity. Cloud computing is also called as utility computing which deliver software, infrastructure, platform as a service in pay as you use model to the consumer [1][2]. There are many definitions for cloud computing. According to National Institute of Standard and Technology(NIST) which is generally accepted standard, ”cloud computing is a model for enabling convenient, on demand, network access to a shared pool of configurable computing resources (such as networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimum management efforts or service provider interaction”. Cloud service models include Saas, Paas, Iaas and deployment models include public, private, hybrid and community-shared infrastructure for specific community. Cloud computing is a kind of ; it has evolved by addressing the QoS (quality of service) and reliability problems. Cloud computing provides the tools and technologies to build data/compute intensive parallel applications with much more affordable prices compared to traditional parallel computing techniques.[9]

Cloud computing shares characteristics with:

- Client–server model — *Client–server computing* refers broadly to any distributed application that distinguishes between service providers (servers) and service requestors (clients)[10]
- Grid computing — "A form of distributed and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks."
- Mainframe computer — Powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as: census; industry and consumer statistics; police and secret intelligence services; enterprise resource planning; and financial transaction processing.[11]
- Utility computing — The "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity.[12][13]
- Peer-to-peer — A distributed architecture without the need for central coordination. Participants are both suppliers and consumers of resources (in contrast to the traditional client–server model).

A few related term are mentioned here. A computing *cluster* consists of a collection of similar or identical machines that physically sit in the same computer room or building. Each machine in the cluster is a complete computer consisting of one or more CPUs, memory, disk drives, and network interfaces. The machines are networked together via one or more high-speed local area networks. Another important characteristic of a cluster is that it's owned and operated by a single administrative entity such as a research center or a company. Finally, the software used to program and manage clusters should give users the illusion that they're interacting with a single large computer when in reality the cluster may consist of hundreds or thousands of individual machines. Clusters are typically used for scientific or commercial applications that can be parallelized. Since clusters can be built out of commodity components, they are often less expensive to construct and operate than supercomputers .Although the term *grid* is sometimes used interchangeably with cluster, a computational grid takes a somewhat different approach to high performance computing. A grid typically consists of a collection of heterogeneous machines that are geographically distributed. As with a cluster, each machine is a complete computer, and the machines are connected via high-speed networks. Because a grid is geographically distributed, some of the machines are connected via wide-area networks that may have less bandwidth and/or higher latency than machines sitting in the same computer room. Another important distinction between a grid and a cluster is that the machines that constitute a grid may not all be owned by the same administrative entity. Consequently, grids typically provide services to authenticate and authorize users to access resources on a remote set of machines on the same grid. Because researchers in the physical sciences often use grids to collect, process, and disseminate data, grid software provides services to perform bulk transfers of large files between sites. Since a computation may involve moving data between sites and performing different computations on the data, grids usually provide mechanisms for managing long-running jobs across all of the machines in the grid.

2. CHARACTERISTICS:

Cloud computing exhibits the following key characteristics:[30]

- **Agility** improves with users' ability to re provision technological infrastructure resources.
- **Application programming interface** (API) accessibility to software that enables machines to interact with cloud software in the same way that a traditional user interface (e.g., a computer desktop) facilitates interaction between humans and computers. Cloud computing systems typically use Representational State Transfer (REST)-based APIs.
- **Cost**: cloud providers claim that computing costs reduce. A public-cloud delivery model converts capital expenditure to operational expenditure.[14] This purportedly lowers barriers to entry, as infrastructure is typically provided by a third party and does not need to be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is fine-grained, with usage-based options and fewer IT skills are required for implementation (in-house).[15] The e-FISCAL project's state-of-the-art repository[16] contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.
- **Device and location independence**[17]enable users to access systems using a web browser regardless of their location or what device they use (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.
- **Virtualization** technology allows sharing of servers and storage devices and increased utilization. Applications can be easily migrated from one physical server to another.
- **Multitenancy** enables sharing of resources and costs across a large pool of users thus allowing for:
  - **centralization** of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
  - **peak-load capacity** increases (users need not engineer for highest possible load-levels)
  - **utilisation and efficiency** improvements for systems that are often only 10–20% utilised.[18]
  - **Reliability** improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery[19]
  - **Scalability and elasticity** via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis in near real-time[20][21](Note, the VM startup time varies by VM type, location, os and cloud provider[51]), without users having to engineer for peak loads.[22][23][24]
  - **Performance** is monitored, and consistent and loosely coupled architectures are constructed using web services as the system interface[25][26]
  - **Security** can improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels.[27]Security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford to tackle.[28] However, the complexity of security is greatly

increased when data is distributed over a wider area or over a greater number of devices, as well as in multi-tenant systems shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

- **Maintenance** of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

There are many cloud vendors which offer their services upon some monitory cost. Cloud computing proved to be beneficial for enterprises [4]. Some of the big cloud infrastructure/service providers are Amazon [5], Salesforce [6], Google app engine [7] and Microsoft azure [8].

Various Cloud computing approaches use parallelism to improve the computational performance of applications. The Google MapReduce framework is particularly good at this so long as the problem fits the framework. Other approaches to high performance computing have similar constraints. It's very important for developers to understand the underlying algorithms in their software and then match the algorithms to the right framework. If the software is single-threaded, it will not run faster on a cloud, or even on a single computer with multiple processing cores, unless the software is modified to take advantage of the additional processing power. Along these lines, some problems cannot be easily broken up into pieces that can run independently on many machines. Only with a good understanding of their application and various computing frameworks can developers make sensible design decisions and framework selections.

3. VARIOUS RESEARCH AREAS

Although much progress has already been made in cloud computing, there are a number of research areas that still need to be explored. Issues of security, reliability, and performance should be addressed to meet the specific requirements of different organizations, infrastructures, and functions.[29]

Security

As different users store more of their own data in a cloud, being able to ensure that one user's private data is not accessible to other users who are not authorized to see it becomes more important. While virtualization technology offers one approach for improving security, a more fine-grained approach would be useful for many applications.

Reliability

As more users come to depend on the services offered by a cloud, reliability becomes increasingly important, especially for long-running or mission critical applications. A cloud should be able to continue to run in the presence of hardware and software faults. Google has developed an approach that works well using commodity hardware and their own software. Other applications might require more stringent reliability that would be better served by a combination of more robust hardware and/or software-based fault-tolerance techniques.

Vulnerability to Attacks

If a cloud is providing compute and storage services over the Internet such as the Amazon approach, security and reliability capabilities must be extended to deal with malicious attempts to access other users' files and/or to deny service to legitimate users. Being able to prevent, detect, and recover from such attacks will become increasingly important as more people and organizations use cloud computing for critical applications.

Cluster Distribution

Most of today's approaches to cloud computing are built on clusters running in a single data center. Some organizations have multiple clusters in multiple data centers, but these clusters typically operate as isolated systems. A cloud software architecture that could make multiple geographically distributed clusters appear to users as a single large cloud would provide opportunities to share data and perform even more complex computations than possible today. Such a cloud, which would share many of the same characteristics as a grid, could be much easier to program, use, and manage than today's grids.

Network Optimization

Whether clouds consist of thousands of nodes in a computer room or hundreds of thousands of nodes across a continent, optimizing the underlying network to maximize cloud performance is critical. With the right kinds of routing algorithms and Layer 2 protocol optimizations, it may become possible for a network to adapt to the specific needs of the cloud application(s) running on it. If application level concepts such as locality of reference could be coupled with network-level concepts such as multicast or routing algorithms, clouds may be able to run applications substantially faster than they do today. By understanding how running cloud applications affects the underlying network, networks could be engineered to minimize or eliminate congestion and reduce latency that would degrade the performance of cloud-applications and non-cloud applications sharing the same network.

Interoperability

Interoperability among different approaches to cloud computing is an equally important area to be studied. There are many cloud approaches being pursued right now and none of them are suitable for all applications. If every application were run on the most appropriate type of cloud, it would be useful to share data with other applications running on other types of clouds. Addressing this problem may require the development of interoperability standards. While standards may not be critical during the early evolution of cloud computing, they will become increasingly important as the field matures.

4. APPLICATIONS

Even if all of these research areas could be addressed satisfactorily, one important challenge remains. No information technology will be useful unless it enables new applications, or dramatically improves the way existing applications are built or run. Although the effectiveness of cloud computing has already been demonstrated for some applications, more work should be done on identifying new classes of novel applications that can only be realized using cloud computing technology. With proper instrumentation of potential applications and the underlying cloud infrastructure, it should be possible to quantitatively evaluate how well these application classes perform in a cloud environment. Experimental software engineering research should be conducted to measure how easily new cloud-based applications can be constructed relative to non-cloud applications that perform similar functions. [29]

5. CONCLUSION

In this paper motivation and suggestions for additional research has been provided. As more experience is gained with cloud computing, the breadth and depth of cloud implementations and the range of application areas will continue to increase. Like other approaches to high performance computing, cloud computing is providing the technological underpinnings for new ways to collect, process, and store massive amounts of information. Based on ongoing research efforts, and the continuing advancements of computing and networking technology, cloud computing is poised to have a major impact on our society's data centric commercial and scientific endeavors.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Communications of the ACM, vol. 53, pp. 50–58, April 2010. [Online]. Available:*http://doi.acm.org/10.1145/1721654.1721672*.

[2] M. Creeger, "Cloud computing: An overview," *ACM* Queue, vol. 7, June 2009.

[3] Jericho Forum, "Cloud cube model: selecting cloud formations for secure collaboration," Version 1.0, April, 2009, San Francisco CA, USA: The Open Group. [Online]. Retrieved(Aug.30,2011):http://www.opengroup.org/jerich o/cloud_cube_model_v1.0.pdf

[4] M. D. de Assuncao, A. di Costanzo, and R. Buyya, "Evaluating the cost benefit of using cloud computing to extend the capacity of clusters," in Proceedings of the 18th International Symposium on High Performance Distributed Computing (HPDC '09), Jun. 2009, pp. 141– 150.

[5] "Amazon elastic compute cloud,"http://aws.amazon.com/ec2/.

[6] "Salesforce's force.com cloud computingarchitecture,"http://www.salesforce.com/platform/.

[7] "Google app engine," https://appengine.google.com/.

[8] "Windows azure platform,"http://www.microsoft.com/windowsazure

[9] Distributed Application Architecture". Sun Microsystem. Retrieved 2009-06-16.

[10] Sun CTO: Cloud computing is like the mainframe".Itknowledgeexchange.techtarget.com. 2009-03-11. Retrieved 2010-08-22.

[11] It's probable that you've misunderstood 'Cloud Computing' until now". TechPluto. Retrieved 2010-09-14.

[12] Danielson, Krissi (2008-03-26). "Distinguishing Cloud Computing from Utility Computing". Ebizq.net. Retrieved 2010-08-22.

[13] Recession Is Good For Cloud Computing – Microsoft Agrees". CloudAve. Retrieved 2010-08-22.

[14] Defining 'Cloud Services' and "Cloud Computing"". IDC. 2008-09-23. Retrieved 2010-08-22.

[15] e-FISCAL project state of the art repository".

[16] Farber, Dan (2008-06-25). "The new geek chic: Data centers". CNET News. Retrieved 2010-08-22.

[17] He, Sijin; L. Guo, Y. Guo, M. Ghanem,. *Improving Resource Utilisation in the Cloud Environment Using Multivariate Probabilistic Models*. 2012 2012 IEEE 5th International Conference on Cloud Computing (CLOUD). pp. 574–581. doi:10.1109/CLOUD.2012.66. ISBN 978-1-4673-2892-0.

[18] King, Rachael (2008-08-04). "Cloud Computing: Small Companies Take Flight". *Bloomberg BusinessWeek*. Retrieved 2010-08-22.

[19] Mao, Ming; M. Humphrey (2012). "A Performance Study on the VM Startup Time in the Cloud". *Proceedings of 2012 IEEE 5th International Conference on Cloud Computing (Cloud2012)*: 423. doi:10.1109/CLOUD.2012.103. ISBN 978-1-4673-2892-0.

[20] Dario Bruneo, Salvatore Distefano, Francesco Longo, Antonio Puliafito, Marco Scarpa: Workload-Based Software Rejuvenation in Cloud Systems. IEEE Trans. Computers 62(6): 1072-1085 (2013)[1]

[21] Defining and Measuring Cloud Elasticity". KIT Software Quality Departement. Retrieved 13 August 2011.

[22] Economies of Cloud Scale Infrastructure". Cloud Slam 2011. Retrieved 13 May 2011.

[23] He, Sijin; L. Guo; Y. Guo; C. Wu; M. Ghanem; R. Han. *Elastic Application Container: A Lightweight Approach for Cloud Resource Provisioning*. 2012 IEEE 26th International Conference on Advanced Information Networking and Applications (AINA). pp. 15–22. doi:10.1109/AINA.2012.74. ISBN 978-1-4673-0714-7.

[24] He, Qiang, et al. "Formulating Cost-Effective Monitoring Strategies for Service-based Systems." (2013): 1-1.

[25] A Self-adaptive hierarchical monitoring mechanism for Clouds Elsevier.com

[26] "Encrypted Storage and Key Management for the cloud". Cryptoclarity.com. 2009-07-30. Retrieved 2010-08-22.

[27] Mills, Elinor (2009-01-27). "Cloud computing security forecast: Clear skies". CNET News. Retrieved 2010-08-22.

[28] http://www.nsa.gov/research/tnw/tnw174/articles/pdfs/TNW_17_4_Web .pdf

[29] http://en.wikipedia.org