Emerging Cloud Computing Paradigm

Abu Sarwar Zamani [1], Md. Mobin Akhtar [2] and Sultan Ahmad [3]

**[1] College of Science , Shaqra University,
Al-Quwaiyyah, Kingdom of Saudi Arabia**

**[2] Al-Qiayaha Community College, Shaqra University
Kingdom of Saudi Arabia**

**[3] College of Computer Engineering & Sciences, Al-Kharj University
Al-Kharj, Kingdom of Saudi Arabia**

Abstract

"Cloud" computing – a relatively recent term, defines the paths ahead in computer science world. Being built on de cades of research it utilizes all recent achievements in virtualization, distributed computing, utility computing, and networking. It implies a service oriented architecture through offering software and platforms as services, reduced information technology overhead for the end-user, great flexibility, reduced total cost of ownership, on de mand services and many other things. This paper is a brief survey based on readings of "cloud" computing and it tries to address related research topics, challenges ahead and possible applications.

Keywords: *Cloud Computing, Cyber infrastructure, Virtualization.*

1. Introduction

Cloud computing is the next generation in computation. Maybe Clouds can save the world; possibly people can have everything they need on the cloud. Cloud computing is the next natural step in the evolution of on-demand information technology services and products. The Cloud is a metaphor for the Internet, based on how it is depicted in computer network diagrams and is an abstraction for the complex infrastructure it conceals. It is a style of computing in which IT-related capabilities are provided "as a service", allowing users to access technology-enabled services from the Internet (i.e., the Cloud) without knowledge of, expertise with, or control over the technology infrastructure that supports them. Email was probably the first service on the "cloud". As the computing industry shifts toward providing Platform as a Service (PaaS) and Software as a S ervice (SaaS) for consumers and enterprises to access on demand regardless of time and location, there will be an increase in the number of Cloud platforms available. But it s eems that Cloud computing cannot save the universe. Cloud computing cannot run for President. Cloud computing is a very specific type of computing that has very specific benefits. But it has specific negatives as well. And it does not serve the needs of real businesses to hear only the hype about cloud computing – both positive and negative. One thing that is hoped to be accomplished with this paper is not only a clear picture of what the cloud does extremely well and a brief overview of them, but also a short survey on their criteria and challenges ahead of them.

2. Background

2.1 Cyber infrastructure

"The comprehensive infrastructure needed to capitalize on dramatic advances in information technology has been termed cyberinfrastructure".The term "cyberinfrastructure" describes the new research environments that support advanced data acquisition, data storage, data management, data integration, data mining, data visualization and other computing and information processing services over the Internet. Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked by high speed networks to make possible scholarly innovation and discoveries not otherwise

possible. Cyberinfrastructure is a term first used by the US National Science Foundation (NSF), and it typically is used to refer to information technology systems that provide particularly powerful and advanced capabilities. In scientific usage, cyberinfrastructure is a technological solution to the problem of efficiently connecting data,computers, and people with the goal of enabling derivation of novel scientific theories and knowledge."Cyberinfrastructure makes applications dramatically easier to develop and deploy, thus expanding the feasible scope of applications possible within budget and organizational constraints, and shifting the scientist's and engineer's effort away from information technology development and concentrating it on scientific and engineering research. Cyberinfrastructure also increases efficiency, quality, and reliability by capturing commonalities among application
needs, and facilitates the efficient sharing of equipment and services".

2.2 Workflows

A workflow is a depiction of a sequence of operations, declared as work of a person, work of a simple or complex mechanism, work of a group of persons, work of an organization of staff, or machines. Workflow may be seen as any abstraction of real work, segregated in workshare, work split or whatever types of ordering. For control purposes, workflow may be a view on real work under a chosen aspect, thus serving as a virtual representation of actual work. The flow being described often refers to a document that is being transferred from one step to another. A workflow is a model to represent real work for further assessment, e.g., for describing a reliably repeatable sequence of operations. More abstractly, a workflow is a pattern of activity enabled by a systematic organization of resources, defined roles and mass, energy and information flows, into a work process that can be documented and learned. Workflows are designed to achieve processing intents of some sort, such as physical transformation, service provision, or information processing. A workflow can be represented by a directed graph that represents data- flows that connect loosely and tightly coupled (and often asynchronous) processing components. One such graph is shown in Figure 1. It illustrates a Workflow as a part of Experiment Builder in Lead Project.

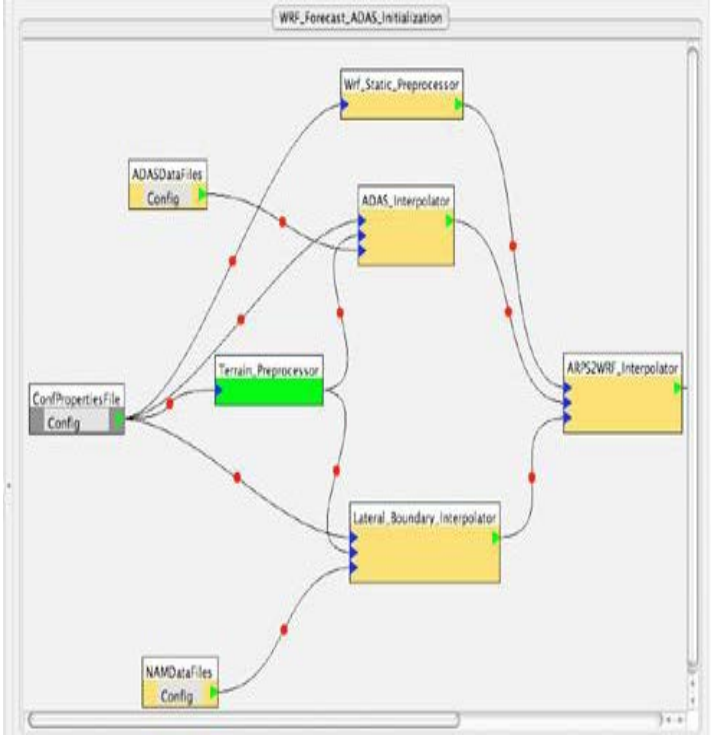

Fig. 1 WRF Forecast ADAS Initialization.

2.3 Virtualization

Virtualization is a framework or methodology of dividing the resources of a computer into multiple execution environments, by applying one or more concepts or technologies such as hardware and software partitioning, time-sharing, partial or complete machine simulation, emulation, quality of service, and many others. It allows abstraction and isolation of lower-level functionalities and underlying hardware. This enables portability of higher-level functions and sharing and/or aggregation of the physical resources. There are lots of virtualization products, and a number of small and large companies that make them. For instance, in the operating systems and software applications space are VMware1, Xen - an open source Linux-based product developed by XenSource2, and Microsoft virtualization products, can be mentioned.

3. Cloud Computing

Cloud computing is a paradigm that focuses on sharing data and computations over a scalable network of nodes. Examples of such nodes include end user computers, data centers, and Web Services. We term such a network of nodes as a cloud. An application based on such clouds is taken as a cloud application. Basically cloud is a metaphor for internet and is an abstraction for the complex infrastructure it conceals. The main idea is to use the

existing infrastructure in order to bring all feasible services to the cloud and make it possible to access those services regardless of time and location. Whether it's called Cloud Computing or On-demand Computing, Software as a Service, or the Internet as Platform, the common element is a shift in the geography of computation. When you create a spreadsheet with the Google Docs service, major components of the software reside on unseen computers, whereabouts unknown, possibly scattered across continents. The shift from locally installed programs to cloud computing is just getting under way in earnest. Shrink-wrap software still dominates the market and is not about to disappear, but the focus of innovation indeed seems to be ascending into the clouds. Some substantial fraction of computing activity is migrating away from the desktop and the corporate server room. The change will affect all levels of the computational ecosystem, from casual user to software developer, IT manager, even hardware manufacturer.

4. Current Works

Currently there are various cloud systems on both academic and industrial world are being built. Following is a brief review of what is undergoing presently.

4.1 Academia

4.1.1 Eucalyptus

Eucalyptus is an open source software framework developed by University of California – Santa Barbara for cloud computing that implements what is commonly referred to as Infrastructure as a Service (IaaS); systems that give users the ability to run and control entire virtual machine instances deployed across a variety physical resources. [19] The current interface to Eucalyptus is compatible with Amazon's EC2 interface, but the infrastructure is designed to support multiple client-side interfaces. EUCALYPTUS is implemented using commonly available Linux tools and basic Web-service technologies making it easy to install and maintain.

4.1.2 Nimbus

The University of Chicago Science Cloud, codenamed "Nimbus", provides compute capability in the form of Xen virtual machines (VMs) that are deployed on physical nodes of the University of Chicago TeraPort cluster (currently 16 nodes) using the Nimbus software. The Nimbus cloud is available to all members of scientific community wanting to run in the cloud.Nimbus supports both WSRF and EC2 interfaces and it can be configured to use familiar schedulers like PBS and SGE to manage VMs.

5. Challenges Ahead

One of the most important challenges ahead is that clouds will always be compared to local machine in the time of usage. It's important for the user to know what he gains of shifting to the cloud. Obviously using services on local machines, the user needs more resources but at least he knows that he has access to his data all the time and he has the data he owns on his local machine. But who is in charge of restoring his data if something happens to the cloud and the fact that the user is not aware of the physical place which his data is stored makes cloud more unreliable for him. Here is a list of issues that cloud computing is currently facing.

5.1 Information Policy

Cloud computing raises a range of important policy issues, which include issues of privacy, security, anonymity, telecommunications capacity, government surveillance, reliability, and liability, among others. At a minimum, users will likely expect that a cloud will provide:

Reliability and Liability: Users will expect the cloud to be a reliable resource, especially if a cloud provider takes over the task of running "mission-critical" applications and will expect clear delineation of liability if serious problems occur.

Security, privacy, and anonymity: Users will expect that the cloud provider will prevent unauthorized access to both data and code, and that sensitive data will remain private. Users will also expect that the cloud provider, other third parties, and

Access and usage restrictions: Users will expect to be able to access and use the cloud where and when they wish without hindrance from the cloud provider or third parties, while their intellectual property rights are upheld.

Here the most important issue is security and the way that the provider has to assure the user of providing it. Also one of the most important aspects of cloud in which academia is more interested is high performance and adding securing will always reduce performance. Thus there is a need to find a way of implementing security with the least effect on performance.

5.2 Provenance Data

Cloud provenance data, and in general meta-data management, is an open issue. Open challenges include: How to collect provenance information in a standardized and seamless way and with minimal overhead – modularized design and integrated provenance recording; How to store this information in a permanent way so that one can come back to it at anytime, - Standardized schema; and How to present this information to the user in a logical

manner – an intuitive user web interface. There are also issues like the scalability, portability of services, cloud interactions, interoperability, fault tolerance, energy cost and the cost of building clouds versus keeping the current systems. Given the proliferation of different virtualization environments, and the variety in the hardware, standardization of image formats is of considerable interest. Some open solutions exist or are under consideration, and a number of more proprietary solutions are here already.

6. Conclusion

Cloud computing is an emerging computing paradigm that is increasingly popular. Leaders in the industry, such as Microsoft, Google, and IBM, have provided their initiatives in promoting cloud computing. However, the public literature that discusses the research issues in cloud computing are still inadequate. In a study of the research literature surrounding cloud computing, I found that there is a distinct focus on the needs of the scientific computing community. Big IT companies are also building their own version of cloud. But still there are many question have left without an answer and indeed the most important one is security. One of the other aspects of the cloud which is left is the social aspect of it. The Cloud is going to happen but which services should be offered on the cloud and for whom. What happens if smaller IT companies start to offer their services on the cloud and no one uses them?! I believe that everything eventually can move to the Cloud. The question is if users are ready for that and if it's the right move and this need must be addressed.

References

[1] An Introduction to Virtualization, http://www.kernelthread.com/publications/virtualization

[2] Lijun Mei, W.K. Chan, T.H. Tse, "A Tale of Clouds: Paradigm Comparisons and Some Thoughts on Research Issues", To appear in Proceedings of the 2008 IEEE Asia-Pacific Services Computing Conference (APSCC 2008), IEEE Computer Society Press, Los Alamitos, CA

[3] M.A. Vouk, "Virtualization of Information Technology Resources", in Electronic Commerce: A Managerial Perspective 2008, 5th Edition y Turban, Prentice-Hall Business Publishing, to appear.

[4] Mike P. Papazoglou, "Service -Oriented Computing: Concepts, Characteristics and Directions", Tilburg University, INFOLAB

[5] "NSF'S Cyberinfrastructure Vision for 21st Century Discovery, NSF Cyberinfrastructure Council, September 26th, 2005, Ver.4.0, pg 4.

[6] R. Buyya, C. S. Yeo, and S. Venugopa, "Marketoriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities" In Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08, IEEE CS Press, Los Alamitos, CA, USA) 2008.

Abu Sarwar Zamani was born on 15th Jan 1982 in Bihar, India. He is working as a Lecturer in College of Science in Shaqra University, Kingdom of Saudi Arabia. He has received his Master of Science in Computer Science from Jamia Hamdard (Hamdard University), New Delhi, India in the year 2007, and later he did Master of Philosophy in Computer Science from Vinayak Mission University, Chennai, India in 2010. He is member of International Association of Computer Science and Information Technology-IACSIT. He has actively attended and published various research papers in National as well as International papers. His current area interests are in Cloud Computing, Networking, Artificial Intelligence, and Data Mining. Apart from that I worked in the field of Software Development in MNC companies.

Md. Mobin Akhtar has received his Master of Science Technology from Jamia Millia Islamia University, Delhi India in 2008 with distinction marks. He has done his B.Sc. (Hons.) in Mathematics in 2003 from Vinoba Bhave University, India. He is currently working as Lecturer in Computer Science & info. System Departement, Community College in ALQuwaiyyah, Shaqra University, Kingdom of Saudi Arabia. His research interests include Cloud Computing, Software Engineering and Data mining. Before starting his teaching career, He spent two years at S.G Innovative Noida (India) working on the XSLT, Regulation Expression and Qube Activity tools. His current interests are in Computer Organization, Artificial Intelligence, and Data Mining. Besides teaching, he enjoys giving tech talks, reading biographies, reading world history and reading about new technology articles on internet. His teaching abilities include innovative skills and extensive use of technology in teaching.

Sultan Ahmad has received his Master of Computer Science and Applications from Aligarh Muslim University, India in 2006 with distinction marks. He has done his B.Sc. (Hons.) in Computer Application in 2002 from Patna University, India. He is currently working as Lecturer in College of Computer Engineering and Sciences in Al-Kharj University, Al- Kharj, Saudi Arabia. His research interests include Cloud Computing, Software Engineering and Knowledge Management. He has presented his research papers in many national and International Conferences & journals. He has a good experience in software development specially for Windows USB Device Driver. He has worked in various administrative capabilities including software / network implementation and testing, customization & training during last 5 years. Besides teaching, he enjoys giving tech talks, reading biographies, reading world history and reading about new technology articles on internet. His teaching abilities include innovative skills and extensive use of technology in teaching.