Sindh Univ. Res. Jour. (Sci. Ser.) Vol.45 (1): 137-140 (2013)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

RapidCompute for Research Community

F. N. MEMON++, Z. U. A. KHUHRO, I. A. KOREJO, M. U. R. MAREE

Institute of Mathematics and Computer Science, University of Sindh, Jamshoro

Received 17th December 2012 and Revised 20th February 2013

Abstract: Cloud computing provides a way to use computing resources as a utility on pay-as-you-go bases. It is getting popular day by day in almost every industry mainly due to two of its benefits - accessibility from anywhere and no investment on hardware resources. From a large organization to an individual, everyone can be benefited from the cloud computing to perform different activities from business to research. Cloud computing is also expected to be a great support in research particularly when high power computing resources are required for some time/experiments but not all the time. Instead of buying/upgrading existing resources for a few experiments, it is better to use cloud for those tasks. Cloud computing provides elasticity so a researcher may extend or shrink the resources as per experiment's needs and pay for those resources only for the time they were used.

Many companies provide their clouds that include Amazon EC2, Microsoft Azure, etc. Recently, a Pakistani company, CyberNet, has introduced the first Pakistani cloud, the RapidCompute, which mainly focuses corporate market. This paper presents the comparison of RapidCompute with the well-known Amazon EC2 particularly for individuals or research activities. It is proposed to the CyberNet to consider the individuals particularly the research community of Pakistan that could motivate them to use Cloud for research activities.

Keywords: Cloud Computing, RapidCompute, Amazon EC2.

1. INTRODUCTION

Cloud computing gives a concept of a pool of computing resources that can be use from anywhere through Internet and on pay-as-you-go bases. Cloud computing gives the concept of providing computing as a utility (Buyya, 2009). From an individual to a big organization, this technology can help everyone. The use of computing resources as a utility is mainly due to the two reasons: (1) rapid growth of processing and storage technologies and (2) success of the Internet (Zhang, 2010). Cloud computing is discussed in detail in many articles (Armbrust, 2010; Chandio, 2013, Wang, 2010, Mell, 2011).

Cloud computing is now considered as a feasible solution that provides an on-demand and flexible computing infrastructure for a large number of applications almost in every field of life including Healthcare (Shimrat, 2009; Rolim, 2010), Astronomy (Hoffa, 2008), Government (Pokharel, 2009), Video transcoding (Chandio, 2013), Education (Ercan, 2010) etc. Its flexibility enables users to extend or shrink the resources as per requirement and pay for those resources only for the time they were used. It has been observed that this technology is also getting popular amongst researchers (Memon, 2010; Memon, 2012).

Cloud computing is a fine solution to use high power computing resources without any investment on buying them. Not only this, users are also free from maintaining these resources. The portability, flexibility, elasticity and other benefits of cloud computing are attracting not only the big organizations but individuals too.

The usage of cloud computing by individuals/researchers could be different than the business industry. Such as for a particular activity or may be for a smaller task of that activity researchers need high power computing resources but might not prefer long-term commitment to use a cloud service.

The research trend in Pakistan is increasing in recent days, for example, research in the field of Bioinformatics is getting popular and recently, the first Pakistani DNA sequence has been mapped (Suhail, 2011). Since the biological data expending very quickly, the state-of-the-art processing and storage technologies are needed to interpret that data. We have earlier explored Amazon Cloud (Memon, 2010; Memon, 2012) to analyse genomic sequences and found it very feasible in situations where computing resources limits our research activities.

++Corresponding author: F. N. Memon, Email: farhatnaureen@gmail.com

RapidCompute is claimed to be the first Pakistani cloud computing service that has been introduced recently by CyberNet; however, they focused on corporate market (Shahid, 2012). The vice president (corporate sales) of CyberNet mentioned, "Companies no longer have to invest in installing servers, air conditioning for their hardware, power setups or engage manpower to run IT related business". From the statement of vice president it is clear that that the company does not focus on individuals. This paper does not discuss the cloud computing in general but it shows the comparison of RapidCompute with one of the well known clouds, the Amazon Elastic Computing Cloud (EC2).

The rest of this paper is as follow. The next section describes the method used to compare the two clouds. The results and discussion are presented in section III. Finally, section IV provides the conclusion.

2. MATERIALS AND METHOD

A similar type of experiment on the two clouds, RapidCompute and Amazon EC2, can help in comparing them. We have already analysed human genomic sequence on Amazon EC2 (Memon, 2012) and the same experiment can be repeated on RapidCompute. The complete description to analyse human genomic sequence can be referred from (Memon, 2012).

The results must be same as the same algorithm and same data are used; however, a number of parameters are considered to compare RapidCompute with Amazon EC2. These parameters include cost/charges of the services, payment method/long-term commitment, Security, public datasets and freedom to design virtual machine image. To compare the two clouds, first of all it is required to select the instance types of the two clouds with same specifications. Hence, we are using small instance type (m1.small) of Amazon EC2 and RC-2G machine of RapidCompute, which is roughly equated to the m1.small instance, to presents the entire comparison.

Cost:

A user has to pay $96.14 for a month to use RC-2G machine for Ubuntu platform. According to this a user pays $0.13 for an hour whereas Amazon charges maximum $0.088 for an hour (depending on the zone) that means if a user uses their machine for 24 hours in a month, it will cost $63.36 which is one third less then the amount charged for RapidCompute.

Similarly Windows users have to pay $116.14 per month ($0.16 per hour on RapidCompute) while Amazon charged $108 for a month (or $0.15 per hour).

Payment Method/Long-term commitment:

Amazon does not require long-term commitment from its users. It works on post-paid method while RapidCompute is pre-paid service. A user of Amazon cloud pay on hourly basis and this payment is deducted after the use of the service. They do not require any extra amount to deposit to use their services.

RapidCompute is a pre-paid service. Depending on the requirement of the service a user plans to use, it first charged the amount of that service for the entire month. This could be suitable for a business organization that is completely based on cloud services but not suitable for the users who need cloud services where their resources limit them to work and they need the services for small period of time. Although the company the RapidCompute charges on hourly basis and returns the remaining balance to the users but it bounds a user to have an amount reserved for a month that is sufficient for a month to use the machine for example for an hour.

Security:

While using Amazon, an independent virtual machine is provided to each user that guarantees high level security (Khan, 2011). CyberNet also claims that their customers are provided with a unique VLAN to protect users' data. Both clouds take the support of Hypervisor.

Freedom:

Users of Amazon cloud are free to create their own AMI (Amazon Machine Image) that can be understood as a file containing everything required to run a job. Users can install any software if required. However, RapidCompute does not allow it users to create their AMI instead it can be done by the company on the request of a user.

Public Datasets:

Use of public datasets is one of the main reasons that motivated us to explore Amazon's cloud. Among the others, some well-known and widely used biological datasets are available at Amazon's cloud as public datasets that can be used free of cost. It also saves uploading/downloading time of that datasets from/to Amazon's S3. So, the researchers particularly bioinformaticians who require biological databases for their research get benefited by the Amazon's cloud**.**

3. RESULTS AND DISCUSSION

It is obvious to see no difference in the results **(Table 1)** because of the use of same algorithm and same data. However, the aim is to compare the above-mentioned features of the two clouds with respect to an individual's perspective.

Table 1: The chromosome name/number, number of possible G-quadruplex forming sequences (PGQFS), size of chromosomes (in bases), and processing time to find PGQFS in each chromosome of human are listed.

Chromosome	No. of Bases in chromosome	No. of possible G-quadruplex forming sequences
1	247,249,719	175,131
2	242,951,149	149,051
3	199,501,827	105,700
4	191,273,063	86,836
5	180,857,866	99,071
6	170,899,992	91,671
7	158,821,424	108,987
8	146,274,826	83,421
9	140,273,252	91,570
10	135,374,737	95,383
11	134,452,384	99,129
12	132,349,534	98,393
13	114,142,980	46,857
14	106,368,585	62,863
15	100,338,915	59,306
16	88,827,254	92,806
17	78,774,742	93,204
18	76,117,153	38,873
19	63,811,651	108,865
20	62,435,964	62,214
21	46,944,323	28,044
22	49,691,432	49,044
X	154,913,754	115,767
Y	57,772,954	9,150

Cloud computing is getting admired by the researcher for their research activities/experiments. It is feasible particularly when computing resources are required for some times but not all the time. In these cases, investment on buying the high power computing resources is not a good idea where they are idle most of the time.

The research community is growing rapidly in Pakistan, so the RapidCompute can get attention of the research community that could be either individual researchers or those who are associated with certain organizations. In later case, researchers may have better financial support from their organizations than individual researchers.

Although RapidCompute charge a bit more for an hour but it is expected that these charges will reduce by time as it happened with Amazon's cloud.

However, to use a machine on RapidCompute, charges of that machine for a month is initially paid to the company even if the machine is required for a short period of time, an hour for example. Although the remaining amount can be returned to the customer but a customer cannot use RapidCompute if s/he has enough amounts to use a machine for the required time but not for a month. At some instance, a user may bear increased charges (on hourly basis) but may not bear to hold an amount for a month for a small service.

Users may also feel restricted if they are not authorized to create their AMI by their own.

It is suggested to the RapidCompute provider to consider the individuals or research community in a way to facilitate them as much as possible to get them attracted towards RapidCompute. By ignoring them, the company may lose a large number of customers. By considering the researchers/individuals it is expected that the research community can significantly enhance the number of users of RapidCompute.

4. CONCLUSION

Cloud computing is being adopted in every type of application. This paper presented a brief overview of cloud computing in general and RapidCompute in particular. RapidCompute is the first Pakistani; claimed by the company. Although it is a good initiative to build a Pakistani cloud however, the company has focused on corporate market. Hence a couple of issues including cost, payment method, and freedom to create users own AMI, can affect the attention of the individuals that can also include research community of Pakistan. RapidCompute is compared with a well-known cloud service (Amazon EC2). RapidCompute is found to be a bit expensive as compared to the Amazon EC2. Secondly, users have to pay for a month even if s/he wants to use the services for an hour. These limitations can stop the individuals to use RapidCompute. It is therefore suggested that the provider shall look at these issues to grab a large number of individuals and research community.

REFERENCES:

Armbrust, M., A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and others, (2010) A view of cloud computing, *Communications of the ACM*, **(53):** 4, 50-58.

Buyya, R., C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, (2009) Cloud computing and emerging IT

platforms: Vision, hype, and reality for delivering computing as the 5th utility, *Future Generation computer systems*, **(25):** 6, 599-616.

Chandio A. A., I. A. Korejo, Z. U. A. Khuhro, F. N. Memon, (2013) Clouds Based Smart Video Transcoding System, Accepted for publication in *Sindh University Res. Jour. (Science Series)*,**(45):** 1, (In Press).

Ercan, T., (2010) Effective use of cloud computing in educational institutions, *Procedia-Social and Behavioral Sciences*, **(2)** 2, 938-942.

Hoffa, C., G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman, and J. Good, (2008) On the use of cloud computing for scientific workflows, *Proceedings: IEEE Fourth International Conference on eScience (eScience'08), IEEE*, 640-645.

Khan, N., A. Noraziah, M. M. Deris, and E. I. Ismail, (2011) Cloud Computing: Comparison of Various Features, *DEIS*, *Springer-Verlag Berlin Heidelberg*, 243-254.

Mell, P., T. Grance, (2011) The NIST definition of cloud computing, *NIST special publication*, 800.

Memon, F. N., A. M. Owen, O. Sanchez-Graillet, G. J. G. Upton, and A. P. Harrison, (2010) Identifying the impact of G-Quadruplexes on Affymetrix 3' Arrays using Cloud Computing, *Journal of Integrative Bioinformatics*, **(7):** 2, 111-112.

Memon, F. N., Z. A. U. Khuhro, M. U. R. Maree, I. A. Korejo, A. P. Harrison, (2012) Analysis with special reference to genomic sequence on Amazon Cloud, *Sindh University research journal (Science Series)*, **(44):** 3, 441-444.

Pokharel, M. and J. S. Park, (2009) Cloud computing: future solution for e-governance, *Proceedings of the 3rd international conference on Theory and practice of electronic governance, ACM*, 409-410.

Rolim, C. O., F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi, and G. S. Salvador, (2010) A cloud computing solution for patient's data collection in health care institutions, *Proceedings: Second International Conference on eHealth, Telemedicine, and Social Medicine (ETELEMED'10), IEEE*, 95-99.

Shahid, J., (2012) Rapid Compute launched in Pakistan, *Dawn.com, SCI-TECH* (30th May, 2012).

Shimrat, O., (2009) Cloud computing and healthcare, *San Diego Physician. Org. Magazine*, 26-29.

Suhail Y., (2011) Genome mapping of first Pakistani, Dawn.com (1st July, 2011).

Velte, T., A. Velte, and R. Elsenpeter, (2009) Cloud Computing, A Practical Approach, *McGraw-Hill Osborne Media, New York, USA*, 1st edition.

Wang, L., G. V. Laszewski, A. Younge, X. He, M. Kunze, J. Tao, and C. Fu, (2010) Cloud computing: a perspective study, *New Generation Computing*, **(28):** 2, 137-146.

Zhang, Q., L. Cheng, and R. Boutaba, (2010) Cloud computing: state-of-the-art and research challenges, *Journal of Internet Services and Applications*, **(1):** 1, 7-18.