Evaluation of Server Performance in Cloud Computing
On Virtual Network Infrastructure: Using Riverbed Simulator

Charles Amo-Asuah
T. I Ahmadiyya SHS
Kumasi, Ghana

Isaac Kofi Nti
Electrical/Electronic Engineering Department
Sunyani Technical University
Sunyani, Ghana

Samuel Akyeramfo-Sam
Department of Computer Science
Sunyani Technical University
Sunyani, Ghana

Abstract: Computing in the cloud changes the way Information Technology (IT) is managed and disbursed, given an improved cost-efficiencies, faster novelty, quicker time-to-market, and the capacity of scaling applications on request (Arora & Gupta, 2012). Most organizations are now opting for web based services, since the use of virtual networking infrastructure in cloud storage does not only reduce costs, eliminates hardware failures and security risks such as theft of physical components in the real world, but also provides a graphical user interface for the topology design allowing for realistic simulation of networks and high performance displaying modules as an added advantage for modern enterprises. This research focuses on examining the performance and relations of distributed systems in servers used for cloud storage on virtual platforms, by analyzing the behavior and efficiency of server performances in three different scenarios in a cloud network using Riverbed Edu 17.5 edition as a virtual network platform simulation tool. Networks with three different server scenarios simulated against time in a network were modelled. The focus was on the server performance, at different loads and processing speeds. The simulation was configured for about 100 workstations in a manner that all the applications and users can access the parameters; database, file transfer protocol, hypertext transfer protocol, and email response times on cloud utilization throughputs in two directions. The results for server performance in efficiency and viability is analyzed and evaluated with conclusions drawn as a guiding principle for organizations and companies considering the use or using servers for cloud storage.

Keywords: Cloud-Computing; Networking; Servers; file-transfer-protocol; cloud-utilization; throughputs; email-response-time; Virtualization

1. INTRODUCTION
Computers have grown with rapid changes in technology, and they have become increasingly faster than ever to the extent that users of microcomputers are said to be under utilizing the potentials of their computer systems both at home and in the offices or work place. This paper centers on virtual network presented as part of cloud service as a model of virtual resource interconnect, where resources are dynamically deployed and migrated in a cloud. Visualization stands for pulling out hardware from existing software. The process of extraction can be done in many forms, by the use of operating systems such as Hyper-V, creating visualizations via terminal systems or by means of App-V and with establishments such as Hewlet Packard and Cisco which runs both network and visualization storage. Organization globally have incorporated the visualization operating systems in the workstations to augment the delivery and the performance of their day to day activities (Desai, 2012). Virtualization can be defined as the act of designing the real virtual multiplicity of an item such as device storage, operating system (OS), resource network and hardware platform with the physical outlook of a computer representing a whole machine incorporating and encompassing both subjective and objective point of view in a non-questionable manner (Turban, et al., 2008). Subjectively, the virtual electronic gadget is whole or complete whereas from an objective perspective, they are just a collection of files which operates on a physical machine. Globally, the virtualization technology is widely accepted and has become part of daily life activities, owing to its attractiveness. The virtualization technology operational techniques permits easy distribution of information between distinct virtual machines (VM), by this means enhancing the maximum use of hardware’s. The flexibility with the use of virtual machines makes its management easier and has also spawned new models of efficiently operating the system (Koponen, et al., 2014). In a more sensitivity view, virtualization is vital in operating cloud computing minus which cloud computing would appear anomalous. Virtualization has made IT cheaper and a global commodity. The solid foundations of cloud computing is built on the prowess of virtualization which has cost effective use of the model with emphasis on scalability and the rate at which the service is demanded (Turban, et al., 2008). The cloud system provides a common platform for easy access, managing and operating resources. Also cloud computing constitute the tendency to use scalable and shared IT platform within the parameters of an internet (Kaufman, 2009). Virtualization technology is the heart of cloud computing. Unlike other archetypes, cloud computing is yet to gain a common grounds regarding its definition due to the fact that its parameters are frequently varying and increasing. However, the US National Institute of Standards and Technology (USNIST) have provided a concise definition as; “Cloud computing is an IT architecture that provide more convenient services to user on their computers which requires minimal supervision from provider to operate (Mell & Tim, 2009). There is the existence of resource sharing among virtual networks. This means that if each product draws on separate protocols, there is the need for discrimination to maintain integrity within the cloud system. Moreover, in order to deliver quality service, adapting bandwidth is important due to the fact that different approaches are required for different bandwidth. There are some IT developers such as Apple, Google and among other who have develop their own cloud systems to enhance user preference. With the problems associated in the use and adoption of traditional methods of network storage and hosting, this paper seeks to assess the technologies, advantages, cost-efficacy and deployment competence of virtualization infrastructure employed in cloud computing systems in today’s computing over traditional platforms.
2. TYPES OF VIRTUALIZATION
There are numerous types of virtualization in computing but this paper seeks mention and explain two of them.

2.1 Server Virtualization
Virtualization is the main and primary accelerator of cloud-computing and cloud-storage (Arora & Gupta, 2012). Virtualization permits the run of multiple applications on a virtual machine within the same physical server as a substitute to running only one application on a physical server. It enable delivering of greater accessibility, scalability, and optimization of resources such as servers and storage (Arora & Gupta, 2012). Carefully analysing Arora & Gupta definition of server virtualization, one can say that the server is said to have so many resources that are available for users to use. Not ever said resources is easily and user-friendly for the virtual IT environs when serving multiple isolated virtual environment with single server.

2.1.1 Types of Server Virtualization
(Baburajan, 2011) states that, in order to virtualize resources or materials, there should be virtual machine (VM) models, OS stage virtualization as well as Para-virtual machine models. The research analyzed and chose amongst alternatives. There are different types of server virtualizations on the IT markets. Management has to decide on the one which best suits the institution’s processes in order to achieve the much needed efficiency gains from server virtualization most importantly on the cost saving advantage that comes along with virtualizing systems.

2.1.2 VirtualMachine Based Virtualization
The virtual machine virtualization server runs on virtual machine monitor or can also use the hypervisor in transmitting command to and receiving information from the CPU and this is done without any extra modifications. Subsequently, VM (Virtual Machine)-based Server Virtualization give protection to the hardware layer, partition can be done within the server while running on separate operating systems without any guest application systems and knowing that they are all running on a mother OS (Saunders, et al., 2009)

2.2 OS Level Virtualization
Operating System level Virtualization is attained through the mother OS running on an isolated system and through the management and controlling of guest OS functionalities. This means that virtual level operating systems draws from the creation of segregated containers in unified physical servers while using each segregated entities to work separately of the partitions (Baburajan, 2011).

Types of Desktop Virtualization
For this research purpose four types of VDI operation are stated namely;

- Hosted (mostly used in the form of a server)
- Centralized
- Remote Synchronization
- Client-Hosted

2.3 Benefits and Drawbacks of Virtualization

- There are numerous benefits to be derived from server and desktop virtualization:
  - It reduces solid containers (servers), hence cutting down cost and maintenance overheads.
  - More space are created with an environment with less physical hardware’s.
  - Spontaneous performance of multiple task on one platform.
  - Risk of losing data is less.
  - Allowing employees to work using a common platform.

Drawbacks

- In desktop Virtualization, when the server goes down all users/client can’t work.
- There is lack of user privacy, thus since all the processing is done by probably less or single servers remotely, everything or any task that a connected terminal does can easily be monitored.
- Sometimes users are restricted as to the number of tasks they can perform on the terminals or virtualized systems.

2.4 Cloud Technology and Cloud Computing
The term cloud computing is a new paradigm in the field of IT that has attracted attention from researchers, organizations such as the banking industry, educational institutions, SMEs among others (Luis, et al., 2008). This model talked about is expanding at a faster rate and attracting various vendors and clients. The faster expansion of cloud computing is being geared by the emanating technologies which gives a clear and rational price use of computing structures and aggregate storage capacity. It lessens the need for huge investment in information Technology infrastructure. This computing model implicates procuring of computer properties with the ability of increasing resource capability of being scaled, on-demand supplying of less no leading IT cost of investing in infrastructure. The advantages in cloud computing is offered through three delivery models called infrastructure-as-a-service (IaaS), Platform-as-a-service (PaaS) and software-as-a-service.
2.5 Cloud Computing Services
Cloud computing services can be grouped into three categories:

2.5.1 Software as a Service (Saas)
In delivering cloud computing as a service, users patronize the right to access the service for being hosted on the cloud system. Currently, Saas is been offered to individuals and organizations by IT giants such as Google, Microsoft, Sales force and among others. The system gives responsibility to users to patch and maintain the OS and the applications software.

2.5.2 Platform as a Service (Paas)
Paas offers cloud users the opportunity to access platforms as well as problem resolving stacks such as programming language, database, web servers and operating systems. Cloud users have the liberty to develop their own applications which operates on the cloud server. Among the popular Paas used are Google App engine, Force.com and others.

2.5.3 Infrastructure as a Service (Iaas)
Iaas offers fundamental storage and computing, which is done to standards as against the provision of network. Organizations and individual users are allowed to maintain and control their own applications, issues regarding storage, operating systems and network connections. Examples of Iaas are as follows; Amazon, GoGrid, 3 Tera etc.

3. MATERIALS AND METHODOLOGY
Riverbed Edu 17.5 was used as a simulation tool

Riverbed Edu version 17.5 was used as a simulation tool, it offers virtualization for development, analyzing and evaluating the capacity and performance of virtual systems which includes application servers and network topologies.

4. Results and Discussions
Three different scenarios were implemented with the simulation tool and tested for an hour on 30mb/s, 2gb/s and 10gb/s loads for 100mb/s bandwidth in each case of the experiments.

In the first scenario, the server was used in the network with four applications and specified number of users (10-60) passing through the cloud and server to generate the required traffic across the system. The second and third instance, the servers were two and three respective with the same applications and users to generate the required traffic across the distributed system.

The email, ftps web application and database executions were assessed and the output parameter measured for all three levels, cloud/router measurement and connection measurement respectively. Each diagram acquired was analysed against the execution measurements and a critical scrutiny was done.

4.1 Database Application Results
The database application generated a lot of traffics when measured against the response time

Database Query Response Time:- This is the slipped by time between the end of a request, inquiry or interest on database server and the start of a request.

4.2 Database Query Response Time–One Server
The tables 1 and 2 shows the database query response time when one server was implemented and figure 2 gives the combined graph in Figure 3 showing the graphical representation and performance of the one server scenario.

<table>
<thead>
<tr>
<th>Table 1 Database query response time with packed size of 50MB (low)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario s</td>
</tr>
<tr>
<td>Single server</td>
</tr>
<tr>
<td>0.26</td>
</tr>
<tr>
<td>Double server</td>
</tr>
<tr>
<td>0.25</td>
</tr>
</tbody>
</table>

Figures:
- Figure 1 Cloud Server on virtual desktop
- Figure 2 Riverbed Edu 17.5 as a Simulation Tool
Table 2 Database query response time with packet size of 250MB (High)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10min</td>
<td>50min</td>
<td>10min</td>
</tr>
<tr>
<td>Single server</td>
<td>1.94</td>
<td>2.36</td>
<td>2.29</td>
</tr>
<tr>
<td>Double server</td>
<td>4.43</td>
<td>6.79</td>
<td>4.43</td>
</tr>
<tr>
<td>Triple server</td>
<td>2.42</td>
<td>2.28</td>
<td>2.61</td>
</tr>
</tbody>
</table>

Figure 3 Database Query Response Time–One Server Scenario

4.3 Database Query Response Time–Two Server Scenario
Tables 1 and 2 row 4 represents the performance of the two server scenario, and the combined graph in Figure 4 shows the graphical representation and performance of the two server scenario.

4.4 Database Query Response Time–Three Server Scenario
In the third scenario, another server was introduced. The tables 1 and 2 Row 5 shows the response time when other applications were diverted to other servers on the network by the LAN.

4.5 Server Database Query Load
Server database load is the time taken for the database server to receive data from users across the network, process it back and forth across the cloud network.

4.6 Server Database Query Load–One Server Scenario
The values in tables 3.3–3.4 showed the server loads across the network when only one server was used. Refer to the combined graph in Fig4.4 for the graphical representation and performance of the one server scenario.

4.7 Server Database Query Load–Two Server Scenario
The result in tables3.3–3.4 showed the load on the server when two servers are used on the network. The combined graph in Fig4.4 showed the graphical representation and performance of the one server scenario.

4.8 Server Database Query Load–Three Server Scenario
In the third scenario, all the four applications were passed through their destination servers and the load analyzed and evaluated on the cloud network. The tables 3 and 4 showed how the loads performed in the network.
Table 3 Server database load with packet size of 50MB (Low)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>0.72</td>
<td>0.72</td>
<td>0.44</td>
</tr>
<tr>
<td>Double server</td>
<td>0.53</td>
<td>0.58</td>
<td>0.36</td>
</tr>
<tr>
<td>Triple server</td>
<td>0.69</td>
<td>0.69</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 4 Server database load with packet size of 250MB (High)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>1.77</td>
<td>0.06</td>
<td>2.08</td>
</tr>
<tr>
<td>Double server</td>
<td>2.64</td>
<td>0.06</td>
<td>2.10</td>
</tr>
<tr>
<td>Triple server</td>
<td>2.18</td>
<td>0.00</td>
<td>0.69</td>
</tr>
</tbody>
</table>

4.9 Results for E-mail Application

Server performance was analyzed and evaluated in this section against the email downloads and uploads response time when the three scenarios were considered. E-mail with packet size of low (50mb), medium (150mb) and high (50mb) were used in link speeds of 10gbps, 2gbps and 30Mbps configured and evaluated against the performance metrics.

4.10 E-mail Download Response Time–One Server Scenario

The tables 5 and 6 shows the e-mail download response time when one server was used on the network and the combined graph in Figure 8 shows the graphical representation and performance of the one server scenario.

Table 5 E-mail downloads response time with packet size of 50MB (Low)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>0.59</td>
<td>0.59</td>
<td>0.71</td>
</tr>
<tr>
<td>Double server</td>
<td>0.76</td>
<td>0.99</td>
<td>0.66</td>
</tr>
<tr>
<td>Triple server</td>
<td>1.03</td>
<td>1.03</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Table 6 E-mail downloads response time with packet size of 250MB (High)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 mins</td>
<td>50 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>1.27</td>
<td>1.23</td>
<td>1.27</td>
</tr>
<tr>
<td>Double server</td>
<td>1.42</td>
<td>1.19</td>
<td>1.44</td>
</tr>
<tr>
<td>Triple server</td>
<td>1.31</td>
<td>1.24</td>
<td>1.33</td>
</tr>
</tbody>
</table>

4.11 E-mail downloads responses time—Two Servers Scenario

The result of the download response time for two servers was used on the network are shown in tables 5 and 6. The result shows an increase in downloads response time when data packet sizes increased.

4.12 E-mail downloads response time—Three Server Scenario

In the third scenario, the email application was diverted to its preferred destinations in the LAN across the routers and cloud to the servers in the network. Table 6 showed the results when email application was analyzed and evaluated against the performance metrics among the three servers in the network.

Figure 7 E-mail downloads response time—Three Server Scenario

From Figure 7, the e-mail response time which is the time taken for the web server to process user requests across the cloud network back and forth, was evaluated against the performance metrics within the three servers. From the graph, it could be deduced that the load had a low value of 0.8secs at high (250mb) and medium (150mb) packet sizes to 1.8secs and then declined to 1.2secs and maintained a constant value throughout the simulation period. It was concluded that the e-mail response time was high when the load on the network was high and vice versa. The combined graph in Figure 8 shows the graphical representation and performance of the three scenarios.

Figure 8 E-mail downloads response times

Figure 8 gives account of all three scenarios for e-mail downloads response times.

4.13 E-mail Upload Response Time

This is the response time taken for data to be uploaded unto a cloud network.

4.14 E-mail Upload Response Time—One Server Scenario

The upload response time was evaluated in this section. The tables in 7 and 8 shows the upload response time when one server was used on the network. The graph in Fig 4.8 showed the performance of the one server scenario.

4.15 E-mail Upload Response Time—Two Servers Scenario

The tables 7 and 8 shows the values when two servers were implemented on the network. A packet latency of 0.05secs was enforced to introduce delay into the system. Refer to the combined graph in Figure 10 for the graphical representation and performance of the two server scenario.

4.16 E-mail Upload Response Time—Three Servers Scenario

In the third scenario, the email application was routed by the routers to the three servers on the network. The results in tables 7 and 8 shows the response time when email was uploaded across the cloud network.
Table 7 E-mail Upload Response Time with packet size of 50MB (Low)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 mins</td>
<td>50 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>1.19</td>
<td>1.00</td>
<td>1.19</td>
</tr>
<tr>
<td>Double server</td>
<td>0.55</td>
<td>0.74</td>
<td>0.60</td>
</tr>
<tr>
<td>Triple server</td>
<td>1.08</td>
<td>0.93</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Table 8 E-mail Upload Response Time with packet size of 250MB (High)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 mins</td>
<td>50 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>1.16</td>
<td>1.39</td>
<td>1.56</td>
</tr>
<tr>
<td>Double server</td>
<td>1.12</td>
<td>1.35</td>
<td>1.11</td>
</tr>
<tr>
<td>Triple server</td>
<td>1.48</td>
<td>1.46</td>
<td>1.38</td>
</tr>
</tbody>
</table>

4.17 E-mail uploads response time for the Scenarios

It is the e-mail response time received by the server when users make a request to the e-mail server back and forth across the network. Combined effects of the three scenarios is as represented in Figure 10.

Figure 10 E-mail Upload Response Time

Figure 10 shows the simulation results when all the three scenarios were considered. It can be seen from graph that the upload response time was high when one server was used on the network than when two and three servers were used.

4.18 Cloud Utilization Performance

Cloud utilization was evaluated against the point to point utilization. Network utilization is the proportion of current network traffic to the maximum traffic that the port can work with. It indicates the bandwidth use in the network. While high network use indicated the network was busy, low network utilizations indicated the network was idle. At the point when system usage surpassed the limit under ordinary condition, it would cause low transmission speed, intermittence and request delays. The tables and 10 shows the link utilization for applications on the network and Fig 4.18 showed the link utilization across the three scenarios.

Table 9 Cloud utilization with packet size of 50MB (Low)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>10gb/s</th>
<th>2gb/s</th>
<th>30mb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 mins</td>
<td>50 mins</td>
<td>10 mins</td>
</tr>
<tr>
<td>Single server</td>
<td>0.38</td>
<td>0.30</td>
<td>0.31</td>
</tr>
<tr>
<td>Double server</td>
<td>0.32</td>
<td>0.26</td>
<td>0.22</td>
</tr>
<tr>
<td>Triple server</td>
<td>0.36</td>
<td>0.29</td>
<td>0.26</td>
</tr>
</tbody>
</table>

From Figure 9, e-mail upload response time had an initial low value of 0.008secs to a high value of 1.155secs due to routing protocols before reducing to an appreciable value of 0.05secs and then maintained a constant value throughout the simulation period.
Based on the findings, the following conclusions were drawn:

1. The general analysis of the results revealed that, the proposed server model was well utilized for improving the user applications.

2. The utilizations demonstrated that, the total performance of database application was improved when the web traffic was given a separate server. Again when there was heavy browsing across the network, the total cloud utilization was increased and thus the server performance was low and vice versa.

3. The general conclusion was that, server performance and network performance are inversely related, which implies that introducing more servers on the network correlates to increase in the network utilization performance but degrades server performance due to packet delays.

4. The overall analysis was that, using one server had an initial low value but a higher value at end of the evaluation than using two or more servers. Using two or more servers had an initial higher value due to routing protocols of applications to destination servers coupled with synchronization protocols of the additional server node on the network but a lower value at the end of simulation than using a single server.

5. Again, when one server in the network goes down, the other servers in the network can serve customers without a shutdown until the faulty server is restored. Samples of Average performance showed a clear distinction of the server performances.

6. It also came to light that, using more than one server in a company’s network had an initial high cost due to purchasing devices and installation overheads but has an overall cost effectiveness, operational efficiency and user performance than using a single server.

7. Nonetheless, at a point where the servers or nodes are beyond more than required on the network, there would be the possibility of poor system performance. This is because of transmission and routing protocol complexities in processing of user requests by the servers causing delays, congestions and consequent packet dropping.

**RECOMMENDATIONS**

By the findings in this research, it is recommended that companies wishing to test their server performance, efficiency and reliability should consider using a virtual simulator. This is because apart from mitigating foreseeable installation problems and avoiding theft of physical components, it is also cheap and has an operational efficiency compared to using real machines.

The following are recommended for feature studies:

1. A number of applications other than database, e-mail, http and ftp should be used to evaluate the performance of the server models which this paper could not cover.

2. Anyone trying to embark on a similar project should give attention to the network security such as firewalls for which this research failed to highlight.

3. The graphical analysis measurements should be done in 10mins and 50mins to give vivid accounts at all times to be able to capture for variations in values due to the rise and fall at certain times which the table measurements fail to do.
6. REFERENCES


