1 PENDAHULUAN

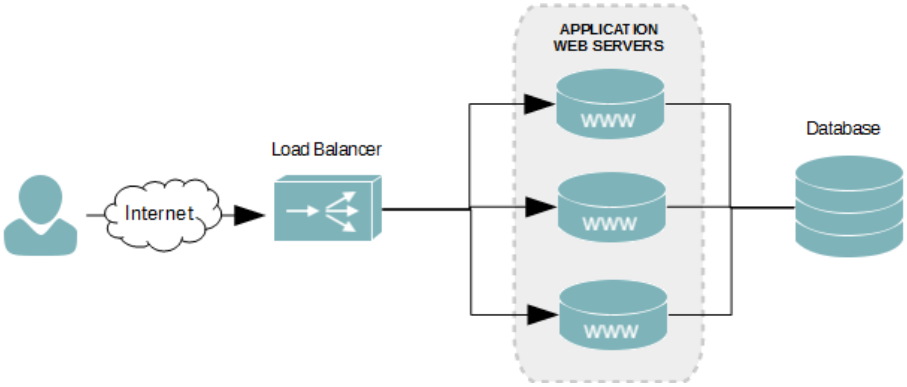
Developments in information technology today is growing very rapidly, this development can be seen from the website uses an information provider. Web server is the main component in the development of a website, so it requires a web server has the reliability, availability, and scalability very well. An information service provider, in this case, is a website that is accessible to millions of users, each information request to the web server will become a burden, and more requests are accepted, also increasing the burden of a web server, a web server is overloaded if it will cause web server to be down. This is of course directly into a disadvantage as information providers and users of information. Load balancing is one of the solutions to overcome this because load balancing can improve reliability, the availability, and scalability (Jindal et-al, 2001),(Bourke, T., 2001). In this study analyzed how the performance difference between a single Web Server with Load Balancing Web Server technology so as to provide clarity difference why Load Balancing can be a solution for a Web server that has a high workload.

With the advent of the Internet, the network now occupies center stage. As the Internet connects the world and the intranet becomes the operational backbone for businesses, the IT infrastructure can be thought of as two types of equipment: computers that function as a client and/or a server, and switches/routers that connect the computers. Conceptually, load balancers are the bridge between the servers and the network, as shown in Figure 1. On one hand, load balancers understand many higher-layer protocols, so they can communicate with servers intelligently. On the other, load balancers understand networking protocols, so they can integrate with networks effectively (Jindal et-al, 2001),(Singh, H., Kumar, S.,2011),(Cardellini, V., Colajanni, M.,1999).

Load Balancer uses some of the same equipment to perform the same task. This makes it possible to do it faster by using only 1 equipment and can ease the workload, and speed up the response time. Load Balancing is the main service and user mediator, where the main service is the server/machine that is ready to serve users. When Load Balancer receives service requests from users, the request will be forwarded to the main server. Usually Load Balancing can smartly determine which server has a lower load and faster response. Can be done for servers that are experiencing problems and can only be distributed to servers that can provide services. This is one of the advantages that load balancers generally have, the service does not allow interference in the user's eyes (Cardellini, V., Colajanni, M.,1999),(Aversa, L., Bestavros, A.,2000),(Salchow, Jr.KJ.,2012). By deploying the load balancer, we can immediately gain several benefits:

1. Flexibility. Load Balancing allows the addition and removal of servers to a site at any time, and the effect is immediate. Among other advantages, this allows for the maintenance of any machine, even during peak hours with little or no impact to the site. A load balancer can also intelligently direct traffic using cookies, URL parsing, static and dynamic algorithms, and much more (Iniewski, K., McCrosky, C., Minoli, D.,2008).

2. High availability. Load Balancing can check the status of the available servers, take any nonresponding servers out of the rotation, and put them in the rotation when they are functioning again. This is automatic, requiring no intervention by an administrator. Also, the load balancers themselves usually come in a redundant configuration,



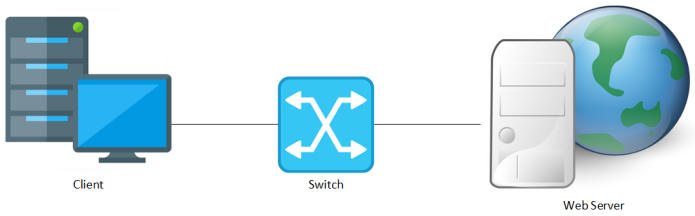
Gambar 1: Load balancing concept

employing more than one unit in case anyone unit fails (Iniewski, K., McCrosky, C., Minoli, D.,2008).

3. Scalability. Since Load Balancing distributes load among many servers, all that is needed to increase the serving power of a site is to add more servers. This can be very economical, since many small- to medium-sized servers can be much less expensive than a few high-end servers. Also, when site load increases, servers can be brought up immediately to handle the increase in traffic (Iniewski, K., McCrosky, C., Minoli, D.,2008).

There are two options to consider when designing load-balancing solutions. The choice of the solution is to use software load balancing or hardware load balancing. Each choice has its own requirements, strengths, and weaknesses. It is up to us to evaluate our business needs, configuration, and growth path so that we can identify optimal solutions to meet needs. And from the type Load Balancing can be divided into 2 types, namely (Bourke, T., 2001),(Cardellini, V., Colajanni, M.,1999),(Ehrhardt, C 2010):

- Load Balancing Software. Where Load Balancing runs on a PC/Server, and the Load Balancing application is installed and needs to be configured before it can function. The advantage is that if there are additional features or additional facilities there is no need to replace the entire load balancing device. The performance of the load balancing process is influenced by the computer device used, it cannot only rely on sophisticated software capabilities. Hardware that can affect the performance of this method is the network card (Network Interface Card) used, the amount of RAM on the device, large and fast storage media, etc. So that the performance of this method is difficult to predict. There are many Load Balancer Software, some of which are the most widely used are Linux Virtual Server, Ultra Monkey, and Network Load Balancing.
- Hardware Load Balancing. Where Load Balancing runs in a device/tool that has been prepared from the factory and is ready for use. Load Balancing Hardware Type is widely used because of its ease. Some Hardware Load Balancers include Cisco System Catalyst, Coyote Point, BIG-IP F5 Network, Baraccuda Load Balancer.



Gambar 2: Single web server topolgy

2 RESEARCH METHODOLOGY

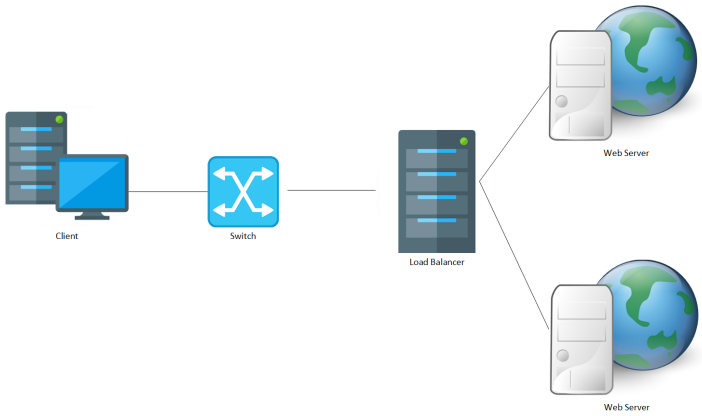
This study performs single processing and parallel processing performance measurements by building two topologies to see the performance of each architecture. The testing process begins with the process of preparation, installation, and configuration of the test topology. The next step is to measure the performance of CPU usage and RAM usage. This research uses Load Balancing Software that is by utilizing Linux Virtual Server installed on the Server that has been prepared.

From Figure 2 and Figure 3 below is a network topology of Single Web Server and Load Balancing Web Server, there is a client that serves to test the website on a web server using assistive applications. At the time of the test, there is a switch that serves to connect the network between the client and the server. And for the single web server using 1 machine as a Web Server, Web Server Load Balancing using 2 machines as a web server and 1 machine functions as a Load Balancer. The design of the network topology is implemented for this study are as follows:

The hardware used in this study consisted of computer servers, client computers, switches and other network devices such as in Table 1 for a Single Web Server and table 2 for a Load Balancing Web Server.

CPU usage (%) and RAM usage (MiB) is done by using an application for resource monitors that run on the Web Server. At this stage of testing is done by running the Stress Meter application Apache JMeter on the client computer to simulate the load request respectively 50, 100, 250 and 500. This testing will be performed using each of the bandwidth is 512 Kbps, 2 Mbps, 5 Mbps and 10 Mbps to measure how much performance difference.

The next test is to measure the Request Response Time and Fail. Response time in question is how much time (milliseconds) when a Web server responds to any requests that come from the client. To Fail Request in question is what percentage of the number of clients who have failed in a request to the Web server (%). The test is performed by using the help of an application using Apache JMeter is executed on the client, the Web server to the load carried each with a request number 50, 100, 250 and 500. Once charged to the Web server is then obtained: Average number of response time (millisecond) and Fail Request (%) of the test results.



Gambar 3: Parallel web server topology

Table 1: Single Web Server Hardware Specifications

Device	Quantity	Specifications
Web Server	1	· Processor Intel Xeon Processor E3-1225 v5 4 Core 3.30 GHz
		· RAM ECC UDIMM 8 GB
		· Hardisk SATA 1 TB
		· 1 LAN Card Giga Ethernet
Client	1	· Processor Intel Core i5 3.0 GHz
		· RAM DDR3 4 GB
		· Hardisk SATA 1 TB
		· 1 LAN Card Fast Ethernet
Switch	1	Unmanageable Fast Ethernet Switch 16 Port

3 PERFORMANCE SINGLE WEB SERVER AND LOAD BALANCING WEB SERVER

In this study, the performance measurement of a single web server and load balancing web server is done with two approaches:

- Measure and compare the performance of CPU usage. CPU time (or process time) is the amount of time for which a central processing unit (CPU) was used for processing instructions of a computer program or operating system, as opposed to elapsed time, which includes for example, waiting for input/output (I/O) operations or entering low-power (idle) mode. The CPU time is measured in clock ticks or seconds. Often, it is useful to measure CPU time as a percentage of the CPU's capacity, which is called the CPU usage [10]
- Measure and compare RAM usage performance.

Table 2: Load Balancing Web Server Hardware Specifications

Device	**Quantity**	**Specifications**
Web Server	2	· Processor Intel Xeon Processor E3-1225 v5 4 Core 3.30 GHz · RAM ECC UDIMM 8 GB · Hardisk SATA 1 TB · 1 LAN Card Giga Ethernet
Load Balancer	1	· Processor Intel Xeon Processor E3-1225 v5 4 Core 3.30 GHz · RAM ECC UDIMM 8 GB · Hardisk SATA 1 TB · 1 LAN Card Card Giga Ethernet
Client	1	· Processor Intel Core i5 3.0 GHz · RAM DDR3 4 GB · Hardisk SATA 1 TB · 1 LAN Card Fast Ethernet
Switch	1	Unmanageable Fast Ethernet Switch 16 Port

Table 3: CPU Usage of Single Web Server and Load Balancing Web Server

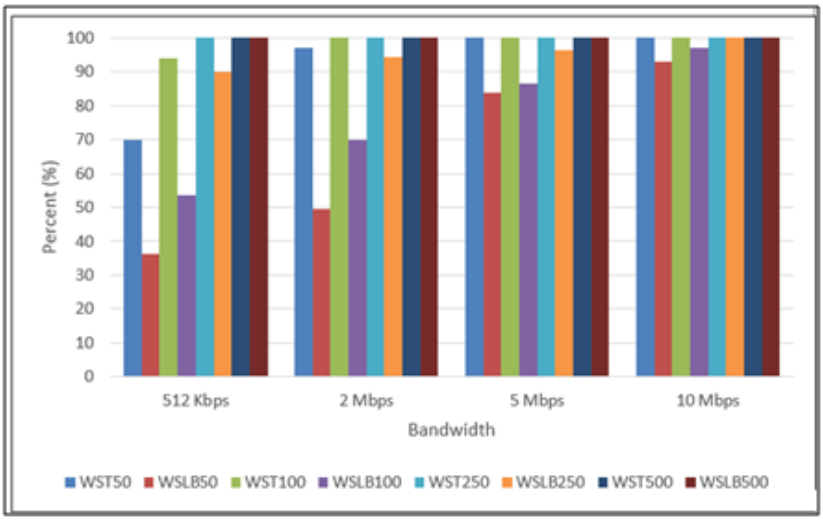
Band-width	Ave. CPU WST 50(%)	Ave. CPU WSLB 50(%)	Ave. CPU WST 100(%)	Ave. CPU WSLB 100(%)	Ave. CPU WST 250(%)	Ave. CPU WSLB 250(%)	Ave. CPU WST 500(%)	Ave. CPU WSLB 500(%)
512 Kbps	69,8	36,4	93,9	53,55	100	89,95	100	100
2 Mbps	97	49,35	100	70	100	94,35	100	100
5 Mbps	100	83,7	100	86,3	100	96,35	100	100
10 Mbps	100	92,85	100	96,85	100	100	100	100

3.1 CPU Usage

From both the results obtained (Table 3 and Figure 4) and the comparison between Single Web Server with Load Balancing Web Server, that the bandwidth is too small to affect the client and server latency indirectly also affect CPU usage. Another thing that is obtained is ideally a single Web server according to the specifications of the hardware that has been mentioned previously is able to handle a load of 100 requests well in some conditions the bandwidth that has been tested. And for Web Server Load Balancing in accordance with the previously mentioned hardware is capable of handling the load is 250 requests well in some conditions the bandwidth that has been tested. Both of these things, it can also be concluded that there is an increase in handling requests from Web Server Load Balancing Web Server Single compared with the magnitude of improvement of approximately 150%.

3.2 RAM Usage

When compared to the use of RAM (Table 4 and Figure 5), in contrast to the increased CPU usage which reached approximately 150%, then this is not the case with the use of



Gambar 4: Comparison single and load balancing web server in CPU usage

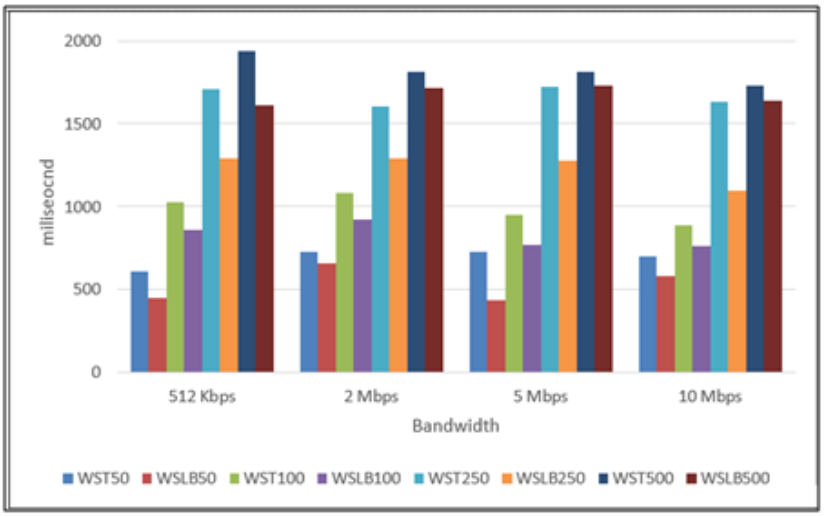
Table 4: RAM Usage of Single Web Server and Load Balancing Web Server

Band-width	Ave. RAM WST 50(%)	Ave. RAM WSLB 50(%)	Ave. RAM WST 100(%)	Ave. RAM WSLB 100(%)	Ave. RAM WST 250(%)	Ave. RAM WSLB 250(%)	Ave. RAM WST 500(%)	Ave. RAM WSLB 500(%)
512 Kbps	611	449	1028	859	1705	1288,5	1936	1610
2 Mbps	729	656,5	1083	925,5	1607	1290,5	1815	1714
5 Mbps	729	438	948	772	1721	1280	1812	1726,5
10 Mbps	700	583	890	761,5	1634	1099,5	1728	1638

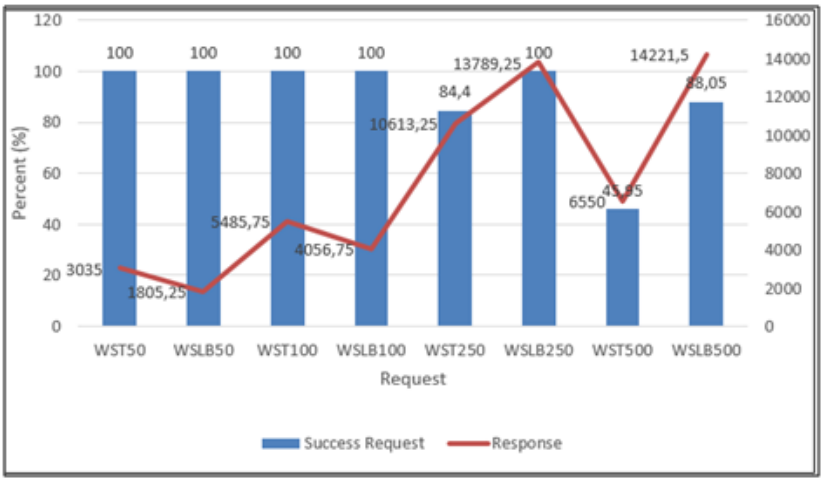
RAM, although the data obtained from the use of RAM Web Server Load Balancing less (better) compared with a single web server, but the difference was not significant. It can be concluded that the use of the RAM between Single Web servers and Load Balancing Web Server did not differ, although the actual physical RAM belongs Load Balancing Web Server has a capacity of 2 times the RAM Single Web Server.

3.3 Average Response Time to Fail Request

From the test results ranging from 50, 100, 250 and 500 requests (Figure 6), each of the stages by limiting the bandwidth used, the test has met the ideal and maximum load, respectively. From this writer tried to draw the conclusion that the CPU load was very role to influence the ability of a Web server to serve the number of requests are coming, along with the RAM usage also become indispensable to assist in the performance of the CPU service requests are coming. From these two parameters, namely the use of CPU and RAM will determine the success of request and response time of a Web server. So if the CPU and RAM utilization is not maximized, it is possible that the web server is still able to serve requests are coming.



Gambar 5: Comparison single and load balancing web server in RAM usage



Gambar 6: Avarage response time

4 CONCLUSION

After testing has been done in this paper, now can be conclusion that, Load Balancing Web Server can serve more requests are coming and have the reponse time better than Single Web server. Load Balancer can divide the workload on both web servers evenly. And, the size of the bandwidth effect on response time, due to the size of the bandwidth affects the latency of a server to a client when accessed. The greater the bandwidth, the better it will be owned by the response time to the server.

Referensi

Jindal, A., Lim, S.B., Radia, S. and Chang, W.L., Sun Microsystems Inc, (2001). Load balancing in a network environment. U.S. Patent 6,327,622.

Bourke, T., (2001). Server load balancing. O'Reilly Media, Inc.

Kopparapu, C.,(2002), *Load Balancing Servers, Firewalls and Caches.* Canada : Wiley Computer Publishing.

Iniewski, K., McCrosky, C., Minoli, D.,(2008), *Network Infrastructure and Architecture Designing High Availability Networks.* New Jersey: John Wiley & Sons, Inc.

Bourke, T.,(2001), *Server Load Balancing.* United States of America : O'Reilly & Associates, Inc.

Singh, H., Kumar, S.,(2011), Dispathcer Based Dynamic Load Balancing on Web Server System. *Int. J. of Grid and Distributed Computing Vol. 4, No. 3, (2011). 90-106.*

Cardellini, V., Colajanni, M.,(1999), Dynamic Load Balancing on Web Server Systems. *J. IEEE Internet Computing.*

Aversa, L., Bestavros, A.,(2000), Load Balancing a Cluster of Web Servers Using Distributed Packet Rewriting. *In Proceedings of IPCCC'2000 The IEEE International Performance, Computing, and Communication Conference, Phoenix, AZ.*

Salchow, Jr.KJ.,(2012), *Load Balancing Nuts and Bolts.* F5 Networks, Inc.

Ehrhardt, C (2010), CPU time accounting. IBM