

# Comparison of distributed cache and centralized cache on web proxy

Parulian<sup>1</sup>, Muhammad Zarlis<sup>2</sup>, Benny BenyaminNasution<sup>3</sup>

<sup>1</sup>University of North Sumatra | University of HKBP Nommensen

<sup>2</sup>University of North Sumatra

<sup>3</sup>State Polytechnic Medan

**Abstract.** *Caching* is an object storage mechanism from a web page that has been accessed. With caching, bandwidth usage will be more efficient and the time to access a web page is getting faster. In caching there are 2 techniques used, namely centralized cache and distributed cache. The purpose of this study is to build a distributed cache system that supports Availability, Reliability, and Scalability. This system will overcome various problems found in the centralized cache, such as single point of failure, unreliable, and not scalable. With the availability of Availability, the system will avoid the problem of single point of failure by implementing a failover mechanism. For the Reliability system will still hold (reliable) accept requests from users, because some proxy servers (parent proxy) will serve requests from users. For Scalability, the system can add proxy servers easily and quickly. To develop the system, the developer takes several approaches such as basic learning about proxy servers and related theories. Next is an analysis to build a system that supports Availability, Reliability, and Scalability. Then the analysis phase is implemented and testing of the system is built. Tests carried out are testing centralized cache systems and testing distributed cache systems. The test results will then be compared to see the performance comparison between centralized cache and distributed cache.

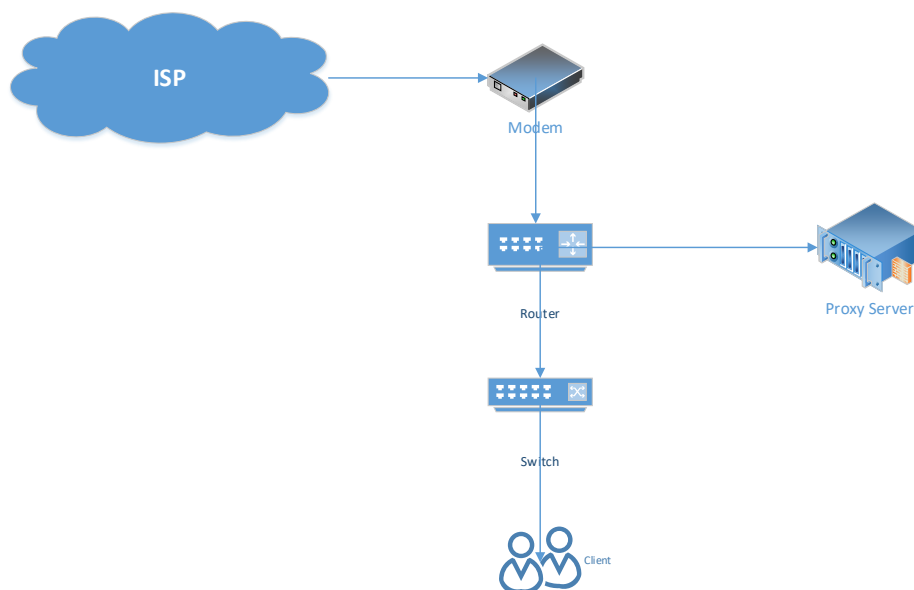
## 1. Introduction

Caching technique is a technique that uses a cache to store the temporary web document, such as HTML pages and images that previously accessed by the user. In caching there are three known system is, without a proxy server, web caching proxy server centralized and distributed caching web proxy server. Commonly used systems are centralized caching proxy server, the caching proxy server centralized technical problems were found to availability, reliability, and scalability that should be owned by a reliable system. Availability problems faced by this technique is when a single web proxy server down and no backup servers to handle user requests, while the possibility of a single problem for reliability web proxy server can not handle the request that a lot of users, and the last problem is scalability, ie when the number of users increases, centralized caching proxy server can not be developed easily. Based on the problems faced by centralized web proxy server will require a solution to overcome these problems, one of them is to use a distributed caching proxy server (Piatek 2004). This technique will distribute to multiple web proxy cache server, so it can support in terms of availability of backup server availability, reliability in terms of handling request, and scalability in terms of system development if the number of users increases. Based on the problems faced by centralized web proxy server will require a solution to overcome these problems, one of them is to use a distributed caching proxy server (Piatek 2004). This technique will distribute to multiple web proxy cache server, so it can support in terms of availability of backup server availability, reliability in terms of handling request, and

scalability in terms of system development if the number of users increases. Based on the problems faced by centralized web proxy server will require a solution to overcome these problems, one of them is to use a distributed caching proxy server (Piatek 2004). This technique will distribute to multiple web proxy cache server, so it can support in terms of availability of backup server availability, reliability in terms of handling request, and scalability in terms of system development if the number of users increases.

## 2. Methodology

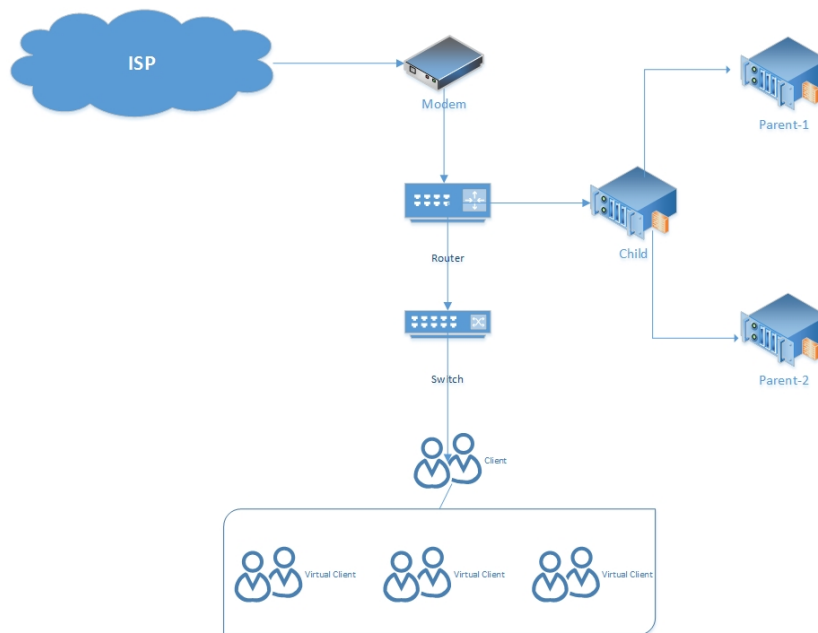
2.1 Centralized Cache Physical Topology describes the interconnection between components is required at the time of testing, the following picture Physical topology used at the time of testing.



**Figure 1.** Physical Topology on Centralized Cache

### 2.2 Physical Topology Distributed Cache

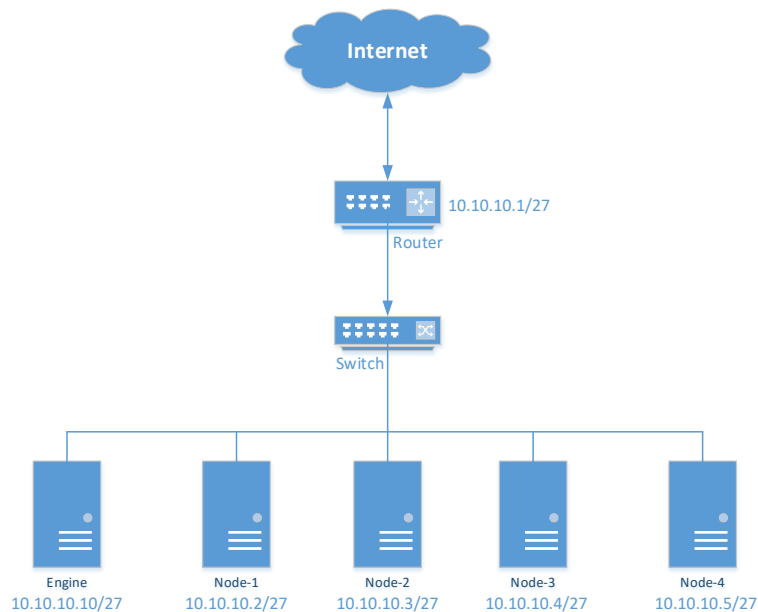
Physical topology describes the relationship among components is required at the time of testing, the following picture Physical topology used at the time of testing.



**Figure 2.**Physical Topology in Distributed Cache

### 3. Discussion

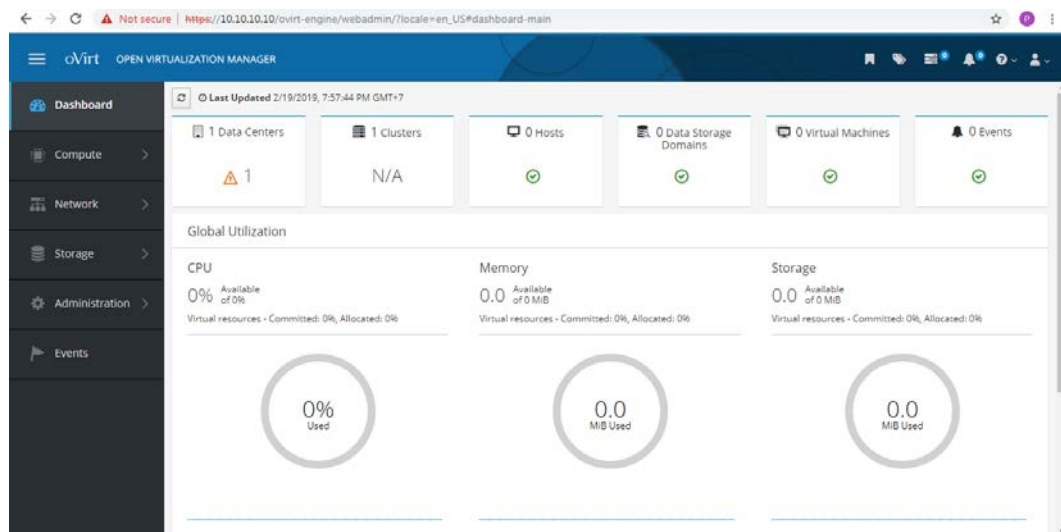
Network topology is used for Server Virtualization is a star topology. All server computers connected to one switch with UTP cable.



**Figure 3.**Network Topology Virtualization

#### 3.1 Operating System Installation and Configuration on the Server Engine

Access Web Administration Portal oVirt engine through a browser by typing in the search server computer IP address in the browser address bar.



**Figure 4.** Display search oVirt Portal

### 3.2 Operating System Installation on any Server Node

Each node of the node 1 to node 4 to install the operating system oVirt. Nodes are computers that are physically different to be combined into one in the system with oVirt. OVirt node will act as a so-called hypervisor or Virtual Machine Monitor function to run multiple operating systems in the host operating system. Hypervisor task is to set the appropriate operating system every turn so as not to interfere with another operating system that is running. The hypervisor set up multiple systems at once at the same time.

### 3.3 Configuration Server Node

Results from the hosts, nodes that have been add to the engine.

Name	Comment	Hostname/IP	Cluster	Data Center	Status	Vir
node1.testis		10.10.10.2	Default	Default	Up	
node2.testis		10.10.10.3	Default	Default	Up	
node3.testis		10.10.10.4	Default	Default	Up	
node4.testis		10.10.10.5	Default	Default	Up	

**Figure 5.** Node add to that already in the engine

#### 3.3.1 Installs NFS Storage on each node oVirt

NFS (Network File System) is a sharing protocol used for shared storage or network resources that are channeled through without seeing the client using the operating system used on each client. Mechanism of Network File System is a distributed file, which is generally implemented in a computing environment centralized resources.

#### 3.3.2 Installing ISO Domain on each oVirt node

ISO domain on the oVirt is a host that contains the ISO storage that is used to install the operating system in the form of ISO. System operation in the form of an ISO used was CentOS 7, ISO is installed on the oVirt node-1.

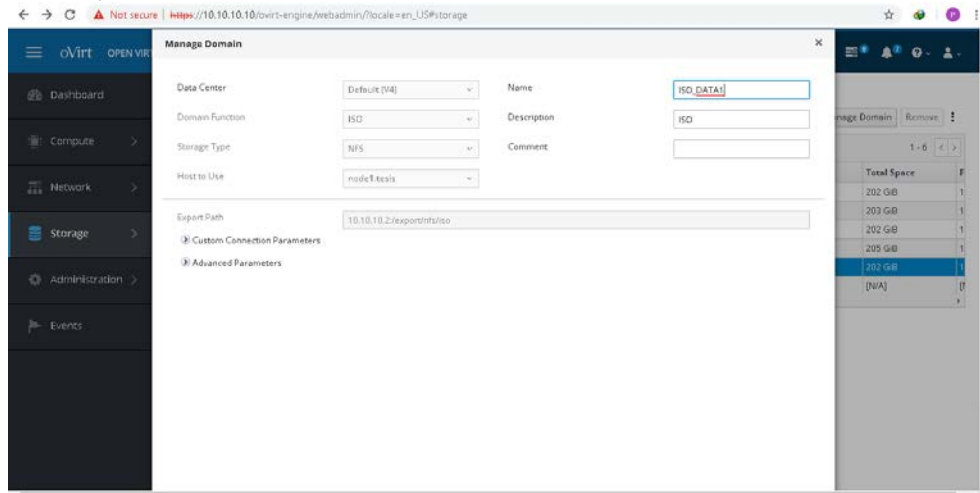


Figure 6. Install ISO

### 3.5 Install Squid Proxy

In each virtual machine proxy-1, 2-proxy virtual machine, virtual machine proxy-3, and virtual machine proxy-4 in the install squid.

### 3.6 Install and Configure a DNS (Domain Name Server) and Reverse Proxy

DNS (Domain Name Server) serves as a translator to hostname or IP address instead. This makes Internet users become easier to access and recall. DNS has five (5) degree level, the Root Level Domains Top Level Domains, Second Level Domains, third Level Domains, and the last hostname. Reverse Proxy is a proxy type that is used as an intermediary between the client and the web server.

### 3.7 Configuring to Forward DNS Zone and Reverse DNS Zone Configuration

*forward zone* is a zone that is useful when the client performs a query with the domain name, then the server will look up the IP address of the domain has been queried or zone domain whose job translating, into IP addresses. *Reverse Zone* Forward Zone is the opposite of that work when the client performs a query on behalf of the IP address, then the server will look up the domain of IP addresses that have been queried and this zone that served to translate IP addresses into domain.

### 3.8 Install and Configure Nginx server 10.10.10.12

Nginx is an open-source web server that is used as a reverse proxy, HTTP Cache and Load Balancer.

### 3.9 Installation and Configuration Nginx On Virtual Machine

Nginx configuration in each virtual machine that serves as a reverse proxy, HTTP Cache, and Load Balancer. The algorithm used in the load balancer nginx configuration is ip\_hash. This is a static scheduling algorithm which forwards the request from the server to the client according to what was requested, the algorithm is very useful to process very large.

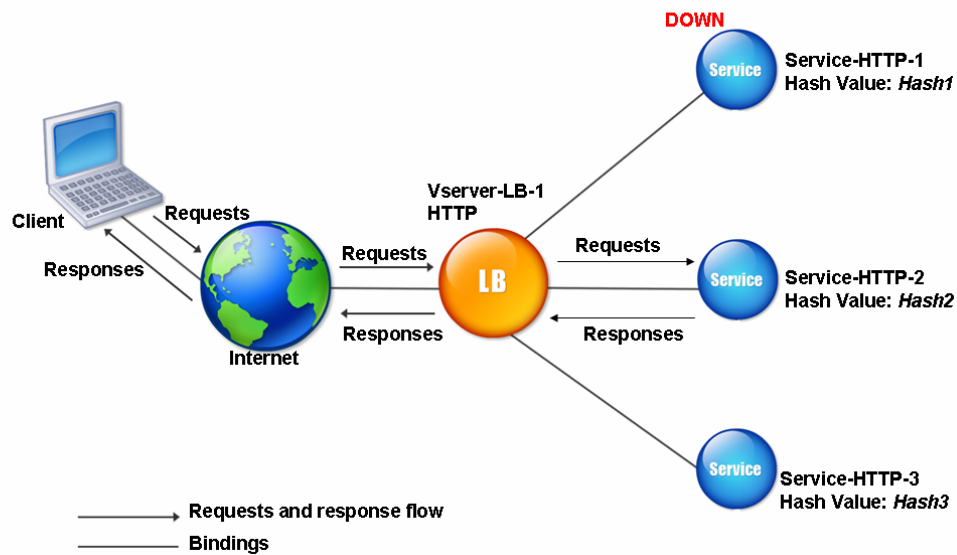


Figure 7. Hash Method

This method can reduce demand latency, response and ensure the utilization of the CPU resources better. This method is also a web server load balancer.

#### 4. Examination

##### 4.1 Distributed Systems Testing Availability

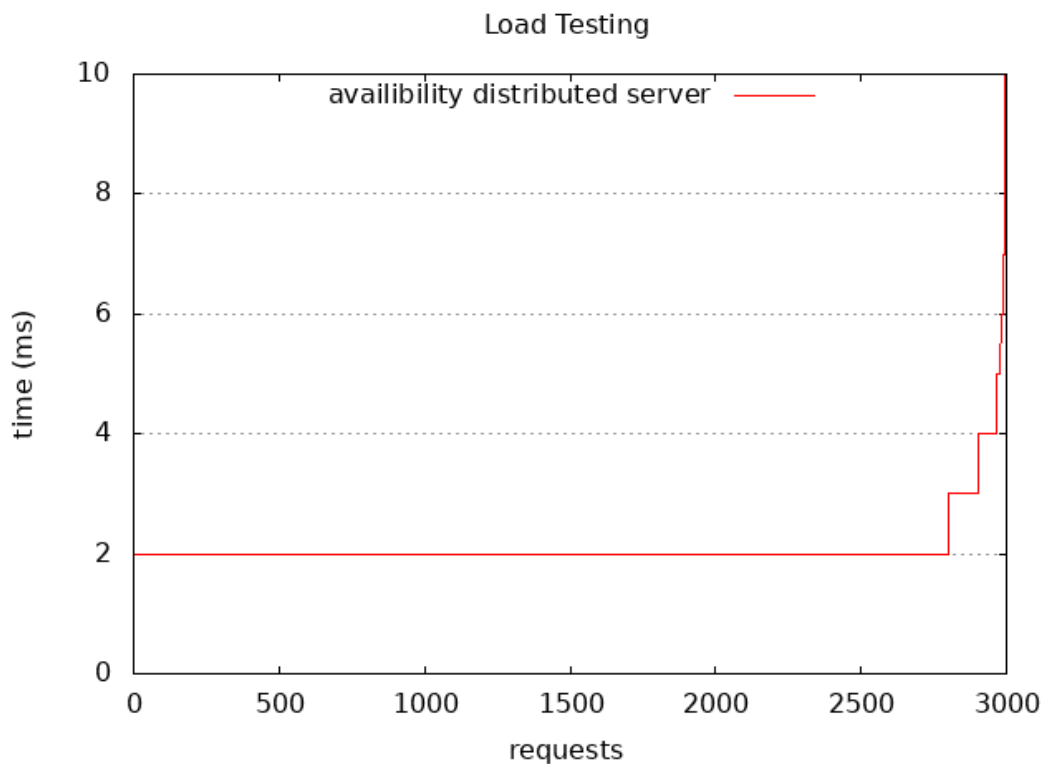
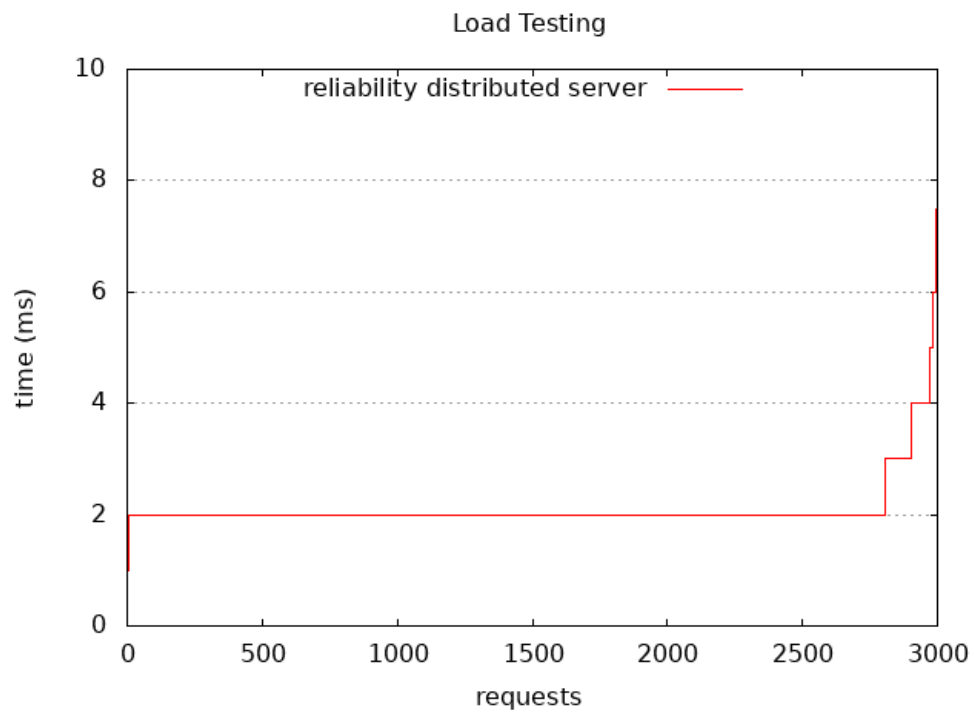


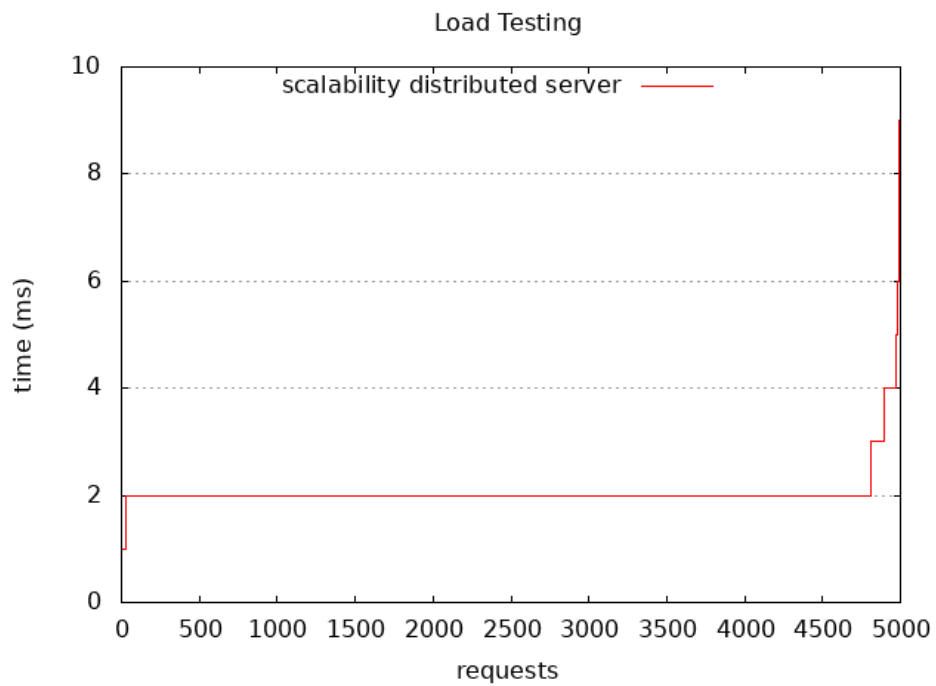
Figure 8. Graph Testing Distributed System Availability The user 3000

## 4.2 Distributed System Testing Reliability



**Figure 9.** Graph Testing Distributed System Reliability 3000 user

## 4.3 Distributed System Testing Scalability

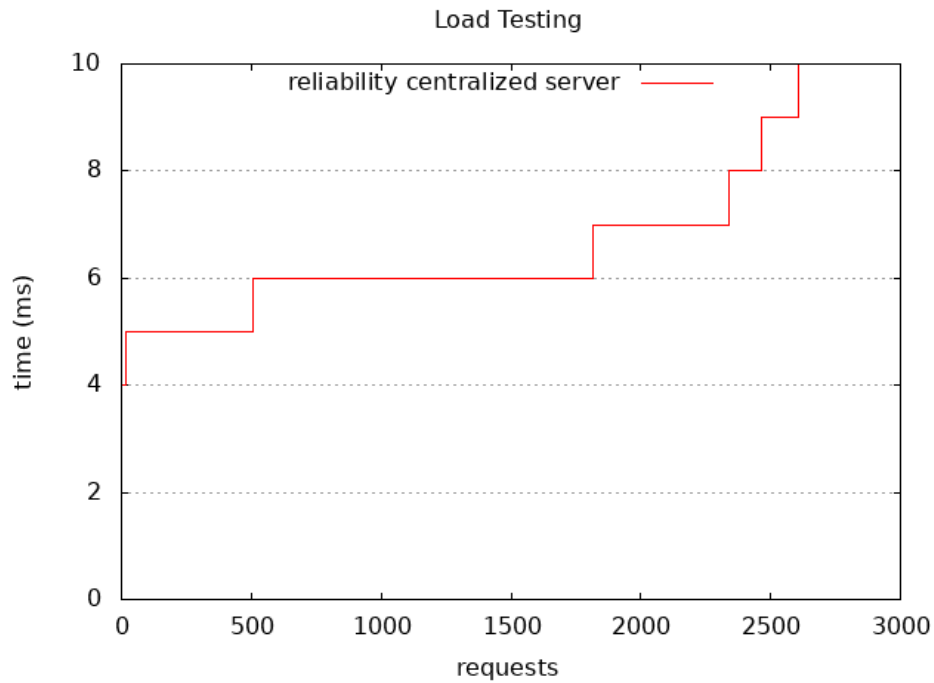


**Figure 10.** Graph Scalability Testing Distributed System 5000 user

## 4.4 Centralized System Testing Availability

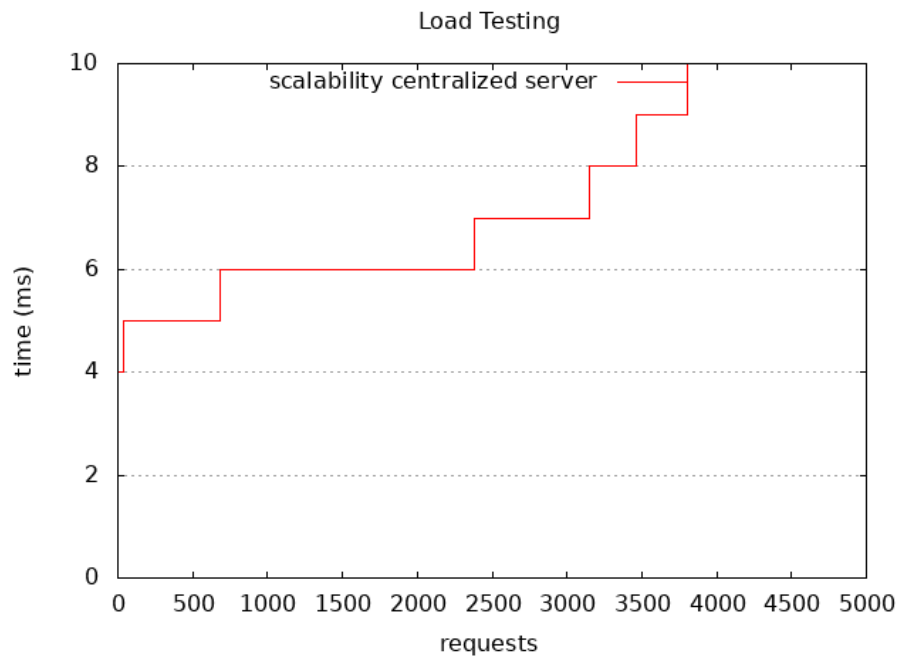
Testing will stop because the connection is lost due to the centralized system there is no back-up.

#### 4.5 Centralized Reliability Testing Systems



**Figure 11.** Graph Reliability Testing System Centralized user 3000

#### 4.6 Centralized Scalability Testing System



**Figure 12.** Graph Scalability Testing System Centralized user 5000



## 5. Comparison

Once centralized and distributed systems built then conducted an experiment to test the system. The test is done to see availability, reliability, and scalability of the system that has been built. The comparison obtained through experiments are as follows:

**Table 1.**Comparative Experiment 1

Experiment 1	Centralized (avg) / ms	Distributed (Avg) / ms
Availiability	X	2
Reliability	7	2
Scalability	7	2

**Table 2.** Comparative Experiment 2

Experiment 1	Centralized (avg) / ms	Distributed (Avg) / ms
Availiability	X	4
Reliability	9	3
Scalability	7	5

**Table 3.** Comparative Experiment 3

Experiment 1	Centralized (avg) / ms	Distributed (Avg) / ms
Availiability	X	5
Reliability	11	3
Scalability	9	4

## 6. Conclusion

After analyzing and conducted experiments on the system is built, then the conclusions obtained are as follows:

1. Based on test results, distributed cache system has been built to meet the criteria for Availability, Reliability, and Scalability.
2. Based on the tests performed, distributed cache system its performance is better than a centralized system cache.
3. The number of users who make requests to the system will affect the response time and error rate of the system.

## References

- [1] Aivaliotis, D., 2013. Mastering Nginx. Birmingham: Packt Publishing.

- [2] Alsaibani, Omar. Load Balance Algorithm Performance 2013. On the Apache Web Server and Nginx With PostgreSQL Database.
- [3] Apache Software Foundation. Apache HTTP Server 2013. Documentation.[http://httpd.apache.org/docs/2.0/mod/mod\\_proxy.html](http://httpd.apache.org/docs/2.0/mod/mod_proxy.html),
- [4] Arifin, Fitriana. 2012. Effectiveness Analysis Bandwidth Using IPCop (Case Study: Energy Technology Center). Stun.
- [5] Bourke, Tony. 2001. Server Load Balancing
- [6] S. Deshpande, Ravale P., and S. Apte 2010. In Centralized Cache Coherence And Distributed Shared Memory Shared Memory Architectures. International Journal on Computer Science and Engineering (IJCSE).
- [7] Eom. Y., Hwang D., & Lee J. 2015. EM-KDE: A Locality-Aware Policy Job Scheduling With Semantic Distributed Caches. J. Parallel Distrib. Comput.
- [8] Foster and Kesselman, 2003. The Grid: Blueprint for a New Computing Infrastructure, Second Edition, Morgan Kaufman
- [9] Hu, Y., Nanda, A. & Yang, Q., 1999. Measurement, Analysis and Performance Improvement of the Apache Web Server. IEEE International Performance Computing and Communications Conference, pp. 261-267
- [10] Listyawan. N, Riadi. I. 2013. Optimizing Storage Video Using Videocahce On Proxy Server (Case Study in figuring Janturan.Net Yogyakarta). Bachelor of Engineering Informatics Journal, Volume 1 Number 2.
- [11] Lutfi. M, Data. M & Yahya. W. 2018. Comparison of Reverse Proxy Caching Performance Nginx and Varnish On Apache Web Server. Journal of Information Technology Development and Computer Science.
- [12] Mekongga, Gernowo&Sugiharto. 2012. The Prediction of Need Computer Network Bandwidth On Through Artificial Neural Network Method of Backpropagation. Journal of Business Information Systems
- [13] Meiyappan, Somasundaram & Kit Wai Chan. Concept of Cache in web proxies.
- [14] Nedelcu. C. Nginx HTTP Server 2015. Third Edition. Packt Publishing.
- [15] Piatek, Michael. 2004. Web Distributed Caching Proxy in a Local Network Environment.
- [16] Radjabaycolle, Pulungan. 2016 PREDICTION OF USE BANDWIDTH NETWORK USING Elman Recurrent Neural. Journal of Mathematics and Applied Sciences.
- [17] Setyawan, Arief. 2014. Analysis of Multi Server Load Balancing Mechanism on the UB Server SIAKAD. EECCIS Journal, Vol 8 No 1.
- [18] Tanenbaum, Andrew S.: "Computer Networks", Pearson Education Asia Pte.Ltd, Fourth Edition, 2003
- [19] Warman, I. & Zahni, A., 2013. To make a reservation Mobile Web Engineering JQuery Based On Cell gem. Momentum Journal, Volume 15, pp. 30-38.
- [20] Zhang X., Wang N., Vassilios G., & Howarth M. 2015. A distributed in-network caching scheme for P2P-like chunk of content delivery. Computer Networks.