Parallel Processing to Improve Performance of e-Learning Server Used Message Passing Interface (MPI) Case Study at SMA Negeri 1 Pekanbaru

Abdul Syukur[a]* and Evizal Abdul Kadir[b]

[a,b]*Teknik Informatika, Fakultas Teknik, Universitas Islam Riau*
Jl. Kaharuddin Nasution, Marpoyan, Pekanbaru, Riau, Indonesia 28284

Abstract

Along with the development progress of science, the number of computer resource demand is also increasing. Some fields that require high-end computing is numerical simulation, the problems of science, mathematics and engineering. Thus the computational problem is not a problem with the presence of this cluster that offers services to perform computational processes that exceed the speed of the server computer that stand alone. Cluster is a collection of stand-alone computers that are connected to one another in a computer network used to perform specific computing process in parallel. Cluster basically using the network, but that is characteristic in particular is how this configuration is used to solve the problem. The problem here is how to increase or to balance server workloads e-learning (web server) and the mechanism Message Passing Interface (MPI).

1. Introduction

Currently the development of computer technology has been growing rapidly, one of which is a networking technology, and one of the development of network technology is Parallel Processing Computing. Parallel processing is the process of computing work by using a programming language that run in parallel at the same time. In general, parallel computing is necessary to improve the computing speed when compared to the use of computing on a single computer.

Cluster (parallel processing) constructed a dedicated cluster specifically for high-performance computing or High Performance Computing (HPC) and server performance balancing e-learning or Load Balancing Cluster (LBC). To build a cluster (parallel processing) that requires at least two computers connected to a private network. The operating system is selected to create a cluster (parallel processing) is the Linux operating system. Linux is chosen because it is considered as an operating system that is easily configured to meet the infrastructure of a cluster (parallel processing) and open source. Linux cluster must have a mechanism parallel processing to run applications that can be run in parallel. Parallel processing mechanisms used are Message Passing Interface (MPI). Message Passing Interface (MPI) is a programming interface that is used to distribute computing to the vertices of the other in a cluster (parallel processing).

SMA Negeri 1 Pekanbaru is one of 132 high schools in Indonesia Patronage Ministry Directorate of high school to become a school Implementing Model SKM-PBKL-

---

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
  E-mail address: abdulsyukur@eng.uir.ac.id

PSB where the school-based learning should apply Information and Communication Technology (ICT) in particular e-learning. With the research on parallel processing and cluster computers, is expected to provide solutions to the problems of computing to improve server performance e-learning SMA Negeri 1 Pekanbaru.

2. Support Theoretical

2.1. Taxonomy Flynn for Computer Architecture

Taxonomy is the most popular computer architecture defined by Michael J. Flynn in 1966, Flynn classification scheme based on the flow of information to the processor. Two streams of information flowing on the processor are instruction and data. The instruction stream is defined as a sequence of instructions processed by the processing unit. The data stream is defined as traffic data exchange between memory and processing unit. According to the classification Flynn, the flow of instructions and data flow can be either singular or plural. Computer architecture according to Flynn can be classified as follows:

a) SISD   : Single Instruction Single Data
b) MISD   : Multiple Instruction Single Data
c) SIMD   : Single Instruction Multiple Data
d) MIMD   : Multiple Instruction Multiple Data

2.2. Computer Cluster

Cluster is a set of computers connected and work together as an integrated computer resources to accomplish a goal. A computer cluster consists of at least Message Passing Interface (MPI) and a regulator workload (Job Scheduler). Message Passing Interface (MPI) task for sending data between computers in a parallel system (usually referred to as nodes or host). Job scheduler such as the name would imply duty to receive tasks from the user and to schedule the job on multiple nodes in a parallel system as needed.
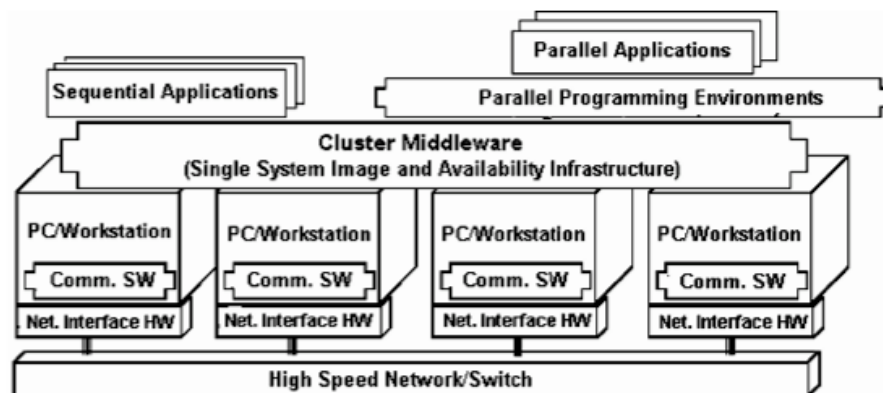

Figure 1. Computer cluster layer.

A system need to build a cluster of four components, namely the two components of hardware and two software components. The first hardware component is a node or computing engines, and the second is a network of interconnected, whereas the first software component is software that can be used by users to develop parallel

applications and the second is the software used to monitor and manage applications running in the cluster.

2.3. System Cluster

Clustering in computer science is a combination of computer resources connected to each other and work together at the same time to perform a task or provide a service. Personal Computer (PC) cluster is a computer system that consists of multiple PC (Personal Computer) which is connectable to a network to perform a task or perform computational work together.

The fundamental mode of the cluster itself is divided into four types, namely:

a) High Performance Computing (HPC) is a type of cluster with high performance achieved with the distribution of tasks or the loads into a set of server computers. High Performance Computing (HPC) is a set of servers that work together at the same time to do something specific task, usually in the form of calculation tasks such weighty, such as simulating the earth, render animated films, and others.

b) Load Balancing Cluster (LBC) is a concept for load balancing or workload to multiple server computers, for example divide the workload web server and database server.

c) High Availability Cluster (HAC) is a type of computer cluster if a server fails serves a particular service, and then the task of the server is automatically thrown to the other server.

d) Grid Computing (GC) grid computing technology is a way of merging the resources owned many computers connected in a network so as to form a unified computer system with large computational resources.

3. Research Methodology

In the methodology of this study it can be concluded that the problems faced is the use of server resources that are not optimal. The methodology can be used to increase a server is to create a Cluster Load Balancing (LBC). With the Load Balancing Cluster (LBC) is expected to address the issue of resource dependency and can improve performance or balance the workload on the server e-learning SMA Negeri 1 Pekanbaru.

a) Preliminary Study

The initial phase will be done to complete this study is a preliminary study, which is studying and researching the issues to be analyzed. The scope of the issues examined or analyzed should be determined beforehand, because without specifying and defining the boundaries of the problem to be studied, then there is never a solution to these problems.

b) Literature

In pursuit of research that has been set, it is necessary to learn some of the literature that will be used in this study. Then studied literature were selected for the literature can be determined which will be used in research. Through the study of literature, we can learn the theories with the use of technology parallel processing that can be applied to improve the performance of server e-learning. Its source may include books, national and international journals, and those associated with parallel processing technology and utilization.

c) Data Collection

To obtain the desired data, the method used in this research is the study of literature or bibliography. Data collected through the study of literature and collect references that had to do with the completion of this study as support research to be conducted.

d) Analysis and System Design

The results in the study of literature, and then performed the analysis phase to determine the need to build a parallel processing both in terms of topology and also the necessary infrastructure. The initial phase will be done is to establish and provide for anything just to make a parallel processing

As for the hardware (hardware) required is three units of computer servers using the Linux operating system Rocks 6.0 (Mamba) and application Message Passing Interface (MPI) as a mechanism that can perform parallel processing to improve performance or balancing the workload of the server computer and use apps ganglia for monitoring processes and computational burden both on each machine computation and computational load in a cluster or all nodes.

a) Implementation

The implementation phase is the last stage of research. This stage prefers the durability and usability of parallel processing in particular to improve performance or balancing the workload of the server computer. At this stage there is a possibility for further development.

b) Results

The results obtained can form an implementation of parallel processing for high-performance computing or High Performance Computing (HPC) and server performance balancing or Load Balancing Cluster (LBC) to be applied to the world of education, government agencies, and companies.

4. Design and Analysis

This section discusses the analysis and design cluster systems that will be used as a parallel processing infrastructure. Clusters are made using three computers with a computer is used as a head node (Front end) and the two others are used as a computer node. Each node uses the Linux operating system Rocks 6.0 (Mamba).

a) Analysis System

For phase analysis of the problems facing our prior audit the network topology that applied previously, it is to know and help us in designing the network topology using computer cluster (parallel processing) that can later be applied to improve or to balance server performance in e-learning. The results of observations conducted in the field as shows in figure 2, the network topology in SMA Negeri 1 Pekanbaru. Based on the scenario of figure 2 can be explained that every computer user either teacher or student environment of SMA Negeri 1 Pekanbaru who wants to access the e-learning intranet either using Personal Computer (PC) of the labor computer is connected to the switch via a cable or using a laptop connected in Wireless Fidelity (Wi-Fi) using the Access Point (AP) which is attached evenly school environment. Based on figure 2, shows that the computer servers of e-learning in SMA Negeri 1 Pekanbaru that exist today still use the stand-alone server. If at the same time the teacher / student actively accessing e-learning which numbered about 900 people then it will slow down the performance of the e-learning server because the workload is too heavy.
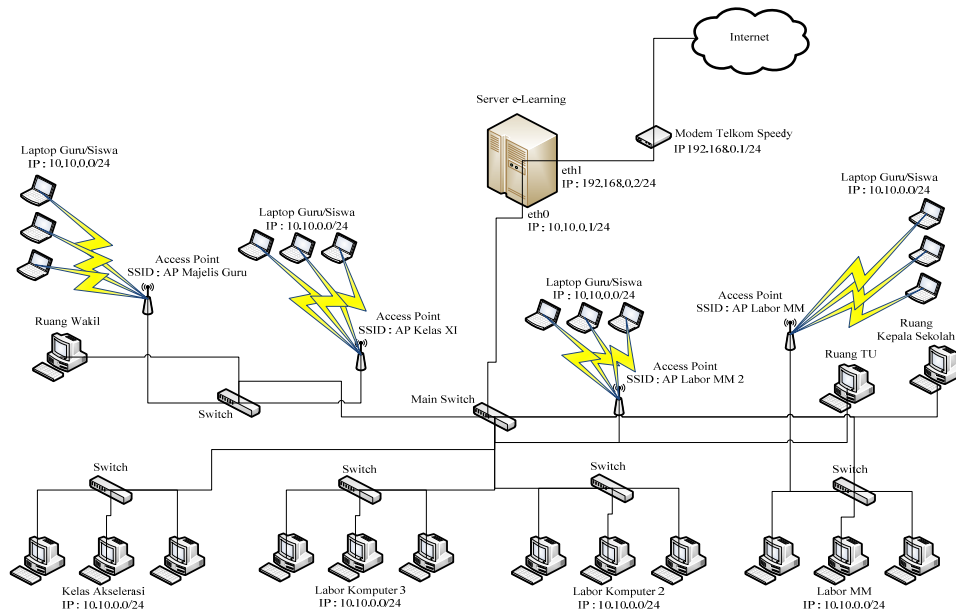
Figure 2. Network topology at SMA Negeri 1 Pekanbaru before used Parallel Processing.

b) Analysis of Cluster System Requirements

Cluster built in this study is a dedicated cluster for parallel computing. Clusters are built is to homogenize heterogeneous operating systems and using different hardware devices on each computer. In building a dedicated cluster for parallel computing needed some requirements indicated by the following points:

i. Cluster must have the ability to distribute the computing process of the machine head node (frontend) to machines in other computing.

ii. Cluster must have a regulatory system data files on each node so as to facilitate centralized system administration.

iii. Cluster must have a global file system that can be accessed by all the nodes for the purpose of accessing the files.

iv. Head node (Frontend) must be able to communicate with the machines node without having to make the authentication process.

Cluster must have a system to monitor the processes and computational burden both on each machine computation and computational load in one cluster or all nodes.

c) Head Node (Frontend)

Head node (Frontend) is a computer in a cluster that is used to delegate tasks or tasks that will be done by the computer nodes. This computer is also used as a place for users to access the cluster so that users can provide computing tasks into clusters. Head node (Frontend) should be accessible through the network, so that the head node (frontend) requires two Ethernet cards, an Ethernet card connected to the public network and the other is connected to a private network or computer node. Besides being a place of access and delegating tasks to the computer node, the head node (frontend) should also run the services include: Network File System (NFS), Network Information System (NIS), Dynamic Host Configuration Protocol (DHCP), Hypertext Transfer Protocol (HTTP), Secure Shell Host (SSH). The services required by the head node (frontend) to integrate the nodes in the cluster. Machines used as a head node (frontend) should have a higher specification than the machines that are used as a computer node. That is because the machine head node computers greater than the machine computer node for the machine head node (frontend) must be running the services which are not run by the computer nodes.

29

d) Computer Node

Computer node is a computer worker who received the task of head node (frontend) and process the task. This computer cannot be accessed over the network because it is connected to a private network. The computer that functioned as a computer node can only be accessed by a head node (frontend) to perform the computing process. Configuration files related to user administration systems such as password, group, and etcetera shadow is located centrally in the head node (frontend). The mechanism of centralized configuration is done using the Network Information System (NIS). NIS enables computers computer node will have a uniform system users with the computer system created on the head node (frontend). Use of the Network Information System (NIS) will certainly ease the system administrator to perform addition or removal of users in a cluster system.

e) Network Topology Design

Cluster network that will be created using standard design outline of rocks cluster that will ultimately result in the expected performance at first. In this study, the cluster that was built in the laboratory SMA Negeri 1 Pekanbaru using two units of the server computer, the computer unit is used as a head node (frontend) and one unit is used as a computer node computer-0-0. To improve performance or to balance server performance e-learning, the total amount of Central Processing Unit (CPU) in the cluster there are three units of Personal Computer (PC). It is intended that the resulting performance is able to provide better results again. The following physical network topology and logical topology created in this study.
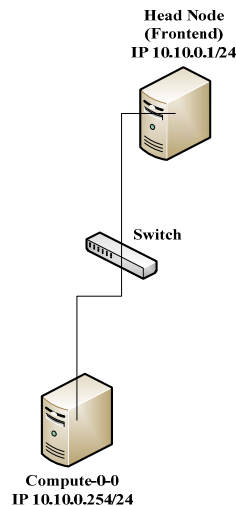

Figure 3. Physical network topology of computer Cluster

Explanation Network Topology Cluster System:

a) In the figure 3 topology, there is one head node computer unit (frontend) or master and one computer unit computer node or a slave.

b) Computer head node (frontend) serves as the main server in serving computer node to perform parallel processing to improve performance or to balance server e-learning.

c) On the head node computer (frontend) providing services to the IP address automatically computer nodes. Service will function after the IP address to configure Dynamic Host Configuration Protocol (DHCP) Server.

d) Network topology design used *Load Balance Cluster* (LBC) method.

In the method of Load Balance Cluster (LBC) concept for load balancing or workload to multiple server computers, for example divide the workload web server and database server. The workload of a network spread across the nodes of this is to improve performance in network computing.
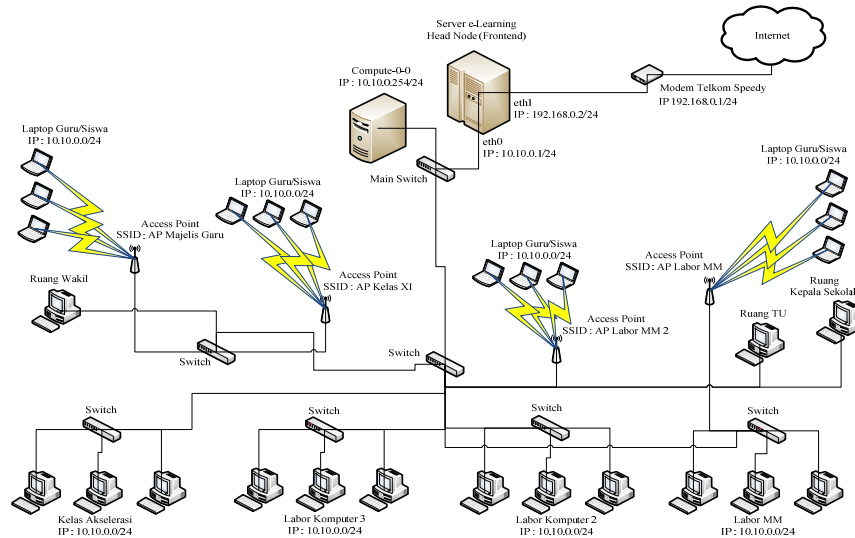

Figure 4. Computer network topology in Cluster (Parallel Processing) at SMA Negeri 1 Pekanbaru.

Based on the figure 4, server computer e-learning are connected in parallel itself of two units of the server computer, wherein the computer units serve as the head node (frontend) and one computer unit used as a computer node, the computer-0-0. Addressing the entire IP address connected in the network will obtain IP address automatically using the IP address of class A with a range 10.10.0.1 - 10.10.0.254.

Every computer user either teacher or student in environment that will access the e-learning intranet either using Personal Computer (PC) of the labor computer is connected to the switch via a cable or using a laptop connected wirelessly (Wi-Fi) using the Access Point (AP) attached evenly school environment. By using the Load Balance Cluster (LBC) is applied based on network topology above would be expected in the future if the teacher or student active simultaneously will access the e-learning will not slow down the performance of the server e-learning for both server computers that share the workload in servicing computer client.

5. Implementation

In order to develop this cluster, several stages to be done both in terms of configuration and installation of software and support package that will be useful for parallel processing. Here are the stages are performed in this study.

a) To determine the hardware requirements (hardware) and software requirements (software).
b) Installation of linux rocks 6.0 (mamba) on the head node computer (frontend).
c) Administration head node (frontend).
d) Installation of computer nodes.
e) Testing the cluster system (parallel processing).
f) Testing with the computer nodes.
g) Testing the client computer that will access e-learning.
h) Testing using a single client computer that access e-learning.
i) The test uses two client computers that accesses e-learning.

j) The test uses three client computers that accesses e-learning.
k) Allocation of memory.
l) Conclusion testing.

A. Hardware Requirement

Here are the specifications of the computer head node (frontend) and computer nodes used in this study.

1) Spesifikasi Komputer *Head node* (*Frontend*)

Berikut adalah spesifikasi komputer head node (frontend).

a) CPU Speed   : 2.40 GHz
b) CPU Model   : Intel(R) Core(TM) i5 2430M CPU
c) HDD Device   : SATA, 500 GB
d) USB Controler : NEC Corporation uPD720200 USB 3.0
e) IDE Device   : DVD-RW Multi
f) RAM   : 4 GB
g) Audio Device   : Intel Corporation 6 Series/C200
h) Ethernet Controller : Realtek Semiconductor Co., Ltd. RTL8101E
i) Motherboard   : Dell

2) Specification of Computer node (0-0)

Berikut adalah spesifikasi komputer Computer node (0-0).

a) CPU Speed   : 2.20 GHz
b) CPU Model   : Intel(R) Pentium(R) Dual Core CPU T4400
c) HDD Device   : SATA, 320 GB
d) USB Controler : Intel Corporation 82801I USB 2.0
e) IDE Device   : DVD-RW Multi
f) RAM   : 2 GB
g) Audio Device   : Intel Corporation 82801I
h) Ethernet Controllerb: Realtek Semiconductor Co., Ltd. RTL8101E
i) Motherboard   : HP

B. Software Requirement
Here is the need of software (software) used in this study.
a) Operation System Linux Rocks 6.0 (Mamba)
b) Message Passing Interface (MPI)
c) Ganglia
d) Apache web server
e) MySQL server

C. Cluster Testing System (*Parallel Processing*)
By default if everything configured the cluster we got it right on the head node computer (frontend) then we can use the application to monitor system ganglia cluster. The following graphic display is done in the study where the following image display is still using a computer stand alone server or has not happened the processes cluster (parallel processing).
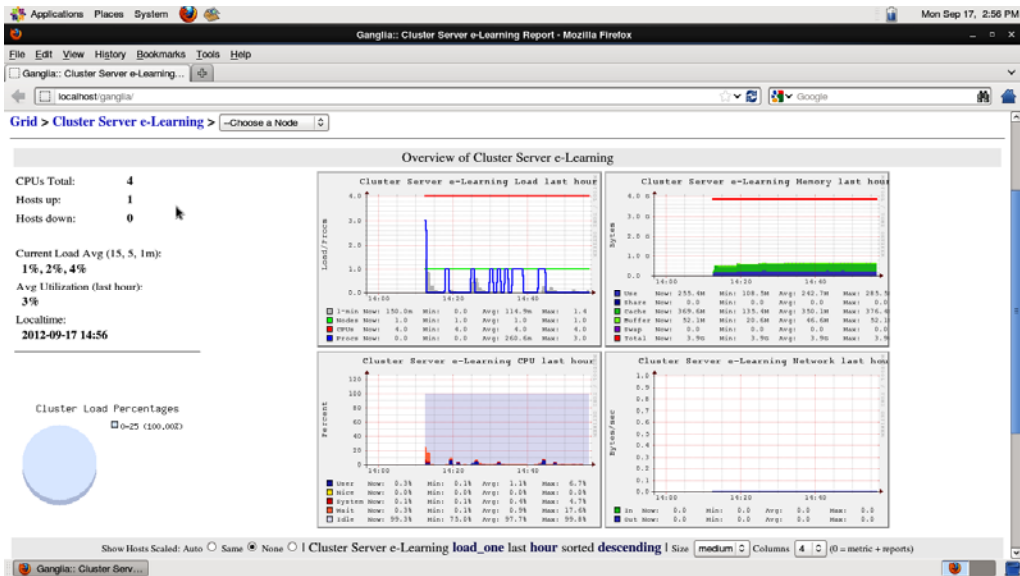
Figure 5. Monitoring System Server Stand Alone at Ganglia

Based on the description of Figure 5 shows a graphic display of a computer processor head node (frontend) that stand alone, meaning that not happened a parallel processing between the computer head node (frontend) to computer nodes. Total processor contained on the caption above, namely four (Quad Core). Host up to one unit that meaning that only the nodes that are running.

D. Cluster System (Parallel Processing)

The following graphic display is done in the study where the following image display has been a parallel processing between the computer head node (frontend) to the computer nodes (computer-0-0).
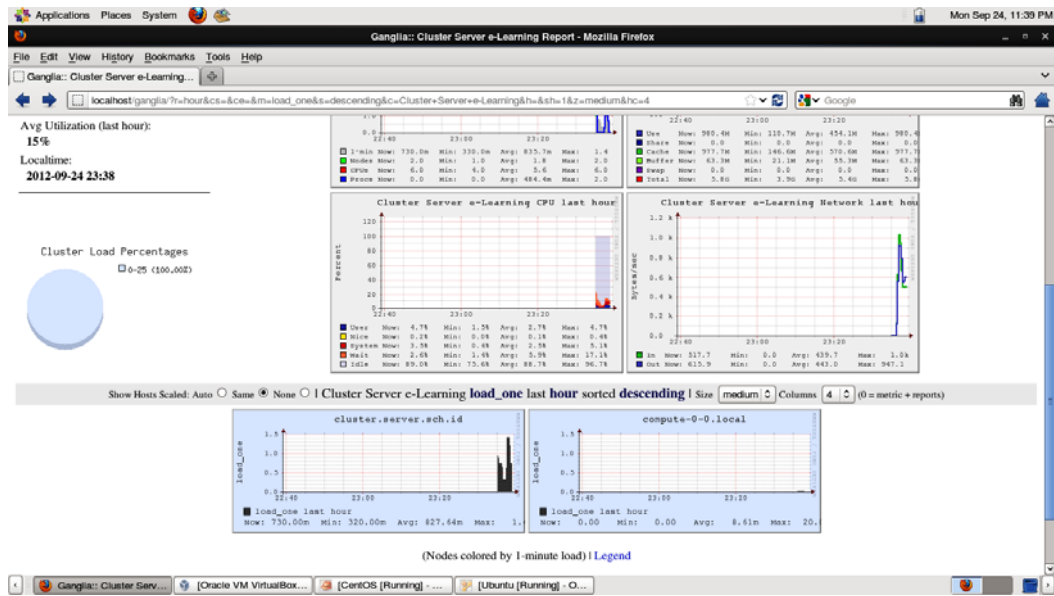

Figure 6. Cluster Monitoring System at Ganglia Used 2 Node

33

Based on information from the graph Figure 6 looks the two nodes shown on application monitoring (ganglia) that cluster.server.sch.id and computer-0-0.local, meaning between the computer head node (frontend) and computer nodes (computer-0 -0) are connected in parallel (cluster). The information that can be seen from the above picture by default will be visible all the resources that exist in the system, based on what tab shows the metric system in monitoring, can select a category. For more detail see Cluster status we can choose the Top Cluster shows information process from each node in the cluster system. Total processor contained on the caption above those six pieces, where the head node computer (frontend) is using a Quad Core processor and computer nodes (computer-0-0) using processor Dual Core. Host up two that meaning there are two nodes that are running.

E. Testing Used Computer Note

Once everything is installed properly it will display the image below, enter login "com0" with password "com0" on computer-0-0. Computer nodes will automatically log them centrally on the head node computer (frontend). Figure 7 shows of computer-0-0 and computer-0-1 were performed in this study.
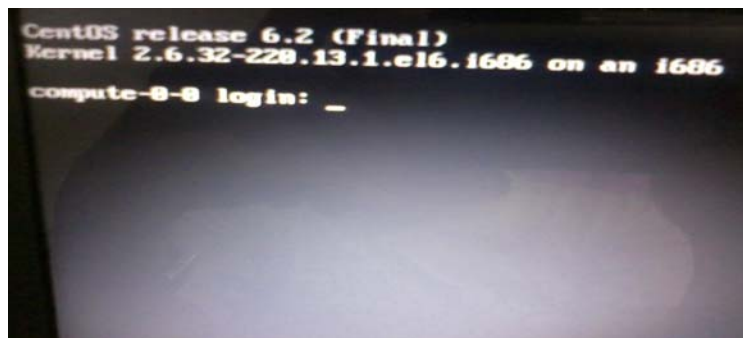

Figure 7. Login display of Computer-0-0

After all installed properly to ensure computer head node (frontend) and the computer nodes are integrated correctly, then we will select the tab "Physical View" in the monitoring system in the ganglia so that the display appears as below.
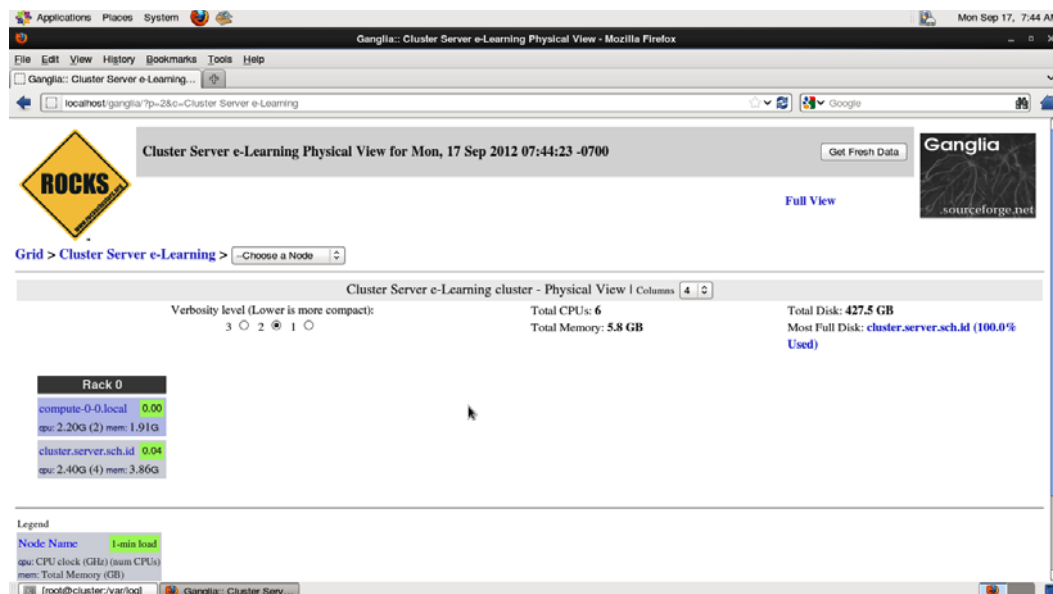

Figure 8. Display of Physical View at Ganglia

34

Based on information shows in figure 8 the head node computer (frontend) and the computer nodes are integrated properly. In other information that can be seen from the description of the image above the number of processors, processor speed, total memory, total hard drive.

F. Testing with *Client Computer* Accessing *e-Learning*

The next stage is to test from a client computer to access e-learning. In this experiment will be accessible using a web browser Mozilla Firefox, from the server side using Apache as the web server and MySQL as the database server of its.

Figure 9. One of user login in e-Learning

Based on information from the figure 9 shows the user Abdul Syukur was login at page e-learning on the computer head node (frontend).

G. Testing Used Two Unit of Client Computer *Client* to Access *e-Learning*
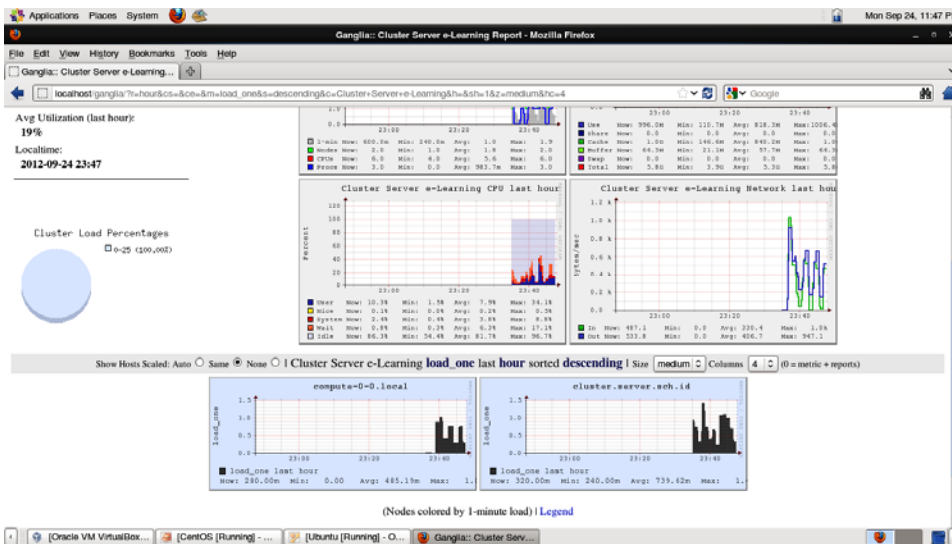

Figure 10. Cluster monitoring system used 2 Unit of client computer to access e-Learning

Based on information from the computer graphic image shows a cluster in figure 10 (parallel processing) where the condition of the computer graphics head node (frontend), which provides e-learning web server being accessed using two units of the client computer. At the time of the client computer to access e-learning server looks graph between cluster.server.sch.id and computer-0-0.local traffic will be increase and balance, reflecting the diversity of traffic that is displayed using three units of the client computer than using two computers client and the client computer unit is accessing e-learning.

H. Memory Allocation

In this case the memory usage will certainly be indispensable in doing parallel processing. Every memory usage will assist in the process of balancing cluster to improve server performance or e-learning. Table 1 is shows the result of the conclusion of use of available memory.

Table 1. Memory used

Node Cluster	1 Node	2 Node	Total
Ukuran Memori	3.86 GB	1.91 GB	5.8 GB

6. Conclusion

The conclusions in this test are:

1. The entire cluster has been able to run as expected.
2. The client computer successfully accesses the e-learning web server that no computer head node (frontend).
3. From the results of testing the use of a single node processor (Cluster Load Percentages) is 100% and the use of two nodes using a processor (Cluster Load Percentages) is 50%. To use cluster.server.sch.id processor 0-25 (50%) and for the use of the processor computer-0-0.local 25-50 (50%).

References

Fatullah Fani, Mutiara MQN A. Benny, Yulianto Chandra, (2004) "Analisa Kinerja *Cluster Linux* Dengan Pustaka *MPICH* Terhadap Perkalian Matriks", KOMMIT 2004, Jakarta.

Harsono Nonot, Sukaridhoto Sritrusta, Subhan Ahmad, Sudarsono Amang, (2006) "*Model Distributed Data Processing* Berbasis *Open Source* Untuk Pengolahan Data Kependudukan Dalam Rangka Implementasi *e-Gov* Indonesia", Prosiding Konferensi Nasional, Bandung.

Hidayat Syarif, (2006) "Pemrosesan Paralel Menggunakan Komputer Heterogen", SNATI 2006, Yogyakarta.

Kurniawan Agus, Arsyad Sagi, (2010) "Desain dan Implementasi Komputasi Paralel Pada *Excel*", SNATI 2010, Yogyakarta.

Lauwali Tomiputra, Kartawidjaja A. Maria, (2008) "Komputasi Paralel Menggunakan *Parallel Virtual Machine* Untuk Peningkatan Kinerja Komputasi", Jurnal Elektro, Jakarta.

Liswandini Intan, Irawan Budhi, Irzama, "Studi Komparatif Antara *Parallel Virtual Machine* (PVM) dan *Message Passing Interface* (MPI) Dengan Memanfaatkan *Local Area Network* (LAN)", Jurnal, Bandung.

Prianto Bayu, (2008) "*Cluster* Komputer Sebagai Pengganti Super Komputer Tunggal Untuk Pemodelan Kimia Komputasi" Jurnal Berita Dirgantara.

Siregar Simon, Ismail, Hendri Rio, (2010) "Pengukuran Kecepatan Proses Pada *High Performance Computing* (HPC) *Cluster Server*", Jurnal, Bandung.

Suwarningsih Wiwin, Nuryani, Arisal Andria, Wirahman Taufik, Masthurah Nurhayati, (2009) "Membangun *Cluster* Dengan Menggunakan *IGOS* Dwi Warna", SNATI 2009, Yogyakarta.

Widyatmoko Amri, Halimurahman, (2006) "Manajemen *Linux Cluster* Menggunakan *Rocks 4.1* dan Integrasi Model *Atmosfer WRF*", Prosiding Semiloka Teknologi Simulasi dan Komputasi Serta Aplikasi, Bandung.