

Building a Parallel Low-cost Computers

Jeffrey John R. Yasay

Assistant Professor

Department of Computer Studies

Tarlac Agricultural University, Camiling, Philippines

Abstract: This study provides guidelines and procedure on how to develop a low-cost parallel clustered computer. It aims to disseminate information in building a high-performance computer using ABC GNU Linux. Different configurations using the command line interface and graphical user interface were conducted. The processes of the computer were determined by the use of graphical representation using Ganglia as a monitoring tool. Rapid application development was used to establish the methodology in creating these parallel computers.

Index Terms – High-performance, low-cost, parallel, computer.

I. INTRODUCTION

Building a super computer has been a dream of every technology enthusiast, in recent years calculation has progressed to unimaginable heights. Larger volumes of data are being collected from a variety of sources, including web pages, social networks, and IoT devices, while more advanced algorithms are being used in new applications such as personalized advertising, user behaviour mining, and visual/audio recognition. To fulfill the demand, service providers such as Google, Microsoft, and Amazon are expanding their datacenter infrastructures. The semiconductor technology, on the other hand, is approaching its physical scaling limit [1]. Computer clusters are the outcome of the convergence of several computing phenomena, such as high-speed networks, low-cost microprocessors, and software for high-performance distributed computing. Rather of managing IO-oriented procedures, computer clusters are employed for computation-intensive tasks [2]. Big data necessitates the advancement of datacenter infrastructure in both hardware and software. The datacenters data processing and computing capacities will be considerably boosted as the volume of structured and unstructured data continues to grow. Although contemporary datacenters provide hardware and data storage, they must also take on additional tasks, such as data acquisition, analysis, and organization. [3]. All machines in homogeneous clusters are supposed to be the same; nevertheless, in heterogeneous clusters, machines have varying computation and consumption power. thus, a single chip's power density can no longer be increased [4]. In many modelling situations, parallel computers are used. Parallelism is determined by machine hardware, with multicore processors containing two or more independent cores and multiprocessor computers containing multiple processing elements within a single machine. Many of today's complicated numerical model programs are already parallelized in a visible fashion for users, have high portability between parallel architectures, have good scalability over a range of processors, and can also be executed in their original serial version on a single computer [5].

II. METHODOLOGY

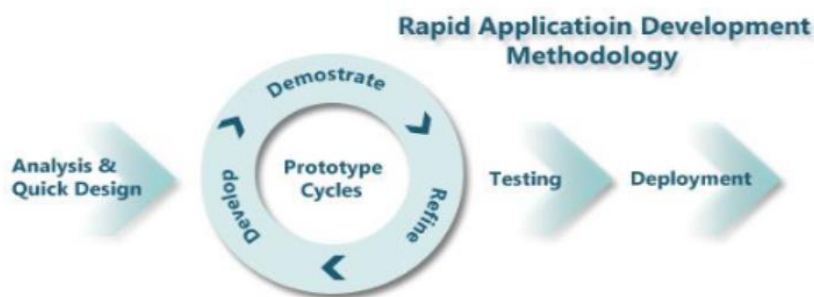


Fig. 1 Rapid application development

Figure 1 depicts the process and methodology in creating a parallel cluster [6].

III. TOOLS

The system was designed using the hardware and software requirements specifications as a foundation.

A. Personal computer (PC) - A personal computer is self-contained and has its own processor. Organizations require a multi-user environment that allows for data sharing as well as the sharing of costly resources like as printers and storage.[7].

B. Fast Ethernet Switch - It can deliver data transfer speeds of up to 100 Mbps, indicating high-speed LAN technology [8].

C. Straight Through - A straight through Ethernet cable is used and the pairs will match up properly. [9].

D. Ubuntu ABC - Based distribution allows to automatically build Beowulf clusters either live or installing the software in the frontend.[10].

To maximize the system's benefits, homogeneous computing is used to connect similar processing cores or units to create a high-performance computer [11]. All nodes, including the master node, adhere to the same requirements.

The next step is to install the software after the machines have been constructed. There are numerous Linux distributions available, and different people favors various distributions for various reasons. As a foundation, use ABC GNU Linux (Ubuntu 9.04) with the default kernel.

IV. SPECIFICATIONS AND TESTING

A. Master Node and Slave Nodes

Processor: Intel Celeron M
 CPU Number: 1
 CPU Speed: 2266 MHz
 Total Memory: 509452 Kb
 Hard Disk Size: 40 Gb
 Built-in LAN Card: Intel UNDI, PXE-2.0 (build 082)
 Copyright © 1997, 1998, 1999 Intel Corporation VIA Rhine II Fast Ethernet Adapter V2.38 (2004/09/15)
 External LAN Card: SiS (Model 2003)

B. Setting up of ABC GNU

1. Boot from the CD-ROM
2. Choose an install mode, press enter then follow the directions on the screen.
3. Select use entire disk to partition the hard disk.
4. Create username and password
5. Install ABC GNU (Ubuntu 9.04)
6. Connect cables to the fast ethernet switch
7. Run/load slave node network boot
8. Test base linear of master and slave nodes

```

master@master-desktop:~$ recon -v clusterhosts
n-1<4633> ssi:boot:base:linear: booting n0 (192.168.0.1)
n-1<4633> ssi:boot:base:linear: booting n1 (192.168.0.13)
n-1<4633> ssi:boot:base:linear: booting n2 (192.168.0.3)
n-1<4633> ssi:boot:base:linear: booting n3 (192.168.0.10)
n-1<4633> ssi:boot:base:linear: booting n4 (192.168.0.8)
n-1<4633> ssi:boot:base:linear: finished
  
```

 Woo hoo!

recon has completed successfully. This means that you will most likely be able to boot LAM successfully with the "lamboot" command (but this is not a guarantee). See the lamboot(1) manual page for more information on the lamboot command.

If you have problems booting LAM (with lamboot) even though recon worked successfully, enable the "-d" option to lamboot to examine each step of lamboot and see what fails. Most situations where recon succeeds and lamboot fails have to do with the hboot(1) command (that lamboot invokes on each host in the hostfile).

```

-----
master@master-desktop:~$
  
```

-).
9. Testing the LAM Boot using command line interface.
- ```

master@master-desktop:~$ lamboot -v clusterhosts
LAM 7.1.2/MPI 2 C++/ROMIO - Indiana UniverResity
n-1<4651> ssi:boot:base:linear: booting n0 (192.168.0.1)
n-1<4651> ssi:boot:base:linear: booting n1 (192.168.0.13)
n-1<4651> ssi:boot:base:linear: booting n2 (192.168.0.3)
n-1<4651> ssi:boot:base:linear: booting n3 (192.168.0.10)
n-1<4651> ssi:boot:base:linear: booting n4 (192.168.0.8)
n-1<4651> ssi:boot:base:linear: finished

```

```

master@master-desktop:~$

```

10. To view and monitor the clustered computer open Ganglia using graphical user interface

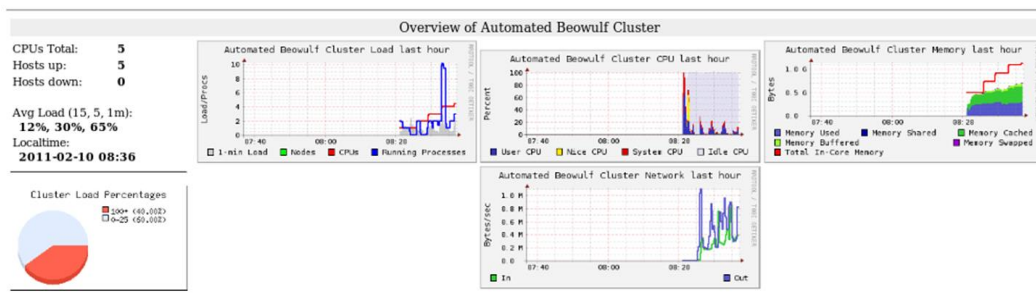


Fig. 2 Cluster view using Ganglia

Figure 2 depicts the system's cluster view. It also implies the cluster's CPU counts and summary graphs. The pie graphic depicts how much of the cluster is active. Each CPU's node processing is depicted by a series of tiny graphs.

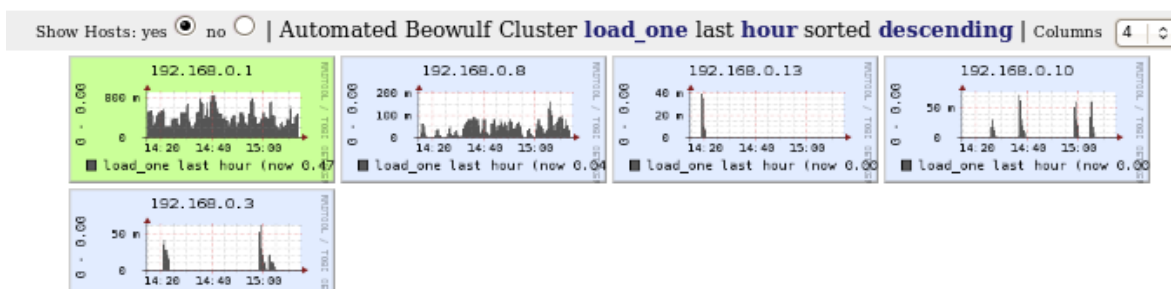


Fig. 3 Cluster view of each node

Figure 3 depicts the cluster view of each node. The master node's IP address is 192.168.0.1, and there are four slave nodes. It also displays the various reading processes of each slave node.

### V. RESULTS

The Ganglia monitoring tool [12] was used as the basis in evaluating and monitoring phase, snapshots of the system were taken every hour.

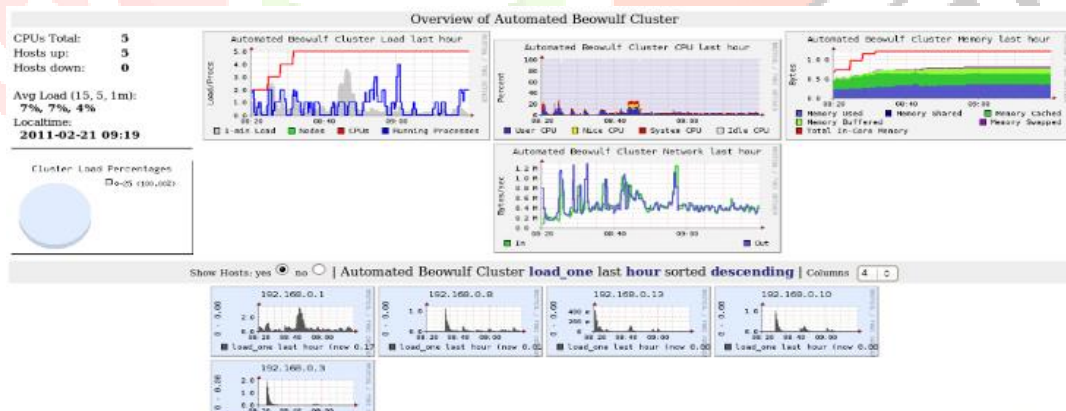


Fig. 4 Overview of Automated Beowulf Cluster

### VI. CONCLUSION

Based on the given results and setup, the clustered computer was successfully developed. It exhibited high-performance based on the figures. Master node and slave nodes were also determined using different interfaces and commands. It is also concluded that the use of this configuration and procedure enhanced the capability of certain computer enthusiast to develop more high-performance parallel computers.

Different set of machine configuration must also be tested to perform higher performance. It is also recommended finding other monitoring tools to evaluate the system's performance. Large data collections are being gathered in various IT systems all around the world. Many of them exhibit a networked aspect and need to be processed and studied as graph structures.

### VII. ACKNOWLEDGMENT

This study would not be possible without the support of the Tarlac Agricultural University Tarlac, Philippines.

## REFERENCES

- [1] Cong, Jason; Huang, Muhuan; Wu, Di; Yu, Cody Hao (2016). [ACM Press the 53rd Annual Design Automation Conference - Austin, Texas (2016.06.05-2016.06.09)] Proceedings of the 53rd Annual Design Automation Conference on - DAC '16 - Invited - Heterogeneous datacenters., (), 1–6. doi:10.1145/2897937.2905012
- [2] Karan Kantharia, Adarsh Chaturvedee, Manas Nikam, Rujul Shringarpure, Disha Bhosle, 2017, Supercomputer using Cluster Computing, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICIATE – 2017 (Volume 5 – Issue 01),
- [3] P. Castoldi, N. Androioli, I. Cerutti, O. Liboiron-Ladouceur, P.G. Raponi, Energy efficiency and scalability of multi-plane optical interconnection networks for computing platforms and datacenters. in: Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012. <http://dx.doi.org/10.1364/OFC.2012.OW3J.4>.
- [4] Mehdipour, Farhad (2016). [Advances in Computers] Energy Efficiency in Data Centers and Clouds Volume 100 || Energy-Efficient Big Data Analytics in Datacenters. , (), 59–101. doi:10.1016/bs.adcom.2015.10.002
- [5] . Torres, R. (2011). Treatise on Estuarine and Coastal Science || Modeling of Estuarine and Coastal Waters. , (), 395–427. doi:10.1016/b978-0-12-374711-2.00216-3
- [6] Rapid Application Development. 2017. Retrieved January 15, 2018 from <http://www.ramsoft.com.au/methodology.php>
- [7] The Editors of Encyclopaedia Britannica. 2017. Personal computer. (December 2017). Retrieved June 8, 2018 from <https://www.britannica.com/technology/personal-compute>
- [8] A. Gunasekaran, (2001). Agile Manufacturing: The 21st Century Competitive Strategy. doi:10.1016/b978-008043567-1/50040-1
- [9] W. Wu, "The Design of a New Network Cabling Experimental Instrument Based on Embedded System." Advanced Materials Research 328, pages 2427-2431, 2011
- [10] Castaos, I & Garrido, Izaskun & Garrido, Aitor & Sevillano, M. (2009). Design and implementation of an easy-to-use automated system to build Beowulf parallel computing clusters. 1 - 6. 10.1109/ICAT.2009.5348420.
- [11] Irimia R, Gottschling M (2016) Taxonomic revision of Rochefortia Sw. (Ehretiaceae, Boraginales). Biodiversity Data Journal 4: e7720.
- [12] Matt Massie. 2013. Monitoring with Ganglia, Sebastopol, CA: O'Reilly.

