

THE AFFORDABLE SUPERCOMPUTER



**HARRISON CARRANZA
APARICIO CARRANZA
JOSE REYES ALAMO**

**CUNY - NEW YORK CITY COLLEGE
OF TECHNOLOGY**

***ECC Conference 2015 – June 14-16, 2015
Marist College, Poughkeepsie, NY***

OUTLINE

INTRODUCTION

WHAT IS COMPUTER CLUSTERING?

WHY USE COMPUTER CLUSTERING

ADVANTAGES OF COMPUTER CLUSTERING

WHAT IS MOSIX?

SIMPLE LOAD TEST

FEATURES OF CLUSTER COMPUTING

DATA RESULTS

CONCLUSION

INTRODUCTION

- Enterprises work to resolve complex calculations that require heavy usage of CPU
- The ideal machine to use is the mainframe
- Though efficient, mainframes are expensive
- Cost effective solution is introduced

WHAT IS A COMPUTER CLUSTER?

- Computer clustering is a technology that allows two or more computers to be interconnected through a LAN
- It allows us to gather resources in order to perform CPU intensive tasks
- High Availability clusters (HA) and High Performance clusters (HP)
- Its functions include parallel processing and load balancing

WHY USE A COMPUTER CLUSTER?

- Enterprises and corporations require the processing power of a supercomputer or mainframe
- Select few are able to afford expensive machines
- Main idea is to maximize efficiency and lower financial cost
- Creating a cluster using existing desktops and laptops, all the unused processing power could be used to improve capabilities
- Clustering solutions are available on Linux due to its reliability, security, and free accessibility

ADVANTAGES OF CLUSTER COMPUTING

- **Cost effective** – small enterprises or corporations could afford it
- **Parallel processing power** – more cost effective than a mainframe of that amount of power
- **Modern networking** – lagging is minimized and it guarantees that data will not bottleneck the system
- **Scalability** – adding more nodes to the network shall give you more power than using a mainframe, which has a fixed processing capacity
- **Redundancy** – if one node fails, the data is transferred to another node whereas a mainframe failing could cause the data to be lost

WHAT IS MOSIX?

- **MOSIX** is an extension to the **Linux** kernel for single system image clustering, known as a kernel patch
- A network of regular computers could be put as one supercomputer
- It is capable of creating a cluster via the network, allowing users to do less work in setting it up
- Once a computer running **Linux** is patched using **MOSIX**, we could gain access to create and join a computer cluster

APPROACH

- Introduce MOSIX as a cost-effective solution for enterprises
- A better option than mainframes for small organizations that lack a huge IT budget
- Perform a load test to demonstrate performance increase of computers inside of a cluster network

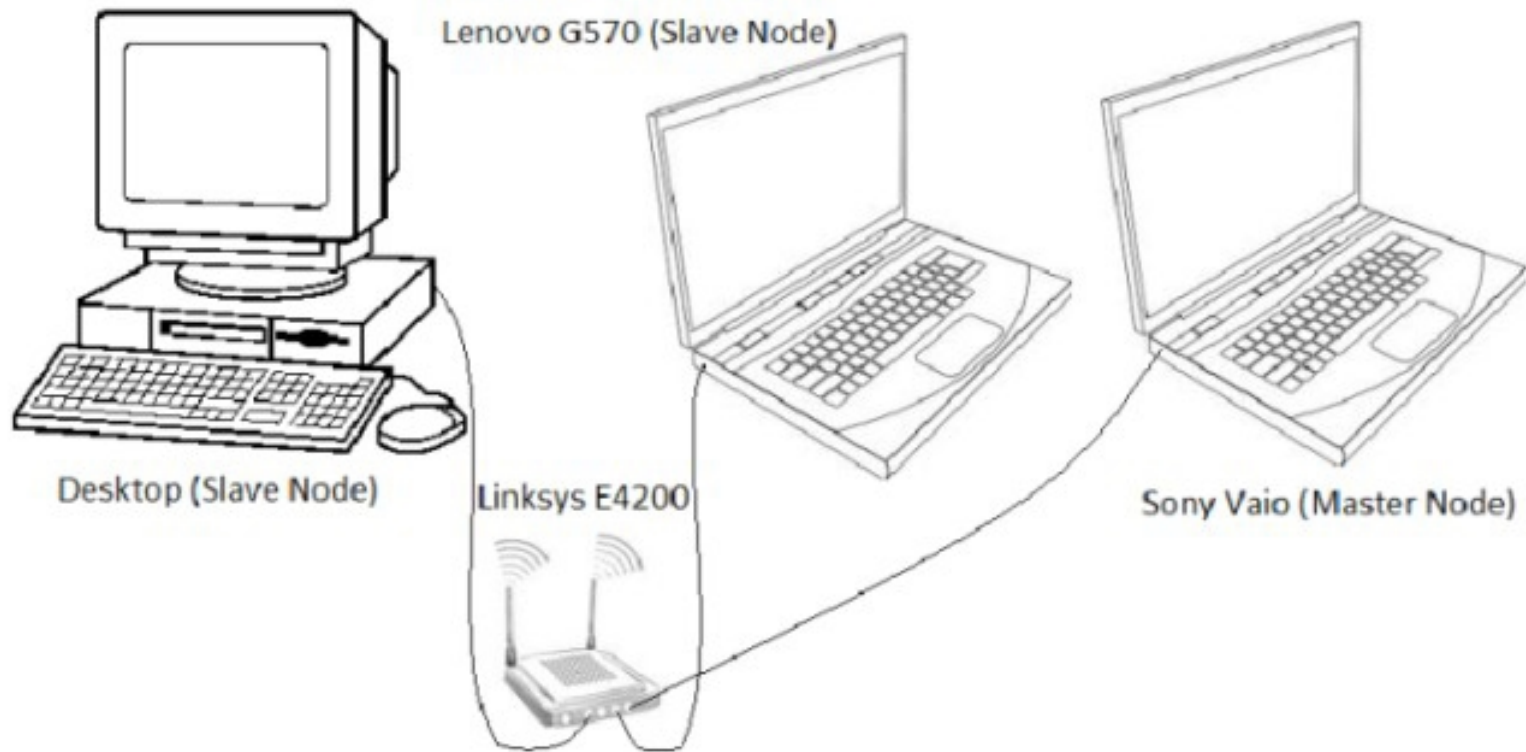
MATERIALS USED

- **Lenovo G570**
- **Intel B940 @ 2.00GHz**
- **3GB RAM**
- **SONY Vaio**
- **Intel Pentium P6100 @ 2.00GHz**
- **4GB RAM**
- **Desktop PC**
- **Intel i5-2500k @ 4.5GHz**
- **16GB RAM**

SIMPLE LOAD TEST

- To perform the load test, each computer runs `Clusterknoppix v3.6`
- It is preinstalled with `openMOSIX`, open source version of MOSIX
- A simple shell script runs 17 times on each system to overload the cluster network
- This application loads system to use its maximum processing power
- The script commands the computer to count from 1 - 10000
- This is performed 10000 times
- `(awk 'BEGIN{for(i=0;i<10000;i++)
For(j=0;j<10000;j++);}'`

LOAD TEST SETUP



Setup requires 3 PCs and a router / switch.

CLUSTERKNOPPIX

The screenshot displays the Clusterknoppix interface, which is a Linux-based environment for cluster management. It consists of three main windows:

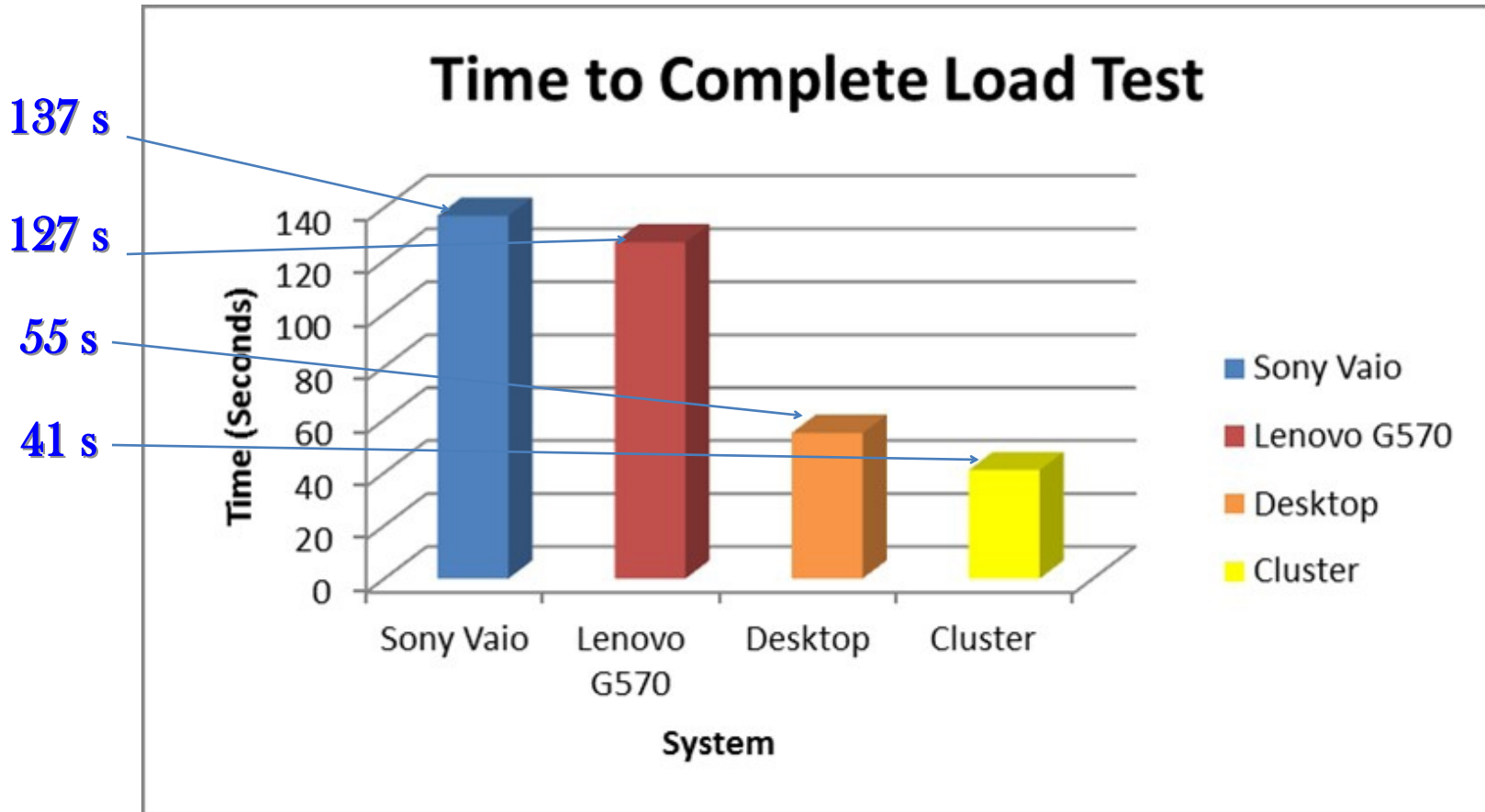
- Terminal Window (Shell - Konsole):** Shows a series of shell commands and their outputs. The commands are used to run a benchmark script (awk) on multiple nodes. The outputs show the completion of the script and the number of iterations completed (e.g., [14] Done, [31] 4697, [32] 4698, [33] 4699, [34] 4701).
- Network Diagram (openMosix/mignon 1.5):** A visual representation of the cluster topology. It shows three nodes (labeled 249, 1, and 250) connected by dashed lines, indicating network links. The nodes are arranged in a circular pattern, with node 1 in the center and nodes 249 and 250 on the left and right respectively.
- Monitoring Dashboard (openMosixview 1.5):** A graphical interface for monitoring the cluster's performance. It displays a table of metrics for each node and overall cluster statistics.

id	clusternodes	load-balancing efficiency	overall load	overall used memory	all memory	all cpu
all	all-nodes	95%	70%	4%	3072 MB	6
1	192.168.0.1	30000	64%	4%	1024	2
249	192.168.0.249	49638	74%	4%	1024	2
250	192.168.0.250	30000	72%	4%	1024	2

started the openMosixmignon

Clusterknoppix comes prepackaged that contains a number of applications to monitor and manage the cluster

LOAD TEST TIMING



Load test results demonstrate the performance increase of a simple task that the cluster network achieves

FEATURES OF CLUSTER COMPUTING

- Openmosix terminal server - is able to boot linux clients through the network via DHCP and PXE
- CDRoms and other media are unnecessary
- Openmosix's auto discover feature automatically joins nodes to the cluster network
- Cluster Management tools - tools such as Userland and openMOSIX are pre-packaged
- Every node in the cluster runs an X server which allows the user to use a GUI interface

DATA RESULTS

- The Sony Vaio received a 70% increase in performance while the Desktop PC only received 25% increase in performance
- The cluster proved to increase performance for all of the computer nodes on the network
- As more computers were added, the scalability of network cluster performance also increased

CONCLUSION

- The openMOSIX extension to the Linux kernel is a cost effective solution to increase processing power
- Building a cluster allows organizations to increase processing power by adding nodes to the network
- If a node is not using its available processing power, then other nodes are able to use it efficiently to complete a work load
- In the event that one node fails, the other nodes are there to acquire the transferred files from that node
- Cluster computing applies modern networking, meaning that data is not going to bottleneck the system

THANK YOU!!

QUESTIONS??