IMPLEMENTATION OF A LOAD BALANCING CLUSTER IN A COMPUTER NETWORK

Ramesh Nair, Salma Shaheen and Victor Raj
School of Computing & Information Technology, INTI College Subang Jaya, Selangor, Malaysia
(victor@inti.edu.my)

ABSTRACT

This paper investigates the implementation of load balancing and process migration on existing servers and workstations that allows cost reduction, stability, scalability, time saving and resource maximization. The concept of load balancing is ideal in situations where computing tasks need to be executed in a fast and efficient manner with maximum availability. Web servers are a typical example of heavily loaded systems, which require high availability. In high traffic sites, web servers tend to slow down if only a single server is used to cater to the requests. Load balancing solution provides an answer to the problem faced by web servers. Load balancing implementation allows the load to be shared amongst existing networked systems configured to carry out load balancing.

This paper also investigates the capability and efficiency of the load balancing technique with a direct test on a cluster/network of computers with different configurations. A load balancing technique will be deployed and it is called Multi-computer Operating System for Unix (MOSIX) (Barak et. al.,1999). MOSIX moves processes around the cluster/network to balance the load, using faster machines first, without user intervention. This means if a node or a server is overloaded beyond a certain threshold level then the process will be transferred to another node or a server.

INTRODUCTION

Load balancing systems have been the subject of substantial research over the past decades. The importance of these systems has remained, from the viewpoint of practice, owing to the phenomenal success of low-cost networks of workstations, large-scale information infrastructures containing replicated servers, and multiprocessor (SMP's) architectures that enable fast, concurrent execution of fine-grained threads. Distributing workload and also the processing capabilities among servers or systems allows the increase of speed, manageability, scalability, flexibility and reliability.

MOSIX Load balancing is defined as the fine-tuning of a computer network, or a disk subsystem in order to distribute the data or process across available resources in a network of computers (Barak et. al., 1999). The MOSIX technology consists of two parts, a Pre-emptive Process Migration (PPM) mechanism and a set of algorithms for adaptive resource sharing. Both parts are implemented at the kernel level, using a loadable module, such that the kernel interface remains unmodified (Barak et. al., 1999).

TESTING METHODS

The load-balancing cluster is benchmarked with programs designed to stress test the systems. The tests are done on CPU, memory and network. An initial test bed consists of

systems without load balancing capabilities. This is to allow comparison to be made on the performance of systems with and without load balancing. Benchmarks are run on the systems and their results are obtained and analysed in this paper.

The tests are conducted in a controlled environment with minimal, medium and heavy load to test the system's overall capability. All the systems used in this work are set up by using Linux operating system with networking options enabled.

Test instruments

Four to five benchmark programs/softwares have been identified for this work to be carried out.

The following benchmark Software are used to test the network and load-balancing cluster, which consists of eight different types of systems.

- Netperf : A Network Performance Benchmark. Rev.2.0 (Netperf, 1995).
- Ntop: Network Top, Network Analysis Tool Rev. 3.0 (Luca Deri, Finsiel, 2000).
- POVRAY: Persistence of Vision Ray tracing utility. (Buck et. al., 1999).
- CacheBench : To show the performance of the CPU Memory (Mucci et. al., 1998).
- NetPipe 2.4: To show the throughput and network performance.
- Nettest: To benchmark the network performance.
- TTCP: Program to test the Networks capability

The tests are done on CPU, memory and network. The Load balancing system that has been identified to be used in the experiment is MOSIX.

TEST RESULTS AND DISCUSSION

POVRAY is a ray-tracing utility used to conduct a CPU bound process test within the systems in the cluster. Initial tests with non-MOSIX systems showed that the systems performed similar to individual nodes in a cluster, as there is no interaction or algorithms available for interactions. The cluster test bed consists of eight nodes with each running identical OS, Linux RedHat 7.1 with kernel release 2.4.10, in variation of resources such as CPU speed, RAM size, Bus Speed and disk space.

With the introduction of the MOSIX cluster technology, the systems that performed slower in the initial tests improved as the load balancing algorithms are put to use. The performance of the slower systems improved according to the highest available resource in the cluster, which is a Pentium III, 600 MHz machine (Delta from table 1). The initial values of the cluster without load balancing are compared and an improvement of performance is noticeable. To a certain extent the slower systems gained more advantages as they performed as well as the high-end systems in the cluster after load balancing is applied. For example from table 1, an increase of 428.18% was noted on Kappa when load balancing is introduced which is a very good improvement. At this stage the network load and the network equipment efficiency has not been tested.

The second test has also been conducted in a similar manner but with additional load. Table 2 compares the systems/clusters performance and obtains the overall performance profile of the cluster.

With the introduction of more CPU bound processes it has been noticed that the introduction of algorithms into the cluster allowed the work units or processes to be moved to the other nodes in the cluster. The MOSIX