ISSN: 2277-9655
Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Performance Evaluation of Cluster Computing

K D.Kavitha[*1], RojaRamani.Adapa[2]
[*1,2] Assistant Professor, TKR College of Engineering & Technology, Meerpet ,Hyderabad, India
davidikavitha2011@gmail.com

Abstract

Cluster Computing addresses the latest results in these fields that support High Performance Distributed Computing (HPDC). In HPDC environments, parallel and/or distributed computing techniques are applied to the solution of computationally intensive applications across networks of computers. A cluster computing is a type of parallel or distributed computer system, which consists of a collection of interconnected stand-alone computers working together as a single integrated computing resource. The key components of a cluster include multiple standalone computers (PCs, Workstations, or SMPs), operating systems, high-performance interconnects, middleware, parallel programming environments, and applications. It assumes that the reader is familiar with the standard commodity hardware and software components such as stand-alone computers, operating systems such as Linux and Windows, and standard communication software such as TCP/IP. There are many applications which can benefit from parallelisation. Employing clusters of computers provides a method to utilise commodity components, minimising cost and and maximising longevity of the individual parts.

Keywords: Introduction, High Performance, MPI

Introduction

Clusters are usually deployed to improve performance and availability over that of a single computer, while typically being much more cost-effective than single computers of comparable speed or availability. Employing clusters of computers provides a method to utilise commodity components, minimising cost and and maximising longevity of the individual parts. Recent advantages in network and workstation performance have made clustered computer systems an interesting alternative to massively parallel computers. There are many operating systems, languages, programming environments and tools, used in design, optimization and debugging of parallel programs. A cluster is a type of parallel or distributed processing system, which consists of a collection of interconnected stand-alone computers cooperatively working together as a single, integrated computing resource. Workstation clusters are a cheap and readily available alternative to specialized High Performance Computing (HPC) platforms.Main benefits of using clusters are:

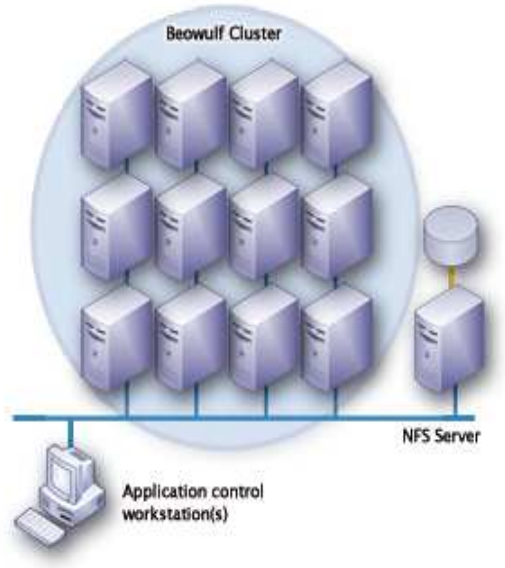
- **System availability** - clusters offer inherent high system availability due to the redundancy of hardware, operating systems, and applications
- **Hardware fault tolerance** – it is based on redundancy for most system components including both hardware and software
- **OS and application reliability** - clusters run multiple copies of the OS and applications
- **Scalability** – it is possible to add servers to the cluster or add more clusters to the network as the need arises
- **High performance** – it is possible to put together a cluster with equal or greater power than a single large machine.

High Performance Computing

High Performance Computing (HPC) allows scientists and engineers to solve complex science, engineering and business problems using applications that require high bandwidth, low latency networking, and very high compute capabilities. Typically, scientists and engineers must wait in long queues to access shared clusters or acquire expensive hardware systems. Using Amazon EC2 Cluster instances, customers can expedite their HPC workloads on elastic resources as needed and save money by choosing from low-cost pricing models that match utilization needs. Customers can choose from Cluster Compute or Cluster GPU instances within a full-bisection high bandwidth network for tightly-coupled and IO-intensive workloads or scale out across thousands of cores for throughput-oriented applications. Faster computers allow us to solve larger problems, and to find solutions more quickly, with greater accuracy, and at a

ISSN: 2277-9655
Impact Factor: 1.852

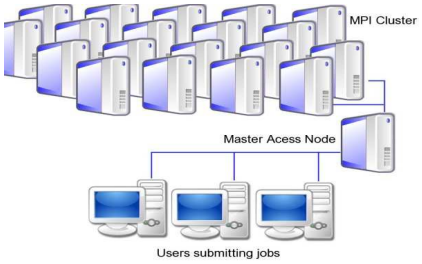
lower cost. All this adds up to a competitive advantage. High performance is a key issue in data mining or in image rendering. Traditional high performance clusters have proved their worth in a variety of uses from predicting the weather to industrial design, from molecular dynamics to astronomical modeling. A multicomputer configuration, or cluster, is a group of computers that work together. A cluster has three basic elements—a collection of individual computers, a network connecting those computers, and software that enables a computer to share work among the other computers via the network. Clusters are also playing a greater role in business.



High Performance Clusters
Fig:1

Traditional high-performance clusters have proved their worth in a variety of uses—from predicting the weather to industrial design, from molecular dynamics to astronomical modeling. High-performance computing (HPC) has created a new approach to science—modeling is now a viable and respected alternative to the more traditional experiential and theoretical approaches. High performance is a key issue in data mining or in image rendering. Advances in clustering technology have led to high-availability and load-balancing clusters. Develop the algorithms for the Sorting that are faster and more accurate than they were ever before. The problem with the huge data sorting is that it takes a lot of time and as a result a large amount of money is required. This huge money is just wasted in the sorting of data and not in any useful work so, it is highly required that this time be as less as possible. So we came up with the some new methods to lower that Complexity. The approach we followed was High Performance Computing.The field of High Performance Computing is the most popular for getting faster results for any application. The applications designed using High Performance Computing concepts are Numerical Integration, Quick Sorting. Application for Numerical Integration is designed for calculating the integration of function cos(x) under the limits 0 to pi/2. Such integration application can be easily executed over several nodes simultaneously so as to reduce the work load on single computer for calculating the integration of function completely.The numerical Integration algorithm can divide a big problem of integration into several smaller problems of integration. For decomposing the problem, the area under the curve of the function has to be divided into rectangles. These smaller segments of problem can be distributed over the cluster (i.e. the nodes) so that they can be executed simultaneously and hence result can be produced in a shorter time than computing the whole problem on a single machine. Using the concept of Message Passing Interface, we can also select the number of processes in which we want our problem to be decomposed and executed. On increasing the number of processes, the result can be generated by the processors more accurately.

Message Passing Interface



Message Passing Interface (MPI)
Fig:2

The Message Passing Interface (MPI) standard defines a programming model for message passing, along with a series of functions to support that programming model.MPI processes are named, and processes send and receive messages in a point-to-point fashion based on process name. Processes can be grouped, and collective communication functions can be used to perform global operations on a group, such as broadcast and synchronization. Message exchanges in MPI can convey the communication context in which a message exchange occurs. MPI even offers a process the ability to probe its environment, and to probe for messages, allowing MPI programs to use both synchronous and asynchronous message exchange. In addition to defining message exchange semantics, the MPI programming model provides explicit support for the construction of parallel programming libraries suitable for execution on a cluster

ISSN: 2277-9655
Impact Factor: 1.852

. Libraries written for the MPI standard are portable, and can be reused by higher-level application software.The chief MPI tools for library construction are communication contexts, process groups, virtual topologies, and cached attributes. An MPI construct called a “communicator” encapsulates all thesefunctions in a reusable fashion. An MPI process group is an ordered collection of processes that defines the scope for process names and for collective communication. The MPI standard limits itself to defining message passing semantics and functions, and to defining theprimitives required for reusable libraries. MPI does not provide an infrastructure for program construction and task management. Those responsibilities are left to MPI implementations.

References

[1] Rajkumar Buyya et.al. High Performance Cluster Computing Vol 1, Prentice Hall PTR, 1999.

[2] Michael D. Kane, John A. Springer, “Integrating bioinformatics, distributed data management, and distributed computing for applied training in high performance computing**”,** Proceedings of the 8th ACM SIGITE conference on Information technology education, Destin, Florida, USA, Pages: 33-36 Year of Publication: 2007 ISBN:978-1-59593-920-3

[3] Ian Gorton, Daniel Chavarria-Miranda, Manoj kumar Krishnan, Jarek Nieplocha, **“**A High-Performance Event Service for HPC Applications”, Proceedings of the 3rd International Workshop on Software Engineering for High Performance Computing Applications, Washington, DC, USA, Page: 1, Year of Publication: 2007, ISBN:0-7695-2969-0

[4] Sadaf R. Alam, Jeffrey S. Vetter, Pratul K. Agarwal, Al Geist, **“**Performance characterization of molecular dynamics techniques for biomolecular simulations”, Proceedings of the eleventh ACM SIGPLAN symposium on Principles and practice of parallel programming, New York, USA, Pages: 59 – 68, Year of Publication: 2006, ISBN:1-59593-189-9

[5] Volodymyr Kindratenko, George K. Thiruvathukal, Steven Gottlieb, “High-Performance Computing Applications on Novel Architectures”, Computing in Science & Engineering, Pages: 13 – 15, Volume : 10 , Issue:6, Nov.-Dec. 2008, ISSN : 1521-9615 [5] James C Phillips, John E. Stone.

[6] IEEE Task Force on Cluster Computing (TFCC),http://www.ieeetfcc.org/.G. Pfister, *In Search of Clusters*, Prentice Hall, 1998.

[7] G. Pfister, “An Introduction to the InfiniBandArechitecture”, *High Performance Mass Storage andParallel I/O*, IEEE Press, 2001. www.infinibandta.org

[8] S. Roosta, *Parallel Processing and Parallel Algorithms:Theory and Computation*, Springer Verlag, 2000.

[9] M. Snir, S. Otto, S. Lederman, D. Walker, J. Dongarra,*MPI: The Complete Reference*, MIT Press, 1996.

[10] D. Spector, *Building Linux Clusters*, O’Reilly, 2000.[30] Thomas Sterling, J. Salmon, D. Becker, and D. Savarese, *How to Build a Beowulf*, MIT Press, 1999.

[11] UC Berkeley NOW project, http://now.cs.berkeley.edu/

[12] Virtual Interface Architecture, http://www.viarch.org/.

[13] B. Wilkinson and M. Allen, *Parallel Programming*,Prentice Hall, 1999.