IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 8, August 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908039 |

Study on Cluster Computing

Kaushal Jha , Sumit Devre

U.G. Student, Department of Information Technology, B.K. Birla College of Arts Science and Commerce(Autonomous), Kalyan, Maharashtra, India

U.G. Student, Department of Information Technology, B.K. Birla College of Arts Science and Commerce(Autonomous), Kalyan, Maharashtra, India

ABSTRACT: A computer cluster is a group of Computers,Which are working together closely so that they form a single computer. The components of cluster are commonly connected to each other through fast Local Area Network (LAN) but not always. Each computer connected through that network in the cluster has an Operating System. A Cluster improves performance and/or availability over that provided by a single computer. Which means they are being much more cost-effective than a single computer of comparable speed or availability. The major objective of cluster computing is Multiple computers can work together to do a particular task in a minimum amount of time by working cooperatively and such strategy can be achieved by transferring the extra loads from busy nodes to free nodes. It can be used in small business with little nodes to fastest computers . In this paper authors will explain the concepts and the principles involved in it.

KEYWORDS : Cluster, Computer, Nodes, System

I. INTRODUCTION

Clustering can be described as an amalgam of parallel, high performance, distributed and high availability computer domains. Alternative computing has undergone many changes since the days of highly expensive and proprietary supercomputers. Many people from academia and industry such as system designers, algorithmic developers, network developers and language designers, standardization forums,faculties and graduate students conduct a lot of cluster computing research. Changes and performance enhancements have also been observed in the mainframe domain for many environments. But these computing environments may not be the most cost-effective and responsive solution to a problem. The underlying network is a dedicated network consisting of low-latency high-speed switches that may be a single switch or a hierarchy of multiples. Over the past decade, cluster technologies have been developed to allow multiple low-cost computers to operate in a coordinated way. Enabling heterogeneity in hardware and software (SW) is extremely important and becomes one of the key issues in cluster research. Before putting together a cluster, a number of considerations must be taken (for example, what type of hardware the nodes will run on, which interconnect to utilize , and what switching architecture to employ). Selecting the right cluster elements requires an understanding of the application and resources required, such asstorage, flow, latency and node counts. The economy, performance and flexibility of computing clusters make cluster computing an interesting alternative to Centralized computer modelsthe patterns and problems associated with the resulting costs, inflexibility and scalability.for commercial applications, cluster can be better used in e-commerce as a super server, which consolidates the web server, ftp server, mail server, database server, etc. Each compute node (computer) can have different characteristics such as a single processor or a symmetrical multi-processor design, and access to various types of storing devices. Clusters aren't commodities in and of themselves, despite the fact that they're built on commodity hardware.

II. HISTORY OF CLUSTER

ARCnet, developed by Datapoint in 1977, was the first commercially available clustering product. It was a commercial failure. In 1984, DEC introduced the VAXcluster operating system for the VAX/VMS operating system. Parallel computing, shared file systems, and peripheral devices were all supported by the ARCnet and VAXcluster products. The goal was to combine the benefits of parallel computing with data consistency and uniqueness. On HP OpenVMS systems running on Alpha and Itanium processors, **VAXcluster**, now **VMScluster**, is still accessible.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

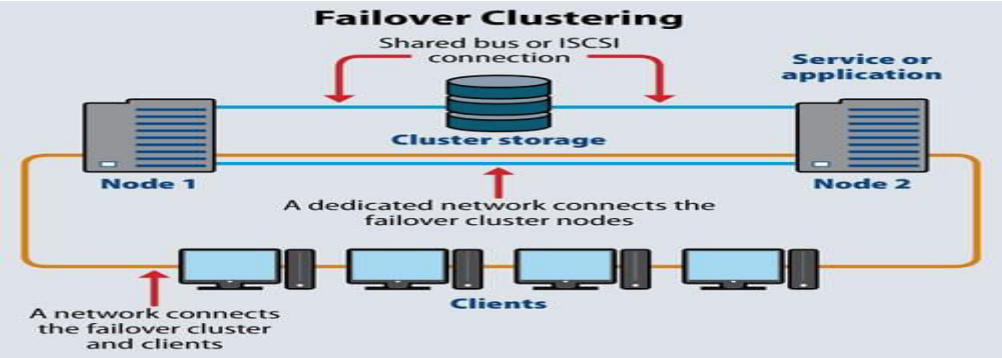
| DOI: 10.15680/IJIRCCE.2021.0908039 |

III. WORKING OF CLUSTERS

1. High Availability or Failover Clusters:

These clusters are designed to provide uninterrupted availability of data or services to the end-user community. The purpose of this cluster to assure that a single instance of an application is only running over on one single cluster at a single time but if and when that cluster member is no longer available, the application will failover to another cluster member. With a high uptime cluster, nodes can be removed. service for maintenance or repairs . In addition, if a node fails, the service can be restored without impacting the availability of the services provided by the group. Although the app will still be available, there will be a performance downturn due to the missing node. Implementation of high availability clusters are best for mission essential applications or databases, mail, files and printers, web or application server.

Unlike distributed or parallel treatment clusters, high availability clusters are transparent and transparent existing stand-alone integration, Conscious non-cluster applications together in one virtual machine needed to enable the network to grow effortlessly to meet business growth demands

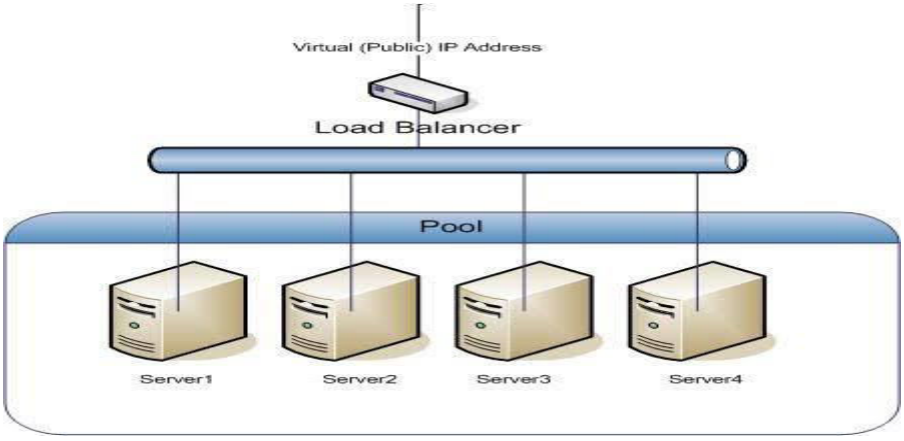


Failover Cluster

2. Load Balancing Cluster :

This type of cluster distributes incoming demands for resources or content between several nodes running the same programs or with the same content. Each node in the cluster is capable of handling queries for the same content or application. In case of failure of a node, the requests are redistributed between the remaining available nodes. This kind of distribution is usually found in a web hosting environment.

High-availability cluster and load-balancing technologies can be combined to enhance reliability, availability and scalability of widely deployed applications and data resources for web, mail, news or FTP services.



Load Balancing Cluster

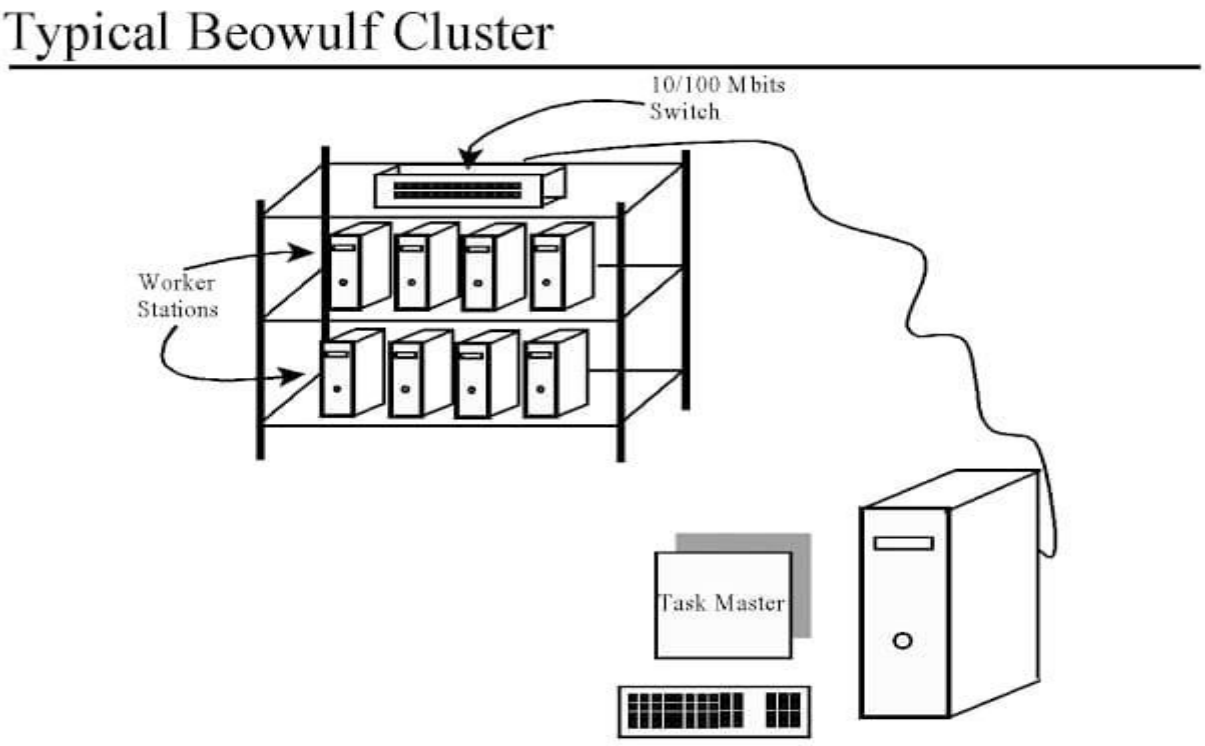
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908039 |

3. High Performance Cluster :

Traditionally, parallel processing was carried out by more than one processor in a specially designed parallel computer. These are systems where several CPUs share a single memory and a simple computer bus interface. Computers may now be networked to form a parallel-processing cluster thanks to advances in high-speed, low-latency switching technologies. Thanks to advancements in high-speed, low-latency switching technologies, computers can now be networked to form a parallel-processing cluster. A parallel cluster is a system that employs multiple nodes to do a single computing or data-mining operation at the same time. Unlike load balancing or high availability clusters, which distribute requests/tasks among nodes and have each node handle the complete request, a parallel environment divides the request into many sub-tasks and distributes them to multiple nodes within the cluster. Parallel clusters are commonly used for CPU-intensive analytical tasks like mathematical computing, scientific analysis, and financial data analysis. Parallel clusters are commonly used for CPU-intensive analytical tasks like mathematical computing, scientific analysis, and financial data analysis.



BEOWOULF CLUSTER

ADVANTAGES AND DISADVANTAGES OF CLUSTER COMPUTING

ADVANTAGES OF CLUSTER :

1. Cluster Computing systems are up to 15 times less expensive than traditional supercomputers when compared to performance.
2. **Reduced Cost**: The cost of off-the-shelf consumer desktops has come down, and this has coincided with an increase in processing power and performance.
3. **Scalability**: It's simple to upgrade and keep up with.
4. Reliability: It is dependable to deal with since the Cluster Computing system continues to function even if sections of it fail.
5. **Processing Capacity**: In most circumstances, a high-performance cluster's parallel processing capability can prove to be more cost effective than a mainframe with similar processing power.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908039 |

6. **Availability**: When a mainframe computer goes down, the entire system goes down. If a node in a computer cluster fails, however, its operations can be transferred to another node in the cluster, ensuring that service is not disrupted.

DISADVANTAGES OF CLUSTER :

1. For cluster computing, it is difficult to manage and organise a big number of machines.
2. It performs poorly when used in non-parallel applications.
3. Physical space requirements are significantly higher when compared to a single server.
4. Power usage rises when compared to a single server.
5. **Skew in Software**: - The most significant disadvantage is software skewness. Skewness refers to the fact that if the software configuration on some nodes differs from that on others, even little deviations can cause a program to fail.

IV. ARCHITECTURE OF CLUSTER COMPUTING

A cluster is a distributed or parallel processing system that consists of a number of interconnected computers that act as a single computing resource.

A computer is a single-processor or multiprocessor system (Workstation, PCs) that includes all of the necessary features such as storage, I/O, and an operating system. A cluster's nodes can be physically separated and connected through a LAN, or they can be gathered together in a single storage.

The following are the components of the architecture:

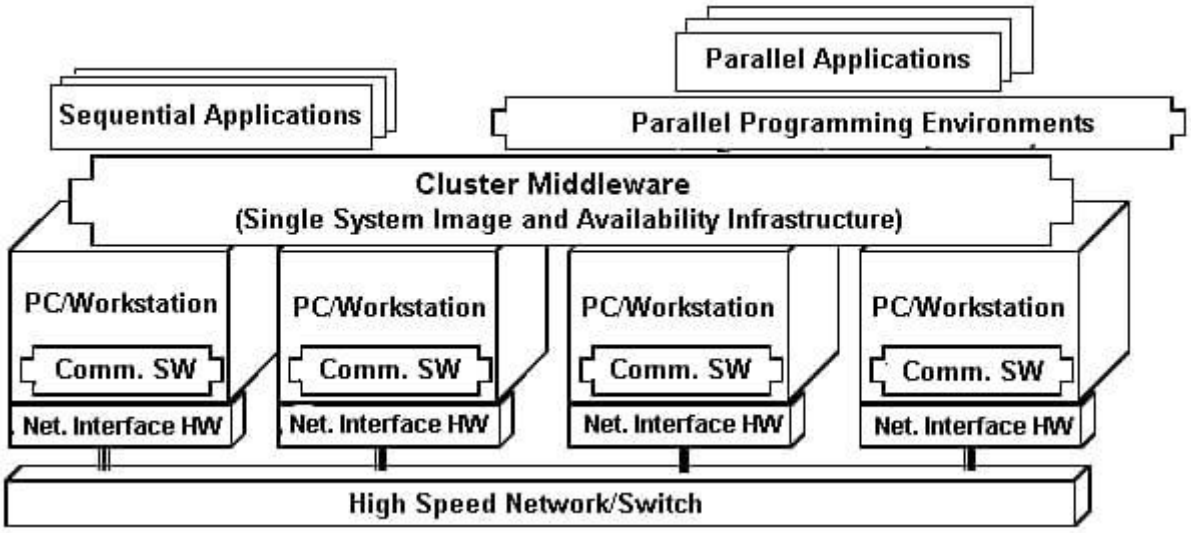
- Multiple High performance computers(PCs and Work Stations)
- Operating System
- High performance Networks
- Network Interface Cards (NIC)
- Communication Protocols and Services
- Cluster Middleware (Single System Image and Availability Infrastructure)

o Hardware

o Operating System Kernel or Gluing Layer

o Applications and Subsystems

- Applications
- Runtime Systems
- Resource Management and Scheduling software
- Parallel Programming Environment and Tools



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908039 |

V. FEATURES OF CLUSTER COMPUTING

1) **Stability** :- The first and most important aspect of cluster computing is its robustness and health in the face of any crashing process. Dynamic reconfiguration can be used to solve any problem.

2) **Performance** :- Cluster computing's performance is mostly determined by memory management and processing. The programmer can change the relevant parameters invisibly to (fine-tune the OS) to meet his own needs.

3) **Extendibility** :- Cluster-specific additions, such as adding user-loadable drivers and raising the intensity of interface documentation, can be simply integrated by the user.

4) **Scalability** :- The qualities of the enclosed nodes, which are driven by the performance characteristics of the interconnectivity, have a big impact on the cluster's scalability. This also includes an OS's resource constraints, such as the maximum size of available address space.

5) **Support** :- Many clever and technically better computer systems have failed in the past owing to a lack of support, such as hardware drivers and middleware. This assistance is readily available in cluster computing. The only disadvantage is the high hardware expenses (because usually dozens of nodes are to be installed). However, it significantly improves cluster applications.

6) **Heterogeneity** :- Clusters offer a dynamic and changing environment since they may be updated with standard hardware as needed or afforded by the user. As a result, a cluster environment does not have to be made up of identical hardware that runs the same operating system.

7) **Manageability** :- Another feature of cluster computing is that it is possible to administer a single system unit even after merging a number of off-the-shelf commodity computers.

VI. RESULTS AND DISCUSSION

1. Issue and Discussion

Cluster networking: There will be significant disparities in the speed with which data is accessed and how individual nodes connect if you mix hardware with different networking technologies. Ensure that all of the computers you intend to include in your cluster have similar networking if it is within your budget.If at all possible, network adapters from the same manufacturer should be used.

Cluster Software : You will have to build versions of clustering software for each kind of system you include in your cluster.

Timing :- Cluster's most bothersome element is this. Because these machines all have various performance profiles, our code will run at varying speeds on each type of node. If a process on one node is waiting for the results of a calculation on a slower node, this can generate major bottlenecks.

2.Behaviour of programming system:

Programming:For data types supported by the least powerful node in our cluster, our code will have to be developed to support the lowest common denominator. The more powerful machines in a mixed machine will have attributes that the powerful machine does not have.

Network selection :- Buses, cubes of varying degrees, and grids/meshes are some of the diverse types of network topologies. One or more network interface cards, or NICs, will be installed in the head-node and computing nodes of our cluster to implement these network topologies.

Speed collection :- Whatever topology you choose for your cluster, you'll want to acquire the fastest network possible within your budget. Fortunately, the widespread availability of high-performance computers has compelled the creation of high-performance networking systems

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908039 |

VII. CONCLUSION

Cluster computing has made its way into enterprise data centres employing clusters of varying sizes due to the economics of cluster computing as well as the flexibility and high performance it provides. As clusters become increasingly popular and prevalent, careful consideration of application requirements and how they translate to network characteristics is crucial to the design and implementation of an optimal and reliable performance solution.

High-performance cluster computing is allowing a new class of computationally intensive applications to solve challenges that were previously too expensive for many businesses. Commodity computers working together to accomplish very complicated, computationally expensive tasks has a wide range of applications in chemistry and biology, quantum physics, petroleum exploration, crash test simulation, computer graphics rendering, and financial risk analysis. Cluster computing, on the other hand, tests the boundaries of server architecture, processing, and network performance.

Cluster computing technologies are continuously changing to satisfy the demands of current, new, and emerging applications, including host protocol stack processing and interconnect technologies. Low-latency switches, protocols, and standards that efficiently and effectively utilise network hardware components have made significant progress.

REFERENCES

[1.] http://academic.csuohio.edu/yuc/hpc00/lect/chapter-B1.pdf
[2.] http://www.buyya.com/cluster/
[3.] http://www.gridrepublic.org/?gclid=CLPnt4-ThbsCFWIC4godynsArg
[4.] http://www.springer.com/computer/communication+networks/journal
[5.] http://www.clustercomp.org/
[6.] http://compnetworking.about.com/od/networkdesign/l/aa041600a.html
[7.] http://www.cloudbus.org/paper
[8.] http://research.amnh.org/scicomp/amnh-computer-clusteric_cluster.pdf
[9.] http://www.buyya.com/csc433/ClusterApps.pdf
[10.] http://www.davidchappell.com/writing/white_papers/Cluster_Computing
[11.] http://www.codeproject.com/Articles/11709/Cluster-Computing

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details