Cluster Computing

Ms. Arshpreet Kaur [#1]
Lecturer in Computer Science Department
[#1]Mata Sahib Kaur Girls College, Gehal (Barnala)
Affiliated to Punjabi University, Patiala
[#1]arshpreetdhillon@yahoo.co.in

Ms. Manpreet Kaur Mann [*2]
Lecturer in Computer Science Department
[*2]Mata Sahib Kaur Girls College, Gehal (Barnala)
Affiliated to Punjabi University, Patiala
[*2]maan_muskan_20@yahoo.co.in

***Abstract*- A cluster has been defined as a group of machines that may be viewed as a single entity for the purposes of control and job assignment. A cluster is a type of parallel or distributed processing system, which consists of a collection of interconnected stand-alone computers co - operatively working together as a single, integrated computing resource. This cluster of computers shares common network characteristics like the same namespace and it is available to other computers on the network as a single resource. These computers are linked together using high-speed network interfaces between themselves and the actual binding together of the all the individual computers in the cluster is performed by the operating system and the software used.**
The clustering applications also assist administrators to remotely control and monitor the performance of each node and the general environment the clusters that might be a factor for its performance. There are a good number of clustering applications in the market today. Most of the applications are geared towards specific type of clustering while there are those that are versatile enough to be adapted to different types of clustering. Choosing the right clustering application really depends on the business need.

***Keywords*-Configuration, high-traffic Web sites, load balance traffic, parallel processing, usability, versatile, Web farm, Web page.**

I. INTRODUCTION

Cluster computing is a form of computing in which a group of computers are linked together so that they can act like a single entity. There are a number of reasons for people to use cluster computers for computing tasks, ranging from an inability to afford a single computer with the computing capability of a cluster to a desire to ensure that a computing system is always available. In gist, clustering is a technique of connecting different computers together so that a more powerful computing device could be achieved. Individually, a computer may not be able to fully harness the application or provide an efficient service but when computers are clustered, they become a more powerful computing device. [2]
A highly efficient cluster will usually have nodes that are physically together. The nodes should be placed in one area so that the highest output could be achieved. However, there are times that there are never enough nodes in one area. For that reason, virtual clusters are created. Different nodes are connected even though they are geographically separated. Virtual clusters have been proven to provide efficiency especially when the clusters are only considered for storage purposes.

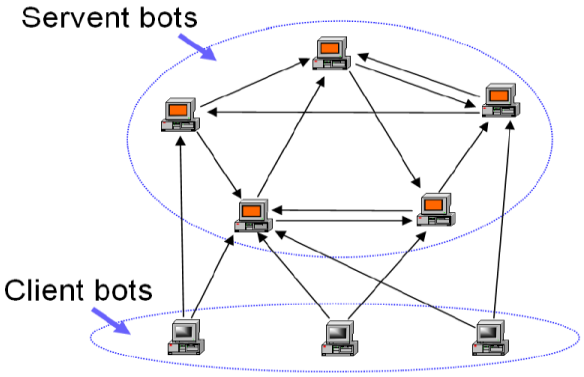

Figure 1: clustering

II. TYPES OF CLUSTERING

There are four forms of clustering wherein each form caters to a specific need.

A. Load Balancing

This type of clustering connects the computers (usually referred to as nodes) wherein each computer provide unified workload. Each node will have a specific workload so that each computer could work efficiently as it will not work on additional function. The advantage of load balancing is the fact that it is always optimized to work while its disadvantage is the higher need of maintenance.

1) Features of Load balancing clustering:

- Transfer of Workload
- Prevention of DoS Attack

B. High Availability

High availability focuses on the nodes ability to be fully available under any circumstances. The functions could be implemented by each node wherein each node becomes a back-up to another node in case something wrong happens. High availability is also referred to as failover clusters or clusters with high redundancy level. The advantage is, of course, on its promise of stability.[6]

1) Features of High Availability clustering:

- Ease of Start/Stop Operation
- Availability of Storage Area Network
- Ability to Save State.

C. Grid Computing

This type of cluster refers to the behavior of nodes. In grid computing, nodes rarely communicate with each other as they only provide an output. The nodes in grid computing often have little to no relationship with each other as each node provide specific function for the completion of a specific problem or computation.[7]

1) Features of Grid Computing Clustering:

- Versatile
- Easy Maintenance

D. Clusters for Computations

Compute clusters requires an active relationship of each nodes so that the right computation could be achieved. Although they could be used for business purposes, these types of computer clusters are usually used for scientific and mathematical purposes as these fields require heavy calculations.

1) Features of Compute Clustering:

- Computation
- Good Communication

III. CLUSTERING APPLICATIONS

The clustering applications also assist administrators to remotely control and monitor the performance of each node and the general environment the clusters that might be a factor for its performance. There are a good number of clustering applications in the market today. Most of the applications are geared towards specific type of clustering while there are those that are versatile enough to be adapted to different types of clustering.

A. MPI as the Standard Clustering Application

Its simplicity and adaptability to different programming language has made MPI a highly favored clustering application. The proper implementation of MPI will enable simplified communication that could be used in simple two-node cluster setting or could be implemented in complicated load balancing to ensure that workload will not be focused on a single node only.

B. Linux Clustering Applications

The open source community has also developed specific applications for clustering. The most popular applications that are freely distributed today could be divided into four groups.

1) The first group refers to service based clustering application.

2) The second group is the applications that will not only distribute requests but will move towards data and functions distribution which could easily react when the clustering environment changes.

3) The third group is the applications that are simple single-system image wherein the clusters are grouped into one for ease of control.

4) The fourth group is the applications that are geared towards application workload control.

Some of the recognized clustering applications are the following:

• *Linux-HA*

• *Beowulf*

• *MOSIX*

• *Kerrighed*

• *OpenSSI*

• *Linux Virtual Server*

C. Windows Compute Cluster Server

The software giant Microsoft Corporation has also developed their own clustering application. This type of clustering application is closely linked to another product from the same company simply called Windows Server.. This cluster server comes with schedulers and other tools for ease of control particularly some functions based on MPI.[1]

D. Choosing the Right Application

Choosing the right clustering application really depends on the business need. However, administrators have to make sure that the clustering application should not be too much for the nodes to handle or would have functions that could be worthless.

IV. CLUSTERS-MONITERING AND MEASUREMENT

Clustering will also require resources and the right manpower as these requires consistent check and monitoring. Information such as temperature, processing speed and available node should be provided to the administrator in real time. Measurement data such as expected performance on various forms of connections, file compatibility and latency should also be available for administrators. Monitoring and measurement tools could be installed by administrators in nodes or file servers so that it could actively monitor the cluster's performance.

A. Avoiding Bloatware

One of the many challenges for administrators in selecting the right monitoring and measurement tools for the clusters is in avoiding bloatware. A bloatware is a valid application for a specific purpose but would have too many functions that are not necessary or useful at all for users. Bloatware could be found in different applications for different industries including tools for clustering. Most applications for clustering come with different functions that are geared for specific clustering environment.

B. Combined vs. Individual Tools

Administrators should always take caution when choosing an application that provides both measurement and monitoring functionality. Administrators will be able to dig deeper on the functionality of their measurement and monitoring tools. As separate tools, administrator will be able to avoid bloatware since they can enable only one application to run whereas combined applications will naturally load two functions at the same time.

C. Reporting Options

The monitoring and measurement tools of today could provide real time feedback to administrators not only in the local area but online as well. When errors are detected, there are monitoring tools that can send out messages to administrators not only via email but even through text messaging. Mobile monitoring of the performance of nodes and clusters is also possible. [9]

V. VIRTUAL MACHINE CLUSTERING

Virtualization could even be implemented in clustering. Instead of physical computers or nodes, virtual nodes are created wherein each node has the same capability except that they are never physical. With the right configuration, clustering could be properly introduced to virtual machines to achieve stability and improvement of functionality.

A. Advantages of Clustering of Virtual Nodes

The advantage of clustering in virtual nodes is often based on its difference between physical nodes. When clustering is administered in virtual nodes, no additional hardware is required if the network administrator just wanted to have one server with multiple nodes inside the server.

Another advantage of virtual nodes is its ability to be totally transparent to the network administrator. Although physical nodes could provide transparency, control and monitoring is a lot easier through virtualization. Data could be easily accessed; control on different nodes is possible and troubleshooting is even easier. The ease of troubleshooting is also another reason why clustering in virtual machines is easy.

B. Disadvantage of Clustering for Virtual Machine

The problem in clustering virtual machines is not exactly on its software and configurations but on the fact that it lacks additional hardware. When the virtual nodes are launched in a single machine, consistency might not be achieved. Since there is no physical back-up, the process could immediately stop and jeopardize the whole operation.[11]

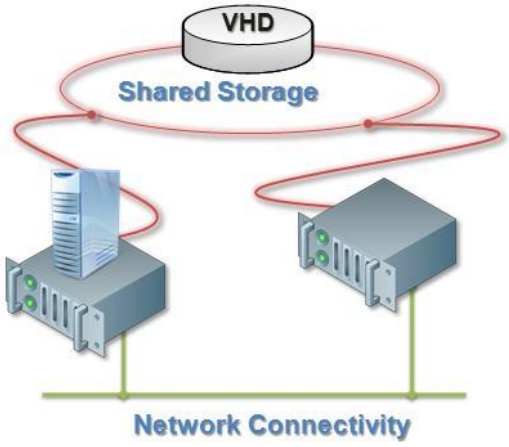

Figure 2: Virtual Machine Clustering

VI. GPU (GRAPHICS PROCESSING UNIT) CLUSTERS

Additional powerful nodes, better infrastructure and connections and updated application to control the nodes are only some of the requirements in improving the computing performance of the clusters. GPU is used primarily for improving videos; they can be harnessed to assist computing in the nodes.

A. Forms of GPU Clusters

The first form of GPU powered cluster is the heterogeneous type of GPU clustering. They could come from different manufacturer and might even have different processing power. The advantage of this form of GPU clustering is the cost efficiency since a powerful GPU could be assisted by less powerful GPU.

The next form of GPU powered cluster is the homogeneous form. In this GPUs installed in the node are uniform in processing power, make and model. Although a little bit expensive for powerful GPUs, they provide the extra processing power every developer expects. Careful consideration should be done when working with GPUs.

B. Connection Requirements

For the GPU to properly work as performance boosters when installed, additional connection is required. Connection will not require additional software although control of nodes with GPU will have a special application which will be outlined later. An additional Ethernet switch or a communications link between GPUs is required. To properly implement GPU powered clustering, the software should be present that is the Right Operating System, drivers, API, Mapping. [3]

VII. CLUSTERING SECURITY

Clustering security is one of the most important factors that need to be considered during clustering. The attacks to clustering of nodes could come in different forms – it could be as simple as a virus wherein its sole purpose is to destroy files or could be a very powerful spyware that can easily hijack the controls of nodes for malicious purposes. It only takes a single security flaw to destroy the entire clustering configuration. Whenever a network opens up a connection to its administrator, it automatically opens itself to different forms of attacks. This is also possible for users who try to access the nodes and stores data.

A. Network Tool Based

Software giants such as Microsoft Corporation and Sun Microsystems and Open Source communities have network tools for proper implementation of different types of clustering. When a new network tool is implemented for clustering, security might not be immediately implemented. But whatever network tool is used, it is important to have the latest version. Different types of attacks have been experienced the past year alone that an update is always necessary to achieve optimal security.

B. Domain Based vs. Local Security

For domain based security for clusters, administrator could easily control the clusters since security is based online. However, domain based security could be easily hacked.

Local security on the other hand boasts of optimal security by localizing administrator credentials for access. It requires a lot of resources especially when they are configured for the first time. [4]

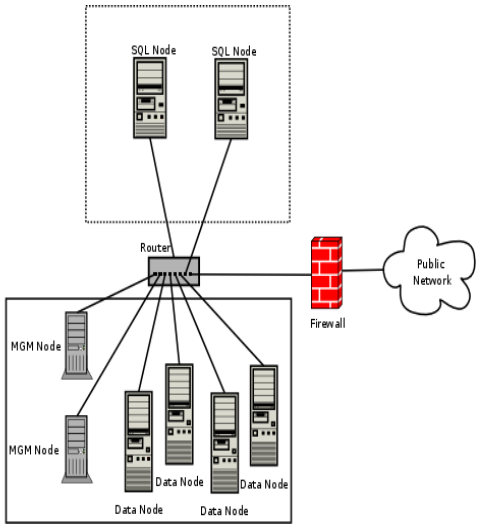

Figure 3: clustering security

VIII. CLUSTERING ON P2P

P2P or Person-to-Person is a type of network wherein the connection is based on different users. The success of P2P is practically based on the number of active users. However, sharing workload through P2P is usually affected by users' bandwidth. The users who are part of the P2P network group are usually linked online. P2P is often an option for those who cannot meet physically to boost their application's performance or faster file sharing. Through online connectivity, multiple users would be able to boost the performance of their application while helping others improve their application. Person to person networking is one of the most cost efficient practices for file sharing.

A. Criticisms of P2P

Unfortunately, P2P has been under a lot of criticisms for many years. Today's main use of P2P is through file sharing which has prompted the use of this type of networking for disputed purposes. Music, movies and applications could be easily shared in P2P network even if the user will not pay for downloaded files. Anyone in the network could upload corrupt files and anyone who tries to download the file could be affected.

B. Hybrid P2P for Proper Clustering

In pure P2P setting, one of the users will have to be the cluster manager which could become the single point of failure. Hybrid P2P has become the standard for P2P especially on file sharing. Through a server wherein data is controlled, additional features such as better access to information including the number of active users could be available for users.[8]

IX. TWO-NODE CLUSTERS

As technology evolve, setting up clusters for small businesses has got a lot more affordable. Businesses today could consider implementing two-node clusters to secure their website or online application. This type of clustering would ensure continuous business operations without having to spend too much on hardware. As the name suggests, the two-node clusters consists of only two-nodes in HA (High Availability) clustering configuration. In gist, the first node will be the primary node while the second node would be the back-up. In case the first node experiences trouble in maintaining functions or data requests, the back-up node will take over.

A. Hardware Requirements

The hardware components to fully implement two-node clusters are minimal but each takes a very important role to ensure efficiency of the cluster. As the name suggest, it will only have two nodes with optimized processing speeds. Its only difference with larger nodes is its ability to function as a storage device which is added to assist the shared storage. The two nodes connect to the shared storage which functions as a server. The components would also need to have interconnection hardware to ensure that each would communicate for stability. Administrators could choose between serial and Ethernet crossover cabling. An optional hardware could be installed with interconnection protocol. A shared disk could be part of interconnection wherein it will serve as an active monitoring tool so that it could alert the administrator if one of the nodes is not working. [5]

X. CLUSTERING THROUGH SERVER FARMS

A single server could even be considered as a single point of failure since an attack to the server could debilitate the entire operation. To prevent downtime from happening, businesses implement server farms. Server farms, as the name suggest, is composed of servers connected together to serve the business interests of providing the right data or faster streaming of information. The processing power of each server is harnessed so that it could push data faster.

A. Challenges to Server Farms

Although server farms ensure faster data processing and prevention of single point of failure, this configuration for businesses is very costly. Another challenge to server farms is the external circumstances. Weather and other natural circumstances could easily destroy the server farm which could easily bring down the operation. This is another challenge in implementing server farms. Different server farms set up in various locations has become a requirement to avoid downtime. This is another financial challenge for those who are planning to set up server farms.

B. Routers and Network Switches

Routers help the users connect to the servers while the network switches controls the server farms to ensure availability of data to its users. These routers often have the capability of handling heavy workload and have advanced controls for network administrators. This also holds true in network switches wherein server farms are monitored and controlled by the network administrators. A back-up server farm could be activated through network switching. [8]

XI. CLUSTER SUPERCOMPUTER

The addition of nodes to build a cluster to support file sharing, functions, applications and websites all point to one thing – building a supercomputer. The collection of nodes that would eventually form a cluster would be part

of something bigger which could easily support almost any type of workload without showing any signs of downtime. The idea of preventing single point of failure, adding monitoring and measurement tools and even considering the external factors that might jeopardize infrastructure are also small parts towards the ultimate goal of building a supercomputer that will not post any problems while computing.

A. Hardware Considerations

The first thing administrators or developers would consider in building a supercomputer for varied interests is to have the right hardware.

1) Processors – building a supercomputer is just including a dual-core processor so that it could provide better processing speed.

2) Inclusion of GPU –. With the addition of right connectors, GPU (Graphics Processing Unit) could easily improve the performance of the cluster.

3) Memory - Processors and GPU would still need the help of an efficient memory (RAM) if they wanted to use the cluster for application processing.

B. Applications for Supercomputer

Even though you have the best hardware for a supercomputer through clustering, you will still need to have the right application to ensure the clusters are working together.

1) OS (Operating System) – The operating system suggests the environment of the cluster and their configuration will suggest how the nodes will interact with each other. Because of the needed flexibility and stability, the most popular OS today for clustering is powered by Linux and UNIX.

2) Applications –MPI (Message Passing Interface) could be used by programming developers to build their clustering application so that it would fit based on their need. Microsoft, Open Source and Linux also have separate offerings but they still include some libraries based on MPI.

3) Tools –Measurement and monitoring tools are very important to be integrated in clusters and nodes since they provide the much needed timely or even real time feedback to administrators in case one of the nodes fails to perform per expectation.[10]

XII. CONCLUSION

The recent advances in high-speed networks and improved microprocessor performance are making clusters or networks of workstations an appealing vehicle for cost effective parallel computing. Clusters built using commodity hardware and software components are playing a major role in redefining the concept of supercomputing .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. This cluster of computers shares common network characteristics like the same namespace and it is available to other computers on the network as a single resource. These computers are linked together using high-speed network interfaces between themselves and the actual binding together of the all the individual computers in the cluster is performed by the operating system and the software used.

REFERENCES

[1] Bader, David; Robert Pennington (June 1996). "Cluster Computing: Applications". Georgia Tech College of Computing. http://www.cc.gatech.edu/~bader/papers/ijhpca.html. Retrieved 2007-07-13.

[2] Pfister, Gregory (1998). *In Search of Clusters* (2nd ed.). Upper Saddle River, NJ: Prentice Hall PTR. p. 36. ISBN 0-13-899709-8.

[3] GPU Clusters for High-Performance Computing

[4] GPU cluster at STFC Daresbury Laboratory

[5] Evan Marcus, Hal Stern: *Blueprints for High Availability: Designing Resilient Distributed Systems*, John Wiley & Sons, ISBN 0-471-35601-8

[6] Chee-Wei Ang, Chen-Khong Tham: *Analysis and optimization of service availability in a HA cluster with load-dependent machine availability*, IEEE Transactions on Parallel and Distributed Systems, Volume 18, Issue 9 (September 2007), Pages 1307-1319, ISSN:1045-9219.

[7] P Plaszczak, R Wellner, *Grid computing*, 2005, Elsevier/Morgan Kaufmann, San Francisco

[8] Ranjan, Rajiv; Harwood, Aaron; Buyya, Rajkumar (1 December 2006), *A Study on Peer-to-Peer Based Discovery of Grid Resource Information*

[9] Rüdiger Schollmeier, *A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications*, Proceedings of the First International Conference on Peer-to-Peer Computing, IEEE (2002).

[10] "The Grid Café – The place for everybody to learn about grid computing". CERN. http://www.gridcafe.org. Retrieved December 3, 2008.

[11] Maurer, Ryan: Xen Virtualization and Linux Clustering