

Study of Cluster, Grid and Cloud Computing

Vinayak Shinde¹, Amreen Shaikh², Chris Donald D'Souza³

Assistant Professor, Computer Engineering, Shree L. R. Tiwari College of Engineering, Thane, India ¹

ME Scholar, Computer Engineering, Shree L. R. Tiwari College of Engineering, Thane, India ^{2,3}

Abstract: Cloud computing is rapidly growing and is gaining popularity across the world because of its numerous features. Cloud computing has evolved from simple networking applications. Its functioning is however based on models of cluster and grid computing. Cluster, Cloud and Grid Computing are practices of distributed and local computing. Cluster consists of computers connected over a LAN, whereas cloud and grid are geographically distributed on a wider scale. This paper strives to compare, contrast and analyze cloud computing with grid computing and cluster computing from various aspects and in conclusion give an understanding of their crucial characteristics.

Keywords: cloud classification, cloud computing, cluster computing, fabric grid, grid computing, high availability cluster.

I. INTRODUCTION

Computing on a whole has undergone innumerable changes over the years. During the earlier stages of technological development high-performance computing was restricted to institutions which could afford the cost of dedicated supercomputers. This created a need for high-performance computing in small scale architecture at a lower cost thus leading to cluster computing. The concept of grid computing which allows the users to have demand based computing followed. The concept of cloud computing followed later, concentrates on the provisioning/de-provisioning of computation, storage and data services to and from the user without him knowing where he is getting these resources from. With widespread use of internet all over the globe, almost anything can be delivered over the internet via cloud computing.

II. CLUSTER COMPUTING

In Cluster computing several nodes are run as a single entity. The different nodes in cluster are usually connected by faster local area networks, thus enabling us to achieve high performance and fault tolerance. High performance is achieved through parallel programming whereas fault tolerance is achieved by connecting multiple nodes to each other. Cluster computing provides faster processing speed, larger storage capacity, better data integrity, superior reliability along with a wider availability of resources.

Cluster computing was invented when users could not fit all their work into a single computer or when they needed a backup of their data. The first commercially available clustering product was ARCnet developed in 1977 by Datapoint DEC later released VAX cluster in 1980's.

A. Cluster Architecture

*Figure 1 shows cluster architecture. The architecture of cluster computing contains the following main components: multiple standalone computers, operating system, high performance interconnects, communication software and different application platforms

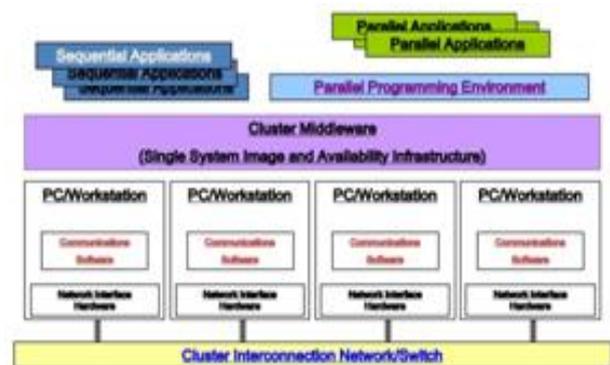


Fig 1: Cluster Computing Architecture

B. Cluster Categorization

- 1) **High Availability Server:** It is placed in a cluster to achieve high availability. If any hardware or software fails, a counter action is taken by the cluster itself and resources are then fetched from other servers to continue with the operation. Resources are constantly monitored in the cluster. A practical failover of software or hardware including the current state of application can be achieved by fault tolerant fail over cluster.
- 2) **Load Balancing:** This architecture distributes the load among various servers in a cluster. The servers are set up in a manner that the traffic is proportionately distributed among all the servers. Due to such architecture the entire network load gets re-distributed without focusing on a single server thus offering better reliability and an improved performance.
- 3) **High Performance:** It uses cluster nodes to perform concurrent calculations. A high-performance cluster allows applications to work in parallel, therefore enhancing the performance of the applications.

C. Advantages of Cluster Computing

- 1) **Manageability:** Cluster combines a large numbers of components to work as a single entity thus reducing the effort, cost and money to manage a large number of components.

2) **Single System Image:** With cluster the user gets the feel that he is working with a single system thus making him think that he only needs to manage a single system image.

3) **High Availability:** Since all the components are replicas of each other, if one component goes down because of any technical reason, then some other component can take over its place, without affecting the user's work.

D. Disadvantages of Cluster Computing:

1) **Programming Complexities:** If the cluster components are different in terms of software from each other, then there might be issues when combining them together to work as a single entity.

2) **Troubleshooting Woes:** Because we are dealing with a single entity image, it is difficult to troubleshoot the problem with an individual element of the cluster

3) **Complex to maintain:** Since cluster computing involves merging of different or same components together with different functionality, a non-professional person may find it difficult to maintain a cluster.

E. Challenges in the Cluster Computing

1) **Middleware:** The middleware has to produce software environments that generate a false impression of a single system image, rather than a cluster of independent computers.

2) **Program:** The applications that run on the clusters must be explicitly written to incorporate the division of tasks between nodes and also take care of the communication between these nodes

3) **Elasticity:** The variance in real-time response when the number of service requests changes dramatically needs to be curtailed.

4) **Scalability:** To meet the additional requirements of a resource without affecting the performance of the system.

III. GRID COMPUTING

In grid computing large-scale science and engineering is done through the interaction of people, various computing resources, information systems, and instruments etc. all of which are geographically and organizationally dispersed. The overall inspiration for grids is to facilitate the routine interactions of these resources in order to support large-scale science and engineering.

Grid computing is a form of distributed computing wherein a super and virtual computer is composed of a cluster of networks, loosely coupled computers. Grid Computing enables the sharing, selection, and conjugation of geographically distributed autonomous resources dynamically at runtime depending on various factors like their availability, capability, performance, cost, and users' quality-of-service requirements [9]. Grids consist of heterogeneous resources thus facilitating flexible, secure, coordinated large scale resource sharing.

A. Grid Computing Concept

The Fig. 2 shows the concept of grid computing in general where various resources are geographically distributed

across the globe or the resources are at dispersed locations connected to a central location i.e. the grid system.

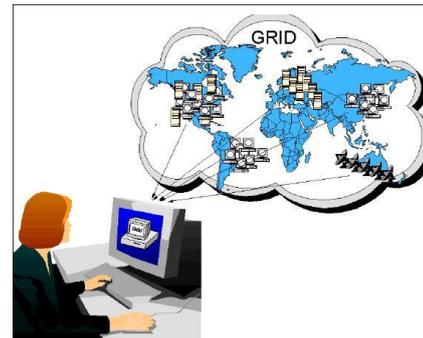


Fig 2: Grid Computing Concept

B. Layered Architecture of Grid

1) **Fabric layer** provides the resources in which shared access is mediated by grid computing.

2) **Connectivity layer** enables the core communication and authentication protocols required for grid specific network functions [4].

3) **Resource layer** defines the protocols, APIs and SDK for secure negotiations, imitations, accounting, monitoring control and payment of sharing operations on individual resources.

4) **Collective layer** it contains protocols and services that are used to capture interactions among a collection of resources

5) **Application layer** contains user applications that operate within VO environment.

C. Advantages of Grid Computing

1) The need to buy large capacity servers is reduced in cases of applications that can be split up and farmed out to smaller capacity servers

2) More efficient in terms of resources utilization.

3) Grid environments are much more modular thus reducing the points of failure.

4) Policies in the grid can be easily managed by a grid software

5) System upgrades can be done without scheduling downtime, and jobs can be executed in parallel thus increasing the performance and hence the throughput

D. Disadvantages of Grid Computing

1) Need for interoperability arises when dissimilar groups want to share resources.

2) Diversity of components, policies, mechanisms. E.g., standard notions of identity, means of communication, resource descriptions etc. might differ

3) Need for sharing infrastructure services to avoid repeated development, installation. E.g., Having one service/port/protocol for remote access to computing instead of one per tool/application

E. Challenges in grid computing

1) **Dynamicity:** Resources in a grid can be owned and managed by more than one organization which can enter/leave the grid at any time thus causing burden on the other resources in the grid.

- 2) Administration: To form a unified resource pool, heavy burden is raised on system administration along with other maintenance work like coordinating local administration policies with global ones.
- 3) Development: Problems like de-composing, distributing to processing elements, and assembling solutions etc. are concerned with methods of writing software to run on grid-computing platforms
- 4) Accounting: Difficulty in finding ways to support different accounting, infrastructure, economic model and application models to cope well with tasks that need frequent communication and are interdependent.

IV. CLOUD COMPUTING

Cloud computing model enables an omnipresent and convenient on-demand network based access to a shared pool of computing resources (e.g., servers, storage or services etc.) that can be rapidly provisioned/released with minimal effort. Cloud computing thus focuses on maximizing the effectiveness of shared resources. Along with being shared by multiple users, cloud resources are also dynamically reallocated as per demand. Cloud computing model offers computing, storage, and software as a utility or as a service. However an internet connection is required to use this model.

A. Cloud Computing General concept

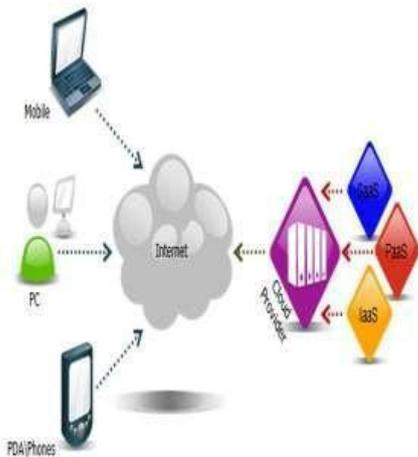


Fig 3: Cloud Computing Concept

The Fig. 3 shows that how different users can connect to the cloud services provided by various service providers by using any device over the internet. Cloud infrastructure includes scalable resources in terms of storage, network, and compute. It also contain virtualized infrastructure

B. Cloud Classification

1) IaaS (Infrastructure as a service) model: The main concept behind this model is virtualization wherein a user has a virtual desktop and consumes various resources like network, storage, virtualized servers, routers etc. supplied by the cloud service provider. Usage fees may be calculated as per hourly usage of CPU, data stored (GB), network bandwidth consumed, network infrastructure used, value added services used e.g., monitoring, auto-scaling etc. Examples: Storage services provided by

Amazon EBS, Amazon S3 etc. Computation services provided by Layered Tech, AmazonEC2 etc.

2) PaaS (Platform as a service) model: It refers to an environment that provides a runtime environment, software deployment framework and component to enable the direct deployment of application level assets or web applications. PaaS is thus a virtual platform where software can be developed, tested and deployed. Hence the entire life cycle of software can be virtually operated on a PaaS. This service model is dedicated to application developers, testers, deployers and administrators. Examples: Microsoft Azure, Google App Engine (GAE), IBM Smart Cloud, Amazon EC2, salesforce.com, jelastic.com etc.

3) SaaS (Software as a service): Under SaaS, delivery model end users can consume software application services directly over the network as per on-demand basis. For instance Gmail can be considered a SaaS where Google is the provider and we the consumers. Examples: Billing services by Arial system, op source. Financial services by Concur, work day. Backup and recovery services by various providers.

C. Advantages of cloud computing

- 1) Cost efficiency: Convenient and scalable charging models have emerged making the cloud a cost effective solution
- 2) Convenience and continuous availability[8]: Public clouds services are available 24x7 irrespective of location
- 3) Backup and Recovery [8]: Since the data resides on the cloud and not on a physical device backup/recovery are more easier and convenient
- 4) Scalability and Performance [8]: Cloud instances are deployed automatically only when needed. Hence you pay only for the applications and data storage you need
- 5) Increased Storage Capacity [8]: The cloud can store much more data compared to a personal computer thus eliminating the worries about running out of storage space

D. Disadvantages of cloud computing

- 1) Internet: A high speed internet connection is required; unavailability of internet would cause unavailability of data.
- 2) Non-Interoperability: Data stored in one cloud cannot be moved to another cloud service provider.

E. Challenges in the cloud computing

- 1) Dynamic scalability: The scheduling delays involved in dynamic scaling of computer nodes are of real concern leading to the need of an effective and dynamic load management system.
- 2) Multi-tenancy: If the number of applications running on a computer node increases, it will reduce the bandwidth allocated to each application thus leading to performance degradation.
- 3) Standardization: Every organization has their own APIs and protocols thus making the integration and interoperability of all the services a challenge
- 4) Debugging and profiling: Parallel and remote debugging has always been a problem

- 5) Security and Privacy: The user never knows where the data gets stored.
- 6) Power: Vast amount of power is consumed in order to meet the needs of various users.

V.CONCLUSION

Through this paper, we have tried to throw light on the three models of present day computing: cluster, grid and cloud computing. Cloud computing is a new technology of computer network, providing the web services at lower cost comparing to normal technique. In this paper we highlight the advantages, disadvantages and challenges in cluster computing, grid computing, and cloud computing.

REFERENCES

- [1] Kaur, Kiranjot, and Anjandeeep Kaur Rai. "A comparative analysis: Grid, cluster and cloud computing." *International Journal of Advanced Research in Computer and Communication Engineering* 3.3 (2014): 5730-5734.
- [2] Omer, Samah Mawia Ibrahim, Amin Babiker A. Mustafa, and Fatema Abdallah Elmahdi Alghali. "Comparative study between Cluster, Grid, Utility, Cloud and Autonomic computing."
- [3] Sadashiv, Naidila, and SM Dilip Kumar. "Cluster, grid and cloud computing: A detailed comparison." *Computer Science & Education (ICCSE)*, 2011 6th International Conference on. IEEE, 2011.
- [4] <http://www.slideshare.net/ajayspi/grid-computing-6366921>
- [5] <https://ppandey07.wordpress.com/2013/10/07/cloud-computing-advantages-disadvantages/>
- [6] http://edutechwiki.unige.ch/en/Grid_computing
- [7] http://edutechwiki.unige.ch/en/Grid_computing
- [8] <http://www.cloudconceptsinc.com/cloud-benefits>
- [9] <http://www.slideshare.net/dibyadip/grid-computing-31446107>
- [10] "ClusterComputing", http://en.wikipedia.org/wiki/Cluster_computing.
- [11] D. Thain, T. Tannenbaum, and M. Livny, "Distributed computing in practice: The Condor Experience", *Concurrency and Computation: Practice and Experience*, 17, pp. 323-356, 2005.
- [12] N. Karonis, B. Toonen, and I. Foster, "MPICH-G2: A Grid-Enabled Implementation of the Message Passing Interface", *Jr. of Parallel and Distributed Computing*, 63, pp.551-563, 2003.
- [13] "Gridbus toolkit", <http://www.gridbus.org>.
- [14] Indu Gandotra, Pawanesh Abrol, Pooja Gupta, Rohit Uppal and Sandeep Singh - Cloud Computing Over Cluster, Grid Computing: a Comparative Analysis - *Journal of Grid and Distributed Computing* - Volume 1, Issue 1, 2011
- [15] <http://utilitygridcomputing.blogspot.com/2008/10/advantages-disadvantages-of-uc.html>
- [16] Radu Prodan and Thomas Fahringer, "ZENTURIO: An Experiment Management System for Cluster and Grid Computing", *Proc. of the IEEE Intl. Conf. on Cluster Computing (CLUSTER'02)*.
- [17] An Introduction and Overview of Cloud Computing - Presented by: Prof Mark Baker - ACET, University of Reading - E-mail: Mark.Baker@computer.org - Web: <http://acet.rdg.ac.uk/~mab> - 19th May, 09