Cluster, Grid and Cloud Computing: A Detailed Comparison

Naidila Sadashiv
Dept. of Computer Science and Engineering
Acharya Institute of Technology
Bangalore, India
sadashiv@acharya.ac.in

S. M Dilip Kumar
Dept. of Computer Science and Engineering
University Visvesvaraya College of Engineering (UVCE)
Bangalore, India
dilipkumarsm.uvce@bub.ernet.in

***Abstract*—Cloud computing is rapidly growing as an alternative to conventional computing. However, it is based on models like cluster computing, distributed computing, utility computing and grid computing in general. This paper presents an end-to-end comparison between Cluster Computing, Grid Computing and Cloud Computing, along with the challenges they face. This could help in better understanding these models and to know how they differ from its related concepts, all in one go. It also discusses the ongoing projects and different applications that use these computing models as a platform for execution. An insight into some of the tools which can be used in the three computing models to design and develop applications is given. This could help in bringing out the innovative ideas in the field and can be explored to the needs in the computing world.**

***Index Terms*—Cluster Computing, Grid Computing, Cloud Computing, Computing Models, Comparison.**

I. Introduction

High-performance computing (HPC) was once restricted to institutions that could afford the significantly expensive and dedicated supercomputers of the time. There was a need for HPC in small scale and at a lower cost which lead to *cluster computing*. The emergence of cluster platforms was driven by a number of academic projects, such as Beowulf [1], Berkeley NOW [2], and HPVM [3]. The popularity of the Internet and the availability of powerful computers and high-speed network technologies has changed the way computers are used. *Grid computing* originated in academia in the mid 1990s with an aim to facilitate users to remotely utilize idle computing power within other computing centers when the local one is busy [1]. Initially, it only referred to a compute grid and had a rather limited audience. However, after years of development the grid gained momentum and came to mean an effective way for coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations. Cloud computing, is a kind of computing model that came into existence around the end of 2007. It provides a pool of computing resources which the users can access through Internet. The basic principle of cloud computing is to shift the computing done from the local computer into the network [4]. This makes the enterprise use the resource which includes network, server, storage, application, service and so on that is required without huge investment on its purchase, implementation, maintenance rather use it for other significant purpose. Resources are requested on-demand with out any prior reservation and hence eliminates over-provisioning and improves resource utilization.

To the best of our knowledge, in the literature, only a few comparisons have been appeared in the field of computing. In this paper we bring out a complete comparison of the three computing models. Rest of the paper is organized as follows. The cluster computing, grid computing and cloud computing models are briefly explained in Section II. Issues and challenges related to these computing models are listed in Section III. Section IV compares these three models from different perspectives. Section V discusses projects and applications. Tools and simulation environment are part of Section VI. Concluding remarks is presented in Section VII.

II. Three Computing Models

In this section, we briefly explain the three computing models: cluster, grid and cloud.

A. Cluster Computing

For many years supercomputer was the leader in the field of computing. But due to some of the problems faced in the area of science, engineering, and business, which could not be effectively dealt with using supercomputers. They were replaced with clusters [5] with the aim of overcoming these problems and also offered a very cheap way for gaining access to potentially huge computing power.

Definition: A cluster is a collection of parallel or distributed computers which are interconnected among themselves using high-speed networks, such as gigabit Ethernet, SCI, Myrinet and Infiniband. They work together in the execution of compute intensive and data intensive tasks that would be not feasible to execute on a single computer. Clusters are used mainly for high availability, load-balancing and for compute purpose. They are used for high availability purpose as they maintain redundant nodes which are used to provide service when system components fail. The performance of the system is improved here because even if one node fails there is another standby node which will carry the task and eliminates single points of failure without any hindrance [5]. When multiple computers are linked together in a cluster, they share computational workload as a single virtual computer. From the users view point they are multiple machines, but they function