Journal of Cases on Information Technology, 16(2), 1-14, April-June 2014

A Case Study of Teaching Parallel and Distributed Computing Topics on a Computer Cluster

Hong Lin, Department of Computer and Mathematical Sciences, University of Houston (Downtown), Houston, TX, USA

EXECUTIVE SUMMARY

This paper presents the establishment of cluster computing lab at a minority serving institution that aims to provide computing resources to support undergraduate computer science curriculum. The computing resources of the cluster are managed by a job distribution environment that allows the users to upload, compile, and run their jobs. The job distribution software distributes the submitted jobs to the computing nodes of the cluster. The authors will present a case study of using this platform to teach parallel and distributed computing topics in the operating system course. The evaluation of the teaching effectiveness is presented thereafter.

Keywords: *Cluster Computing Lab, Computing, Computing Nodes, Computing Resources, Teaching Effectiveness*

ORGANIZATION BACKGROUND

The University of Houston-Downtown (UHD) is a Minority Serving Institution (MSI). It is an undergraduate education institution. The Department of Computer and Mathematical Sciences is one of the departments of the College of Science and Technology. The department has a computer science program that offers BS in Computer Science. The diversity of UHD's student body lies not only in ethnicity but in academic preparedness. With the fast change in computing and information technologies, it is necessary to update the computer science program to accommodate these changes and provide a computing environment to allow students to develop hands-on skills in computation that is current and up to date (Lin, 2008a & 2008b).

In recent years, cluster computing proved itself to be a very efficient way to serve as a high performance computing platform with low cost. It also addresses the flexibility and scalability issues by allowing the upgrading of individual components in the cluster. With the support from an NSF grant, UHD established a computing

cluster to meet educational and research needs. The cluster has four segments, composed of different types of computers acquired in different times. A master server node connects all the segments together to form a bigger cluster or a grid. The portal on the server allows remote access to the computing resources. The portal is accessible from the webpage at http://grid.uhd.edu (Lin, Sirisaengtaksin, & Chen, 2010a & 2010b).

The intent of the presented project is to develop and implement a web interface for the cluster job distribution and multi-platform source handling. The goal of the project is to allow users of the computational cluster to remotely manage their files, and perform application deployment, execution and cluster job scheduling on the cluster. The means of remote access to the cluster resources are provided by the use of a web browser. The project should serve as framework allowing limited platform processing, compilation and execution of C, C++, and Java source code, and its effective distribution on the cluster. The framework can then serve for further expansion and development of modules to handle additional programming languages and platforms (Lin, Nguyen, & Saqer, 2009; Lin, 2007).

Needless to say, the project has an expected impact on utilization of the computational resources provided by the cluster; the ability to access and use the computational resources among faculty members, research personnel, and students remotely from web browser tremendously increases the access to harness the computational power of the cluster. Further, an anticipated outcome of the project is increased student participation in research and increased interest in fields like parallel and distributed programming (PDC), grid and high-performance computing (Lin, Ganov, Kemp, & Gilbert, 2012).

Among the uses of this multi-purpose platform, we present the use of this platform in teaching PDC topics in an undergraduate computer science curriculum. With the user portal of the cluster, students are able to access the computing resources online from their classrooms and experiment on various PDC topics. We present the use of this platform in teaching PDC topics in an operating systems course and evaluate the teaching effectiveness of the relevant course modules.

SETTING THE STAGE

The cluster has four segments, each having sixteen slave nodes and a master node. A master server node connects all the clusters together to form a bigger cluster or a grid. The portal on the server allows remote access to the computing resources. The portal is accessible from the webpage at http://grid.uhd.edu.

Intrinsically, the cluster portal allows the access to compilers and executors and the means to distribute content on the cluster. Through the web interface, a user is able to upload the source code (or any other content) of an application to the directory structure nested in their home directory, compile it, and/or execute it on one or multiple nodes of the cluster. The latter execution is dependent on whether the target application is sequential or parallel in nature. The web interface serves as a remote access tool to the backend workhorse. It takes the needed information from a user, it then creates a compilation and/or executor object, which in turn upon success contacts a job distributor to allocate resources on the cluster and finally dispatch the job onto those resources. The web interface allows the user to monitor the standard streams, and even provide input, if so the target application requires it.

The requirements of the application were set to:

- Provide means of user distinction, through the method of user authentication;
- Provide intuitive navigation through the application;
- Provide facilities for file manipulation, like directory browsing, file uploading and downloading;