Realistic Modeling of Resources for a Large-Scale Computational Grid System

Uei-Ren Chen, Hsiuping University of Science and Technology, Taiwan
Yun-Ching Tang, Hsiuping University of Science and Technology, Taiwan

ABSTRACT

To build a large-scale computational grid resource model with realistic characteristics, this paper proposes a statistical remedy to approximate the distributions of computational and communicational abilities of resources. After fetching the source data of resources including computing devices and networks from the real world, the gamma and normal distributions are employed to approximate probabilities for the ability distribution of grid resources. With this method, researchers can supply the computational grid simulators with required parameters relating to computational and communicational abilities. The experiment result shows that proposed methodology can lead to a better precision according to the error measurement of the simulated data and its source. The proposed method can support existing modern grid simulators with a high-precision resource model which can consider the characteristics of distribution for its computational and communicational abilities in the grid computing environment.

Keywords: Communication, Computational Grid, Grid Resource Model, Resources, Source Data

INTRODUCTION

Computational grid (Foster & Kesselman, 1998) provides us a powerful ability to solving the large sale problem efficiently in modern distributed computing systems. Experiments on the grid computing can usually be performed on the hardware framework or the software simulation. The framework of computational grid can be established by the middle-wares, such as Globus Toolkit (Globus, 2011), Oracle Grid Engine Software (Oracle, 2011), and gLite (2009), etc. We know that the hardware framework is more realistic than software simulation. However, high costs for devices and complex system management result in significant difficulties when grid researchers intend to build a large-scale grid environment. Comparing with building a large sale hardware platform, software simulation is more elastic and economic for researchers at the beginnings of analyzing the grid system.

Related simulation tools for grid computing are MicroGrid (Liu, Xia, & Chien, 2004; Song, Liu, Jakobsen, Bhagwan, Zhang, Taura, & Chien, 2000), GridSim (Buyya & Murshed,
and the grid resource models can be found in these simulators. For simulation tools, the grid resource models have direct influence on simulation result. The common drawback of these grid resource models is usually simple and not realistic enough, so we need some real evidences to make up its completeness, and improving the accuracy of software simulation becomes a critical issue. In our previous work, an Internet-scale grid resource topology generator with aggregative characteristics has been discussed and established (Chen, Chen, Wang, & Lin, 2007). The grid resource topology can be easy generated and integrated with existing grid simulators to perform numerous experiments. Moreover, by using the topology transformation technologies, such as, the PaGrid (Huang, Aubanel & Bhavsar, 2006) or a proposed virtual mesh transformation method on a computational grid (Chen & Lin, 2010), the irregular grid resource topology can be transformed into a virtual mesh. The benefit is that researchers can construct a problem-solving environment on a computational grid so as to solve a large-scale mesh-based problem by traditional parallel processing algorithms (Grama, Gupta, Karypis, Kumar, 2003). The mesh-based problems including matrix multiplication, sorting, fast Fourier transform (FFT) and partial differential equations (PDE), etc., are of fundamental importance to the scientific computing applications (Karniadakis & Kirby, 2003). Before performing these topology transformations, the realistic grid resource topology surely takes a critical position in the experimental simulation. This research puts emphasis on how to build a grid resource topology with realistic characteristics. To carry on the progress of this topic, the characteristic distribution of real grid resource topology such as computational ability and communicational ability will be analyzed and discussed in this paper. The existing grid simulators just simply or randomly configure the resource models without consideration of characteristics in the real grid resources. This effort is unprecedented and is made to rectify this problem.

The remainder of this paper is organized as follows. The next section defines the grid resource model and its essential components. The proposed approximating method is introduced step by step in Section Methodology, and it is illustrated by two examples. In the fourth section, the experiment results are presented with a discussion. The final section draws a conclusion for this research.

GRID RESOURCE MODEL

The computational grid includes a set of resources, network devices, and links used to connect them. Theoretically, the grid resource model can be represented as an undirected connected graph \( G = (R, L) \), where \( R \) is the resource set, and \( L \) is the link set. The resource set consists of computational nodes and routing nodes. The routing network of a grid resource model is composed of routing nodes and links. An example of the grid resource model is given in Figure 1. In this figure, there is a routing network with six routing nodes and nine computational nodes attached to them.

Computational Nodes. In the proposed grid resource model, one of the components, computational node, may be regarded as a supercomputer, a cluster of workstations or computing devices practically. The computational node is denoted as \( c \) with a parameter \( P_c \) called computational ability. In the top 500 supercomputers list (it is hereafter abbreviated Top500) (Top500, 2011), the computing ability of computers are ranked by their \( R_{\text{max}} \) value, which is the maximal LINPACK (2011) performance achieved in GFLOPS. In our research, this value of \( R_{\text{max}} \) in Top500 can be used to represent the concept of computational ability \( P_c \) for each computational node \( c \).

Routing Networks. The routing network \( G' = (R', L') \) is a sub-graph of the grid resource model \( G = (R, L) \), where \( R' \subseteq R \) is the set of routing nodes, and \( L' \subseteq L \) is the set of links that connect the routing nodes in \( R' \).
Grid Transaction Management and Highly Reliable Grid Platform
www.igi-global.com/chapter/grid-transaction-management-highly-reliable/64547?camid=4v1a

Migrating to the Cloud: The Sullivan University Experience
www.igi-global.com/article/migrating-to-the-cloud/149916?camid=4v1a