CAD: An Efficient Data Management and Migration Scheme across Clouds for Data-Intensive Scientific Applications

許慶賢,Alfredo Cuzzocrea,陳世璋
Computer Science & Information Engineering
Computer Science and Informatics
chh@chu.edu.tw

Abstract

Data management and migration are important research challenges of novel Cloud environments. While moving data among different geographical domains, it is important to lower the transmission cost for performance purposes. Efficient scheduling methods allow us to manage data transmissions with lower number of steps and shorter transmission time. In previous research efforts, several methods have been proposed in literature in order to manage data and minimize transmission cost for the case of Single Cluster environments. Unfortunately, these methods are not suitable to large-scale and complicated environments such as Clouds, with particular regard to the case of scheduling policies. Starting from these motivations, in this paper we propose an efficient data transmission method for data-intensive scientific applications over Clouds, called Cloud Adaptive Dispatching (CAD). This method adapts to specialized characteristics of Cloud systems and successfully shortens the transmission cost, while also avoiding node contention during moving data from sites to sites. We conduct an extensive campaign of experiments focused to test the effective performance of CAD. Results clearly demonstrate the

improvements offered by CAD in supporting data transmissions across Clouds for data-intensive scientific applications.

Keyword：Data Management