A New Electronic Commerce Architecture in the Cloud

Guigang Zhang, Tsinghua University, Beijing, China
Chao Li, Tsinghua University, Beijing, China
Sixin Xue, Renmin University of China, Beijing, China
Yuenan Liu, Renmin University of China, Beijing, China
Yong Zhang, Tsinghua University, Beijing, China
Chunxiao Xing, Tsinghua University, Beijing, China

ABSTRACT

In this paper, the authors propose a new electronic commerce architecture in the cloud that satisfies the requirements of the cloud. This architecture includes five technologies, which are the massive EC data storage technology in the cloud, the massive EC data processing technology in the cloud, the EC security management technology in the cloud, OLAP technology for EC in the cloud, and active EC technology in the cloud. Finally, a detailed discussion of future trends for EC in the cloud environment is presented in this paper.

Keywords: Architecture, Cloud Computing, Cloud Database, Cloud File System, Electronic Commerce, HUABASE Database, Semantic Analysis

INTRODUCTION

Electronic commerce is becoming more and more important in the 21st century. Many traditional trade activities have been migrated into the internet. Electronic commerce is playing a more and more important role all over the world. According to the report of electronic commerce research and development (Aliresearch, 2011), the e-commerce transactions accounted for a total GDP of China in 2011 is 12.1%. And it will arrive at the 20% in the next five years. Now, more users have become the participants of electronic commerce. They buy and sell products from the EC web sites. The electronic commerce has undergone three stages. The first stage is from the 1990 year to 2000 year. In this stage, the electronic commerce concept begins to be formed. The most important form is that the companies publish their commerce information in their web pages, the buyers or sellers can find the commerce information from these web pages, and they can complete their commerce trade offline. The second stage is from the 2001 year to 2010 year. In this stage, the electronic commerce has a very big de-
velopment. Lots of electronic commerce web sites occur. The national trade and international trade have gotten a very big development. Lots of traditional companies have developed their electronic commerce, too. The main electronic business models (He, Jennings, & Leung, 2003) have B2B, B2C, B2G, and C2C. The most important electronic commerce web sites have Ebay, Alibaba, Taobao, etc. In the ten years, the EC has become the most important commerce activity. The third stage is from the 2011 year to now. In this stage, the electronic commerce has entered into a new era. We can name it the new electronic commerce era. The traditional electronic commerce has little intelligence. It only need help the sellers and buyers can complete their trades online or offline. However the new electronic commerce has more intelligence. It not only needs to help all buyers and sellers to complete their trades, but also can help the buyers and sellers can analyze their markets. At the same time, all the sellers and buyers want to have more interactivities each other and wish exchange more information each other.

In the new electronic commerce stage, the cloud computing will play an important role. Lots of cloud technologies will be applied in the EC. The mainly technologies include the cloud storage technology, the big data processing technology, the intelligence computing technology in the cloud, how to ensure the electronic commerce security in the cloud and so on.

In order to storage big data for all kinds of electronic commerce applications, lots of cloud databases and distributed file systems occur. As we know, in the traditional electronic commerce, the data only include the structured data, and all these structured data will be stored into the relational databases such as the Oracle, My SQL, DB2, and SQL Server so on. However, in the new electronic commerce stage, lots of structured data, semi-structured data and non-structured data should be stored all together. In order to store all these non-structured data, lots of distributed file systems occur. The most important distributed file systems are Google’ GFS (Ghemawat, Gobioff, & Leung, 2003), Hadoop’s HDFS (Ghoting & Pednault, 2009), KFS (Gomez, 2010), Taobao FS (Taobao, 2010), and Hystack (Beaver, Kumar, & Li, 2010). All these distributed file system can run on top of millions of machines and they can manage all files automatically. In addition, lots of semi-structured data need to be storage, too. And so, lots of cloud databases have been developed, too. The most important cloud databases have the Google’s BigTable (Chang et al., 2006), Hadoop’s HBase (Jianling & Qiang, 2010), HyperTable (2012), Facebook’s Cassandra (Avinash & Prashant, 2010), and Amazon’s SimpleDB (Amazon, 2009) so on. For example, the Google uses the BigTable to storage the index information of all web pages files in the internet. All these cloud databases uses the key/value mechanism.

In order to process massive data in the cloud environment, lots of new technologies occur, too. Google MapReduce is the most important massive data processing framework for Google’s all kinds of applications. Each Google search will use the MapReduce to process massive data. Based on the Google MapReduce, the Hadoop MapReduce emerges. It is the open source version of Google MapReduce. However, the MapReduce (Biswanath, Herbach, & Basu, 2009; Lin, 2010; Kovoor, Singer, & Lujan, 2010; De Kruijf & Sankaralingam, 2009) still cannot process some applications well with the loop conditions. And so, lots of others big data processing frameworks emerge such as the Haloop and Twister. The Haloop (Bu, Howe, & Balazinska, 2010) and Twister (Jaliya, Hui, & Bingjing, 2010) are the iterative MapReduce. They can satisfy these applications with loop conditions well. Although these massive data processing frameworks can process data quickly, they have little computing abilities. Unlike the relational databases, they can’t execute complete computing such as complex queries, complex join and some other complex computing. In order to improve the computing ability, lots of technologies attempt to combine the databases and these frameworks together. The HadoopDB (Abouzeid, Bajda, & Abadi, 2009) is the most famous product which hopes to combine the MapReduce and the Parallel re-
A New Electronic Commerce Architecture in the Cloud
Guigang Zhang, Chao Li, Sixin Xue, Yuenan Liu, Yong Zhang and Chunxiao Xing (2012). Journal of Electronic Commerce in Organizations (pp. 42-56).
www.igi-global.com/article/new-electronic-commerce-architecture-cloud/72999?camid=4v1a