Thoughts after reading Google Bigtable

It is the last paper of the three google paper, and it is time limited, so I can only look through it, and may not be quite right about what I said, I just share some of my feelings.

I'm quite interested in the title of this paper, ”Bigtable”. BigTable is Google's distributed structured data storage system. It was put into use around April 2005. The design and implementation of this system took seven years.

In 2004, Google released this paper, which is a programming model based on GFS system for processing and generating large data sets. As a computing system, MapReduce consists of map and reduce. In GFS, the master divides the data files to be processed into multiple files and distributes them to each computer for reading and processing. Finally, the reduce function combines all the processing results and outputs them. The computing power of personal PC is idle in idle time, and the system can make use of these idle computing power. The parallel processing time of thousands of devices is also relatively short.

At the beginning of this article we describe what BigTable is, its meaning in Google and its application in life. Currently, its existence has achieved a wide range of objectives, such as applicability, scalability, high performance and availability.

The last section begins with a detailed description of the system components. The data model is described in detail, including the API client and the main components of BigTable. GFS stores protocols and data files. The internal data storage file takes over the format of Google sshable. It also relies on high availability and a serialized distributed locking service component called Chubby.

It also describes BigTable, which consists of three main components: a library linked to the client program, a master server and several tablets.

Finally, the practical application of BigTable is discussed. The company said that from August 2006, 388 untested Bigdale clusters run on different server clusters within Google, with total around 24500 tablet server. He also talked about the role of Google Analytics,

It can be accessed through Google Earth's customer-specific client software.

When I read this article, I welcome the growing power of science and technology in the world. I also admire the researchers who have contributed to the darkness of global technology. They involve young people in scientific research and let us experience the benefits of technology.

After reading all of the three paper, I found that the three paper are highly interrelated. The publication of three papers has brought about the era of big data, and Google has taken the lead in this era. Today, we are still not satisfied with the old data models and data processing methods. More and more massive data are generated every day in the world, so we need more efficient and lower cost processing methods.

I have to say that I really have learned something after reading the three google paper, and I think it deserves reading a second or even third time, just like the teacher said, it is quite benefit for us to read some classic paper.