

# Combined Application of Cloud Computation Technology and Business Intelligence

Fengyu Wang, Aimei Yan, Lifang Yang

Third Research and Development Institute of Smart Grid  
Guodiantong Corporation, State Grid Information & Telecommunication Corporation  
Beijing in China, 100761  
fengyu\_wang@sgcc.com.cn, blue\_drop0813@163.com, funy-2008@163.com

**Abstract**—In paper, in order to use the data produced in the development process of epc system network reasonably, Cloud Computing (CC) and the related technologies are presented, and the conception and architecture of Business Intelligence (BI) are introduced. Then the platform architecture of BI based on CC technology, which is also named as the platform architecture of “CC & BI”, is put forward and compared with traditional BI, which obtains the advantages of architecture of “CC & BI”. Finally, it looks into the beautiful future of the applications of “CC & BI”.

**Keywords**—cloud computation; business intelligence; combined application

## I. INTRODUCTION

Business Intelligence (BI) system [1] provides the technologies and methods to analyze data quickly, transform present data of enterprises into knowledge, and help managers to make sensible decisions. It can extract and organize available information to help users to make timely and accuracy decisions to promote management, marketing and company development by fully using present business information and modern techniques. Also, BI can improve financial performance and service to sustain competing merit and bring more economics benefits for company through analyzing and holding new opportunities and finding potential threaten.

Cloud computation (CC) [2] is commended as “New Generation Industrial Revolution”, which is an information technological revolution in the world. CC technology is characterized as large scale, high elastic computing, unlimited extending storage, high data security, and high cost performance and so on, all of the advantages provide good platform for data storage and computing for smart grid [3][4], so as to be applied in the fields such as monitoring the condition of power grid, power grid emulation, power grid planning, handling massive information of intelligent power consumption [5], transaction in the power market and visualization management of power grid, and so on.

In recent years, “CC & BI” is applied in the domestic and overseas. The successful instances of the famous enterprises at home and abroad fully proved that the combined application of CC & BI is reasonable. For example, China Mobile developed

cloud computation system based on the open source software. The version of BI mobile of UFIDA Company met the requirement of CC. And compared with traditional BI, one of the advantages of BI Silver Spotfire of Tibco is that its users can set new application with no back-end basic equipment, so that it saves much time. The applications of CC in the telecom and BI analysis were also the examples of “CC & BI”. Moreover, “CC & BI” is also applied in Facebook and Taobao, and so on.

CC technology has the ability of handling large scale of data, and analysis mining technology and method supplied by BI system would achieve the fast transformation of data to knowledge. Therefore, it's important to research combined application of “CC technology & BI”.

## II. BASIC PRINCIPLE OF CC AND BI

### A. Basic principle and related technologies of CC

The basic principle of CC [6] is to distribute compute into massive distributed computers, not into local computer or remote server, and the runtime of data center in the enterprises is more similar with epc system network, so that enterprises can switchover resources into applications in need and access computers and storage system according to the requirements. CC integrates lots of data resources and processing unit resources stored in the distributed computers to work together. CC is a new method to share basic architecture that can ally the huge system resources to provide different services.

CC technology includes virtualization, auto deployment, extension of application scale, distributed file system, distributed data base, unstructured data storage and distributed computing. CC generates some excellent distributed file system and cloud storage service, so as to overcome the system bugs resulted by single nodal point bug, and to achieve the storage of large scale of files. The distributed storage system based on the distributed file system makes the massive structured data stored in it unstructured. MapReduce programming model based on the most typical distributed computing mode of cloud platform would divide large task into many mini-sub-tasks, which are dispatched and computed in lots of computing nodal points, so as to obtain the capacity handle huge data on the platform.

### B. Basic Principle of BI and Complete Architecture

BI is usually understood as a tool which transforms existing data into knowledge to help managers to make sensible decisions. BI [7][8] is a data driven decision support system (DSS) that combines data gathering, data storage, and knowledge management with analysis to provide input to the decision process. Technically, it integrates ETL (Extract, Transformation, and Load), data warehousing, data mining

(DM) [9][10] and data visualization, such as online analytical processing (OLAP). In the recent years, with BI technology continuously perfect, more and more enterprises applied BI to maintain and improve its competitiveness, and BI suppliers and institutes at home and abroad have carried on basic research and applied research of BI extensively.

The complete architecture of BI is shown as Fig.1.

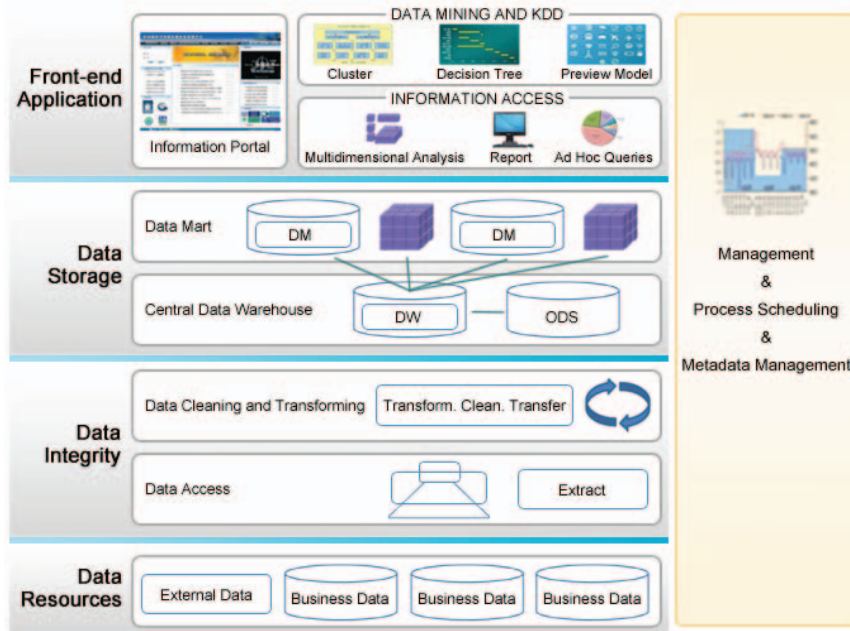


Figure 1. The complete architecture of BI.

The architecture of BI can be divided into four layers according to different forms of data, and 7 links according to processing and application process of data. The questions from data resources to central data warehouse (DW) by ETL, then to data mart by classifying, or to dimensional database, are the questions about data organization, and from interlayer or dimensional database to terminal users are about front application realization. However, the system process scheduling control and metadata management is penetrating the data processing link in the whole system.

### III. PLATFORM ARCHITECTURE OF BI BASED ON CC TECHNOLOGY

At present, with the development of informative construction of enterprise, enterprises has collected and stored massive basic data. It is an important question for the next development of smart grid that how to use the huge data reasonably and fully. CC technology has the ability of handling large scale of data, and analysis mining technology and method supplied by BI system would achieve the fast transformation of data to knowledge. Therefore, it's important to research combined application of "CC technology & BI".

#### A. Introduction of Architecture of "CC Technology & BI"

The architecture of BI based on CC technology is put forward in paper, which is subdivided into 7 layers, shown as Fig.2.

- (1) Data source layer: It comprises all key data in need.
- (2) Data integrated layer: It integrates various data from data source, handles and transforms the heterogeneous data of data source layer to form unified connected data view, then divides into different topic data according to business needs. Meanwhile, another function of the layer is to reduce the impact of process of data extraction on each business system of data source layer.
- (3) Data storage layer and data analysis layer: The key function of the two layers is to achieve CC of system, data from ODS is stored in the HDFS supported by Hadoop in the form of text file, and reach the parallel handles of multi-nodal-point and multi-task by Hive + MapReduce. It can improve computing ability so as to reduce the time of data process.
- (4) Data cushion layer: It can analyze the cushion and summary of topical data and build a series of data summary and select views from different angles. On the one hand, it

simplifies the logical complexity of front display; on the other hand, it improves corresponding data speed so as to reduce the time of data access.

(5) Display layer: The independent developed tool of front display could generate characterized charts more flexibly according to business needs, which is of good extension.

(6) Business application layer: It is mainly various business applications supported by the architecture.

The physical architecture figure of “CC & BI” is shown in Fig.3.

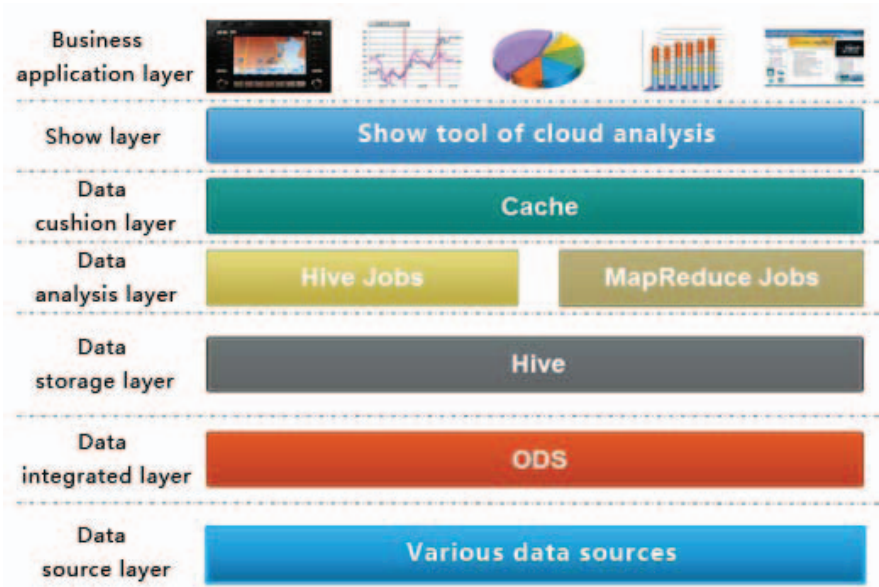


Figure 2. Platform architecture of BI based on CC technology

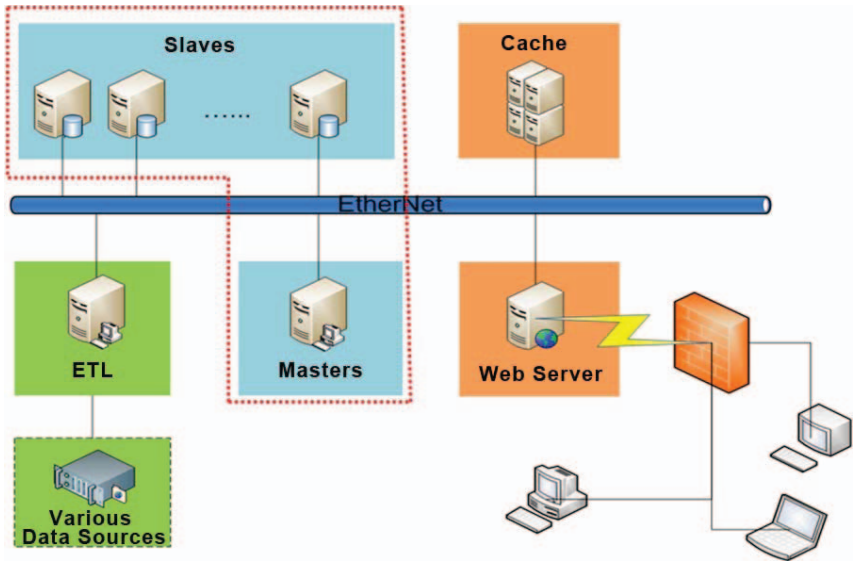


Figure 3. Physical architecture figure of “CC & BI”

B. Comparison of Traditional BI Architecture and One Based on CC

The traditional BI architecture generally adopts mature method that includes ETL tools, softwares of data warehouse

and report forms tool, whose disadvantages are mainly listed as belows.

(1) Support computing of data of TB level, but no support computing of data of PB level and parallel computing.

(2) Weak extension of hardware. Because mature software of data warehouse are stand-alone at present, not distributed. It can't deal with problem simply by increasing the number of machines when data in the data warehouse increasing over the maximum of the machine, so the machine with better function and more expensive will substitute for it.

(3) High function server of high cost.

The architecture of BI based on CC adopts multi-nodal-point distributed structure, whose advantages are as follows:

(1) Support computing of data of PB level and parallel computing.

(2) High security. Because the architecture have two backup copies at least for each file, that's to say, at least 3 same copies of data are stored in the system.

(3) High efficiency data storing and computing. Because the parallel computing adopted in the architecture can divide a task into several sub-tasks and distribute to several machine to work together, so as to improve efficiency data storing and computing.

(4) Support extension of hardware. The cloud data warehouse in the architecture supports distributed hardware deployment, and it will have no impact when adding or deleting random hardware nodal point. When magnifying the number of system data, it only needs to add hardware of low cost.

(5) Low system cost. Because hardware needed in the system can use common servers or PC, its hardware is with low cost.

#### IV. CONCLUSIONS

One of the development directions of BI market is its combined application of CC technology. Enterprises will improve the interesting on Hadoop framework of open source software, which supports more analysis and data types, to compare and analyze larger scale of data, and to simplify the complex analysis. To apply CC technology in the analysis of BI can improve the efficiency of data analysis, help enterprises to accommodate the rapid changing market, and provide accordance for inventing new products. Hadoop, which is

combined BI & CC, is adopted by many enterprises to handle huge logs, click flow analysis, simulate the finance, handle files, Web indexes and so on, and the data warehouse usually adopts Hive. Moreover, the huge data generated with the rapid development of smart grid in the power field can be dealt with by combined method of BI & CC.

Therefore, we should deal with all the technique questions possibly appeared in the future research, so as to make the applications of CC technology more perfect.

#### REFERENCES

- [1] M. L. Gonzales, K. Bagchi, G. Udo, et al., "Diffusion of Business Intelligence and Data Warehousing: An Exploratory Investigation of Research and Practice," *Proceedings of the 44th Hawaii International Conference on System Sciences*, pp. 1-9, 2011.
- [2] Y. X. Li, C. H. Yang, J. X. Tian, "Computation mode of information construction for medium and small businesses based on cloud computation", *Hebei Enterprises*, vol. 6, pp. 70, 2009.
- [3] Q. P. Wang, C. N. Lin, X. H. Yang, et al., "Scheme of Intelligent Community Based on Distributed Generation and Micro-Grid," *International Conference on Power System Technology*, pp. 1-6, 2010.
- [4] K. Thananunsophon, B. Mangalabruks, Y. Fujii, et al., "Community Monitoring and Security using an Intelligent Camera in EAT Smart Grids," *Procedia Engineering*, vol. 8, pp. 332-336, 2011.
- [5] J. S. Wen, H. W. Song, X. N. Liu, et al., "Research on Data Acquisition System of Embedded Intelligent Uptown," *Journal of Shijiazhuang University*, vol.12, pp.74-79, 2010.
- [6] X. C. Chen, "Study on application of cloud computation in technical support system of smart grid dispatching", *East China Electric Power*, vol. 38, pp. 800-804, 2010.
- [7] K. Rezaei, A. Ansarinejad, A. Haeri, et al., "Evaluating the Business Intelligence Systems Performance Criteria Using Group Fuzzy AHP Approach," *UKSim 13th International Conference on Modelling and Simulation*, pp.360-364, 2011.
- [8] L. y. Tang, Z. W. Ni, Z. J. Wu, L. Wang, "A Conceptual Architecture for Business Intelligence as a Service (SaaS BI)," *Fourth International Conference on Intelligent Computation Technology and Automation*, pp. 1025-1028, 2011.
- [9] J. Y. Wu, "Survey of the Application of Data Mining and Data Warehouse Technology in the Electronic Commerce," *Journal of Fujian Institute of Education*, vol. 4, pp. 126-128, 2010.
- [10] Y. H. Li, "The Study of Applying Data Mining Technique on Customer Relationship Management of Third Party Logistics," *Market Modernization*, vol.1, pp.125-126, 2009.