International Journal of Database Theory and Application
Vol.8, No.6 (2015), pp.333-342
http://dx.doi.org/10.14257/ijdta.2015.8.6.30

# The Green Project of Data Management for Industrial Integration

Yiming Tian [1], Keyu Zhu [2, 3], Ruoyu Pan [2, 3] and Changyong Liang [2, 3]

*[1] School of Computer & Information Hefei University of Technology, Anhui Transport Interconnection Management Center, Hefei, China*
*[2] Institute of Computer Network Systems Hefei University of Technology, Key Laboratory of Process Optimization and Intelligent Decision-making, Ministry of Education, Hefei, China*
*[3] Anhui Engineering Technology Research Center for Key Technologies & Equipment of IOT of Highway Traffic*
*Email: benzlion@163.com*

***Abstract***

*The green project fulfills the requests of industrial management and application by sharing data rather than isolating. The data center with energy-efficient equipment is correspondingly effective in industrial service, management and construction of information. This study suggests a green data center in the two-level framework as a basic factor in industrial integration with potent procedures. In the case of transportation, it processed validating data to resolve logistics and emergency operations.*

***Keywords:*** *industrial integration, data center, service and management, the green mode*

## 1. Introduction

Data sharing and business collaboration are rituals of in formalization. The connectivity of communication systems is prerequisite to industrial integration process involving basic data unity, information isolated islands and intensive management. The traditional data centers in the collection of data and facilities perform well for less collaborative enterprises. However, industrial integration demands the elaboration of data systems in operation, under construction and to be built, a tough part of application integration. Green development mechanism [1] claimed in the Eleventh Five-year Plan, the policy for low carbon growth and energy efficiency management, is supposed to result in a sustainable prospect for industrial enterprise and mechanism.

This study [2] examined the network oriented data center and the service oriented data center. Infrastructure has varied from various storage media including tabulating cards, tape, hard drives and disk arrays, to cloud devices. Accordingly, the architecture and the relevant software applications, such as network database, relational database, virtualization, automatization and service oriented computing, came out. Figure 1 illustrates the development.

ISSN: 2005-4270 IJDTA

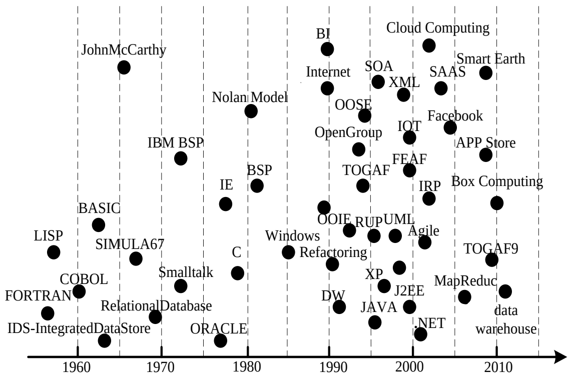


**Figure 1. The Development**

Roth estimated energy consumption by operation parameters of servers in internet data centers (IDC) installed in every continent, which was a rough calculation. Mitchell Jackson [3] measured that of data centers with advanced instruments offering sufficient statistics, which was relatively accurate.

## 2. Green Resource Management and the Green Mode

### 2.1. Green Resource Management

The resource of low carbon and energy efficiency is the heart of data center construction. The green resource, an overall service/management-directed plan, sets infrastructure norms.

Data centers, involved in performances such as collection, processing, storage and transferring of the required data, are the key to industrial management and application. Informalization corresponding to the intensive production reform accounts for specific communication platforms and standards in individual industries. The thoroughly reviewed plan with management and service goals, which holds true to green principal, is prior in decision making in the same industrial relations, absolving the accumulation of information isolated islands and hardware.

According to Green Grid, the green criteria of facilities are power usage effectiveness (PUE) and data center efficiency (DCE) [4], which are given by:

$$PUE=\frac{TotalFacilityPower}{ITEquipmentPower}$$

$$DCE=\frac{ITEquipmentPower}{TotalFacilityPower}$$

TotalfacilityPower is the total energy used by the data center consisting of numerous devices such as power supply module, cooling system, calculation nodes and storage nodes, lighting system and fire control system. ITEquipmentPower is the processing equipment loads and the energy consumed by accessory equipment including KVM and work stations.

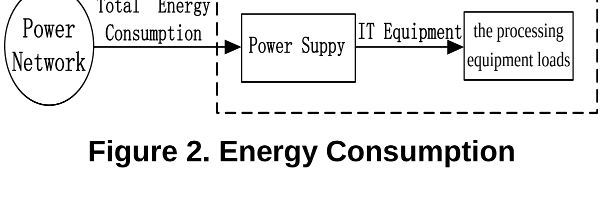


**Figure 2. Energy Consumption**

The amount depends on the accurate load of precision cooling system and IT device integration. Green Grid modified the equation.

 

$$PUE = CLF(Cooling\ Load\ Factor) + PLF(Powe$$

CLF is the ratio of the total cooling load to ITEquipmentPower, while PLF is the ratio of the power supply module load to ITEquipmentPower.

DCPE, a third green norm, is given by:

$$DCPE = \frac{UsefulWork}{TotalFacilityPower}$$

The input and the output are the only variables in the equation. As the black box, the data center is in short of evidence to validate UsefuleWork. Blazek [5] conducted the quantitative research on MB/kWh parameter.

QoS and particular operations decide the other green norms, such as Energy Consumption Rating (ECR), Energy Efficiency Rating (EER), Energy-Response Time Product(ERP) and Energy Proportionality Index (EPI) [6-8].

The cooling system load enormously accounts for the energy consumption. The prevailing optimization measure is to reduce physical constraints of facilities, for instance, Scofield invented wet bulb economizers [9] with flexible humidity requirement and scientists study air distribution for air flow adjustment of air conditioners. [10-11]

Inspired by the solar database [12], Deng Wei reviewed the data centers powered by the renewable energy [13], which is the alternative for big data processing.

## 2.2. The Green Mode

The green mode, the management mechanism responsible for the high efficiency of facilities such as internet equipment, storage devices, servers and software, is the expected target of data center construction.

In the case of internet, the energy-saving management mode is formed by the Table1.

**Table 1. Energy-Saving Strategy**

| Criteria | Category | Instruction |
|---|---|---|
| **time scale** | online/offline | upgrade frequency of strategy |
| **(online) range** | partial/whole | communication traffic of mission completion |
| **TCP/IP** | links, networks, transferring, application, cross layers | cooperation substances of strategy enforcement |
| **Inputting** | instantaneous, historical, predictable | algorithm acquisition and adaptation |
| **methodology** | flow analysis, modeling, simulation | research and effectiveness |

The ACPI specification [15] by Intel, Microsoft and Toshiba defines the dynamic power management of computer system in terms of the global state and the power state of equipment and processors. Node level energy-saving is feasible on fundamental facilities. Optimized systems and energy reservation are viable to protocols and evaluations. What is widely discussed is the optimization in multiple constraint conditions. Linear and heuristic algorithm offer the answer [16-18]. According to load and energy consumption, Dou Hui suggested the optimization strategy [19]. Ye Kejiang of Zhejiang University examined energy-saving/emission-reduction models of cloud computing [20].

With software updates lagging far behind hardware upgrade to physical servers, there are immense physical resources taken less and least advantage. The solution lies in virtualization, which allows immediate distribution of server resources by priority to the system urgently demanding [21]. Being part of the server consolidation preferred in data center construction, the technology is beneficial to hardware cost saving and power saving. The cloud infrastructure is virtualization-based structure.

In March 2011, the European Grid Infrastructure (EGI) published the report of InSPIRE Project (Integrated Sustainable Pan-European Infrastructure for Researchers in Europe), which was ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, to provide a sustainable, reliable e-Infrastructure that can support the needs for large-scale data analysis. Earlier that year, Federal Cloud Computing Strategy by the US government reported barriers to effective communication such as the low facilities utilization, individual resource requirements, the repeated system constructions, ineffective operating environment management and long purchasing cycles. The federal government allocated a quarter of 80 billion-dollar IT budget for cloud computing research. Chinese government decided on the industrial policy of high-tech services aiming at cloud computing (in Guiding Opinions of the General Office of the State Council on Accelerating the Development of the High-tech Service Industry by the National Development and Reform Commission, the Ministry of industry and the Ministry of science and technology). The first batch of pilot cities, including Shenzhen, Hangzhou and Wuxi, took the lead in cloud utilization, for instance, the Sea of Clouds Plan in Shanghai and Auspicious Cloud Computing Project in Beijing.

Cloud computing, as green computing, is the hardware as well as system software in a data center. Clouds are classified as public (open for public use) and private (operated solely for a single organization). The computing allows dynamic deployment, reallocation and real time monitoring of virtualization and storage resource pools, and provides storage, processing and platforms in accordance with QoS standard for users.

The model in this study is a two-level SOA-structure data center with service bus in industrial applications to process data with business intelligence tools.

## 3. The Two-Level Data Center

### 3.1. Target Fixation

Industrial integration and other requests including network bandwidth determine a two-level module. A data center is not accessible in any case. On the basis of frequency and importance, Level I center extracts the basic data, the key indicators and the result set data. By business and regional demand, Level II center process the business data and the regional result set data.

Level I center, in which collecting, processing, storing and transferring occur, is the intersection of industrial management.

Level II center of the processed data of a specific industry or a sub industry submits the basic data of the result set to Level I center.

### 3.2. The Structure of the Two-Level Data Center

The required Level II center of management embranchment and business embranchment offers data on the exchange platform for Level I center, which is illustrated by Figure 3.

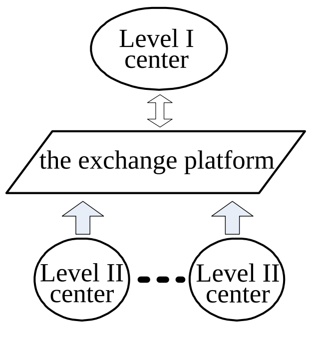


**Figure 3. The Two-Level Data Center**

The two-level data center processes data categorized by local and industrial requests, operating for industrial integration but consuming less energy.

### 3.3. The Data Flow

Data source consists of business data and statistics of industries, such as labors, materials, projects, assets, plans and security. Business data aggregate in the corresponding embranchment of Level II center and Level I center. Basic data in the management embranchment of Level II center are those in the unified format in the ODS (operational data store) of Level I center.

Level II center consists of communication systems and sub systems resulted from business development and the elaboration of the existing units. Figure 4 illustrates the structure.

## 4. The Structure of Level I Center

### 4.1. The Structure

As proven by examinations, a secure, standardized Level I center supports business development in multiple aspects, which is illustrated by Figure 5.

The industrial application is a comprehensive system of data centers of SOA (service oriented architecture) and ESB. The former structure provides the unified users' access. Other necessary techniques are ETL, relational database, spatial database, data management, data exchange, data warehouse, in-line analysis and data mining.

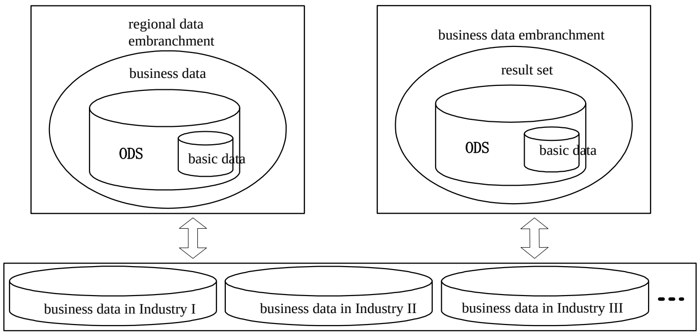


**Figure 4. Level II Center and the Data Source**

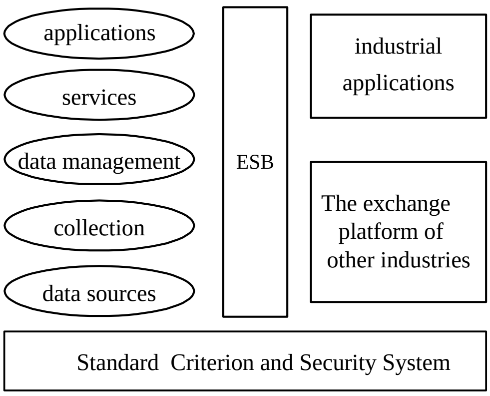


**Figure 5. The Structure of the Level I Center**

### 4.2. Multiple Aspects of Basic Support System

Data sources are available statistics in diverse units of the current and in-construction systems in the data center. The information dispersed in specific electronical or paper medium is either structured or unstructured such as videos and graphics. Data sources provide basic data for aggregation, storage and presentation.

Data collection，the process of gathering and measuring information，is the direct data exchange with desktop databases (access, excel, txt, xml) and text files. It is accessible to SOA structure, direct open databases and text input/output in various formats (word, excel). The required data interfaces are available. Collection methods are ETL (extract, transform, load), online-offline collection and data aggregation and exchange on the service bus of web service.

Data management is the intensive processing of business data in operational data store (ODS), database management of data on specific platform or data warehouse (DW) management. The key factors of data center are the exact ODS, DW, data mart (DM) and metadata for mass data. Relevant system applications on ODS and data exchange with between comprehensive application platforms and enterprise service bus (ESB) open to SOA take place. Data management allows online analytical processing (OLAP) and data mining. Figure 6 illustrates data management.

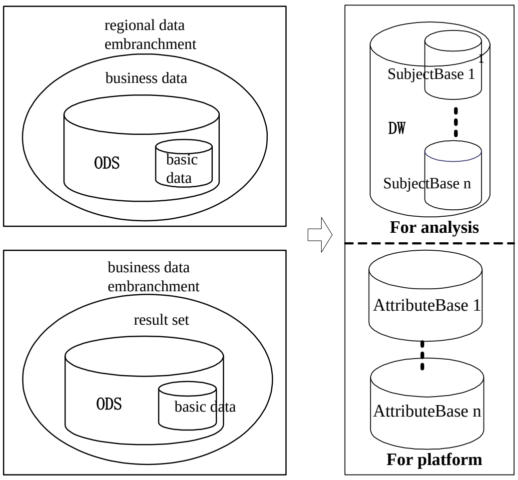


**Figure 6. Data Management**

Services are OLAP and data mining for intelligent decisions and data analysis. Figure 7 illustrates services.

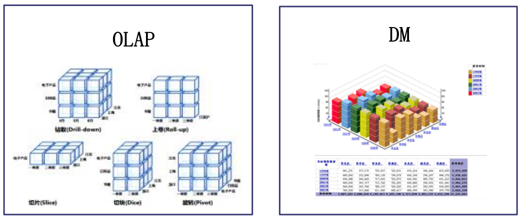


**Figure 7. Services**

Most applications are of the view to ODS and business intelligence (BI). The others are of data exchange between data centers on service bus and industrial systems or inter-industrial exchange.

 

### 4.3. Green Management Platform

Green storage reform is essential. Traditional storage technique only meets particular information needs. This explains massive isolated islands of storage systems in the existing data centers. Energy consumed by memory ram and cooling devices, along with additional sources for system maintenance, is in demand. Life circle management of data centers relies on green storage.

Network-attached storage (NAS), a synthesis of hardware and software, allows access to storage devices by infrastructure of IP network. Remote access depends on the distributed storage. NAS provides storage aggregation and sequential services. The storage infrastructure is the optimized network entity, submitting massive raw data accessible to client servers to application servers. SAN network consists of servers and storage devices of IP/Ethernet network, which supports transparent life circle management and automatic data distribution to the corresponding storage section and consumes less energy.

Virtualization integration is highly effective in management and tracing and server maintenance. Heterogeneous database system is an inter-industrial management platform to configure source integration, which is time-saving, cost-saving, hardware saving, energy and cooling saving and improves resource utilization.

Virtualization platform, with enhanced virtualization and data management , is the solution to barriers of data center construction, such as highly heterogeneous facilities, low resource utilization rate, difficult maintenance or unsustainability of business, Figure 8 illustrates the virtualization.

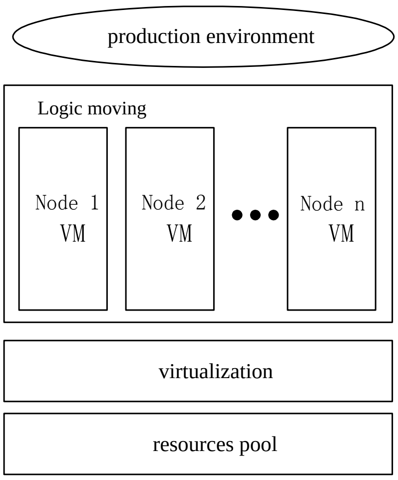


**Figure 8. Virtualization Platform**

## 5. Green Management of Transportation

The transportation in Anhui Province, charged by the Provincial Department, is the object of this study. City and county highway bureaus, and other industry authorities including TMB (transportation management bureau) and MB (marine board), empowered by provincial department, specifically undertake transportation management. Administrative organizations, such as the provincial department and city/county highway bureaus, process administration data and some comprehensive result sets while industry authorities process business data. The informalized transportation system consists of information platforms of highways and freeways management, road transport management and maritime management, sequentially civil aviation management and postal management get involved.

The transportation data center collects business data, data in all files or submitted manually, and industrial data. Level I data center additionally collects network-based information, road ancillary information, road maintenance information and emergency information. All data are on the data exchange platform.

Applications in the provincial data center are data analysis and presentation of transportation in Anhui, and various business applications.

The application of transportation data analysis and presentation result from data in ODS, business data of services and data of ESB. The other application result from data of ESB and ETL.

ESB submits data to applications of transportation rather than other industries. Provincial transportation application collects ESB data for data analysis and presentation. Chart 8 illustrates the application allows ESB data accessible to data center.

To meet the transportation needs, city/county bureaus set Level II data center, respectively indicating sub regional center and sub business center. The two-level database vital to industrial integration resolves logistics and emergency with green data.

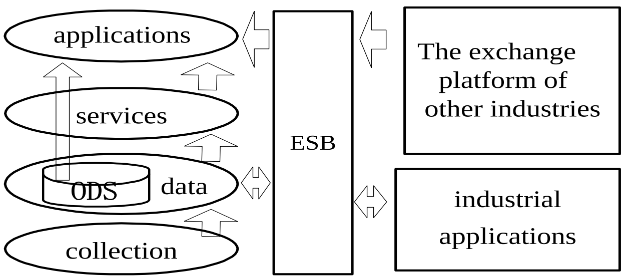


**Figure 9. Provincial Transportation Data Center (Level I Data Center)**

## 6. Conclusions

Green data centers are a complex of information engineering and industrial development. Among hardware, software, application and structure, priority to any single factor invites disadvantage rather than progress, which decides industrial informalization integration. We defined green database management accessible to industrial integration and suggested a two-level model of efficient procedures available to services and management. Evaluation criteria, such as the optimal tradeoff between energy consumption and efficiency, are potential to be precise [22]. Advanced techniques such as sensing, transmission and application, requires elaboration of data centers, a major section of IoT (internet of things).

## References

[1] H. Angang, "Global climate change and green development in China", Journal of the Party School of the Central Committee of the C.P.C, vol. 14, no. 2, **(2010)**, pp. 5-10.

[2] W. X. Lin, C. Ming, F. J. Hua, H. G. Min and L. Z. Yi, "Architecture of the Data Center Network", Journal of Software, vol. 24, no. 2, **(2013)**, pp. 295-316.

[3] M. Jackson, J. J. Koomey, B. Nordman and M. Blazek, "Data Center Power Requirements: Measurements from Silicon Valley", Energy–The International Journal, vol. 28, no. 8, **(2003)**, pp. 837-850.

[4] G. Grid, "Green grid metrics: Describing Datacenter Power Efficiency of High-density and Low-density Spaces", THERMES, Santa Fe, NM, **(2007)**.

[5] M. Blazek, H. Chong, W. Loh and J. G. Koomey, "Data Centers Revisited: Assessment of the Energy Impact of Retrofits and Technology Trends in a High-Density Computing Facility", Journal of Infrastructure Systems, vol. 10, no. 3, **(2004)**, pp. 98-104.

 

[6] Ceuppens L., Sardella A. and Kharitonov D., "Power saving strategies and technologies in network equipment opportunities and challenges, risk and rewards", Proceedings of the 2008 International Symposium on Applications and the Internet (SAINT. 08). Turku, Finland, **(2008)**, pp. 381-384.
[7] Gandhi A., Gupta V., Harchol B. M and Kozuch M. A., "Optimality analysis of energy-performance trade- off for server farm management", Performance Evaluation, vol. 67, no. 11, **(2010)**, pp. 1155-1171.
[8] P. Mahadevan, P. Sharma, S. Banerjee and P. Ranganathan, "A power benchmarking framework for network devices", Proceedings of the 8th International IFIP Networking Conference (NETWORKING. 09). Aachen, Germany, **(2009)**, pp. 795-808.
[9] M. Scofield, P. E. Thomas and S. Weaver, "P.E. Using Wet-Bulb Economizers Data Center Cooling", ASHRAE Journal, August **(2008)**.
[10] K. C. Karki, A. Radmehr and S. V. Patankar, "Use of Computational Fluid Dynamics for Calculating Flow Rates Through Perforated Tiles in Raised-Floor Data Centers", International Journal of Heating, Ventilation, Air-Conditioning, and Refrigeration Research, vol. 9, no. 2, **(2003)**, pp. 153-166.
[11] C. D. Patel, R. Sharma, C. E. Bash and A. Beitelmal, "Thermal Considerations in Cooling Large Scale High Compute Density Data Centers", Proceedings of Itherm2002, Eighth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, San Diego, May **(2002)**.
[12] C. Cheng, H. Bingsheng and T. Xueyan, "Green databases through integration of renewable energy", Proceeding of CIDR, **(2013)**.
[13] D. Wei, L. F. Ming, J. Hai and L. Dan, "Leveraging Renewable Energy in Cloud Computing Datacenters: State of the Art and Future Research", Chinese Journal of Computers, no. 3, **(2013)**, pp. 582–598.
[14] Bianzino A. P., Chaudet C., Rossi D. and Rougier J. L., "A survey of green networking research", Proceedings of IEEE Communications Surveys & Tutorials, no. 99, **(2010)**, pp. 1- 18.
[15] Hewlett-P. I., "Advanced configuration and power interface specification", ACPI Specification Document, Revision4.0a, April 5, **(2010)**.
[16] Andrews M., Anta A. F., Zhang L. and Z. Wenbo, "Routing for energy minimization in the speed scaling model", Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM'10). San Diego, USA, **(2010)**, pp. 1-9.
[17] Shang Y., Li D. and Xu M., "Energy-aware routing in data center network", Proceedings of the 1st ACM SIGCOMM Workshop on Green Networking. New York, NY, USA, **(2010)**, pp. 1-8.
[18] Rao L., Liu X. and Liu W., "Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment", Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM'10). San Diego, U SA, **(2010)**, pp. 1-9.
[19] D. Hui, Q. Yong, W. P. Jian and Z. K. Yu. "Workload Scheduling Algorithm for Minimizing Electricity Bills of Green Data Centers", Journal of Software, vol. 25, no. 7, **(2014)**, pp. 1448-1458.
[20] Y, K. Jiang, W. Z. Hui, J. X. Hong, H. Q. Ming, "Chinese Journal of Computers", vol. 35, no. 6, **(2012)**, pp. 1262-1285.
[21] "Integration of clouds and virtualization of clouds and virtualization into the European production infrastructure", https://documents.egi.eu/public/RetrieveFile?docid=258&version=8&filename=EGI-D2.6-258-final.pdf.
[22] L. Chuang, T. Yuan and Y. Min, "Green Network and Green Evaluation: Mechanism", Modeling and Evaluation, Chinese Journal of Computers, vol. 34, no. 4, **(2011)**, pp. 593-612.