We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Cloud Computing

Chien Wen Hung

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.77283

Abstract

Cloud computing was a cloud technology pioneered by Amazon for a long time due to its software technology that is based on the online shopping platform. After Google, Microsoft also follow up, and this technology, in fact, already exists in our lives, and applications continue to expand, become an integral part of life. With the rapid development of the Internet and the demand for high-speed computing of mobile devices, the simplest cloud computing technology has been widely used in online services, such as "search engine, webmail," and so on. Users can get a lot of information by simply entering a simple instruction. Further cloud computing is not only for data search and analysis function, but also can be used in the biological sciences, such as: analysis of cancer cells, analysis of DNA structure, gene mapping sequencing; in the future more Smart phone, GPS and other mobile devices through the cloud computing to develop more application service.

Keywords: cloud technology, smart phone, GPS

1. Introduction

As early as 1983, Sun Computer proposed the concept of "network as a computer", opening up the direction of thinking and development.

In 2006, Amazon introduced "resilient cloud services" and decentralized architecture technologies to provide limited-service web services.

In 2006, Eric Schmidt, Google's chief executive, put forward the concept of "cloud computing," laying another new era in computer development.

In 2007, Google and IBM, in cooperation with prestigious universities in the United States, started to develop "Cloud Services" software and hardware technology on campus and provided school professors and students to develop large-scale research projects on the Internet.

In 2008, Yahoo, Hewlett-Packard, Intel and the United States, Germany and Singapore jointly launched a large-scale research and development platform for cloud computing to build 6 data center research centers. On average, each data center is equipped with 2500 processors and is actively developing cloud service technologies.

In 2008, Dell officially applied to the U.S. Patent and Trademark Office for a "cloud computing" patent application. In the meantime, large names such as Fujitsu, Red Hat, Hewlett Packard, IBM, VMware, and NetApp compete in R & D.

In 2010, NASA teamed up with major computer vendors such as Rackspace, AMD, Intel, Dell and Microsoft to develop cloud computing technologies.

The meaning of cloud computing is to store the data stored in the local computer and store it in the cloud website. The calculation is made by the local computer and handed over to the cloud computing website. Users do not need to worry about hardware devices, system installation, applications, just open the cloud page, you can perform various types of data storage and computing.

The basic characteristic of cloud computing is "computing in the cloud", that is, building a large-scale data center by combining multiple useful websites to satisfy any data storage and problem computing; meanwhile, users are not required to worry about their own hardware and software facilities. Cloud site consider all possible troubles. As long as the user open the web page, send information to complete the operation.

Cloud computing can be described as "network computer", make full use of the Internet function to connect multiple useful sites, the formation of cloud sites, providing users with data storage and problem computing; users no longer worry about local storage devices and computing applications, do not have to worry about computer professional knowledge, through the Internet to connect to the cloud site, you can send data to the cloud storage page, the cloud application can solve the problem.

2. Cloud computing features

The basic characteristics of cloud computing is "computing in the cloud," that is, to meet:

1. Multiple large-scale data centers and a large number of processors: the combination of a number of useful websites and a large number of large-scale data centers and a large number of processors to meet any data storage and problem computing.

2. Cloud service: users do not need to worry about hardware devices, no troubles to install the system, no need to worry about applications, cloud sites consider all possible troubles, design execution web pages, users can simply open web pages to store data, computing data, delivery data.

If you want to create an all-encompassing cloud site, in some functional environment may be a waste of time, in order to properly apply, we can cloud site as:

1. Public cloud: public cloud is the most common way to deploy cloud computing. Cloud resources are owned and operated by third-party cloud service providers and delivered over the Internet. In the public cloud, all hardware, software, and other supported infrastructure is owned and managed by the cloud provider. In the public cloud, you share the same hardware, storage, and networking devices with other organizations or cloud rental users. You use a web browser to access services and manage your account. Common cloud deployments are commonly used to provide web-based email, online office applications, storage, and test and development environments.

    Public cloud benefits:

    a. Low cost—no need to purchase hardware or software, just pay for the services you use.

    b. No maintenance—maintenance is provided by service providers.

    c. Scalability—unlimited on demand to meet your business needs.

    d. High reliability—a wide range of server networks to ensure that from failure.

2. Private cloud: private clouds are comprised of cloud computing resources that are used exclusively by a single organization or organization. Private clouds can be physically located within the organization's local data center or hosted by third-party service providers. But in the private cloud, services and infrastructure are all maintained on the private network, with hardware and software exclusive to your organization. In this way, the private cloud enables organizations to more easily customize their resources to meet specific IT needs. Private clouds are widely adopted by government agencies, financial institutions, and other medium- and large-scale organizations that need to handle business-critical operations in pursuit of better control of their environment.

    The advantages of private cloud:

    a. More flexibility—your organization can customize the cloud environment to meet specific business needs.

    b. Improved security—no sharing of resources with other people, resulting in greater levels of control and security.

    c. High scalability—private clouds can still afford the scalability and efficiency of public clouds.

3. Hybrid cloud: combining both private and public clouds gives organizations the best of both worlds. In hybrid cloud, data and applications move between private and public clouds, making your enterprise more resilient and have more deployment options. For example, you can use a large number of projects with low security requirements (such as web-based email) for the public cloud, and use confidential, business-critical items such as

financial reports for private cloud (or other on-premises basic structure). In hybrid cloud, you can also choose Cloud Load Balancing. This is when an application or resource is executing in a private cloud until demand increases (such as seasonal activities such as online shopping or tax returns), during which time an organization can "load balance" the public cloud to apply other operations Resources.

The advantages of hybrid cloud:

a. Control—your organization can maintain the private infrastructure of confidential assets.

b. Resiliency—you can take full advantage of other resources in the public cloud, if needed.

c. Cost-effective—you can take advantage of changes to the public cloud to pay for additional computing power only when you need it.

d. Easily-switch to the cloud with less effort because you can move in and out-gradually introducing workload over time.

3. Cloud computing advantages and disadvantages

Cloud computing advantages [4].

1. In R & D and maintenance, because of the huge resources, the solution to the problem can be rapidly established and deployed to reduce the barriers of information problems and save the R & D and maintenance costs of new methods.

2. In terms of exchanging messages, users find problems and solve problems. Because cloud terminals are hubs, users can easily communicate with each other and exchange views with each other so that they can collaborate to solve problems and give play to their wisdom and ability.

3. In terms of slim and light development, users do not need to have large memory capacity and powerful computing functions, such as tablets and mobile phones, because of data storage and functional computing in the cloud.

4. In the system update, due to functional hardware and software are in the cloud site, where there is a newly developed system software, as long as the site updated, immediately put into use, users no longer have to worry about update and installation problems.

5. In terms of functional applications, it is easy to meet the needs of users by combining multiple useful websites with diverse functional backgrounds, providing highly competitive application functions and enhancing users' application capabilities and vendor reputation.

6. In terms of ease of use, cloud computing is a fast service, and users can operate anytime, anywhere as long as they are in an Internet-enabled area and open a webpage.

Cloud computing disadvantages [4].

1. There must be network connectivity between the cloud website and the user. Therefore, users in the area without the network cannot share any cloud function. If the network speed is slow, it will also affect the efficiency of the function and cannot handle difficult problems.

2. In order to build a cloud website, in addition to requiring huge funding, more personnel and equipment are needed. In the future, maintenance requires funds and talents. So running a cloud site has a very heavy burden on it.

3. Cloud computing is a new era of computer technology, and we are now using computer systems and methods, but also after a long time to build a little bit by bit, if you immediately give up straight to the cloud, will waste the previous investment; if not updated system will not keep up with new technology.

4. For users, participation in cloud computing can make the operation easy and effective, but it may also be a commercial trap. Once it enters the cloud, it will abandon its own capabilities and devices if it relies on the cloud. If the cloud system aggravates the payment, at this time can only be allowed to ask.

4. Cloud application status

Cloud computing industry can be divided into three categories: cloud software, cloud platform, cloud devices.

1. Cloud software (SaaS): to break the monopoly of the previous situation of manufacturers design, interested parties can develop their own design, propose a wide range of software services.

2. Cloud platform (PaaS): research and develop operating system platform, provide software developers to design cloud software and serve the general public via the Internet. As a result, operating system platforms and non-manufacturers with ample human and material resources cannot afford to participate. Currently, there are: Google, Yahoo!, Microsoft, Apple.

3. Cloud equipment (IaaS): the basic equipment (such as IT systems, databases, etc.) are systematically integrated to make it work together to provide the maximum storage space for data and provide the fastest execution time for the operation. The current participants: IBM, Dell, Sun, Hewlett-Packard, Amazon.

Enterprises in cloud service environment providing customized customer service have become popular. However, quick and proper understanding of customer needs to provide customized services should be a priority of companies. Therefore, the proposed customer service for customized applications and the use of data mining techniques to collect information from a large number of promotional products to meet customer needs will result in proactive customized information to the customer to save time searching for products. The system can enhance the competitiveness of enterprises and increase corporate profits.

1. SaaS is a model that provides software through the Internet. Instead of purchasing software, users can use the Web-based software to manage their business activities without having to maintain the software. The service provider manages and maintains the software, for many small businesses, SaaS is the best way to adopt advanced technology, which eliminates the need for companies to buy, build and maintain infrastructure and applications. In recent years, the rise of SaaS has given traditional packaged software vendors real pressure.

2. PaaS is a platform as a service (Platform as a service). Platform as a service is a cloud computing service that provides a computing platform and a solution stack as a service. At the typical level of cloud computing, the platform as a service layer is between software as a service and infrastructure as a service.

   Platform-as-a-service provides the ability for users to deploy and create cloud infrastructure to clients, or to use programming languages, libraries, and services. Users do not need to manage and control the cloud infrastructure, including the network, servers, operating systems or storage, but need to control the upper application deployment and application hosting environment.

   PaaS takes the software development platform as a service and delivers it to the user as a softwareas-a-service (SaaS) model. Therefore, PaaS is also an application of the SaaS model. However, the emergence of PaaS can accelerate the development of SaaS, especially to accelerate the development of SaaS application speed.

3. Infrastructure as a service (IaaS) is the software that consumers use to process, store, network and various basic computing resources, deploy and execute operating systems or applications, and so on. Clients can deploy and run processing, storage, networking, and other basic computing resources at will, without the need to purchase network devices such as servers and software. They cannot control or control the underlying infrastructure, but can control operating systems, storage devices, deployed applications.

5. Cloud platform

5.1. Google Cloud platform

Google (Google) to develop Gmail, Google Docs, Google Talk, iGoogle, Google Calendar and other online applications, the establishment of basic cloud computing platform. General users use the browser to connect to the designated website platform, you can edit the file, and then online archive. In the company did not complete the file, go home from work can be connected to the Internet to continue, Google Spreadsheet graphical online spreadsheets can be defined formula fill in the numerical calculation, Google Cloud Computing website, these work has nothing to do with the performance of the computer we use Only the internet connection speed is a problem.

In order for cloud sites to process large numbers of users and large amounts of information in parallel. The Google Cloud assumes that every information system can fail at any time, so use software layers to create fault tolerance and standardize machines. As the amount of data increases, the performance of the cloud system can only be achieved by continuously expanding

the number of machines and equipment without modifying the original application. For example, the software layer includes three technologies: Google File System (GFS), distributed database Google BigTable (GBT), and Google MapReduce (GMR).

Google has been successfully developed and welcomed by users of the cloud site:

a. Gmail: each account provides 15 GB of high-capacity storage space, effectively control spam and provide normal operation assurance and security.

b. Google Calendar: a web-based calendar application that increases the productivity of individuals or groups of users, helps reduce the cost of work, and enhances the division of labor. Agenda management, scheduling, sharing of online calendar and calendar synchronization move.

c. Google Docs: provides the execution environment for word processing, trial execution, briefing and processing at anytime and anywhere, supports up to 1G unlimited file types upload (virtual hard disk) and instant sharing. Through the webpage, users are provided multiplayer editing files at the same time.

d. Google Talk: groups effectively interact and communicate, provide group mail communication, facilitate content sharing, and quickly search for files. Share calendars, documents, websites and videos.

e. Google Group: the joint vertical and horizontal integration of the intranet, the establishment of a safe and effective community team project.

f. Security: provide the best security solutions such as video, private files.

5.2. Yahoo Cloud platform

Yahoo! successfully developed four major technologies for cloud infrastructure:

a. Establishing cloud storage for structured and unstructured data.

b. Establishing large-scale decentralized data computing and storage.

c. Providing cloud data cache and proxy Function.

d. To provide advanced rapid data processing services.

The ultimate goal is to complete Online Serving, so that developers can complete the development environment online to establish and accelerate product and service development time.

Working with the Apache Software Foundation to develop Hadoop, a cloud-based operating system, Hadoop is a distributed computing environment written in Java that provides a wealth of data. Hadoop's architecture is based on the concept of BigTable and Google File System It is similar to the cloud computing architecture used internally by Google.

5.3. Microsoft Cloud platform

On the cloud computing platform, Microsoft has developed the most complete applications, including:

a. Cloud computing service application (Windows Azure, SQL Azure) in the Internet data center.

b. Enterprise online cloud service application (Microsoft Online Services).

c. Enterprise Server (System Center), providing customers the freedom to choose their own solution or mix different solutions.

In order to lead Google, Yahoo! and other existing cloud platforms, Microsoft has also released a self-developed cloud computing platform Azure Services Platform, using the operating system Windows Azure.

5.4. Apple cloud platform

The Apple cloud platform is slightly different from the previous three platforms. When used by users, they must be downloaded to the local user device before being turned on.

Apple's approach is not to treat the cloud as a platform to solve all issues. Instead, the cloud is viewed as a central monitoring station to monitor the user's operation. The main reasons why Apple handles cloud services in this way are as follows:

a. Apple does not trust the quality of current network data delivery, especially the streaming quality provided by mobile phone carriers.

b. Apple does not participate in other systems to share their achievements, as far as possible limited to consumers using Apple devices only to use this service.

c. Apple to reduce replay functionality load, allowing users to play through the cloud, but also if necessary Device download, replay at any time.

Apple cloud platform operating system to cloud data flow rather than control the organization of information transfer. Apple cloud platform operating system applications, music, media, files, messages, photos, backups, settings and other centralized storage in the cloud. iCloud supports all iOS devices. When Apple users use this system to upload files, iCloud automatically backs up the purchased music, applications, files, photos and system devices to other cloud devices and synchronizes them to other Apple devices.

6. Cloud computing security

According to the Cloud Security Alliance (CSA), cloud computing services may encounter seven major information security issues:

1. Unscrupulous people using cloud computing technology to engage in unscrupulous resource services.

2. User operations Interface and cloud computing services with information security concerns.

3. Insiders interested parties, the use of illegal ways to get the cloud client's resource content.

4. Cloud data sharing may cause resource sharing error or interference situation.

5. When the data are all concentrated in the cloud resources platform, may lead to data leakage problems.

6. Cloud client account and password authentication information was intentional tampering.

Information security threats, in addition to man-made attacks, include information system attacks and internal staff operations. There are three steps in protection measures that must be paid special attention to:

1. To strengthen the security protection and setting of hardware and software.

2. Enhance the information system security monitoring and internal operation auditing mechanism.

3. Timely adjustment of information system settings and data backup.

After the cloud computing was put forward, the network attacker also continued to devote himself to attacking the virtual machine. For the self-protection of cloud computing, the following information security recommendations were made:

1. Do not trust the network is a guarantee always safe and secure.

2. Login information on the Internet too detailed, easy to leak personal privacy issues.

3. Whether the data will affect the normal work affairs when the data is placed on the cloud platform.

7. Cloud system virtual technology and pivotal technology

In the cloud system, the most basic software technology is "Virtual Machine" technology. Virtual Machine, or VM for short, is to simulate the host operating system. In recent years, VM technology has been gradually used to simulate different operating systems to support various host needs of cloud users. This is not only the Platform as a Service, the operating system virtualization into network services) core essence of technology. For example, we can host different versions of the Linux operating system, FreeBSD operating system and MS-Windows operating system in the cloud system. The cloud system users can be based on their needs, were connected to the cloud host MS-SQL database services or MySQL database services. For cloud system maintainers, only one cloud host needs to be maintained. However, cloud users do not need to care about the real operating system that the cloud host executes, as long as they can access the data services they need. In this way, the cloud-based SaaS can be completed and the application functions can be virtualized into network services the core technology.

Cloud server virtual host can be roughly divided into two types of server host and client host, server host which is the real hardware computer operating system to host the implementation of the virtual host operating system; and the client host is virtual out of the host system. As long as the hardware capacity permits (memory capacity and hard disk capacity), each server host can execute many different client hosts, each Guest can represent a host, and has an independent IP address and different network Road service. Of course, a more complex cloud architecture can be used as the host of the Guest host, and then host the implementation of different versions of the virtual host operating system, and thus form a "virtual host inside the virtual host." Similar to the nested virtual host architecture (Nest Virtual Machine).

7.1. Pivotal technology—Google File System (GFS)

Google File System (GFS), mainly to deal with the rapid increase in cloud computing data. GFS has all the features of a distributed file system, including storage efficiency, scalability, reliability and reusability, large GFS distributed file system can be composed of hundreds of hard drives, without the use of high-end expensive Storage equipment can maintain the file storage quality. And with fault-tolerant capabilities, GFS easily recovers corrupted files through fault-tolerant detection and auto-recovery in GFS, even in the event of an operation.

When the GFS client client application requests a GFS host server to create a file request, the GFS host server will cut the request into blocks of 64 MB in size, which are then allocated to the lower layer for processing. The file system in order to ensure the safety of the file will automatically copy the block data to at least three backup actions and save them to the hard disk in the ext2 format. Finally, the result of the reassembly process is delivered to the GFS client application, which is the GFS operational flow [1]. As shown in **Figure 1**.

GFS operation process, GFS is mainly used to store metadata, the main namespace file, the partition namespace partition and each file is mapped to the block record location, usually about 64 MB in size, the average of each file only 100 characters, and the real data is stored on the server. In this way, you do not need to worry about the risk of inconsistent backup data of the three lower copies if the server is damaged in an unexpected way, and administrators can also make backups via the remote mechanism so as to be more secure.
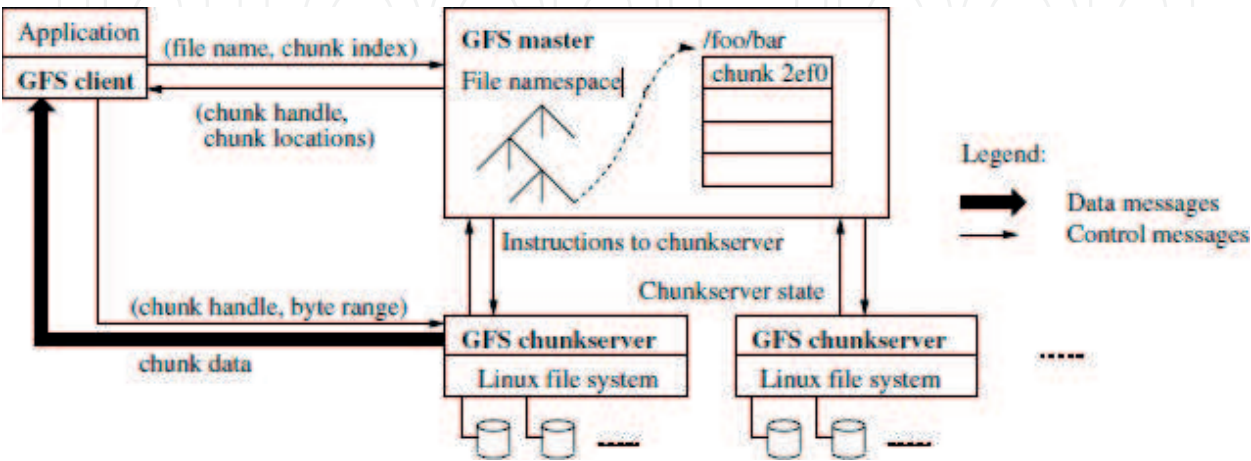

Figure 1. Google File System (GFS) [1].

In addition, 64 MB huge block space has several advantages. First, he reduces the need for communication between the client and the server because the same block space can be processed with only one read for immediate information, and the block storage design can reduce the workload of the search service. Second, because the block space is large, the client usually only needs to operate multiple times in the same block space, which can reduce the number of server switches and reduce the load of network traffic. Third, large block storage space can also reduce the amount of metadata that the main server needs to store. Because of the reduced processing load, you can speed up the main server's connection by placing frequently read metadata on the main server's memory instead of on the hard disk.

7.2. MapReduce

MapReduce is actually a simple programming model. As long as the map and reduce functions are used in programming, the information system helps map to sort out the available data from the original data and classify the available data, and then use the reduce program to simplify the usable data.

The purpose of MapReduce is to run large-scale computer data and the implementation of decentralized computing, for a large number of data to do parallel computing. So the whole structure of MapReduce is composed of two functions Map and Reduce. When the program inputs a large group of key or value, the Map function will automatically disassemble the key or value of many groups, and then Reduce Function simplifies the data content and then merges with the same Key value of the pairing, and finally produce the analysis of the data results [2]. As shown in **Figure 2**.
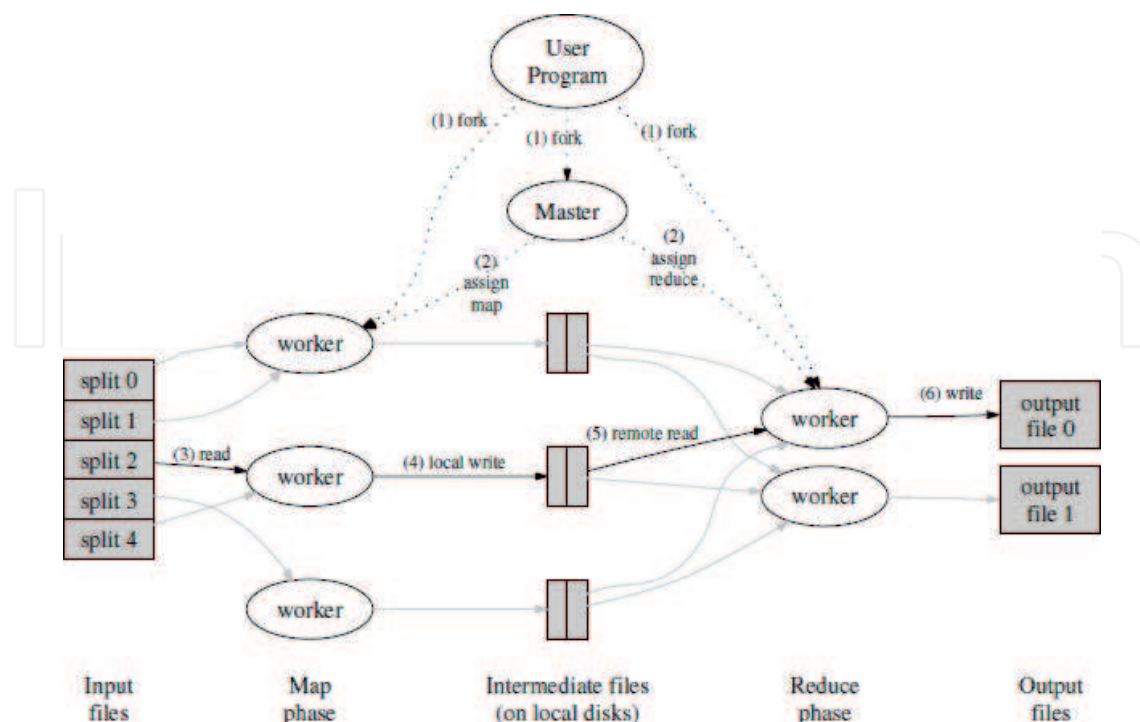

Figure 2. MapReduce [2].

7.3. BigTable

BigTable distributed database query data used in the language of Google Query Language, is a database similar to the SQL command language. Because BigTable can perform complicated analysis and query functions with MapReduce technology architecture, it is especially suitable for integrated data table access up to 10 TB or more.

BigTable storage for the column-oriented, and the traditional relational database row-oriented way, the benefits of the use of column-oriented is very convenient in the new data, each data can be stored in a fixed field of objects. In addition BigTable index of information is divided into two kinds of row key and column key, and can be any string, so more than the traditional relevance of the database with high compression ratio, high read performance.

BigTable the first field is the Row key, the second field is the Column key, the third field is the timestamp as a data index, the three parameters corresponding to the data is stored as a string .

Each basic unit of storage stored in a bigTable has Timstamp, which allows multiple versions of the same stored data over time. And users of the cloud application can specify which data to keep. Since the data type of the timestamp is int64, each storage unit can use the difference of one millionth of a second for storing, so as to completely prevent data from being overwritten [3]. As shown in **Figure 3**.
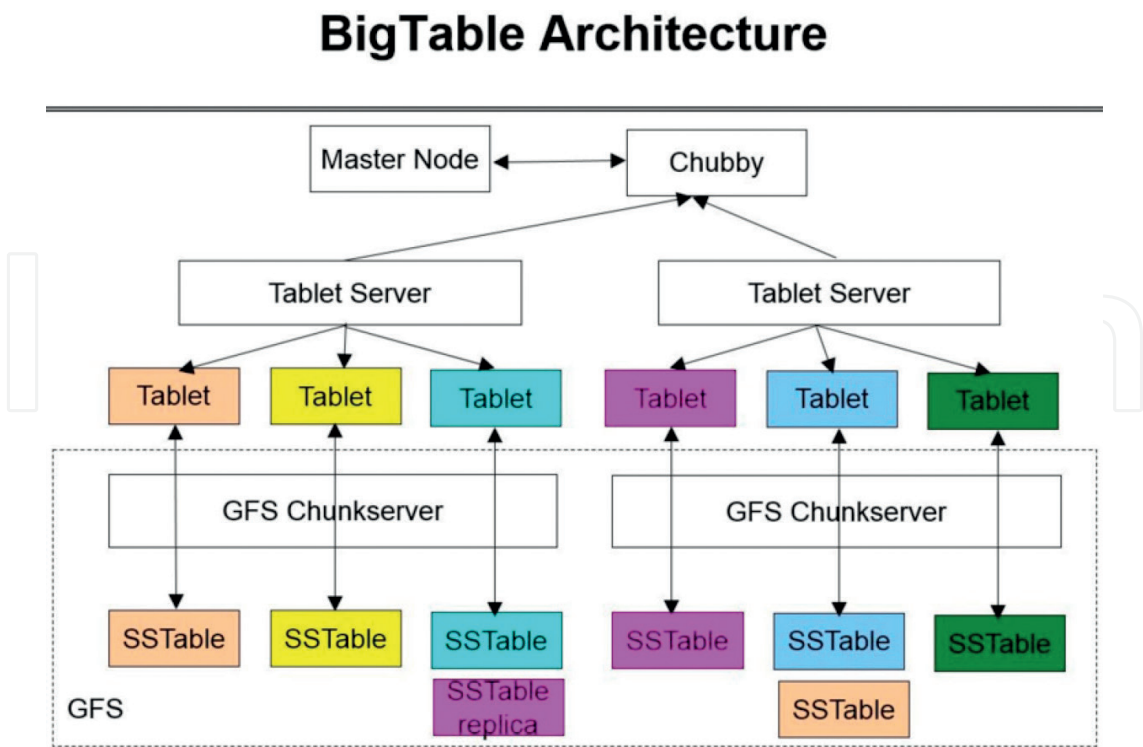

Figure 3. BigTable [3].

8. Case examples

This case example uses PaaS architecture to collect customer consumption data and to analyze the interactions of purchased goods. This study aims to construct a cloud service customized product selection information system to provide a reference to the industry.

The customized products framework of the system is shown in **Figure 4**, in which the relational database management system is used to conduct the data mining which consists of three steps:

The Customized products cloud Information System is shown in **Figure 4**, where the relational database management system is used for data mining that consists of four steps:

Step 1: input the system code by certification and open the databases, which consist of the customer database, the retail mall database, and the transaction database.

Step 2: analyze the data according to the database.

Step 3: transfer the Marketing Knowledge Database to the relevant branches and to the customers.

8.1. Association rule mining

The association rules are proposed by Agrawal and Srikant [5]. Association rule mining is widely used for analyzing the product items purchased by consumers. It is also used to support sales promotion and marketing segmentation. The association rule is represented by X→Y where X and Y are a set of items. This rule means that the transaction records in a business database that contain X tend to contain Y. A large number of valid algorithms for mining association rules have been proposed [7].
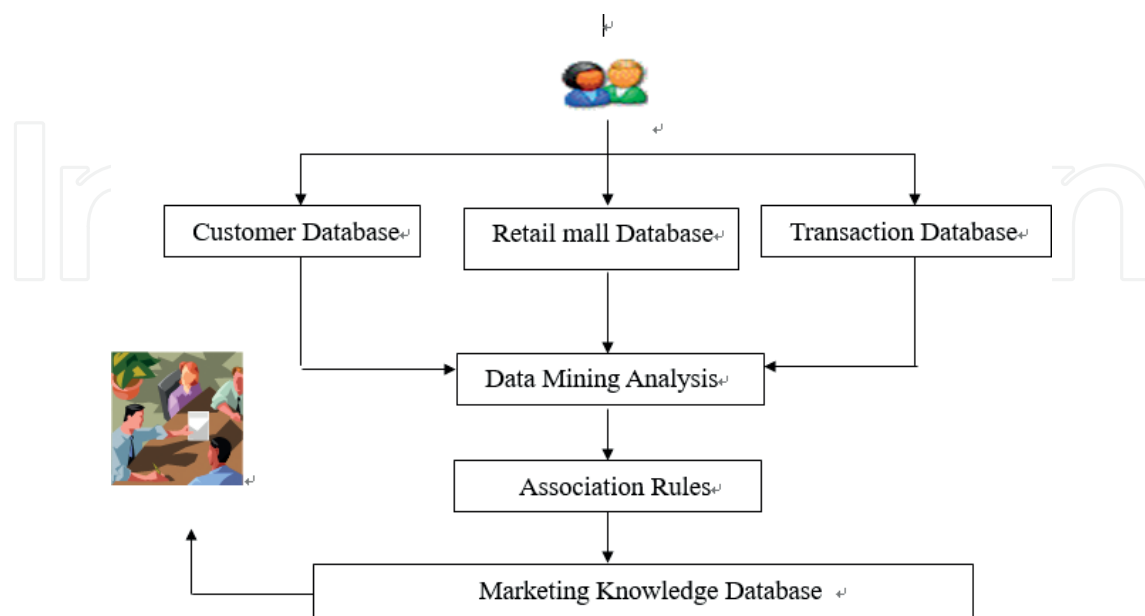

Figure 4. The customized products customized products cloud information system.

In this study, a mining system to detect customer behavior is proposed. The association rules from relational database design are utilized to mine consumer behavior.

In the consumer purchase of computers and memory, for example, the association rules are as follows:

Computer → Memory [support = 25%, confidence = 75%].

The formula shows that 25% of the entire transaction database will buy a computer and memory, while 75% of the total customers who purchase computers will buy them together with memory. The following steps are used to determine the association rules:

1. First, find a collection of high-frequency items (large item set). This collection of support must be greater than the user customized minimum support (minimum support).

2. Second, use a collection of items produced by high-frequency generating association rules.

Currently, many algorithms can identify the high-frequency items associated with a collection of rules, such as Apriori [6] and DHP [8]. Apriori is the most commonly used and best-known algorithm; therefore, its data analysis will be used in this study.

8.2. Customized products analysis

From our interviews with the company marketing managers, the customized products design is the fundamental promotional tool for the firm. The customized products is designed by the marketing office based on decisions from department meetings. For example, more customized products is dedicated to facial care products and decorations products.

The customized products is designed and produced by the marketing office based on aggregated information from head office. The data was collected from June 2015 to May 2016. The database of the system consists of three major parts, namely, customer data, product data, and transaction data.

In this chapter, we propose a cloud service information system prototype using data mining techniques to help enterprises find suitable promotional products for each customer in the cloud customer database.

Step 1: create a basic cloud customer database.

The main characteristic attribute in the collection of user data. It records all possible influence factors of the customer buying behavior attribute data. The cloud customer database is composed of two parts, namely, basic information and preference category. A1 to Ai-1 are the basic properties of the cloud customer information, such as gender, education, salary, and others. Ai records the user's preferences for product categories, such as facial cleanser, Shampoo, and so on.

Step 2: cloud pre-processing of customer databases and program code conversion.

This step is mainly for the retrieval of the required fields from the cloud customer database. Then, the cloud customer attribute data are matched with the users to do the coding to attribute on the clustering.

Step 3: association rules.

The first level of association rules uses the product department as decision variables, For example, we let pc, fc, hc, sb, denote the Personal cleaning products department, the Facial care products department, the Home cleaning products department and the Snack biscuits products department, respectively. The equation $P_P(pc \cap fc = Type1 | T)$ denotes the probability that the customers will buy products from the Personal cleaning products department and Facial care products department; where type1 and T denote the product combination type1 and customer, respectively. The detailed results are shown in **Table 1**.

The decision variable in the second level association rules is product brand. Herein, three product brands are sold by the Facial care products department (i.e., NIVEA Facial cleanser coded by fc3, Biore Facial cleanser coded by fc1, and Deep Clean Facial cleanser coded by fc2), first product brand is sold by the Home cleaning products department (i.e., Tide Laundry detergent coded by hc2) and second product brands are sold by the Personal cleaning products department (i.e., Biore Body Wash coded by pc2). The equation $P(pc2 \cap fc3 = Type1)$ denotes the probability that the customers will buy pc2 and fc3 simultaneously. Please see **Tables 2** and **3** for more detailed results.

8.3. Marketing knowledge database

According to the mining results, the promotion projects are to design some special customer catalogs for the customer in accordance with the customer's preference for the product brand to enhance product sales.

Product classification as decision variable (step 1)		
Product combination code	Product combination (%)	rank
$pc \cap fc$	56.2	1
$pc \cap hc$	38.6	2
$fc \cap sb$	25.8	3
$fc \cap hc$	16.5	4
$hc \cap sb$	9.8	5

Table 1. Results of section classification as the decision variable (first level).

Product brand as decision variable (step 2)		
Product combination code	Product combination (%)	rank
$pc1 \cap fc3$	82.1	1
$pc2 \cap hc1$	71.5	2
$fc2 \cap sb5$	62.8	3
$fc2 \cap hc3$	53.6	4
$hc6 \cap sb2$	46.3	5

Table 2. Results of product brand (second level).

Decision variable	Combined name of products	Rank
Personal cleaning products department And Facial care products department	a Biore Body Wash paired with a Biore Facial cleanser (85.6%)	1
Personal cleaning products department And Home cleaning products department	a Biore Body Wash paired with a Tide Laundry detergent (72.6%)	2

Table 3. Results of product brand combine (third level).

Figure 5. (Biore Body Wash with and Biore Facial cleanser).

Herein, the special customer promotional products are designed following special customer promotional products 1 and 2 and 3 (SCPP 1 and 2 and 3).

SCPP 1: combine the sale, offering a 15% discount to all customers, of Biore Body Wash with and Biore Facial cleanser, which rank No. 1 in product sales of **Tables 2** and **3** respectively. **Figure 5** shows the description in the special customer promotional products 1.

SCPP 2: combine the sale of Biore Body Wash and tide laundry detergent, which rank No. 2 in product sales of **Tables 2** and **3** respectively, offering a 20% discount to all customers. **Figure 6** shows the description in the special customer promotional products 2.

SCPP 3: combine the sale of Biore Body Wash with the own brand products, offering a 25% discount to all customers. **Figure 7** shows the description in the special customer promotional products 3.

Figure 6. (Biore Body Wash and tide laundry detergent).

Figure 7. (Biore Body Wash with the own brand products).

Therefore, when these three indicators by policy makers are found, the needs of individual customer's promotional products shall be considered in the current business situation. The main objective of the promotion is to determine the appropriate weight of each index. The system helps enterprises to efficiently deliver the customized commodity product promotion, as well as really meet the business situation and needs.

9. Conclusion

In this study, association rules were used to identify links between customer profiles and products purchased. It provides marketing managers with a useful tool to rapidly search for

valuable information based on customer transaction cloud information, and rapidly establish marketing strategies to enhance sales and profit.

This study provides a cloud service information system that improves the sales of case companies' products by changing the original promotion methods. It can help supervisors and employees to provide them with useful knowledge through the new system presented here to better decide on promotional activities.

After marketing the customized product promotion catalog, the data mining system helps companies understand the market acceptance of the product and understand the product sales, thus repairing its marketing strategy and increasing the sales of the product.

In addition, the data mining system can also amend the design of the customized product catalog based on market sales data and design a customized product catalog for different periods. Therefore, the effective use of data mining systems can help understand the customer's buying behavior, so that decision makers can formulate optimal policies.

Author details

Chien Wen Hung

Address all correspondence to: cwhong@mail.cnu.edu.tw

Department of Information Management, Chia-Nan University of Pharmacy and Science, Tainan, Taiwan

References

[1] Ghemawat S, Gobioff H, Leung S-T. The Google File System. 2003. http://static.google-usercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/gfs-sosp2003.pdf

[2] Dean J, Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. 2004. https://static.googleusercontent.com/media/research.google.com/zh-TW//archive/mapreduce-osdi04.pdf

[3] Chang F, Dean J, Ghemawat S, Hsieh WC, Wallach DA, Burrows M, Chandra T, Fikes A, Gruber RE. Bigtable: A Distributed Storage System for Structured Data. 2006. https://static.googleusercontent.com/media/research.google.com/zh-TW//archive/bigtable-osdi06.pdf

[4] Elsenpeter RC. Cloud Computing, A Practical Approach. McGraw Hill: Osborne Media; 2009

[5] Agrawal R, Srikant R. Mining sequential pattern. In: Proc. of the 11th International Conference on Data Engineering. Vol. 12. 1995. pp. 3-14

[6] Aaker DA. Measuring brand equity across products and markets. California Management Review. 1996;**38**(3):103-120

[7] Anand SS, Patrick AR, Hughes JG, Bell DA. A data mining methodology for cross-sales. Knowledge-Based Systems. 1998;**10**:449-461

[8] Leonard H, Eduardo LR, Stefan V. A cloud brokerage approach for solving the resource management problem in multi-cloud environments. Computers & Industrial Engineering. 2016;**95**:16-26