

Research on planning and design of private cloud computing data center

Wei Guanghui^{1,a}

¹ChongQing College of Electronic Engineering, computer college, ChongQing, 401331, china
^aqdbiq@163.com

Key words: private cloud data center server network system storage system

Abstract. cloud computing server and storage system play a very important role in the cloud computing data center. The performance of the server and the storage system will directly determine the performance of the entire cloud computing system. From the cloud computing server design principles, considerations, selection, determine the number and the storage system analysis and design of several aspects, the study of the planning and design of cloud computing server. Finally, the results of the program in CloudSim simulation, the results show that the design scheme can achieve a high cost performance, can achieve the requirements of private cloud users to cloud computing service performance.

Introduction

Private cloud computing enterprise data center data center compared with the traditional one, the former can support dynamic and flexible infrastructure, reduce information technology (IT) the complexity of architecture, make all kinds of resources integration and standardization of IT. Given its data security, quality of service, resource management and other advantages, but also can be seamlessly integrated with the existing IT system, so more and more attention by the enterprise [1]. But the traditional enterprise data center because of the lack of unified planning for the deployment of resources, resulting in heterogeneous computing resources and storage resources to effectively integrate and extend, resulting in a waste of resources, management, operation and maintenance difficulties and high costs of malpractice. Although the use of storage area network (SAN) to build a shared storage architecture, establish a unified cluster job management system, create a resource management platform can alleviate the problem of low rate and complexity of management and maintenance and other resources to a certain extent, but it is still difficult to achieve a comprehensive resource sharing, for infrastructure expansion and service the deployment of 2. How to integrate these resources, seamless integration with the existing IT system, system solution for large amount of data, computing task intensive and complex process to provide flexible and efficient becomes an urgent problem to be solved.

In the construction of private cloud, there are a lot of factors need to be considered, in which the private cloud computing server is used in taking into account in [3], because the enterprise data storage and computing will depend on the cloud server, with the continuous development of enterprises, the business will also follow the rapid development and change, and IT the system itself has high complexity and unpredictability, some change is difficult to predict, therefore, to build the company's private cloud must be rational planning of each link to the cloud server system led, to ensure that the cloud computing data center in ensuring the premise of economy, more reliable, safe and efficient operation.

Server planning and design principles and considerations

In the design of the private cloud computing server, it needs to follow some basic principles. These principles mainly include the principle of practicality, the principle of standardization and the best price to price ratio [4]. Medium, practicality is the most important, private cloud computing system is in real-time, reliability and security requirements of processing enterprises to solve the growing data industry, in order to achieve these requirements, build a private cloud environment,

users of the system can meet the practical function is very important. Here, special attention should be paid not only to emphasize the leading and use the latest technology products; the principle of standardization for the purpose of the use of all kinds of products and manufacturers supporting the world, for the enterprise private cloud software installation, hardware update, system upgrades and system data exchange with the outside world is convenient; the best price is between advanced and the applicability of the system to solve the balance between maximizing the private cloud computing enterprise can obtain investment returns and returns.

Private cloud computing data center server system with specific business types, characteristics of the large amount of data, in the choice of the server, the main factors should be considered include computation ability, memory capacity, I/O capacity and expansibility. The operation ability mainly consider the real-time response capability to accept business in business at the peak, to ensure that when the system is busy not to slow memory; real-time data collection, in the server, are required for temporary storage of intermediate results through memory, sufficient memory to improve the performance of the system is the ability to deal with a large number of I/O; the data needs the ability of real-time storage; scalability is an important characteristic in cloud computing system to improve system processing ability, is used to solve the business increasing load shunt, real-time business processing ability.

Planning for the number of servers in the private cloud

In the planning and design of the server, we know that the computing power of the server and I/O is an important criterion for the performance of the server, so the number and size of the server needs to focus on the rules of design.

At present, the comprehensive evaluation system of server performance evaluation method for maximum impact is applied to the Transaction Processing Performance Council(Tpm) is created by dozens of members of the organization, TpmC is testing a single server or cluster system with specific operating system, the ability of processing transactions per minute the characteristics of the case of the database, the higher the value that the processing power of the server is strong. TpmC value calculation principle is to calculate the performance of a single server, that is, to ensure that the work of a single server can meet the needs of the normal operation of the system, the formula as shown in the formula 1:

$$TpmC = \frac{M \times T \times S \times L \times (1 + F)}{C} \quad (1)$$

In Formula 1, M said the system peak processing ability; T represents the number of transactions each transaction completed; S represents the actual business operation compared to the standard preparation of TpmC benchmark operating environment complexity ratio; L computer deviation value and the value of experience published by C; for the host CPU rate, general server CPU utilization should be less than 80%, otherwise it will produce a system bottleneck; F said to the next 5 years of business development for redundancy.

The estimated number of private cloud computing data center server should be combined with some server server vendors are TmpC models to value and specific business system TmpC demand value, the number of needs for the $N_{servers}$ server, the server's TmpC value is $TmpC_{server-one}$, the private cloud computing data center TmpC of total demand for $TmpC_{total}$ computing needs the number of servers as shown in equation 2:

$$N_{servers} = \text{int}\left(\frac{TmpC_{total}}{TmpC_{server-one}}\right) + 1 \quad (2)$$

Server CPU resource planning

Database is an important part of information system government unit, and it is the core of data processing and information management system. To solve the database and storage of computer information processing in the process of large amounts of data to effectively organize the problem,

can reduce the data redundancy, data sharing, data security data security and efficient retrieval of data and processing in database system. For the OLTP business class database system, the server computing ability in general can use Transaction Processing Performance Council (TPC) released the TPC-C formula of the corresponding measure, the database server computing requirements as shown in equation 3:

$$TPC - C = \sum (M1 * M2 / 1 - M3) \quad (3)$$

In Formula 3, M1 per minute business transaction amount, transaction volume ratio of M2 as the standard, is a current online transaction business system is equivalent to the number of standard tpmC transaction value, M2 value should be in the range of 5-15 M3 for the system resource redundancy rate and general 20%-40%. WEB users demand can be in accordance with the international organization according to the SPECweb2005 evaluation standard beta, Web application service performance requirements as shown in equation 4:

$$SPEC = (User_{total} * rate_{online} * NUM_{http}) / (1 - rate_{redundancy}) \quad (4)$$

The SPEC WEB application server SPEC Web2005 value, namely the server to the client's processing ability, $User_{total}$ represents the total number of users, $rate_{online}$ said users online rate, NUM_{http} said the average number of HTTP requests online user initiated $rate_{redundancy}$, expressed as the redundant rate. According to the requirements of the TpmC and the reservation of the system, the computing power and the number of the required CPU can be calculated.

Server network resource planning

Cloud computing large-scale operations, to the traditional network architecture and traditional application deployment are brought to the challenge of [5]. The traditional network planning and design based on high reliable ideas, forming a mesh network redundancy structure complex, structured mesh physical topology to maintain high reliability, fault tolerance and performance has a great advantage, is the general design rules. The traditional network challenges are mainly the following points: the traditional network complexity in the actual operation, management personnel undertake a very tedious work; cloud computing platform under the deployment of virtual machines on the same physical servers, server utilization increased from 20% to 80%, the server port traffic increased dramatically and put forward higher requirements on network performance.

Through the use of cloud computing, the new generation of network in solving the growing giant computing tasks, can be dynamic, flexible, flexible way to achieve the network deployment. In the cloud computing platform, virtual machine migration between physical servers, in order to avoid the virtual machine migration after routing concussion and modify the network planning, the migration is usually carried out in the two layer in the domain, so the cloud computing platform requires a higher performance for the migration, two layer domain more network environment to provide protection. By analyzing the challenges of cloud computing to the traditional network infrastructure, we can deal with it from two aspects:

One is to build a high performance, high reliability network, so as to meet the pressure of cloud computing network;

Two is to build a virtual network to meet the cloud computing in the virtual machine deployment, migration, as well as the security strategy for the implementation of the network proposed flexibility, security requirements.

Cloud storage resource planning

The construction of private cloud computing data center, involves the integration of storage resources, storage platform carrying the enterprise all the important information and data, virtual machine system data storage platform, so the design must have high reliability, security services 7 * 24 hours; secondly, the storage platform must have high performance high pressure, fully meet the business needs, guarantee the smooth operation of the enterprise application system; in addition, the storage platform must be intelligent, can dynamically adjust the resource, automatic tuning,

optimization of I/O processing, and have good compatibility, the need for linkage with virtualization software, virtual machine to complete the relevant operation, provide the overall performance of the system; furthermore, the storage platform to have sustainable good expansion ability, meet the development needs of business in the coming years. In order to make the non structured data and file type data, at the same time can be stored in the unified platform, therefore requires the use of SAN+NAS unified storage, storage resources in storage from Input Output Per Second (IOPS) in terms of bandwidth and backup performance and storage design.

In cloud storage, the data backup mechanism is also very important because the backup strategy can resist the user error and some cases of software and hardware failures. In the cloud computing data center, virtual machine application virtualization platform increasingly, more than 40% enterprises have deployed the application of virtualization, and more than 30% of users plan to use cloud computing system in the short term. The virtual application platform to the core of the increasingly serious data loss, the main backup for a large number of customers of data virtualization platform, the virtual machine is difficult in data backup, the specific performance: the number of duplicate data, increasing the time required for data backup. Therefore, it is needed to design a system for the central design of the private cloud computing system to improve the efficiency of duplicate data backup.

The commonly used storage technology including Direct Attached Storage (DAS), Network Attached Storage (NAS) and Storage Area Network (SAN) three, because DAS technology affect the scalability of the system, so in the private cloud computing data center, should focus on using NAS or SAN technology.

In the NAS storage system, storage system through the I/O bus comrades attached to a server, but through the network interface and the network directly connected directly, similar to a dedicated file server, as shown in Figure 1 the principle of NAS:

The advantages of NAS are: easy deployment, can be used for efficient file sharing tasks; drawback is the request using the file, the poor performance of the equipment, and the storage rate is not high.

SAN is a high-speed, dedicated to the storage system of the network, generally independent of the computer LAN and the existence of. SAN can provide a dedicated communication channel for any one of its hosts and any storage device. According to the different ways of connection, SAN has two kinds of FC-SAN and IP-SAN.

FC-SAN is composed of disk array RIAD connection fiber channel, through the SCSI protocol to achieve data communication, the network topology is scalable, through the fiber channel with high transmission rate of the direct connection. Schematic diagram FC-SAN 2:

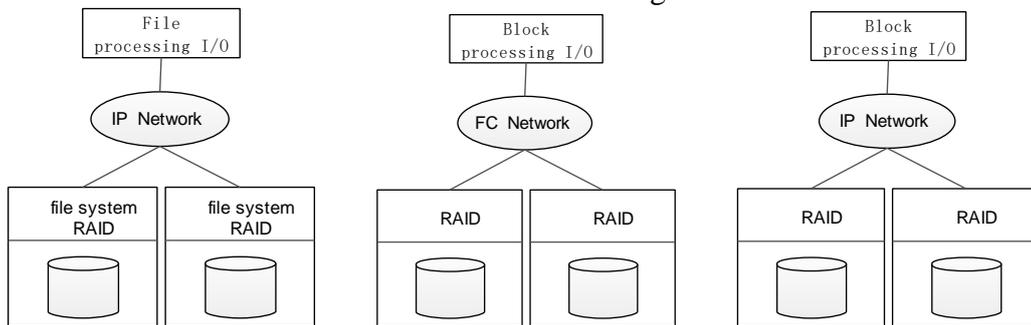


Figure 1 NAS schematic diagram

Figure 2 FC-SAN schematic diagram

Figure 3 IP-SAN schematic diagram

The advantage of FC-SAN is the farther the distance between servers and storage devices with high reliability, high performance, high scalability, improve the utilization rate of the storage device; the disadvantage is the need for optical fiber as transmission medium, the initial cost is higher.

IP-SAN is a fast data storage technology based on iSCSI interface in Gigabit Ethernet, which can implement SCSI protocol on IP network. The principle of IP-SAN is shown in Figure 3.

The advantage of IP-SAN is the high utilization rate of storage equipment, low initial investment; the disadvantage is that the performance is not as good as FC-SAN.

Summary

In the planning and design of private cloud computing data center, according to the specific needs of private cloud business applications, the number of servers from the CPU operational planning, demand planning, cyber source and storage resource planning and other aspects of overall design, to balance the relationship between system performance requirements and system of budget constraints, and ultimately achieve the best price.

Reference

- [1] Wei Xing,Zhang Jian-jun,Shi Lei. Dynamic Active Servers Allocating Policy for Cloud Computing Data Centers[J]. Journal of Electronics & Information Technology,2015.37(8):2008-2012.
- [2] Cao Jun-wei, Li Ke-qin and Stojmenovic I. Optimal power allocation and load distribution for multiple heterogeneous multi-core server processors across clouds and data centers[J]. IEEE Transactions on Computers, 2014, 63(1): 45-58.
- [3] YOU Xiao-ming. Design and Implementation of New Type Cloud Computing Server[J]. Computer Engineering,2011,37(11):40-42.
- [4] Li Shuangquan, Wang Yanwei. Research on the Server Virtualization Technology in Cloud Computing[J]. Designing Techniques of Posts and Telecommunications,2011,11:27-33.
- [5] Guo Yuan-xiong and Fang Yu-guang. Electricity cost saving strategy in data centers by using energy storage[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(6):1149-1160.