Applied Mathematics, Modeling and Computer Simulation
C.-H. Chen et al. (Eds.)

doi:10.3233/ATDE230942

Resources Optimization Research of Service Oriented Architecture Based on Hyper Converged Infrastructure Cloud Computing Platform

Yang Yang[1], Jia Junjun, Lv Hua
China National Petroleum Corporation Tubular Goods Research Institute

Abstract. Nowadays Cloud Computing (CC) has been to a period of flourishing development in the worldwide. Hyper Converged Infrastructure (HCI) has been leading software defined datacenter of CC to a new stage. Service Oriented Architecture (SOA) in HCI CC platform provides the agile and abstract distributed IT architecture mode, and it will dramatically change the service style and efficient of software defined datacenter. The revised simplex method is used to solve the optimization question of resources schedule in SOA based HCI CC platform, and the result shows at least 25% profit promotion, while in medium or large-scale CC platform the promotion will be greater.

Keywords. Cloud computing, hyper converged infrastructure, service oriented architecture, revised simplex method

1. Introduction

Cloud Computing belongs to the most fabulous technologies of the 21st century [1], and it has established itself as a fundamental aspect within modern internet infrastructure [2]. HCI as the leading software defined datacenter technology of CC, has been to the main increasing impulse of global CC market. According to the report published by Allied Market Research, the global HCI market is expected to hit $33.16 billion by 2026, registering a CAGR (Compound Annual Growth Rate) of 30.7% from 2019 to 2026 [3].

CC platforms are widely adopted by enterprise, serving as the IT infrastructure today [4]. It represents a paradigm shift in provisioning on-demand computational resources underpinned by datacenter infrastructure. So it is used by providers to deliver IT services as a utility: customers have access to on-demand service enforced by a Service Level Agreement.

2. HCI

HCI is a software-defined, unified system that combines all the elements of a traditional datacenter: storage, compute, networking and management. This integrated solution uses

[1] Corresponding author, Yang YANG, China National Petroleum Corporation Tubular Goods Research Institute, Xi'an, Shaanxi, China; Email: yyyang911@163.com.

software and x86 servers to replace expensive, purpose-built hardware [5]. It combines compute virtualization storage and networking in a single cluster [6]. Furthermore, it contains restore component, snapshot technology, duplicate data deletion, online data compression and multiple unit devices polymerize through network to compose seamless scale-out modules within a unified resource pool. CC platform as virtual datacenter through HCI will be stable and reliable, resource flexible and architecture extendable.

2.1 Advantages

HCI uses integrated virtual datacenter management interface for computing, storage and network resources dispatch and delivery. The resources are integrated to the pools, while the working system will not occupy the single machine but share the hardware resource with some other ones. So the pools become dispatching elastically and distributing agility. The services provided by HCI virtual datacenter are standard and safety.

2.2 HCI Hardware Preparation

With information technology rapid developing, compute, storage and network resources are needed tremendously, and the hardware virtualization technology is promoted to resolve the coupling and sharing problem of devices and resources. In HCI platform at least three physical machines are deployed to realize the virtualization. According to the computing scale and requirements more physical machines can be integrated to the CC platform. The physical machines are installed double network interface card and connected with double 10 Gigabit switches to guarantee the network speed not to become transition bottleneck. A Gigabit switch is connected as management switch. The network configuration guarantees bandwidth and speed while considering economy. The devices are listed in table 1.

Table 1. The devices of HCI CC platform

Device Type	Performance Parameters	Redundancy Design	Quantity
HCI Physical Machine	CPU: Intel(R) Xeon(R) Silver 4210 CPU Memory:4* 64GB DDR4 STORAGE:4* 480GB SSD & 8*2TB SAS (RAID)	power network interface card	6
10G Switch	exchange capacity≥640Gbps forwarding performance≥476Mpps port delay < 2.3us(electrical interface)	device redundancy	2
1G Management Switch	exchange capacity≥ 128Gbps forwarding performance≥ 96Mpps port delay < 3.3ms(electrical interface)	singleton	1

2.3 Software Installation

2.3.1 Vmware Esxi

VMware ESXi is an enterprise-class, bare-metal hypervisor dedicated to providing the state-of-the-art private cloud infrastructures. VMware ESXi uses a technique called lazy

restore that loads the memory image from disk where the virtual machine (VM) runs [7]. The small footprint provides numerous advantages, such as reduced attack surface, less code to find bugs, shorter deployment times, and multiple deployment methods [8]. The methods of install ESXi are the interactive way, scripts, auto deploy and CLI based.

2.3.2 Vcenter Server Appliance

vCenter Server brings the advantages of the client-server architecture to the ESXi host and to VM management. It is an application that serves as a centralized management tool for ESXi hosts and their respective VMs. It is pre-configured Linux virtual machine and flexible and extendable virtualization management platform of vSphere, to promote the integrated control and visibility of every levels in virtual foundation architecture.

2.3.3 Vcenter Management Database

vCenter Server requires databases to store and organize server data. vCenter Server support MS SQL, Oracle and Postgre SQL. The vCenter Server system must have a Data Source Name. This requirement applies to all supported databases. The procedure is installing the ODBC drivers and using administration application to create the DSN.

2.3.4 VcENTER DEPLOYMENT

vCenter allows to centrally manage hosts from either a physical or a virtual Windows machine, and enables the use of advanced features such as VMware Distributed Resource Scheduler, high availability, and VMware vMotion. The installation procedure of vCenter is to configure a continuously interactive dialogue boxes. If a bundled database is installed, such as SQL server Express instance (for small-scale deployments), this database is suitable for up to 5 hosts and 50 virtual machines. Then create a standalone VMware vCenter Server Instance for the hosts and virtual machines. Now, we can use vSphere client or web client to connect to the VMware vCenter.

2.3.5 Software Defined Storage

Software-defined storage (SDDC) is a new generation of storage system [9]. It is the automation and pooling of storage through a software control plane, and the ability to provide storage from industry standard servers. This offers a significant simplification to the way storage is provided and supported, and also paves the way for storage on industry standard servers at a fraction of the cost. HCI realize real meaning SDDC.

2.3.6 Vsan

A distributed port group is configured to bearing VSan data flow. A single ESXi machine can use the distributed port group by creating VMKernal interface. In HCI all of the ESXi can visit the shared distributed VSan storage after putting VSan into use. Disks groups can be managed in VSan management interface and used for deploying virtual machine. Till then the HCI based cloud computing platform is constructed.

3 SOA

SOA is composed by specified functional services and which can communicate with each other, and it is an agile and abstract distributed IT architecture mode. It is a strategy technical framework, and impels enterprise inner and outer systems to disclose and visit well defined service and information band by the services, and is abstracted to flow layers and combined applications for new solutions.

3.1 Service Mode

The central objective of a service-oriented approach is to reduce dependencies between "software islands," which basically comprise services and the clients accessing those services [10]. The purpose of this architecture is to address the requirements of loosely coupled, standards-based, and protocol-independent distributed computing, mapping enterprise information systems appropriately to the overall business process flow [11]. The problems can be solved by SOA through integrate the fine granularity to CC services abstracted by coarse granularity. CC has many modes of services stack, and the most important are as follows: IaaS (Infrastructure as a service), PaaS (Platform as a service), SaaS (Storage as a service), Daas (Database as a service), AaaS (Application as a service): Also called as software as a service, provides software services in cloud computing through networks.

Though integration with SOA in CC, we achieve the aim of unifying IT services governance and cloud infrastructure, and realize the SOA in CC, to solve the integration of infrastructure and service mode fundamentally.

3.2 Service Resource Optimization

The services stack in CC has many forms, and different service occupies different cpu, memory, storage, network and display resources, so the resources schedule for the total profit optimization becomes a general issue. Sivadon, S. Ezhildevi discussed about the algorithms and methods to solve optimization of resource provisioning cost in cloud computing [12] [13]. According to the question, we built an optimization model to solve the resources schedule question for the total max profit, and we tested according to the HCI CC resources and SOA modes and obtained the satisfied effect.

3.2.1 Optimization Model

Set the SOA CC platform has m types of different and quantitative resources, and it can provide service mode type of n with different service value V. when allocating resources, it is considered that how to effectively use the quantitative resources to obtain max profit. Upon the point, it is set that:

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}=[A1,A2,\dots\dots, Am]T=(P1,P2,\dots,Pn)$$

aij is quantity of resource Ai occupied by a kind of service Bj (i=1,2,…,m; j=1,2,…,n, m≤n), Ai=(a_{i1},a_{i2},…,a_{in}), Pj=$[a_{1j}\ a_{2j}\ \cdots\ a_{mj}]^T$; B=(b1,b2,…, bm)^T, bi is quantity of resource Ai; V=(v1,v2,…,vn)^T, value of service mode Bj; X=(x1,x2, … , xn), xj is quantity of service mode Bj (To be quantified).

And also,

(1) resources allocated should not exceed the total resource quantity of SOA CC platform;

(2) quantity of service mode is positive.

So the question is to be:

$$\begin{cases} max\ Z = \max VX \\ s.t.\ \mathrm{Ax} = \mathrm{B} \\ X > 0 \end{cases}$$

3.2.2 Using Revised Simplex Method to Solve the Linear Programming

According to the above condition, it is a typical linear programming question, and classical extreme value theory is inability, only simplex can solve the question effectively [14]. In order to calculate on computer, we use revised simplex method promoted by Dantzig in 1954 to solve the question [15]. The method is based on matrix and less computation. The steps are as follows:

Set: B=(P_{j1}, P_{j2}, …, P_{jm}), $B^{-1}=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$, $xb=B^{-1}b=[b_{10}\ b_{20}\ \cdots b_{m0}]^T$, N={1,2,…,n}, D={j1,j2,…,jm}, G=N\D.

(1) Calculate Simple type multiplier $\pi^T = c_B^T B^{-1}$;

(2) Calculate $b_{0j} = \pi^T P_j - c_j (j \in G)$. If $b_{0j} \geq 0 (j < s, j \in G), b_{0s} < 0 (s \in G), go\ to\ 3)$; else (if all $b_{0j} \geq 0 (j \in G)$), then stop, the optimal solution is:

$$\begin{cases} x_{ji}^* = b_{i0}, i = 1,2,\cdots,m \\ x_j^* = 0\ (other\ components) \end{cases}$$

(3) Calculate the vector:

$$B^{-1}P_s = \begin{bmatrix} b_{1s} \\ b_{2s} \\ \cdots \\ b_{ms} \end{bmatrix}$$

If all $b_{is} \leq 0\ (i = 1,2\cdots,m)$, then stop, the question has no optimal solution; else, go to 4)

(4) Calculate $\theta = \min\limits_i \left\{ \frac{b_{i0}}{bis} \middle| b_{is} > 0, 1 \leq i \leq m \right\}, j_r = \min\limits_i \{ j_i | \frac{b_{i0}}{b_{is}} = \theta, b_{is} > 0, 1 \leq i \leq m \}$

(5) Form elementary transformation matrix Ers;

(6) Calculate $\overline{B^{-1}} = E_{rs}B^{-1}, x_{\overline{B}} = E_{rs}x_B$, using $\overline{B^{-1}}$ to replace B-1, and $\overline{G} = N\backslash\{j_{1s},\cdots,j_{r-1,s},j_{r+1,s},\cdots,j_m\}$.

Thus to calculate in limited times, we can get the optimal solution or judge of no optimal solution.

3.2.3 Soa Schedule Calculation: A Case Study

According to the above method, considering the SOA CC resources limitation, we can calculate the optimal schedule of SOA and obtain the maximum profit. For example, CC platform contains cpu, memory, storage resources (m=3), and provide 5 modes of services (n=5). Every resource quantity is b=(1000GHz 6000GB 1000TB), and the services occupy resources as the table 2.

Table 2. The resources quantity occupied by every service mode

Service & Resource	Service 1	Service 2	Service 3	Service 4	Service 5
CPU (GHz)	4	8	12	16	8
MEMORY (GB)	8	16	24	32	8
STORAGE (TB)	0.5	2	3	4	10
SERVICES VALUES	1	1.8	3.5	6	4

To calculate as the above steps, we get the optimal solution of X= [0,0,0,13.5,17.3] in the CC platform. The total hosts are 30. If we allocate resource equally, X= [6,6,6,6,6], and the total profit is Z=97.8. While in actual condition, as least 3 to 5 hosts should be scheduled for every service mode. Generally, we allocate hosts with less value. So X= [3,3,3,13,8], and the total profit is Z=122.9. The optimization profit is added to 25% in contrast.

The result is based on a small scale CC platform. If optimization method is applied to medium or large scale platforms such as resources type more than 5 and services modes more than 10, the optimal profit by the method will be greater than average or random schedule conditions.

4 Conclusion

In this paper installation and deployment of CC platform based on hyper converged infrastructure is detailed introduced, including VMware virtualization software suits and VSan deployment process to realize management and maintenance. SOA is promoted to achieve an agile and abstract distributed IT architecture mode in the HCI CC platform. Since different service mode occupies different resources, in order to get optimal solution for maximum profit we use revised simplex method to solve the linear programming question. The result shows that in a small scale cloud computing platform the profit will increase at least 25%, and it is predictable that in large scale platform the optimization profit will be much bigger than the average or random total profit.

ACKNOWLEDGMENTS

This work is supported by Scientific Research and Technology Development Project of China National Petroleum Corporation Limited. Fund Topic Name: Research on Digital Construction and Application of Petroleum Standard Resources, Grant No: 2022DQ0108-17(JT)

REFERENCES

[1] Ayesha Nasir, Tahir Alyas, Muhammad Asif, Muhammad Naeem Akhtar. Reliability Management Framework and Recommender System for Hyper-converged Infrastructured Data Centers. 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). IEEE, 2020, p. 987

[2] B. Primas, P. Garraghany, D. W. McKee, J. Summersz, J. Xu. A Framework and Task Allocation Analysis for Infrastructure Independent Energy-Efficient Scheduling in Cloud Data Centers. 2017 IEEE International Conference on Cloud Computing Technology and Science (CloudCom). IEEE, 2017, p. 1
[3] Allied Market Research. September 16, 2020. http://www.globenewswire.com/news-release/2020/09/16/2094548/0/en/Global-Hyper-Converged-Infrastructure-Market-Is-Expected-to-Reach-33-16-Billion-by-2026-Says-AMR.html
[4] Hao Wen. Improving Application Performance in the Emerging Hyper-Converged Infrastructure. 2019.03. Chapter 1, p. 1
[5] Carlos Melo; Jamilson Dantas; Paulo Maciel. Availability Models for hyper-converged cloud computing infrastructures planning. International Journal of Grid and Utility Computing, 2020 Vol.11 No.2, pp. 196–208.
[6] Rizqie Joko Prabowo, Achmad Nizar Hidayanto, Puspa Indahati Sandhyaduhita. The Determinants of User's Intention to Adopt Hyper-Converged Infrastructure Technologies: An Integrated Approach. 2018 International Conference on Information Technology Systems and Innovation (ICITSI). IEEE, 2018, p. 306
[7] Irene Zhang, Tyler Denniston, Yury Baskakov, Alex Garthwaite. Optimizing VM Check pointing for Restore Performance in VMware ESXi. Proceedings of the 2013 USENIX conference on Annual Technical Conference.
[8] Marshall, Nick Brown, Mike Fritz, G. Blair Johnson, Ryan. Planning and Installing VMware ESXi. Issue Date: 2018. p. 23-46
[9] Ming-Jen Huang, Chun-Fang, Huang. Architecting a Software-Defined Storage Platform for Cloud Storage Service. 2015 IEEE International Conference on Services Computing.
[10] Michael Stal. Using Architectural Patterns and Blueprints for Service-Oriented Architecture. IEEE, 2006, p.55
[11] Mike P. Papazoglou · Willem-Jan van den Heuvel. Service oriented architectures: approaches, technologies and research issues. The VLDB Journal (2007) 16:389–415, p. 382
[12] Sivadon Chaisiri; Bu-Sung Lee; Dusit Niyato. Optimization of Resource Provisioning Cost in Cloud Computing. IEEE, 2012, p.164
[13] S.Ezhildevi, H.Shabuddeen. Optimizaiton of Resrouce Provisioning Cost in Cloud Computing. 2014 International Journal of Innovations in Scientific and Engineering Research (IJISER).
[14] Ao Te-gen. Analysis of the formation and development of the simplex method. Journal of Northwest University (National Science Edition). Oct., 2012, Vol. 42, No. 5, p.863
[15] Dantzig G B, Orden A, WOLF P. Generalized Simplex Method for Minimizing a Linear Form Under Inequality Restraints [J]. RAND Report RM-1264, Santa Monica The RAND Corporation, 1954,5:183-195.