Advanced Science and Technology Letters
Vol.49 (SoftTech 2014), pp.89-94
http://dx.doi.org/10.14257/astl.2014.49.19

Storage Sizing Issue of VDI System

Jeongsook Park[1], Cheiyol Kim[1], Youngchang Kim[1], Youngcheol Kim[1], Sangmin Lee[1] and Youngkyun Kim[1]

[1] Electronics and Telecommunications Research Institute
218 Gajeong-ro, Yuseong-gu, Daejeon, 305-700, KOREA
{jungsp, gauri, zerowin, kimyc, sanglee, youngkim}@etri.re.kr

Abstract. VDI system is a complicated one which has many components such as hosts, centralized storage and network and demands PC-class performance at the same time. Centralized management of all resources complicates the VDI system even more, and especially storage is considered as the critical bottleneck point. Therefore we focused on storage sizing issue of VDI system with IOzone and IOmark which measure system performance in terms of IOPS and response time. IOmark is widely used to measure storage performance in VDI environment in the absence of standardized tools, which measures response time with workload specified to VDI system. The results showed that conventional calculation method of storage size couldn't meet the performance well. We analyzed causes of it and suggested various methods to overcome the low performance.

Keywords: VDI, storage sizing, IOPS, response time, IOmark

1 Introduction

Virtual Desktop Infrastructure (VDI) has been recently one of key technologies in the IT industry, which has many benefits such as acquisition of security by centralized data management, saving energy, and so on. These aspects can appeal to IT businesses and make them consider the introduction of a VDI system to their companies.

However, in spite of the merits, VDI has several limitations of cost, performance and techniques which make them to hesitate before. Among them, storage I/O performance is the most critical bottleneck point. If enough resources allocated to the system correctly, user dissatisfaction will be followed because VDI system is to provide PC environment to each end user with centrally managed on the server side. End users don't endure slower response time than their expectations, and consequently the system will be a headache.

Therefore we focus on storage sizing issue of VDI system with IOzone and IOmark which measure system performance in terms of IOPS and response time. IOmark is widely used to measure storage performance in VDI environment in the absence of standardized tools, which measures response time with workload specified to VDI system [1-2].

ISSN: 2287-1233 ASTL

This paper consists of 5 sections. Section 1 is an introduction describing about storage sizing issue of VDI system. In section 2 we explain VDI workload characteristics which have influence on system performance and several issues to be resolved. In section 3 system configuration and experimental methods are described. In section 4 testing results are described in detail. We conclude our paper in the last section.

2 VDI Workload Characteristics and Storage

Understanding the characteristics of VDI workload is the first step in developing a VDI system because VDI workload has its own unique feature. VDI workload issuing from each VM in the steady state may be trivial, ranging from 3-7 IOPS to 20~30 IOPS [3]. Though loads fluctuate in situations, but 50 IOPS is enough even in the worst case. However, it becomes burden to the system as the number of active users increases because the summation of trivial loads becomes not trivial. Specifically workload from VMs is almost reads or writes to the storage and it makes storage the most critical bottleneck point. Reports of workload characteristics says that more than 80% of traffic from VMs, in some case more than 90%, are small random writes [4]. It can degrade storage performance. But this will be acceptable in terms that total load to the system in a steady state will be even.

However, system performance may be deteriorated severely by I/O storms which are short-term and heavy load such as simultaneous logins at the commute time, simultaneous virus scans and Windows updates. Workload in I/O storms may amount to about 4~6 times of in steady state and 10 times in the worst case, so it can influence heavily on system performance deterioration [5].

To accommodate this situation, we have to calculate system size including CPU, RAM and storage in order to meet I/O storms. But it is not desirable because workload in a steady state is low and waste system resources and even doesn’t guarantee stationary services. Therefore, it is desirable to build the system size based on steady state and implement additionally mechanisms to accept high traffic conditions.

3 System Configuration and Testing Methods

3.1 Storage Size Calculation

For testing, we simply assume that the system supports 100 office users requiring 50GB dedicated storage and Windows 7 OS (i.e. demanding 2G RAM and 2 CPUs for Windows 7) respectively. The first thing to do is to calculate system size to support the size of users. Storage size, the most important resource among them, is generally calculated using required IOPS per user type, network capacity, disk IOPS by disk RPM, write penalty of RAID type, the ratio of read and write and so on [6].

$$\# \text{ of disks} = (\text{Total ROPS} + (\text{Total WOPS} * \text{RAID Penalty})) / \text{Disk speed} \quad (1)$$

According to information, workload of an office user ranges in 8~25 IOPS [3] and even in case of 100 users, 800~2,500 IOPS is required. It requires 82Mbps on the network and 1GbE is thought to be enough for 100 office users. When we use a 750GB 7,200 RPM SATA2 disks, RAID10 configuration and 4:6 of R:W ratio, 17 to 53 disks are required to support 100 office users with 8 to 25 IOPS. Higher the ratio of R:W, more disks are necessary.

3.2 System Configuration

Based on the calculation, we construct the system with 12 hosts, 1 MDS and 1 DS shown in Figure 1.

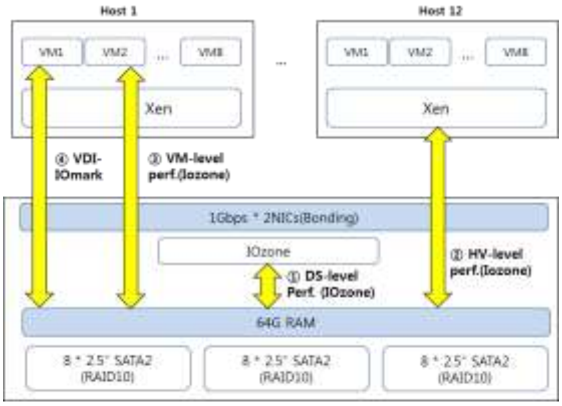

Fig. 1. System configuration and testing issues

1 CPU and 1GB RAM is for the hypervisor because we want to exclude the performance of a hypervisor. Network has 2Gbps capacity with bonding of 2 NICs. DS has 64GB RAM and 17TB storage by RAID10 of 24 750G SATA2 disks. 17TB is enough space for 100 VMs. Moreover 24 SATA2 disks are supposed to support more than minimum IOPS of 100 office users. System specification for experiment is shown in Table 1.

Table 1. System specification

MDS(1)	OS: CentOS 6.2 64 bit, CPU: 4Core 2CPUs(HT, 16Cores), Memory: 64GB
DS(1)	Disks: 24 750GB SATA2 Disks with RAID10 (RAID options: Stripe size=256KB, write through, read ahead, direct disk cache enable) OS:CentOS 6.2 64bit, CPU: 4Core 2CPUs(HT, 16Cores) Memory: 64GB RAM
Host(12)	Hypervisor: 1 CPU, 1G RAM VM: 2 CPUs, 2G RAM (totally 8VM per host)

3.3 Testing Methods

The experiments are performed in two steps. Firstly basic performance is measured in terms of DS-level, hypervisor-level and VM-level. After that, response time is measured with IOmark which is an official tool to measure response time of a VDI system.

4 Experimental Results

4.1 DS-level

DS-level testing is to understand the basic I/O performance of RAID10 disks in a DS node. We use IOzone and measure aggregated I/O performance of sequential and random read/writes to the disks. Parameters for this test are the number of threads, varying values of 3, 12, 24 and 48, 1G file per thread, and I/O size with 4K~128K for sequential I/O and 4K for random I/O.
Result of the sequential I/O in A of Figure 2 showed that aggregated performance increase as I/O size grows and read performance was better than write because of write penalty of RAID10. Result of the random I/O in D of Figure 2 showed that random read performance was increased linearly as the number of threads grew, but random write performance remained uniformly because of write penalty of RAID10.

4.2 Hypervisor-level

Hypervisor-level testing is to understand the network performance between hosts and DS by measuring the minimum performance between DOM0 and DS (i.e., cold I/O), maximum performance (i.e., warm I/O) and aggregated sequential and random performance.

We used IOzone cluster tool in hypervisors in 12 hosts. Parameters for the test are 8 threads per host emulating 8VMs and I/O size of 128KB, the maximum size of kernel FUSE, to 1G file per thread for sequential I/O performance. For warm I/O performance we created 500 M file per thread because RAM size of DS is 64GB. For random cold I/O performance, we created 1G file per thread in order to make the total size of cached data overflow the RAM size.

Result of the sequential I/O in B of Figure 2 showed that network performance was 220MB/s, which was supposed as the maximum in 2Gbps network, and testing more than 2 hosts utilized of the maximum network performance. We analyzed this phenomenon was caused by the factors that the maximum size of kernel FUSE is 128KB and data upload/downloads were possible through already established sessions between client and the file system. Result of random warm I/O in E of Figure 2 was much better than random cold I/O, where warm I/O was maximal when the number of hosts was 4 and performance dropped more than 8 hosts. Result of cold I/O dropped under conditions of more than 4 hosts. The cause of this was analyzed as increased

IOPS by I/O to the disk. But even in the worst case, random performance resulted in more 3,000 IOPS owing to the RAM cache of DS.

4.3 VM-level

VM-level testing is to calculate the maximum bandwidth and IOPS which a single storage server can provide.

We measured at one VM and then at all of VMs with IOzone. At a VM, we tested random I/O performance within the range of 2GB after creating 4GB file. To measure the aggregated performance, we registered a measuring job in the job scheduler in every Windows7 in advance to start at the same time and after the testing we accumulated the results from 100VMs and got the aggregated and averaged performance.

We anticipated that the performance of a single VM would be warm performance because all of data were cached in the DS memory. The result of sequential performance in C of Figure 2 showed write was lower than read because memory of a VM with size 2G was smaller than the file size of 4G and this caused ceaseless overhead between VM memory and DS. The result of random performance at a VM in F of Figure 2 showed that write was much better than read because write used the RAM of the VM as write-back cache. The result of aggregated performance showed that random reads were started at the same time but random write was not because random read was finished at different time by resource contention at the hypervisor and DS levels resulting in random writes at different times and different performances depending on every VM. But even in this environment aggregated performance showed 80 IOPS, higher than the required maximum IOPS, 25.

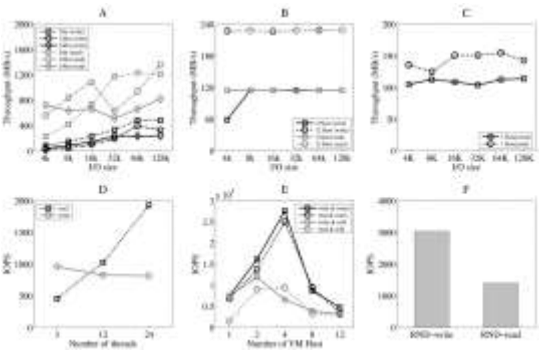

Fig. 2. I/O performance with IOzone

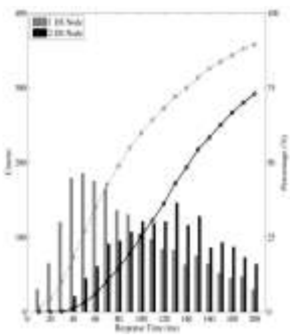

Fig. 3. Performance with IOmark

4.4 Response time

IOmark is a practical and useful tool in VDI environment because it replays synthetic workload of real VDI traffics. Testing with IOmark is to measure the maximum performance at the aspect of response time which the file system can provide.

Through it we can understand the maximum number of users which our system can guarantee. We used 50ms as a metric of response time, where response time in PC environment is known as within 50ms.

We used the size of users as 96 because a job in one VM can model workload of 8 users. To disperse the load as much as possible, we installed IOmark at three hosts, 2 VMs per a host. And we tested response time with steady state workload.

The light line in Figure 3 was the result. The performance was too low and the ratio of response time within 50ms was only 3%. So we attached one more DS with the same system specification. But the result could not show the performance of our expectation. The ratio of response time within 50ms was about 27%. We thought that minimally over than 80% can suffice the users of the VDI system. Theoretically, 48 disks were supposed to suffice the performance which we expected under the condition of 20 IOPS per user but the experimental result totally different.

5 Conclusions

In this paper, we measured IOPS and response time to decide whether the system configuration based on prevalent storage size calculation suffices the goal of supporting 100 office users for VDI system. Differently from the calculation, the measurement of response time couldn't meet our expectation. This may be caused by transient peak workload periods which occur not infrequently. We expect that this phenomenon can be mitigated by introducing a cache between hosts and the storage because the cache can absorb most part of peak workload characteristics.

Acknowledgment. This work was supported by the IT R&D program of MKE/KEIT, KOREA. [10041730, A Cloud File System Development for Massive Virtual Desktop Infrastructure Services].

References

1. http://www.evaluatorgroup.com/iomark, "Getting Started Guide with VDI-IOmark", Evaluator Group, December 30, 2011.
2. http://www.vdi-iomark.org/sites/default/files/VDI-IOmark_theory_of_operation.pdf, "VDI-IOmark: Theory of operationsGuide", Evaluator Group, January 2012.
3. http://www.vmware.com/files/pdf/view_storage_considerations.pdf, "Storage Considerations for VMWare Horrizon View 5.2", White Paper, VMware.
4. http://www.tdeig.ch/kvm/VDI/VDI_Storage.pdf, "VDI & Storage: Deep Impact", PQR, September 2011.
5. J. McLeod e t al., "5,000-Seat VMware View Deployment," NetApp, Aug. 2011.
6. http://www.emc.com/collateral/software/white-papers/h11096-vdi-sizing-wp.pdf, "Sizing EMC VNX Series for VDI Workload", White Paper, EMC, September 2012.

