International Journal of Grid and Distributed Computing
Vol. 9, No. 6 (2016), pp.259-268
http://dx.doi.org/10.14257/ijgdc.2016.9.6.24

Workload Characterization on a Cloud Platform: An Early Experience

Zujie Ren[1*], Jinxiang Dong[1], Yongjian Ren[2], Renjie Zhou[2] and Xindong You[2]

[1]College of Computer Science Zhejiang University
[2]School of Computer Science Hangzhou Dianzi University
[1]renzju@gmail.com,djx@zju.edu.cn ,[2]{rjzhou,youxindong}@hdu.edu.cn

Abstract

Understanding the characteristics of cloud workloads is the key to making optimal configuration decisions and improving the system throughput. However workload characterization of cloud, especially in a large-scale production environment, has not been well studied yet. To gain insights on cloud workloads, we collected a one-week workload trace from a 100-node cloud cluster which hosts 1082 virtual machines. We characterized the workload at the granularity of virtual machines and physical nodes, respectively. We concluded with a set of meaningful observations. The results of workload characterization are representative and generally consistent with cloud cluster for public IaaS service providers, which can help other researchers and engineers understand the performance and VM characteristics of the cloud in their production environments.

Keywords*: workload characterization; Cloud; IaaS;*

1. Introduction

The last decade has witnessed a rapid research progress in the field of cloud computing. Many large-scale cloud platforms such as Amazon EC2 and Microsoft Azure , have been implemented and widely used in practice.

Understanding the characteristics of cloud workloads is the key to making optimal configuration decisions and improving the system throughput. Cloud cluster operators can optimize the scheduling policies and allocate resources more effectively under diverse workloads. However, workload characterization of Cloud, especially in a large-scale production environment, has not been well studied yet.

To gain insights on cloud workloads, we collected a one-week workload trace from a 100-node cloud cluster. It is a public IaaS cloud platform. The trace comprises a one-week period of data, covering 1082 VMs running on 100 nodes. VMs handle diverse workloads, ranging from virtual hosts to application services, such as mail and web servers.

Our analysis reveals a group of workload characterization and provides five direct implications. We believe that workload characterization is useful for cloud cluster operators identify system bottlenecks and further optimize the performance. A deep understanding of the VM characteristics and physical workload helps us get more detailed information about cloud resource usage, which contributes to develop an optimal system configuration policy.

The contributions of this paper are listed as follows.

- We collect a one-week workload trace from a 100-node production cloud cluster. The trace includes 1082 instances, which are representative and common for a public cloud service provider. It is a complement to current public workload repository, especially on the aspect of IaaS cloud.

ISSN: 2005-4262 IJGDC

- We conduct a comprehensive analysis of the workload trace at the granularity of VM and physical node. The main observations and direct implications are presented in Table I. These findings can help other researchers and engineers better understand the performance and job characteristics of the cloud in their production environments.

The rest of this paper is organized as follows. Section II provides a brief introduction of cloud environment, and then gives an overview of the cloud cluster. Section III discusses the summary of the trace, including the information extracted and the items of the logs, and presents a detailed analysis of these logs at the granularity of VM and physical nodes. Section IV discusses related work and we conclude this paper in Section V.

Table 1. Summary of Observations and Implications

Observations	Implications
O1: Pre-defined scheduling mechanism is basically feasible and can deploy VM evenly.	The algorithm of VM placement can be optimized for cloud from the perspective of load balancing and energy.
O2: For users, approximately 80% of the demand for VM memory is less than 4G, the general demand for VM memory is concentrated between 2G and 4G.	Cloud system need to be optimized for allocating VM memory below 4G memory, which will be very beneficial.
O3: VM's computing tasks are concentrated during the daytime, for cloud systems operator.	Arranging the task of upgrading cloud system at the time when VM is not active is better.
O4: Average CPU utilization for all nodes is not high in the cluster, the determining factor is not the number of running VM.	Data centers need to establish appropriate standards for resource utilization, CPU usage is not the only indicator.
O5: Different applications running in the VM cause the physical nodes' diversity.	Putting VM's application types as one of sale parameters is beneficial for scheduling.
O6: Although VMs are distributed more evenly, but the physical node load are obvious differences.	There is the need to introduce consolidation system to schedule running VMs.
O7: Overall resource utilization is low in the cluster and fluctuate periodically.	It is better to shut down some idle nodes to save cost through migrating VMs. At everyday's peak time, it is convenient to identify performance bottlenecks and optimize the underlying system.

2. Background

To facilitate the understanding of cloud workload analysis in this paper, this section first gives a brief overview of the cloud cluster and its architecture. Then, we describe the VM states and transitions in the cloud.

Generally, a cloud cluster is composed of numerous physical servers, racks, routers, databases, etc. They work together to guarantee the cloud service in a stable state. In practice, the servers making up the cloud cluster will be virtualized in order to maximize the resources utilization of the physical computers, which distinguishes cloud from the traditional data centers. Each of these nodes runs multiple VMs, which can be directly offered to the tenants for different use or provided as well as other services such as PaaS or SaaS.

The cluster's running statistic trace was collected over one week period from Apr. 11 to Apr. 17, 2013. During this period, except for the monitor nodes, master nodes, and network control nodes, the number of the compute nodes which are provisioning VM was 100 and the number of VM in the cluster was 1082. One node can support 1 to 29 Vms, the average VM number was 10. When we started to collect the trace, most of the VMs were already running in the cluster.

In order to accurately record the cluster's running states, we developed a daemon process running in every node to monitor the nodes' workload and the VMs' workloads, and logged the data in real time. The VM data collection interval was 60-second, and the items consist of CPU utilization rate, memory size, write IOPS and read IOPS. The four workload characteristics can help us understand the tenants' behavior from the perspective of computing and IO. The nodes workload sampling interval was 5-minute, the data items consist of CPU information, which include *idle* (%), *sys* (%), *usr* (%), *wait* (%), free memory, write IOPS, read IOPS, and power consumption. The relationship between the nodes workload and the VM load can be revealed.

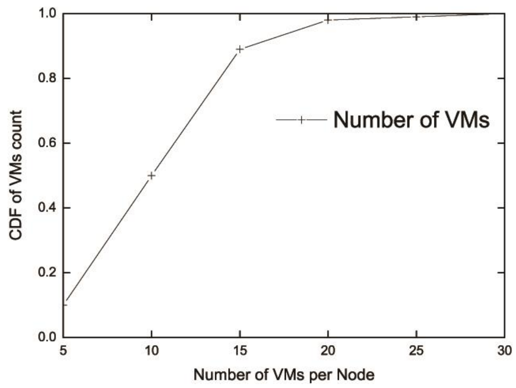


Figure 1. Distribution of VM per Node

3. Trace Analysis

In this section, we first describe VM statistics, including VM distribution in the cluster, and VMs running states. Then, we present nodes statistics, including CPU utilization, memory usage, I/O operation. At last, we analyze influence of VM on a physical node, especially in scheduling and destroying process. Our empirical observations are italicized at the end of the corresponding paragraph.

A. *Analysis of distribution and scheduling of VM*

Public cloud service providers deploy data centers in different areas so that local users' request transmission delay can be reduced. In general, they put the clusters in the same area as a so-called *big region* to schedule VMs. There are many data centers in one area, and many clusters in one data center. The data trace was sampled among the clusters. When a user issues a *create vm* request, the request passes many processes and at last reaches to one physical node. Operators set *weight* value for every cluster, the *weight* represents for the scheduling priority. The higher the *weight*, the cluster is more preferential to schedule. When there are multiple clusters having the same *weight*, the cloud OS considers the *inventory* value which refers to the number of the specific configuration of the VMs in the each *zone* in one cluster. The *zone* refers to the VM data storage unit, and one cluster is divided into multiple *zones*. According to the reported *inventory* value from each cluster, the control node selects the cluster with higher *inventory* value, then selects the *zone* with higher value too and sends *create vm* command to the *zone* finally. In the cluster, each physical node has a field called *health* to represent the health degree of capabilities on provisioning VM. Due to the complexity of the cloud environment, when the request is scheduled to the physical node, some scheduling functions may fail. So the *health* value adds one when succeeds and subtracts one when fails.

In the cluster, each physical node has a field called *health* to represent the health degree of capabilities on provisioning VM. Due to the complexity of the cloud environment, when the request is scheduled to the physical node, some scheduling functions may fail. So the *health* value adds one when succeeds and substracts one when fails.

Figure 1 depicts the cumulative distribution function (CDF) of VM distribution. As shown in the figure, the distribution of VM in the cluster fits the log-normal distribution. The number of VM in every node was relatively even. Most physical node provisioned about 10 VMs, more than 90% of the number of VMs on physical machines was less than 15. We also find that the node with more VMs had many small size VMs, and the less one often had large size VMs. Observation 1. *Pre-defined scheduling mechanism is basically feasible and can deploy VM evenly. Some previous works research the algorithm of the VM scheduling from the perspective of load balancing and energy [3], [4]. Therefore, the algorithm of VM placement can be optimized for cloud.*

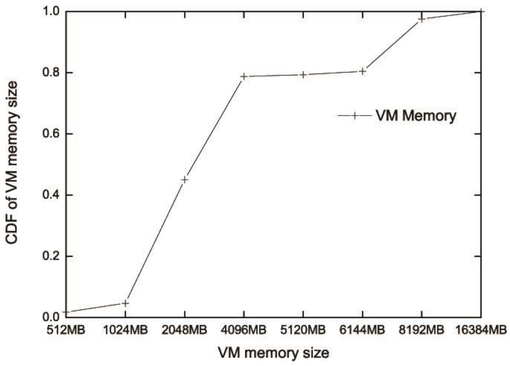


Figure 2. Cumulative distribution of VM memory.

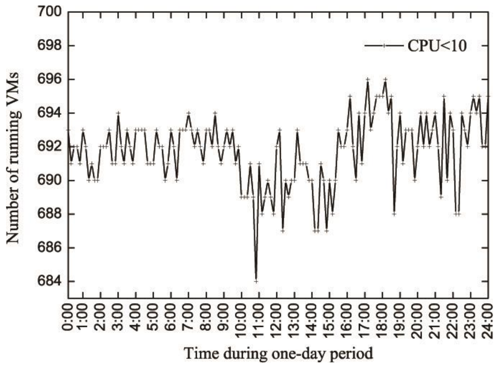


Figure 3. VM count for each 10-minute interval during one day.

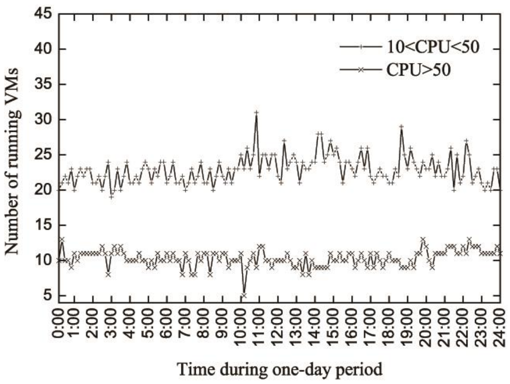


Figure 4. VM Count for each 10-Minute Interval During one Day.

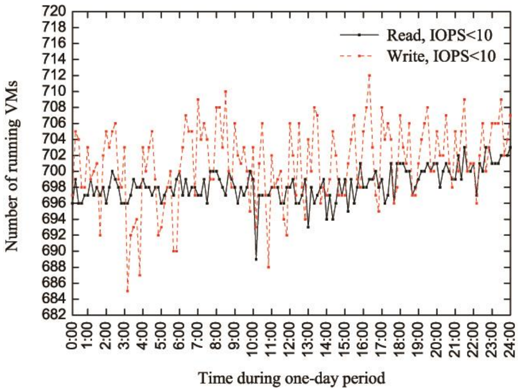


Figure 5. VM Count for each 10-Minute Interval during One Day

B. VM running state analysis

In the cloud cluster, VMs are sold to different users. So every VM encapsulates the different specific applications. VMs' memory occupation are changeless to the physical node, so we show the VM memory distribution in Figure 2. In general, running VM states can be divided into compute-intensive and I/O-intensive [5], but the partition method is blurry. Compute-intensive requires a lot of CPU computing resources, while I/O intensive consumes a lot of disk I/O resources and is very sensitive to I/O operation. When the same type of applications in the VMs on the same physical node, the resources will be competed fiercely, resulting in performance degradation. But when different types of applications are on the same nodes, they will get less resource competition pressure and not induce too much impact on the performance. Therefore, it is necessary to monitor the cluster VMs running statistics.

Figure 2 shows the cumulative distribution of VM memory size in the cluster. The VM memory size type is fixed by cluster operator as shown in Table II. VM memory capacity affects physical nodes' CPU wait time, because if the memory assignment is not sufficient, it will cause that VM's I/O operations increase, which translates to Dom0's I/O operation. Observation 2. *For users, approximately 80% of the demand for VM memory is less than 4G, the general demand for VM memory is concentrated between 2G and 4G, so cloud system need to be optimized for allocating VM memory below 4G memory, which will be very beneficial.*

Because of VMs' CPU utilization are widely distributed, we divided them into 3 groups, which are less than 10, 10 to 50 and more than 50. The data sampling interval was 10 minutes. Figure 3 shows all the VMs' CPU utilization rate's statistical distribution which is less than 10. The number of VMs declines in daily time, which corresponds to people's work habit. Figure 4 shows all the VMs' CPU utilization rate's statistical distribution which is between 10 and 50 and more than 50. We find that the VMs with low CPU utilization account for the majority in the cluster during one-day period, the number is nearly 700. They could be considered in the running state but not active, that is, there are no specific applications in that VM. But the number of VMs with relatively high CPU utilization is fluctuating around 10. The number of active VMs falls down to 26 at about 7:00 am and reaches the peak at about 11:00 am which exceeds 40. As shown in Figure 4, the number of running VM has a rise and fall process, but overall the number of VMs in the cluster is in a steady state, the fluctuation is not obvious. Observation 3. *VM's computing tasks are concentrated during the daytime, for cloud systems operator, arranging the task of upgrading cloud system at the time when VM is not active is better.*

Similarly, we divided IOPS which reflects the VM I/O operations into 3 groups, which are less than 10, 10 to 100 and more than of 100 to show VMs I/O statistical distribution more clearly. The data sample interval was 10-minute. Figure 5 shows all the VM IOPS statistical distribution which is less than 10. The fluctuation of write operation is larger than read's. Figure 6 shows all the VMs' IOPS statistical distribution which is between 10 and 50 and more than 50. We see that the fluctuation of VM's read operation is less than write operation and write operation's IOPS values is larger than read operation's. By comparing these two figures, we can observe that the VMs' read operation is the main I/O operation. But most of the VMs have done almost little I/O operation or a small I/O operations.

When a tenant creates a VM with 2 vCPU (virtual CPU), the CPU usage limit set for that particular VM might be 1, 0.5 or even 0.1 of the physical processor core. Similarly, a host server with 8 GB physical memory may be committing 12 GB or even 16 GB virtual memory to VMs. Understanding the VM resource consumption pattern helps developers design the over-commit (over-commit refers to the practice of committing more virtual resources to customers than the actual resources available on the underlying physical clusters) parameters to balance between the number of VM and VM' performance in the cluster. So it is necessary to set the over-commit parameter according to the VM's workload characteristics.

C. Analysis of physical resource utilization

In this section, we give a detailed analysis of the physical resource utilization statistics of cloud cluster nodes, including CPU, memory usage, disk I/O situations.

Figure 9 shows the average *cpu usr* percentage for all nodes in the cluster during one week. It occupies a large proportion of the CPU utilization. The CPUs in all compute nodes are basically uniform. The X-axis represents the day from 11/4/2013 to 17/4/2011, and the Y-axis represents the *cpu usr* utilization ratio. We find that the average *cpu usr* utilization ratio fluctuates between 10% and 15%. Compared with the previous VMs' CPU statistics, although there is a large number of VMs in the cluster, the overall CPU usage of physical nodes is low. The reason is that the VMs' CPU workload is not heavy. It exceeds 15% at peak and falls down to 2% in the lowest case and fluctuates periodically every day. Observation 4. *Average CPU utilization for all nodes is not high in the cluster, the determining factor is not the number of running VM. Data centers need to establish appropriate standards for resource utilization, CPU usage is not the only indicator.* Figures 7 and 8 shows the average *cpu usr and cpu sys* for all nodes in the cluster during one week, respectively. We observed that the average *cpu usr* ratio mainly ranges from 10% to 12%, while *cpu_sys* utilization ratio is between 8% and 10%. The high rate is due to VM's I/O operation and the virtual network mechanism in Xen, which cause large system calls in the host node.

Figure 9 shows the average *cpu wait* percentage for all nodes in the cluster during one week. The ratio is basically less than 6%, so the bottleneck of the cluster nodes is not I/O operations which don't waste much time for CPU. On the other hand, we can see that VM's memory allocations are sufficient and I/O operation is relatively less. Comparing the three figures, we can see that *cpu wait* is not relevant with *cpu usr* and *cpu sys*. We find that some nodes' *cpu wait* value is a higher ratio, while some nodes are very low. This can be solved by migrating VM from higher workload node to lower ones or through *Cgroup* tool to control the read and write frequency. Considering the server's IOPS often becomes a scarce resource, we can consider storage device such as SSD and PCIE [6]. The rate of *cpu usr* and *cpu sys* are relatively higher, but *cpu wait* rate is low. There is positive correlation between *cpu sys* and *cpu usr*. *cpu wait* is affected by the VM's I/O operation. Observation 5. *Different applications running in the VM cause the physical nodes' diversity. Putting VM's application types as one of sale parameters is beneficial for scheduling.*

About CPU utilization, we know that the following equation: *usr (%) + sys (%) + wait (%) = total usage*. According to the above statistics, it was observed that CPU utilization is relatively stable in this cluster and is at a low state. Although the cluster can run normally, but it did not make the full use of all computing resources. The number of active VMs whose CPU utilization rate is larger than 10, only accounted for 3% of all the VMs, hence, inactive VMs have little effect on the physical machine's CPU utilization.

Figure 10 and Figure 11 present the average I/O read and write operations for all nodes in the cluster during one week, respectively. The two figures reflect the nodes' I/O situations. We see that read throughput is less than the write throughput. The read operation's fluctuation is consistent with the *cpu wait*, while the write operation is not affected. From the log files of VMs' states, we observed that the fluctuation is due to the VM's I/O operations increase.

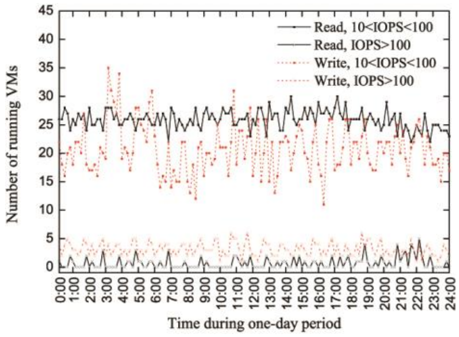


Figure 6. VM Count for Each 10-Minute Interval During One Day.

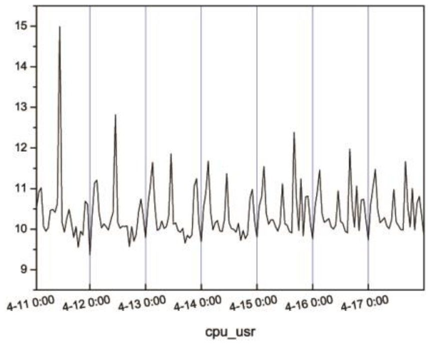


Figure 7. cpu usr Percentage

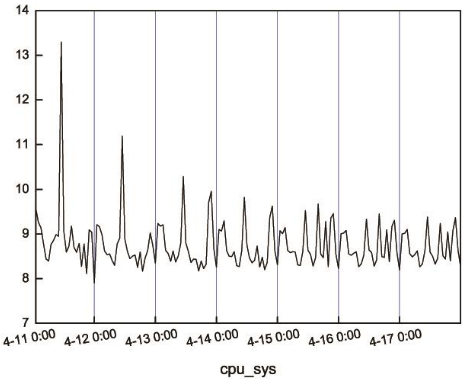


Figure 8. cpu sys Percentage

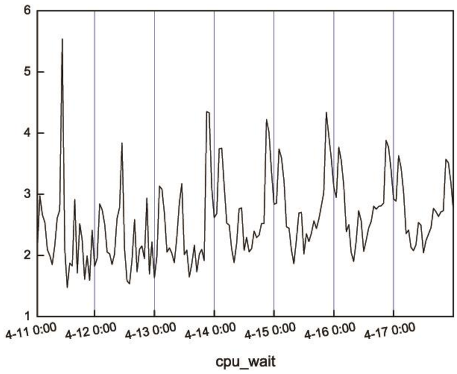


Figure 9. cpu wait Percentage

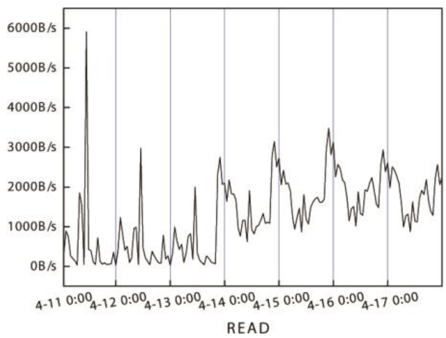


Figure 10. Read Rate

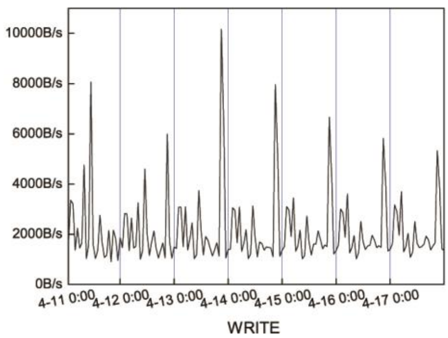


Figure 11. Write Rate

4. Related Work

Workload characterization studies are useful for cloud operators to identify system bottlenecks and figure out solutions for improving performance. Many previous efforts have been made in different areas, including network system [7], [8], storage system [9], [10], [11], Web server [12], [13], and HPC cluster [14]. Both network and storage subsystems are key components for the cloud system. Our trace analysis also refers to their research methods to analyze cluster resource utilization.

Several studies [15], [16], [17] have been conducted for workload analysis in grid environments and parallel computer systems. They proposed various methods for analyzing and modeling workload traces. However, the workloads characteristics and scheduling policies in grid are much different from the VMs in a cloud system.

Mishra et al. [18] focused on workload characterization that includes behavior characteristics of CPU and memory. The Yahoo Cloud Serving Benchmark [19] focused on characterizing the activity of database-like systems at the read/write level. Ren et al. [20] conducted an analysis about the Hadoop's job and task characterization, such as job arrive pattern and task waves, based on the Hadoop logs from the Tabao's internal Hadoop cluster. Their research work focused on finding jobs' running feature and exploring potential performance optimization approaches based on the historical workloads. These works are instructive to our work. Our study uses a similar statistical method to show the behavior at the granularity of VMs and physical nodes. Jiang [21] analyzed and compared the performance of the VM in different public IaaS providers in China.

However, these studies do not analyze many of the workload characteristics discussed in this paper. For example, we report not only the VM statistics but also node statistics. We also derive some direct implications from the observations of the cloud workload trace. They are helpful for cloud operators to improve system performance.

5. Conclusion

In this paper, we presented the analysis of cloud trace derived from a production cloud cluster. The trace covers 1082 VMs and 100 nodes running log files over a one-week period, which are representative and common for a public cloud service provider. We conduct a comprehensive analysis of the workload trace at the granularity of VM and physical node, respectively. Some main observations and their direct implications are concluded. These findings can help other researchers and engineers understand the performance and job characteristics of cloud in their production environments.

Acknowledgments

This research is supported by National Post doctor Science Foundation of China (2013M541779), the National Key Technology R&D Program under Grant (No.2014BAK14B04), and NSF of China (61300033). Renjie Zhou is supported by NSF of China (61300211). Xindong You is supported by NSF of China (61402140).

References

[1] AWS. [Online]. Available: http://aws.amazon.com/
[2] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," ACM SIGOPS Operating Systems Review, vol. 37, no. 5, **(2003)**, pp. 164–177.
[3] A. N. Tantawi, "A scalable algorithm for placement of virtual clusters in large data centers," in Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS), 2012 IEEE 20th International Symposium on. IEEE, **(2012)**, pp. 3–10.
[4] W. Xiaoli and L. Zhanghui, "An energy-aware vms placement algorithm in cloud computing environment," in Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference on. IEEE, **(2012)**, pp. 627–630.
[5] M. Rosenblum and T. Garfinkel, "Virtual machine monitors: Current technology and future trends," Computer, vol. 38, no. 5, **(2005)**, pp. 39–47.
[6] G. P. Chen and J. S. Bozman, "Optimizing i/o virtualization: Preparing the datacenter for next-generation applications," White Paper sponsored by Intel Corporation, **(2009)**.
[7] D. Ersoz, M. S. Yousif, and C. R. Das, "Characterizing network traffic in a cluster-based, multi-tier data center," in Distributed Computing Systems, 2007. ICDCS'07. 27th International Conference on. IEEE, **(2007)**, pp. 59–59.
[8] V. Paxson, "Empirically derived analytic models of wide-area tcp connections," IEEE/ACM Transactions on Networking (TON), vol. 2, no. 4,**(1994)**, pp. 316–336.

[9] L. N. Bairavasundaram, A. C. Arpaci-Dusseau, R. H. Arpaci-Dusseau, G. R. Goodson, and B. Schroeder, “An analysis of data corruption in the storage stack,” ACM Transactions on Storage (TOS), vol. 4, no. 3, **(2008)**, p. 8.

[10] E. Thereska, A. Donnelly, and D. Narayanan, “Sierra: practical powerproportionality for data center storage,” in Proceedings of the sixth conference on Computer systems. ACM, **(2011)**, pp. 169–182.

[11] F. Wang, Q. Xin, B. Hong, S. A. Brandt, E. L. Miller, D. D. Long, and T. T. McLarty, “File system workload analysis for large scale scientific computing applications,” in Proceedings of the 21st IEEE/12th NASA Goddard Conference on Mass Storage Systems and Technologies, **(2004)**, pp. 139–152.

[12] E. Hern´andez-Orallo and J. Vila-Carb´o, “Web server performance analysis using histogram workload models,” Computer Networks, vol. 53, no. 15,**(2009)**, pp. 2727–2739.

[13] W. Shi, R. Wright, E. Collins, and V. Karamcheti, “Workload characterization of a personalized web site and its implications for dynamic content caching,” in Proceedings of the 7th International Workshop on Web Caching and Content Distribution (WCW02), **(2002)**.

[14] A. Iamnitchi, S. Doraimani, and G. Garzoglio, “Workload characterization in a high-energy data grid and impact on resource management,” Cluster Computing, vol. 12, no. 2, **(2009)**, pp. 153–173.

[15] K. Christodoulopoulos, V. Gkamas, and E. A. Varvarigos, “Statistical analysis and modeling of jobs in a grid environment,” Journal of Grid Computing, vol. 6, no. 1, **(2008)**, pp. 77–101.

[16] E. Medernach, “Workload analysis of a cluster in a grid environment,” in Job scheduling strategies for parallel processing. Springer, **(2005)**, pp. 36–61.

[17] B. Song, C. Ernemann, and R. Yahyapour, “User group-based workload analysis and modelling,” in Cluster Computing and the Grid, 2005. CCGrid 2005. IEEE International Symposium on, vol. 2. IEEE, **(2005)**, pp. 953–961.

[18] A. K. Mishra, J. L. Hellerstein, W. Cirne, and C. R. Das, “Towards characterizing cloud backend workloads: insights from google compute clusters,” ACM SIGMETRICS Performance Evaluation Review, vol. 37, no. 4, **(2010)**, pp. 34–41.

[19] B. F. Cooper, A. Silberstein, E. Tam, R. Ramakrishnan, and R. Sears, “ Benchmarking cloud serving systems with ycsb,” in Proceedings of the 1st ACM symposium on Cloud computing. ACM, **(2010)**, pp. 143–154.

[20] Z. Ren, X. Xu, J. Wan, W. Shi, and M. Zhou, “Workload characterization on a production hadoop cluster: A case study on taobao,” in Workload Characterization (IISWC), 2012 IEEE International Symposium on. IEEE, **(2012)**, pp. 3–13.

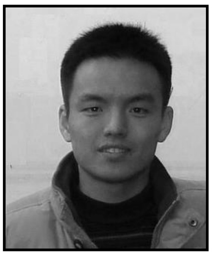
[21] Q. Jiang, “Virtual machine performance comparison of public iaas providers in china,” in Cloud Computing Congress (APCloudCC), 2012 IEEE Asia Pacific. IEEE, **(2012)**, pp. 16–19.

Authors

Zujie Ren is an associate professor of Computer Science and Technology at the Hangzhou Dianzi University in China. He is currently working at the cloud computing research institute of Hangzhou Dianzi University. He received his PhD degree in 2010. His research interests include distributed systems, cloud computing, and massive data processing.

Jinxiang Dong is a Professor in College of Computer Science and Technology, Zhejiang University, China. He was a Visiting Professor at University of Southern California from 1993 to 1994. His research interests include computer graphics,artificial intelligence, database, information processing, security, advanced manufacturing and automation, etc.

Yongjian Ren is a Professor in Hangzhou Dianzi Unviersity, China. He received his first Phd degree in 1989 from Zhejiang University, China, and the second Phd degress in 1998 from Florida Atlantic University, USA.He Now he is the CEO of Hangzhou Inforcore Data Sci-Tech Enterprise. His research interests include data storage, security, and cloud computing.



Renjie Zhou is an assistant professor in School of Computer Science and Technology at Hangzhou Dianzi University, Hangzhou, China. He received his Ph.D. degree from Harbin Engineering University, Harbin, China, in 2012. He was a visiting scholar in the Department of Electrical and Computer Engineering at the University of Massachusetts at Amherst. Currently, he is working at the cloud computing research institute of Hangzhou Dianzi University. His research interests include analysis of online social networks, and network security.

Xindong You is an associate professor of School of Computer Science and Technology, Hangzhou Dianzi University, China. She is in Key Laboratory of? Complex Systems Modeling and Simulation, Ministry of Education. Before joining Hangzhou Dianzi University, she was a PhD candidate in Northeastern University from 2002 to 2007. She received her PhD degree in 2007. Her current research areas include distributed computing, Cloud Storage, Energy Management, Data Replica Management, etc.

