

Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

I/O Performance of Virtualized Cloud Environments

Permalink

<https://escholarship.org/uc/item/67z2q3qc>

Author

Ghoshal, Devarshi

Publication Date

2011-11-12

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

I/O Performance of Virtualized Cloud Environments

Devarshi Ghoshal
Indiana University
Bloomington, IN 47405
dghoshal@cs.indiana.edu

R. Shane Canon
Lawrence Berkeley National
Lab
Berkeley, CA 94720
scanon@lbl.gov

Lavanya Ramakrishnan
Lawrence Berkeley National
Lab
Berkeley, CA 94720
lramakrishnan@lbl.gov

ABSTRACT

The scientific community is exploring the suitability of cloud infrastructure to handle High Performance Computing (HPC) applications. The goal of Magellan, a project funded through DOE ASCR, is to investigate the potential role of cloud computing to address the computing needs of the Department of Energy's Office of Science, especially for mid-range computing and data-intensive applications which are not served through existing DOE centers today. Prior work has shown that applications with significant communication or I/O tend to perform poorly in virtualized cloud environments. However, there is a limited understanding of the I/O characteristics in virtualized cloud environments. This paper will present our results in benchmarking the I/O performance over different cloud and HPC platforms to identify the major bottlenecks in existing infrastructure. We compare the I/O performance using IOR benchmark on two cloud platforms - Amazon and Magellan. We analyze the performance of different storage options available, different instance types in multiple availability zones. Finally, we perform large-scale tests in order to analyze the variability in the I/O patterns over time and region. Our results highlight the overhead and variability in I/O performance on both public and private cloud solutions. Our results will help applications decide between the different storage options enabling applications to make effective choices.

Categories and Subject Descriptors

B.8.2 [Performance Analysis and Design Aids]; C.4 [Performance of Systems]: Performance attributes

General Terms

Performance

Keywords

I/O benchmarking, cloud, virtual environments

1. INTRODUCTION

Data is a critical component of next-generation scientific processes. Scientific processes are generating and analyzing large data sets to derive scientific insights. Cloud computing technologies have largely evolved to process and store large data volumes of web and log data. In the last few years, there has been an increasing interest in evaluating the use of cloud technologies to meet the needs of scientific applications. Several groups have run both standard benchmark suites such as Linpack and NAS [3, 14, 16, 21], and network performance tests [22]. Previous results have shown that the communication-intensive applications do poorly in these environments. However there is limited understanding of the I/O performance in virtualized cloud environments. Understanding the I/O performance is critical to understand the performance of scientific applications in these environments.

The goal of the Magellan project is to evaluate the ability of cloud computing to meet the needs of DOE workloads. We have previously benchmarked various cloud platforms to quantify their performance for scientific applications [7]. In this paper, we evaluate the various cloud I/O offerings and understand their performance characteristics relative to current HPC centers where these scientific applications run.

I/O is commonly used in scientific applications: to store output from simulations for later analysis; for paging data for applications that need to process data more than what can fit in system memory; and for checkpointing to save the state of application in case of system failure. HPC systems are typically equipped with a parallel file system such as Lustre or GPFS that can stripe data across large numbers of spinning disks connected to a number of I/O servers to provide scalable bandwidth and capacity. These file systems also allow multiple clients to concurrently write to a common file while preserving consistency. On systems such as NERSC, often there are two file systems available: local and global. Local file systems accessible on a single platform typically provide the best performance whereas global file systems simplify data sharing between platforms. These filesystems are tuned for achieving high performance that is desired by these scientific applications. Thus it is critical to understand the I/O performance that can be achieved on cloud platforms in order to understand the performance impact on scientific applications that are considering these platforms.

In this paper, we evaluate a public cloud platform and the private cloud platform available on the Magellan testbed. We select Amazon as our public cloud platform since it is currently the most popular Infrastructure-as-a-Service

(IaaS) cloud offering. We benchmark three instance types - small, large and Cluster Compute, the specialized HPC offering. The Magellan virtual machine testbed runs the Eucalyptus 2.0 cloud software stack on top of KVM and is representative of current day private cloud offerings. We used IOR [6] benchmarks and a custom timed benchmark for analyzing the I/O performance on clouds. We compare the performance of different instance types, both local and block store and different availability regions on Amazon to understand the spectrum of I/O performance. Specifically, we make the following contributions:

- We compare the I/O performance of various storage options in virtual machines and instance types provided through private cloud software stacks as well as commercial providers such as Amazon,
- We also study the variability of the performance over multiple runs and multiple time periods on Amazon.

The rest of this paper is organized as follows. In Section 2 we discuss related work. We present our methodology for evaluation in Section 3 and our performance results in Section 4. We discuss the results in Section 5 and present our conclusions in Section 6.

2. RELATED WORK

Several groups have looked at the feasibility of cloud computing and benchmarked various aspects of clouds. However there has been no previous work analyzing the performance of I/O in virtualized cloud environments. Here we summarize related work that looks at the performance aspects of cloud computing.

Various groups have evaluated cloud offering such as Amazon EC2 and hypervisors such as Xen and KVM [13]. Previous work has focused on using standard benchmarks such as Linpack, NAS Parallel Benchmarks and other microbenchmarks [3, 14, 16, 7, 19]. The performance of Xen and KVM environments for scientific applications has been studied [4, 15, 22, 24]

Various efforts have proposed benchmark applications and appropriate parameters for measuring the I/O performance [1, 20, 23]. But there have been no previous work showing any results benchmarking the I/O performance on virtualized environment in clouds.

Application groups have looked at the performance and cost of porting specific application pipelines to Amazon EC2 cloud [2, 5, 10, 11, 9, 17]. In previous work we examined the usefulness of cloud computing for e-Science applications [12, 18].

In this paper we evaluate the performance of I/O that impact the overall performance of the applications running in virtual machines on cloud.

3. METHODOLOGY

Previous work has shown that virtualized cloud environments impact the performance of tightly coupled applications [3, 14, 16, 7, 19]. However studies conducted on Amazon EC2 provide limited understanding of the causes of the performance decrease due to the black box nature of these cloud services. We use the Magellan testbed to further characterize the virtual machine performance since we have control over its environment. The performance impact has

been suspected to be from I/O and networking aspects of these virtualized resources. In this paper, we measure the I/O performance on a range of Amazon resources and the Magellan testbed to understand the impact of various storage options on the performance of applications.

We measure I/O performance using standard and custom benchmarks on the cloud platforms mentioned above over different dimensions. We use the IOR (Interleaved or Random) benchmark to compare the I/O performance across all platforms. In addition, we developed a timed I/O benchmark that records the I/O performance over a period of time at pre-determined intervals to assess variability. We also measure the performance of various storage options on virtual machines and on Amazon record the performance across two availability regions.

3.1 Benchmarks

We executed two sets of benchmarks to analyze and gather statistics of I/O performance on different virtualized cloud environments: a) IOR b) Timed Benchmark

3.1.1 IOR

IOR (Interleaved or Random) is a popular benchmarking tool for understanding the I/O performance of high-performance parallel file systems on HPC systems. IOR can create a variety of IO patterns using several IO interfaces (POSIX, MPI-IO, HDF5) with a variable number of clients. IOR provides the capability to set both block and transfer sizes of the files to be read and written. We experimented with the block-size and transfer-size and all results reported in this paper were executed for a block-size of 100G and a transfer-size of 2M. The I/O type can also be set through a flag to ensure that the I/O operations are all direct, instead of the default buffered I/O helping us tease out the performance impact from buffering.

3.1.2 Timed Benchmark

The Timed benchmark is a custom benchmark designed to measure the I/O performance of a file-system over a period of time. The two major factors controlling the timed benchmark, apart from the block and transfer size, are the total time duration and the frequency interval at which the performance is measured. The timed benchmark has been tested to give equivalent results to that of IOR, with very little overhead. This benchmark helps in understanding the variability of I/O performance on different instance-types over time.

3.2 Machine Description

We perform our experiments on Magellan and Amazon instances. Amazon is the leading cloud provider for IaaS-style cloud computing today. Magellan is representative of private cloud computing platforms.

3.2.1 Magellan

All the experiments were performed using the Magellan compute resources at NERSC. Magellan is a 720 node IBM iDataPlex cluster. Each node has two quad-core Intel Nehalem processors running at 2.67 GHz, 24 GB of RAM and two network connections: a single Quad Data Rate (QDR) Infiniband network connection and a GiB ethernet connector. The IB network is locally a fat-tree with a global 2D-mesh.

Table 1: Amazon EC2 Instance Types- Architecture. Source: <http://aws.amazon.com/ec2/instance-types/>

Inst-type	API name	CPU Family	Expected I/O Performance	EC2 Compute Units	Memory (GB)	Local Storage (GB)
small	m1.small	Intel Xeon E5430	Moderate	1	1.7	160
xlarge	c1.xlarge	Intel Xeon E5410	High	20	7	1690
cc	cc1.4xlarge	2 x Intel Xeon X5570	Very High	33	23	1690

Our virtual machine environment is based on Eucalyptus 2.0.5, an open source software platform that allows organizations to clouds. Our Eucalyptus installation uses Kernel-based Virtual Machines (KVM) as a hypervisor. We modified Eucalyptus to use virtio for disk access. Virtio is a Linux standard for network and disk device drivers which enables virtual machines to get high performance for network and disk operations. We use the KVM option for the emulated e1000 NIC for the network. All codes were compiled with GCC version 4.1.2 and used version 1.4.2 of OpenMPI.

All tests were performed on instance type c1.xlarge that is configured with 8 CPUs/20G memory/20G disk. The guest OS is CentOS release 5.5 (Linux kernel 2.6.28-11). For the MPI-IO tests, we created a virtual cluster where the head node mounts a block store volume that has all the application binaries and data. The block store volume is mounted on the other virtual machines via NFS. All the virtual machine communication traffic goes over the Ethernet network.

Apart from the VM setup, the tests were also executed to benchmark against a high-speed NERSC file-system, Global Scratch. This file system uses IBM's GPFS and has a peak performance of approximately 15 GB/sec. This result was used to understand the I/O performance overhead of virtualized environments compared to what is accessible today to scientific applications in HPC centers.

3.2.2 Amazon

The tests on Amazon's EC2 cloud were performed on three instance types: m1.small (Small Instance), c1.xlarge (High-CPU Extra Large Instance), and cc1.4xlarge (Cluster Compute Quadruple Extra Large Instance). These types were carefully selected to cover the spectrum of resource types (and hence I/O performance) provided by Amazon EC2. The guest OS for the small and xlarge instances were Fedora 8 (Linux kernel 2.6.18-xenU-ec2-v1.0), 32 and 64-bit respectively and for the cluster-compute (cc) instance it was CentOS release 5.4. The configuration of Amazon instances is summarized in Table 1. The compute power in Amazon EC2 is calculated in terms of EC2 Compute Unit which provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor. The connectivity of the cc instances are over a 10 Gigabit Ethernet network to provide low latency and high bandwidth to the instances within a cluster. The table shows the advertised I/O performance expected for each of these instances.

In all cases we ran the benchmarks three times and report the best result. In most cases the three measurements were in close agreement with each other. In other cases, we study the variability aspects in greater detail through our timed benchmark.

3.3 Evaluation Criteria

In our evaluation we try to cover the breadth of parameters that can affect the I/O performance on virtualized environments.

Direct and Buffered I/O. We measure both buffered I/O and direct I/O to understand the effects of buffering caches. We identify how the I/O operations (both read and write) perform based on whether the I/O is buffered or direct. Buffered I/O uses buffer caches to prefetch and store the data in anticipation of the application asking for it resulting in performance gains. Direct I/O does not use caches and hence the memory is freed as soon as the I/O is complete. Thus, direct I/O reduces the CPU overhead associated with I/O by eliminating the data copy between the kernel buffer and the user buffer.

Virtual Machine Instance Types. The commercial cloud vendors provide instances with different machine configurations. As shown in Table 1, the I/O performance advertised by Amazon on each of these instances is also different. This makes it important to analyze the I/O performance to determine the cost-benefit model enabling applications to make the appropriate choice based on the I/O performance they might desire. We measure the performance on three different instance types.

Storage Devices. Virtual machines available on Amazon and through private cloud software solutions such as Eucalyptus provide multiple storage options. Virtual machines have access to non-persistent local disk on the instance. In addition, an instance might mount a block level storage volume that persists independently from the life of an instance. Amazon Elastic Block Store is expected to be highly available, highly reliable storage volume suitable for applications requiring a database, file system, or access to raw block level storage. Additionally, Amazon S3 is an Internet simple storage service that offers a highly-scalable, reliable, and low-latency data storage infrastructure through a web service interface. We benchmark the I/O performance on local ephemeral devices and EBS volumes. We do not consider Amazon S3 in our benchmarking since it has limited applicability to our applications. Thus, our benchmarking effort analyzes the effects of shared vs dedicated and local vs networked disk storage in virtual machines.

Availability Regions. Amazon provides the ability to place instances in multiple locations or regions that are geographically dispersed. In the US, Amazon provides services in two regions - US East (Virginia) and US West (Northern California) with slightly different price points. Cluster compute instances are only available currently in the US East region.

Shared File-system. MPI is used extensively in high-performance computations where multiple processes write onto a shared file-system. An important analysis would be to measure the performance of an application running mul-

multiple instances and sharing a file-system. Scientific applications that require a shared file system often use an EBS volume shared across instances through NFS.

Time of run. Cloud virtual instances often share the underlying resources with other virtual instances thus resulting in variable performance. Hence, it becomes important to analyze the I/O patterns over a period of time. We perform a study over a few days and over a large number of instances to understand this variability.

4. PERFORMANCE RESULTS

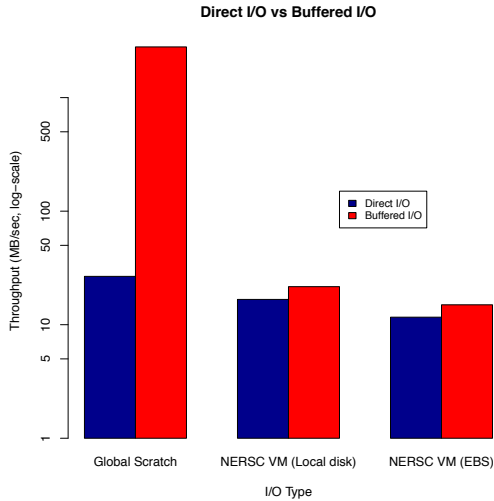


Figure 1: IOR Results: Comparison of Direct vs Buffered I/O on NERSC systems

We use the selected benchmarks and configuration parameters and conducted multiple tests to capture the variation and trends of I/O performance on the Magellan testbed and the Amazon cloud infrastructure. Specifically, we compare the effects of buffer caching and IOR results across all platforms and we evaluate the effects of MPI-IO in these virtualized environments. We also conducted large-scale tests and 24 hour runs to understand the variability experienced on these platforms.

4.1 IOR Results

In this section we report our results from the IOR benchmark. For all tests a block size of 100G and a transfer size of 2M were used.

4.1.1 Buffer Caching

Figure 1 shows the performance of direct I/O and buffered I/O on different configurations on the Magellan testbed. Buffer caching does not seem to be enabled on Amazon instances. As expected, the global-scratch filesystem at NERSC shows significant performance gains with buffer-caching enabled. High performance file-systems such as GPFS on NERSC are optimized to provide high performance using buffer-caching, etc. However, there was only a very small increase in performance on the NERSC VMs from buffering. Thus, buffer caching seems to have minimal or no impact on virtualized resources. The remaining results in our evaluation only considers direct I/O.

4.1.2 Comparison Across All Platforms

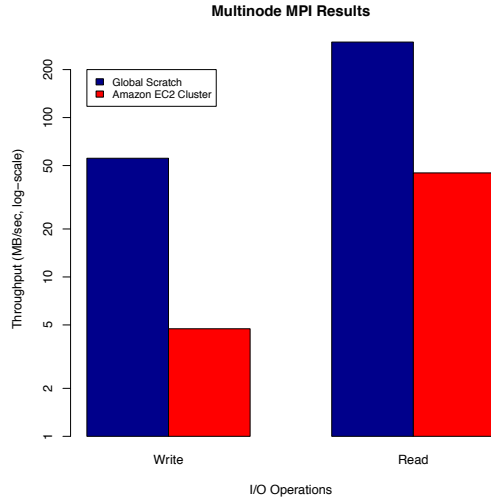


Figure 4: Multinode MPI Shared Filesystem Results on NERSC global scratch and Amazon Cluster Compute instances

Figure 2 shows a comparison of I/O performance on all the virtualized cloud environments. The summarized performance results show the impact of network and disk bandwidth and latencies on the I/O performance. The I/O performance on Amazon instances perform better than the VMs at NERSC. These differences could be from the differences in underlying hardware as well as limited performance tuning done on the NERSC virtual machine systems. Note that this performance is significantly lower than what is achievable on NERSC scratch file systems ($> 2500MB/s$).

The local disk on the instance performs better than the EBS volumes on the VMs at NERSC. On Amazon, the I/O performance on the local ephemeral disks on the small instances is better than writes to the EBS volumes over the network. This may be due to the low network bandwidth for small instances. For the extra-large instances the EBS volumes show better performance than that of the local disks. This may be attributed to the fact that there is higher network bandwidth associated with these instance-types possibly resulting in better I/O operations on the networked EBS volumes. However the difference is minimal. Clearly, EBS performance is better with cluster-compute instances than the other two instance-types (for EBS volumes) which is possibly due to the availability of 10 Gigabit Ethernet network on cc instances. The read operations show similar trends except for the cluster-compute instances that shows significantly lower than expected. I/O device virtualization requires a virtual machine monitor (VMM) or a privileged VM for every I/O operation and hence, there is an additional level of abstraction which degrades the performance in VMs. We also observed a fair amount of variability in the VM performance that we try to understand better through the timed benchmark.

4.1.3 Effect of Region

As described in the previous sections, Amazon’s instances are located in various places and are divided into regions.

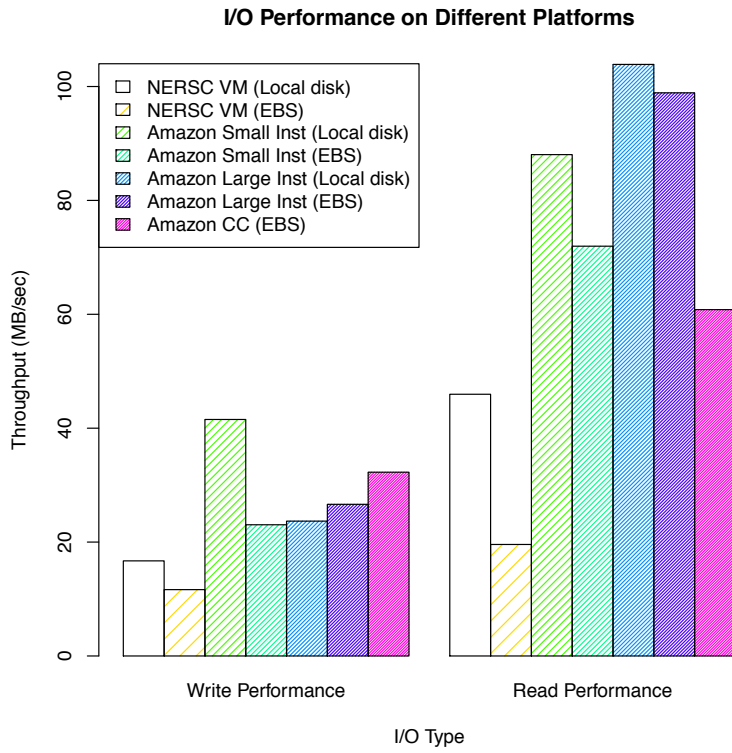


Figure 2: IOR Results: I/O Performance on All Platforms

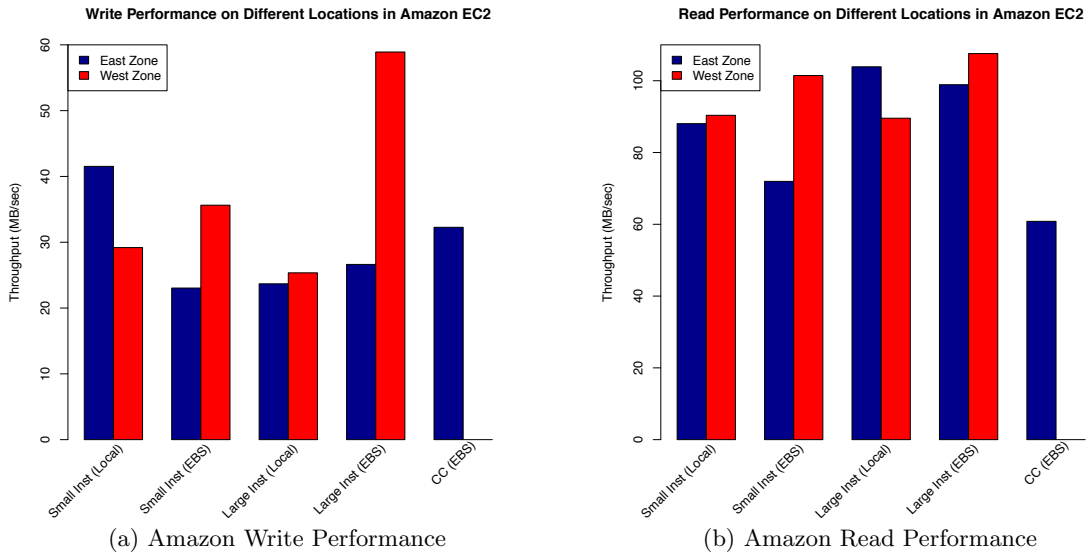
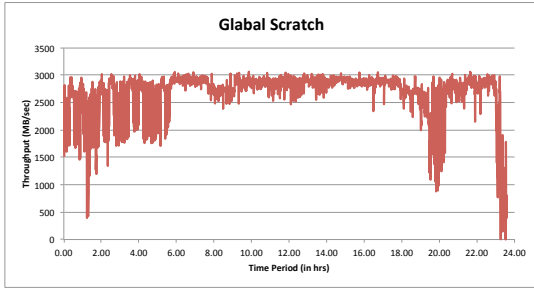
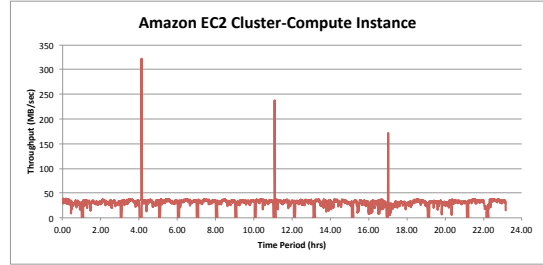


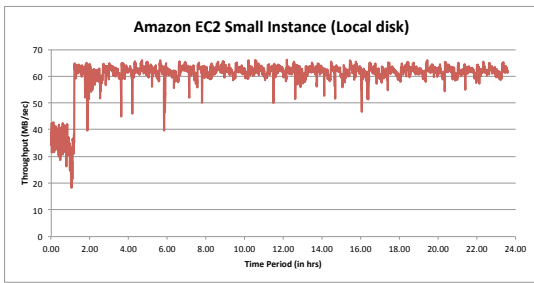
Figure 3: IOR Results: Comparison of Amazon platforms



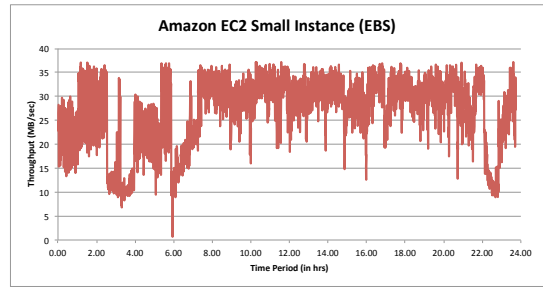
(a) Global Scratch (Buffer caching enabled)



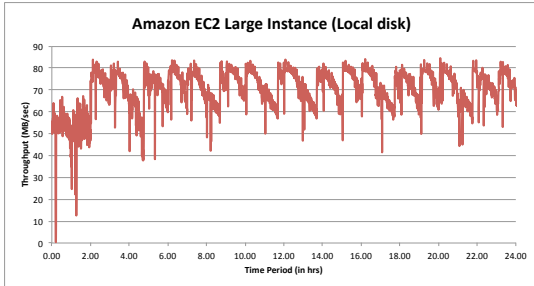
(b) Amazon EC2 Cluster-Compute Instance



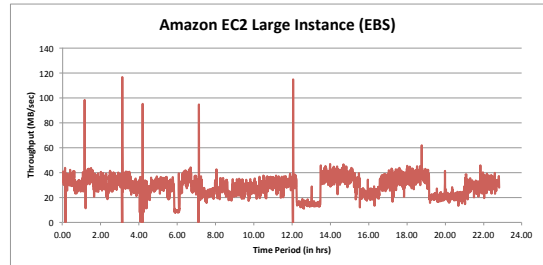
(c) Amazon EC2 Small Instance (Local disk)



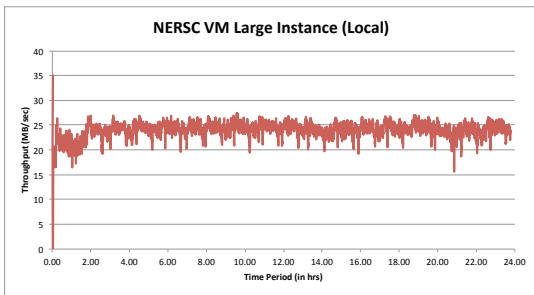
(d) Amazon EC2 Small Instance (EBS)



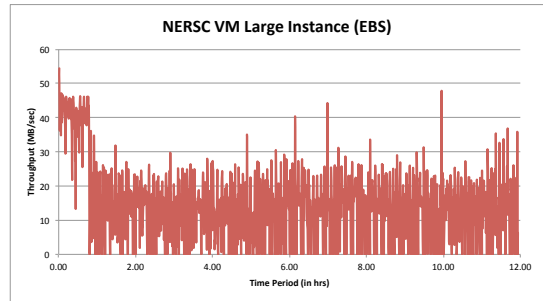
(e) Amazon EC2 Large Instance (Local disk)



(f) Amazon EC2 Large Instance (EBS)



(g) NERSC VM (Local disk)



(h) NERSC VM (EBS)

Figure 5: Timed Benchmark Results

The graphs in Figure 3a and Figure 3b show performance variations due to the location of instances in different regions. In general, the instances on the west zone showed slightly higher performance than the east zone. Note that cluster compute instances are not available in the west zone. In two cases, the east zone showed better performance than west zone: in writes on small instance (local) and reads in large instance (local disk). This can be attributed largely to the variability we see in performance in Amazon instances that we discuss in the next section.

4.1.4 Multi-node

A large number of high-performance scientific applications use MPI, performance analysis on virtualized cloud environments when a shared file-system is used by multiple processes is important to measure. Figure 4 depicts the performance comparison between the NERSC Global Scratch filesystem and a shared EBS volume mounted with NFS across ten nodes on an Amazon EC2 cluster compute instance type. NFS is not a parallel file system, however that is routinely used in virtual machine environments where there is need for a shared file system [7, 8]. If multiple instances executing MPI tasks share a single EBS volume to perform I/O, there is high resource contention over limited network which degrades the overall performance. On the other hand, the NERSC global scratch is a parallel file system that is designed and optimized for such concurrent access. Hence, the nodes in an HPC cluster over a high-speed shared filesystem perform better than the instances on the cloud as shown in Figure 4.

4.2 Performance Variability

It is critical to understand the I/O performance variability over time and instances. Thus we used our timed benchmark to measure the performance on instances over a 24 hour time period and over a large number of instances over one hour periods.

4.2.1 24 hour results

The 24 hour cycle tests enable us to study potential changes in performance from typical usage patterns such as peak vs non-peak hour usage of cloud services. The results in Figure 5 shows the timed benchmark results on Amazon and NERSC systems. Figure 5a shows I/O performance on the Global Scratch filesystem at NERSC. The benchmark was run for a complete 24 hour duration starting at 12 noon PDT. Since, the shared filesystem is being accessed by many users during the daytime, the performance degrades at both end of the spectrum. After the peak processing hours, there is significant improvement in I/O performance due to limited contention for the shared resources. The I/O performance graphs for the timed benchmark on various virtual environments show more periodic patterns with occasional spikes. The reason for the occasional spikes is not known at the current time and will be investigated in future work. Figure 5g shows a 24-hour run of a test starting at 12 midnight PDT. The results suggest that there is more consistent I/O performance on local disk of a VM instance. The graphs showing the results for EC2 instances clearly signify that the I/O performance doesn't follow a specific pattern over time but most of them are either consistent or periodic over time. Figure 5c and Figure 5e show the I/O performance on local disks for different instances. The occasional

drop in performance may be due to the sharing of underlying resources. On the other hand, Figure 5d, Figure 5f and Figure 5b show the performance graphs on EBS volumes over the small, large and cluster-compute instances respectively. EBS volumes are attached to every instance separately over the network and hence there is no sharing of disks or volumes. However, our results show that there are certain spikes in the graphs which may be attributed to the network traffic or the behavior of the hypervisor.

4.2.2 Large Scale Tests

Large-scale tests were run to gauge the variability of performance for multiple instances over a certain time-period. Each test was conducted on 25 Amazon EC2 small instances on local disk in each region (US-East and US-West), in the same time period, over different days. All the tests were executed on a peak usage period of cloud resources on both East and West zones. We selected the time-period between 11:00 AM and 2:00 PM PDT (i.e., 2:00 PM to 5:00 PM EDT). The timed benchmark was used to capture readings at regular intervals for one hour to understand the variability among instances. The histogram plots and the Kernel Density plots for the tests are shown in Figure 6. The histogram plots show the frequency distribution of throughput performance and the density plots using the Gaussian kernel show the corresponding probability density function for the throughput variance.

These large scale tests reveal important characteristics of I/O performance on the US-East and US-West regions. Mean performance in the East zone has a throughput of 30 MB/sec. The area of the graph in Figure 6c with the throughput variance between 10 MB/sec - 20 MB/sec and 35 MB/sec - 45 MB/sec fall under less than one standard deviation from the mean. The peak performance in East zone is 60 MB/sec. The performance in the West zone shown in Figure 6d varies significantly with the peak performance going upto as high as 88 MB/sec. The mean performance in the West zone is between 10 MB/sec - 20 MB/sec which shows a high standard deviation in the West zone.

5. DISCUSSION

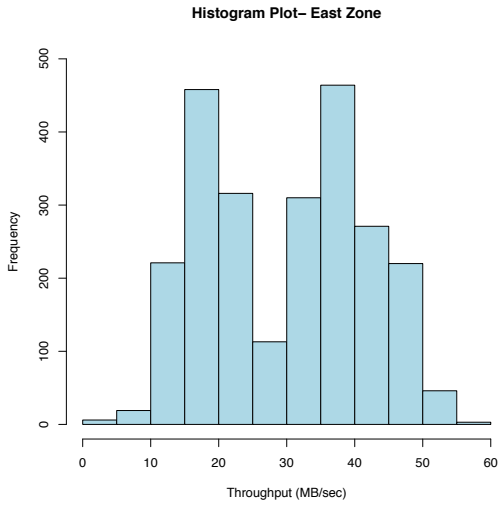
We performed an extensive I/O benchmarking study comparing various cloud platforms and service offerings. The performance on virtual hosts tends to show a fair amount of overhead and variability due to the contention and sharing of underlying resources.

5.1 Buffer Caching

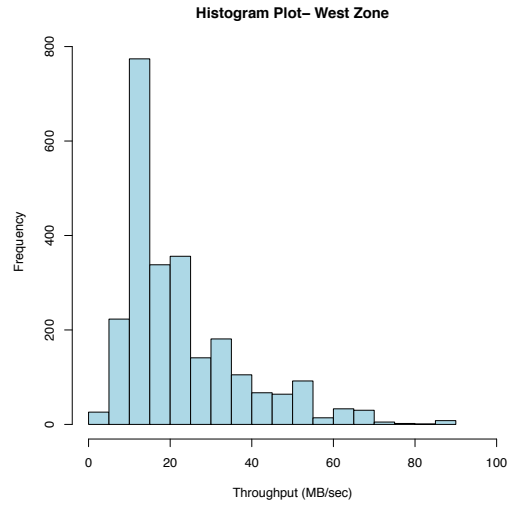
Based on our testing, buffer caching does not seem to be enabled on virtualized resources. Scientific applications running in HPC centers are able to use high performance filesystems such as GPFS that show significantly higher peak performance than what is seen in today's virtual environments. This can have a significant impact on overall application performance for I/O intensive applications.

5.2 Storage Options on VM

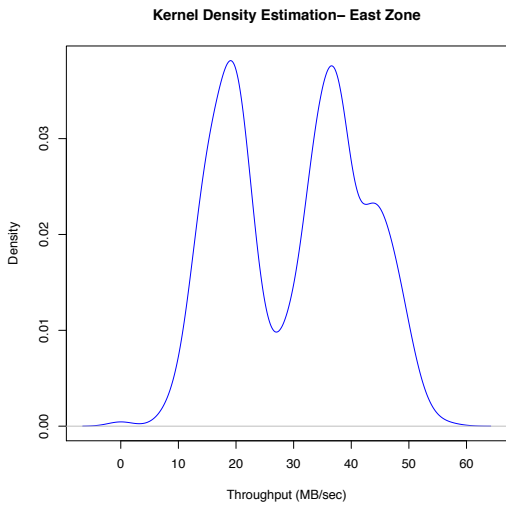
A virtual machine has ephemeral local disk and has the option to mount an elastic block storage volume. Typically, the performance of the local disk tends to be slightly higher than the EBS corresponding volumes. This is especially true in the Amazon small instances that are bandwidth limited. At least during the duration of our tests we noticed that



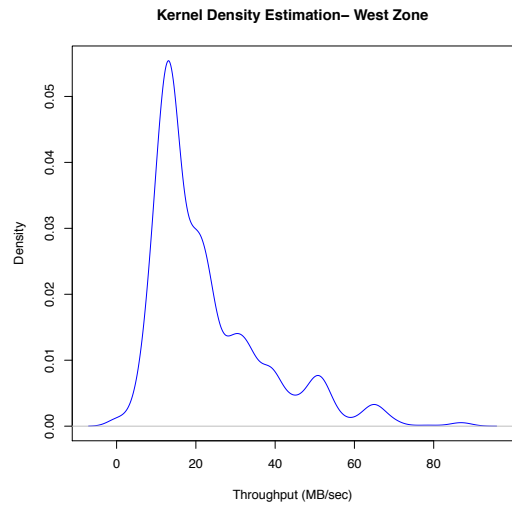
(a) Large Scale Test on Amazon US-East Region



(b) Large Scale Test on Amazon US-West Region



(c) Large Scale Test on Amazon US-West Region



(d) Large Scale Test on Amazon US-West Region

Figure 6: Large Scale Test Results Histogram and Kernel Density Plots

the larger instance types showed lesser advantage with the local disks possibly due to the increased network bandwidth available in the large instances. However applications might need to do a performance-cost analysis of local disk vs EBS for their particular application [8].

5.3 Instance Type

The advertised I/O performance on the instance type is expected to get better with larger instances. However in our limited testing, we encountered situations where we were able to get better I/O performance on the small instance local disk than the large instance. However our tests also show that the small instances do tend to show a fair amount of variability and hence more extensive testing might be needed to capture these behaviors over time. However, the EBS performance definitely seemed to improve with the instance types possibly due to the better network bandwidth available to the large and the cc instances.

5.4 Availability Regions

We observed that the west zone performance was better than the performance of the east zone. The west zone VMs on Amazon have slightly higher price points possibly explaining the better performance. However our large-scale tests also show that the west zone has a higher standard deviation than the east zone.

6. CONCLUSIONS AND FUTURE WORK

The I/O performance results clearly highlight that I/O can be one of the causes for performance bottleneck on virtualized cloud environments. Performance in VMs is lower than on physical machines, which can be attributed to an additional level of abstraction between the VM and the hardware. Also, it is evident from the results that local disks perform better than the block store volumes. Block store volumes can be mounted on multiple instances and used as a shared filesystem to run MPI programs but the performance degrades significantly. Thus these may not be suitable for all the MPI applications that compose the HPC workload. Our large scale tests also capture the variability associated with virtual machine performance on Amazon.

Scientific users planning to run in cloud environments need to understand the effects of the I/O performance for their application I/O access patterns. Applications need to account for the variability in performance in their application design. Additionally, there is a need to investigate mechanisms to provide high-performance file systems in virtual machines for scientific applications that might require it.

We performed an exhaustive study comparing the performance of Amazon and the performance on the Magellan testbed. However, more work is needed to understand the effect of different hypervisors and performance tuning at lower levels. Additionally, additional tests will be needed to understand the variability experienced on virtual systems under different workloads and instance types.

7. ACKNOWLEDGEMENTS

This work was funded in part by the Advanced Scientific Computing Research (ASCR) in the DOE Office of Science under contract number DE-AC02-05CH11231. The authors would like to thank Iwona Sakrejda and Tina Declerck for

system support and Krishna Muriki for guidance on using Amazon resources.

8. REFERENCES

- [1] J. C. Brustoloni and P. Steenkiste. Effects of buffering semantics on i/o performance. In *Proceedings of the Second USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, pages 277–291, 1996.
- [2] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good. The cost of doing science on the cloud: the montage example. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, pages 1–12. IEEE Press, 2008.
- [3] C. Evangelinos and C. Hill. Cloud Computing for parallel Scientific HPC Applications: Feasibility of running Coupled Atmosphere-Ocean Climate Models on Amazon’s EC2. *ratio*, 2(2.40):2–34, 2008.
- [4] M. Fenn, M. Murphy, and S. Goasguen. A study of a KVM-based cluster for grid computing. In *Proceedings of the 47th Annual Southeast Regional Conference*, pages 1–6. ACM, 2009.
- [5] S. Hazelhurst. Scientific computing using virtual high-performance computing: a case study using the Amazon elastic computing cloud. In *Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology*, pages 94–103. ACM, 2008.
- [6] IOR Benchmark. <http://dl.acm.org/citation.cfm?id=1413413>.
- [7] K. Jackson, L. Ramakrishnan, K. Muriki, S. Canon, S. Cholia, J. Shalf, H. Wasserman, and N. Wright. Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud. In *2nd IEEE International Conference on Cloud Computing Technology and Science*, 2010.
- [8] K. R. Jackson, K. Muriki, L. Ramakrishnan, K. J. Runge, and R. C. Thomas. Performance and cost analysis of the supernova factory on the amazon aws cloud. *Scientific Programming*, 19(2-3):107–119, 2011.
- [9] K. Keahey. Cloud Computing for Science. In *Proceedings of the 21st International Conference on Scientific and Statistical Database Management*, page 478. Springer-Verlag, 2009.
- [10] K. Keahey, R. Figueiredo, J. Fortes, T. Freeman, and M. Tsugawa. Science clouds: Early experiences in cloud computing for scientific applications. *Cloud Computing and Applications*, 2008, 2008.
- [11] K. Keahey, T. Freeman, J. Lauret, and D. Olson. Virtual workspaces for scientific applications. In *Journal of Physics: Conference Series*, volume 78, page 012038. Institute of Physics Publishing, 2007.
- [12] J. Li, D. Agarwal, M. Humphrey, C. van Ingen, K. Jackson, and Y. Ryu. eScience in the Cloud: A MODIS Satellite Data Reprojection and Reduction Pipeline in the Windows Azure Platform. In *Proceedings of the 24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)*, Atlanta, GA, April 19-23, 2010.

- [13] A. Menon, J. R. Santos, Y. Turner, G. J. Janakiraman, and W. Zwaenepoel. Diagnosing performance overheads in the xen virtual machine environment. In *Proceedings of the 1st ACM/USENIX international conference on Virtual execution environments*, VEE '05, pages 13–23, New York, NY, USA, 2005. ACM.
- [14] J. Napper and P. Bientinesi. Can cloud computing reach the top500? In *Proceedings of the combined workshops on UnConventional high performance computing workshop plus memory access workshop*, pages 17–20. ACM, 2009.
- [15] Nathan Regola and Jean Christophe Ducom. Recommendations for Virtualization Technologies in High Performance Computing. In *2nd IEEE International Conference on Cloud Computing Technology and Science*, 2010.
- [16] S. Ostermann, A. Iosup, N. Yigitbasi, R. Prodan, T. Fahringer, and D. Epema. An early performance analysis of cloud computing services for scientific computing. *Delft University of Technology, Tech. Rep.*, 2008.
- [17] M. Palankar, A. Iamnitchi, M. Ripeanu, and S. Garfinkel. Amazon S3 for science grids: a viable solution? In *Proceedings of the 2008 international workshop on Data-aware distributed computing*, pages 55–64. ACM, 2008.
- [18] L. Ramakrishnan, K. R. Jackson, S. Canon, S. Cholia, and J. Shalf. Defining Future Platform Requirements for e-Science Clouds. In *Proceedings of the ACM Symposium on Cloud Computing (SoCCi)*. ACM, 2010.
- [19] J. Rehr, F. Vila, J. Gardner, L. Svec, and M. Prange. Scientific computing in the cloud. *Computing in Science and Engineering*, 99(PrePrints), 2010.
- [20] H. Shan, K. Antypas, and J. Shalf. Characterizing and predicting the i/o performance of hpc applications using a parameterized synthetic benchmark. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, SC '08, pages 42:1–42:12, Piscataway, NJ, USA, 2008. IEEE Press.
- [21] E. Walker. Benchmarking amazon EC2 for high-performance scientific computing. *USENIX Login*, 33(5):18–23, 2008.
- [22] G. Wang and T. E. Ng. The impact of virtualization on network performance of amazon ec2 data center. In *Proceedings of IEEE INFOCOM*, 2010.
- [23] B. Wolman and T. M. Olson. Iobench: a system independent io benchmark. *SIGARCH Comput. Archit. News*, 17:55–70, September 1989.
- [24] L. Youseff, R. Wolski, B. Gorda, and C. Krintz. Evaluating the performance impact of xen on mpi and process execution for hpc systems. In *Proceedings of the 2nd International Workshop on Virtualization Technology in Distributed computing*, page 1. IEEE Computer Society, 2006.