Evaluating Interconnect and Virtualization Performance for High Performance Computing

Lavanya Ramakrishnan R. Shane Canon Krishna Muriki
Iwona Sakrejda Nicholas J. Wright
NERSC, CRD, IT, Lawrence Berkeley National Lab
Berkeley, CA 94720
{lramakrishnan,scanon,kmuriki,isakrejda,njwright}@lbl.gov

ABSTRACT

Scientists are increasingly considering cloud computing platforms to satisfy their computational needs. Previous work has shown that virtualized cloud environments can have significant performance impact. However there is still a limited understanding of the overheads and the type of applications that might do well in these environments. In this paper, we detail benchmarking results that characterize the virtualization overhead and examine the performance of various interconnect technologies. Our results show that virtualization and less capable interconnect technologies can have a significant impact upon performance of typical HPC applications. We also evaluate the performance of the Amazon Cluster Compute instances and show that it performs approximately equivalently to a 10G Ethernet cluster at low core counts.

Categories and Subject Descriptors

B.8.2 [**Performance Analysis and Design Aids**]; C.1.4 [**Parallel Architectures**]

General Terms

Performance

Keywords

Performance, cloud, virtualization, interconnects

1. INTRODUCTION

The Infrastructure-as-a-Service (IaaS) type of cloud computing environment today typically consists of a group of machines connected over an ethernet connection, often running some kind of virtualization software. Amazon EC2 is a leading commercial example of this type of cloud service that is available today. Early studies have benchmarked scientific applications on commercial platforms [1, 2]. These studies show that tightly coupled applications perform poorly in virtualized environments such as Amazon EC2. However there is limited understanding about the type of cloud environments (e.g., inteconnects, machine configuration etc) that might benefit science applications. HPC resource providers are interested in understanding the performance tradeoffs of using cloud platform solutions such as virtualization in supercomputing centers. The Magellan project, funded by the DOE, is evaluating cloud solutions to understand their suitability for scientific computing.

PMBS'11, November 13th, 2011, Seattle, Washington, USA.
ACM 978-1-4503-1102-1/11/11.

Today's cloud environments are typically based on commodity hardware and use TCP over Ethernet to connect the nodes. However, compute resources found at HPC centers typically have much higher performing interconnects, such as InfiniBand, that provide higher bandwidth and lower latency. Thus the networking performance of current cloud environments is quite different from that at HPC centers.

In this paper, we address the question of what must virtualized cloud environments provide in order to be beneficial for HPC applications. We use standard benchmarks such as HPC and a subset of benchmarks from the NERSC6 benchmark suite based on our earlier work to capture the application behavior in virtualized environments [1]. Specifically, we make the following contributions in this paper,

- We evaluate the performance impact of each of the layers in the hierarchy of protocols encountered by scientific applications, i.e., InfiniBand, TCP over InfiniBand, TCP over Ethernet and TCP over 10 G and 1G Ethernet. Our results show that the performance differences between the interconnects increases with increasing network contention. This leads to increased time to solution for scientific applications when many nodes are simultaneously communicating.
- We show that using virtualization has significant impact upon application performance. For example, at the lowest count the virtualization layer causes a 1.6× and 2.5× performance decrease for PARATEC and MILC, respectively, compared to the same hardware configuration without using virtualization.
- We evaluate the performance of Amazon Cluster Compute instances, the specialized HPC offering. Our results show that at low core counts the performance is comparable to a non-virtualized 10G Ethernet cluster. As the core count increases and the network contention increases, the results drop below those of a 10G cluster.

2. RESULTS

Figure 1 shows the performance of PARATEC and MILC benchmarks [1] for each of the different interconnect technologies and protocols. The performance for each is shown relative to the Infiniband (IB) performance at the lowest core count. This allows us to show both the relative performance

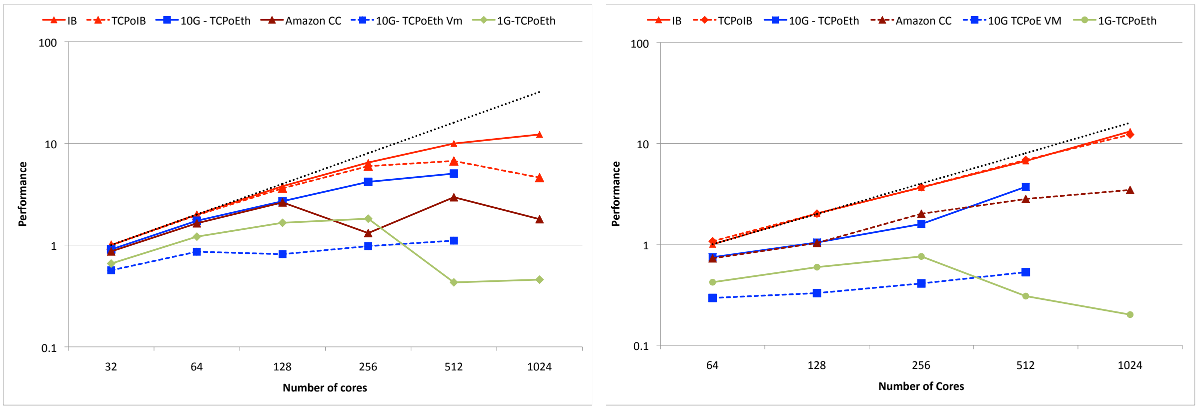

Figure 1: Performance of a) PARATEC and b) MILC plotted on a log-log scale as a function of core count using several different interconnect technologies and/or protocols. The dotted line represents ideal scaling based upon the IB performance at the lowest core count.

differences between the technologies as well as how they affect parallel scaling. Figure 1a shows that for PARATEC the Infiniband results are the fastest at all core counts. The change in protocol from IB to TCP, represented by the TCPoIB line, only affects performance at the higher core counts, 512 and 1024, where it makes the performance 1.5× and 2.5× slower than the native IB, respectively. The 10G TCPoEth performance is within 10% of Infiniband at 32 cores but drops to about 2× slower at 512 cores. As expected 1G TCPoEth shows worse performance than the 10G. The 10G TCPoEth VM results are by far the poorest. They show that the overhead of virtualization, at least as configured on the Magallen cluster, is significant. At 32 cores the performance is 1.75× worse than IB. As the core count increases the performance does not increase and the 10G VM line is not parallel to the 10G one, which indicates that the performance degradation due to virtualization is increasing as a function of core count. The Amazon CC results mirror the 10G TCPoEth ones at low core counts but at 256 cores and above the performance of the Amazon CC starts decreasing, as compared to the 10G TCPoEth one, presumably either because the performance of the 10G switch on the Magellan cluster is greater than that on the Amazon cluster or due to virtualization overheads at higher core counts similar to the 10G VMs on Magellan. Figure 1b shows that for MILC, Infiniband is also the fastest at all core counts. In this case the TCPoIB results are indistinguishable from the native IB ones. The 10G TCPoEth is minimally 35% slower at all core counts. The performance of 1G TCPoEth is about 2.5× slower than IB at 64 cores and is about 4.8× slower at 256 cores. Again, above 256 cores the interconnect simply cannot keep up. The 10G TCPoEth VM results are by far the poorest. The Amazon CC numbers almost exactly mirror the 10G TCPoEth ones because at these core counts the MILC application is scaling better.

3. CONCLUSIONS

Virtualized cloud computing environments promise to be useful for scientific application that need customized software environments. However, earlier studies have shown that virtualization has a significant performance impact for scientific applications. In this paper we analyzed the performance of a number of different interconnect technologies to understand the performance tradeoffs. Our results show that while the differences between the interconnects is small at lower core counts, the impact is significant even at the midrange size of problems of 256 to 1024. While the bandwidth is slightly impacted at higher concurrency, the latency takes a much higher hit. Scientific applications tend to run at higher concurrencies to achieve the best time to solution. Thus to serve the needs of these applications, we need "good" interconnects.

In addition to the networking fabric and the protocol, the virtualization software stack imposes an additional overhead to the performance of the applications. This overhead also is impacted by the communication pattern of the application and the concurrency of the run. The higher bound on the performance on virtual machines today is what is achievable with TCP over ethernet clusters. Our experiments show that the availability of InfiniBand interconnects on virtual machines would boost the performance of scientific applications by reducing the communication overheads. The full paper can be found in [3].

4. ACKNOWLEDGEMENTS

This work was funded in part by the Advanced Scientific Computing Research (ASCR) in the DOE Office of Science under contract number DE-C02-05CH11231.

5. REFERENCES

[1] Jackson, K., Ramakrishnan, L., Muriki, K., Canon, S., Cholia, S., Shalf, J., Wasserman, H., and Wright, N. Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud. In *2nd IEEE Intl, Conf. on Cloud Computing Technology and Science* (2010).

[2] Napper, J., and Bientinesi, P. Can cloud computing reach the top500? In *Proceedings of the combined workshops on UnConventional high performance computing workshop plus memory access workshop* (2009), ACM, pp. 17–20.

[3] Ramakrishnan, L., Canon, R.C., Muriki, K., Sakrejda, I. and Wright, N.J. Evaluating Interconnect and Virtualization Performance for High Performance Computing. *SIGMETRICS Performance Evaluation Review 40(2)* (2012)