Impact of Cloud Computing on Distributed Memory Models

Daniel G. Bueno*, Erik S. Candeia*, Bruno P. Masquio*, Douglas N. de Oliveira*, Caio V. Ortu*, Mauro D. L. Tosi*

Eunice E. Santos†, John Korah†, Vairavan Murugappan†

Motivation

In recent years, cloud computing has become a cost effective and popular platform for scientific applications. Most of these applications utilize distributed/parallel algorithms. Therefore, understanding the performance of these parallel algorithms on cloud platform becomes critical.

- Parallel models (such as LogP, Bulk-Synchronous Parallel) are primarily designed for a cluster environment and do not consider virtualization or resource sharing.
- In this work, we explore the effects of virtualization and network traffic on LogP [1] parameters.
- This analysis will help to provide insights into understanding the performance of parallel algorithms on cloud platforms and also help in the development of performance models for cloud computing.

Background

LogP [1]:

- *Latency (L)*: the time taken by a message to travel between two processors
- *Overhead (o)*: the time spent by the processor in sending/receiving a message
- *Gap (g)*: time interval between two consecutive messages
- *Processors (P)*: number of processors

Cloud Computing:

It provides various on-demand computing resources on a pay-for-use basis over the internet.

- Service Models: *IaaS* (Infrastructure as a Service), *PaaS* (Platform as a Service), *SaaS* (Software as a Service), *XaaS* (Anything as a Service)
- Deployment Models: *Public* - anyone can use; *Community* - a group of entities can use; *Private* - one entity can use; *Hybrid* - Public and Private combined
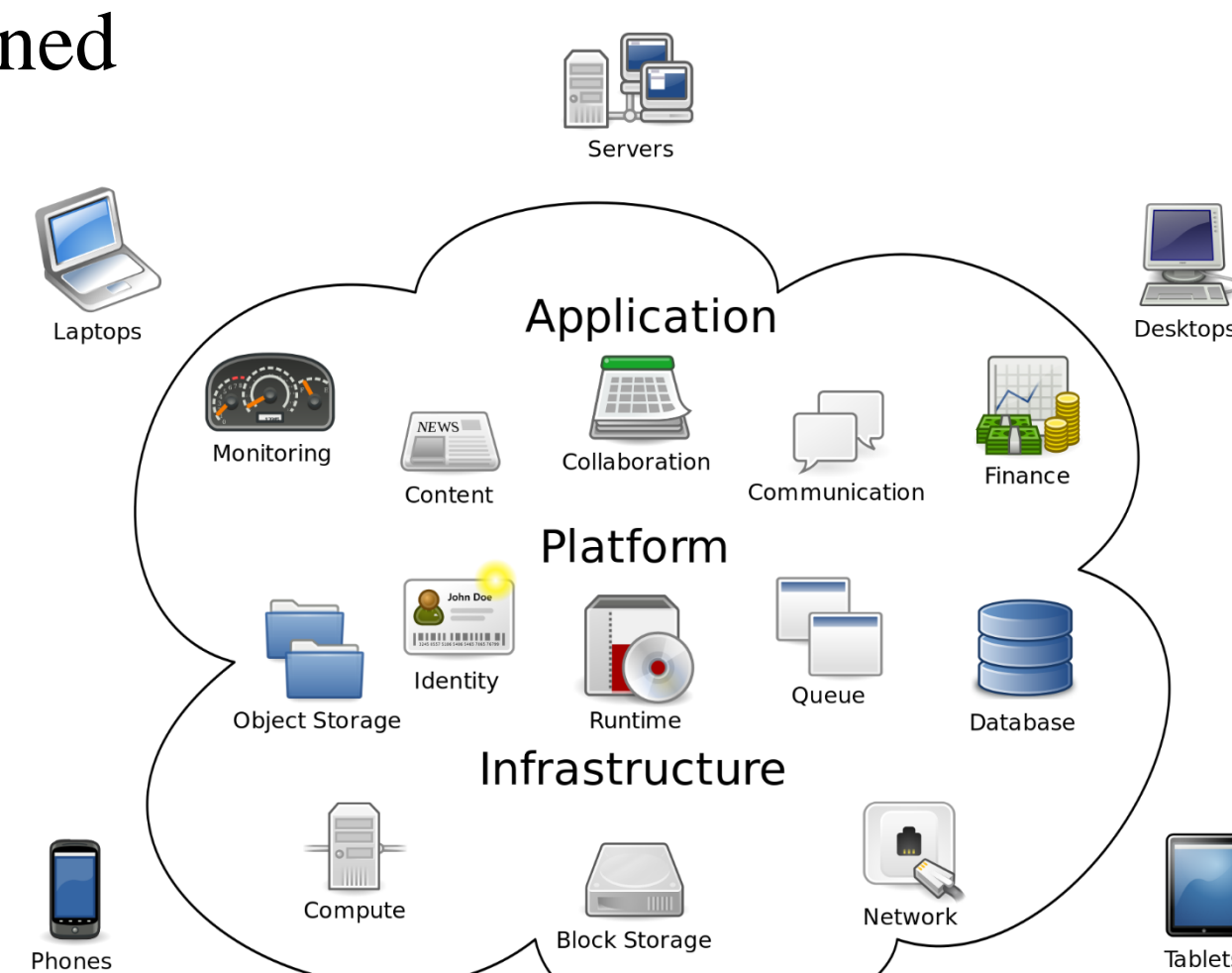

Figure 1: Cloud Service Model example.

Experimental Setup

Cloud System:

- Apache CloudStack™ 4.8 [3]
- KVM Hypervisor

Cloud Infrastructure:

- 1 Management Server and 2 Host Machines
- Each machine has 32 GB of memory, 2 processors with 8 cores each (Xeon E5 1.8 GHz), 1 TB HDD.
- Linux CentOS

LogP:

- The parameters were measured based on "Fast measurement of LogP parameters for message passing platforms." [2] which is one of the many different ways to measure LogP parameters. We used the Message Passing Interface library.

Traffic Simulator [4]:

- *Network Traffic*: HTTP based file hosting service with one server and ten clients were used to simulate network traffic
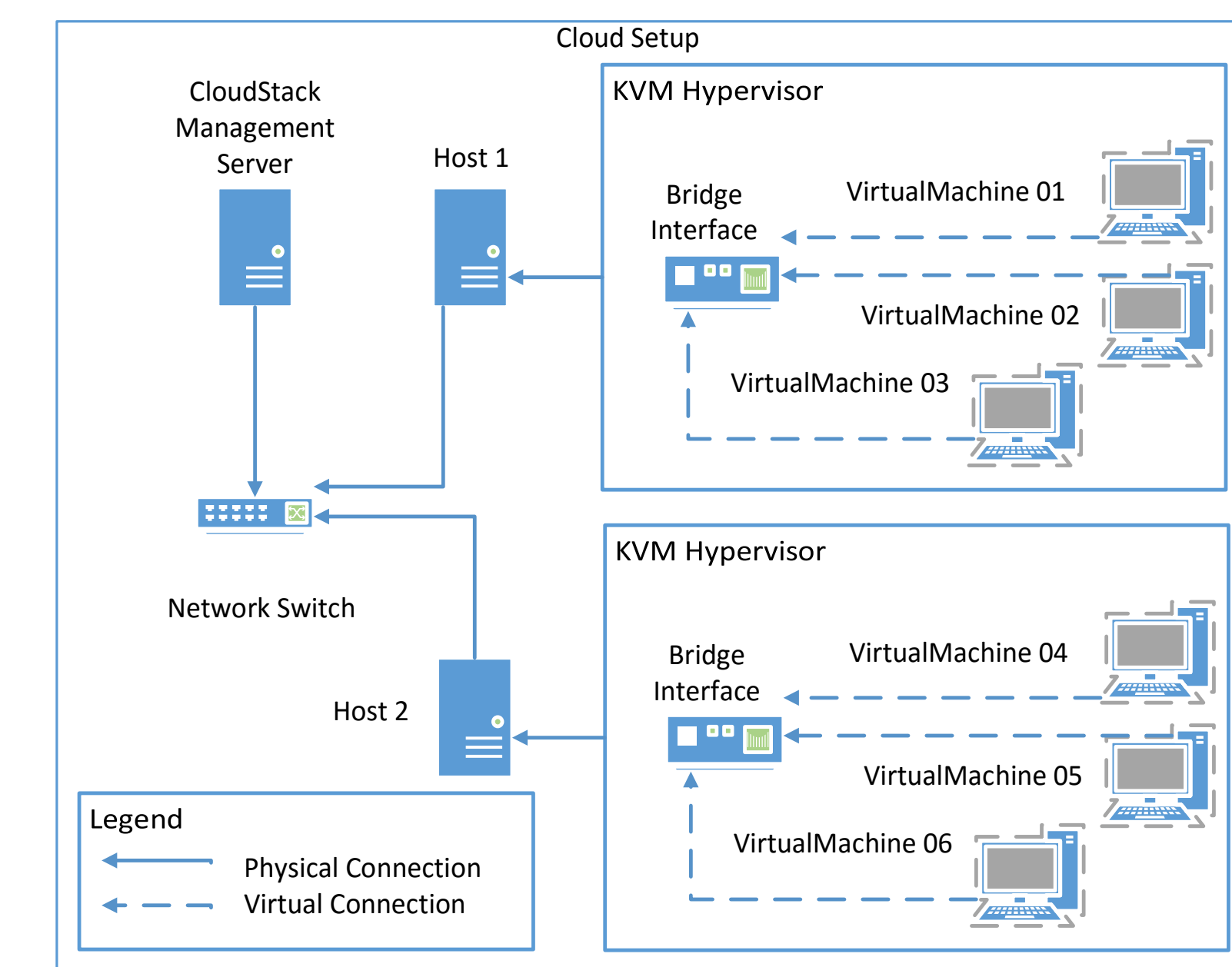- *CPU Traffic*: CPU load generator with Linux *stress* command was used to simulate multiple applications and users

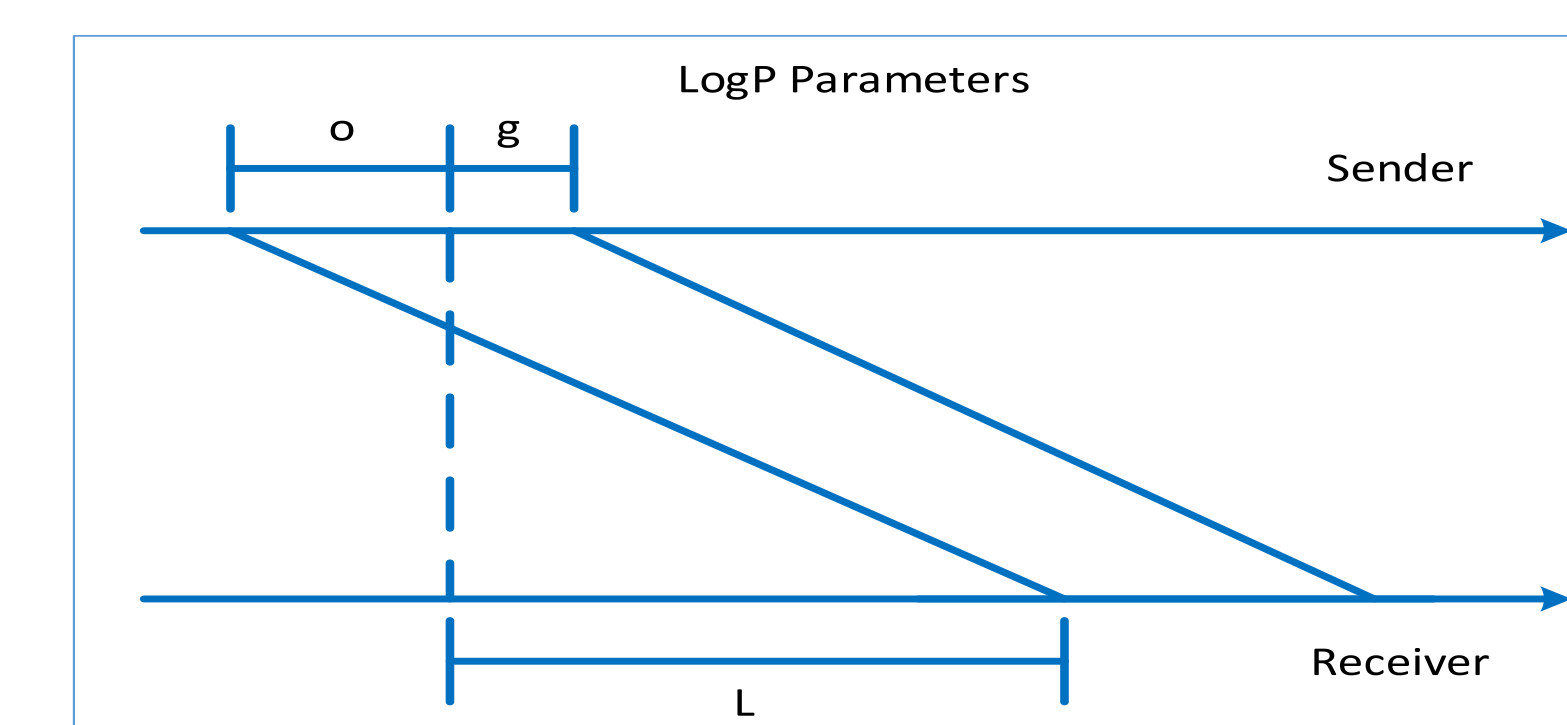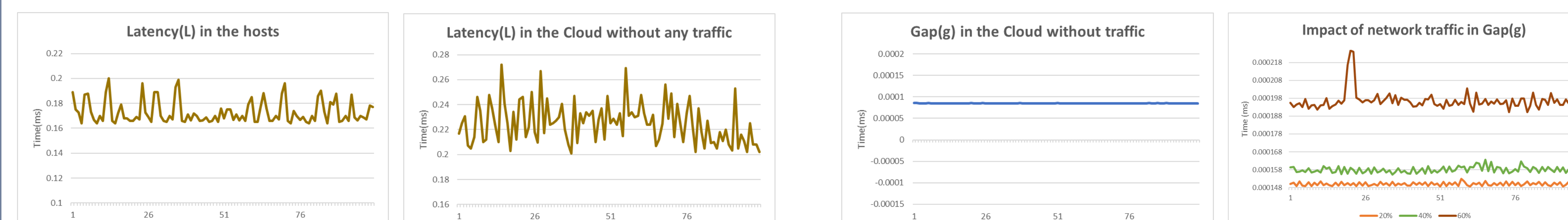Figure 2: Components of Our Cloud Testbed

Figure 3: LogP message transmission

Results

- Virtualization increases the variations observed in latency across multiple runs of the same experiment.
- This can be attributed to the additional software layer (virtual switch) and resource contention.

- Network traffic from other VMs substantially increases network contention and reduces available bandwidth.
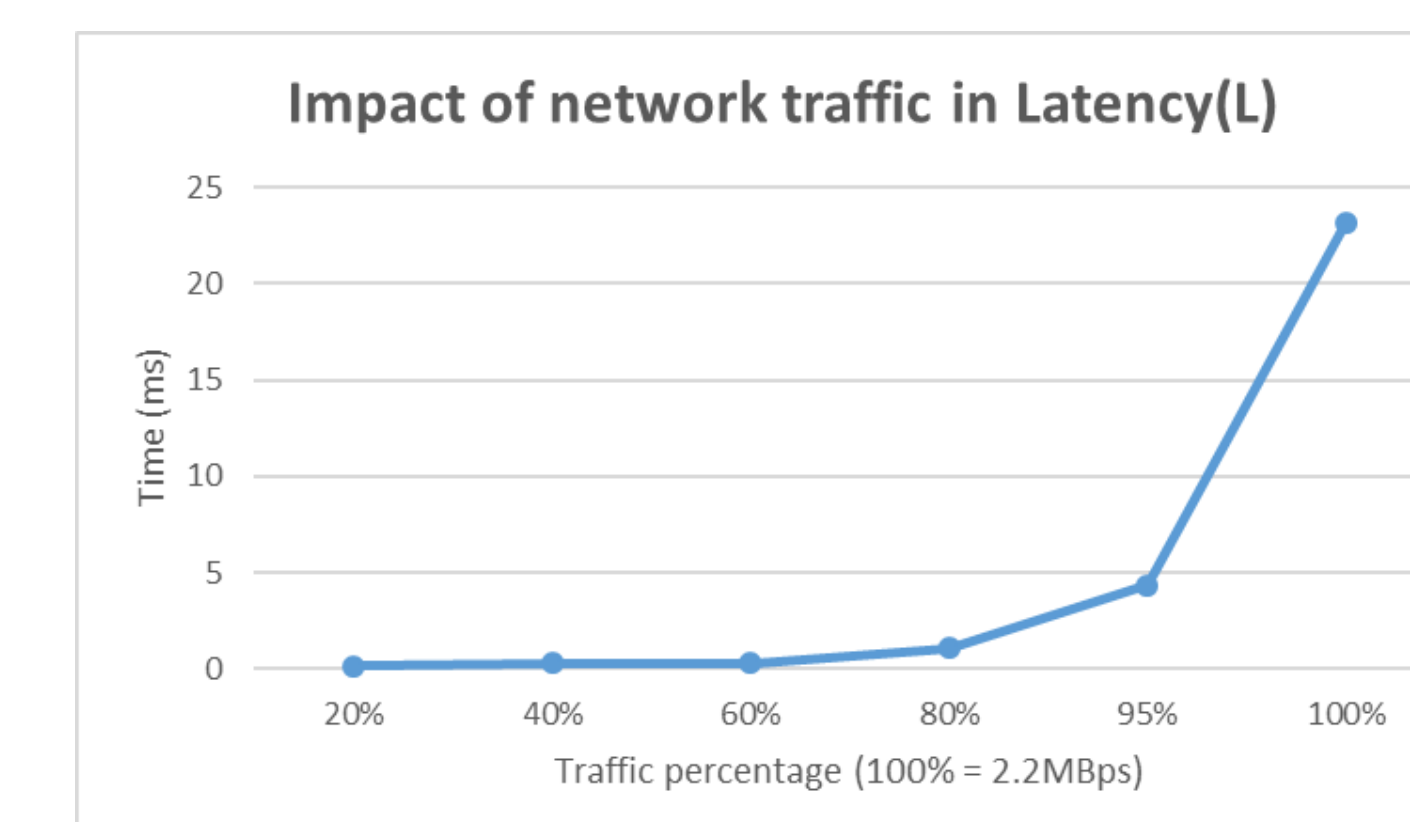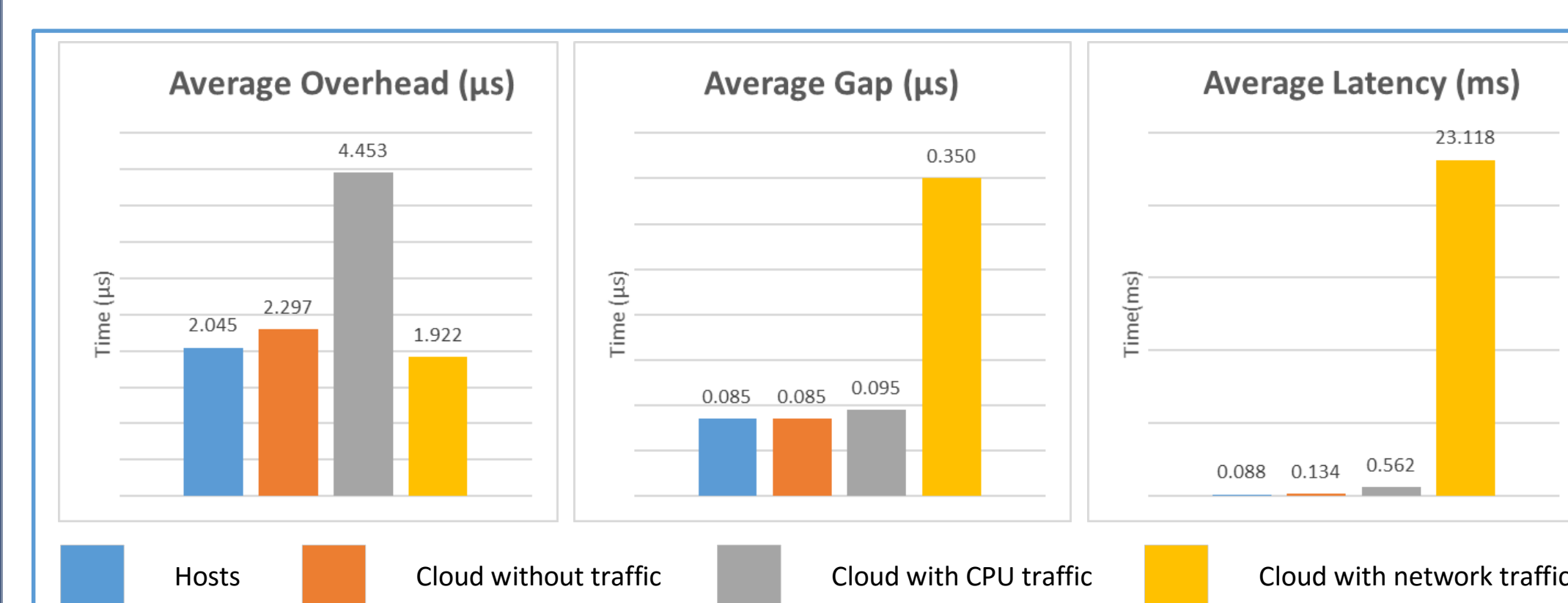
- Additional virtual layer involved in VMs increase the overhead and Latency cost.
- CPU usage from other VMs impacts the overhead
- Both Latency and Gap is affected considerably on a system with high network traffic.

- System with maximum network usage affects the latency drastically.
- The rate of increase in latency is less on a system with up to 80% network usage.

Conclusion

- Results show that the additional virtual layer on the VMs affects network parameters such as latency and bandwidth.
- Moreover, resource contention among the virtual machines within the same hypervisor affects network performance and causes drastic variations across multiple runs of the same experiment.
- Even though the average increase in latency on a system with low traffic (up to 60%) is less, the variations observed across different runs shows the uncertain nature of cloud environment.
- These are initial results that provide some insights into understanding the performance of parallel algorithms on cloud platforms and also help in the development of more accurate performance models for cloud computing.
- In future work, we will compare this method for measuring LogP parameters with other methods to come up with a generic methodology for cloud environments.

References

[1] D. Culler, R. Karp, D. Patterson, A. Sahay, K. E. Schauser, E. Santos, R. Subramanian, and T. Von Eicken. "LogP: towards a realistic model of parallel computation." vol. 39, no. 11. ACM, 1993.

[2] Kielmann, Thilo, Henri E. Bal, and Kees Verstoep. " Fast Measurement of LogP Parameters for Message Passing Platforms" In International Parallel and Distributed Processing Symposium, pp. 1176-1183. Springer Berlin Heidelberg, 2000.

[3] "CloudStack Installation Documentation," CloudStack, accessed June 15, 2016, http://docs.cloudstack.apache.org/projects/cloudstack-installation/en/4.8/.

[4] X. Cheng, C. Dale and J. Liu, "Statistics and Social Network of YouTube Videos," Quality of Service, 2008. IWQoS 2008. 16th International Workshop on, Enschede, 2008, pp. 229-238.

Affiliations

* Brazilian Scientific Mobility Program

† Illinois Institute of Technology

Acknowledgements