Towards Analytical Performance Modeling for HPC on the Cloud

Abdallah Saad, Ahmed El-Mahdy
PCL, CSE
Egypt-Japan Univ. of Sci. & Tech. (E-JUST),
Alexandria, Egypt
{abdallah.saad,ahmed.elmahdy}@ejust.edu.eg

Hisham El-Shishiny
IBM Center for Advanced Studies in
Cairo, Egypt
shishiny@eg.ibm.com

With the current trends towards utilizing cloud computing it becomes interesting to explore the utility of cloud computing in the area of high-performance computing. Such area significantly differs from other typical cloud applications in its intensive computation and communication demands. This contradicts with the underlying sharing concept of resources, exploited by cloud computing.

In this work, we present work-in-progress for evaluating the performance of the standard HPC SPECMPI benchmark on a private, dedicated, cloud system. We construct a simple, yet efficient, analytical model that taking into consideration the application communication and computation patterns, and the underlying physical computing and communication infrastructure. The developed model is a closed-queuing network model that accounts for the contention on physical processing cores and the communication interconnect, and predicts performance.

Figure 1 presents our closed-queuing network system for modelling the execution of an MPI program on the cloud. We consider the MPI program to have a fixed number of cycles; each cycle is a sequence of computations that ends with a communication operation. When running on a number ‘N’ of processes, the cycles are distributed uniformly over the processes.

The queuing network models the execution of a single cycle; a job represents a process. Initially, the jobs reside into the MPI Jobs Pool; they enter the system through the IN port, and eventually exits through the OUT port and return back to the pool; this indicates a completion of an execution cycle; the job is scheduled immediately for another cycle, and this process repeats.

The system models the underlying physical cloud server; for this initial work we consider a simple two-node cluster system. A cluster contains a CPU and a network device (NET). To model the execution and communication behaviour we utilize the concept of visit ratios, utilised in the mean value analysis of queuing networks.

We consider the response time as the performance metric for this experimentation section. For an initial investigation we analyse the ‘104.milc’, '126.lammps' and '137.lu' benchmarks of SPECMPI2007 with different workload configurations. We consider a two compute- nodes cluster system each with 2 Intel Xeon E5645 2.40 GHz CPUs.

Figure 2 shows a strong correlation of r^2 of 0.8, 0.56 and 0.8 respectively for all considered results using our model. The x-axis provides the normalised queuing system response time (cycle time) in seconds. The y-axis provides the actual execution times.

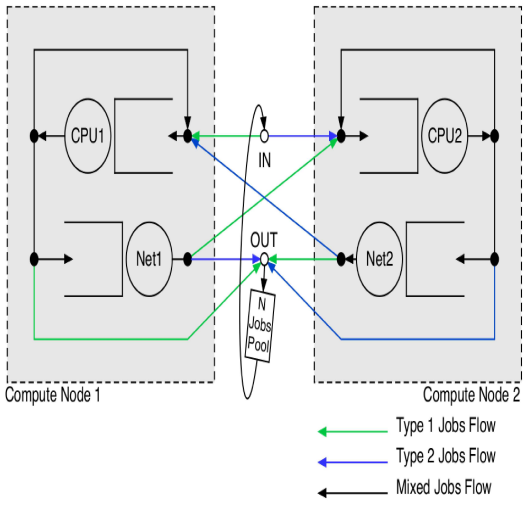

Figure 1: The Proposed Queuing Model.

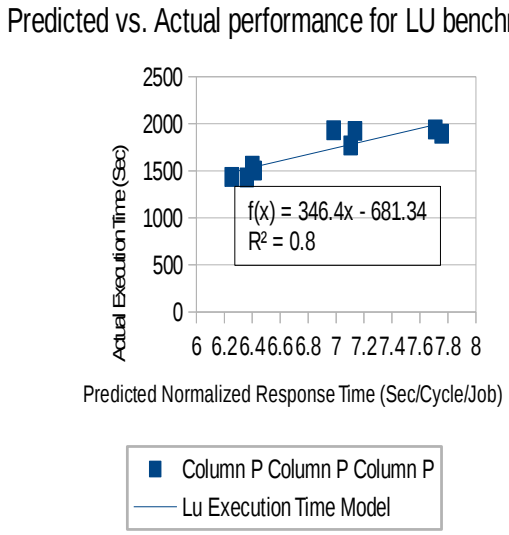

Figure 2.a: Model prediction Vs Actual measures of 137.lu benchmark.

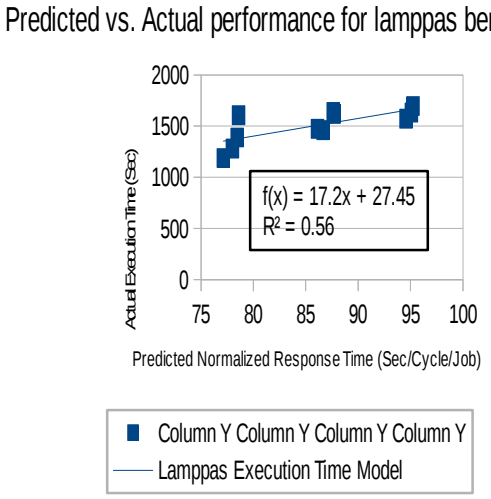

Figure 2.b: Model prediction Vs Actual measures of 126.lampps benchmark.

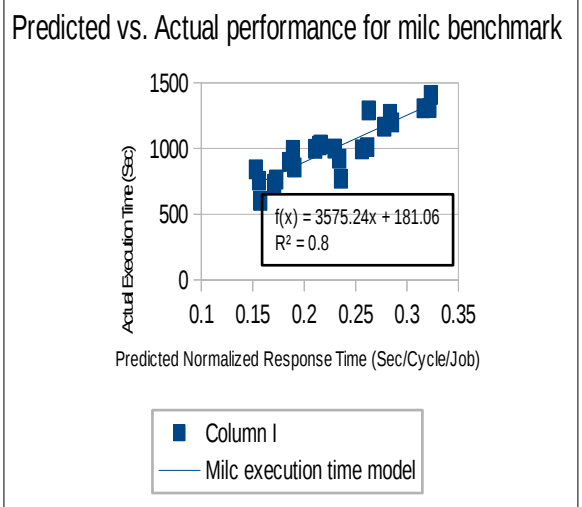

Figure 2.c: Model prediction Vs Actual measures of 104.milc benchmark.

For future work, we are considering many aspects that include: experimenting on a large-scale system to study the scalability of the model across many benchmarks and many nodes on a large-scale cloud system; extending the model to account for interference with other workloads on the cloud; and finally extending the model to account for virtual environment overheads and virtual machines migrations.