Leveraging Commercial Clouds As Adjunct To On-Premises HPC Cluster Resources

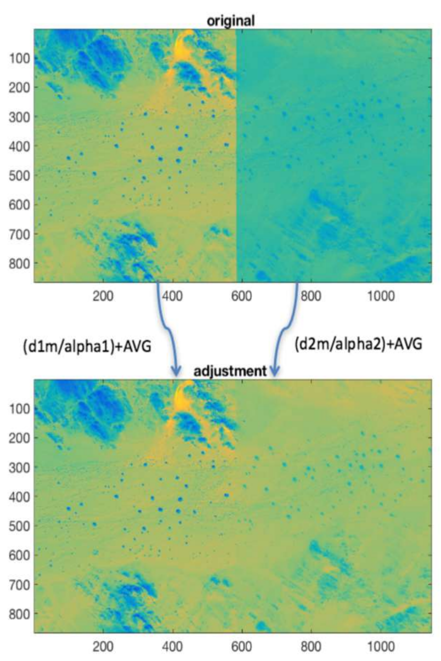
Paulo Paz[1] Hoot Thompson[2]

[1]New Jersey Institute of Technology, [2]Code 606.2 NASA Center for Climate Simulation, NASA GSFC

Introduction

Cloud Computing has long been used to supplement on-premises resources. More recently Cloud Bursting has become a relatively new technology that is being made available to the High-Performance Computing Community. Its appeal lies in its ability to either run jobs in the cloud when demand for on-premises resources have exceeded their capacity or pay on-demand for computing power without the costs associated with setting up the infrastructure.

Recently the NCCS worked on a project involving Cloud Bursting which was used as a motivation behind this study. NCCS's goal was to use satellite imagery data for counting trees and shrubs to estimate biomass and carbon uptake in the Sub-Saharan African region. The initial workflow burst onto Amazon Web Services (AWS) ; orthorectified mosaicking shown right, removed the effect of image perspective and terrain on the images and then stitched the images together [1]. The following processing workflow then carried out the actual algorithm for counting. Using AWS resources for Cloud Bursting, the NCCS was able to reduce their estimate of 10 months for dedicated dedicated on-premises resources to approximately 1 month.



As a result our objective for this project is two-fold; establish a suite of benchmarks to:

1. Verify that the computing power of a Public Cloud is comparable to that of on-premises resources.
2. Using software from Adaptive Computing Enterprises Inc. and from Microsoft (previously Cloud Computing) demonstrate how to configure a Cloud Bursting environment, and study how responsive a Public Cloud Provider is to elastic computing.

Benchmarks

The typical operations performed in a High-Performance Computing Environment will be mimicked both in Microsoft's Azure Cloud Computing Platform and on-premises here at NCCS's supercomputing cluster Discover.

- High-Performance LINPACK Benchmark [2]– solves a dense system of linear equations on a distributed memory cluster.
- OSU Latency and Bandwidth Benchmark [3] – uses MPI send and receive for various data sizes to determine the response time and the maximum sustained data rate respectively.
- IOR Benchmark [4] to measure parallel file system I/O performance for the underlying Lustre file system.
- Stream benchmark [5] for measuring sustainable CPU memory bandwidth and its respective computation rate.

Discover	Azure
• Intel Xeon Haswell 2.6 GHz processor cores • 2 14-core processors per node • 128 GB of memory per node • Interconnect: Infiniband FDR	• Intel Xeon Haswell V3 3.2 GHz processor cores • 2 8-core processors per node • 112 GB of memory per node • Interconnect: Infiniband FDR

Results

High-Performance LINPACK

The problem size between platforms is standardized to use 94% of the memory usage of the node, and then scaling the number of nodes to examine the speedup (Figure 1). Then the problem size is scaled *per node* examining how the well the platforms scale in GFLOPS, and against their theoretical peak.

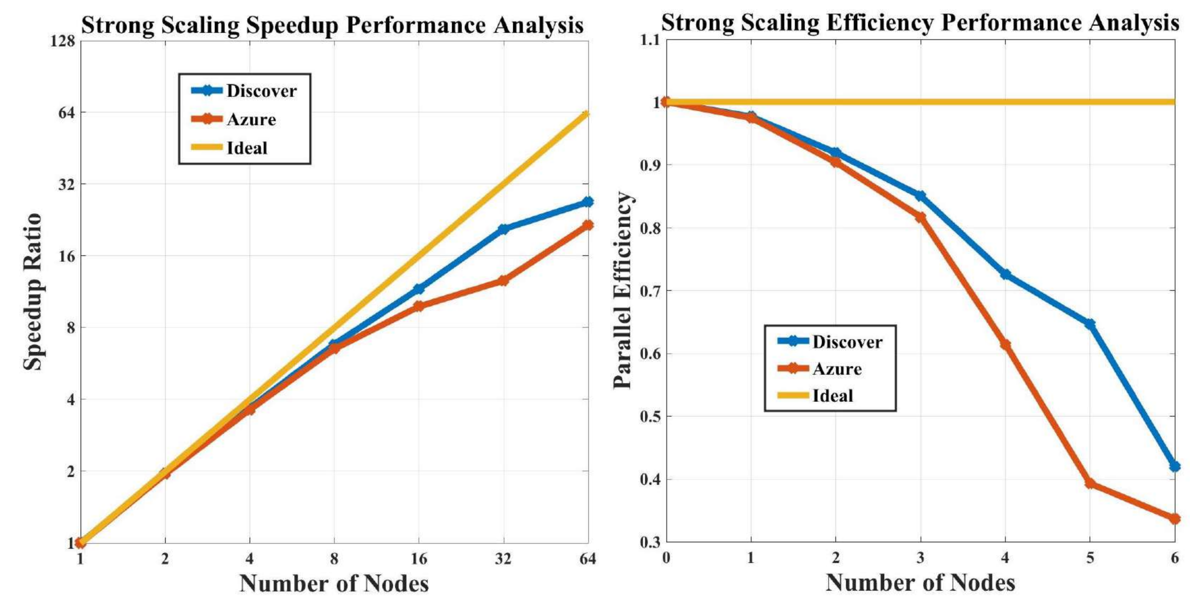


Figure 1. HPL for a fixed problem size.

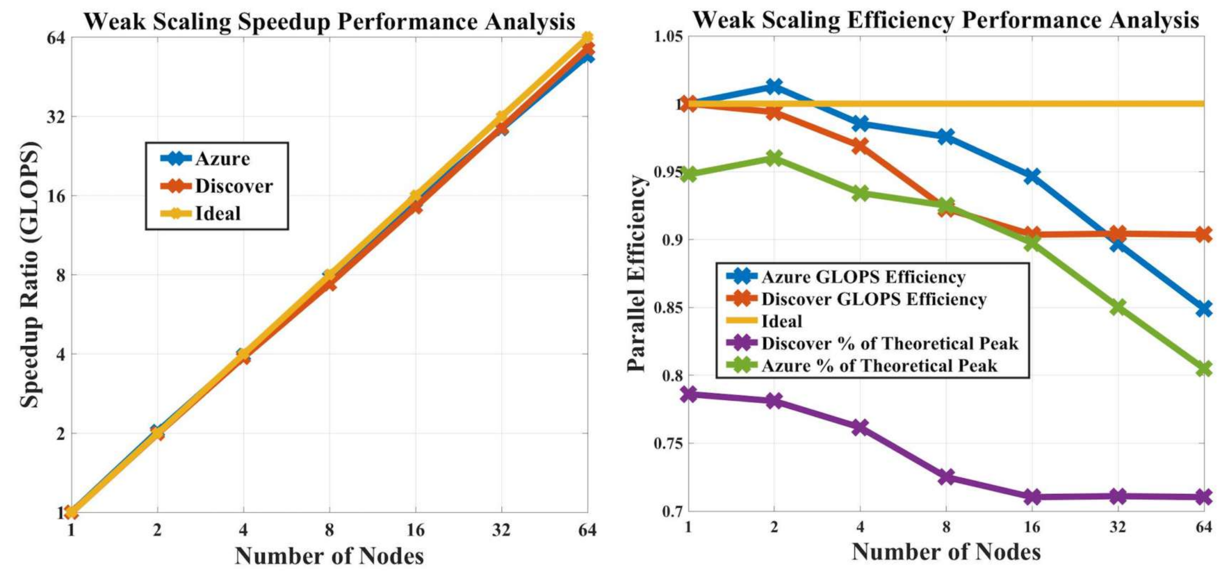


Figure 2. HPL for a problem size that is scaled to the number of nodes.

OSU Micro Benchmarks

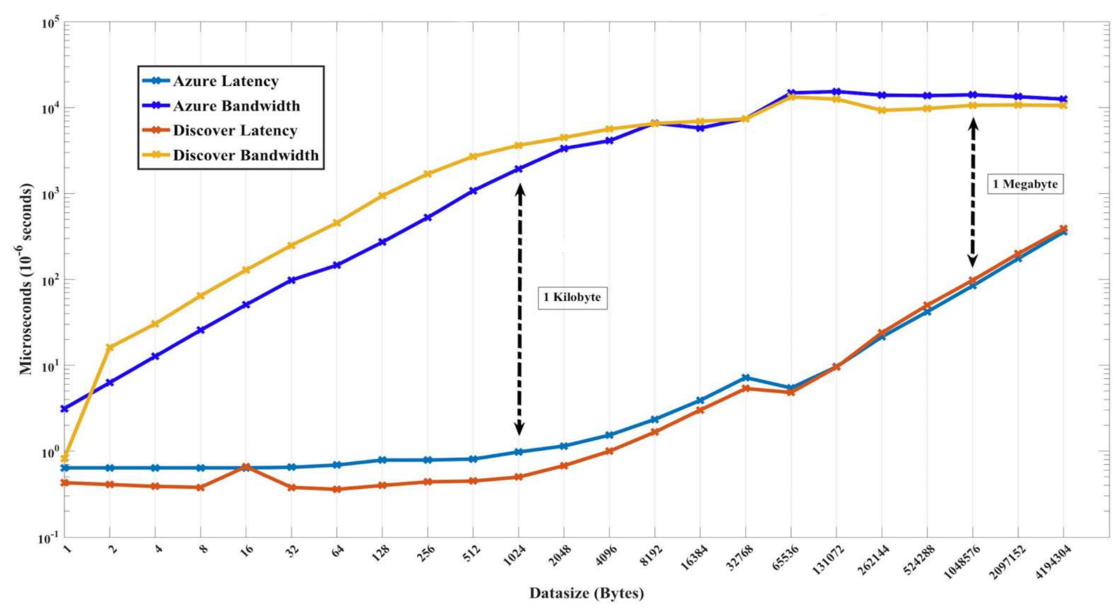


Figure 3. Point-to-point MPI communication with blocking and non-blocking communication for latency and bandwidth respectively.

Stream Benchmarks

With the CPU speed of the fastest microprocessors increasing faster than the speed of memory devices, an increasing fraction of workloads will be dominated by memory access and transfer times rather than compute time. Machine balance is defined as the ratio of peak floating ops/cycle to sustained memory ops/cycle; a ratio of approximately 1 is considered optimal.

4 operations are performed on Discover (gray) and Azure (white); their best MB/s rate is recorded (on an 80m size array to bypass the cache) in addition to their respective machine balances.

	Copy	Scale	Sum	Triad
Best Rate MB/s	94499	94298	100750	100697
	75462	66528	72039	71759

Machine Balance	11.83
	11.68

IOR Filesystem Benchmark

Figure 4. Topology of the filesystem used on Azure.

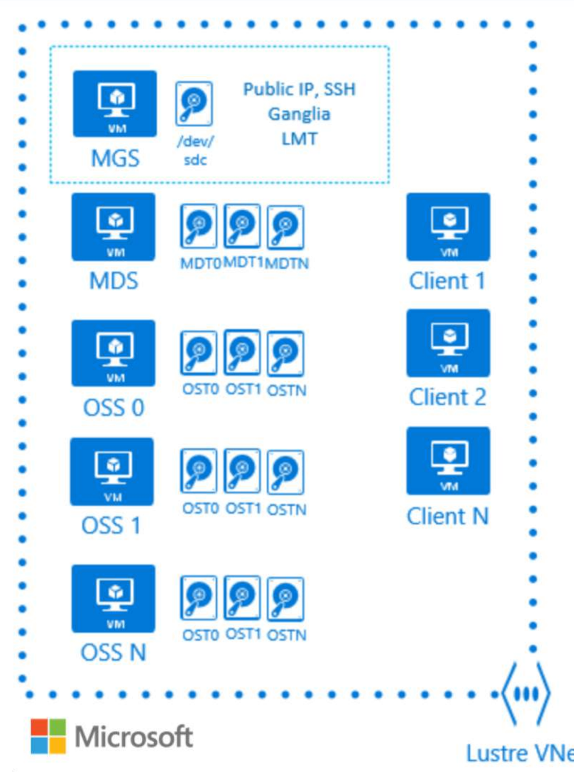


Figure 5. Transfer size=2MB. Aggregate read and write to Azure's Lustre filesystem.

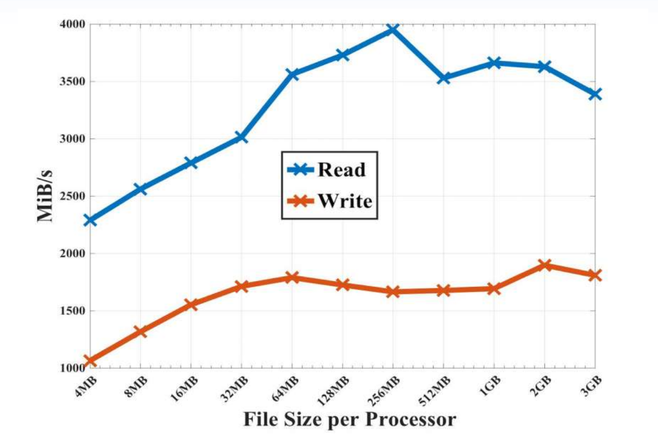
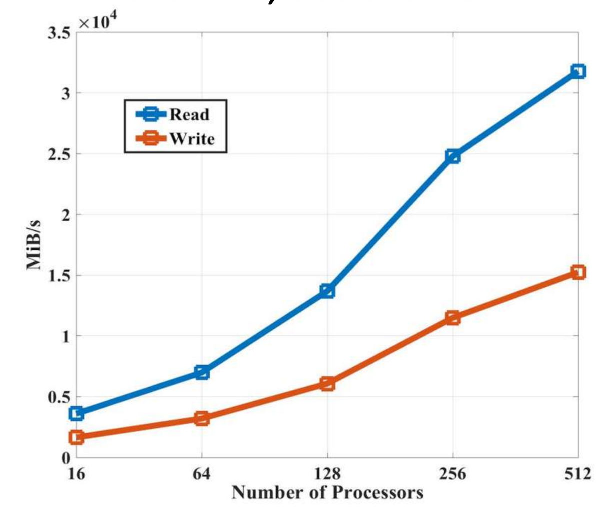


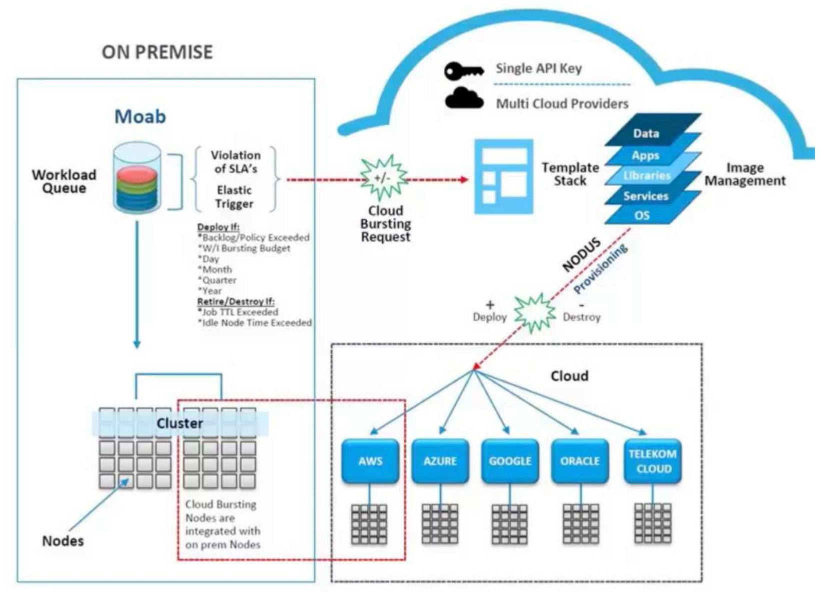
Figure 6. Scaling read and write for transfer size=2MB, blocksize=1GB



- There are 4 Object Storage Servers (OSS) in our environment which provides I/O via network to the clients.
- Each OSS has 4 OSTs which is the actual storage device that stores the stripes of files.

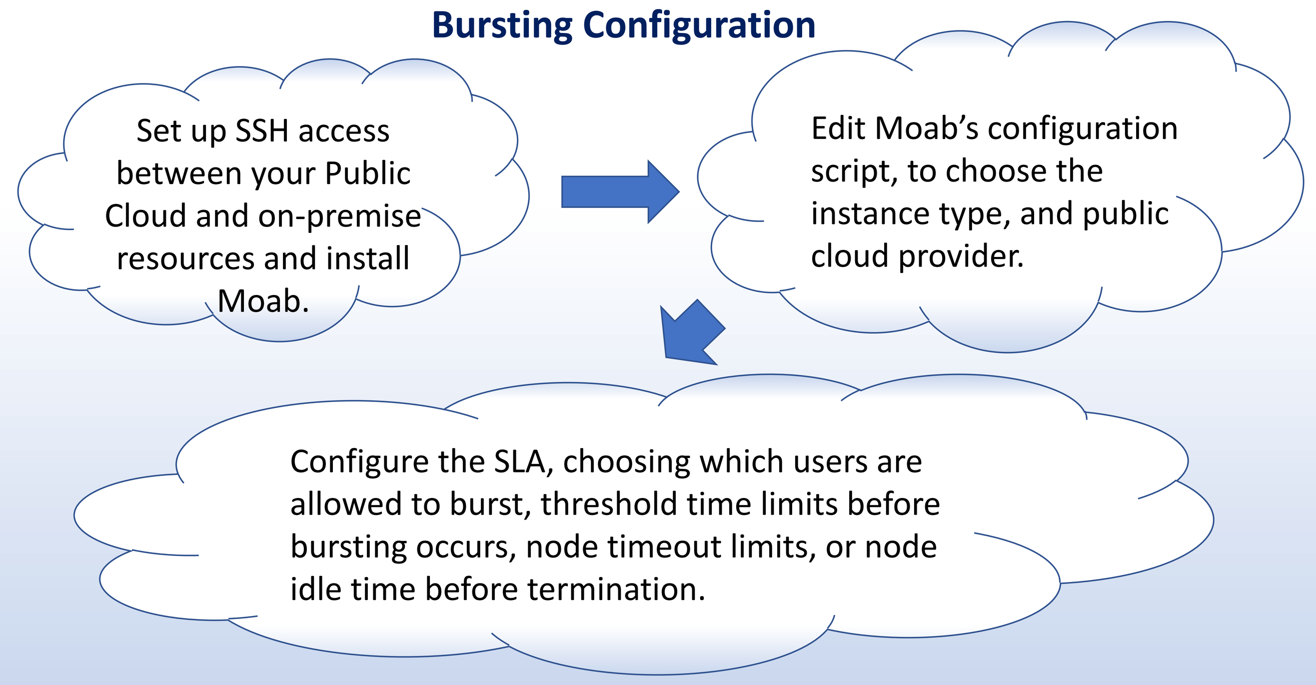
Cloud Bursting

Cloud Bursting is a computing environment that is configured to redirect jobs to a Public Cloud when on-premises resources have reached their computing capacity.



- Adaptive Computing's Moab/NODUS software is used to test cloud bursting. [6].
- Establish a test environment in AWS with 1 server and 2 compute nodes (4 processors each) acting as your on-premise resources.
- Violation of any of the service level agreements (SLA) automatically triggers bursting.

Bursting Configuration



Conclusion

It was verified that Azure cloud resources are comparable to that of our on-premise cluster Discover and demonstrated how both behaved under typical HPC conditions. In addition, we conducted a representative demo of Cloud Bursting on AWS using software from Adaptive Computing and future work will focus on benchmarking the elasticity of a public cloud provider.

References

[1] Requa, Michael, et al. *Using Cloud Bursting to Count Trees and Shrubs in Sub-Saharan Africa*.
[2] "High-Performance LINPACK" http://www.netlib.org/benchmark/hpl/index.html
[3] "OSU Micro Benchmarks" http://mvapich.cse.ohio-state.edu/benchmarks/
[4] "IOR Benchmark" https://github.com/LLNL/ior
[5] McCalpin, John D.: "STREAM: Sustainable Memory Bandwidth in High Performance Computers", a continually updated technical report (1991-2007), available at: http://www.cs.virginia.edu/stream/
[6] "Moab/NODUS Cloud Bursting" http://www.adaptivecomputing.com/moab-nodus-cloud-bursting/