1

AzureBench: Benchmarking the Storage Services of the Azure Cloud Platform

Dinesh Agarwal and Sushil K. Prasad

-By: Pooja Shrivastava & Sruthi Jogi
Student no: 1750220, 1750193

2

INTRODUCTION

- Introduction to open-source benchmark suite, AzureBench for Windows Azure cloud platform.
- Cloud computing effectively saves the eScience developer the hassles of resource provisioning but utilization of these resources will be questionable if it can not meet the performance expectations of deployed applications
- Among all clouds, the emerging Azure cloud from Microsoft remains a challenge for HPC program development both due to lack of its support for traditional parallel programming support such as message passing interface MPI and map-reduce and due to its evolving APIs
- Interesting feature such as open-source generic application framework can be a starting point for application development over azure

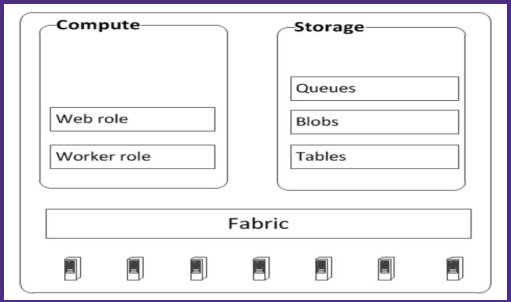
3

Introduction Contd…

- Microsoft azure to be the platform of high performance computing HPC.
- The emerging cloud platforms, such as Microsoft's Azure - with their potential for large scale computing and storage capabilities, easy accessibility by common users and scientists, on demand availability, easy maintenance, sustainability, and portability - have the promise to be the platform of choice for HPC applications
- Cloud computing provides users with quick access to a large-scale HPC cluster without having to worry about the setup, maintenance, or initial investment..

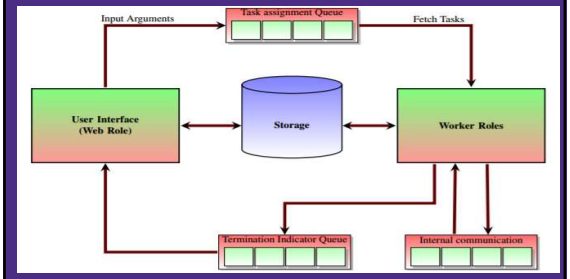
4

Typical programming artifacts of Windows Azure platform



5

AN APPLICATION FRAMEWORK FOR SCIENTIFIC APPLICATIONS ON WINDOWS AZURE CLOUD PLATFORM



6

EVOLUTION OF BENCHMARKING OF STORAGE SERVICES IN AZURE PLATFORM

- Cloud storage services are gaining great popularity and evolving rapidly to address customers' requirements.
- Cloud storage solutions provide easy access to virtually unlimited storage and do not involve any set up cost.
- Service providers storage resources are made available with a simple 'pay-as-you-go' billing model

BACKGROUND AND LITERATURE

7

- Cloud Computing is typically perceived as a set of shared and scalable computing resources, located all over the world and available on-demand over a network.
- The resources are made available to end users as various services such as "Platform as a service" (PaaS), "Infrastructure as a service" (IaaS), and "Software as a service" (SaaS).
- The key idea here is to abstract the provisioning mechanism at a level, where users can avail of these resources dynamically without burdening themselves with either the availability or the maintenance

UNIVERSITY of WASHINGTON

BENCHMARK EXPERIMENTS AND TIMING CHARACTERISTICS

8

- Windows Azure storage services partition the stored data across several servers to provide enhanced scalability.
- The limit on a storage account is 100 TB.
- There are additional limits on scalability targets. Windows Azure storage services can handle up to 5,000 transactions per second.
- The maximum bandwidth support is up to 3 GB per second for a single storage account.
- Exceeding any of the specified limits result in the failure of a role instance.

UNIVERSITY of WASHINGTON

KEY CONTRIBUTIONS

9

- 2010 - Hill et al -extends a preliminary study on providing a comprehensive scalability assessment of Azure platform's storage services using up to 100 processors.
- The assessments of Azure platform provide updated, realistic performance measurements as we utilize the APIs released after significant changes were made to the Azure cloud platform since 2010.
- The importance of Azure platform has been recognized by industry as well as academia as is evident from the rare partnership of National Science Foundation (NSF) and Microsoft in funding scientific research on Azure cloud.

UNIVERSITY of WASHINGTON

EVOLUTION OF STORAGE SERVICES OF THE AZURE CLOUD PLATFORM

10

- The basic ideas of these rules is to distinguish between logic and data, which helps in making maintainable parts independent.
- Many rules engines are designed that are used to reduce the cast of designing, developing, and delivering software .

  For example:
- For instance, in a traditional wireless sensor and actuator network, different kinds of embedded sensors are deployed that are used to collect environmental data.

UNIVERSITY of WASHINGTON

AUTHOR'S CONCLUSION

11

- Azure Bench - an open source benchmark suite for Windows Azure platform's storage services - along with experimental details to analyse the performance capabilities of Azure cloud platform.
- The comprehensive performance evaluation of Windows Azure platform's storage services - Table, Blob, and Queues. We also present a generic framework along with pointers for HPC application development on Azure.

UNIVERSITY of WASHINGTON

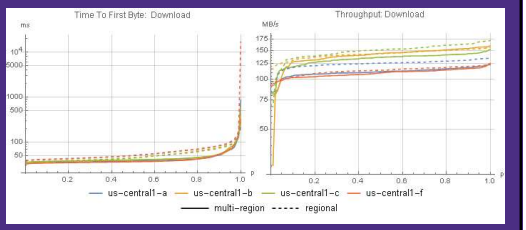
- **Automating Big data Benchmarking and Performance analysis**

12

UNIVERSITY of WASHINGTON

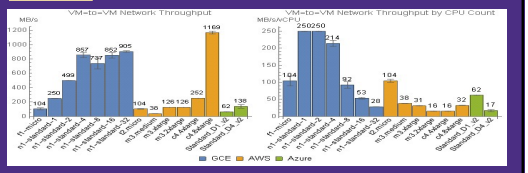
AWS S3 vs Google Cloud vs Azure:Cloud Storage Performance

13



Measured intra-datacenter network throughput for various VM's.

14



CRITIQUE :STRENGTHS

15

- Standardized Service Contract
- Discoverability
- Abstraction
- Lake Of State
- Loose Coupling
- Reusability
- Autonomy
- Composability
- Tools Storage
- Scalability
- The generic framework has reduced the processing time for many time-consuming applications.
- Maximum throughput can be achieved.

CRITIQUE: WEAKNESSES

16

- **Performance Challenges:**
- Customers lacks the tools to verify the services work as expected.
- Lacks functional correctness, service availability, reliability of the tools.
- Security challenges for applications running on multiple servers.

CRITIQUE :EVALUATION

17

- **About the paper**
- Survey paper submitted by Georgia State University in 2012
- Cited 20 times.
- Explains the various performance issues faced by developers dealing with different vendors for the cloud services and resolves its issues.
- Provides a clear report of the scalability and bottlenecks of the Azure storage services.
- Paper acknowledges some of the challenges for HPC program development due to its traditional parallel programming support.
- The cost effectiveness when involved with the parallel access to the various storage services are outside the paper.

GAPS

18

The few limitations which were not discussed by the author

- The verification tools used for improving the performance of the storage services are not feasible with everybody.
- A big number of benchmarks is actually used to evaluate each single feature apart or a global overall performance view.
- The Azure benchmark suite also face challenges such as TPC in that do not generate comparable benchmarking results.

19

FUTURE WORKS

- **Additional services are needed:**
- To compare performance evaluations involving local drives, caches, and SQL Azure database.
- Includes resource provisioning times and application deployment timings.
- Also to incorporate benchmarking suited for other cloud offerings by different vendors.

UNIVERSITY of WASHINGTON

20

Queries?

Any Questions?

UNIVERSITY of WASHINGTON

21

Thankyou !!!

UNIVERSITY of WASHINGTON