Analysis of Variance in Big Data Cloud Networks in Azure Cloud

Jithin Babu(113271021) Satyanarayana Murthy Kaja(113278723) Ashley Lobo(113218828)

1 Abstract

According to many cloud practitioners and performance engineers, the impact of Performance variability is huge in the transfer of big data but in top systems conferences, the research communities have regularly disregarded variability when running experiments in the cloud. Performance variability is an important factor in the cloud as it directly affects the usage and throughput. In this project we replicate the measurements in the paper [1]. The paper mainly focused on AWS, GCP and HPC cloud platforms. We extend the idea of the paper to determine the impact of variability on the Azure cloud platform. We replicate the measurements mentioned in the GCP cloud to verify our measurement setup. We experimented with different scenarios in Azure on regions and Vnets that were not mentioned in the paper. The key factor we have used for performance measurement is measuring bandwidth and latency. To measure Bandwidth we used iperf3 and to measure latency we used tshark which is a command tool for Wireshark. We have observed a significant variance in bandwidth and latency for different scenarios discussed in further sections.
Github Link for resources: https://github.com/murthykaja/FCNProject
Presentation Link:
https://docs.google.com/presentation/d/1zo7ev0wg6HYN9uZrJVsJc8XM4p-AL9_Bj42KvEPPMmI/edit?usp=sharing

2 Introduction

Performance variability is a well-known issue in the cloud and has been studied from the early days of the cloud in the market. This is an important issue as this causes a direct effect on cost and reliability.

For Systems like big data, it is very crucial to perform high-performance measures with higher reliability [4]. As it is very easily available and instances are created on-demand, thus reliability plays a crucial role for all cloud competitors. Hadoop and spark engines are deployed on-demand in most cases. As Big data technologies like spark (communicate with nodes through spark context) rely mostly on networking [5] between virtual machines, the performance variance has an exponential effect.

As the project mainly focuses on measuring variability in Azure for big data systems, from the official documentation of Microsoft we can understand that Azure provides an on-demand service called HDInsight[6] which is PaaS (Platform as a service). Further analysis on the HDInsight service shows that it provides the user with a user-specific cluster (group of virtual machines) with a predefined set of installed software like HDFS, Spark, Kafka, etc. The cloud provider promises Quality of Service and fairness but even then we find variability in some scenarios.

The documentation provided by Azure has some of the decisions which are well documented but others like the ones we have evaluated in the report are not. Unfortunately, these differences are causing more effect on variance on performance in Azure.

3 Setup

To measure bandwidth and latency we need to first set up an environment. Following are the two cloud services we have used and their experimental set up configurations.

Azure:

As we discussed earlier, big data services used in azure are from HDInsight PaaS. Thus from the documentation provided by microsoft[7] we can determine the virtual machines used as clusters in HDInsight service are F series virtual machines. We have conducted a measurement study on F1 virtual machines with configurations in table1.

Region	East US or Central US
Security type	Standard
Image	Ubuntu Server 20.04 LTS - Gen2
Size	Standard_F2s_v2 - 2vcpus 4GiB memory
Inbound ports	SSH(22), HTTP(80), HTTPS(443)
OS Disk type	Premium SSD
NIC network security group	Basic
Accelerated networking	Disable
Boot diagnostics	Disable

Table 1 : Parameter to configure Azure F series Virtual Machine

Bandwidth measurement

We have considered the following scenarios for the measurement study:

1) Experiment 1: Same Region Different Vnet
    a) 5 seconds run and 30 seconds sleep
    b) 10 seconds run and 30 seconds sleep
2) Experiment 2: Different Region Different Vnet
    a) 5 seconds run and 30 seconds sleep
    b) 10 seconds run and 30 seconds sleep
3) Experiment 3: Same Region Different Vnet
    a) 5 seconds run and 30 seconds sleep
    b) 10 seconds run and 30 seconds sleep

Latency measurement

We have considered the following scenarios for the measurement study:

Experiment 4: Same Region Same Vnet

a) 10 seconds run and 30 seconds sleep

For each of the above experiments, we used 2 virtual machines, one being the client and the other being the server. For experiment 2 we have considered 2 regions: East US and Central US regions on Azure. For experiment 1 & 3 we have considered the East US region on Azure. Other technologies installed on the Virtual machines are Iperf3[2] and TShark[3] measurement tools. iPerf3 is a tool for active measurements of the maximum achievable bandwidth on IP networks. It supports tuning of various parameters related to timing, buffers and protocols (TCP, UDP, SCTP with IPv4 and IPv6). For each test it reports the bandwidth, loss, and other parameters. TShark is a terminal oriented version of Wireshark designed for capturing and displaying packets when an interactive user interface isn't necessary or available. It supports the same options as wireshark

GCP:

To measure in GCP, we use the DataProc BigData Cluster. The following configuration was used for both client and server

Region	us-central1
Zone	us-central1-a
Autoscaling	Off
Dataproc Metastore	None
Scheduled deletion	Off
Master node	Standard (1 master, N workers)
Machine type	n1-standard-4
Number of GPUs	0
Primary disk type	pd-standard
Primary disk size	500GB
Local SSDs	0
Worker nodes	2
Machine type	n1-highcpu-4
Number of GPUs	0
Primary disk type	pd-standard
Primary disk size	500GB
Local SSDs	0
Secondary worker nodes	0

Our setup runs for 5 seconds followed by a 30 seconds gap. Like the Azure Setup, we use Iperf3 for the DataProc Cluster. We receive data such as bandwidth, loss, and other parameters using Iperf3. This setup was run for 3 continuous days.

4. Validating measurement setup

To test our measurement code, we considered the GCP platform and ran the same measurement setup. From the below graphs we can see that both the paper[1] and our experiment result showed variability in bandwidth. Though variability is found in our experiment, it is not equivalent to the one in GCP. This is because we used fewer cores compared to paper[1]. Another reason could be that the paper ran the experiment for 3 weeks, while we ran it for 3 days (due to budget constraints).
The maximum variability we observe is 1.8 Gbits/sec. The bandwidth varies between 5.6 Gbits and 6.7 Gbits over 2 days. The variability according to the paper[1] is 2.8 Gbits per second for GCP

Type for 5 Seconds run time in GCP	Min Bandwidth	Max Bandwidth	Average Retransmission
Same Region Different Zone	4.89 GBits/sec	6.69 GBits/sec	44

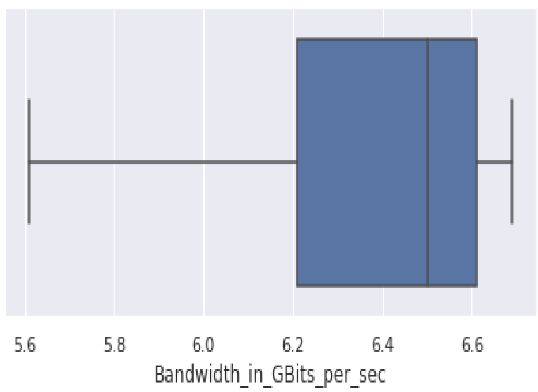


Fig 1 : GCP Client after elimination outliers

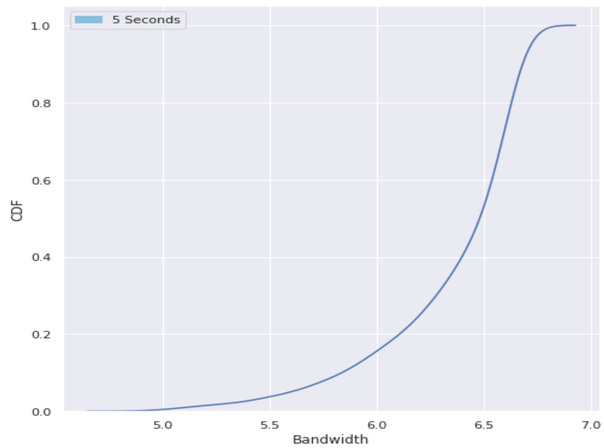


Fig 3 : GCP Bandwidth CDF

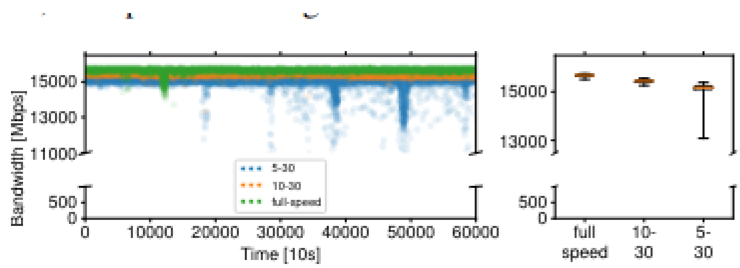


Fig2. GCP variance Diagram from base paper

5. Observation and Analysis

5.1. Bandwidth

The maximum variability is observed for the same region and different vnet set up and is 2.215 Gbits/sec. The minimum variability is observed for the different region different vnet setup and is 0.627 Gbits/sec. These values are observed for the 5 seconds run and 30 minutes wait run-time.

Incase of the 10 seconds run and 30 minutes wait run, the maximum variability is seen for same region different vnet and is 1.43 Gbits/Sec. The minimum variability is seen for Same Region, Same Vnet and the value is 0.56 Gbits/Sec.

Type for 5 Seconds run time	Min Bandwidth	Max Bandwidth	Average retransmissions
Same Region Different Vnet	0.605 GBits/sec	2.82 GBits/sec	113
Different Region Different Vnet	0.443 GBits/sec	1.07 GBits/sec	16
Same Region Same Vnet	1.38 GBits/sec	2.51 GBits/sec	101

Type for 10 Seconds run time	Min Bandwidth	Max Bandwidth	Average retransmissions
Same Region Different Vnet	1.28 GBits/sec	2.71 GBits/sec	130
Different Region Different Vnet	0.111 GBits/sec	1.11 GBits/sec	34
Same Region Same Vnet	2.18 GBits/sec	2.74 GBits/sec	117

5.1.1 Scenarios

From the figure and table we can observe there is variability in bandwidth.

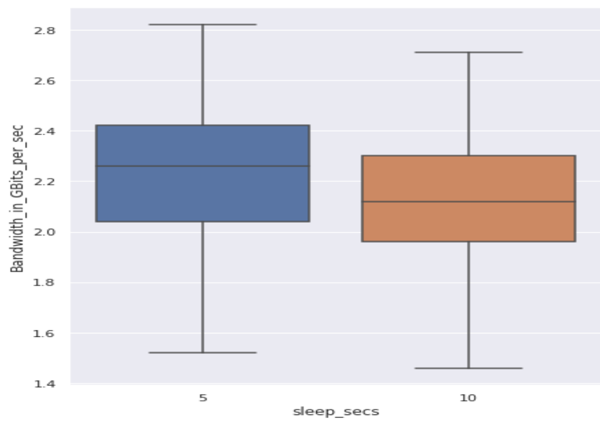


Fig 4. Same Region Different Vnet after removing outliers

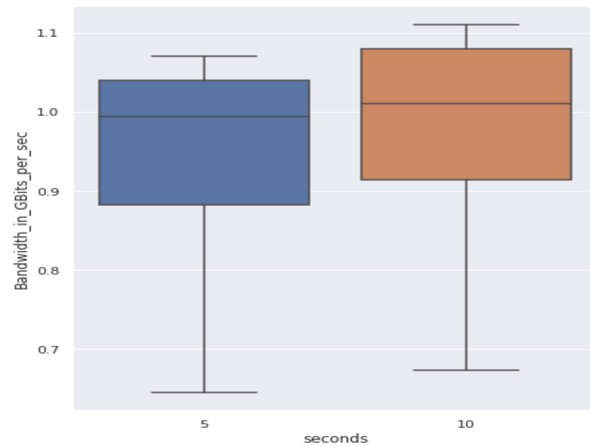


Fig 5. Different Region Different Vnet after removing outliers

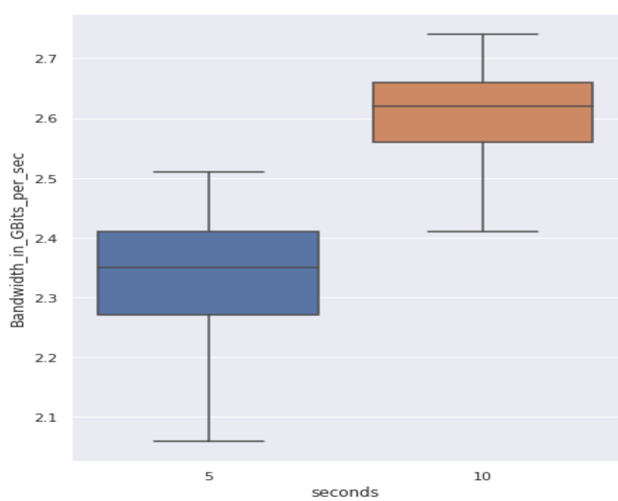


Fig 6.Same Region Same Vnet after removing outliers

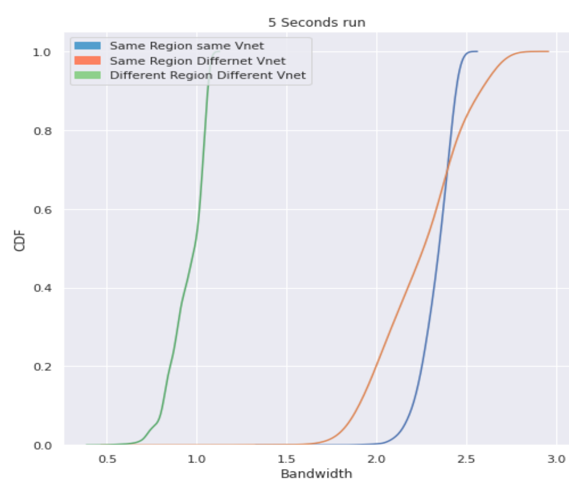
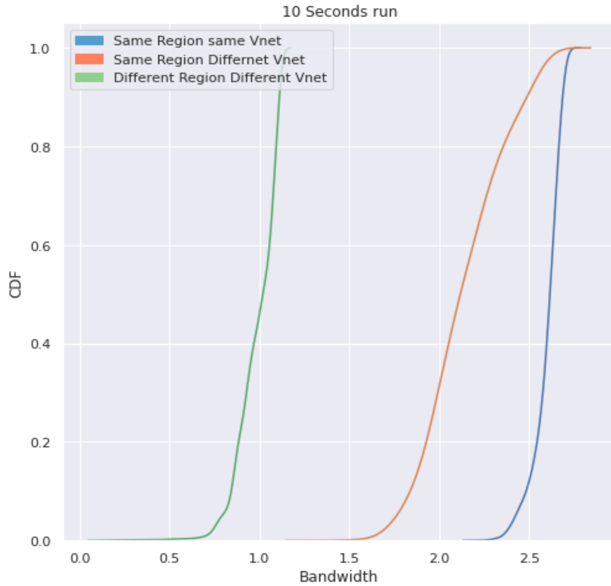


Fig 7 : CDF for 5 Seconds Bandwidth run



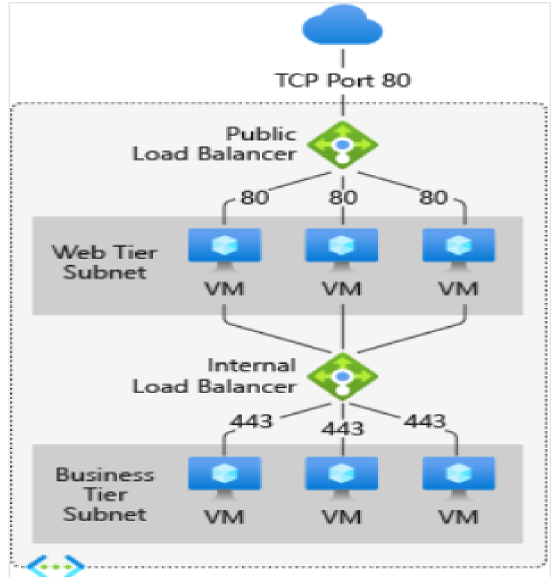


Fig 8 : CDF for 10 Seconds Bandwidth run   Fig 9: topology of virtual machines in same vnet

5.1.2 Bandwidth Analysis:

Below are the explanations of all the findings from the above scenarios.

- When VMs are in Different regions we observed that overall bandwidth has been reduced. This could be due to the design of azure for default settings. We can give additional bandwidth but the main purpose of our study focuses on default configurations.
- For higher bandwidths there is a significant increase in bandwidth variability. Also we can observe that average retransmissions also increased. As we are seeing retransmissions there could be an occurrence of network bottlenecks and thus for higher bandwidths there are more retransmissions and variability. We have also observed that Azure uses TCP cubic during our iperf testing as a transport layer protocol, thus high average retransmissions could occur due to congestion in the network.
- Another interesting finding is about the increase in bandwidth variability in scenarios of virtual machines on different vnet while compared to virtual machines on the same vnet. This is because Virtual machines are on different vnets. Azure allows all resources in the same vnet to have a secure connection for communication thus reducing the number of hops to connect to a virtual machine in the same vnet in figure 9. On the other hand, for Virtual machines in different vnets we need to route between different vnets and hence the network overhead increases, which could result in an increase in chances of facing a network bottleneck or a congestion in the network while routing. This also explains the higher retransmissions for virtual machines in different vnets.
- While comparing the 10 second stream of iperf test with 5 second stream of iperf test experiment across all the scenarios we can see that the average retransmissions has increased. As we increase the time to transmit continuously, there could be a chance of flooding the network and occurrence of congestion, thus leading to higher bandwidth variability.

5.2. Latency:

We calculated latency on the same F series Azure clusters that were used for calculating bandwidth. In this experiment we only considered 10-second streams of iperf tests in which we transfer data for 10 seconds and wait for 30 seconds. This process was repeated for two days. From the data obtained we plotted the below CDF and box plot. From the below plots we can observe the variability in latencies.

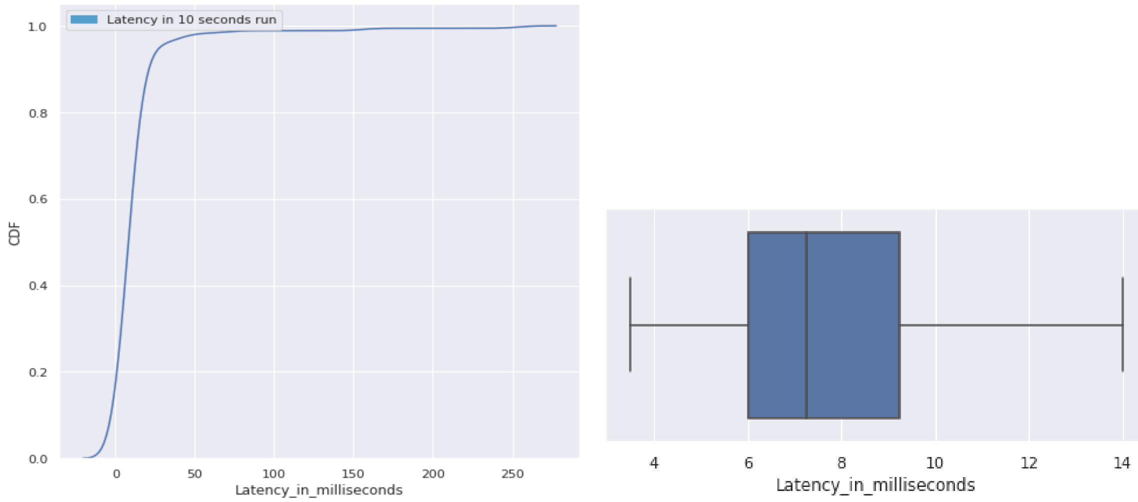


Figure 10: CDF and boxplot for latency observed in 10 seconds stream of iperf test

In the CDF we can observe that the maximum latency was 250 ms. This behaviour is because after transferring data at full speed, the bandwidth of the VM throttles and hence the latency increases.

6. Key Observations:

Below are the key observations and additional findings from the measurement study:

- Using the Accelerated network feature (please refer to the appendix), bandwidth can be increased. But that doesn’t negate the impact of variability
- We observed that the US-East region has a slightly greater bandwidth when compared to the Asia region. This could be because the US-East region is most popular and they have a larger network footprint.
- While testing variability for instances in different regions, we noticed that the iperf test terminated frequently, which can be the future scope of the project.

7. Conclusion:

Big Data operations are moved to the cloud as it provides on-demand instances.Variability plays an important role as big data technologies like spark which internally uses spark context to transfer data across nodes and perform computation. We extended the idea of the paper to determine the impact of variability on the Azure cloud platform. From our Google Cloud Experiments, we successfully validated our measurement setup. For the Bandwidth test we performed on Microsoft Azure, across all the scenarios we observed a variability in bandwidth. Our observations also state that when there is an increase in bandwidth, there is an increase in retransmissions and variability. We also found that there is a decrease in bandwidth when virtual instances are in different regions. Another interesting observation was that when virtual machines are in different Vnet, variability and retransmissions increased when compared to both virtual machines on the same Vnet. Incase of latency, we see that latency varied between 3ms to 14ms with the maximum variability going all the way to 250 ms due to bandwidth throttling.

8. Reference:

[1] Alexandru Uta1, Alexandru Custura1, Dmitry Duplyakin2, Ivo Jimenez3, Jan Rellermeyer4, Carlos Maltzahn3, Robert Ricci2, Alexandru Iosup1. In *Is Big Data Performance Reproducible in Modern Cloud Networks?*

[2] https://github.com/esnet/iperf

[3] https://www.wireshark.org/docs/man-pages/tshark.html

[4] A. Trivedi, P. Stuedi, J. Pfefferle, A. Schuepbach, and B. Metzler. Albis: High-performance file format for big data systems. In *2018 USENIX Annual Technical Conference (USENIX ATC 18)*, pages 615–630, 2018.

[5] https://data-flair.training/blogs/learn-apache-spark-sparkcontext/

[6] https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-overview

[7] https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-selecting-vm-size

[8] https://docs.microsoft.com/en-us/azure/virtual-network/create-vm-accelerated-networking-powershell

9. Appendix:

1) The Accelerated Network feature enables the user to increase the performance parameters like bandwidth and decrease latency & retransmission which is provided by Azure. This enables single root I/O virtualization (SR-IOV) to a VM, greatly improving its networking performance. SR-IOV allows a device to separate access to its resources, thus reducing network throttle. We can infer this from figure11.

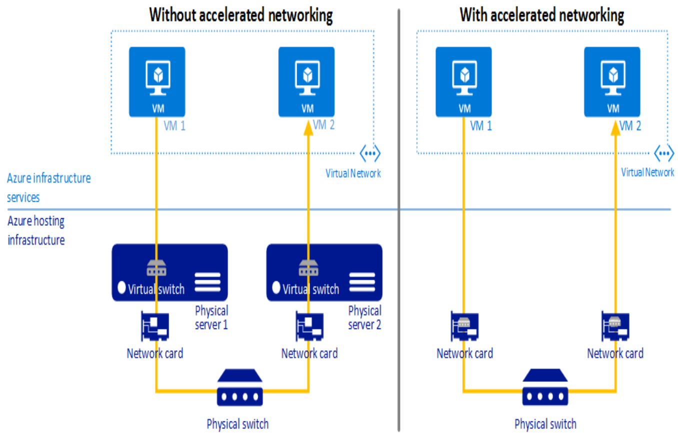


Figure11: Design of communications between 2 virtual networks in Accelerated Network and non Accelerated Network

For all the above experiments the costs incurred were as follows.
a) Cost for running GCP instances $50
b) Cost for running Azure instances $150
Keeping the cost in mind we could not do an in-depth measurement on other scenarios like enabling accelerated networks or testing on different regions apart from US-East and US-Central.