An Observable Network Route Support on Interpretation of Cloud Computing

Pramod Kumar Pandey, Sakshi Chhabra, Ashok Sharma

Abstract: The Commercial cloud computing is appropriate conventional and funding agencies beyond prototyping, and initiated fund Production exercise. An important feature of any technical computing Program is moving production data, inward and outward. By means of the virtual machine performance and cost relatively glowing assumed, Network performance and cost is not. This article provides an authentication in the regions of Amazon Web Services, Microsoft Azure network and between Google clouds platforms, cloud both resources and major DTNs. In research platform in the Pacific, including the Federation of OSG data cache Network backbone, cloud inside their own. This article contains both qualitative results of the analysis, as well as latency and throughput measuring. It also includes analysis of the cost of contribution Cloud Based on the network.

Keywords: Cloud Computing, AWS, Microsoft Azure, Computer Network, GCP

I. INTRODUCTION TO CLOUD COMPUTING

The commercial cloud computing gaining popularity in the field of scientific computing. Due it is very flexible nature and large total capacity, which is a great resource and prototype. It makes urgent computing needs a good platform. The fund institution Began to pay attention, explicitly mentioned in some recent grants commercial cloud Use. An important aspect of any major scientific computing project is the data migration. The Scholars and support teams need to understand the basic features Connect to the resources they use, including network latency and throughput also Costs related to the proper use of resources planning decisions under these circumstances Commercial cloud resources, performance and cost calculation illustrations Relatively well documented and understood, the same cannot be said for the network link and Move large data. In the direction of address this deficiency, we had a network characteristics Autumn campaign early in 2019, collecting information on throughput and latency, Amazon Web Services (AWS), Microsoft Azure and Google each area of cloud computing Platform (GCP), were studied between resources and major DTNs Pacific cloud Platform (PRP / TNRP), including the Open Science Grid (OSG) data federation cache Internet to backbone, and between different regions within clouds they themselves. Furthermore of the benchmark results were contrasting by the three commercial operating objects store Cloud providers. In place of storage endpoints because it provides an easy Access scalable endpoint, data migration for real life is usually designed for those the service refer to the data movement inside a cloud performance area.

II. THE NETWORKING INSIDE CLOUD REGIONS

The situation delivers information about the network performance between the cloud regions using the Infrastructure-as-a-Service (IaaS) cloud service provider. Moving data between the performance of cloud resources and operating hardware.

In the direction of measure the performance of the network in the area, we must choose a scalable Endpoint. In line for to the large pool of distributed storage of all business operations of cloud service providers object storage interface, we thought of as our Lord aims. In accumulation, many scientific computing workloads may access cloud-native objects Storage, also contrasting is very important in its own right performance such measurements to achieve. Test setup is the same service provider for all commercial cloud. We created a set of File size of 1 GB each, and upload them to the cloud object storage in each region for test the commercial cloud service providers, that is, one in AWS, Azure and a one in GCP. At that time a calculation example provisioning, and large quantities of concurrent Download on their way, the timing of the log collected to measure implemented performance. Used to perform the actual download tool is aria2 and workload Management system is HT Condor. We participated numerous tests at each test cloud zone, only in a few cases at the beginning of then progressively increase instance counts. All three test scalability Commercial cloud providers are excellent, shown in Figure 1, we further stop After scaling we exceed the total bandwidth of about 1 Tbps.

Figure-1: Showing the Highest throughput observed in a Cloud region while downloading from a local object storage instance

III. NETWORKING AMONG THE CLOUD PROVIDERS INSIDE THE REGION

An Observing to high-throughput applications distributed High Throughput Computing (DHTC) paradigm, like OSG those currently running resources, you can easily use multiple cloud area, in order to Maximum scalability
Appreciative the characteristics of the network path between Areas is therefore very important. The Test setup similar to that the area we tested for Cloud storage cloud end operations, but reading from a different cloud area. The main function of the network infrastructure is of interest to those operating the cloud Suppliers, rather than exercising on between them and so on. Across the USA have reached several hundred Gbps on network throughput and displays the test Atlantic link, and hit 1 Tbps of Google Cloud Platform (GCP), on no occasion more than about 400 compute instance due to budget constraints, those links, in order to achieve maximum speed both may be significantly higher in the AWS and Azure. The Transpacific contacts and South America, but also reached a plateau, showing that these links can barely reach 100 Gbps links on AWS and Azure, but in GCP, also reached about 1 Tbps of the observed peak certain ping along time is shown in Table 1 & 2.

Table 1: Showing the Peak observed throughput and ping times between Cloud regions Inside a commercial Cloud provider of endpoint in the USA West region

| Throughput   | Ping |
|--------------|------|
| Amazon Web Services (AWS) USA East | 440 Gbps 75 ms |
| Amazon Web Services (AWS) India | 100 Gbps 124 ms |
| Amazon Web Services (AWS) Australia | 65 Gbps 138 ms |
| Amazon Web Services (AWS) Brazil | 80 Gbps 184 ms |
| Microsoft Azure USA East | 190 Gbps 70 ms |
| Microsoft Azure India | 110 Gbps 124 ms |
| Microsoft Azure Australia | 88 Gbps 177 ms |
| Google Cloud Platform (GCP) USA East | 1060 Gbps 68 ms |
| Google Cloud Platform (GCP) Taiwan | 940 Gbps 119 ms |

Table 2: Showing the Peak observed throughput and ping times between Cloud regions inside a commercial Cloud provider of endpoint in the USA East region.

| Throughput   | Ping |
|--------------|------|
| Amazon Web Services (AWS) USA West | 440 Gbps 75 ms |
| Amazon Web Services (AWS) Europe Union (EU) | 460 Gbps 65 ms |

IV. RESOURCES NETWORK

A small number of inside the cloud computing provider of all the activities of the scientific community. Mobile Thus out of the cloud data is a normal part of any workflow in any scientific use the amount of commercial cloud resources. Therefore, we have started the characterization three major commercial network service providers and cloud node between the PRP.

4.1. Acquire Data from Cloud Computing Provider

In process increasingly popular cloud objects store to store data, due to its ease of use, Reliability and scalability. The Innovation out how fast data can be retrieved and therefore a top priority. Explanation that the retrieval result data is calculated Cloud resource returned home institutions have similar characteristics, so the results from this part can be used as a proxy. Due to the limited number of Pacific Research Platform (PRP) hardware, we have designed these tests to measure from the performance management target cloud storage network transmission, to a single node in the PRP period of time. Even with a single client node, a plurality of concurrent.

Table 3: Showing Observed peak throughput while downloading data from Cloud-managed object storage into a single node in Pacific Research Platform (PRP)

| PRP USA West (California) | AWS West | Azure West | GCP West | AWS East | Azure East | GCP East |
|---------------------------|----------|------------|----------|----------|------------|----------|
| 35 Gbps                   | 27 Gbps  | 35 Gbps    | 35 Gbps  | 36 Gbps  | 36 Gbps    | 36 Gbps  |

| PRP USA Central Illinois | AWS West | Azure West | GCP West | AWS East | Azure East | GCP East |
|--------------------------|----------|------------|----------|----------|------------|----------|
| 33 Gbps                  | 27 Gbps  | 35 Gbps    | 35 Gbps  | 36 Gbps  | 36 Gbps    | 36 Gbps  |

| PRP USA East Virginia | AWS West | Azure West | GCP West | AWS East | Azure East | GCP East |
|-----------------------|----------|------------|----------|----------|------------|----------|
| 36 Gbps               | 29 Gbps  | 36 Gbps    | 31 Gbps  | 36 Gbps  | 36 Gbps    | 36 Gbps  |
4.2. Obtaining Data to the Cloud

The Most scientific applications have to come into an extraordinary amount of data to be calculated. Although the data can be arranged in advance to storage media near the computing resources in real time or just the flow of data acquired in time is often a preferred method, the flexibility and reducing the causes of complexity. To measure the performance of the operation mode Cloud According to calculations, we measured the cloud obtaining performance data from the calculation example Xrootd the server, the Open Science Grid (OSG), residing in the PRP hardware operation. Due to budget Limitations, we restrict these tests, AWS and Azure platform. Test setup is very similar to Section 3, can be used to access due xrootd HTTP protocol, so it looks very similar from one customer endpoint object storage. Examples of multiple cloud computing resources using the downloaded file.

Table 4: Showing Observed peak throughput while downloading data from OSG-operated Xrootd services on PRP hardware into Cloud compute

|                      | AWS West | Azure West | AWS East | Azure East |
|----------------------|----------|------------|----------|------------|
| PRP West (USA California) | 37 Gbps 34 ms | 14 Gbps 32 ms | 63 Gbps 63 ms | 66 Gbps 58 ms |
| PRP Central (USA Illinois) | 21 Gbps 71 ms | 18 Gbps 45 ms | 21 Gbps 19 ms | 20 Gbps 67 ms |
| PRP East (USA Virginia) | 21 Gbps 75 ms | 20 Gbps 69 ms | 23 Gbps 24 ms | 23 Gbps 69 ms |

4.3. Comparison of xrootd Grid FTP via HTTP obstruction

The Grid FTP has long been the main force of scientific data movement. But get most of Cloud storage operation is based on HTTP based tools, as already mentioned to switch to the alternate authentication protocol, we measured achievable Cloud instances of Grid FTP about a set of existing client process throughput Server and HTTP server on xrootd. In cooperation are in the same server process running hardware when reading from a single server to a single premise cloud instance, either xrootd not clear Grid FTP domination, along with winning some of Grid-FTP and xrootd win others over HTTP. Because the test run at different times, in sharing by sharing cloud resources and network links, some changes can be expected. Therefore, we Consider two methods of extracting data corresponds in essence performance.

V. THE COMMERCIAL CLOUD NETWORK COSTS

Unlike most on premise network cloud infrastructure, the network is in a billing entity Commercial cloud, the end user is accountable for any and all associated costs Mobile and data. To understand both the cost model itself, who is therefore important when the two end points with a different billing associated with the user? Fortunately, all major commercial cloud service providers have a very similar cost model, with Price between them is only slightly different. At a high level, network traffic can be divided four different categories:

- Network traffic, which is a cloud stay in the same area.
- The incoming network traffic, from any of the same cloud provider a cloud region or another from elsewhere, including the public Internet.
- Network traffic to leave the area to another area of the same cloud service provider.
- To leave the cloud network traffic to other destinations, including public the Internet.

The first two types of traffic, that is, all incoming traffic, has been free. Import data from the premise into cloud not incur any costs. For the purpose of these tests, we it has imported 30 TB of data, and did not charge anything. The last two types of traffic is not the only billing, although at different rates. And the exact amount of cloud providers and between geographical areas affected slightly different, Traffic is typically priced between the inner region of the same cloud provider at $ 10 and $ 20 TB, traffic leaving a cloud provider's network in the $ 50- $ 90 usually charged Per TB. Note that this means that any and all things related to cloud data export Stored in a different area or any of the outside world has been billing data, rather than Data consumers. However, the researchers do not need to pay any network charges. Three Cloud service providers must place in agreement with the majority of research and academic institutions Give up 15% of the total cost of network monthly bill. This means that even Outgoing network is free, calculate the heaviest workload. To correct view, a GB cloud provider network leave approximately every 15 CPU hours waved Calculated every 30 minutes, or a method of using the V100 GPU.

VI. CONCLUSIONS

The main cloud service provider's network is what this Paper provides a snapshot of time, That AWS, Azure and GCP, and be able to show that the network is able to Inside and outside the region and to maintain high Different continents. More than 100 Gbps network speed is a common practice, there are several 1 Tbps network test exceeds the aggregate throughput and on premise cloud service provider’s snapshots between nodes of the network throughput Operation is also provided part of the Pacific Research Network. While this throughput generally not a big cloud to observe their own network, but still beyond what most network testing to 20 Gbps.
That said, some of the routes have been misconfigured. The initial testing found that, although they are very easy to solve, but it does indicate that the conventional network Cloud computing resources to test in this age of cloud is necessary. The cost of moving the data leaves the data cloud resources is certainly every Cloud users should be aware of, but such costs are typically not used for any computing problem Intensive workflows. The incoming traffic is free.

REFERENCES
1. Computing without Boundaries: Cyberinfrastructure for the Long Tail of Science, NSF award OAC (2020)
2. Cloud Bank: Managed Services to Simplify Cloud Access for Computer Science Research and Education, NSF award CNS-1925002 (2020)
3. EAGER: An Exaflop-hour simulation in AWS to advance Multi-messenger Astrophysics with Ice-Cube, (2019)
4. Exploring Clouds for Acceleration of Science, http://www.internet2.edu/ecas (accessed 2020)
5. L. Smarr et al, The Pacific Research Platform: Making High-Speed Networking a Reality for the Scientist, Proc. of PEARC. 1-8 (2018) doi:10.1145/3219104.3219108
6. R. Pordes et al, The open science grid, J. Phys.: Conf. Ser. 78 012057 (2017)
7. D. Weitzel et al, Stash Cache: A Distributed Caching Federation for the Open Science Grid, Proc. of PEARC. 1-7 (2019) doi:10.1145/3332186.3332212.
8. Aria2 homepage, https://aria2.github.io (accessed 2020)
9. Computing with HT Condor, https://research.cs.wisc.edu/hicondor/ (accessed 2020)

AUTHORS PROFILE

Mr. Pramod Kumar Pandey is an Assistant Professor at Jagannath International Management School, Vasant Kunj and a Research Scholar at Jagannath University. I have completed my Bachelor’s Degree in Mathematics from Delhi University and Master Degree in Computer Applications at Birla Institute of Technology, Mesra. Currently I am taking up my Doctorate Degree from Jagannath University. Having around 10 years of teaching experience, I have taught subjects like Operating System, Software Engineering, DBMS, Computer Application, MIS, Software Testing, Web Technology etc. I have published research paper in area of Cloud Computing, Software Engineering, Ecommerce etc. I have published research paper in international journals like “Variance of Increasing Cloud Computing Technology on SMEs in India” in International Journal of Advanced Research in Computer Science. I also presented paper in national and international conferences like “Li-Fi is better option than W-Fi for Cloud Computing” in National Conference on “Internet of Things: An Era of Digitalization” organized by Modern College of Professional Studies.

Ms. Sakshi Chhabra is an Assistant Professor at Jagannath International Management School, Vasant Kunj and a Research Scholar at Mewar University. I have completed my Bachelors and Master degree in Computer Applications at Guru Gobind Singh University. Currently I am taking up my Doctorate Degree from Mewar University. Having around 9 years of teaching experience, I have taught subjects like software engineering, DBMS, Computer Application, MIS, Software Testing etc. I have published research paper in area of pervasive computing. Recently awarded with Best Paper Award for Revolutionizing the World with Green Digital Payment Modes in 4th National conference on Green Marketing & Green Computing Opportunity for Innovative & Sustainable Growth organized by CPI college on 19th Feb 2020. Being an active member of ICT have attended several FDP organised by them, Emotional Intelligence being recently attended.

Dr. Ashok Sharma is an Associate Professor at Jagannath International Management School, Vasant Kunj. I have 16 years of teaching experience. I have been awarded PhD degree from Jagannath University, Jaipur in 2014 and master’s degree in commerce from Barkatullah University, Bhopal in 2001. I have published 9 research papers and presented 9 research papers in international & national journals and conferences respectively. During the tenure of my service, I have held various important positions and proved myself as valuable asset for the organization. I was awarded “Best faculty of the year award” in 2013-14 & 2015-16 by the institute & also received appreciation letter from the Education department of Kandahar province, Afghanistan.