Comparing the Network Performance of AWS, Azure and GCP

Archana Kesavan, ThousandEyes
Agenda

• Report Genesis & Overview
• Research Methodology
• Research Findings
• Summary & Recommendations
• Q&A
HOW DO WE MAKE DECISIONS ABOUT THE CLOUD TODAY?
Typical Cloud Decision Factors

- Global Presence
- Pricing
- Services
- Performance
What’s Been Missing from Cloud Performance Data?

- No network architecture
- No loss or jitter
- No predictability analysis
- Single provider
- Lack comparisons
- Limited geos
- No multi-cloud

DEPTH
TIME
BREADTH

Snapshots not trends
No network architecture
No loss or jitter
No predictability analysis
Data Collection Methodology

- Software probes (monitoring agents) emulate user-data by triggering a “test”
- Customized Traceroute
  - TCP Traces. Resistant to ICMP issues caused by rate-limiting routers.
- Bidirectional
  - Basic traceroute is directed (source-to-destination)
  - Bidirectional tests to account for variation in forward and return routes

MONITORING AGENT

- Bi-directional Tests
- Loss Latency Jitter
- Network Path Data

TCP Traces
Data Collection Methodology

- A single “test” gathers different types of data.

- Network Path Data:
  - Geo-located Layer 3 hops from source to destination including AS paths that allows us to determine connectivity architecture to the cloud providers.
  - Per hop loss and latency.
  - DSCP re-marking across the path.
  - MPLS tunnels.

- End-to-End Network Metrics:
  - Loss, Latency and Jitter.
  - Helps understand user-experience to cloud providers and within cloud providers.
Data Processing Methodology

• Software probes are well provisioned globally with continuous access. Allows easy detection of connectivity issues or local-faults

• Tests are run periodically, at an interval of 10 minutes

• Data gathered over a period of 30 days (07.01.2018 – 07.31.2018) to accommodate for extraneous events such as outages
  • Note: Zero outages occurred during the data collection timeframe

• Data generated is continuously exported to a cloud-based platform for analysis and trend-detection

• Network metrics generated compared across all three providers
  • Latency differences across providers prominent
  • Loss and Jitter were negligible
Test Scope

- **End User Measurements**
- **Inter-AZ Measurements**
- **Inter-Region Measurements**
- **Multi-Cloud Measurements**
End User Measurements

27 Global Locations

1485 Tests

55 Regions

15 AWS

25 Azure

15 GCP
Inter-AZ Measurements

- Inter-AZ tests are measured within a single cloud provider
- Availability zones are mapped independently for each account. Data samples from multiple AZ pairs analyzed to discount for exceptions.
- AWS: 4 regions
- Azure: 1 region (Concept of AZ relatively new for Microsoft)
- GCP: 4 regions
Inter-Region Measurements

15 REGIONS
Multi-Cloud Measurements

Azure

25 REGIONS

aws

15 REGIONS

Google Cloud Platform

15 REGIONS
Data Points Gathered

30 DAYS

10 MIN-INTERVALS

40M Unique Measurements

Loss
Latency
Jitter
Network Path

160M Unique Data Points
Inter-AZ performance is reliable and consistent.
Indicates robust regional backbone for redundant multi-AZ architectures.
Inter-Region Performance

- Inter-region network connectivity stays “within” provider network
- Performance baselined with Internet averages to reflect relative performance
- Most region pairs perform well but exceptions exist
How to Choose Inter-Region Pairs?

<table>
<thead>
<tr>
<th>Region Pair</th>
<th>Bi-directional Latencies (ms) from Sydney, Australia (Primary Region)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AWS</td>
</tr>
<tr>
<td>Tokyo</td>
<td>109.59</td>
</tr>
<tr>
<td>Singapore</td>
<td>174.65</td>
</tr>
<tr>
<td>Mumbai</td>
<td>228.48</td>
</tr>
</tbody>
</table>

10% faster than Baseline  | Baseline  | 10%-30% slower than Baseline  | 30% slower than Baseline
Where is End User Performance Strong?

- Comparable performance in North America and Western Europe
- No significant packet loss or outages in a 4-week period
Exception: GCP is 3X Slower from Europe to India

- GCP exhibited 3x the network latency in comparison to AWS and Azure.
- Circuitous route from Europe via North America to India on the GCP backbone.
Why is GCP Slower to Mumbai?

Users in Europe take a sub-optimal route to India

Traffic moves through the GCP backbone through the United States
No direct route from Europe to India

Why is GCP 3x Slower to Mumbai?
Exception: AWS Performance Anomaly in Mumbai

AWS exhibits the largest standard deviation in network latency in Asia

Considerably higher latency variation with AWS
AWS in Asia is Less Predictable

BI-DIRECTIONAL LATENCY VARIATIONS

Hosting Region: Tokyo
Hosting Region: Singapore
Hosting Region: Mumbai

AWS in Asia is Less Predictable
Why AWS Exhibits Less Predictability

Not all clouds are created equal
Architectural and connectivity differences have an impact on performance
AWS Loves the Internet

Traffic from users across the globe traverse the Internet longer on AWS deployments.

Traffic from Singapore enters the AWS backbone in Dallas, Texas.
Traffic from Singapore enters the MSFT backbone in Singapore
AWS Address Performance Instability With New Service

- AWS Global Accelerator for improved performance
- Pay AWS more $$ to ride the AWS backbone instead of the Internet
- Cloud Trend Alert: Monetization of the backbone
Multi-Cloud Performance

- AWS, Azure and GCP directly peer with each other
- Negligible packet loss and jitter
- Traffic between clouds does not leave their backbones

<table>
<thead>
<tr>
<th>Metric</th>
<th>AWS - Azure</th>
<th>Azure - GCP</th>
<th>GCP – AWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter</td>
<td>0.43ms</td>
<td>0.29ms</td>
<td>0.50ms</td>
</tr>
<tr>
<td>Packet Loss</td>
<td>0.01%</td>
<td>0.01%</td>
<td>0.01%</td>
</tr>
</tbody>
</table>
# Summary of Learnings

**NOT ALL CLOUDS ARE CREATED EQUAL**
Architectural and connectivity differences have an impact on performance

- AWS has less performance stability in Asia

**GEOGRAPHICAL PERFORMANCE VARIES**
Do not assume uniform performance across the globe and use data to test, validate and optimize

- GCP is 3x slower from Europe to India

**INTRA-CLOUD PERFORMANCE IS STELLAR**
Robust internal backbone supports redundant and distributed application architectures

- Inter-AZ performance between 0.5 ms – 1.5 ms

**MULTI-CLOUD IS SAFE**
Symbiotic coexistence between the Big 3 and stable multi-cloud performance

- AWS, Azure and GCP peer directly with each other
Recommendations

- Use data, not your gut to guide your cloud investment decisions.

- Find your “best” and ”worst” region pairs before deciding on your cloud architecture.

- Consider your organization’s tolerance to Internet exposure and evaluate risk vs $$.

- Factor in inter-AZ and inter-region performance. Your app performance and user-experience relies on it.

- Trust, but verify. Avoid assumptions in the cloud.
Our Research Vision

2018
160M

2019

Expand to Include Other Public Cloud Providers

Express Route, Direct Connect

Global Hot Spots
Thank You

For more PCPBR information, please go to
https://www.thousandeyes.com/research/public-cloud