Cloudyn

eBook

Who moved my cloud?

Part II: What's behind the cloud vendors AWS, GCE and Azure?

Part II of an eBook series covering cloud infrastructure and platform fundamentals not to be missed when preparing to make an organizational move to the cloud. In this part, AWS vs. GCE vs. Azure are covered for DevOps and IT Managers.

Intro

Cloud vendor competition is heating up to include a pricing war and new feature offerings in the battle for your business. To better assist in the decision making process, this eBook offers a concise breakdown of three cloud vendors: Amazon Web Services (AWS), Google Compute Engine (GCE) and Windows Azure (Azure).

Amazon is the clear leader with over eight years in the market with Azure following. While not directly behind, Google is still considered a visionary in the cloud arena based on their offering, go2market, enhanced performance and global infrastructure.

Contents

What's in a cloud vendor?

Despite Amazon's significant head start, other players have entered the market, making their mark, moving up the ladder and influencing pricing, features and expectations. Azure has been catching up with their cloud solution, Google entered in late 2013 with its Google Cloud Platform offering, and other players are also present, providing broad or niche solutions for different businesses.

Below is a high-level comparison of their features, global reach, pricing and more.

GCE vs. AWS vs. Azure: features and cost

Instance types

EC2 provides Current Generation instances split into **five families** with a total of 23 instance types, and Previous Generation instances split into **six families** with a total of 14 instance types; GCE splits its instances into **four families** with a total of 15 instance types; Azure provides only **two families** and 13 instance types.

Load balancer

Amazon's load balancing offering is divided for internal requirements: the Elastic LB inside regions and DNS service between regions, while also providing for external requirements, Auto-scaler to assist in case of external traffic spikes. Each is individually priced and added per company's need.

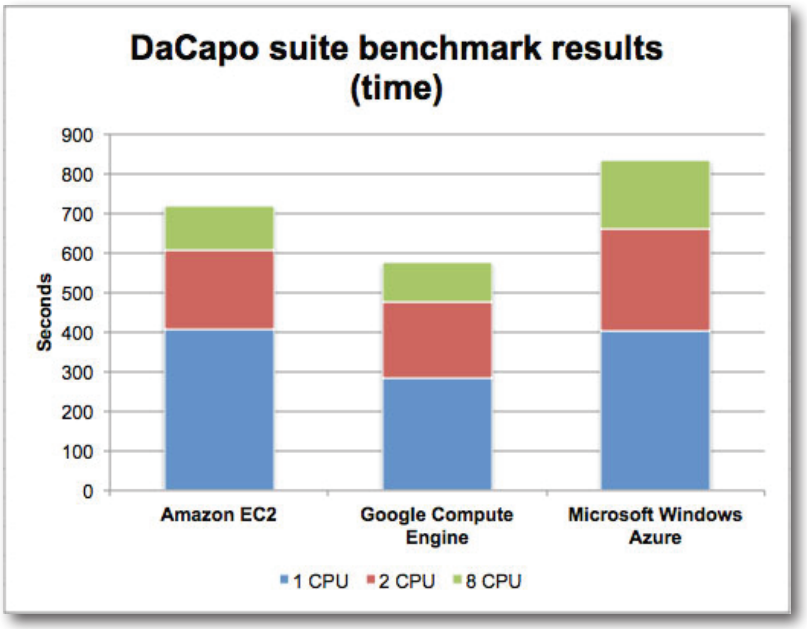
Google's solution and Azure's traffic management support both internal and external requirements listed above with one built-in feature (Azure with the Standard VM version, not Basic).

Of course, all three provide a solution, which is priced separately per requirement.

I/O speed

When configuring a deployment, there are many combinations and variations that could affect speed, such as CPU size and number of instances. While speed is important, it is only one parameter of many that should be considered in vendor selection, so benchmark test results should be taken with a grain of salt.

Overall, based on a test conducted by InfoWorld (image below), of the three vendors, Google is fastest, Azure slowest, and Amazon between the two.



Block storage

AWS and Azure persistent disks support up to 1TB, whereas GCE supports as high as 10 TB.

Global network

AWS has the widest range of region availability with more data centers and access points. They run using their own connectivity within the regions, yet by an open Internet connection between separate regions.

Google's has narrower availability and access, yet runs through its own regional fiber network, both within and across regions, increasing speed and scalable performance.

While Azure assigns quality of service to its traffic, it doesn't own any network infrastructure, so connection between regions runs over open Internet connection like AWS.

GCE vs. AWS pay models

Calculating instance cost is a complex task which requires looking at compute power, networking needs and attempting to calculate project lifecycle needs for proper capacity planning, resource and personnel requirement, and more. The various tools provided by the three vendors aim to make this process easier, but it still leaves many in the dark.

Below is a summary of the billing and price models:

Billing policies

AWS charges cloud usage on an hourly basis, meaning that even if a full hour of usage wasn't completed, the price will be rounded up. GCE and Azure both charge based on one minute intervals, with Google having a ten minute minimum to start.

AWS: the three pay models

AWS instance types and prices break into three categories: On-demand, Reserved Instance (RI) and Spot Instances.

On–demand

The most common instance type purchased, and the most expensive with the highest cost per hour charged for use (compared to the other two AWS options). Full price is paid for utilization and doesn't include cost of upgrading network speed.

With no long-term commitments required or upfront fees, it is the most flexible option which may be ordered on the fly when needed.

Reserved instances

Upfront fees are paid to "reserve" an instance for one or three year periods of time. Amazon discounts the hourly usage price in return for the long term commitment. This model deviates from the pay-as-you-go model, and some companies view the reserved instance as Capex as opposed to Opex, which undermines their original motive to migrate to the cloud.

Spot instances

A spot instance is acquired through bidding in an auction with Amazon determining the price based on demand. This instance is significantly cheaper than RI instances, but provides zero control over the lifespan of the instance provided.

GCE pricing models

On-demand

The most common instance type, with an hourly fee charged per specific use. Unless other GCE options are utilized, it is also the most expensive option.
This type allows for utmost flexibility, as resources may be added or cancelled at any time, and no upfront fees or long term commitments are required.

Sustained use

A retroactive discount is received at the end of the month for using the instance for an extended period of time. For example:

- A business will be charged for the full on-demand rate of an instance that runs for 25% of the month or less.
- For the next 25% of the month, a business will be awarded a 20% discount off the on-demand rate.
- For the following 25% of the month, the business will be awarded a 40% discount off the on-demand rate.

(These figures are as of April 1, 2014)

Inferred instances

The calculation of the sustained use discount includes inferred instances. Google defines an inferred instance as, "a combination of multiple, non-overlapping instances of the same instance type, in the same zone, into a single instance for billing."

Why GCE's Sustained Use Instance may be more attractive than AWS's RI

Unlike AWS's RI instance, Google's Sustained Use instance does not require any upfront costs as the discount is applied retroactively at the end of the month. Additionally, businesses are not committed to use a specific instance for any specific amount of time. This does, however make it more difficult to calculate a project's cost in advance as it is unclear until the end of the month.

Windows Azure pricing models

On-demand instance

The most common and expensive instance type, with hourly costs listed and calculated based on total minutes. No short or long term commitment is required, so high flexibility is available to increase or decrease instances as needed.

Six or Twelve month terms

Companies looking for discounted pricing can order six or twelve month terms, which offer a reduced price on each hour of use. There is a monthly model, as well as a pre-paid, the latter of course further reduces the cost with the upfront payment requirement.

	AWS	GCE	Azure
Features			
Instances: families and types	5 families 23 types	4 families 15 types	2 families 13 types
Load balancer	Elastic LB (internal) DNS service (internal) Auto-scaler (external)	Combined solution	Traffic management combined solution
I/O speeds	Second place	First place	Third place
Block storage	1 TB	Up to 10 TB	1 TB
Global network	Wide regional offering, most data centers and access points	Narrower regions, centers and access, faster per connection	Global datacenter infrastructure
Connection	Open Internet	Regional fiber network	Open Internet
Billing			
Pricing	Per hour – rounded up	Per minute – rounded up (minimum 10 minutes)	Per minute – rounded up commitments (pre-paid or monthly)
Models	On-demand, Reserved, Spot	On-demand, sustained-use	On-demand, short term commitments (pre-paid or monthly)

Conclusion

Companies preparing for a cloud migration must do their homework to find the best route for them. Each vendor has its own sweet spot for particular deployments and it is down to the deploying company to figure out which KPIs are most important to them and select a cloud vendor accordingly. Which one will fit and integrate with business needs best, as well as how it meets business' short and long term goals.