Serverless Computing, Artificial Intelligence, and VMware Collaboration: Amazon is the Public Cloud Vendor to Beat!

Introduction

55% year over year revenue growth and approximately $13 billion in 2016 total revenue (over three times the size of its closest competitor, Microsoft Azure) demonstrate Amazon Web Services' (AWS) strong momentum. Amazon claims that new customers are mostly attracted by two core advantages: rapid provisioning speed of easily consumable services and cost savings. Research by Enterprise Management Associates (EMA) confirms this claim and EMA continues to be impressed by Amazon's pace of innovation and bold investments in future technologies, such as serverless computing, artificial intelligence and the Internet of Things. Exactly this innovation-centric strategy made Amazon.com the sustained and undisputed market leader in online retail, e-readers and ebooks.

The sheer number of significant fourth-quarter announcements (28), combined with this year's size of the AWS re:Invent show in Las Vegas, NV, (32,000 attendees, 562 sessions, and 996 speakers) demonstrate that Amazon will not take their foot off the gas. This EMA Impact Brief provides an overview of the most strategically significant announcements and what they mean for customers and competition.

The VMware Collaboration: Bringing vSphere Customers into the AWS Universe

Later in 2017, customers will be able to run their vSphere-based applications inside Amazon's data centers. VMware will offer this option to its customers so they can rapidly create a hybrid or public cloud while still using the same VMware management and provisioning tools. VMware's upcoming Cross-Cloud Architecture will tie the vSphere and the AWS universe together.

> **VMware's upcoming Cross-Cloud Architecture will tie the vSphere and the AWS universe together.**

EMA expects that running vSphere on Amazon infrastructure will soon replace vCloudAir, which never got off the ground (despite big investments and bigger expectations). Traditional vSphere and vRealize Automation users will have access to VMware's kryptonite (serverless computing) in the form of Lambda.
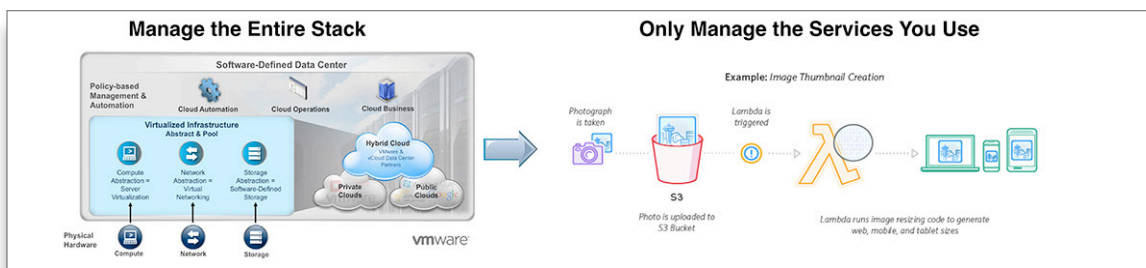

Figure 1 The Evolution from Managing Infrastructure (VMware Stack to the left) to Fully Focusing on Business Logic (Amazon serverless infrastructure to the right), Sources: VMware, Amazon

EMA believes that this will end in a one-way street, where customers modernize their vSphere-driven applications with AWS Lambda, AWS Step Functions, AWS Machine Learning, AWS Aurora Database, and other AWS services that all have one thing in common: they do not require server, storage, and networking infrastructure management. "Customers simply run their code, execute their queries, or design entire business applications without having to worry about architecting a secure, scalable, and well-performing environment," says Dr. Constantin Soeldner of Soeldner Consult, an IT consultancy that partners with both Amazon and VMware. "Deploying and managing this underlying infrastructure is VMware's core business today."

The Future is Serverless: Amazon is Betting Big on Lambda

More and more AWS offerings are charged based on usage, such as Lambda, Step Functions, Athena, Glue, Lex, and Polly. Basing offerings on usage will save customers a significant amount of CAPEX and OPEX by reducing overprovisioning.

> "Customers simply run their code, execute their queries, or design entire business applications without having to worry about architecting a secure, scalable, and well-performing environment," says Dr. Constantin Soeldner of Soeldner Consult, an IT consultancy that partners with both Amazon and VMware.

While this new way of charging for services means that Amazon will benefit less from overprovisioning, it will make AWS offerings more accessible (often POCs can be completed within minutes or hours) and more attractive from a cost perspective. This further enhances Amazon's competitive position against traditional infrastructure vendors, such as VMware, Nutanix, and HPE.

AWS Greengrass - Bringing the Edge to the Customer

AWS Greengrass enables connected devices to collect, process, act upon and store sensor data input while temporarily offline. AWS Greengrass core software runs on Linux OS with ARM or x86 CPUs and at least 128GB of RAM and enables connected devices to run Lambda functions locally, even when temporarily offline.

An AWS customer can use the IoT Device SDK to let other devices interact locally with a device running AWS Greengrass. The Greengrass Core stores "shadows" (cached copies) of each locally networked device, enabling all devices to interact by calling a central REST API to access either live or cached data from a virtual version of this device, if the respective device is offline. Once a device comes back online, the Core updates the corresponding shadow with the latest data (configuration data, sensor data, activity logs, etc).
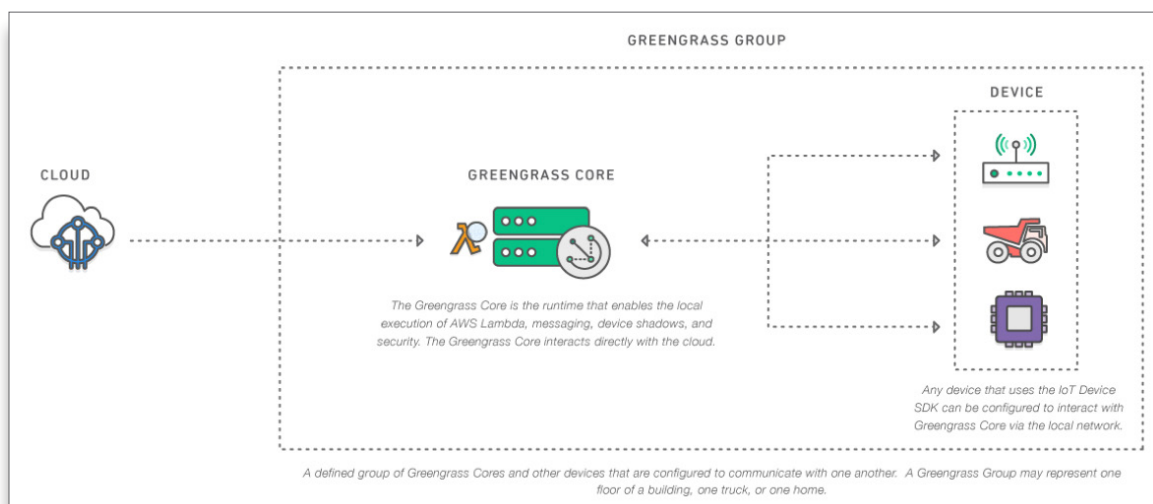

Figure 2 - The Greengrass Core Enables Applications to Interact with Intermittently Connected Smart Devices, source: Amazon

The Greengrass Core synchronizes data and communicates with local devices and AWS while operating under intermittent Internet connections. The secret sauce is an intelligent caching mechanism that ensures only essential data is sent back to the cloud, asynchronously if needed. When the Internet goes out, local devices running Greengrass can continue to operate, communicate and interact with their environment.

AWS Step Functions - Respond to HTTP Request via Lambda

AWS Step Functions is another new and serverless AWS offering that enables end users to create and modify applications by stitching together reusable components. These components can be Lambda functions, container-based microservices, or standard enterprise applications. AWS Step Functions offer a graphical user interface where sequential, parallel, or branching logic is visualized as a flowchart. AWS Step Functions record execution details, including time stamps and input/output variables. Customers can now easily offer complex AWS Lambda- and EC2-driven applications via HTTP requests to their internal business units or to their customers, as a free or paid service.

> **AWS Step Functions is another new and serverless AWS offering that enables end users to create and modify applications by stitching together reusable components.**
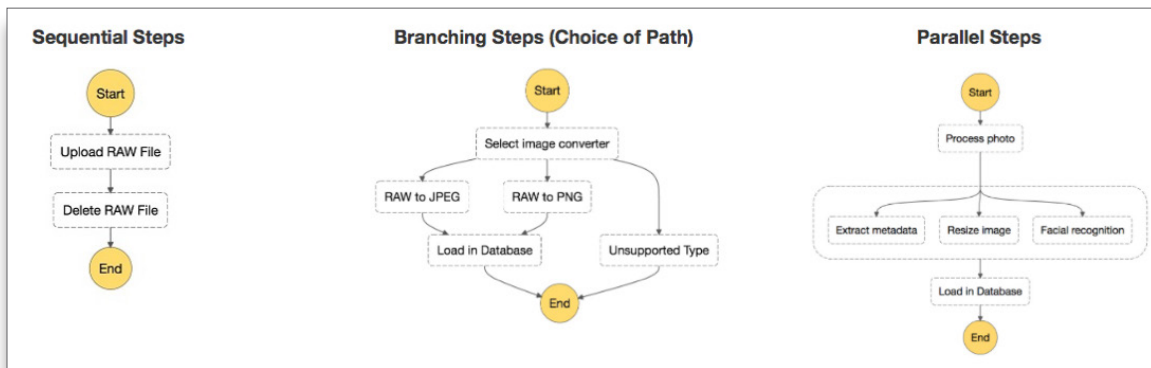

Figure 3 - AWS Step Functions - Define Application as a Series of Steps, Source: Amazon

AWS Glue - Simplified and Serverless ETL

AWS Glue is a database integration service using AWS Lambda functions with visual design tools to extract, transform, and load (ETL) data from Amazon S3, Amazon RDS, and Amazon Redshift. AWS Glue is serverless and offers a high degree of intelligence and automation to minimize coding and manual data processing when configuring ETL tasks.

The Future is Intelligent: Machine Learning and Artificial Intelligence

In the race up the stack, Amazon added three artificial intelligence and machine learning-driven services to the AWS portfolio: Amazon Lex, Amazon Polly, and Amazon Rekognition.

Amazon Lex - Powered by the same Deep Learning Technologies as Alexa

Amazon Lex offers Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU) interfaces as a service to build text and voice chatbots. Amazon Lex enables developers to easily create conversational interfaces for their apps without having to worry about creating the infrastructure and configuring the complex algorithms typically needed to enable this type of capability. Amazon Lex can trigger AWS Lambda functions that utilize other Amazon Web Services or query SaaS applications such as Salesforce or Zendesk.
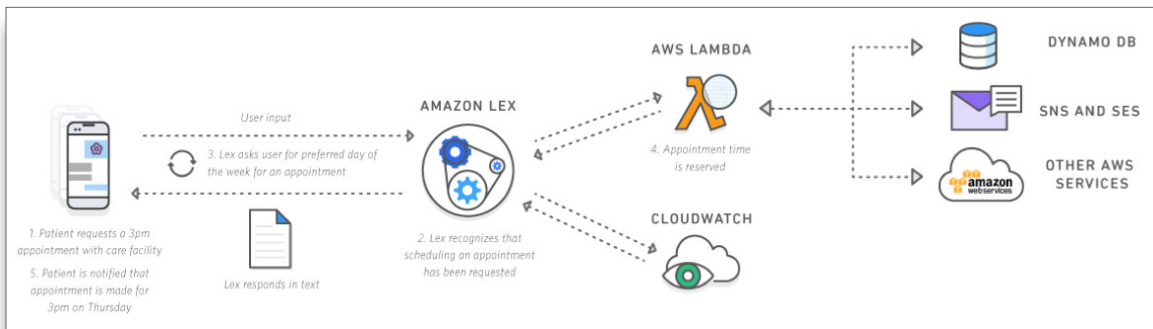
Figure 4 - AWS Lex Leverages Lambda and other Services to Create an Appointment Booking System, Source: Amazon

Home automation with Amazon Echo is the crossover point between Amazon.com and AWS, where Amazon is leveraging Alexa to offer AWS users a turn-key artificial intelligence service that rapidly and constantly learns from millions of Amazon Echo owners using their devices on a daily basis. Echo users are now "working for Amazon" by training the Alexa service through daily use. This reinforced learning approach not only teaches Alexa to better understand questions asked in plain language, but Alexa and Amazon Machine Learning services also find out more about the specific needs of different end user groups and individual end users. For example, Alexa knows that you are a football fan and which team you watch on TV, ordered merchandise from, and bought tickets for. Therefore, the next time you ask "Alexa, player stats for Von Miller" and Alexa cannot hear you properly because your baby is screaming in the background, she will leverage what she knows about your interests and search for football players first and then for Denver Broncos players whose names sound similar to what Alexa received via the microphone. If there's a Broncos player that sounds similar to the fragments she heard of "Von Miller," she will serve up this player's stats. Or, if you say "Alexa, buy football tickets for Sunday," Alexa will know that you are most likely talking about Broncos tickets and that you usually want three of them, as you typically take both of your kids. But if you ask "Alexa, buy tickets for Legoland next summer," Alexa may warn you that one of your kids will still be under the minimum age for most rides in the summer.

> **But if you ask "Alexa, buy tickets for Legoland next summer," Alexa may warn you that one of your kids will still be under the minimum age for most rides in the summer.**

This example illustrates how difficult it will be for the competition to catch up with Amazon in NLU and machine learning.

Amazon Polly

Amazon Polly converts text to speech that sounds like a human voice and lets developers create applications that talk, enabling new categories of voice computing products. Amazon Polly is used by Amazon Lex to convert responses to verbally asked questions into speech that can be streamed directly to the end user's device.

Amazon Rekognition

Amazon Rekognition enables developers to easily add image analysis into applications. It uses deep neural network models to detect and label images by extracting objects, scenes, and faces. This transforms image content into text that can be used to trigger AWS Lambda functions (e.g. a doorbell camera can use your image as a second authentication factor in addition to your smartphone).

AWS Snowmobile: Ingesting Data by the Truckload

This new service includes a 45-foot truck coming to the customer data center to download large file repositories (medical records, imaging, other media, etc.) to ship them to the nearest Amazon data center for upload to the AWS S3 object store or the AWS Glacier data archiving service. AWS Snowmobile shows that Amazon expects customers to migrate vast quantities of data to AWS storage. Of course, this data is then best consumed by AWS offerings, such as the new Athena pay-per-query service.

Figure 5 - The 45-foot Amazon Snowmobile Truck, source: Amazon

Security

As more and more enterprise workloads "flow" to AWS, security becomes a key concern. Two recently launched services aim to address some of the key security issues traditionally affiliated with public cloud in general.

AWS Organizations

AWS Organizations refers to the long-anticipated capability to manage multiple AWS accounts under one corporate account. Customers can still have the separation of applications or of development, staging, production, and disaster recovery accounts, but now both the business and the IT organization can enforce central governance and billing, with near-instant auditability of all environments.

AWS Shield - DDoS Protection

AWS Shield is Amazon's response to the rising threat of distributed denial of service attacks (DDoS) on the application layer. AWS Shield, available in Standard (free) and Advanced editions, supports AWS Elastic Load Balancing (ELB), AWS CloudFront, and AWS Route 53. Rules can be configured by customers or by the Amazon response team.

EMA Perspective

With these recent announcements, Amazon makes it hard for the competition to catch up in the race up the stack. Amazon's strategy is aligned around serverless computing with AWS Lambda at its core. Offering machine learning, artificial intelligence, and an advanced IoT platform in a serverless pay-as-you-go manner will hook customers quickly. VMware may experience firsthand how additive AWS serverless offerings are when making them available to vSphere customers later on in 2017.

Especially, financiers will look at AWS first when it comes to creating new business applications. AWS critics within the customer organization will have a harder time justifying the OPEX and CAPEX that would be needed to stand up to this type of solution on-premise.

If Amazon now closes its gaps in terms of enabling its monitoring and management solutions to include non-Amazon infrastructure and services, and if Amazon EC2 Container Service (ECS) stops relying on EC2 clusters that customers have to pay for whether they are used or not, Amazon will be hard to beat in the race in absorbing the lion's share of the 70%-80% of corporate workloads that will ultimately end up in the public cloud.