

re: Invent 2019 Keynote - Andy Jassy

65,000 attendees

Theme - innovators who have re-invented customer experience on the AWS tech platform

Transform/re-Invent customer experience

1. Don't Wait

Transformation is about leadership

- Senior team alignment (and conviction)
- Top down aggressive goal
- Train your builders
- Don't let paralysis stop you before you start

2. Don't stop me now..

"check box heroes"

IT Spend ~ 3% cloud 97% on prem
AWS 47.8% of cloud market

Compute - instances, containers, network

* Multi-cast teaching today

Instance innovation : AWS Nitro system - offload hypervisor functions from server to Nitro chip

Nitro also moves off control domain to chip

→ Re-invented virtualization layer

ARM based instances - Graviton chip from Amazon Labs

x86, R6g, C6g New instance type based on ARM graviton (v2 chip)
40% price/perf over x86 Graviton 2

Machine Learning - Inferences/Predictions majority of the cost (80%-90%) vs training

* Launch Inf 1 instances for EC2 backed by new Inferentia chips

EC2
Initial support EC2, EKS, Sagemaker

TensorFlow
PyTorch
MXNet

* Amazon Fargate for Amazon EKS

40% of new container customers to AWS choose Fargate

* Nitro Enclaves

2020 Transaction Banking - Digital Platform

Building for Developers

About 40 minutes
into the keynote

- Goldman Sachs
- published some APIs

3. Is That All You Get For Your Money?

on prem vs cloud

Mainframe decan - redo as micro services

- more on a per workload basis

Oracle

Windows

on Side Server

4. The Hunger Keeps on Growing Data

Data Silos to Data Lakes

S3 - most popular for Data lakes

* IAM Access Analyzer

* Amazon S3 Access Points

- Customized access permission rules for each of your applications

- Simplifies managing data access at scale for applications to access data sets

* Amazon Redshift - concurrency scaling

* Federated query - redshift, S3, relational db

* Data Lake Export

* Materialized Views

Many features released this year

* Redshift RA3 Instances with Managed Storage

Scale compute and storage separately

Hard to cost effectively scale without compromising compute

Compute not keeping up with increases in storage performance

* AQUA (advanced query accelerator) for ~~Amazon~~ Redshift
→ Moves the compute to the storage

- custom hardware to accelerate compression, encryption
- do the compute on the raw storage

Great companies - keep listening to their customers and innovating on their behalf

* UltraWarm

- warm storage layer for elastic search
- better durability, backed by S3

* Amazon Managed Apache Cassandra Service

Right Tool for the right job to save money, but best cap for your customers

5. I've worked 900 miles - Machine Learning

- Early Days: working with clunky tools

- 80% of TensorFlow in the cloud runs on AWS

- Big AWS investment in scaling ML across hundreds of GPUs, etc

"Right tool for the right job"

PyTorch and MXNet results too

AWS Services

ML Services - SageMaker

Partners - Experts

* SageMaker Studio - web IDE for everything needed for a ML project
one click notebooks with elastic compute

* Amazon SageMaker Experiment

* SageMaker Debugger

Concept Shift - changes where model performance declines

* SageMaker Model Monitor

* SageMaker AutoPilot

Auto ML with transparency / visibility / control

* Amazon Fraud Detector

- * Amazon CodeGuru - ML service to review your code
- add CodeGuru to your pull request reviewers

ML static analysis, ML profiling

* Contact Lens for Amazon Connect

- transcription, sentiment analysis, long periods of silence, etc.
- search over transcripts
- dashboards
- mid 2020 - real time dashboards

* Amazon Kendra - enterprise search with ML and NLQ

5. Break or Through to the other side - Help customers deal with
- Help to move some of your workload: "these need to stay on prem"

* AWS Outposts GA

ECS EBS EKS
EMR VPC RDS S3

Two floors - AWS on top
VM control plane

* AWS Local Zones

ECS FSx EBS
EBS VPC RDS

workloads in certain geographies that need single digit latencies

How should customers think about SG and how to leverage it?

- network latency going to 10ms
- more processing at the edge

low latency, massive throughput

* AWS Wavelength Build apps with single digit latency...

Perse - city aggr gate
(AWS)

- JMWSS
- we can do it better

Once in a life time chance to re invent