



# Amazon Web Services: An Introduction

[http://jkinney.s3.amazonaws.com/  
NASA\\_AWS\\_Overview.pdf](http://jkinney.s3.amazonaws.com/NASA_AWS_Overview.pdf)

Jamie Kinney

[jkinney@amazon.com](mailto:jkinney@amazon.com)



# Agenda – Day 1

- Introductions
- Discussion: Why are we here?
- AWS Service Overview and Drill-Downs:
- Hands-on Tutorial #1
- Scientific Computing in the Cloud
- Strategies for Enterprise Cloud Adoption

# Agenda – Day 2

- Oversight and Governance of Cloud Services
- Hands-on Tutorial #2
- AWS Support Services
- Ecosystem – Vendors, Tools, and Help
- Security
- Discussion: Accreditation, FEDRAMP, and the Road Ahead

# Your Instructors

Thomson Nguy

NASA Account Manager

[tnguy@amazon.com](mailto:tnguy@amazon.com)

+1.703.371.8541

Jamie Kinney

Solution Architect focused  
on the scientific community

[jkinney@amazon.com](mailto:jkinney@amazon.com)

+1.206.265.9439



# Why are we here?

---

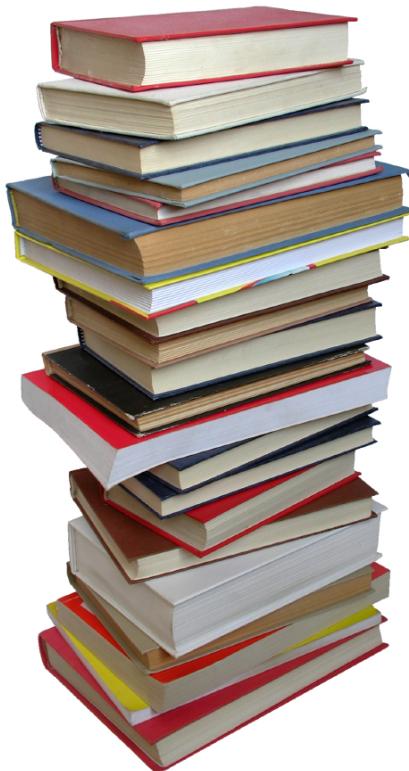
# Tell us about yourself

- Who are you?
- Where do you work?
- What do you do for NASA?
- What would you like to learn in the next 2 days?
- How would you like to use the cloud?

# AWS Service Overview and Drill-Downs

- Compute
- Storage
- Database
- Messaging
- Content Delivery
- Deployment, Management, and Monitoring

# How did Amazon Get into Cloud Computing?



# On-Premise Infrastructure is Costly & Complex

Large Capital Expenditures

Patching Software

Scaling down as needed

Contract negotiation

Prices too high for IT products

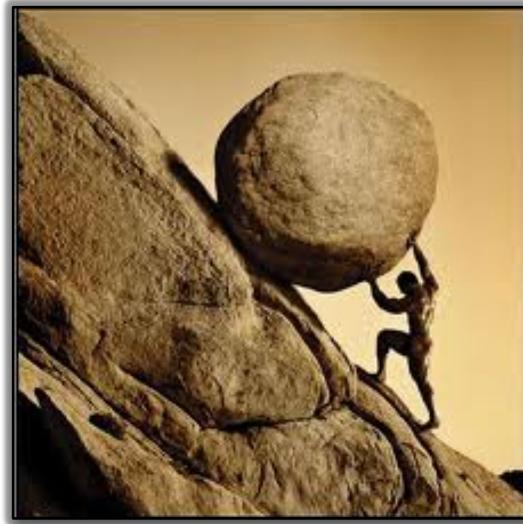
Underutilized IT Assets

Out of Datacenter Space

Slow IT Deployments

Scaling up quickly

Managing physical growth



“IT spends 80% of its time and resources keeping the lights on”

**Gartner**



# No Up-Front Capital Expense

## On-Premise

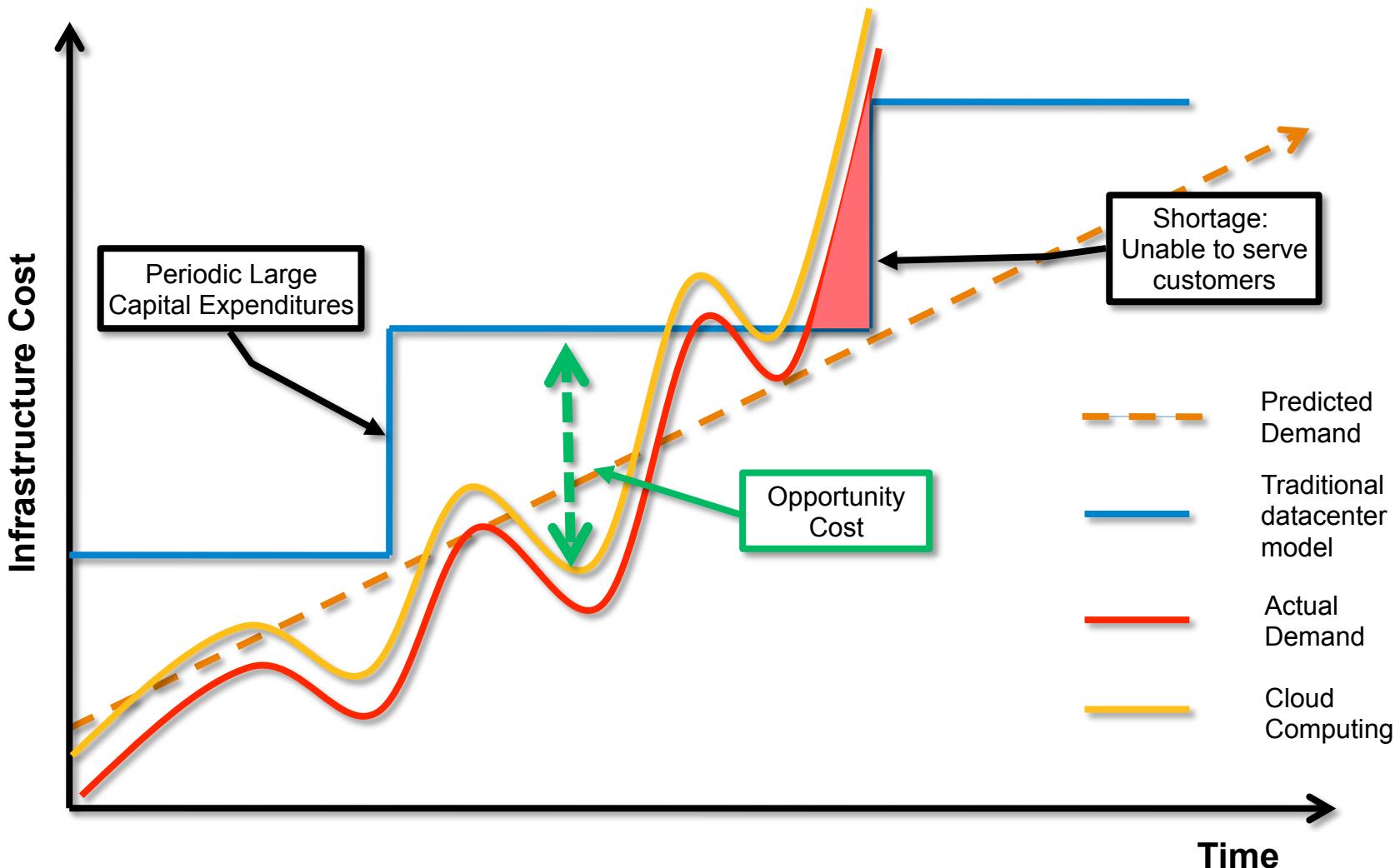
Physical Space  
Cabling  
Power  
Cooling  
Networking  
Racks  
Servers  
Storage  
Certification  
Labor

## Cloud Computing

\$0  
to Get Started



# Pay Only for What You Use

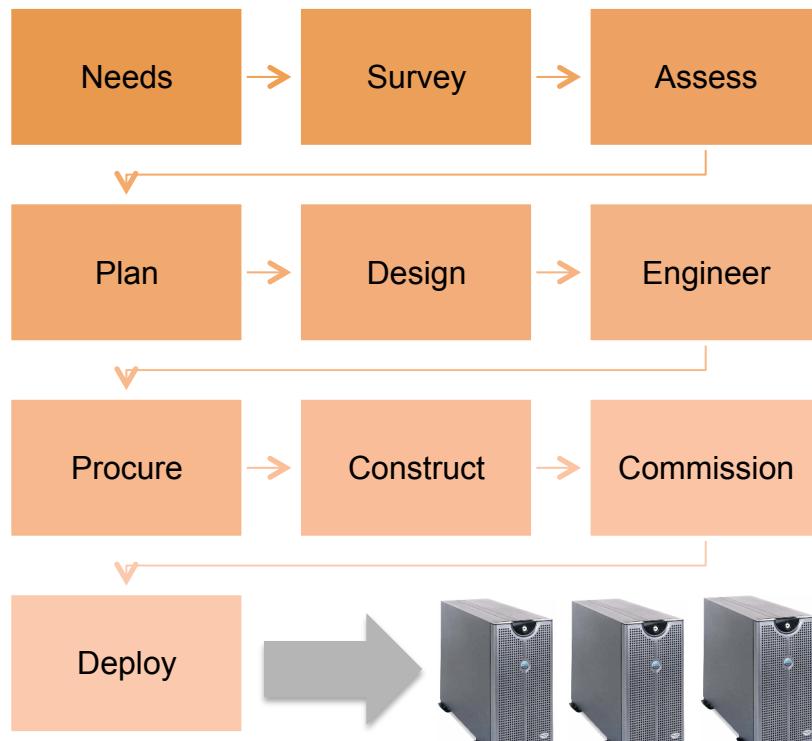


Deploy

# Traditional versus Self-Service Infrastructure

## On-Premise

Build new environments can be complex and slow



## Cloud Computing

New infrastructure can be deployed in minutes



New Development Environment



New Test Environment



New Environment in Japan



Add 1,000 Servers



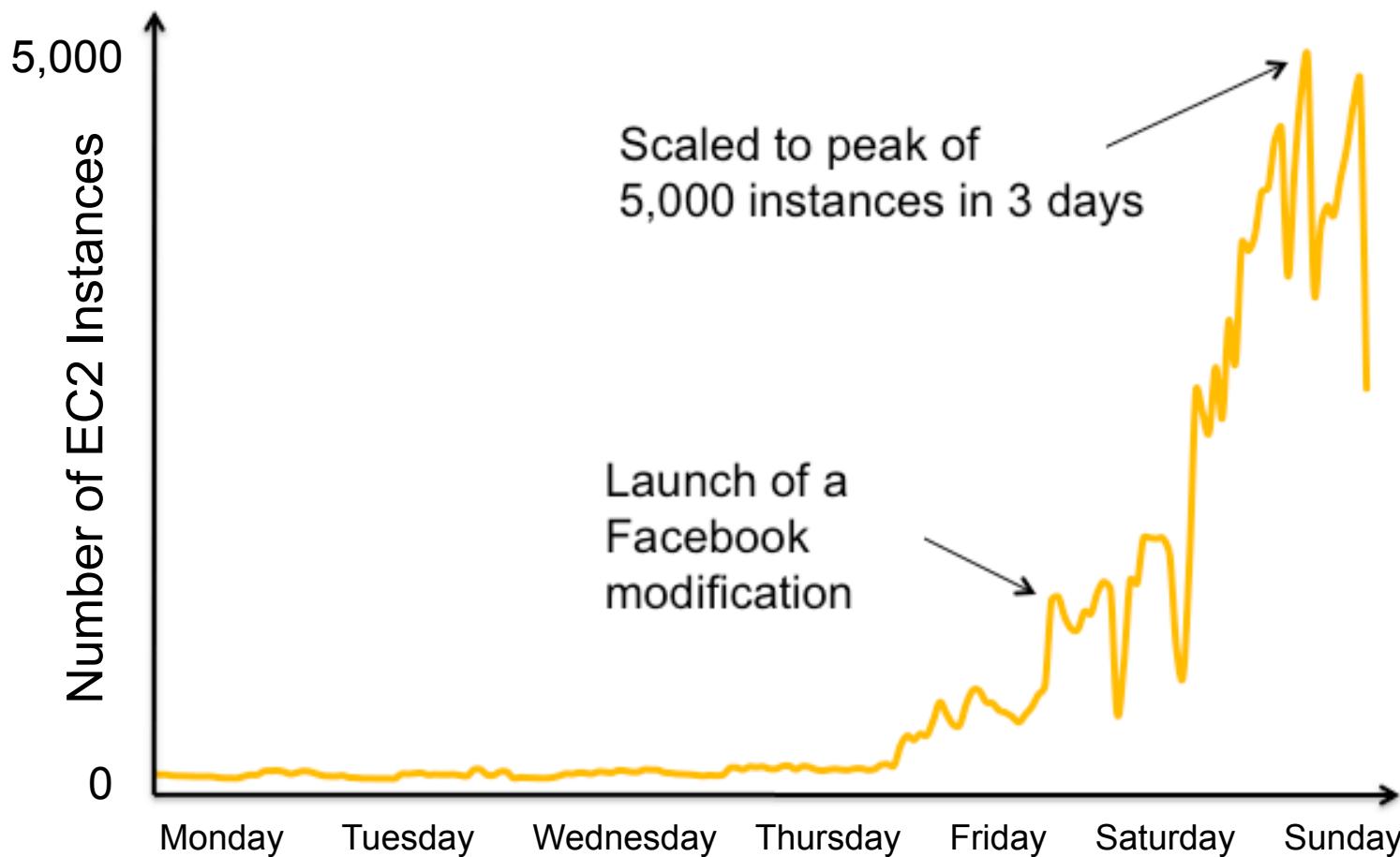
Remove 1,000 Servers



# Easily Scale Up and Down



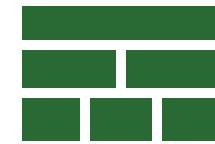
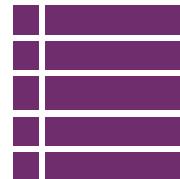
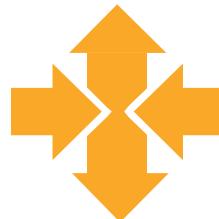
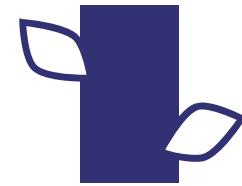
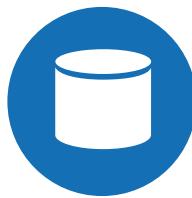
Internet Video App on Amazon EC2:  
From 50 to 5,000 servers in 3 days



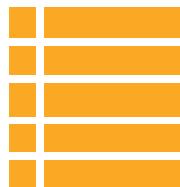
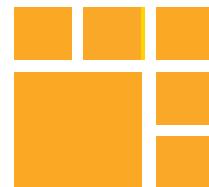
# Cloud Computing is More Than Just Virtualization

	Cloud Computing	On-Premise Virtualization
No Up-Front Capital Expense	✓	✗
Low Cost	✓	✗
Pay Only for What You Use	✓	✗
Self-Service Infrastructure	✓	✓
Easily Scale Up and Down	✓	✗
Improve Agility & Time-to-Market	✓	✗

# Overview of Services



Networking



Compute

Storage

Database

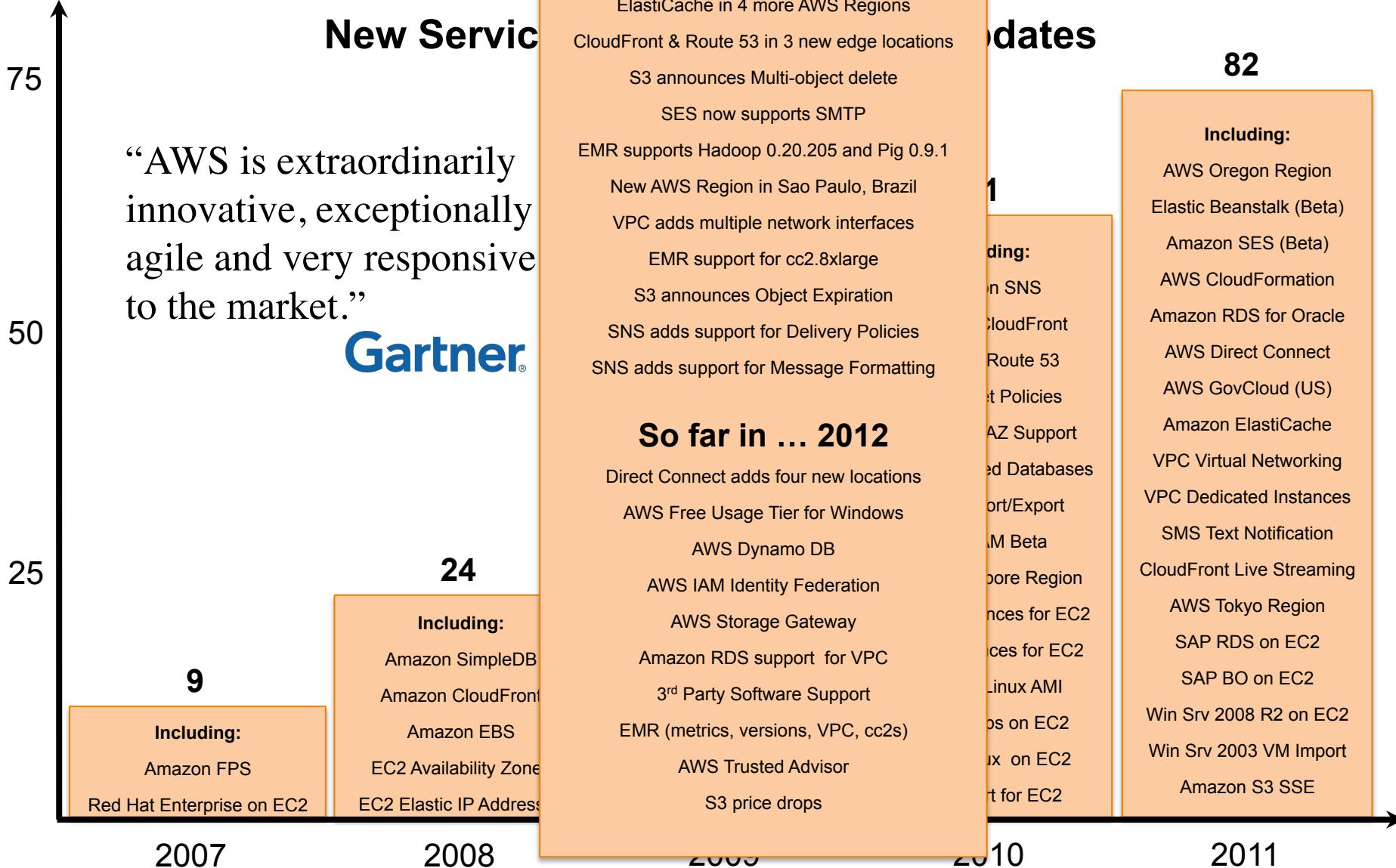
Messaging

Content Delivery

Monitoring



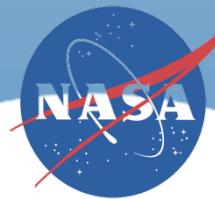
# AWS Pace of Innovation





## Common Use Cases

- Web site hosting
- Application hosting/SaaS hosting
- Internal IT application hosting
- Content delivery and media distribution
- High performance computing, batch data processing, and large scale analytics
- Storage, backup, and disaster recovery
- Development and test environments



# Website Hosting / Application Hosting

## Challenge

Recovery and Transparency Board needed a platform for their website that was scalable, secure, could be quickly deployed, and saved tax payer money

## Solution

RATB chose a FISMA-compliant cloud computing solution based on Amazon Web Services

- deployed applications:
  - Microsoft Sharepoint for web Content Management
  - Business Objection SAP for BI

## Benefit

- Avoided Capital expense, and added capacity to scale up and down based on demand
- Saved \$750k per year in first year and additional dollars from existing solution



*"By migrating to the public cloud, the Recovery Board is in position to leverage many advantages including the ability keep the site up as millions of Americans help report potential fraud, waste, and abuse. The Board expects savings of about \$750,000 during its current budget cycle and significantly more savings in the long-term."*

- Vivek Kundra, CIO, United States

# Website Hosting / Application Hosting

## Challenge

The US Treasury needed to develop a new Treasury.gov website that could provide over 100 organizations within the Department the ability to manage and update their content. At the same time, they needed to roll out new Web 2.0 features to better engage with their constituents.

## Solution

Treasury chose a cloud computing solution based on Amazon Web Services to support over 11 new websites from Treasury

- Deployed capabilities:
  - Microsoft Sharepoint for web Content Management
  - Multi-faceted Search
  - Integration with Social Networking tools

## Benefit

- Avoided Capital expense, and added capacity to scale up and down based on demand
- Time to deployment



*"Treasury's decision to move its flagship site to a public cloud infrastructure reflects the Administration's commitment to closing the IT gap between the public and private sectors by leveraging the power of technology. Use of cloud computing increases cost effectiveness, improves efficiency and provides greater flexibility, as the private industry sector has proven. This is exactly the kind of game-changing technology required to do more with less."*

- Vivek Kundra, CIO, United States

# Mission Data Processing

## Challenge

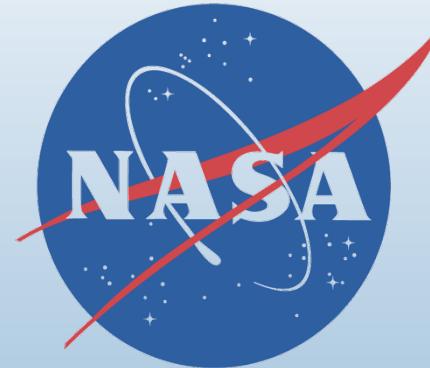
Because of the latency of data transmission from and to Mars, during a 2 hour window, it took mission planners 90 minutes to process telemetry data from the Mars Rover, 20 mins to decide where to move the Rover to, and 10 mins to upload the data.

## Solution

NASA-JPL, loading their custom software application on EC2, was able to horizontally scale the number of virtual machines supporting the data processing.

## Benefit

- Reduced data processing time from 90 minutes to 15 minutes using parallel processing
- Increased mission planning time, resulting in high quality scientific observations



# NASA – MISSION DATA PROCESSING

## Daily Mars Rover Data Processing Window

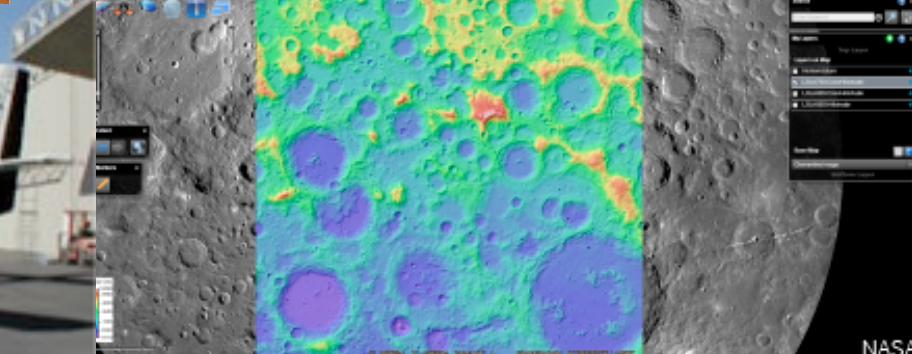
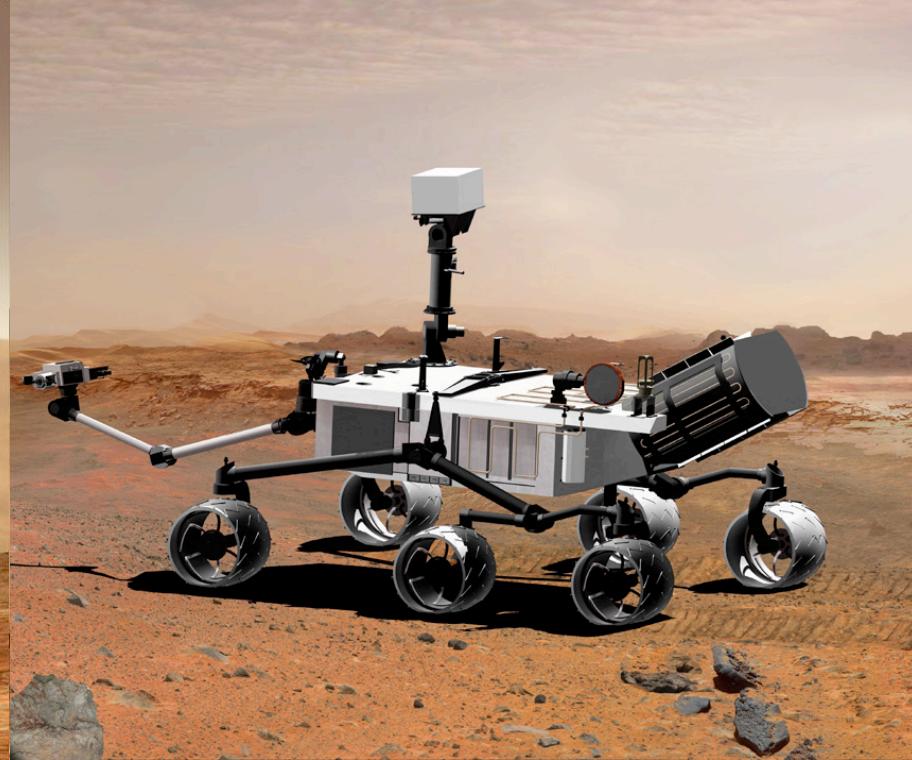
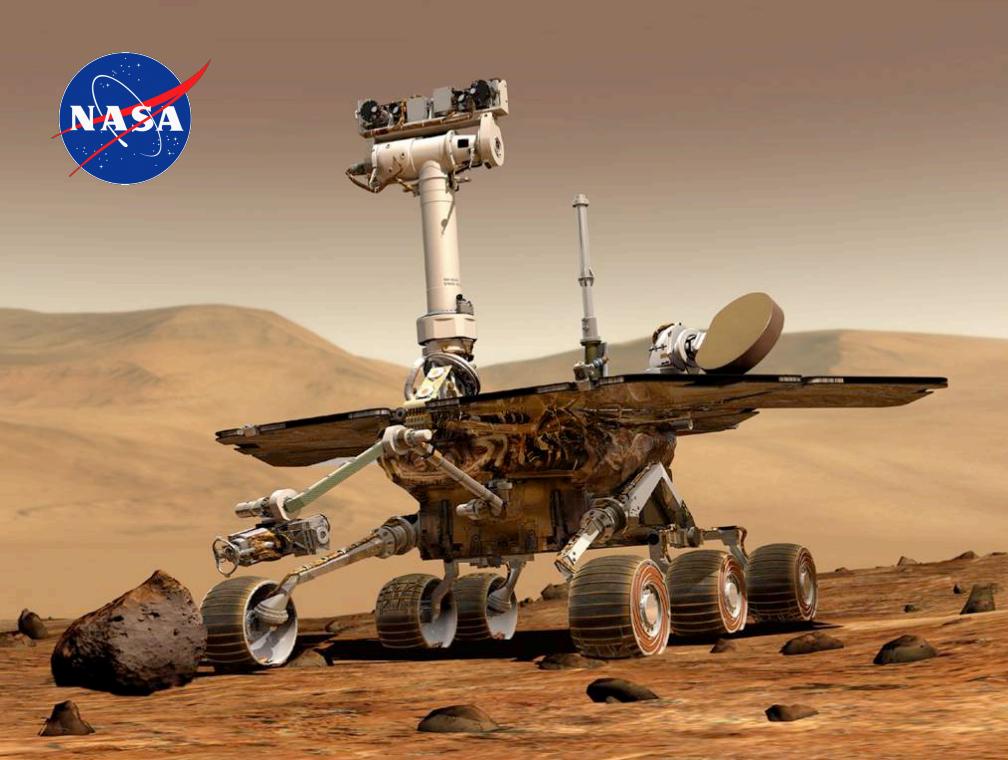
Pre-cloud:

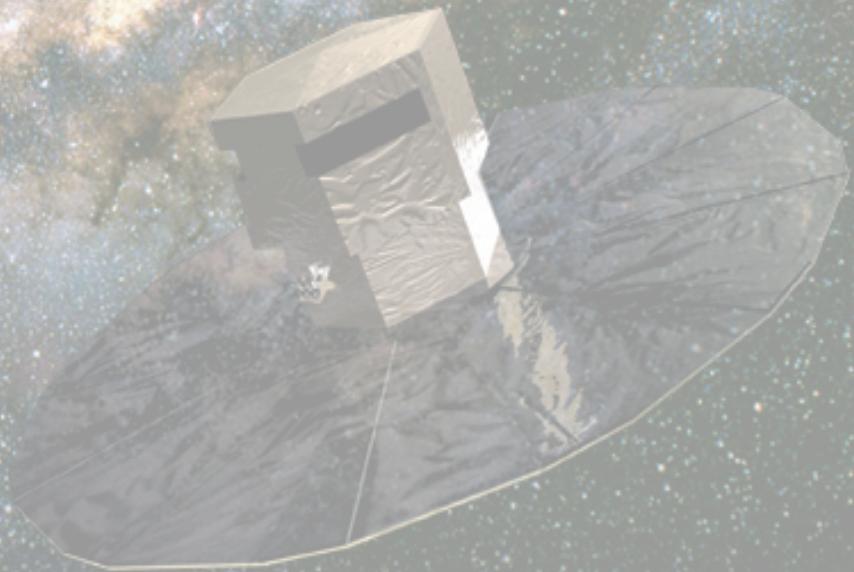
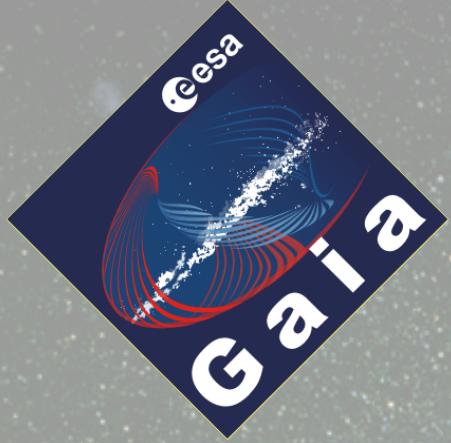


Cloud:



Increase available mission planning time from  
15 minutes to 105 minutes!





esa



# AWS Platform

## Your Applications

### Management & Administration

#### Identity & Access

AWS IAM  
Identity Federation  
Consolidated Billing

#### Web Interface

Management Console

#### Monitoring

Amazon CloudWatch

#### Deployment & Automation

AWS Elastic Beanstalk  
AWS CloudFormation  
Simple Workflow Service

### Application Platform Services

#### Content Distribution

Amazon CloudFront

#### Messaging

Amazon SNS  
Amazon SQS  
Amazon SES

#### Parallel Processing

Elastic MapReduce

#### Libraries & SDKs

Java, PHP, Python,  
Ruby, .NET

### Foundation Services

#### Compute

Amazon EC2  
Auto Scale

#### Storage

Amazon S3  
Amazon EBS  
Amazon StorageGateway

#### Database

Amazon RDS  
Amazon SimpleDB  
Amazon ElastiCache  
Amazon DynamoDB

#### Networking

Amazon VPC  
Elastic Load Balancing  
Amazon Route 53  
AWS Direct Connect

### AWS Global Infrastructure

#### Availability Zones

#### Regions

#### Edge Locations

# AWS Global Infrastructure

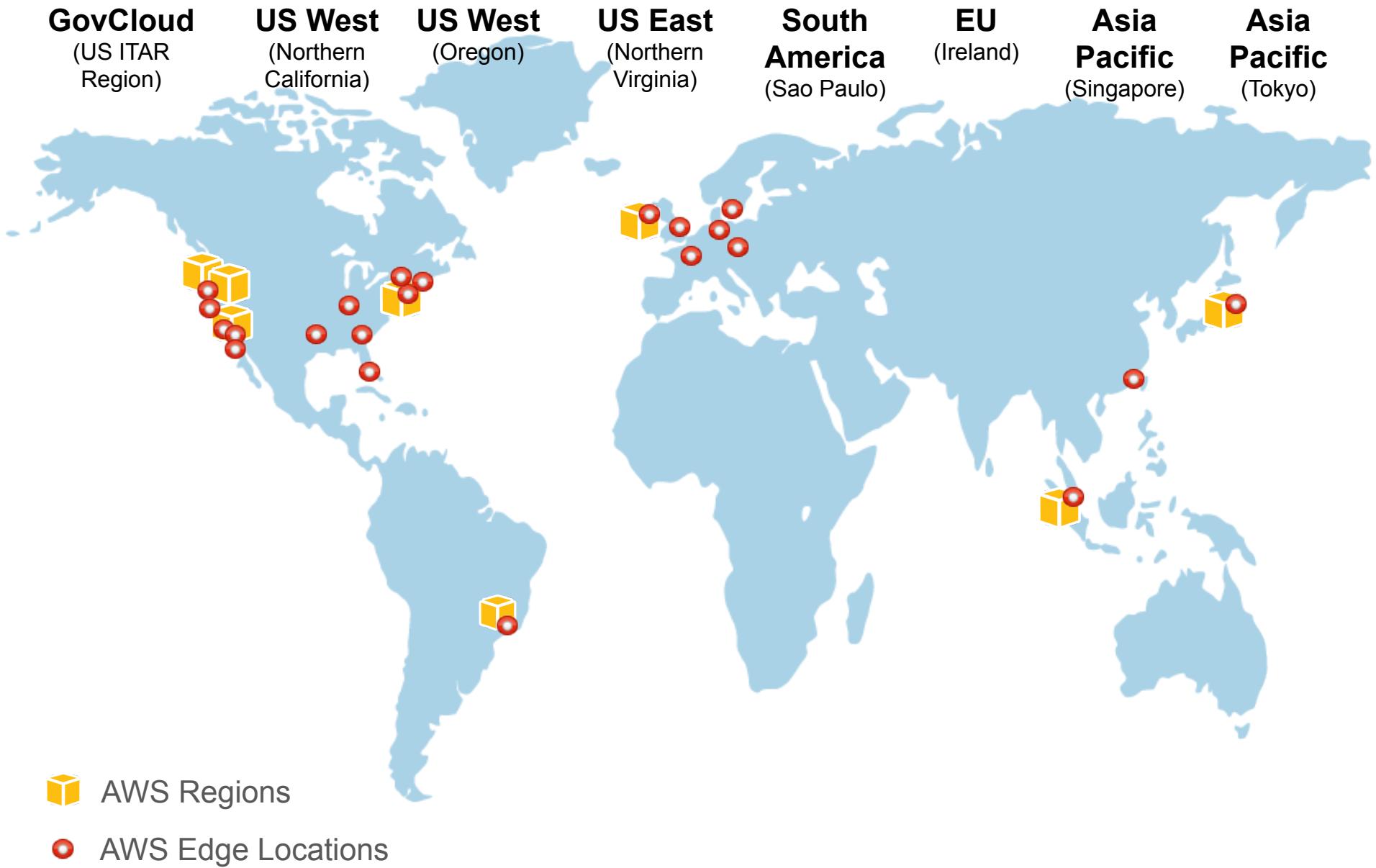
AWS Global Infrastructure

Availability Zones

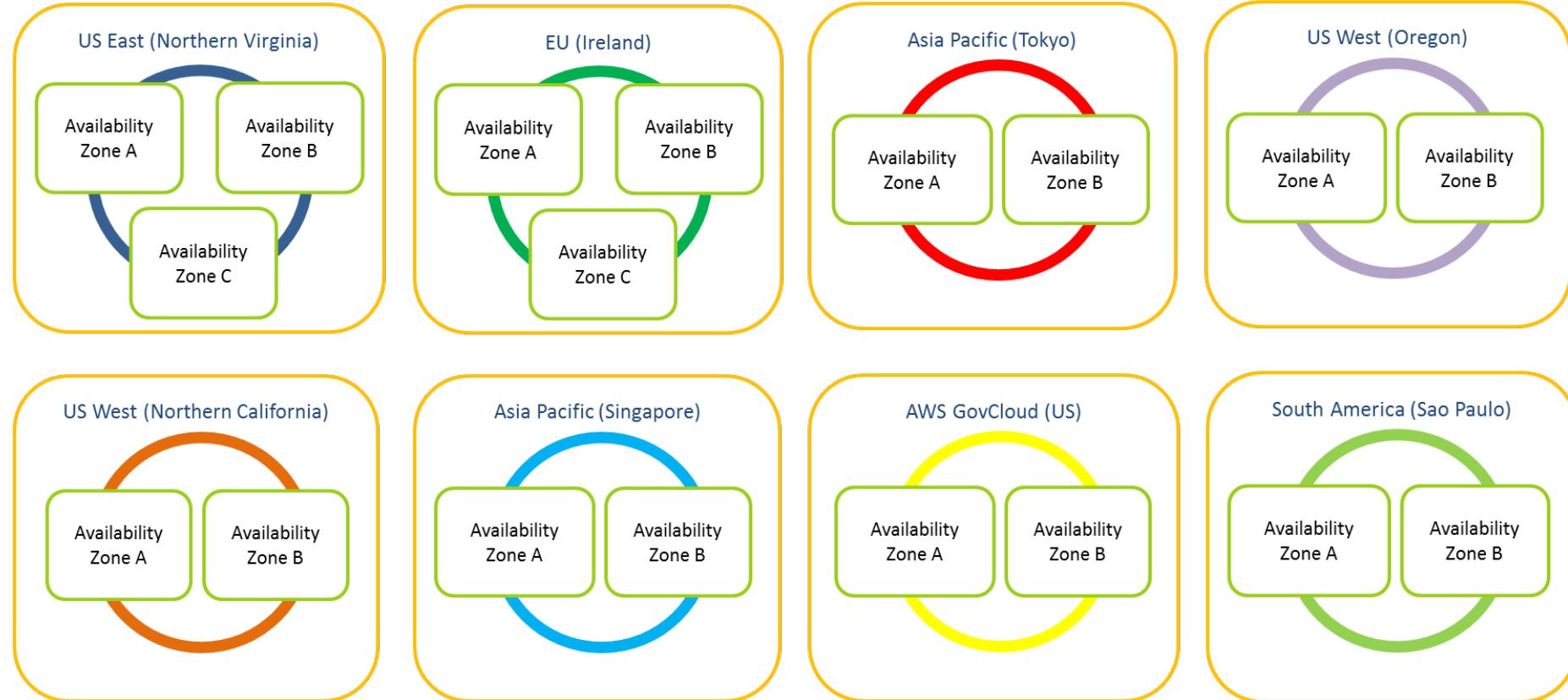
Regions

Edge Locations

# Global Infrastructure for Global Enterprises



# AWS Regions and Availability Zones



Customer Decides Where Applications and Data Reside

# Built to Enterprise & Gov't Security Requirements

## Physical

- Datacenters in nondescript facilities
- Physical access strictly controlled
- Must pass two-factor authentication at least twice for floor access
- Physical access logged and audited

## Hardware, Software & Network

- Systematic change management
- Phased updates deployment
- Safe storage decommission
- Automated monitoring and self-audit
- Advanced network protection systems

## Certifications and Accreditations

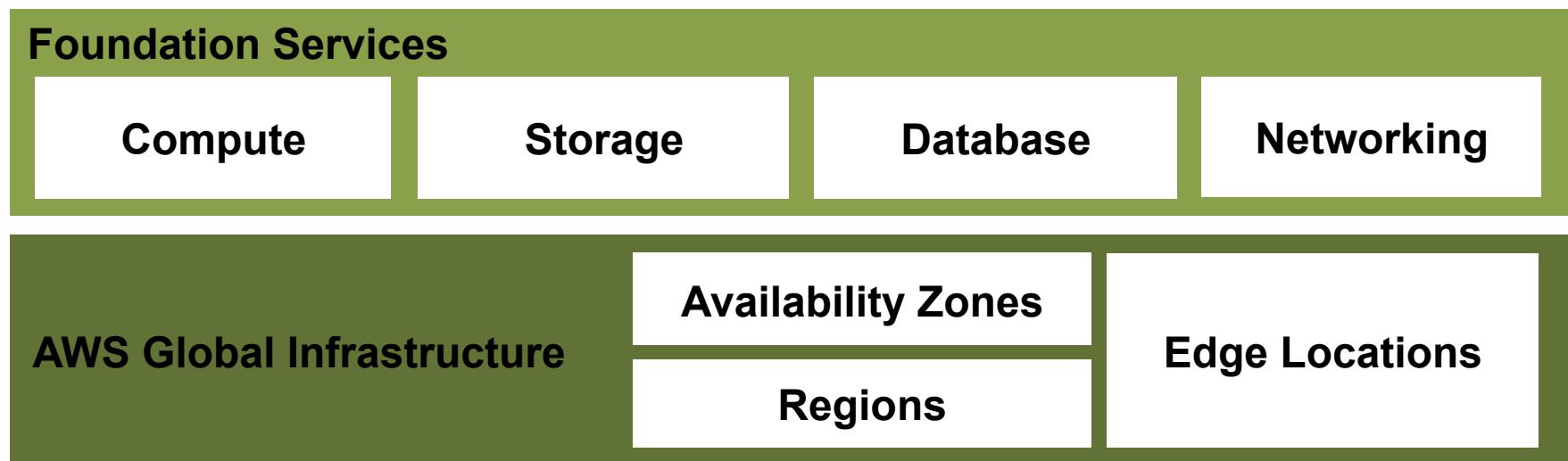
- ISO 27001
- SSAE 16 / ISAE 3402 / SOC1 (formerly U.S. standard SAS-70 Type II)
- FISMA Moderate Controls; ITAR region
- HIPAA applications certified on AWS
- Payment Card Industry (PCI) Data Security Standard (DSS) Level 1
- DIACAP Controls

## Security & Compliance Resources

- Security & Compliance Center: <http://aws.amazon.com/security>
- Security Overview & Best Practices
- AWS Risk & Compliance Whitepaper
- Creating HIPAA Compliant Applications



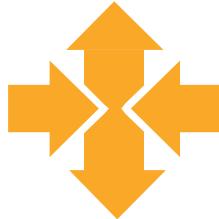
# AWS Foundation Services



# Compute



Virtual Servers

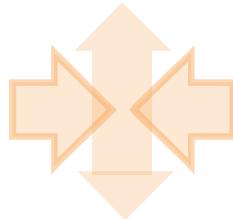


Auto Scaling Servers

# Compute

## Amazon Elastic Compute Cloud (Amazon EC2)

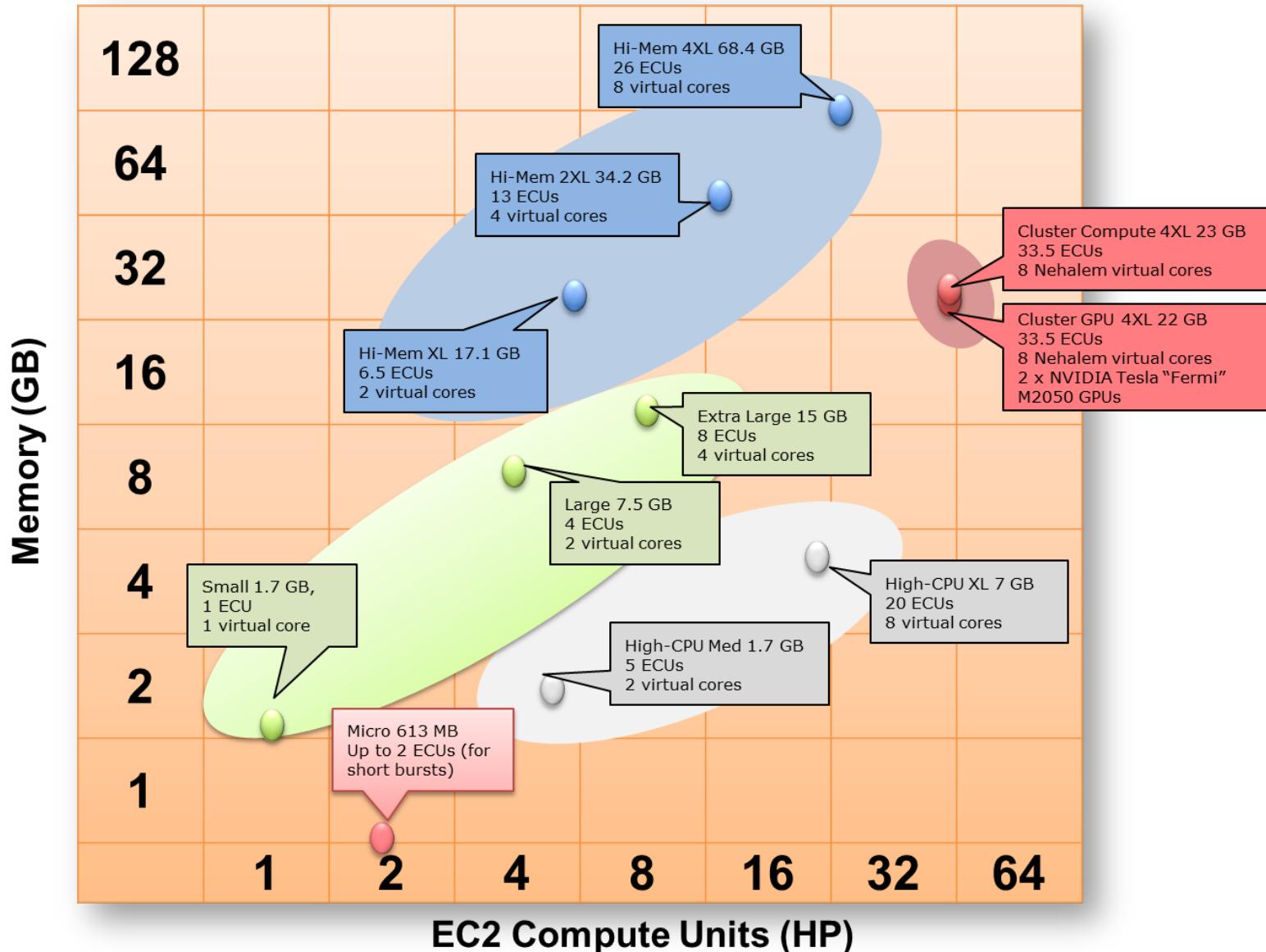
EC2



### EC2 Instances = Virtual Servers

- Resizable compute capacity in 13 instance types
- Reduces the time required to obtain and boot new server instances to minutes or seconds
- Scale capacity as your computing requirements change
- Pay only for capacity that you actually use
- Choose Linux or Windows
- Deploy across Regions and Availability Zones for reliability
- Flexible networking (NAT/classic, VPC, Elastic IPs)
- Support for virtual network interfaces that can be attached to EC2 instances in your VPC

# Compute: Amazon EC2 Instances



One EC2 Compute Unit (ECU) provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor

2 \* Intel Xeon ES-2670  
“Sandy Bridge” Architecture  
16 cores w/ HT  
60.5 GB RAM  
3.4 TB disk  
HVM  
**cc2.8xlarge**

**Each day**, AWS adds the  
equivalent server capacity to power Amazon  
when it was a global, \$2.76B enterprise  
(circa 2000)

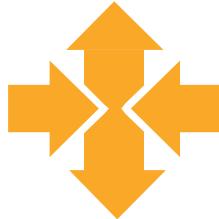


# How to Launch an EC2 Server Instance

- **Select the Instance Size** – Small, Large, Cluster Compute, etc.
- **Select the Operating System Type:**
  - Windows 2003, 2008, Red Hat, CentOS, Suse, etc.
- **Define the firewall rule settings (Security Groups)**
- **Create or use existing Security Key-Pair (RSA-256)** - to restrict access to the server
- **LAUNCH! (Takes less than 1 minute for everything)**

# Compute

EC2



## Auto Scaling Servers

- Client Defined Business Rules
- Scale your Amazon EC2 capacity automatically once you define the conditions (may be 1000's of servers)
- Can scale up just a little...doesn't need to be massive number of servers (may be simply 2 servers)
- Well suited for applications that experience variability in usage
- Set minimum and maximum scaling policies
- Alternate Use is for Fault Tolerance

# Storage



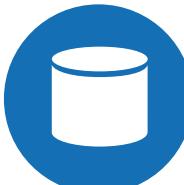
Very Large Data Store



Think of as “Virtual Disks”



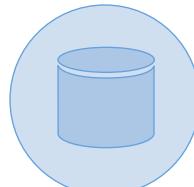
Bulk Move of Data into & out-of AWS



On-Premise Backup/COOP to AWS

# Storage

## Amazon Simple Storage Service (Amazon S3)



### Web-scale Internet Storage

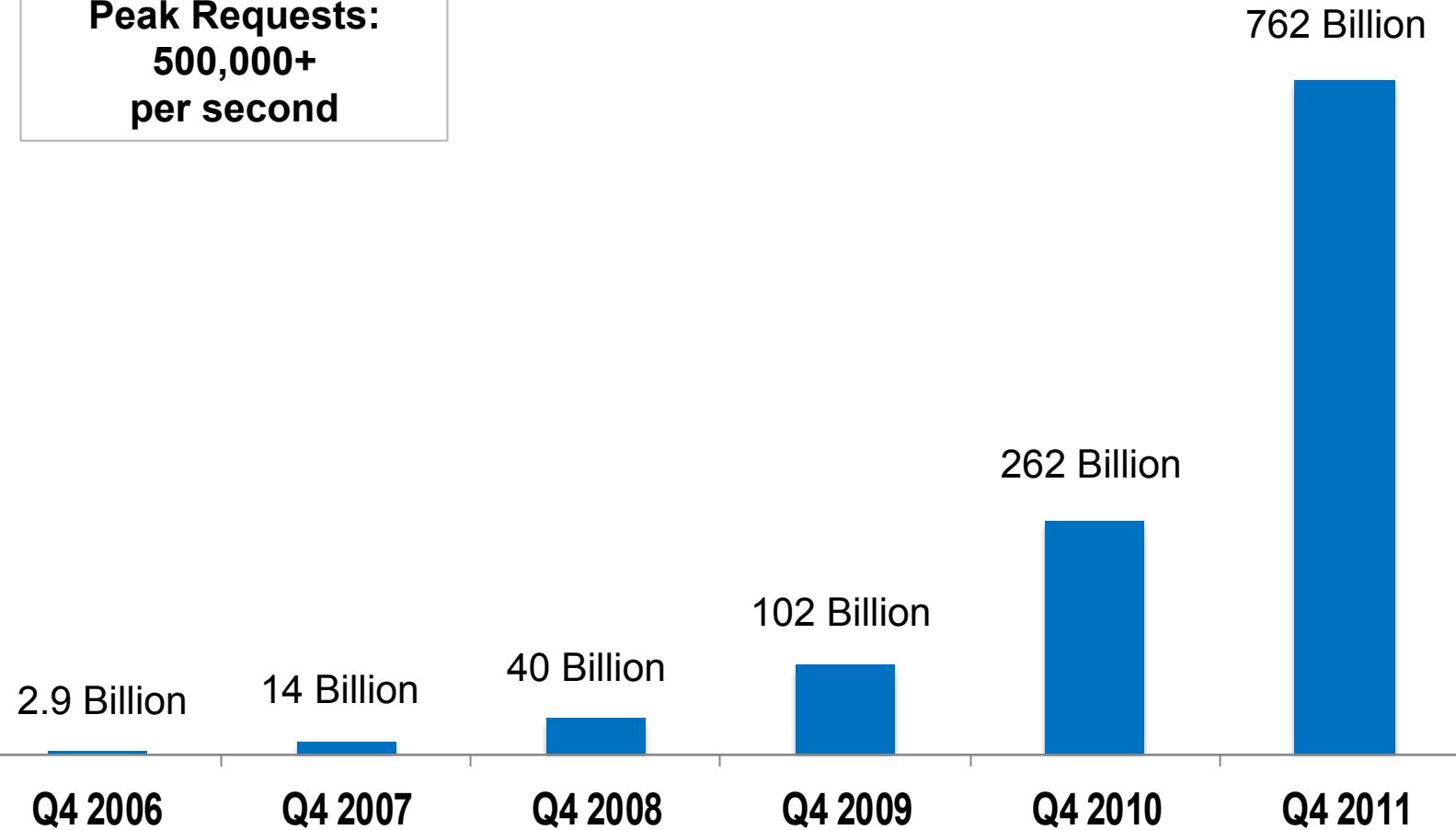
- A “Bucket” is equivalent to a “folder”
- Able to store unlimited number of Objects in a Bucket
- Objects from 1B-5 TB; no bucket size limit
- Highly available storage for the Internet (object store)
- HTTP/S endpoint to store and retrieve any amount of data, at any time, from anywhere on the web
- Highly scalable, reliable, fast, and inexpensive
  - Over 762 Billion objects stored
  - Peak requests 500,000+ per second

### Ideal Use Cases:

- Static web content – often used with CloudFront CDN
- Source and output storage for large-scale “Big Data” analytics
- Backup, archival, and DR storage that is always “live”

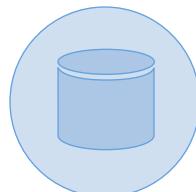
# The Cloud Scales: Amazon S3 Growth

**Peak Requests:**  
500,000+  
per second



**Total Number of Objects Stored in Amazon S3**

# Storage

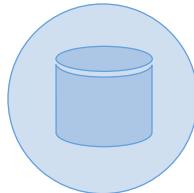
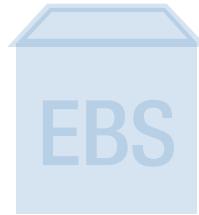


## Amazon Elastic Block Store (EBS)

### EBS Volumes = Virtual Disks

- Use for persistent storage
- Can use to create RAID configuration for a server
- Off-instance block storage that persists independently
- Storage volumes for use with Amazon EC2 instances
  - create, attach, backup, restore and delete
- Can be attached to a running Amazon EC2 instance and exposed as a block device for raw or formatted (filesystem) access
  - Volumes behave like unformatted block devices for Linux or Windows instances
- Ideas use cases:
  - OS Boot device / root file system; secondary volumes/filesystems
  - Typical basis for database storage
  - Raw block devices for RAID, some databases

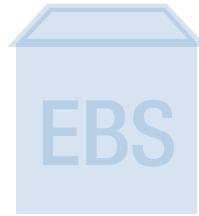
# Storage



## AWS Import/Export

- Accelerates moving large amounts of data into and out of S3 or EBS
- Transfers your data directly onto and off of USB or SATA storage devices shipped to AWS with manifest file
- Final copy uses high-speed datacenter network

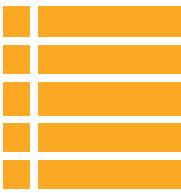
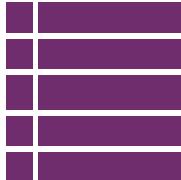
# Storage



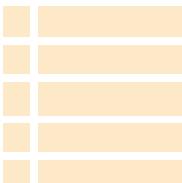
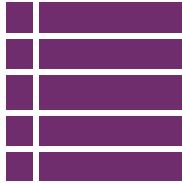
## AWS Storage Gateway

- Storage gateway service connects an on-premise software appliance with cloud-based storage
- On-premises software appliance solution to store data on Amazon S3's storage infrastructure
- Exposes standard iSCSI interface to on-premises applications, while maintaining low-latency data access
- Data in Amazon S3 stored as Amazon EBS snapshots for local & EC2-based recovery
- Use Cases
  - Backup/Restore on-premise data
  - Set up a test/dev environment with production data
  - Migrating applications to the cloud
  - On-premise DR/COOP to AWS

# Database



# Database



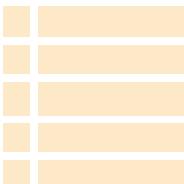
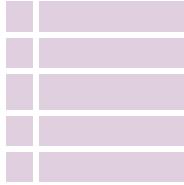
## Amazon SimpleDB

- Core database functions of data indexing and querying of text data
- No schema, automatic indexing
- Eliminates the administrative burden of data modeling, index maintenance, and performance tuning
- Real-time lookup and simple querying of structured data
- Use cases:
  - Metadata storage -- often used in conjunction with S3
  - Structured, fine-grained data needing query
  - Data needing flexible schema

- RDBMS comparison:

SimpleDB		RDBMS
Domains	↔	Table
Item	↔	Row
Attributes	↔	Column *
Values	↔	Values

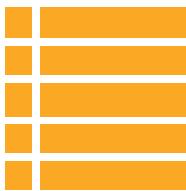
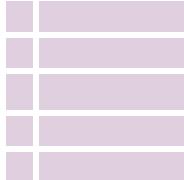
# Database



## Amazon Relational Database Service (RDS)

- Fully-managed, tuned MySQL or Oracle 11g database
- Cost-efficient and resizable capacity
- Manages time-consuming database admin tasks
- Code, applications, and tools you already use today work seamlessly
- Automatically patches the database software and backs up your database
- Flexible Licensing: BYOL or License Include

# Database

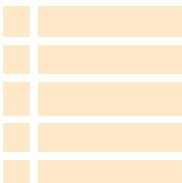
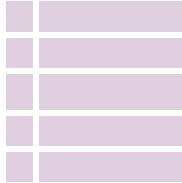


## Amazon ElastiCache

- Fully-managed, distributed, in-memory cache
- Memcached compliant cache cluster on-demand
- Manages patching, cache node failure detection and recovery
- Simple APIs calls to grow and shrink the cache cluster
- Seamlessly caches in front of SimpleDB or RDS instances
- Integrated with CloudWatch and SNS for monitoring and alerts



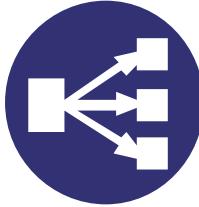
# Database



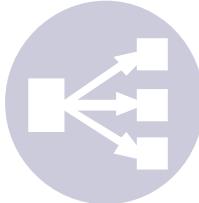
## Amazon DynamoDB

- Fully managed NoSQL database.
- Eliminates the administrative burden of data modeling, index maintenance, and performance tuning.
- Durability and high-availability - stores data on Solid State Drives (SSDs) and replicates it synchronously across multiple AWS Availability Zones in an AWS Region.
- Scalability - With AWS Console, you can grow your DynamoDB table from 1,000 to 100,000 writes per sec.
- See video: <http://www.youtube.com/watch?v=oz-7wJJ9HZ0>

# Networking



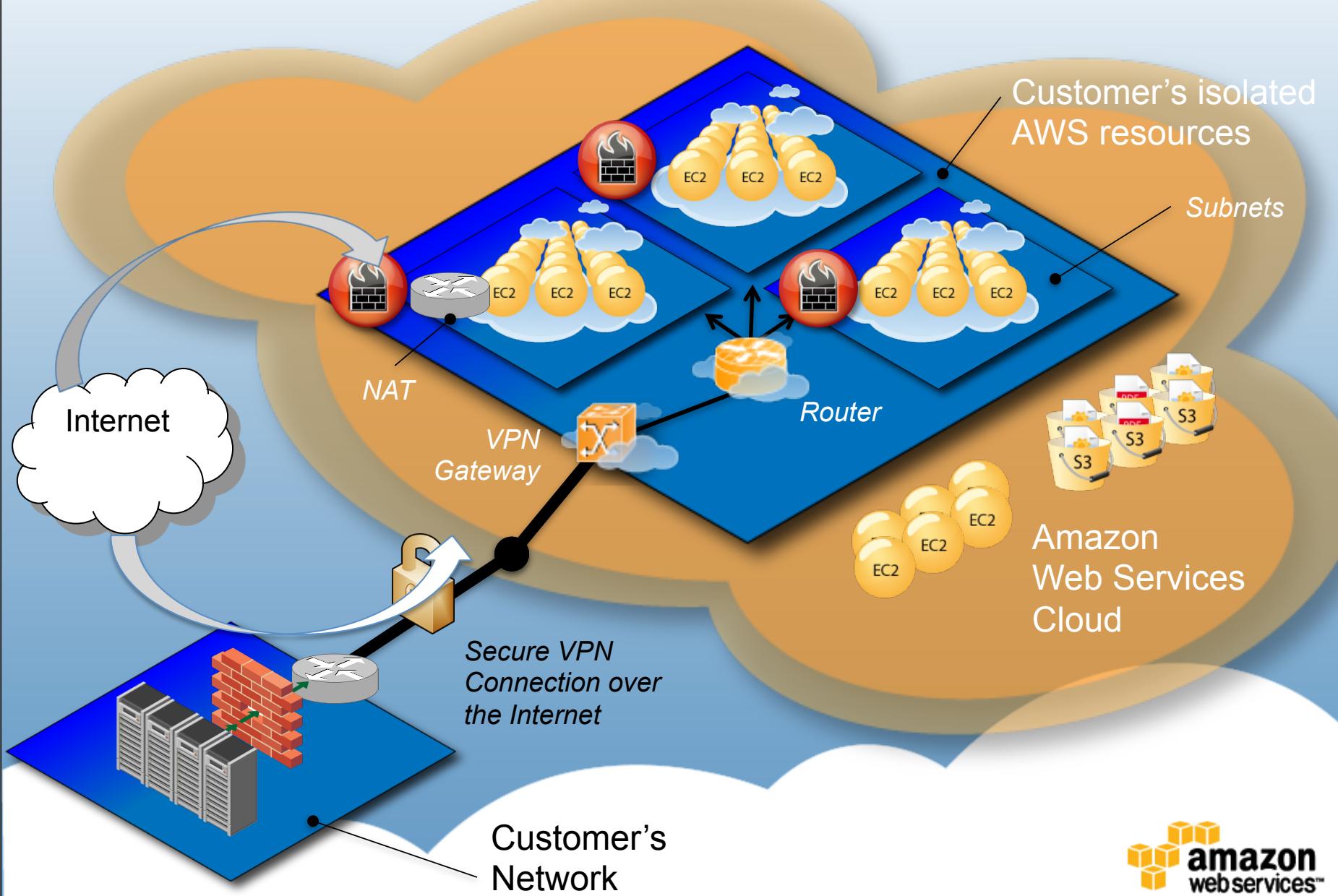
# Networking



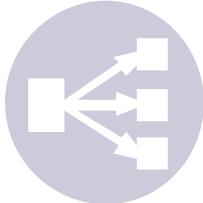
## Amazon Virtual Private Cloud (Amazon VPC)

- Secure and seamless bridge between a company's existing private network and the AWS cloud
- Connect existing infrastructure to a set of isolated AWS compute resources via a Virtual Private Network (VPN) connection
- Bring your own address space and extend existing management capabilities
  - Critical capability: further drill-down on VPC later

# Amazon VPC Architecture



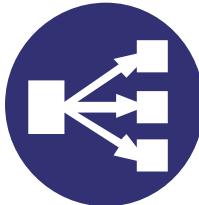
# Networking



## Amazon Route 53

- Route end users to Internet applications
- Answers DNS queries with low latency by using a global network of DNS servers
- Deep integration with other AWS services (ELB, EC2 NAT/EIP, etc.)

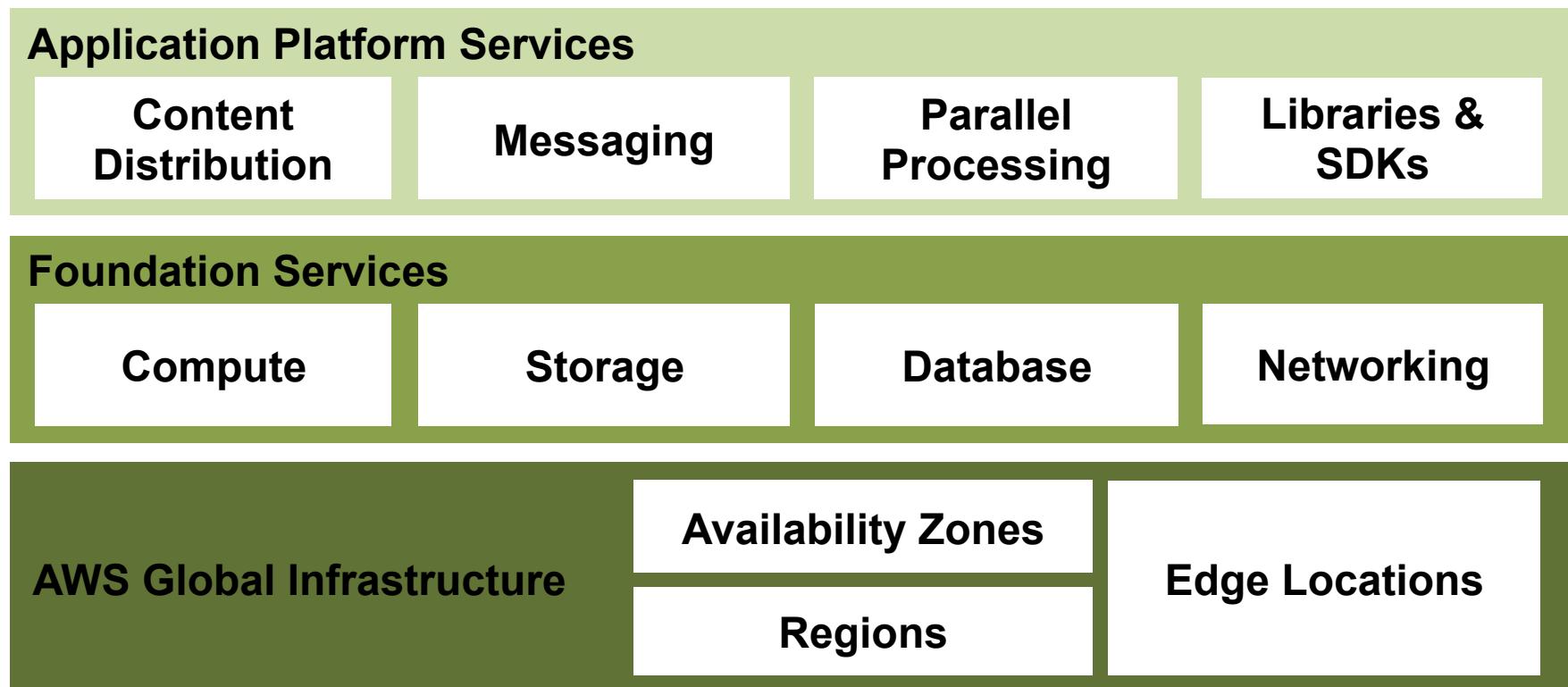
# Networking



## Amazon Elastic Load Balancing

- Supports the routing and load balancing of HTTP, HTTPS and generic TCP traffic to EC2 instances
- Supports health checks to ensure detect and remove failing instances
- Dynamically grows and shrinks required resources based on traffic
- Seamlessly integrates with Auto-scaling to add and remove instances based on scaling activities
- Single CNAME provides stable entry point for DNS configuration

# AWS Application Platform Services



# Content Distribution



## Amazon CloudFront

- Web service for content delivery
- Distribute content to end users with low latency, high data transfer speeds, and no commitments
- Delivers your content using a global network of 26 edge locations
- Supports download, streaming and live streaming
  - Key features: RTMP Streaming, HTTPS Delivery, Private Content for HTTP & Streaming, Programmatic Invalidations, Detailed Logs for HTTP & Streaming, Default Root Object
- Use Cases: Video and Rich Media, Online Gaming, Interactive Agencies, Software Downloads, Static Websites
  - Static web content that must be delivered to global user base at
  - Highest bandwidth / Lowest latency / Lowest cost

# Messaging



# Messaging

## Amazon Simple Notification Service (SNS)



- Set up, operate, and send notifications
- Publish messages from an application and immediately deliver them to subscribers or other applications
- Multiple protocol support: subscribe to e-mail, HTTP, SMS or SQS



# Messaging



## Amazon Simple Email Service (Amazon SES)

- Bulk and transactional email-sending service
- Eliminates the hassle of email server management, network configuration, and meeting rigorous Internet Service Provider (ISP) standards
- Provides a built-in feedback loop, which includes notifications of bounce backs, failed and successful delivery attempts, and spam complaints

# Messaging

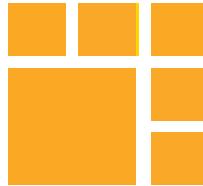


## Amazon Simple Queue Service (SQS)

- Reliable, highly-scalable hosted message queuing service
- Move data between distributed components of their applications
- SQS provides Durable but Temporary storage
  - Messages must be text-only, up to 64KB
  - Messages stay in queue for up to 14 days



# Parallel Processing



## Amazon Elastic MapReduce (EMR)

- Managed Hadoop 0.20.205 infrastructure
- Reduces complexity of Hadoop management
  - Handles node provisioning, customization, and shutdown
  - Tunes Hadoop to your hardware and network
  - Provides tools to debug and monitor your Hadoop clusters
- Provides tight integration with AWS services
  - Optimized for Amazon Simple Storage Service (S3)
  - EC2 integration with automatic re-provisioning on node failure
  - Cluster monitoring/alarming through CloudWatch
- Leverages significant operational experience
  - Monitor thousands of clusters per day
  - Use cases span from University students to Fortune 50

# Libraries & SDKs

Your choice of programming language (Java, PHP, Python, Ruby, .NET) and mobile platform (Android, iOS)

The Developer Centers contains sample code, documentation, tools, and additional resources to help you build applications on Amazon Web Services.

- <http://aws.amazon.com/java/>
- <http://aws.amazon.com/mobile/>
- <http://aws.amazon.com/php/>
- <http://aws.amazon.com/python/>
- <http://aws.amazon.com/ruby/>
- <http://aws.amazon.com/net/>



# AWS Management & Administration

## Management & Administration

Administration Console

Identity & Access

Deployment

Monitoring

## Application Platform Services

Content Distribution

Messaging

Parallel Processing

Libraries & SDKs

## Foundation Services

Compute

Storage

Database

Networking

## AWS Global Infrastructure

Availability Zones

Regions

Edge Locations

# AWS Management Console

One-stop shop to manage your AWS services

The screenshot shows the Amazon EC2 Console Dashboard. At the top, there's a navigation bar with links to various AWS services like S3, VPC, CloudWatch, and Route 53. Below the navigation bar is a "Getting Started" section with a yellow box containing text about launching an instance and a "Launch Instance" button. To the right is a "My Resources" section showing statistics for Running Instances, EBS Volumes, Key Pairs, Load Balancers, Elastic IPs, EBS Snapshots, Security Groups, and Placement Groups. The "Service Health" section displays the current status of Amazon EC2 and its availability zones across four regions: us-east-1a, us-east-1b, us-east-1c, and us-east-1d, all of which are operating normally. A "Scheduled Events" section indicates no events are scheduled. On the left, a sidebar titled "Navigation" lists categories such as EC2 Dashboard, Scheduled Events, Instances, Images, Elastic Block Store, Network & Security, and more.

# Identity & Access Management (IAM)

IAM enables customers to create and manage users in AWS's identity system

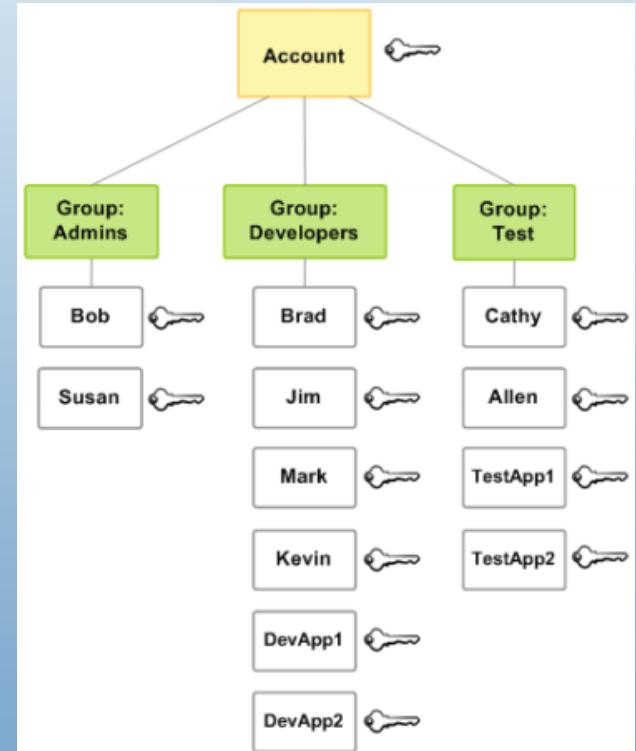
- Identity Federation with local directory is an option for enterprises

Very familiar security model

- Users, groups, permissions

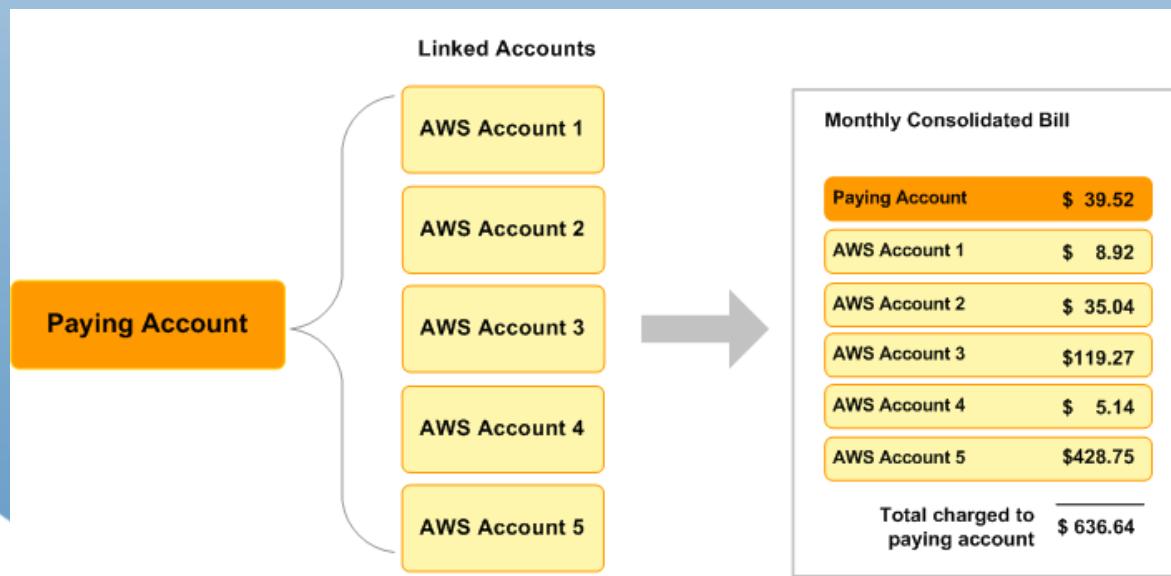
Allows customers to

- Create users
- Assign individual passwords, access keys, multi-factor authentication devices
- Grant fine-grained permissions
- Optionally grant them access to the AWS Console
- Organize users in groups

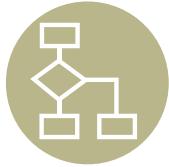
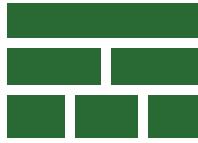
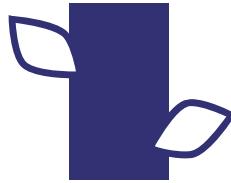


# Consolidated Billing with IAM

- Allows you to get one bill for multiple accounts
- You can easily track each account's costs and download the cost data in CSV format
- You may be able to reduce costs by combining usage from all the accounts to qualify for volume pricing discounts



# Deployment and Management

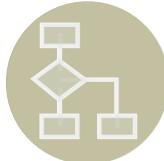
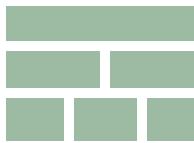


# Deployment and Management

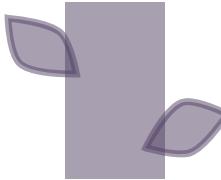


## AWS Elastic Beanstalk (beta)

- Simply upload your application
- Automatically handles the deployment details of capacity provisioning, load balancing, auto-scaling, and application health monitoring
- Retain full control over the AWS resources powering your application

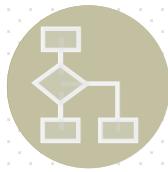


# Deployment and Management

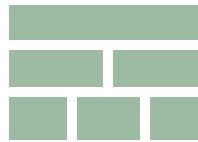
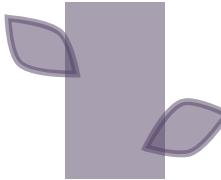


## AWS CloudFormation

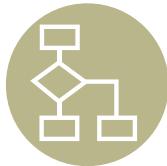
- Create templates of stack of resources
- Deploy stack from template with runtime parameters
- CloudFormer supports generating templates from running environments



# Deployment and Management



## Simple Workflow Service



- Maintains distributed application state
- Tracks and retains workflow executions
- Ensures consistency of execution history
- Holds and dispatches tasks
- Provides control over task distribution



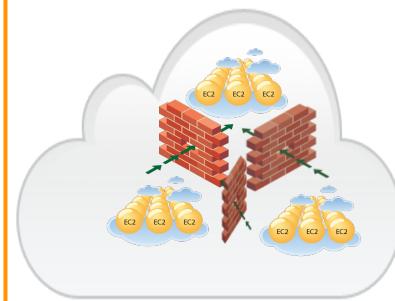
# Monitoring



## Amazon CloudWatch

- Visibility into resource utilization, operational performance, and overall demand patterns
- Metrics such as CPU utilization, disk reads and writes, and network traffic
- Accessible via the AWS Management Console, web service APIs or Command Line Tools
- Add custom metrics of your own
- Alarms (which tie into auto-scaling, SNS, SQS, etc.)

# AWS Networking & Security



## Network Security Groups

Flexible, named network access policy objects attached to VMs, dynamically updateable, enforced by the hypervisor



## Amazon Virtual Private Cloud (VPC)

Create logically isolated AWS subnets in the AWS cloud and connect them with your corporate datacenter via a secure VPN connection

## Single Tenant Physical Nodes



## Dedicated Instances

Run your virtualized operating systems and apps in a "single tenant per physical node" model within the AWS infrastructure



## AWS Direct Connect

Dedicated, private, fast Layer 2 connectivity between your datacenter and the AWS cloud; no traffic on the Internet (available in some regions)

# Thank You!

Jamie Kinney

[jkinney@amazon.com](mailto:jkinney@amazon.com)

[http://jkinney.s3.amazonaws.com/  
NASA\\_AWS\\_Overview.pdf](http://jkinney.s3.amazonaws.com/NASA_AWS_Overview.pdf)

