



Introduction to Amazon Cloud & EC2 Overview

Yasser Quraishi, Solutions Architect - yquraish@amazon.com

Corey Harris, Solutions Architect – chj@amazon.com

John Gilbert, Snr. Manager, Solutions Architecture - jgilbaws@amazon.com

Cody Hulsey, Account Manager – codyhuls@amazon.com

Agenda

- Introduction to AWS Cloud
- Global Reach, Regions and Availability Zones
- EC2 Overview
- EC2 Details

What is AWS?

AWS provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers millions of businesses in over 190 countries around the world.

Benefits

- Low Cost
- Elasticity & Agility
- Open & Flexible
- Secure
- Global Reach



What sets AWS apart?



Experience

Building and managing cloud since 2006



Service Breadth & Depth

175+ services to support any cloud workload



Pace of Innovation

History of rapid, customer-driven releases



Global Footprint

22 regions, 69 availability zones, 199+ edge locations



Pricing Philosophy

77 proactive price reductions to date



Ecosystem

Thousands of consulting/system integrator & technology partners

Pricing Philosophy

High volume / low margin businesses are in our core DNA

Trade CapEX for
variable expense

Pay for what
you use

Our economies of
scale provide us
with lower costs

77 price
reductions
since 2006

Pricing model
choice to support
variable and stable
workloads

On-demand
Reserved Instances
Spot

Save more money as
you grow bigger

Tiered pricing
Volume discounts
Custom pricing

Customer obsessed



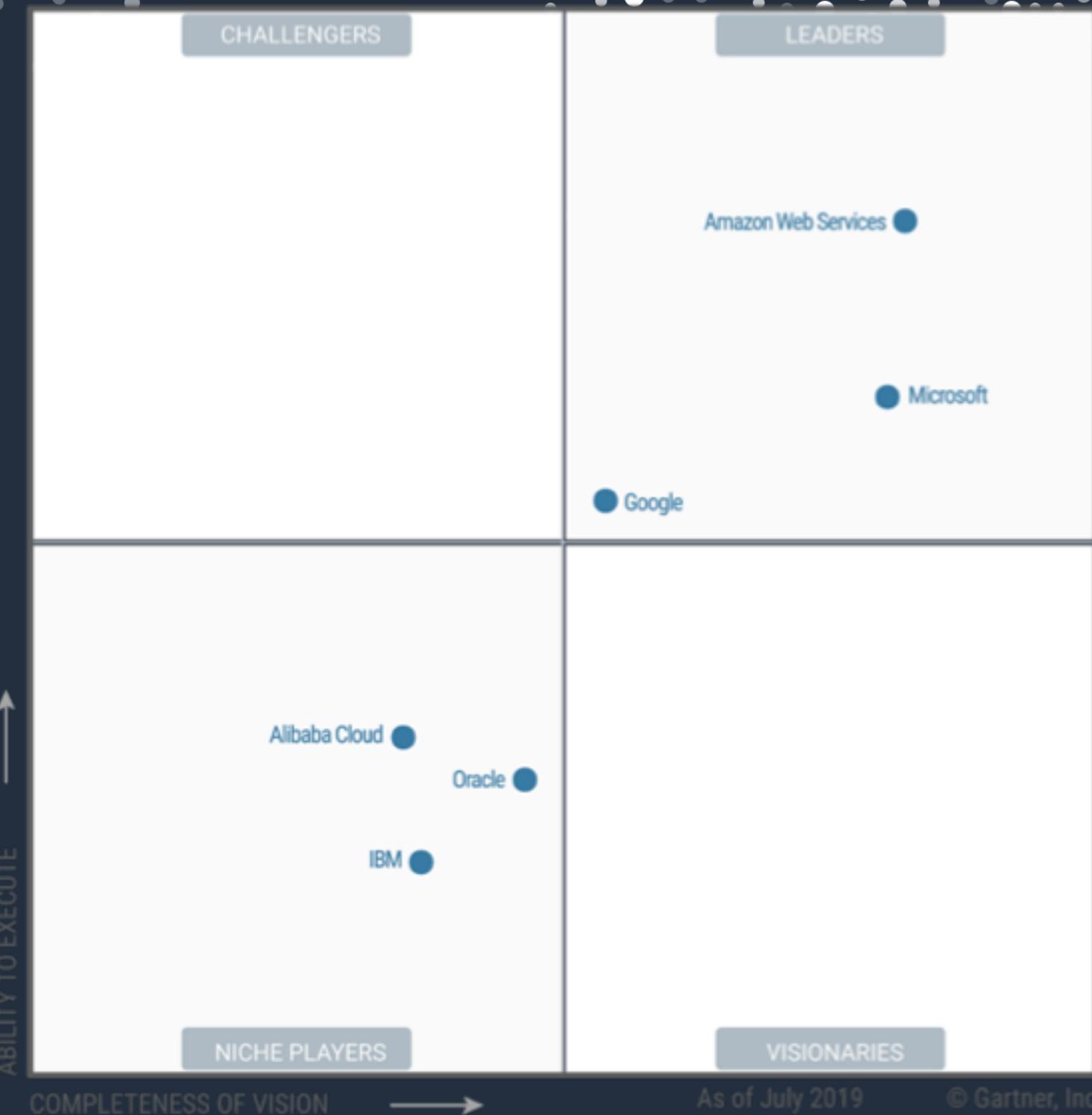
90%
of roadmap originates with customer requests
and are designed to meet specific needs



"Performance, reliability, and responsiveness are fundamental to our customer experience, and T3 instances help us to deliver on that customer promise while also controlling our costs."

—Heroku

Figure 1. Magic Quadrant for Cloud Infrastructure as a Service, Worldwide



AWS Recognized as
a Cloud Leader for the
9th Consecutive Year

Gartner, Magic Quadrant for Cloud Infrastructure as a Service, Worldwide, Raj Bala, Bob Gill, Dennis Smith, David Wright, July 2019. ID G00365830. Gartner does not endorse any product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose. The Gartner logo is a trademark and service mark of Gartner, Inc., and/or its affiliates, and is used herein with permission. All rights reserved.

1

AWS Global Reach

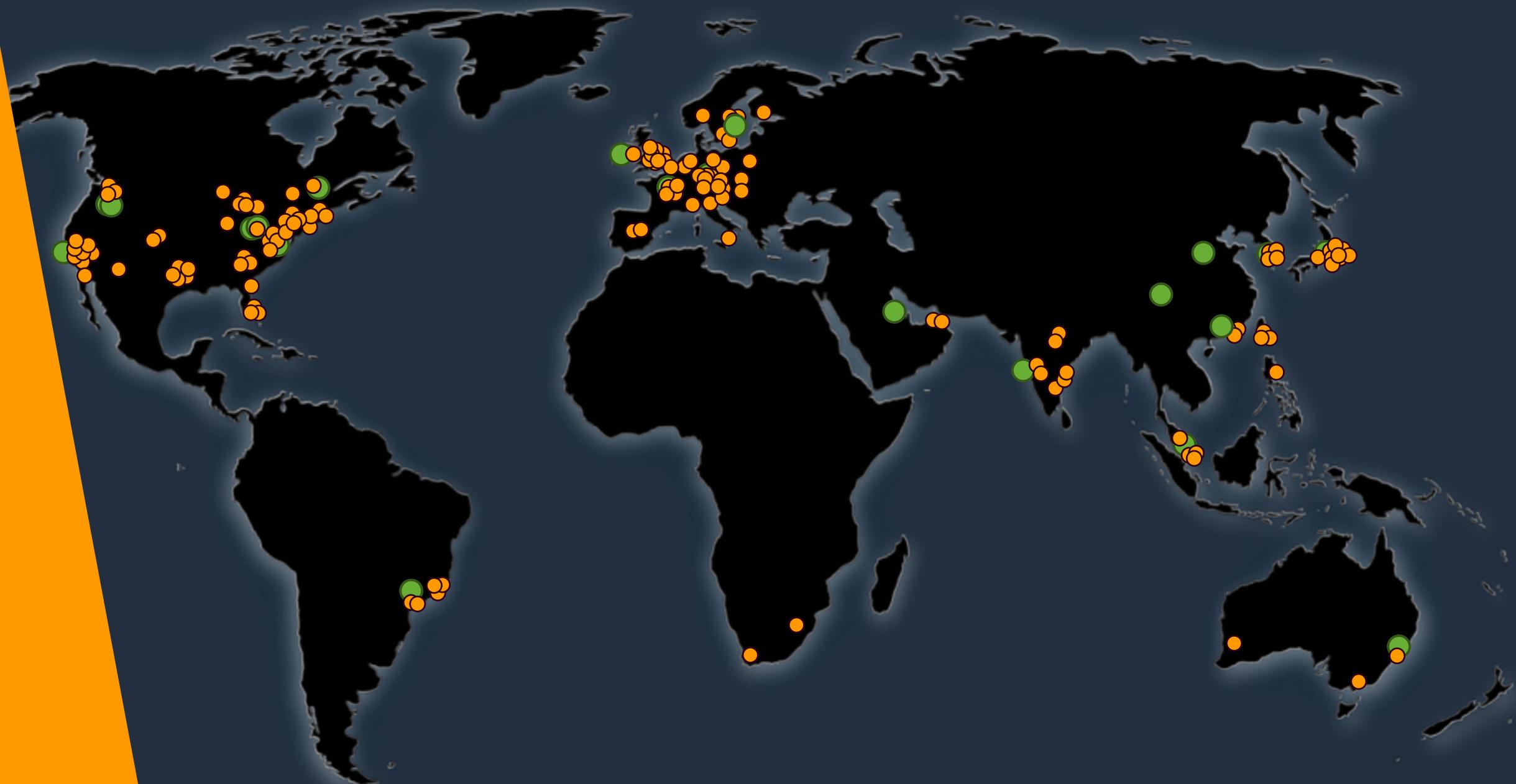
22

Regions



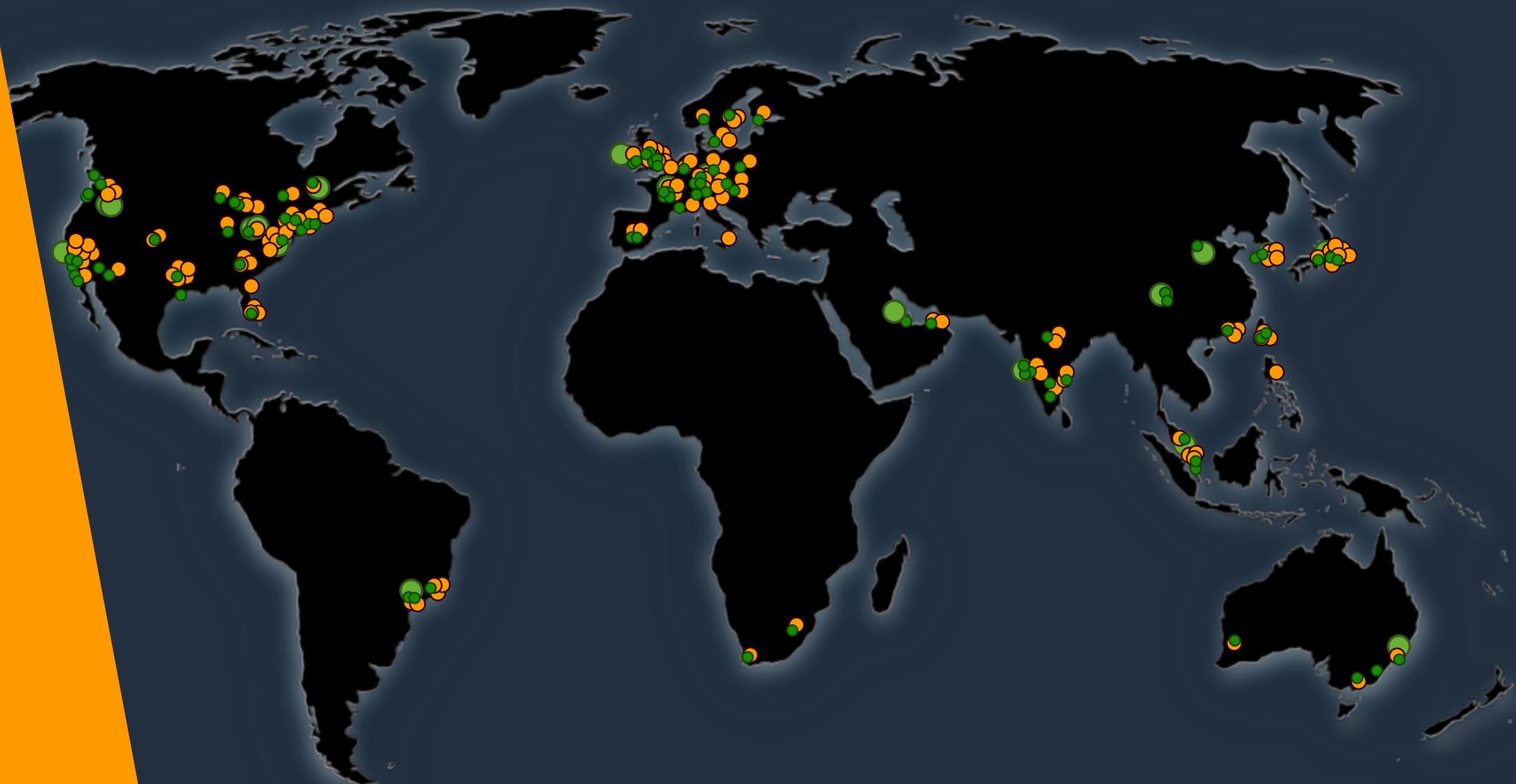
210

Amazon
CloudFront
Points of
Presence



100

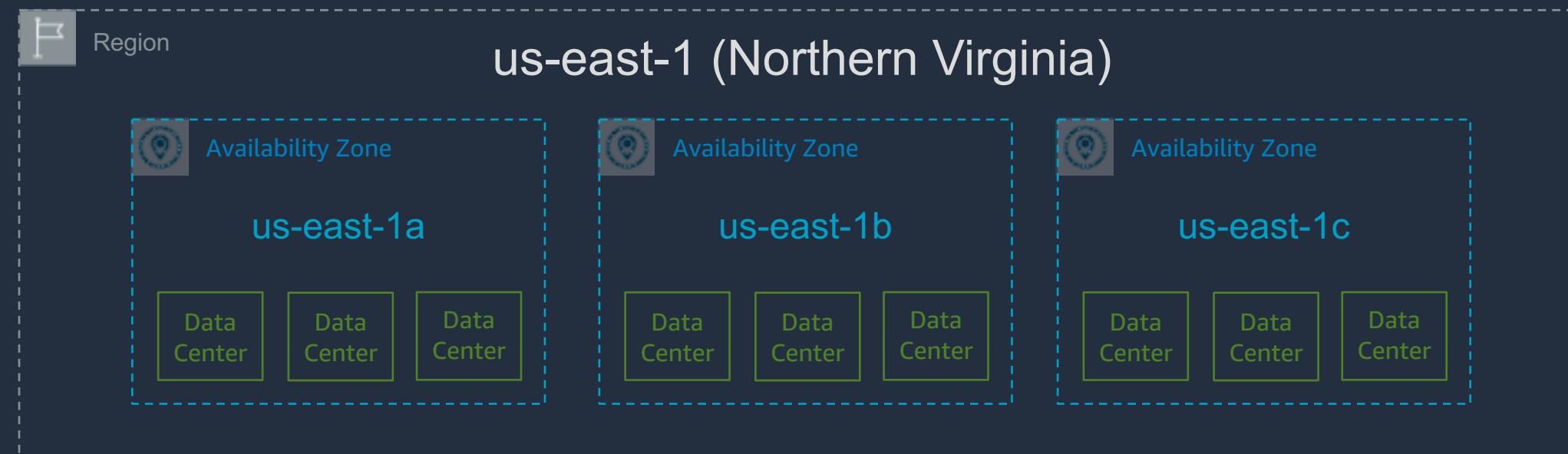
AWS Direct
Connect
locations



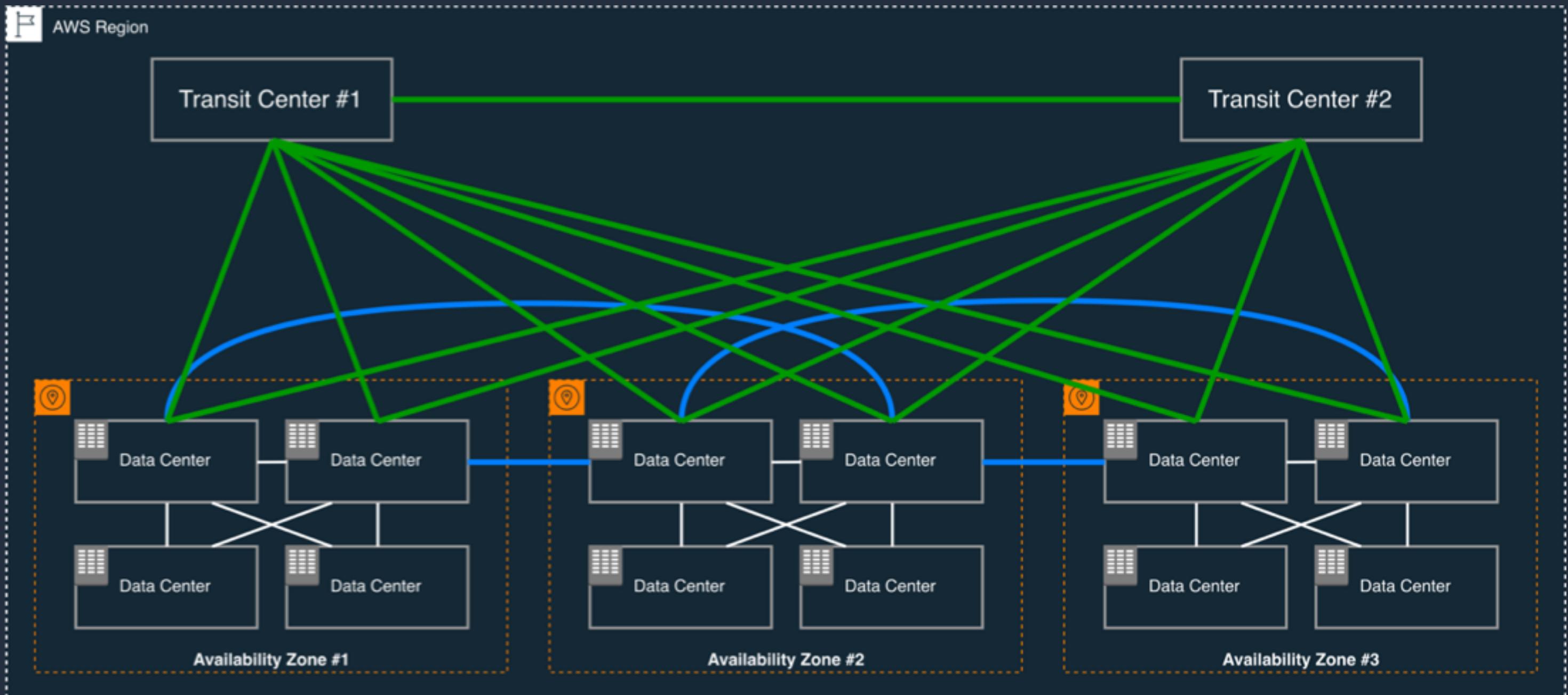


Availability Zones

- A region is comprised of multiple Availability Zones (typically 3)
- Fully independent partitions on isolated fault lines, flood plains, and power grids
- Each AZ: redundant power and redundant dedicated network
- Each AZ: typically multiple data centers
- Between AZs: high throughput, low latency (<10ms) network
- Between AZs: physical separation < 100km (60mi)

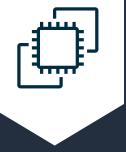


Availability Zones

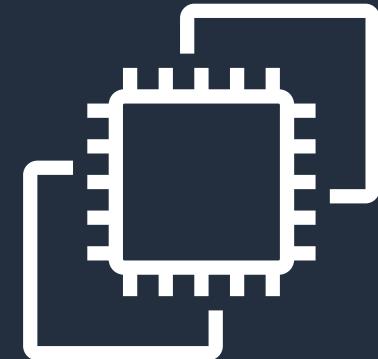


2

EC2 Overview



Choices for Compute



Amazon EC2

Virtual server instances
in the cloud



Amazon ECS, EKS, and Fargate

Container management service
for running
Docker on a managed
cluster of EC2

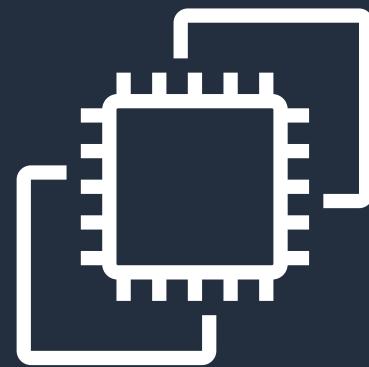


AWS Lambda

Serverless compute
for stateless code execution in
response to triggers



Amazon EC2



Amazon EC2

Linux | Windows

Arm and x86 architectures

General purpose and workload optimized

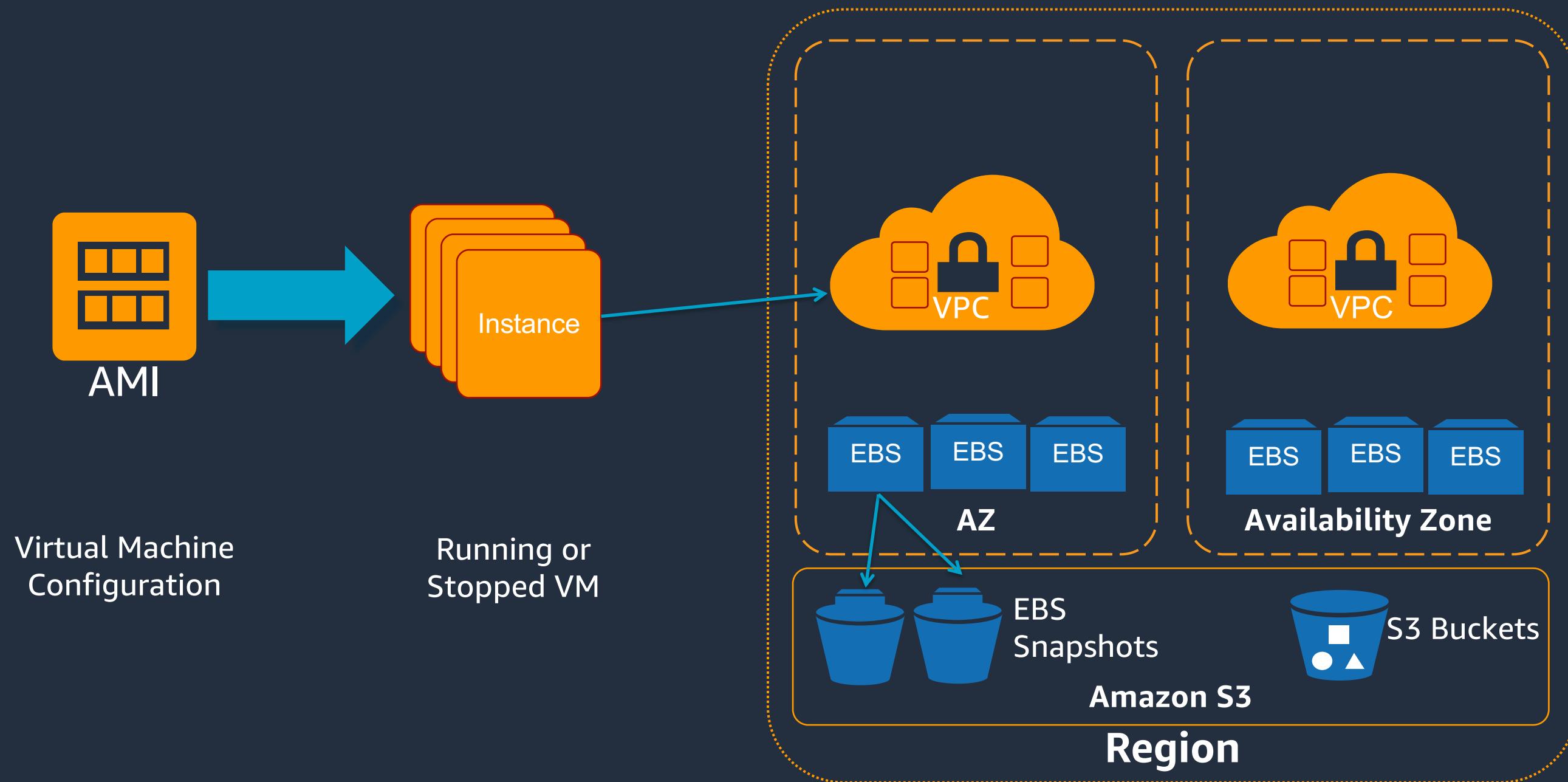
Bare metal, disk, networking capabilities

Packaged | Custom | Community AMIs

Multiple purchase options: On-demand, RI, Spot



EC2 Terminology





Instance sizing





Resource allocation

- All resources assigned to you are dedicated to your instance with no over commitment*
 - All vCPUs are dedicated to you
 - Memory allocated is assigned only to your instance
 - Network resources are partitioned to avoid “noisy neighbors”
- Curious about the number of instances per host?
 - See “Dedicated Hosts Configuration Table” for a guide.

*Again, the “T” family is special



Choose your processor and architecture



Intel® Xeon® Scalable
(Skylake) processor



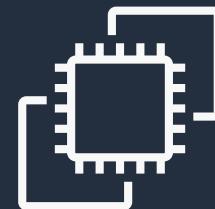
NVIDIA V100
Tensor Core GPUs



AMD EPYC processor



Amazon ARM based
Cloud Processor



FPGAs for custom
hardware acceleration

Right compute for the right application and workload



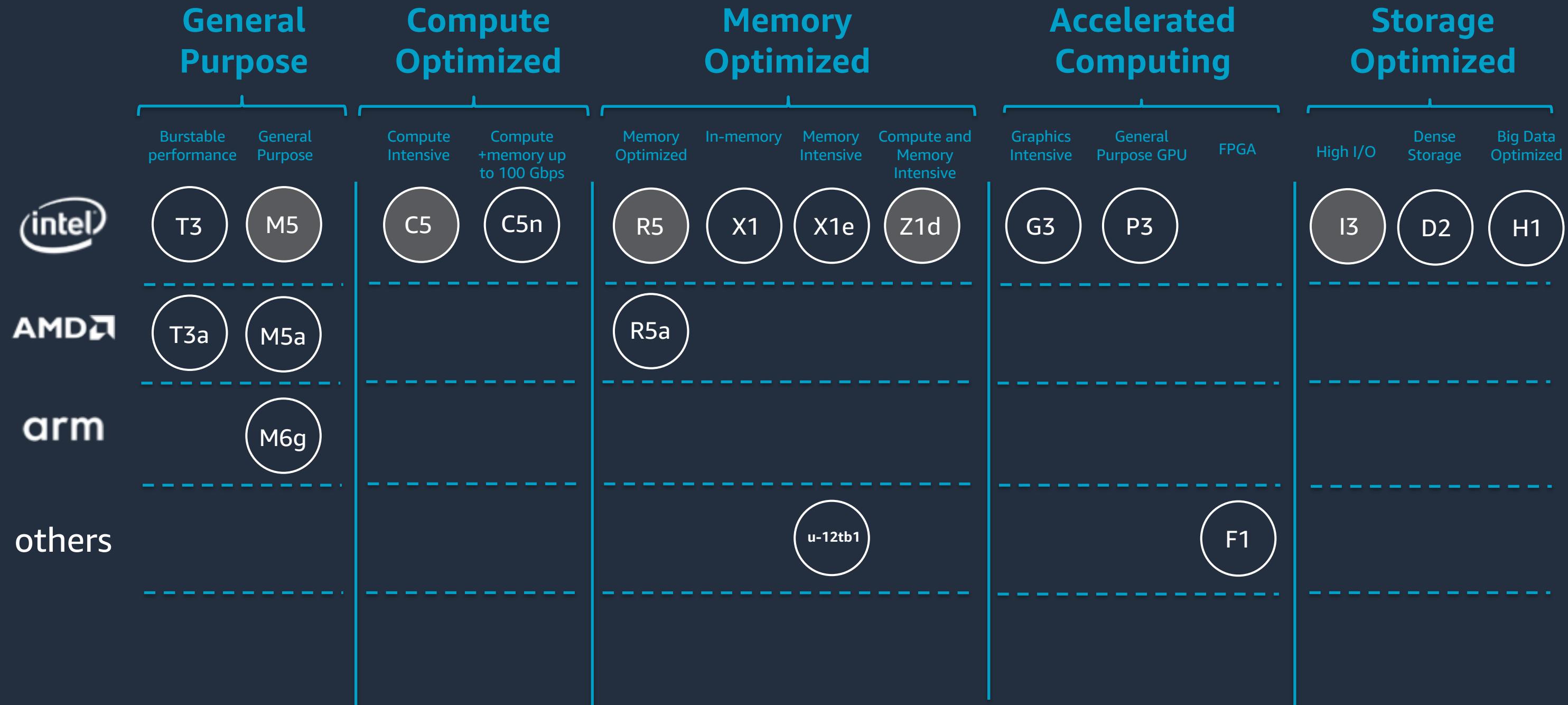
EC2 Naming Explained

Instance generation





Instance Types





EC2 Operating Systems Supported

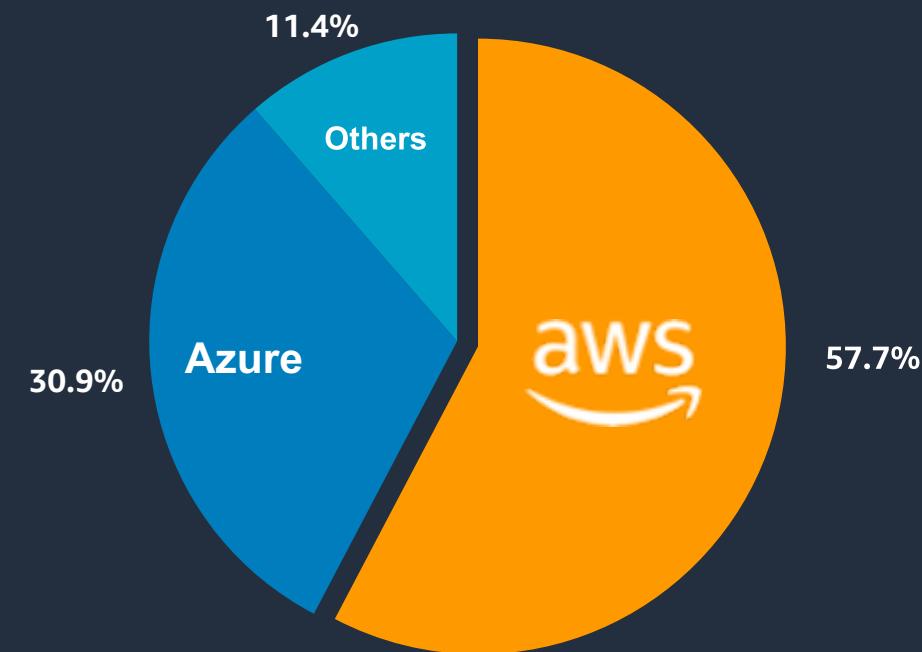
- Windows 2003R2/2008/2008R2/2012/2012R2/2016/2019
- Amazon Linux
- Debian
- Suse
- CentOS
- Red Hat Enterprise Linux
- Ubuntu



for more OSes see: <https://aws.amazon.com/marketplace/b/2649367011>



Windows Licenses by Cloud Provider



Note: Includes Windows instances deployed in the public cloud IaaS market during 2017 Source: IDC estimates, 2018

IDC, Windows Server Operating Environment Market Update, Doc # US44217118, Aug 2018

https://d1.awsstatic.com/analyst-reports/IDC_Slide_WindowsonAWS_JM181015.pdf

What is an Amazon Machine Image (AMI)?



Provides the information required to launch an instance

Launch multiple instances from a single AMI

An AMI includes the following

- A template for the root volume (for example, operating system, applications)
- Launch permissions that control which AWS accounts can use the AMI
- Block device mapping that specifies volumes to attach to the instance



Choosing an AMI

AWS Console

Step 1: Choose an Amazon Machine Image (AMI)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. You can select an AMI provided by AWS, our user community, or the AWS Marketplace; or you can select one of your own AMIs.

Quick Start

My AMIs

AWS Marketplace

Community AMIs

Free tier only

Amazon Linux 2 AMI (HVM), SSD Volume Type - ami-04681a1dbd79675a5

Amazon Linux 2 comes with five years support. It provides Linux kernel 4.14.2.26, Binutils 2.26, and the latest software packages through extras.

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

Select

64-bit

Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type - ami-0f8ba91c07771967

The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

Select

64-bit

Red Hat Enterprise Linux 7.5 (HVM), SSD Volume Type - ami-6871a115

Red Hat Enterprise Linux version 7.5 (HVM), EBS General Purpose (SSD) Volume Type

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

Select

64-bit

AWS Marketplace

aws marketplace

View Categories | Migration Mapping Assistant | Your Saved List | Sell in AWS Marketplace | Amazon Web Services Home | Help

Operating Systems (336 results) showing 1 - 10

Categories: All Categories, Infrastructure Software, Operating Systems

Filters: Vendors (clickwerk Ltd [84], Amazon Web Services [84], Center for Internet Security [28], Thinking Software, Inc. [18], CentOS.org [8], Technology Leadership Corporation [3], Plesk [9], Canonical Group Limited [8], SmartAMM [7], Cloud Linux [6]), Show more

Operating System: All Windows, All Linux/Unix

Software Pricing Plans: Free [104], Hourly [212], Monthly [3]

CentOS 7 (x86_64) - with Updates HVM

★★★★★ (34) | Version 1805_01 | Sold by CentOS.org

This is the Official CentOS 7 x86_64 HVM image that has been built with a minimal profile, suitable for use in HVM instance types only. The image contains just enough packages... Linux/Unix, CentOS 7 - 64-bit Amazon Machine Image (AMI)

CentOS 6 (x86_64) - with Updates HVM

★★★★★ (33) | Version 1805_01 | Sold by CentOS.org

This is the Official CentOS 6 x86_64 HVM image that has been built with a minimal profile. The image contains just enough packages to run within AWS, bring up an SSH Server... Linux/Unix, CentOS 6 - 64-bit Amazon Machine Image (AMI)

Debian GNU/Linux 8 (Jessie)

★★★★★ (36) | Version 8.7 | Sold by Debian

Debian is a computer operating system composed of software packages released as free and open source software primarily under the GNU General Public License along with other... Linux/Unix, Debian 8.6r1 - 64-bit Amazon Machine Image (AMI)

CentOS 6.5 (x86_64) - Release Media

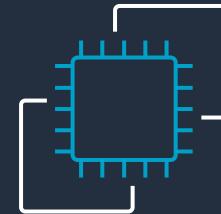
★★★★★ (35) | Version 6.5 - 2013-12-01 | Sold by CentOS.org

This is the Official CentOS 6.5 x86_64 image that has been built with a minimal profile. The image contains just enough packages to run within AWS, bring up an SSH Server...

Use the AMI ID to launch through the API or AWS Command Line Interface (AWS CLI)

```
aws ec2 run-instances --image-id ami-04681a1dbd79675a5 --instance-type c4.8xlarge --count 10 --key-name MyKey
```

Choice of accelerators for specialized workloads



Elastic Graphics

Easily add graphics acceleration to your EC2 instance

Configure right amount of graphics acceleration for your workload

Accelerate application for fraction of cost of standalone graphics instances



Elastic Inference

Reduce deep learning inference costs by up to 75%

Easily attach fractional sizes of a full GPU instance to EC2 or SageMaker instances

Scale inference acceleration up or down as needed with EC2 Auto Scaling

Purchasing Options

On-Demand

Pay for compute capacity by **the second** with no long-term commitments

For Spiky workloads or to define needs



Reserved Instances

Make a 1 or 3-year commitment and receive a **significant discount** off On-Demand prices

For committed utilization



Spot Instances

Spare EC2 capacity at **savings of up to 90%** off On-Demand prices

For time-insensitive or transient workloads
Need to be Fault-tolerant, stateless



Savings Plans

Commit to a \$/h spend and **share discount** across compute options and regions

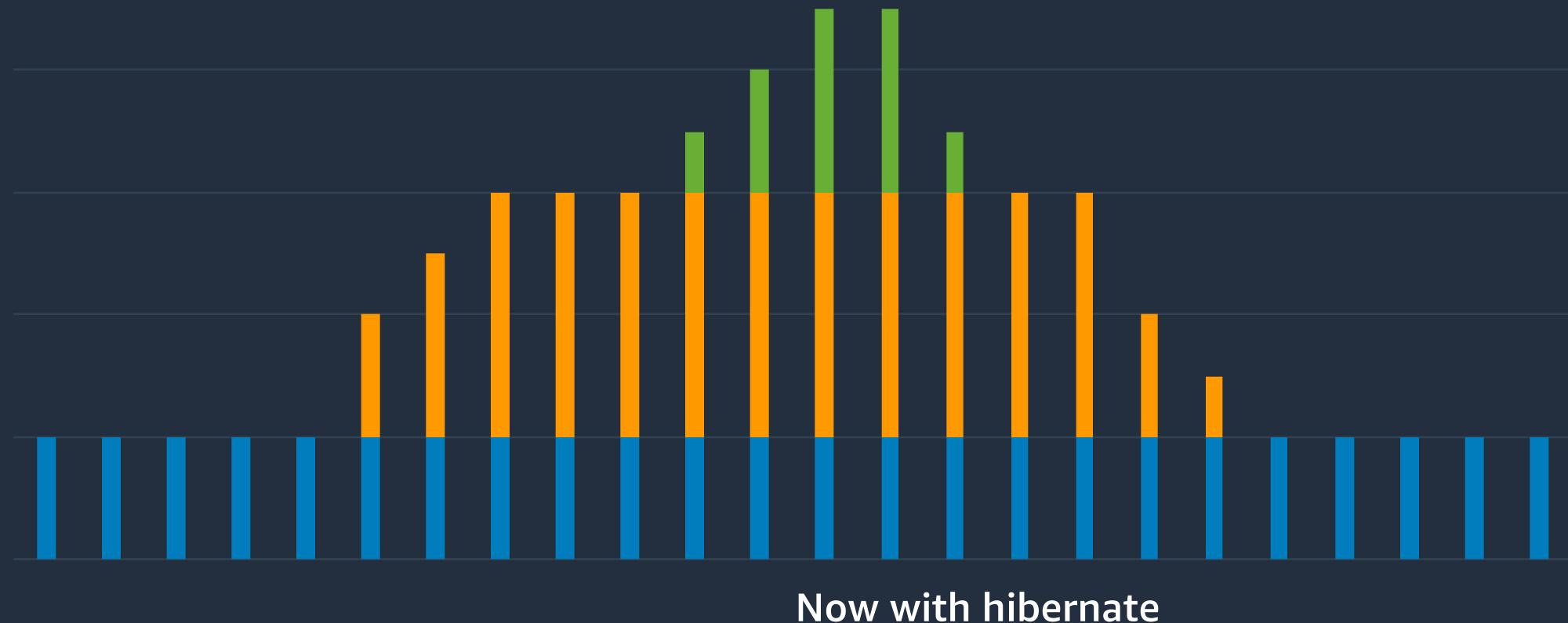
For committed utilization



To optimize EC2, combine all three purchase options!



Simplify capacity and cost optimization



Scale using
Spot,
On-Demand,
or both

Use **Reserved Instances**
for known/steady-state
workloads

AWS services make this easy and efficient



Amazon EC2
Auto Scaling



EC2 Fleet



Amazon Elastic
Container Service



Amazon Elastic
Container Service
for Kubernetes



AWS
Thinkbox



Amazon
EMR



AWS
CloudFormation

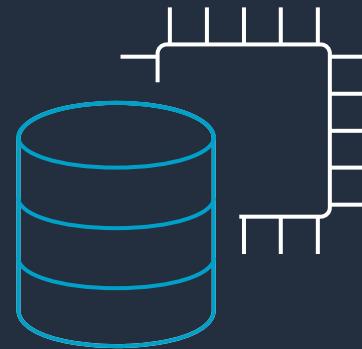


AWS Batch

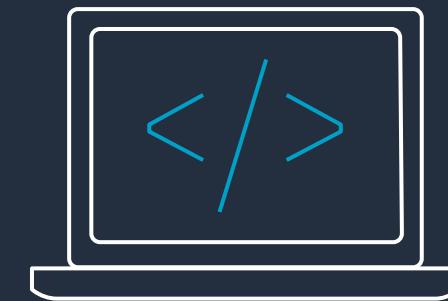


Hibernate Amazon EC2 Instances

Maintain a fleet of pre-warmed instances to quickly get to a productive state



Available with Amazon EBS-backed instances



Use familiar Stop and Start APIs



Memory data saved in EBS root volume



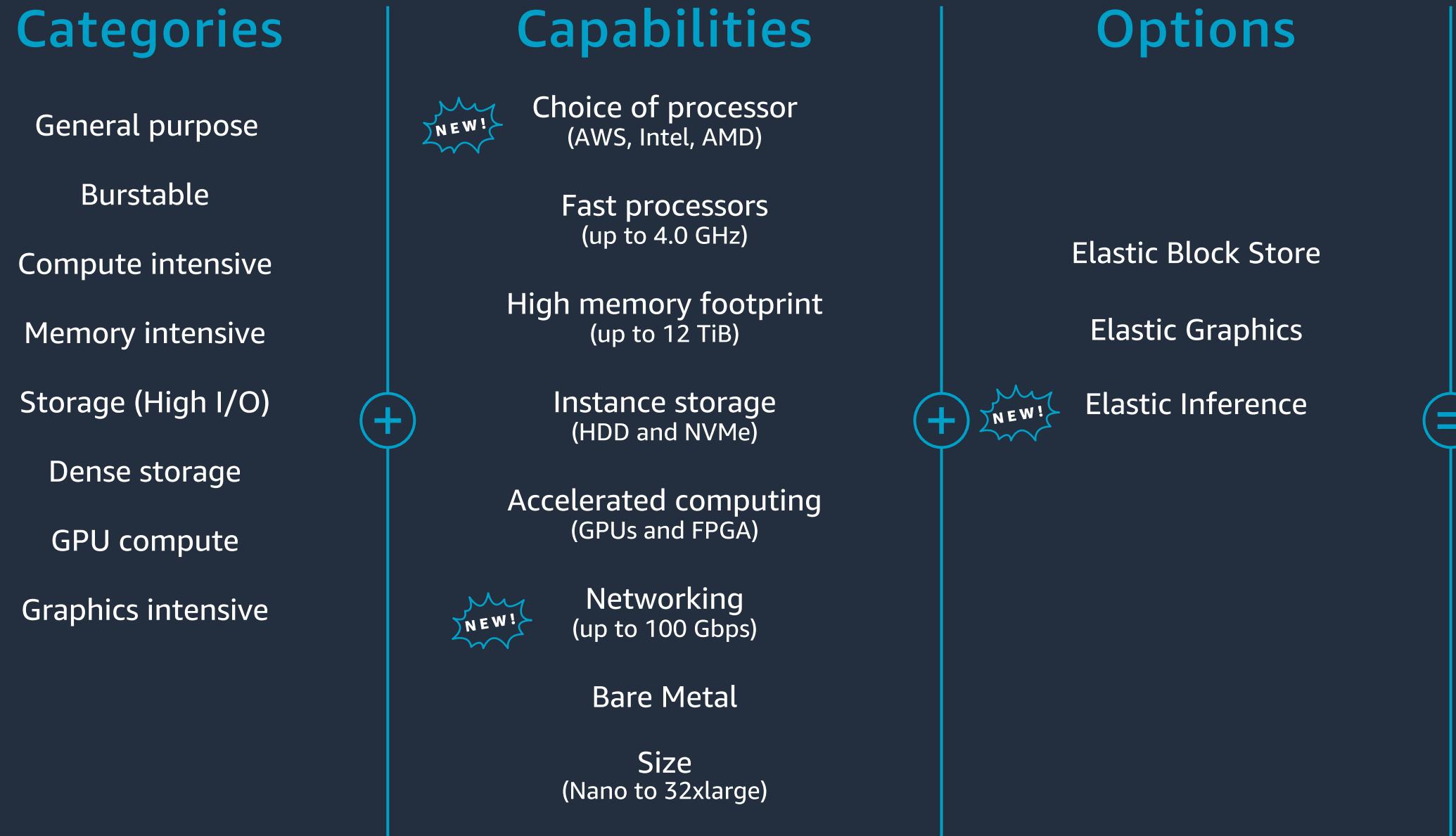
RAM contents are encrypted on EBS

Its just like closing and opening your laptop!

Applications can pick up right where it left off



EC2 Options

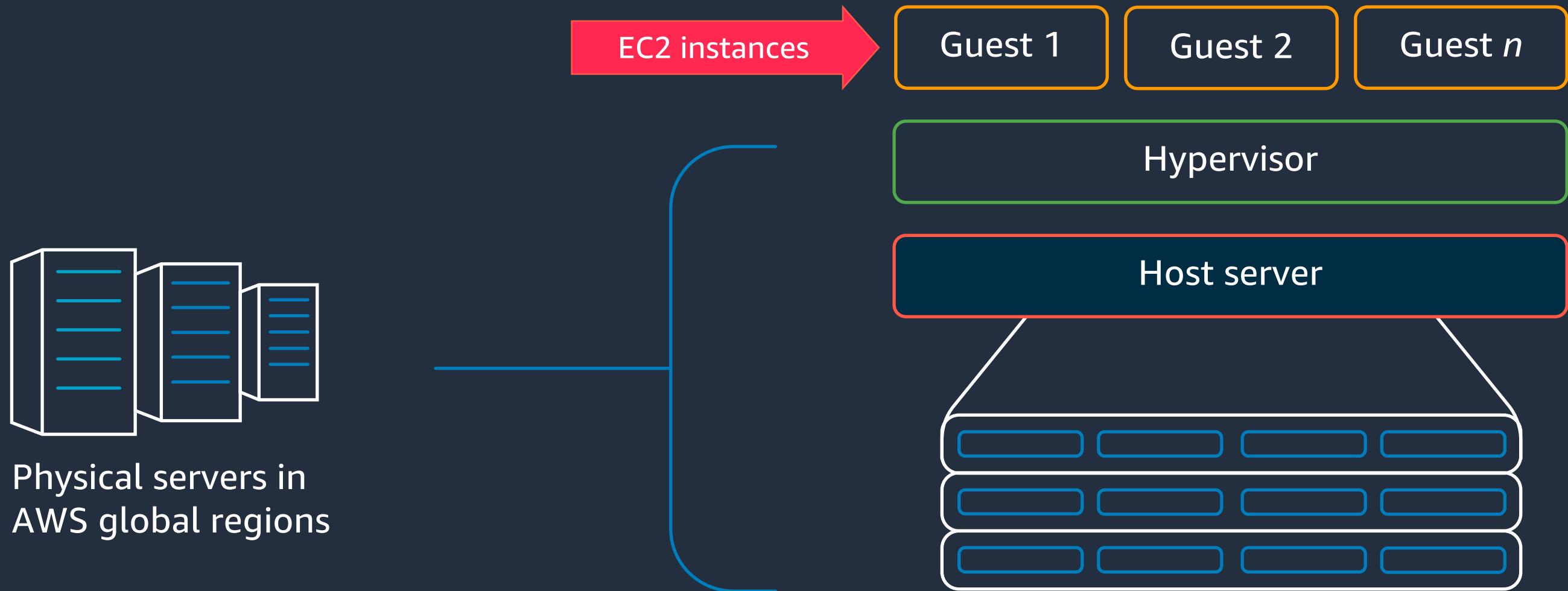


200+
instance types
for virtually
every workload
and business need



EC2 Design

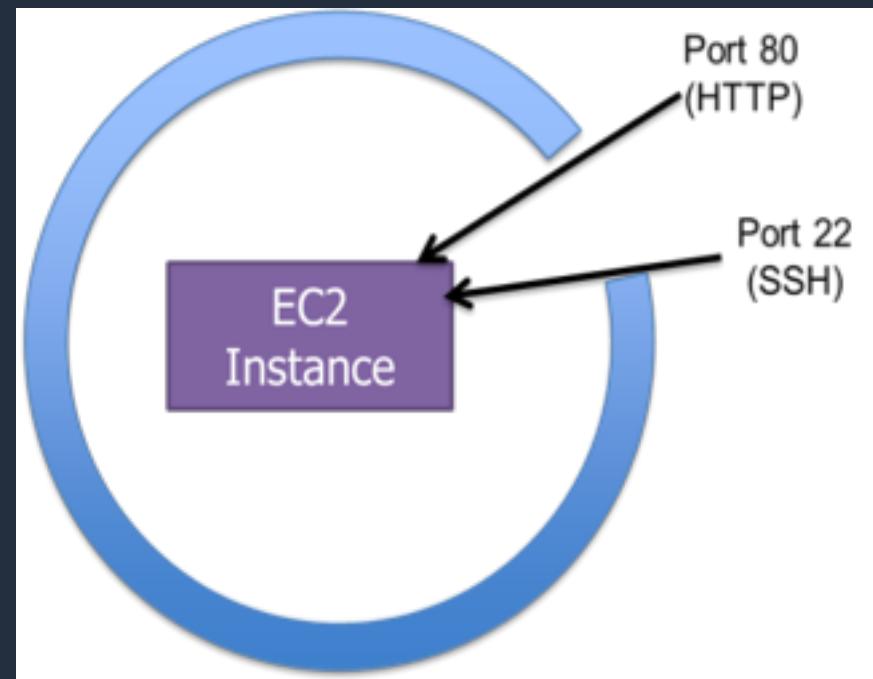
EC2 Host Virtualization



EC2 Security Groups

Security Group Rules

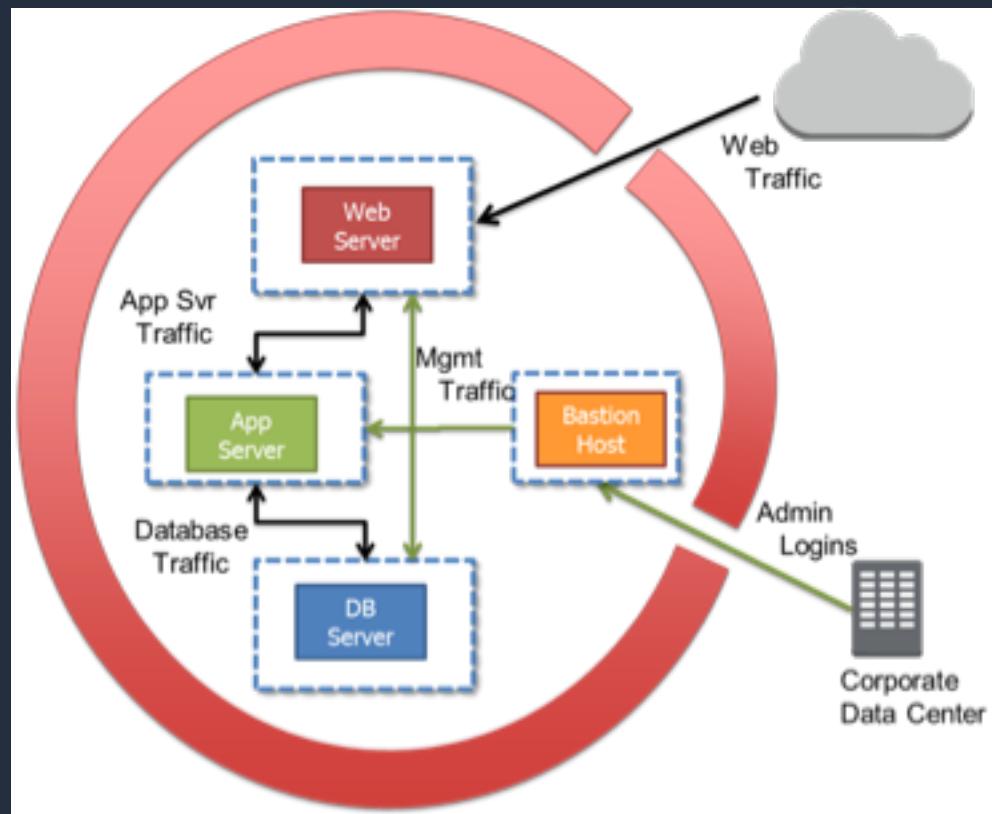
- Name
- Description
- Protocol
- Port range
- IP address, IP range, Security Group name



Tiered EC2 Security Groups

Hierarchical Security Group Rules

- Dynamically created rules
- Based on Security Group membership
- Create tiered network architectures



"Web" Security Group:

TCP 80 0.0.0.0/0

TCP 22 "Mgmt"

"App" Security Group:

TCP 8080 "Web"

TCP 22 "Mgmt"

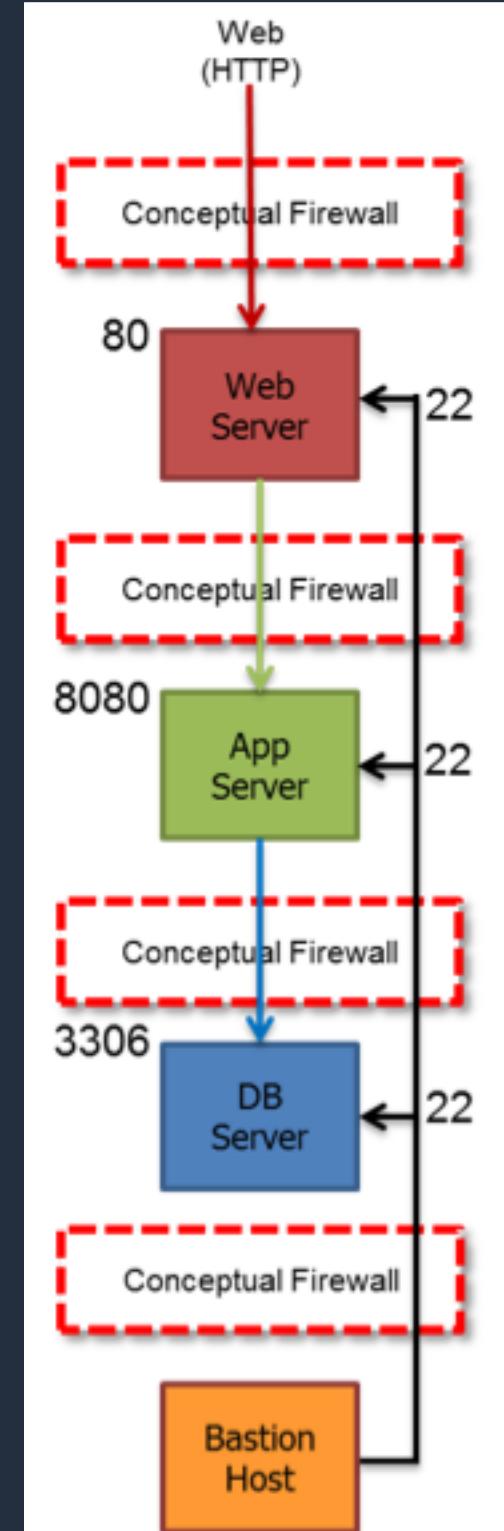
"DB" Security Group:

TCP 3306 "App"

TCP 22 "Mgmt"

"Mgmt" Security Group:

TCP 22 163.128.25.32/32



EC2 IP Addressing

Default VPC	Virtual Private Cloud
Dynamic Private IP	Dynamic or Static Private IP Address
Dynamic Public IP	None by default (can be created with publicIP=true)
Optional Static Public IP (EIP)	Optional Static Public IP (EIP), BYOIP
AWS-provided DNS names <ul style="list-style-type: none">• Private DNS name• Public DNS name	AWS-provided public DNS lookup AWS-provided private DNS names Customer-controlled DNS options

EC2-Specific Credentials

EC2 key pairs

- Linux – SSH key pair for first-time host login
- Windows – Retrieve Administrator password

Standard SSH RSA key pair

- Public/Private Keys
- Private keys are not stored by AWS

AWS approach for providing initial access to a generic OS

- Secure
- Personalized
- Non-generic (NIST, PCI DSS)

“Public Half” inserted by Amazon into each EC2 instance that you launch



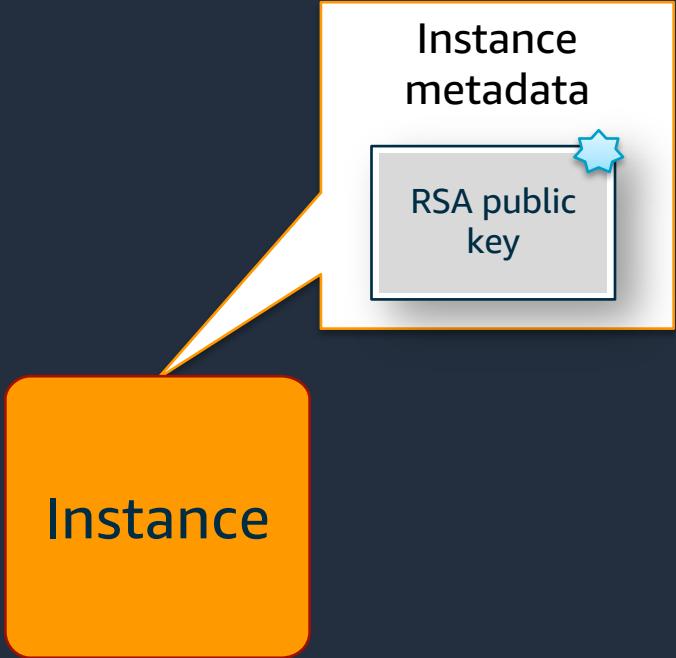
“Private Half” downloaded to your desktop



EC2 Instance access and Key Pairs

Linux launch (first boot)

- Public key made available through metadata
- Public key inserted into `~/.ssh/authorized_keys`
- User connects with SSH using their private key



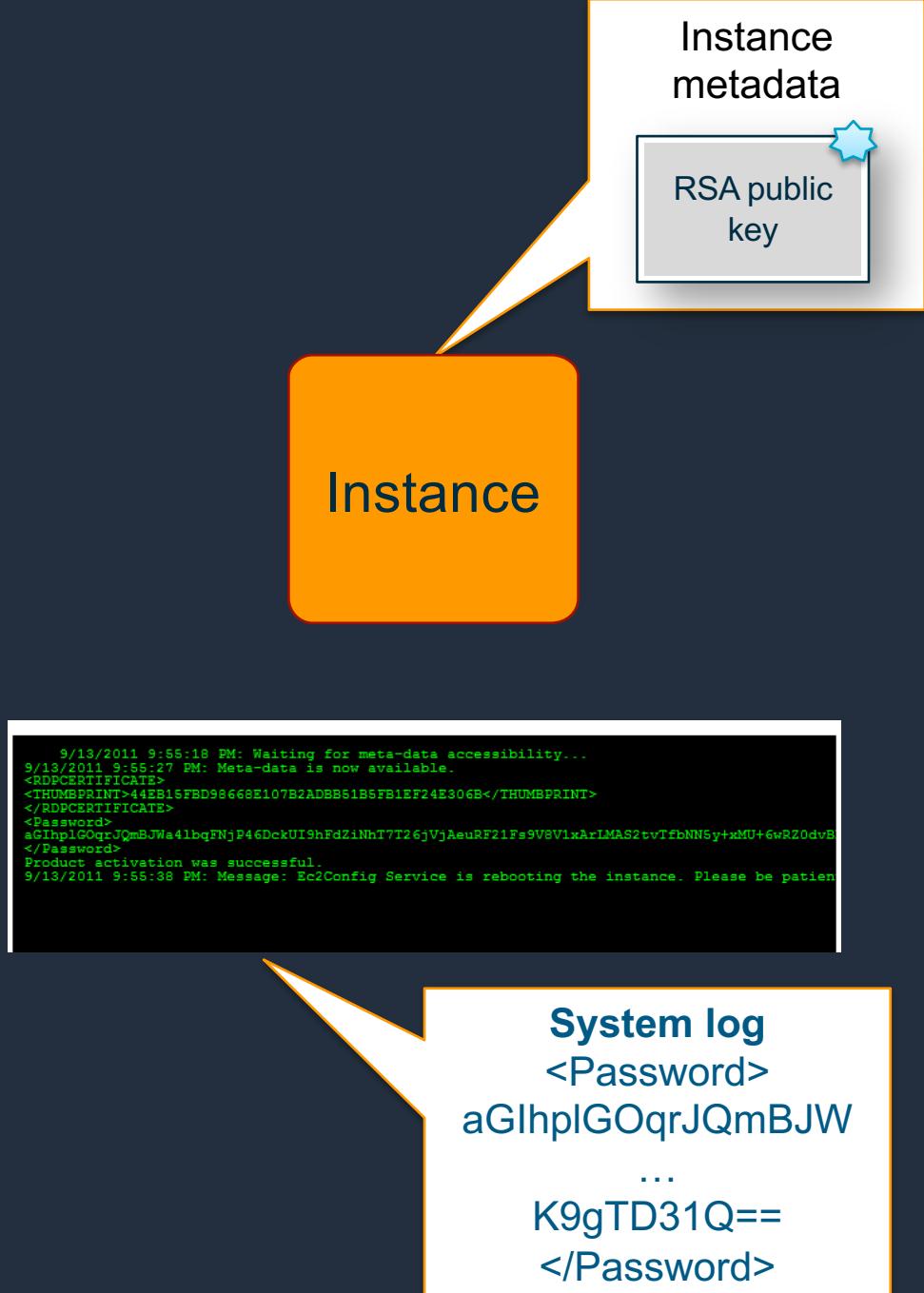
EC2 Instance access and Key Pairs

Linux launch (first boot)

- Public key made available through metadata
- Public key inserted into `~/.ssh/authorized_keys`
- User connects with SSH using their private key

Windows launch (first boot sequence)

- Public key made available through metadata
- Random Administrator password
- Password encrypted with public key
- User decrypts password with their private key



Instance Metadata

<http://169.254.169.254/latest/meta-data/> contains a wealth of info

- ami-id
- ami-launch-index
- ami-manifest-path
- block-device-mapping/
- hostname
- instance-action
- ★ **instance-id**
- instance-type
- kernel-id
- local-hostname
- local-ipv4
- mac
- network/
- ★ **placement/availability-zone**
- profile
- public-hostname
- public-ipv4
- public-keys/



Any Questions?

