



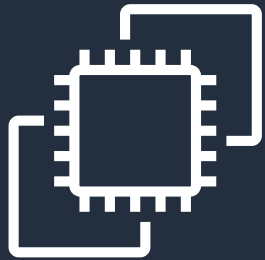
Introduction to EC2



Agenda

- EC2 Overview
- EC2 Details

Choices for Compute



Amazon EC2

Virtual server instances
in the cloud



Amazon ECS, EKS, and Fargate

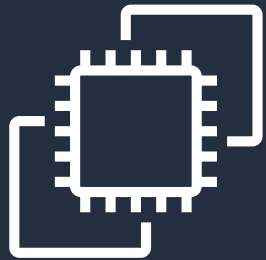
Container management service
for running
Docker on a managed
cluster of EC2



AWS Lambda

Serverless compute
for stateless code execution in
response to triggers

Amazon EC2



Amazon EC2

Linux | Windows

Arm and x86 architectures

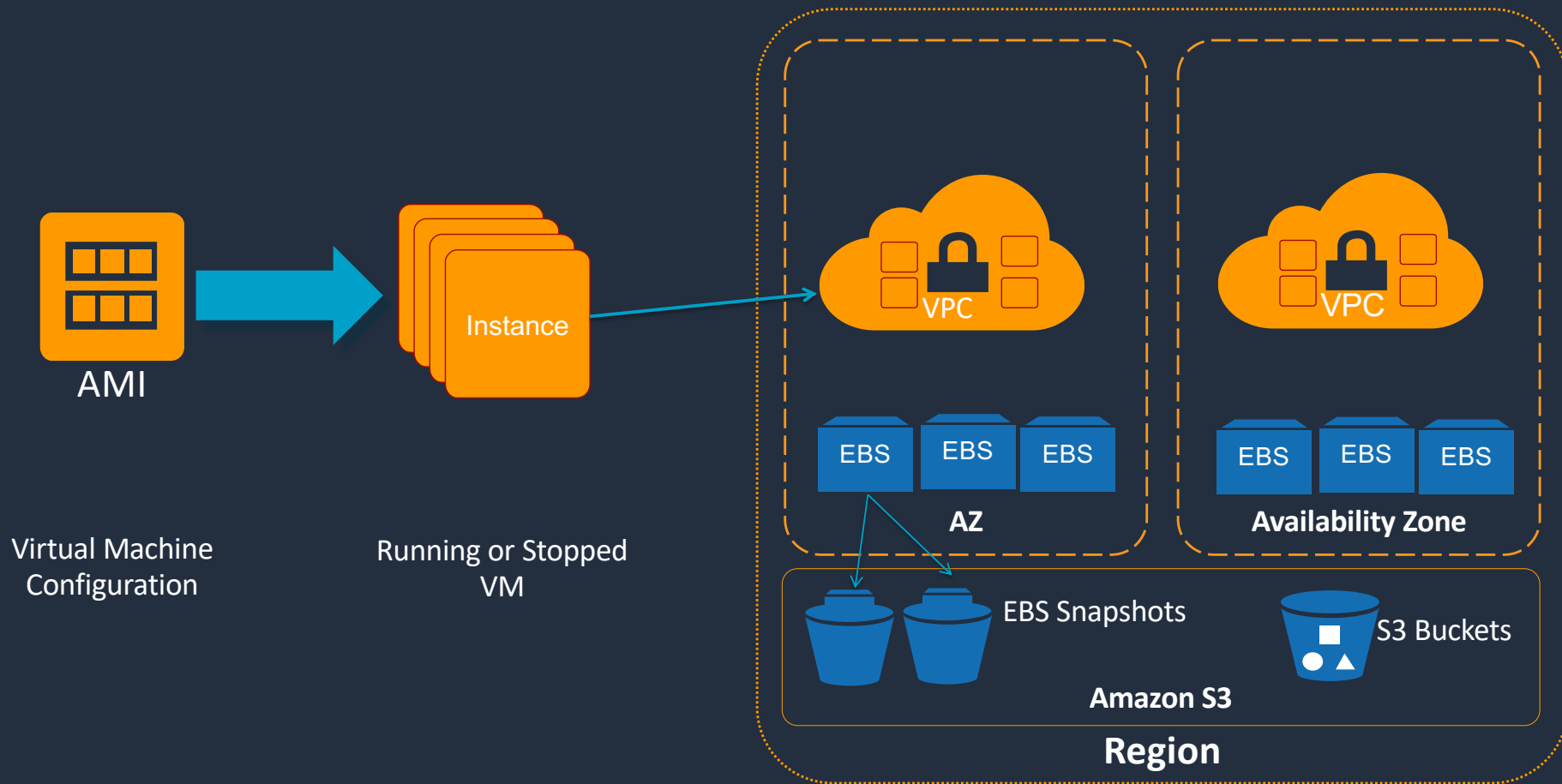
General purpose and workload optimized

Bare metal, disk, networking capabilities

Packaged | Custom | Community AMIs

Multiple purchase options: On-demand, RI, Spot

EC2 Terminology





What's a virtual CPU? (vCPU)

- A vCPU is typically a hyper-threaded physical core*
 - Divide vCPU count by 2 to get core count
 - On Linux, "A" threads enumerated before "B" threads
 - On Windows, threads are interleaved
-
- Cores by Amazon EC2 & RDS DB Instance type:
<https://aws.amazon.com/ec2/virtualcores/>

** CPU Optimizing options allow disabling hyperthreading and reduce number of cores*

Memory and Storage



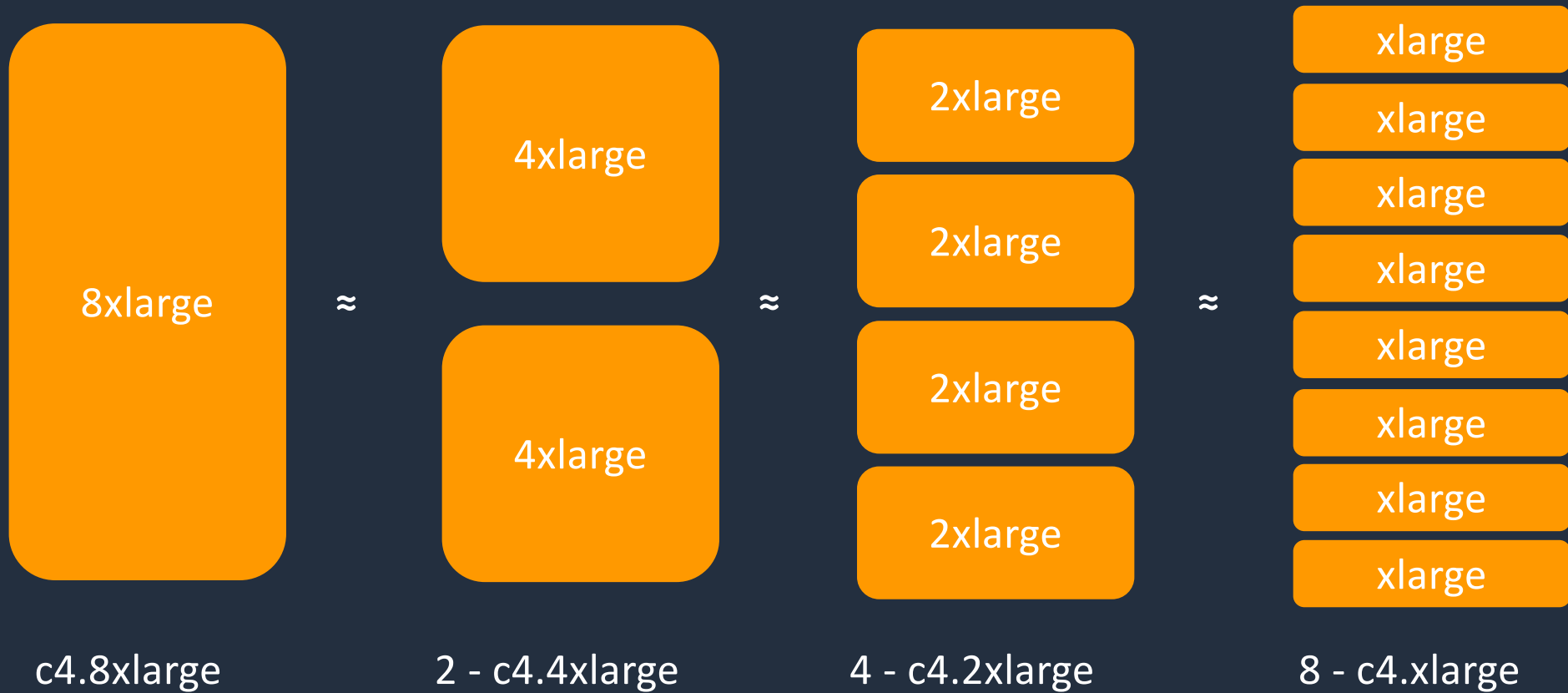
What's a GiB?

- Memory is presented as GibiBytes (GiB) and not Gigabytes (GB)
- 256 GiB = 275 GB

What about storage?

- Storage is independent of compute
- You allocate drives known as EBS volumes
- Max 16 TiB per volume
- Some instance types provide physically attached (ephemeral) storage

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



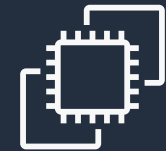
NVIDIA V100
Tensor Core GPUs



AMD EPYC processor



AWS Graviton
Processor (arm)



FPGAs for custom
hardware acceleration

Right compute for the right application and workload

EC2 Naming Explained



Instance generation

c5n.xlarge

Instance
family

Attribute

Instance size

Instance Types



General Purpose

Compute Optimized

Memory Optimized

Accelerated Computing

Storage Optimized

Burstable performance

General Purpose

Compute Intensive

Compute +memory up to 100 Gbps

Memory Optimized

In-memory

Memory Intensive

Compute and Memory Intensive

Graphics Intensive

General Purpose GPU

FPGA

High I/O

Dense Storage

Big Data Optimized



T3

M5

C5

C5n

R5

X1

X1e

G3

P3

F1

D2

H1

Local storage (NVMe SSD)

M5d

C5d

R5d

Z1d

I3



T3a

M5a

R5a

metal

M5m

c5m

R5m

u-12tb1

Z1dm

I3m

others

A1

M6g

C6g

R6g

P3dn

I3en

arm

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>

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

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 1: Choose an Amazon Machine Image (AMI)

Cancel and Exit

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 < 1 to 35 of 35 AMIs >

- Amazon Linux 2 AMI (HVM), SSD Volume Type -**
ami-04681a1dbd79675a5
64-bit
Free tier eligible
Amazon Linux 2 comes with five years support. It provides Linux kernel 4.14 tuned for optimal performance on Amazon EC2, systemd 219, GCC 7.3, Glibc 2.26, Binutils 2.29.1, and the latest software packages through extras.
Root device type: ebs Virtualization type: hvm ENA Enabled: Yes
- Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type -**
ami-0ff8a9100777f867
64-bit
Free tier eligible
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
- Red Hat Enterprise Linux 7.5 (HVM), SSD Volume Type -**
ami-6871a115
64-bit
Free tier eligible
Red Hat Enterprise Linux version 7.5 (HVM), EBS General Purpose (SSD) Volume Type
Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

AWS Marketplace

aws marketplace Hello, duff

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

Categories: All Categories Infrastructure Software Operating Systems

Filters: Vendors: clckwrk Ltd (84), Amazon Web Services (84), Center for Internet Security (20), Thinking Software, Inc. (13), CentOS.org (9), Technology Leadership Corporation (9), Plesk (9), Canonical Group Limited (8), SmartAMI (7), Cloud Linux (6), Show more

Operating System: All Windows, All Linux/Unix

Software Pricing Plans: Free (104), Hourly (212), Monthly (3)

Operating Systems (336 results) showing 1 - 10

- CentOS 7 (x86_64) - with Updates HVM**
★★★★★ (58) | 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)**
★★★★★ (86) | 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 8.6+1 - 64-bit Amazon Machine Image (AMI)
- CentOS 6.5 (x86_64) - Release Media**
★★★★★ (55) | 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...

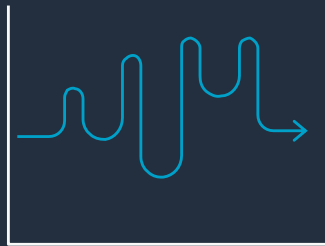
Use the AMI ID to launch through the API or AWS Command Line Interface (AWS CLI)

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

Amazon EC2 purchase options

On-Demand

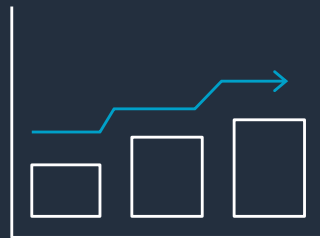
Pay for compute capacity by **the second** with no long-term commitments



Spiky workloads, to define needs

Reserved Instances

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



Committed and steady-state usage

Savings Plan

Same great discounts as Amazon EC2 RIs with **more flexibility**



Committed flexible access to compute

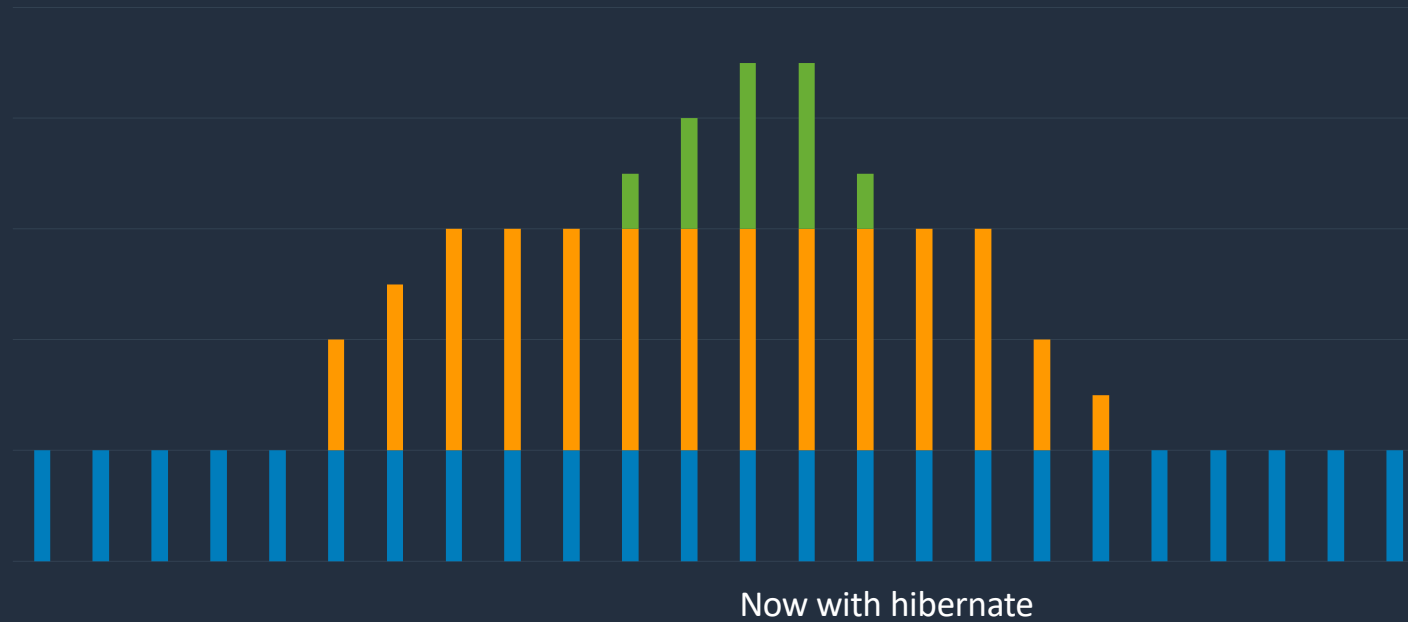
Spot Instances

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



Fault-tolerant, flexible, stateless workloads

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



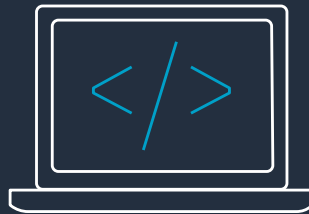
AWS Batch

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!

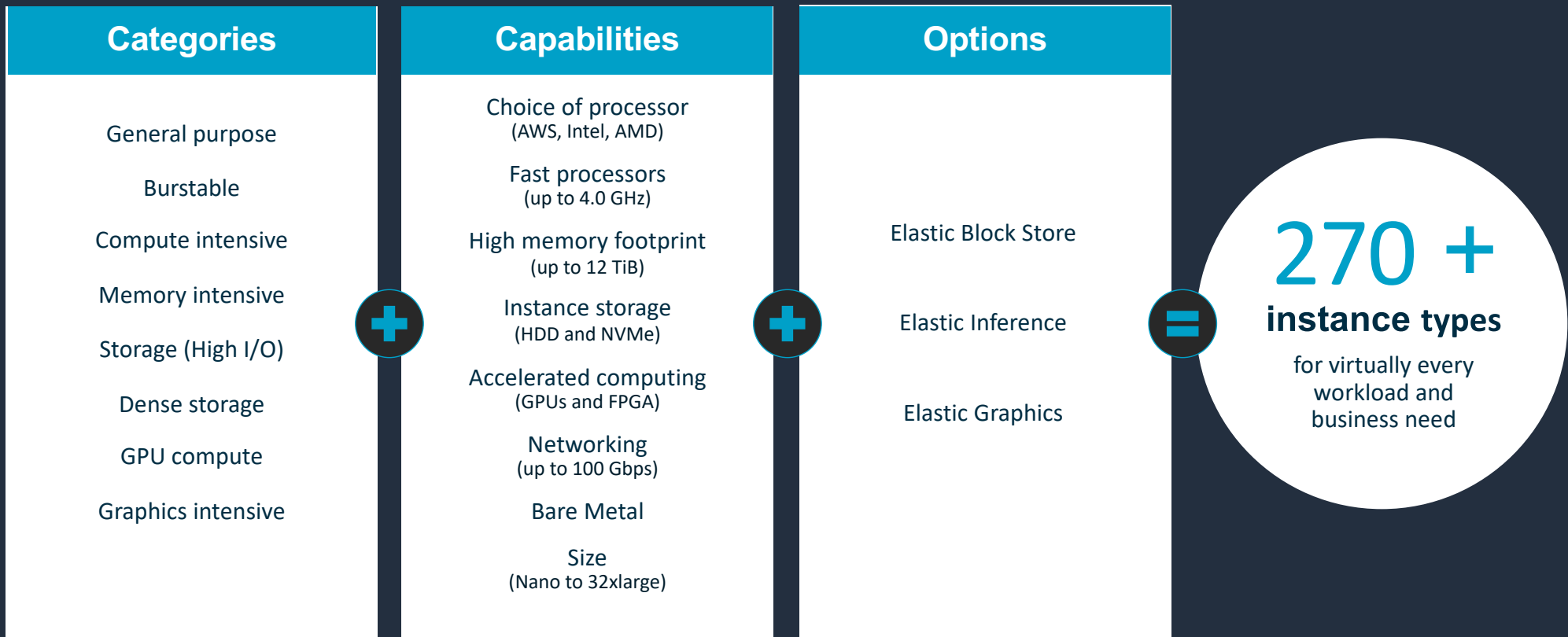
Applications can pick up right where it left off

270+ instances across 42 instance Families

270 +

2017

Broadest and deepest platform choice



Broadest choice of processors



Second generation of
Intel® Xeon processor



AMD Rome



Graviton

AWS Graviton2 Processor

Enabling the best price/performance for your cloud workloads

Graviton Processor



First Arm-based processor available in major cloud



Built on 64-bit Arm Neoverse cores with AWS-designed silicon using 16 nm manufacturing technology



Up to 16 vCPUs, 10 Gbps enhanced networking, 3.5 Gbps EBS bandwidth

Graviton2 Processor



7x performance, 4x compute cores, and 5x faster memory



Built with 64-bit Arm Neoverse cores with AWS-designed silicon using 7 nm manufacturing technology



Up to 64 vCPUs, 25 Gbps enhanced networking, 18 Gbps EBS bandwidth

AWS Graviton2 based instances

Up to 40% better price-performance for general purpose, compute intensive, and memory intensive workloads.

M6g

Built for: General-purpose workloads such as application servers, mid-size data stores, and microservices.

C6g

Built for: Compute intensive applications such as HPC, video encoding, gaming, and simulation workloads.

R6g

Built for: Memory intensive workloads such as open-source databases, or in-memory caches.

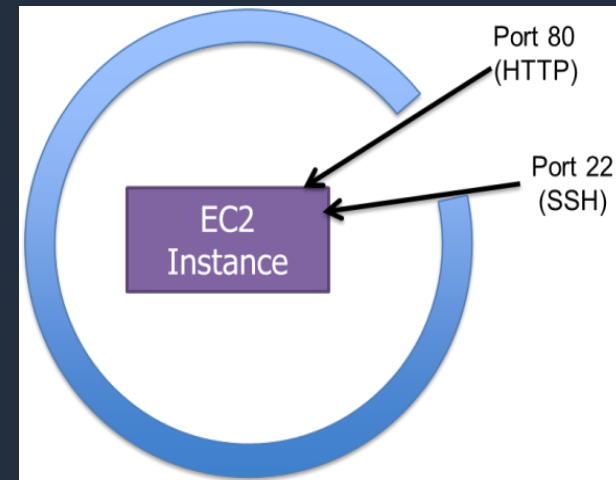
Launched in 2020

Local NVMe-based SSD storage options also available in general purpose (M6gd), compute-optimized (C6gd), and memory-optimized (R6gd) instances

EC2 Security Groups

Security Group Rules

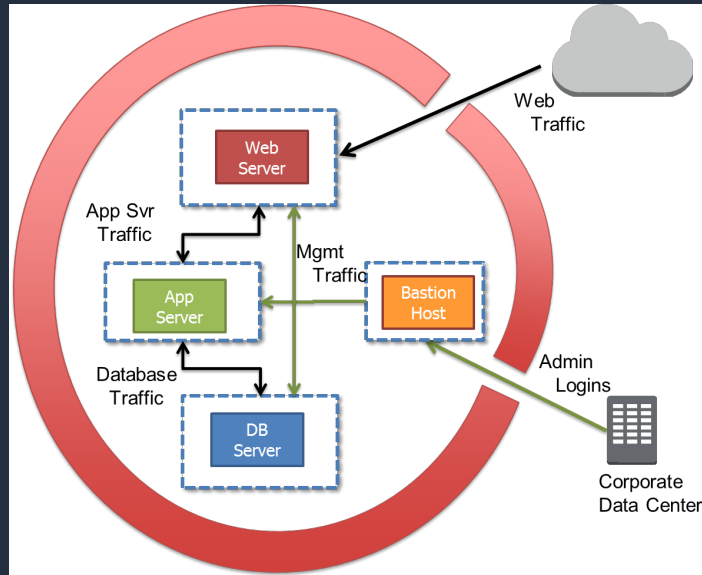
- Name
- Description
- Protocol
- Port range
- IP address, IP range, Security Group name



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:

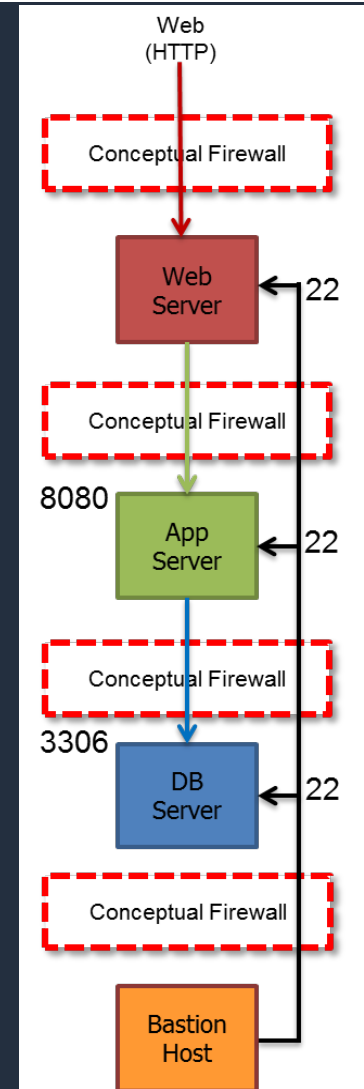
TCP 8080 “Web”
TCP 22 “Mgmt”

“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

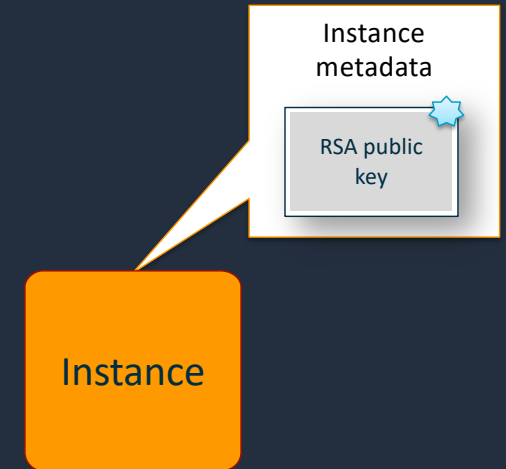


“Private Half” downloaded to your desktop

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**



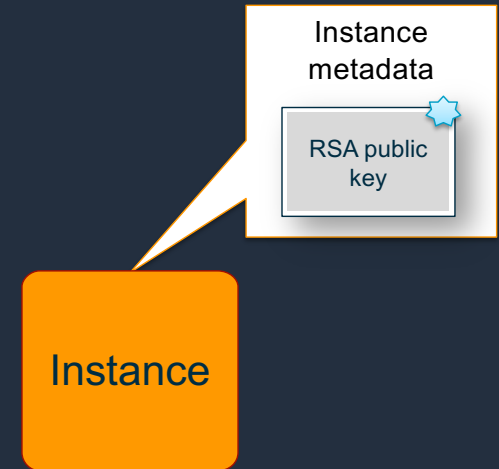
EC2 Instance access and Key Pairs

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
- Sysprep
- Random Administrator password
- Password encrypted with public key
- User decrypts password with their **private key**



```
9/13/2011 9:55:18 PM: Waiting for meta-data accessibility...
9/13/2011 9:55:27 PM: Meta-data is now available.
<RDPCERTIFICATE>
<THUMBPRINT>44BB15FBD98668E107B2ADBB61B5FB1EF24E906B</THUMBPRINT>
</RDPCERTIFICATE>
<Password>
aGIhplGOqrJQmBJW41bqFNjP46DokUI9hFdZiNh77T26jVjAuuRF21F9V8V1aArLMS2cvTfbNNSyxdU+6wR20dv8
</Password>
Provision activation was successful.
9/13/2011 9:56:38 PM: Message: Ec2Config Service is rebooting the instance. Please be patient.
```

System log
<Password>
aGIhplGOqrJQmBJW
...
K9gTD31Q==
</Password>

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?

