Amazon Elastic Compute Cloud (EC2)

Amazon EC2

Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity (Virtual Machine) in the AWS cloud.

Why EC2


Amazon EC2
Available in different locations
Choice of instance families with differing resource ratios
Extensive list of supported operating systems & software
Import and export virtual machines
Integrated with other AWS Services
Purchasing options for cost optimization

List of Supported Operating Systems

- RedHat Linux
- Windows Server
- SuSE Linux
- Ubuntu
- Fedora
- Debian
- Cent OS
- Gentoo Linux
- Oracle Linux
- FreeBSD

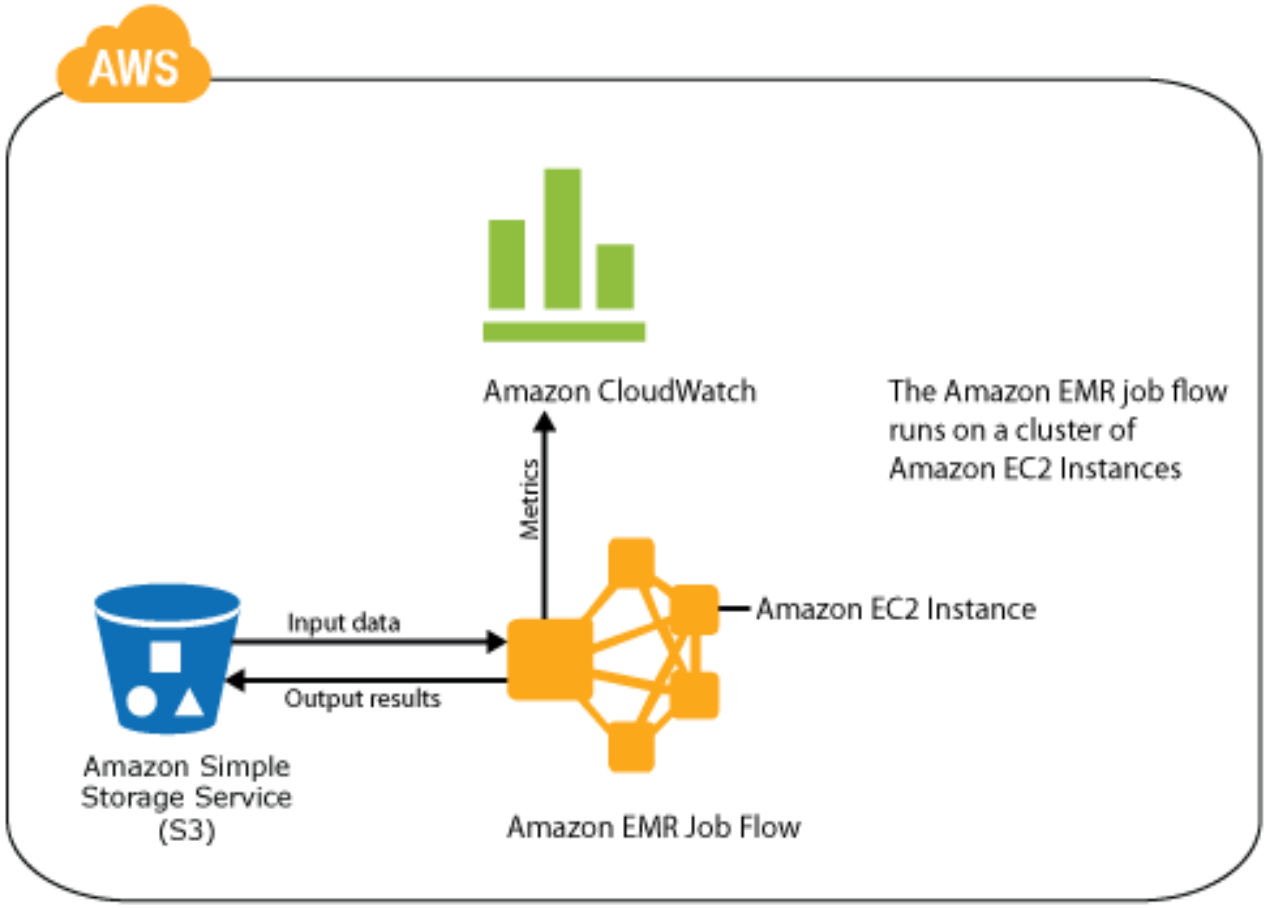
List of Supported Software

(100+)
Business Intelligence products

(250+)
Application Stacks products

(225+)
Content Management Products

EC2 Integration with other AWS Services



EC2 Instance Type

- Instance types comprise varying combinations of CPU, memory, storage, and networking capacity and give you the flexibility to choose the appropriate mix of resources for your applications.

General Purpose Optimized

T2 - T2 instances are Burstable Performance Instances that provide a baseline level of CPU performance with the ability to burst above the baseline.	•High Frequency Intel Xeon Processors with Turbo up to 3.3GHz •Burstable CPU, governed by CPU Credits, and consistent baseline performance
M4 - M4 instances are the latest generation of General Purpose Instances. This family provides a balance of compute, memory, and network resources	•2.4 GHz Intel Xeon® E5-2676 v3 (Haswell) processors •EBS-optimized by default at no additional cost
M3 – M3 provides a balance of compute, memory, and network resources, and it is a good choice for many applications.	•High Frequency Intel Xeon E5-2670 v2 (Ivy Bridge) Processors* •SSD-based instance storage for fast I/O performance

EC2 Instance Type (Cont'd)

- Compute Optimized

C4 - instances are the latest generation of Compute-optimized instances, featuring the highest performing processors and the lowest price/compute performance in EC2.	•High frequency Intel Xeon E5-2666 v3 (Haswell) processors optimized specifically for EC2 •EBS-optimized by default and at no additional cost
C3 - High performance front-end fleets, web-servers, batch processing, distributed analytics, high performance science and engineering applications, ad serving, MMO gaming, and video-encoding.	•High Frequency Intel Xeon E5-2680 v2 (Ivy Bridge) Processors •Support for Enhanced Networking

- Memory Optimized

X1 - Instances are optimized for large-scale, enterprise-class, in-memory applications and have the lowest price per GiB of RAM among Amazon EC2 instance types.	High Frequency Intel Xeon E7-8880 v3 (Haswell) Processors Lowest price per GiB of RAM
R3 - R3 instances are optimized for memory-intensive applications and offer lower price per GiB of RAM.	•High Frequency Intel Xeon E5-2670 v2 (Ivy Bridge) Processors •SSD Storage

EC2 Instance Type (Cont'd)

- Storage Optimized

I2 - I2 provide very fast SSD-backed instance storage optimized for very high random I/O performance, and provide high IOPS at a low cost.	•High Frequency Intel Xeon E5-2670 v2 (Ivy Bridge) Processors •SSD Storage
D2 - D2 instances feature up to 48 TB of HDD-based local storage, deliver high disk throughput, and offer the lowest price per disk throughput performance on Amazon EC2.	•High-frequency Intel Xeon E5-2676v3 (Haswell) processors •HDD storage

EC2 Instance Addressing

Every instance is assigned:

- Private IP address
- Private DNS name(*.internal)
- Public DNS name(*.amazonaws.com)
- Public IP address

- An Elastic IP address is a static IP address designed for dynamic cloud computing. An Elastic IP address is associated with your AWS account.
- Elastic IP address is a public IP address, which is reachable from the Internet. If your instance does not have a public IP address, you can associate an Elastic IP address with your instance to enable communication with the Internet

Amazon EC2 Key Pairs

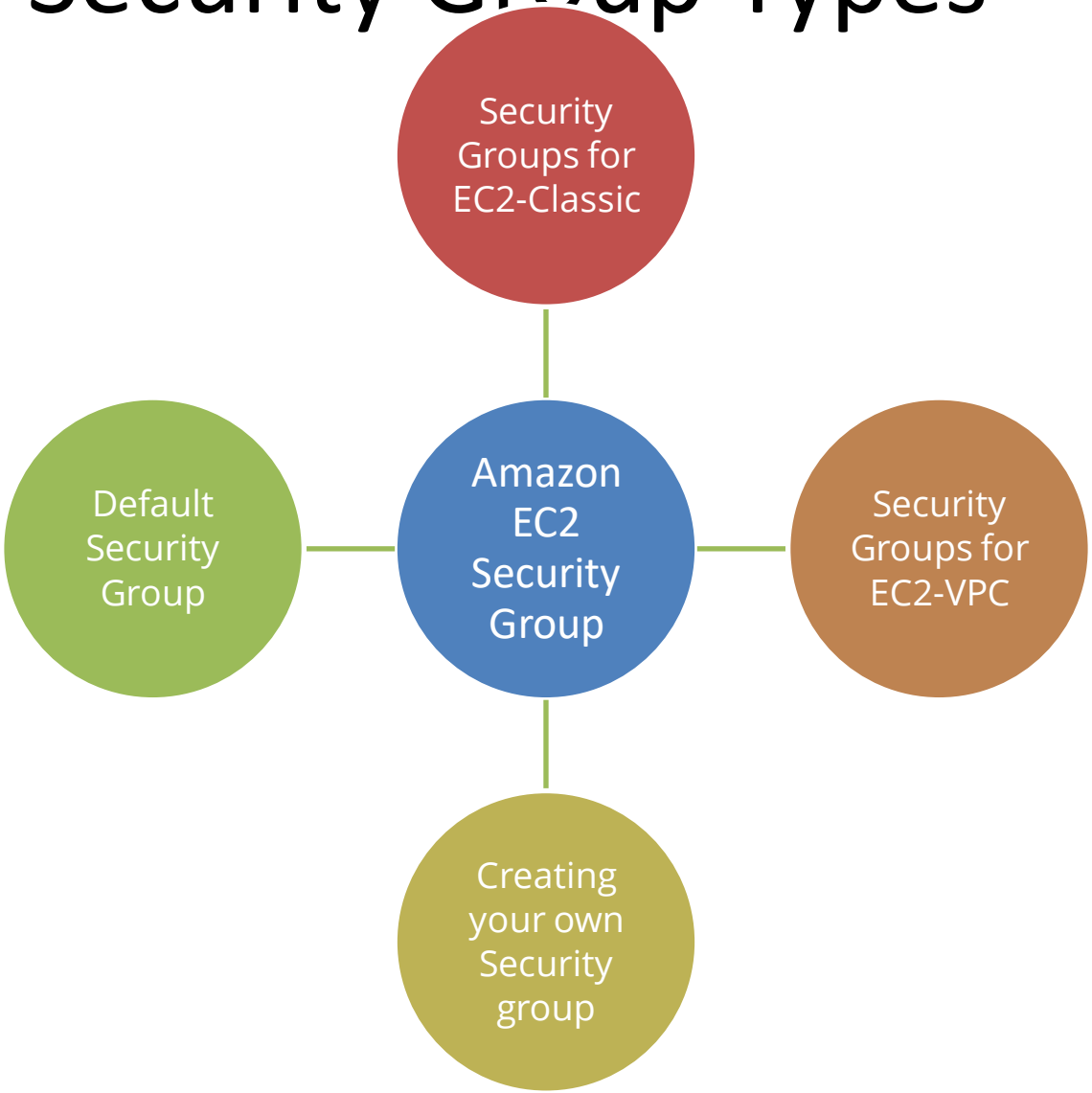
Public–key cryptography :This methodology uses a public key to encrypt a piece of data, such as a password, then the recipient uses the private key to decrypt the data.

- Amazon EC2 uses public–key cryptography to encrypt and decrypt login information.
- The public and private keys are known as a key pair.
- To log in to your instance, you must create a key pair, specify the name of the key pair when you launch the instance, and provide the private key when you connect to the instance.
- Linux instances have no password, and you use a key pair to log in using SSH
- Windows instances, you use a key pair to obtain the administrator password and then log in using RDP.

Security Group

- A Security Group acts as a firewall that controls the traffic allowed to reach one or more instances
- When you launch an instance ,you assign it one or more Security Groups
- You can add rules to each Security Group that control traffic for the instance
- The new rules are automatically applied to all instances to which the security group is assigned
- They are designed to control who can communicate(initiate communication) with your machine
- The default Security Group in each AWS region(that's automatically created for your account by Amazon) does NOT allow any incoming connections
- Therefore, you will need to create a new Security Group that will be used for your production instances or modify default group

Security Group Types



Amazon Elastic Compute Cloud (EC2)

Amazon EC2

- **Resizable** compute capacity
- Complete control of your computing resources
- **Reduced time required** to obtain and boot new server instances

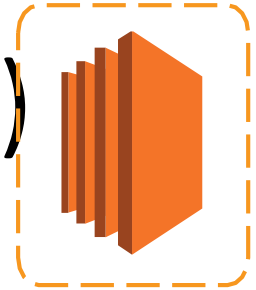
Amazon EC2 Facts

- **Scale capacity** as your computing requirements change
- Pay only for capacity that you actually use
- Choose **Linux** or **Windows**
- Deploy across **AWS Regions** and **Availability Zones** for reliability
- Use **tags** to help manage your Amazon EC2 resources

Launching an Amazon EC2 Instance via the Management Console

1. **Determine the AWS Region** in which you want to launch the Amazon EC2 instance.
2. **Launch** an Amazon EC2 instance from a pre-configured Amazon Machine Image (AMI).
3. **Choose an instance type** based on CPU, memory, storage, and network requirements.
4. **Configure** network, IP address, security groups, storage volume, tags, and key pair.

Amazon Machine Image (AMI) Details



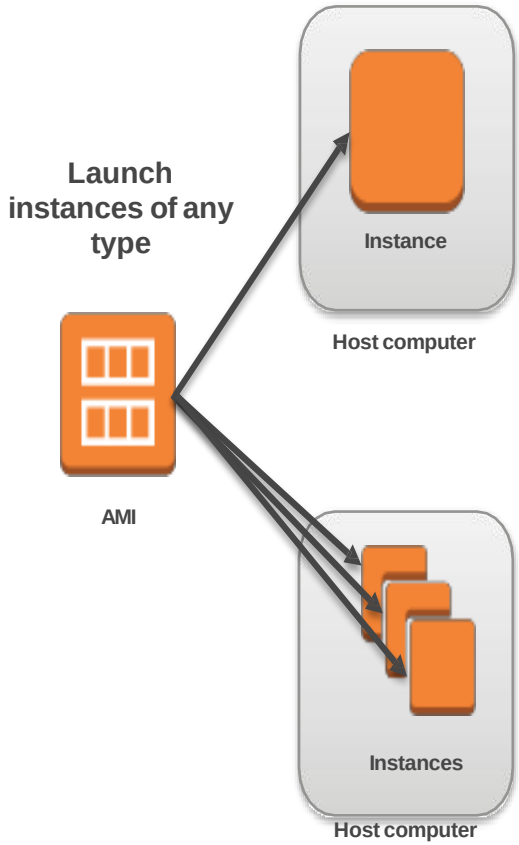
An AMI includes the following:

- A template for the **root volume** for the instance (for example, an operating system, an application server, and applications).
- **Launch permissions** that control which AWS accounts can use the AMI to launch instances.
- A block device mapping that specifies the **volumes to attach** to the instance when it is launched.

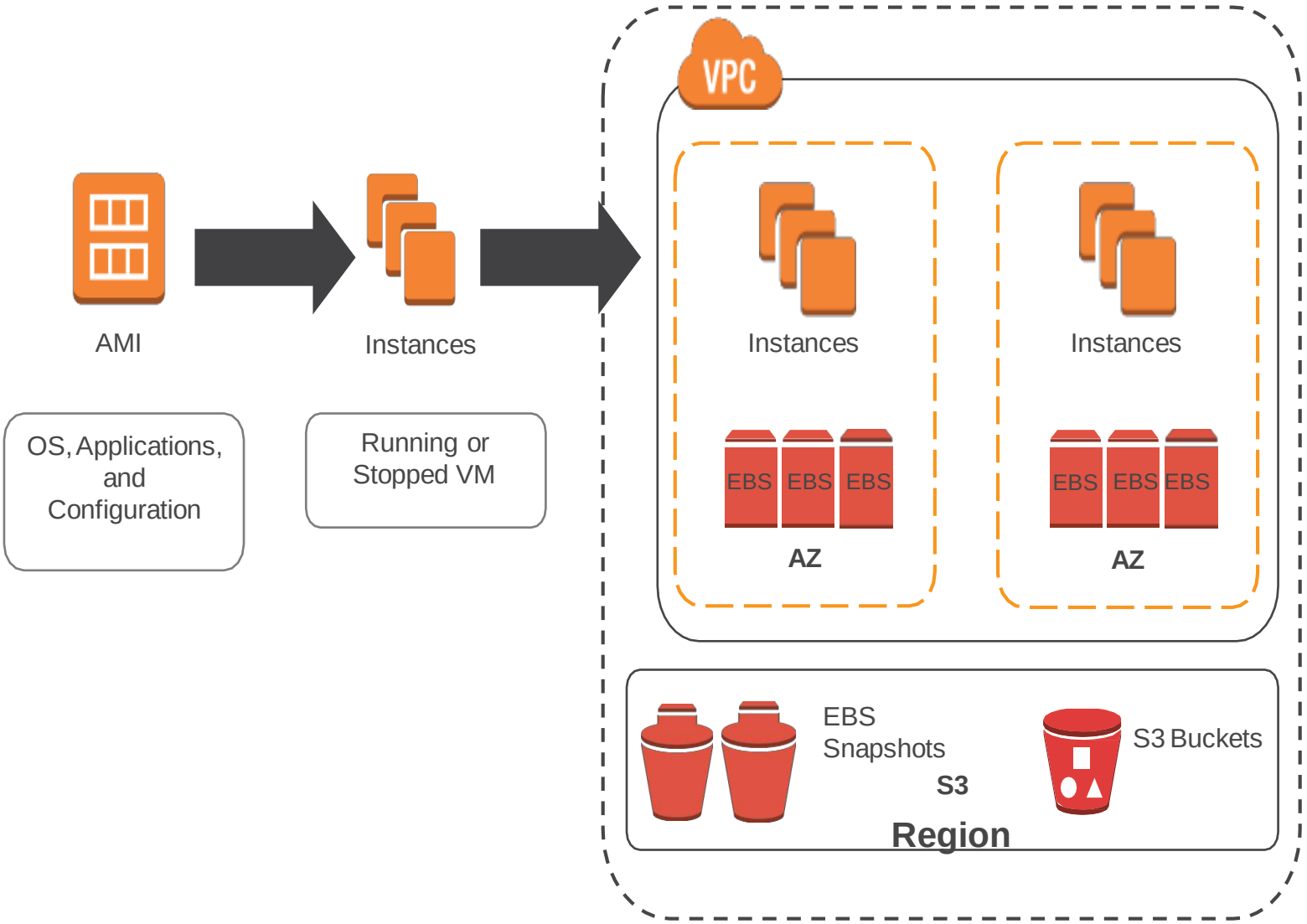
Instances and AMIs

Select an AMI based on:

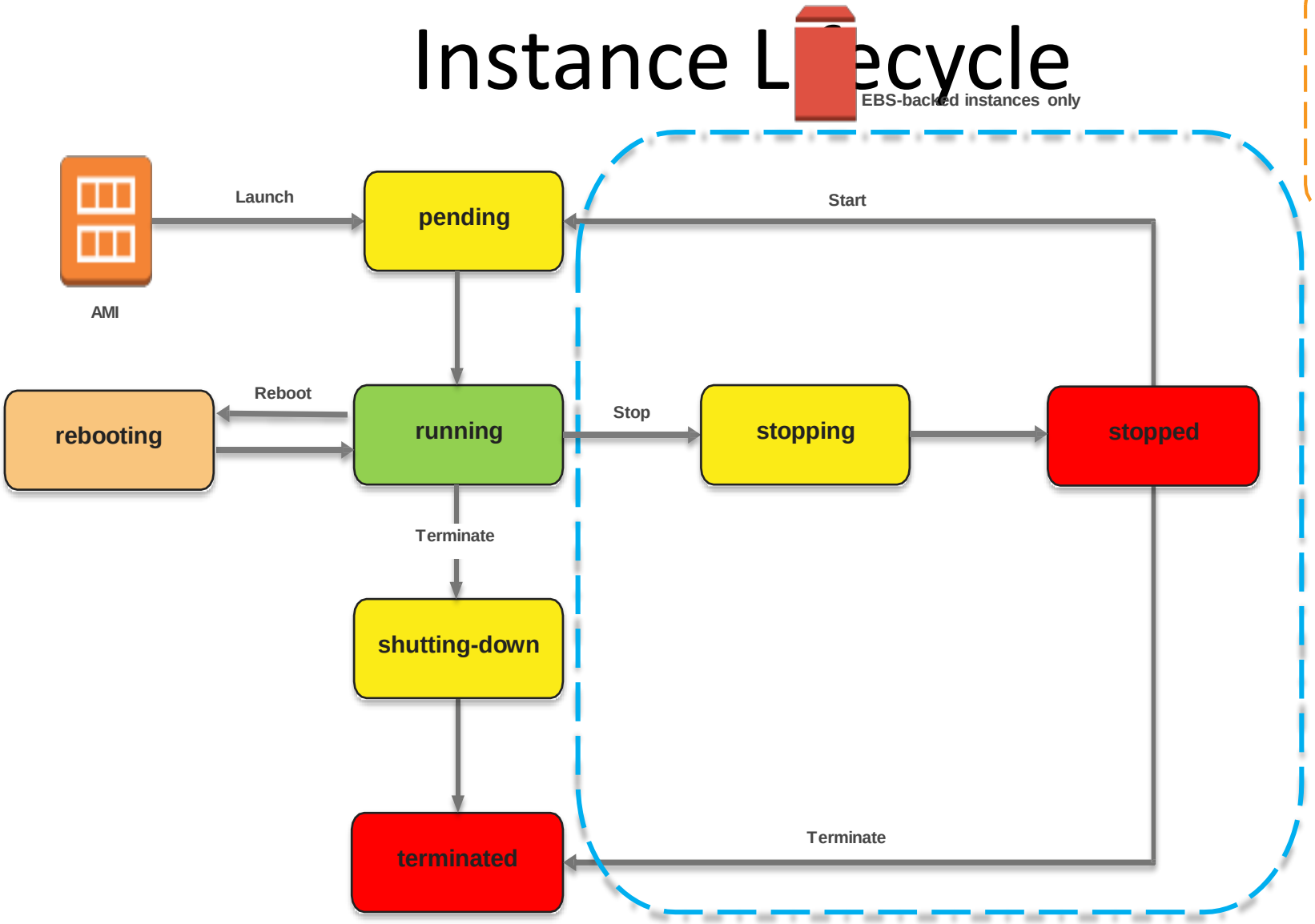
- Region
- Operating system
- Architecture (32-bit or 64-bit)
- Launch permissions
- Storage for the root device



Amazon EC2 Instances

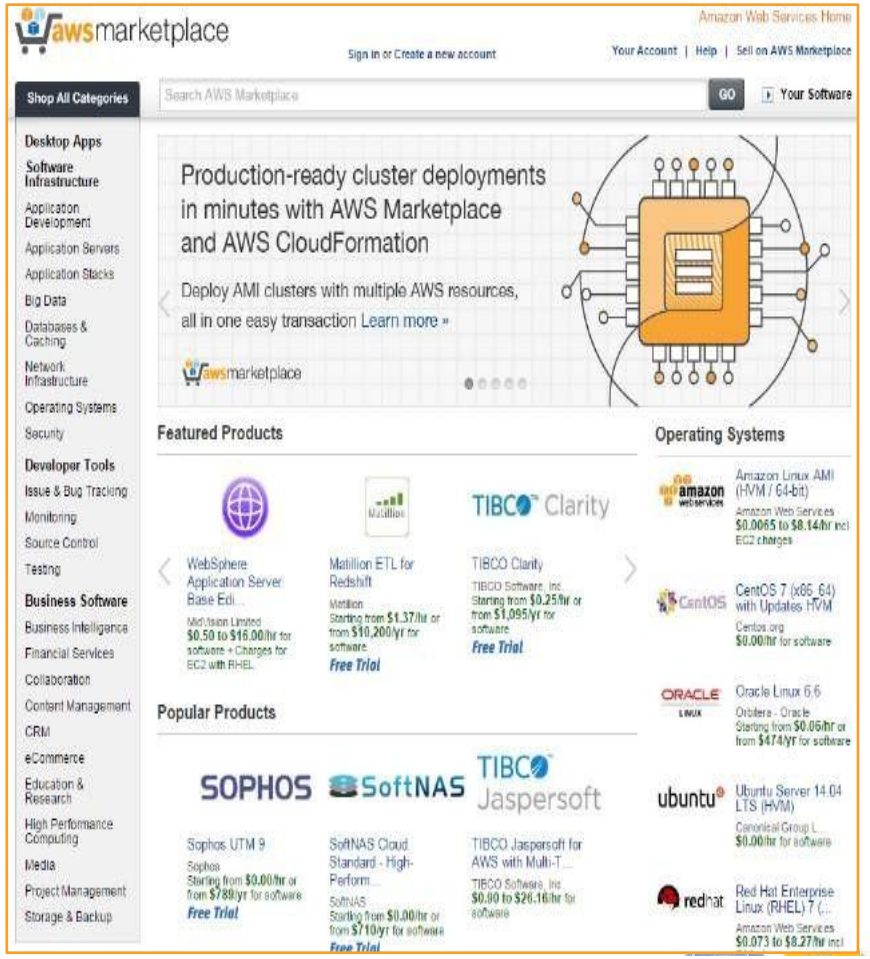


Instance Lifecycle



AWS Marketplace – IT Software Optimized for the Cloud

- Online store to discover, purchase, and deploy IT software on top of the AWS infrastructure.
- Catalog of **2700+** IT software solutions including Paid, BYOL, Open Source, SaaS, and free-to-try options.
- Pre-configured to operate on AWS.
- Software checked by AWS for security and operability.
- Deploys to AWS environment in minutes.
- Flexible, usage-based billing models.
- Software charges billed to AWS account.



Includes [AWS Test Drive](#).

https://aws.amazon.com/marketplace

Choosing the Right Amazon EC2 Instance

AWS uses Intel® Xeon® processors to provide customers with high performance and value. EC2 instance types are optimized for different use cases, workload requirements and come in multiple sizes.

Consider the following when choosing your instances:

- Core count
- Memory size
- Storage size and type
- Network performance
- CPU technologies

Get the Intel® Advantage

Intel's Haswell microarchitecture on new X1, C4, D2, and M4 instances, with custom Intel® Xeon® v3 processors, provides new features:

Haswell microarchitecture can boost existing applications performance by **30% or more** for better workload performance and faster response times.

Newer Hardware Assisted technologies, such as Intel® AVX2.0 instructions, can double the floating-point performance for compute-intensive workloads and provide additional instructions for compression and encryption

X1 Instance - Tons of Memory

The X1 instance:

- Features up to 2TB of memory and 100 vCPU.
- Uses Intel E7 v3 Haswell processors.
- Is designed for demanding enterprise workloads, including production installations of SAP HANA, Microsoft SQL Server, Apache Spark, and Presto.

Intel® Processor Technologies

Intel® AVX: Provides dramatically better performance for highly parallel HPC workloads such as *life science engineering, data mining, financial analysis*, or other technical computing applications. AVX also enhances *image, video, and audio* processing.

Intel® AES-NI: Enhance your security with these new encryption instructions that reduce the performance penalty associated with encrypting/decrypting data.

Intel® Turbo Boost Technology: Provides more computing power when you need it with performance that adapts to spikes in your workload.

Intel Transactional Synchronization (TSX) Extensions: Enable execution of transactions that are independent to accelerate throughput.

P state & C state control: Gives you the ability to individually tune each cores performance & sleep states to improve application performance.

AWS EC2 Instances with Intel® Technologies

AWS Instance Type	High Memory X1	Compute-Optimized C4	Storage-Optimized D2	General Purpose M4	Memory-Optimized R3	IO-Optimized I2	Graphics-Optimized G2	Burstable Performance T2
Intel Processor	Intel Xeon E7-8880 v3	Custom Intel Xeon E5-2666 v3	Custom Intel Xeon E5-2676 v3	Custom Intel Xeon E5-2676 v3	Intel Xeon E5-2670 v2	Intel Xeon E5-2670 v2	Intel Xeon E5-2670	Intel Xeon Family
Intel AVX	AVX 2.0	AVX 2.0	AVX 2.0	AVX 2.0	Yes	Yes	Yes	Yes
Intel AES-NI	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Intel Turbo Boost	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intel TSX	Yes	No	No	No	No	No	No	No
Per core P- and C-state control	No	Yes (8xlarge only)	No	No	No	No	No	No
SSD Storage	EBS Optimized by	EBS Optimized by	No	EBS Optimized by	Yes	Yes	Yes	EBS only

Current Generation Instances

Instance Family	Some Use Cases
General purpose (t2, m4, m3)	• Low-traffic websites and web applications • Small databases and mid-size databases
Compute-optimized (c4, c3)	• High performance front-end fleets • Video-encoding
Memory-optimized (r3)	• High performance databases • Distributed memory caches
Storage-optimized (i2, d2)	• Data warehousing • Log or data-processing applications
GPU instances (g2)	• 3D application streaming • Machine learning

Instance Metadata

- Is **data** about your **instance**.
- Can be used to **configure or manage** a running instance.

Retrieving Instance Metadata

To view all categories of instance metadata from within a running instance, use the following URI:
http://169.254.169.254/latest/meta-data/

On a Linux instance, you can use:

- `$ curl http://169.254.169.254/latest/meta-data/`
- `$ GET http://169.254.169.254/latest/meta-data/`

All metadata is returned as text (content type text/plain).

Instance User Data

- Can be passed to the instance **at launch**.
- Can be used to perform common **automated configuration tasks**.
- Runs scripts after the instance starts.

Adding User Data

- You can specify user data when launching an instance.
- User data can be:
  - Linux script – executed by **cloud-init**
  - Windows batch or PowerShell scripts – executed by **EC2Config** service
- User data scripts run once per instance ID by default.

User Data Example Linux

User data shell scripts must start with the #! characters and the path to the interpreter you want to read the script.

```
#!/bin/sh
yum -y install httpd

chkconfig httpd on
/etc/init.d/httpd start
```

Install Apache web server
Enable the web server
Start the web server

User Data Example Windows

```
<powershell>
Import-Module ServerManager

Install-WindowsFeature web-server, web-webserver
Install-WindowsFeature web-mgmt-tools

</powershell>
```

Import the Server Manager module for Windows PowerShell.

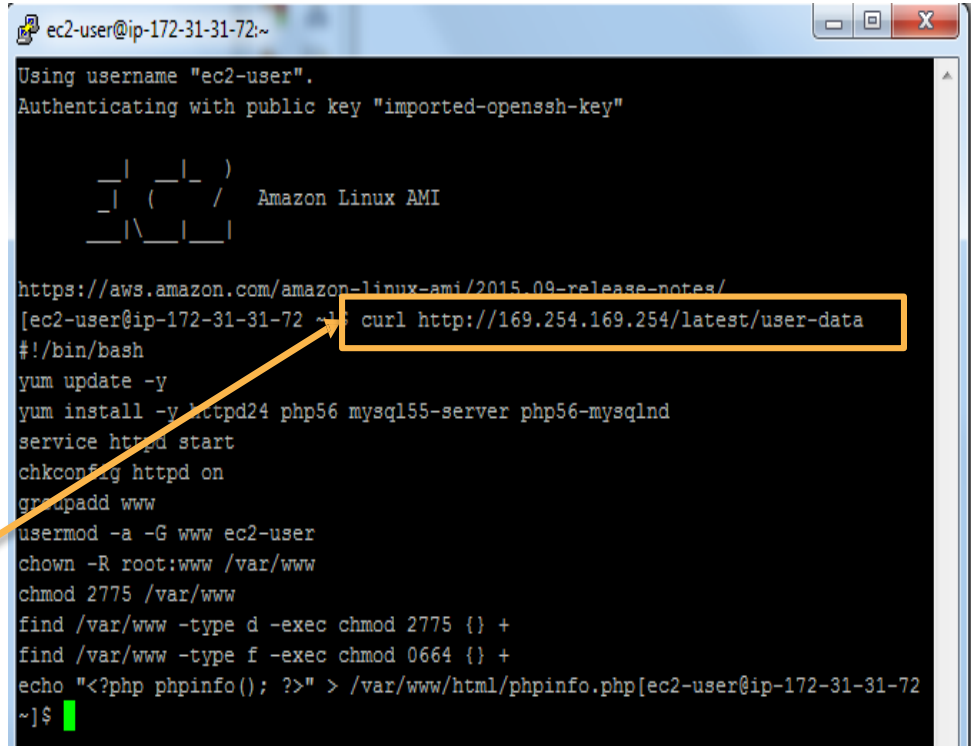
Install IIS
Install Web Management Tools

Retrieving User Data

To retrieve user data, use the following URI: http://169.254.169.254/latest/user-data

On a Linux instance, you can use:

```
$ curl http://169.254.169.254/latest/user-data/
$ GET http://169.254.169.254/latest/user-data/
```



Amazon EC2 Purchasing Options

On-Demand Instances

Pay by the **hour**.

Reserved Instances

Purchase, at a significant **discount**, instances that are **always available**

1-year to 3-year terms.

Scheduled Instances

Purchase instances that are **always available** on the specified **recurring schedule**, for a one-year term.

Spot Instances

Bid on **unused instances**, which can run as long as they are available and your bid is above the Spot price.

Dedicated Instances

Pay, by the hour, for instances that run on **single-tenant hardware**.

Dedicated Hosts

Pay for a physical host that is **fully dedicated** to running your instances.