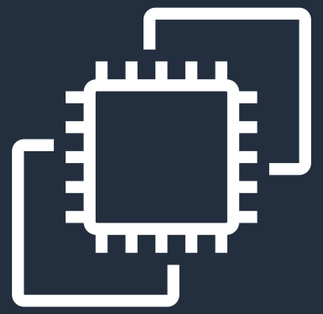
Elastic Compute Cloud (EC2) Overview

Scott Bateman

Nov 17, 2020

Choices for Compute



Amazon EC2

Virtual server instances
in the cloud

Amazon ECS,
EKS, and Fargate

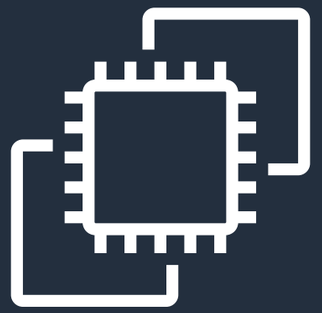
Container management service
for running
Docker on a managed
cluster of EC2

AWS Lambda

Serverless compute
for stateless code execution in
response to triggers

aws

Amazon EC2



Amazon EC2

Linux | Windows

Arm and x86 architectures

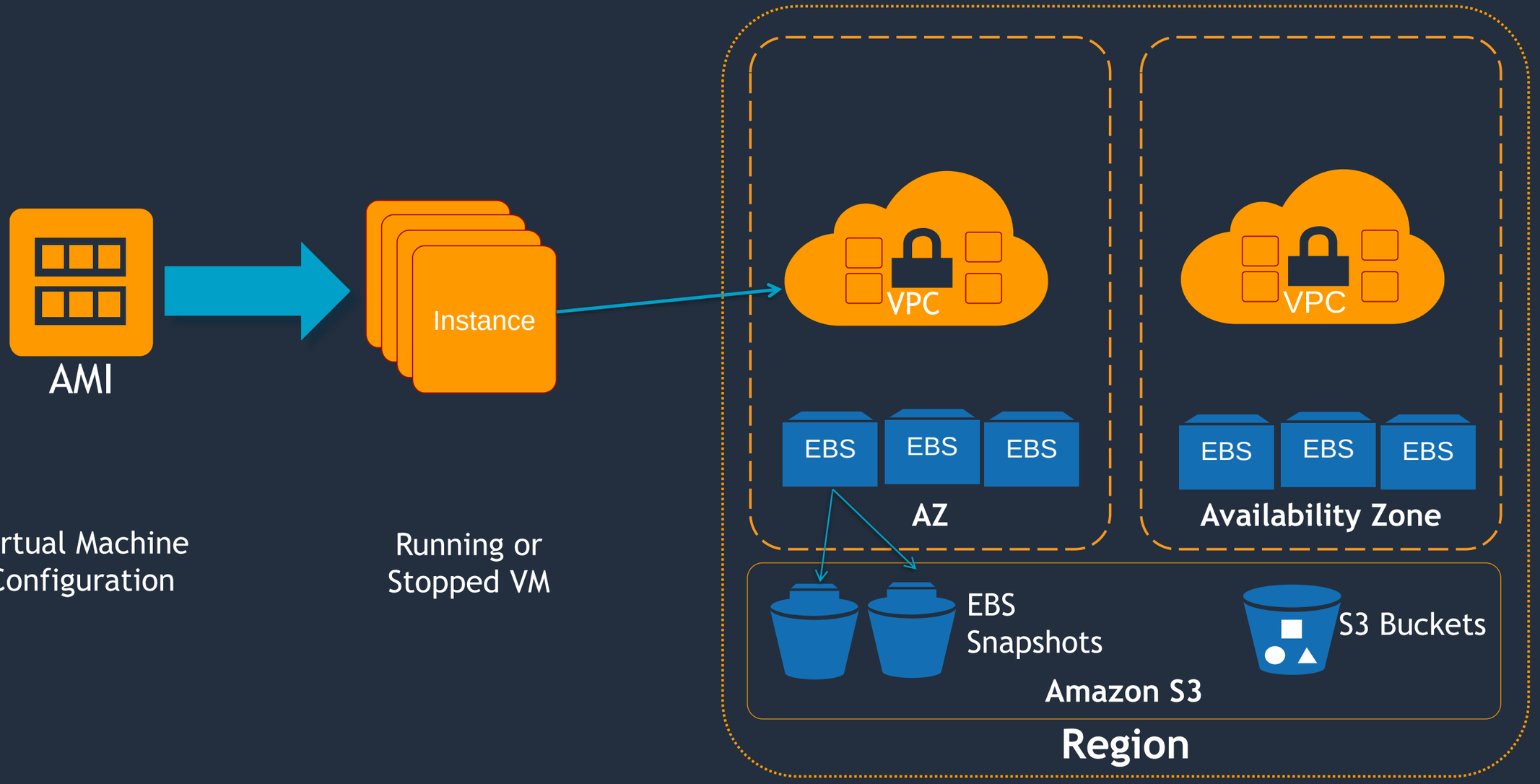
General purpose and workload optimized

Bare metal, disk, networking capabilities

Packaged | Custom | Community AMIs

Multiple purchase options: On-demand, RI, Spot

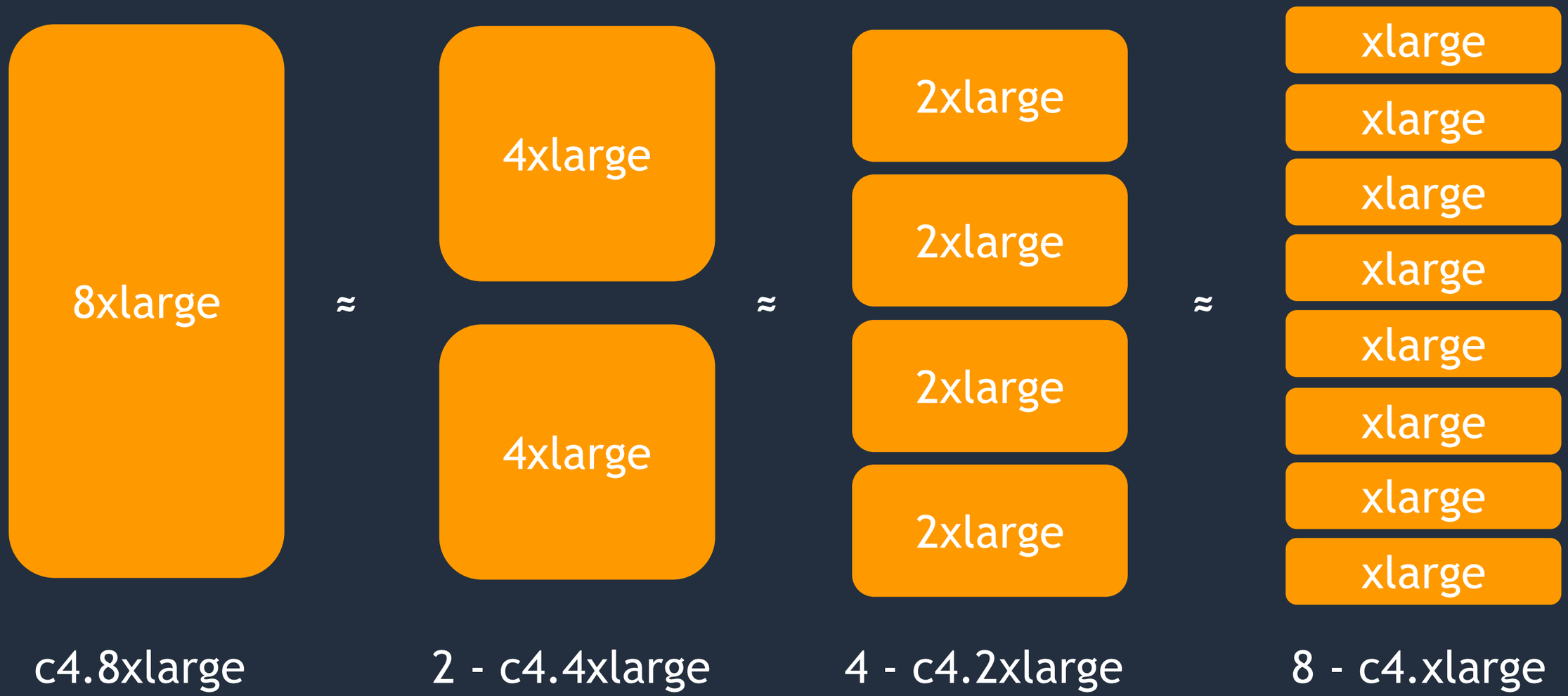
EC2 Terminology



Virtual Machine Configuration

Running or Stopped VM

Instance sizing



c4.8xlarge | 2 - c4.4xlarge | 4 - c4.2xlarge | 8 - c4.xlarge

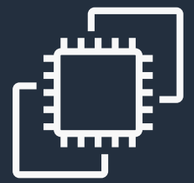
Choose your processor and architecture

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

aws

EC2 Naming Explained

Instance generation

c5n.xlarge

Instance family

Attribute

Instance size

aws

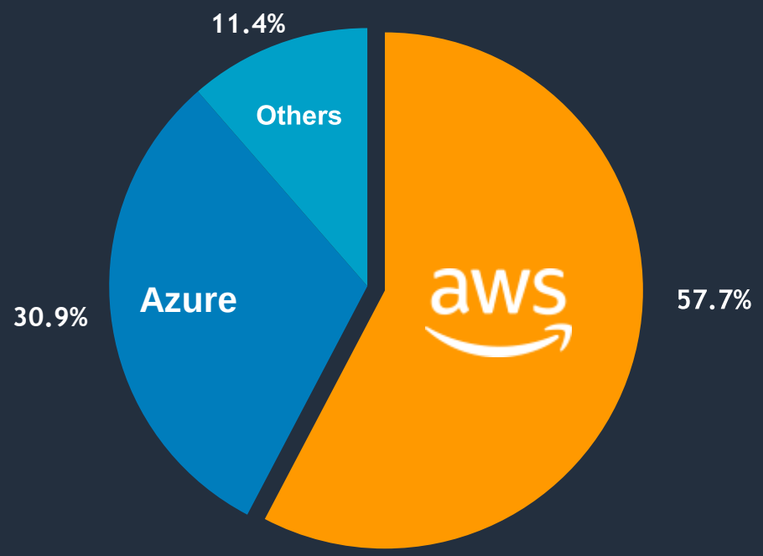
Instance Types

	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
intel	T3	M5	C5	C5n	R5	X1	X1e		G3	P3	F1		D2	H1
Local storage (NVMe SSD)		M5d	C5d		R5d			Z1d				I3		
AMD	T3a	M5a			R5a									
metal		M5m	c5m		R5m		u-12tb1	Z1dm				I3m		
others	A1	M6g	C6g		R6g					P3dn		I3en		

arm

aws

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

aws

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

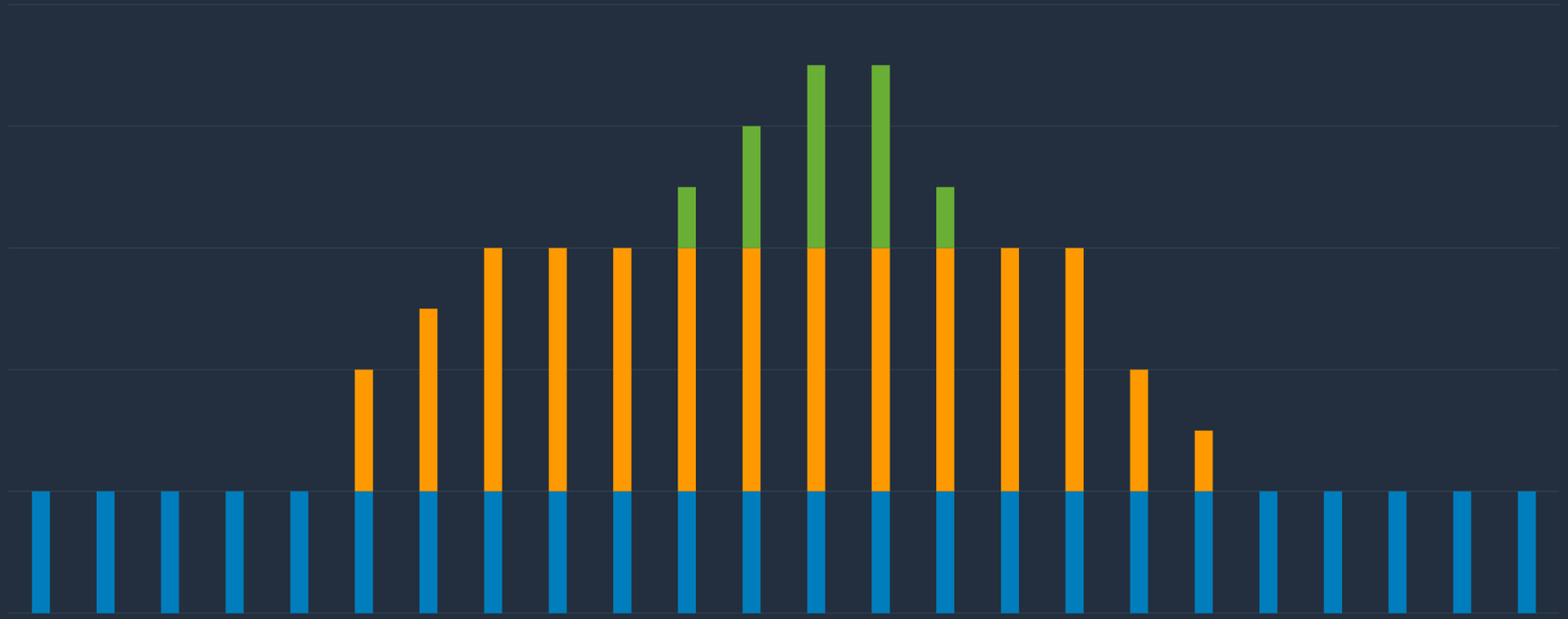
Spot Instances

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

Fault-tolerant, flexible, stateless workloads

aws

Simplify capacity and cost optimization



Now with hibernate

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

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

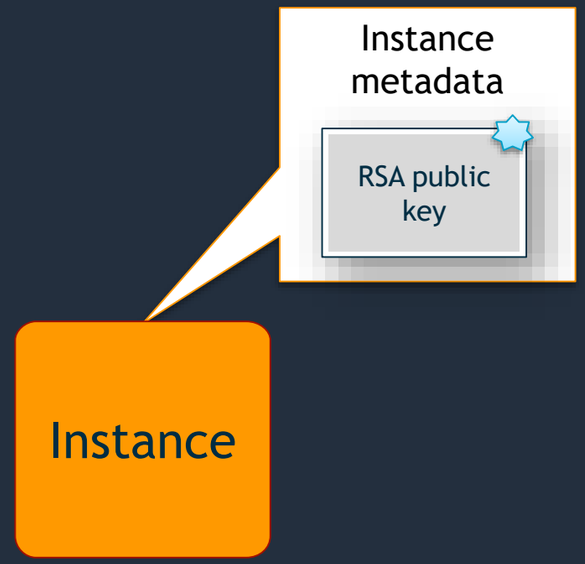
"Private Half" downloaded to your desktop

aws

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**



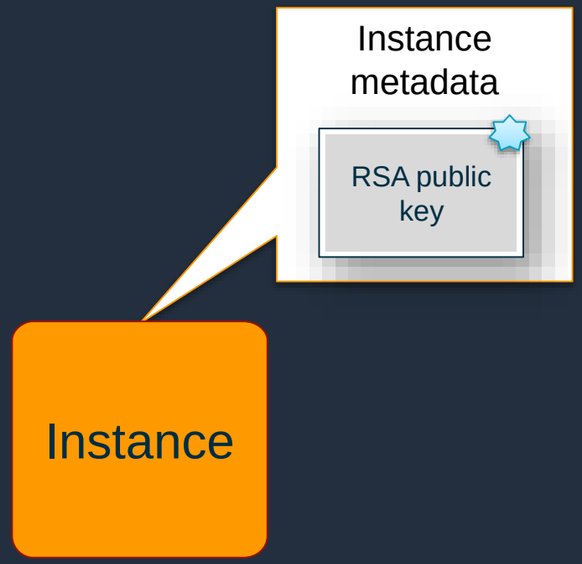
EC2 Instance access and Key Pairs

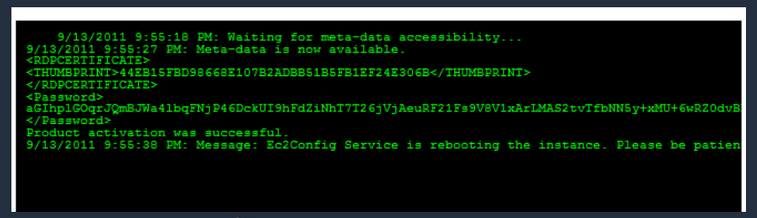
Linux launch (first boot)

- **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**





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

Any Questions?

aws