AWS Compute - EC2

Thomas Le Moullec. AWS Solutions Architect
September 29th 2020

Agenda

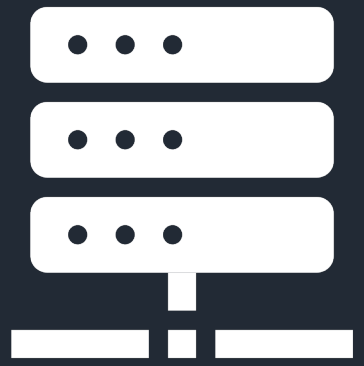
- What is Compute?
- Compute offering
- Amazon EC2 (Virtualization, Resource allocation, Configuration)
- Purchase Plans
- Capacity and Provisioning
- Network and Security

What is Compute ?

What is Compute ?

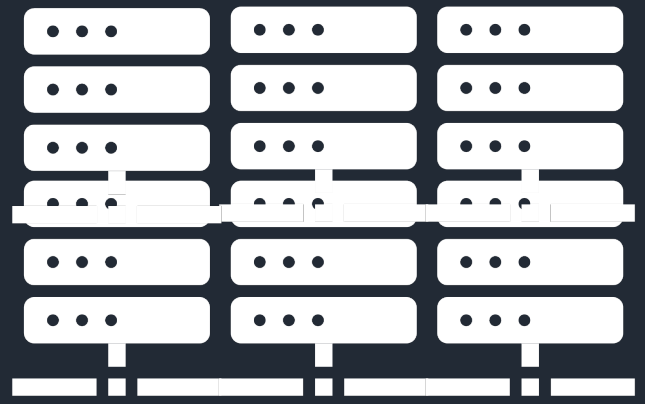
- CPU, RAM , OS, Hard Drive, Network Card, Firewall

What is Compute ?



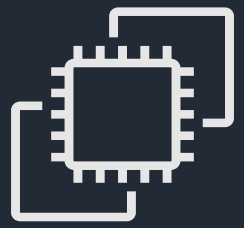
Large “Instance”

- CPU, RAM , OS, Hard Drive, Network Card, Firewall



x100

Compute offering



Amazon EC2

Virtual server hosting

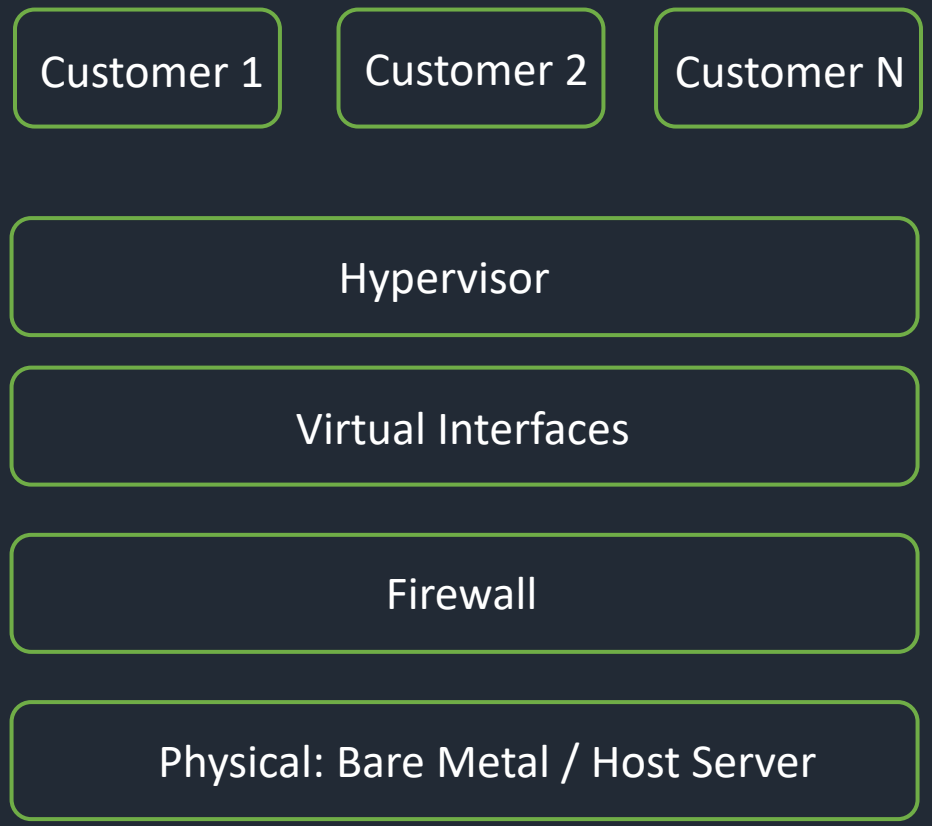
Amazon ECS, EKS, Fargate

Container management

AWS Lambda

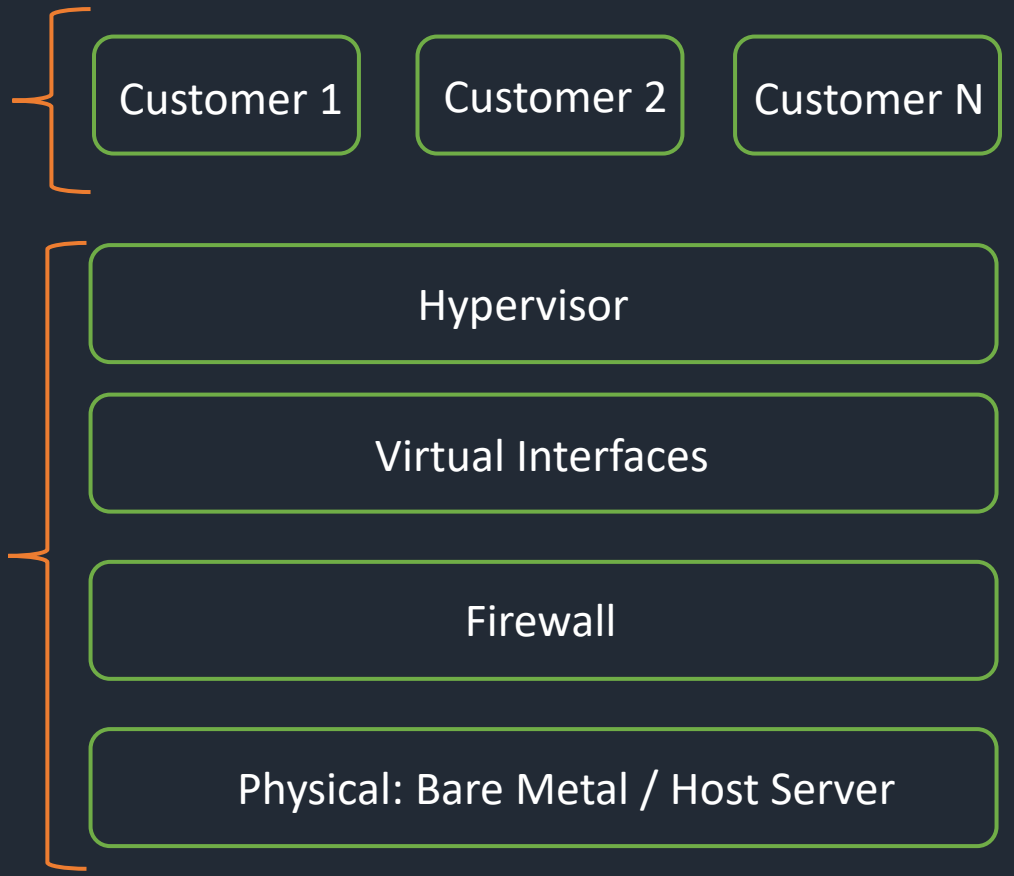
Serverless computing

Virtualization



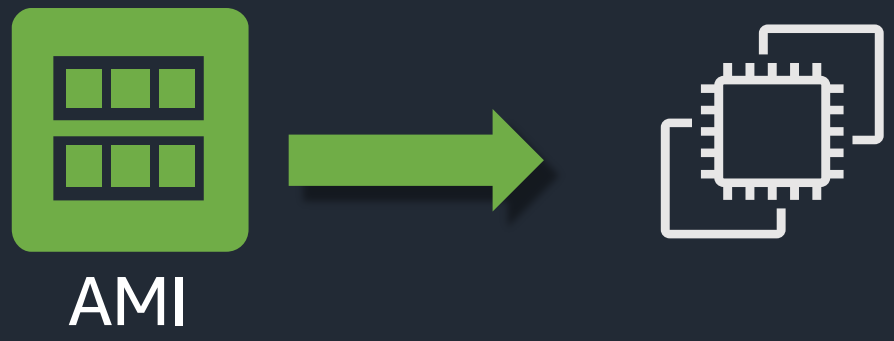
Virtualization

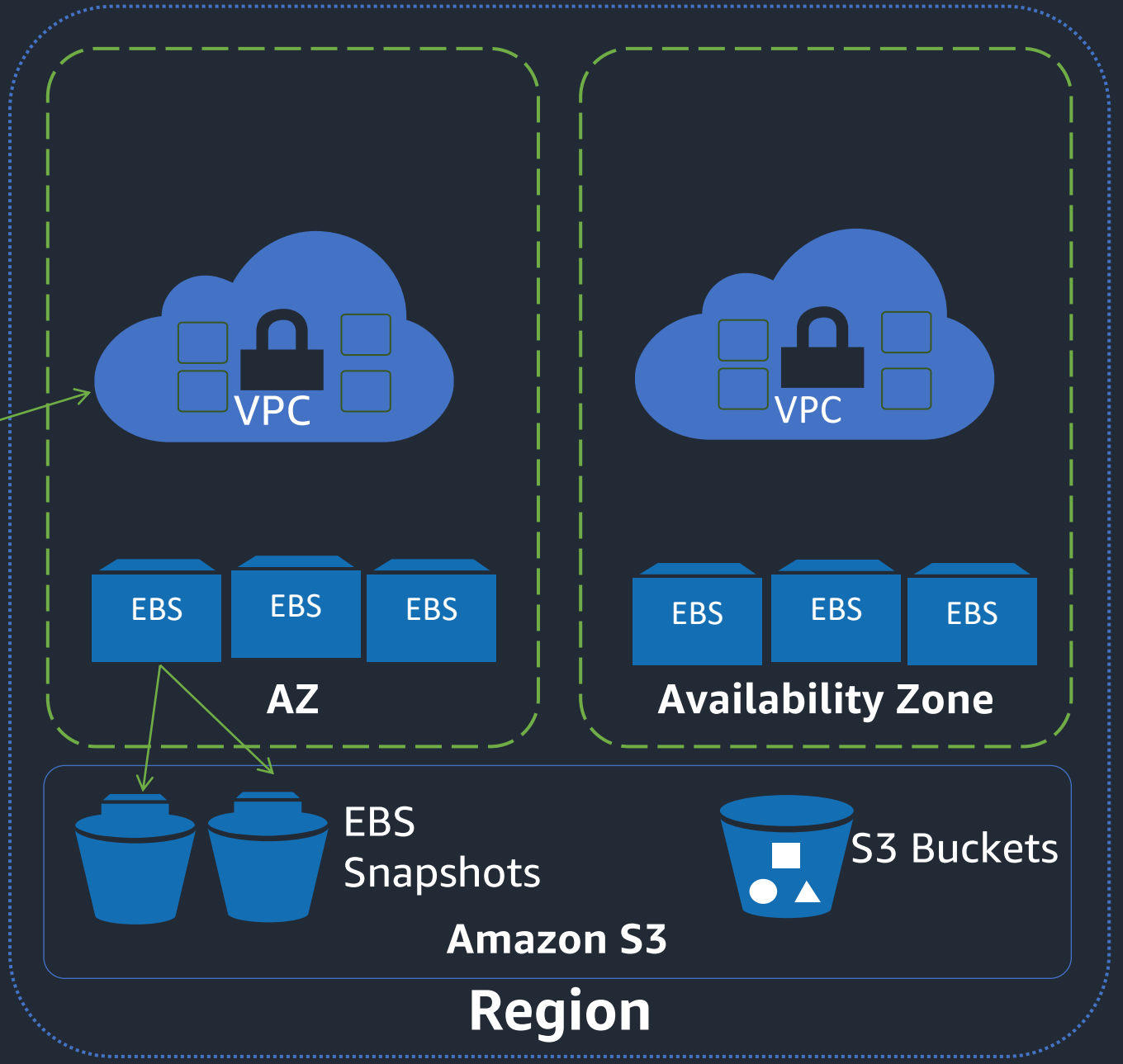
Customer Responsibility:
User-Data, Applications and Guest OS running on EC2 Instances



AWS Responsibility:
Servers located in AWS data centers

Amazon EC2





EC2 – Resources Allocation

vCPU

EC2 supports multithreading
Each thread is represented as a virtual CPU (dedicated vCPU)
An instance has a default number of CPU cores with 2 threads per core

Memory

Dedicated memory in GibiBytes (GiB) and not Gigabytes (GB)
1000GB = 1024GiB

Storage

Independent of Compute
Attach virtual disk drives with Elastic Block Storage (EBS)
Max size of one EBS volume is 16TiB

EC2 – Instance Types

Family

A for General Purposed ARM based
T for General Purpose with burst / Turbo
M for General Purpose (Most Scenarios)
C for Compute optimized
D for Dense Storage (48TB Local)
I for Storage I/O (NVMe Local)
H for Storage HDD (high throughput)
HS for High Storage
X for Extra-Large Memory (~4TB DRAM)
U for Ultra-high RAM
R for Random-Access Memory (RAM)
F for Accelerated Computing (FPGA)
G for Accelerated Computing – (GPU)
P for Accelerated Computing – (GPU)

Additional Capabilities

a for AMD CPUs
d for Directly-Attached Instance Storage (NVMe)
e for Extra Capacity (Storage or RAM)
n for Networking Optimized - Bandwidth
g for AWS Graviton CPUs – ARM based

Size

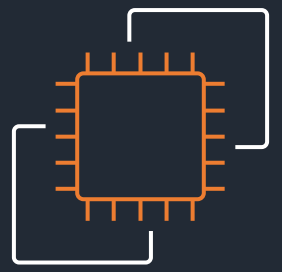
- nano, micro
- small, medium, large, xlarge
- 2xlarge, 3xlarge, 4xlarge, 8xlarge, 9xlarge, 12xlarge , 16xlarge, 18xlarge, 24xlarge, 32xlarge
- metal

c5d.8xlarge

Generation

The higher the newer / better

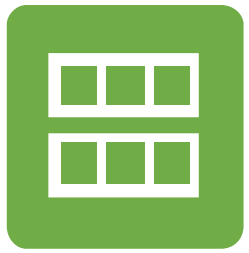
EC2 – Instance Options



Elastic Graphics

Elastic Inference

EC2 – Images and Licenses



AMI

EC2 Image Builder

EC2 – Economy

On Demand

Pay as you go (per second granularity)

EC2 – Economy

On Demand	Pay as you go (per second granularity)
Spot	Spare compute capacity, save up to 90%

EC2 – Economy

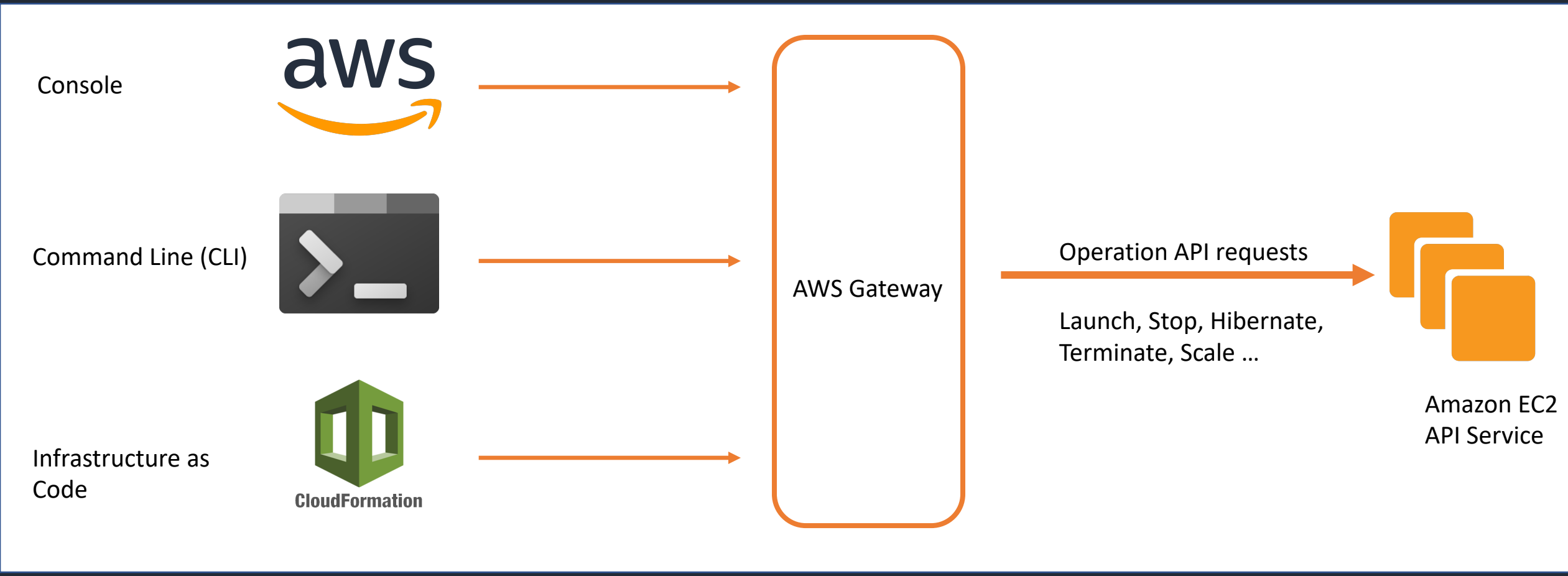
On Demand	Pay as you go (per second granularity)
Spot	Spare compute capacity, save up to 90%
Savings Plans	Up to 72% in exchange for a commitment

EC2 – Economy

On Demand	Pay as you go (per second granularity)
Spot	Spare compute capacity, save up to 90%
Savings Plan	Up to 72% in exchange for a commitment

- Focus budgets on business differentiators, not management of infrastructure.
- Helps reduce the Total Cost of Ownership: Reduce need of Large Capital Expenditures, invest in capacity you need

EC2 – Operations



EC2 – Capacity & Provisioning

AWS
CloudFormation

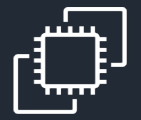
Amazon Elastic
Container Service
for Kubernetes

Amazon Elastic
Container Service

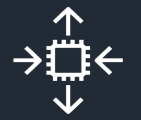
Amazon
EMR

AWS Batch

Amazon
Sagemaker

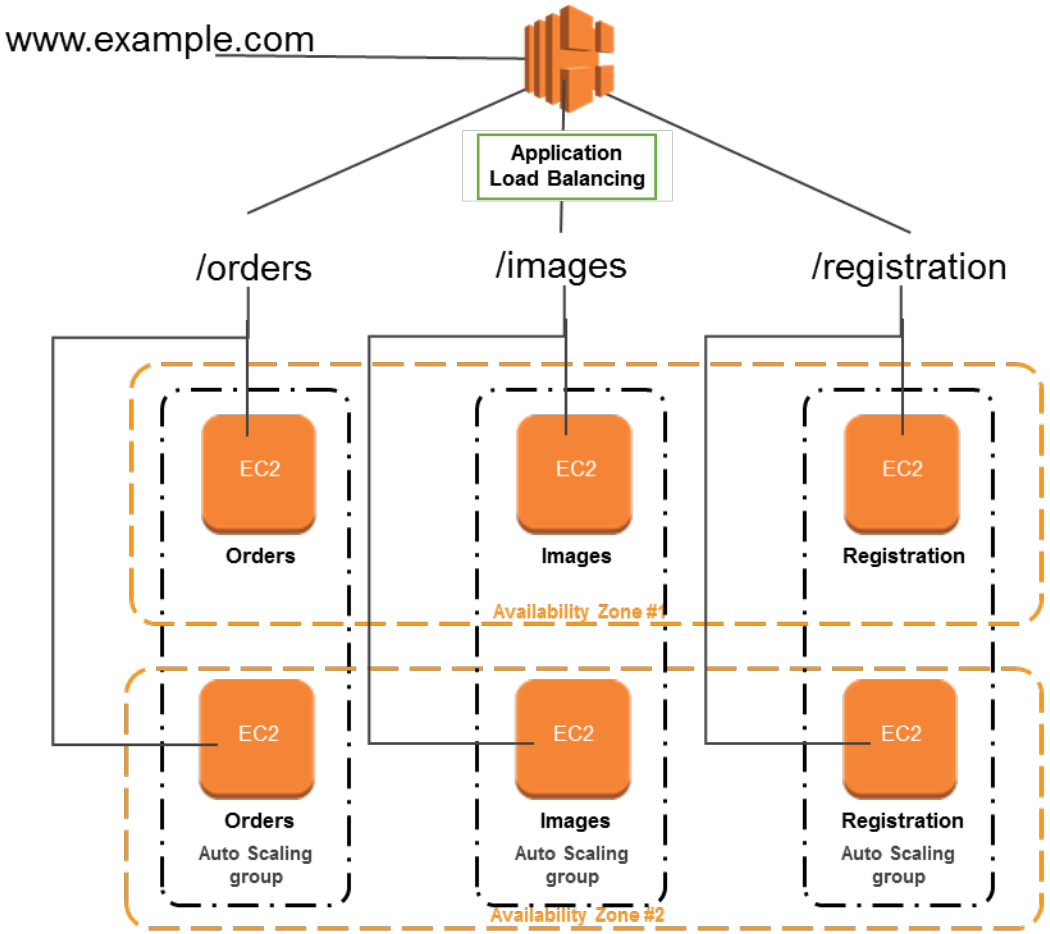


EC2 Fleet

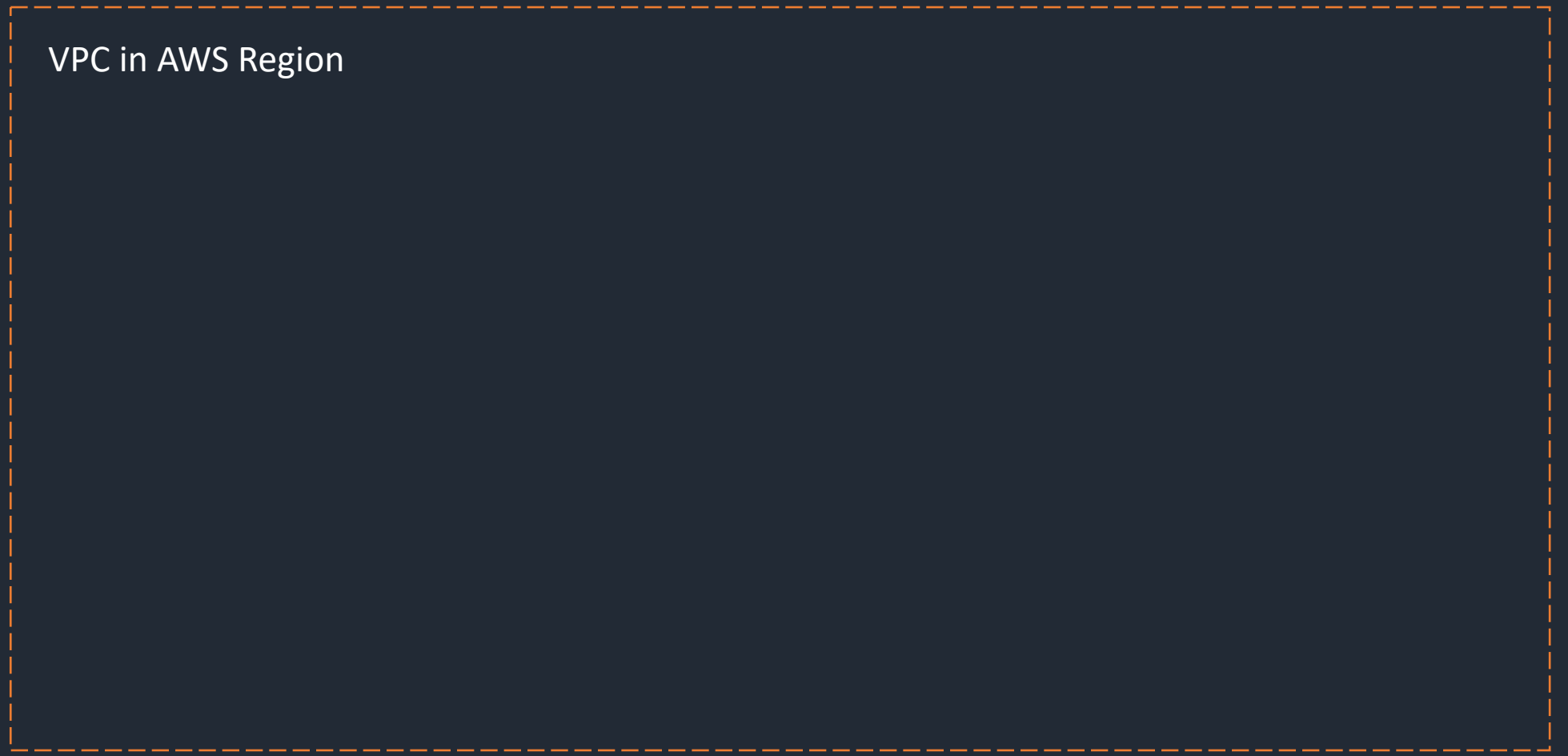


Amazon EC2
Auto Scaling

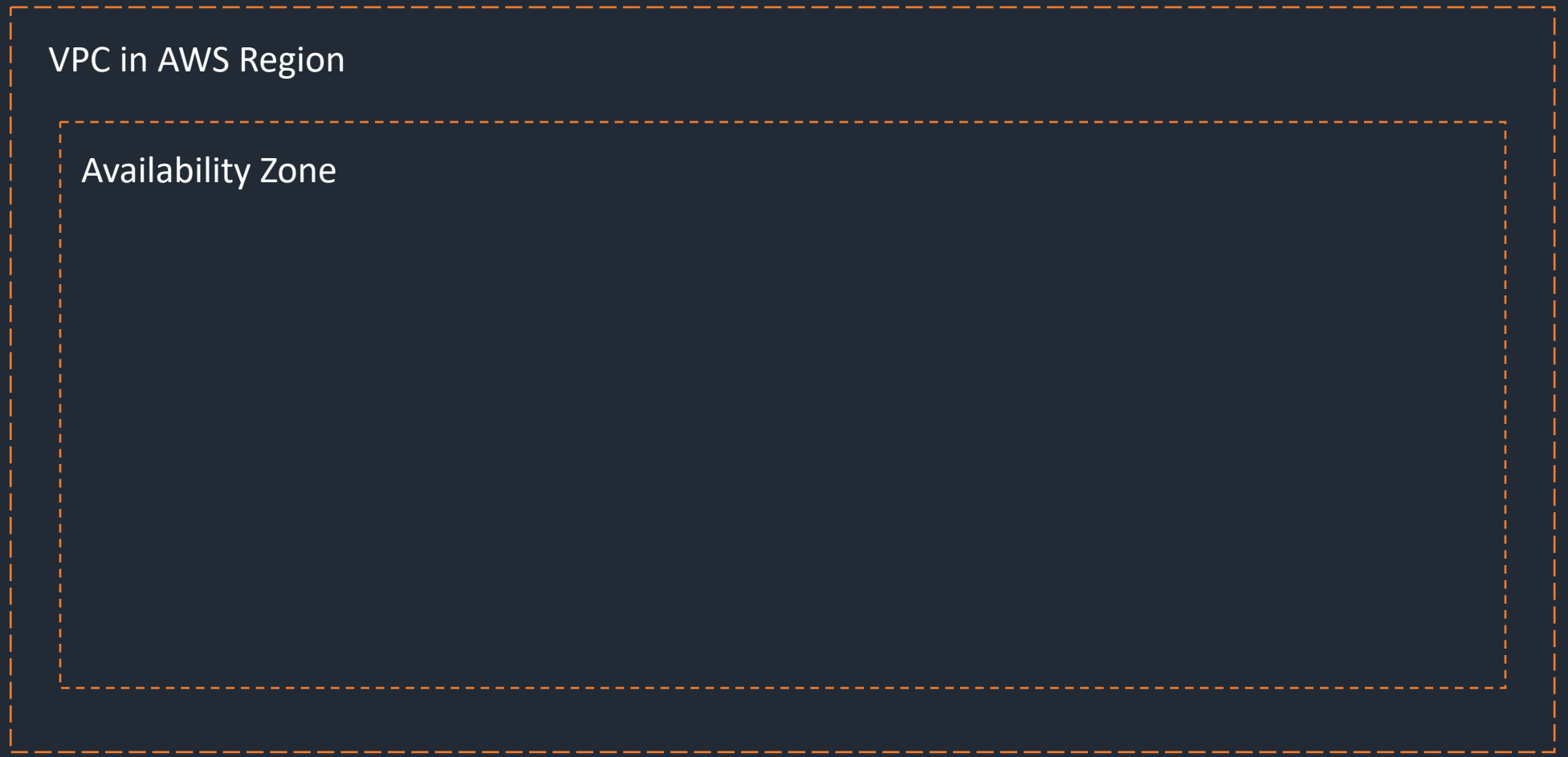
EC2 – Scaling



EC2 – Networking & Security

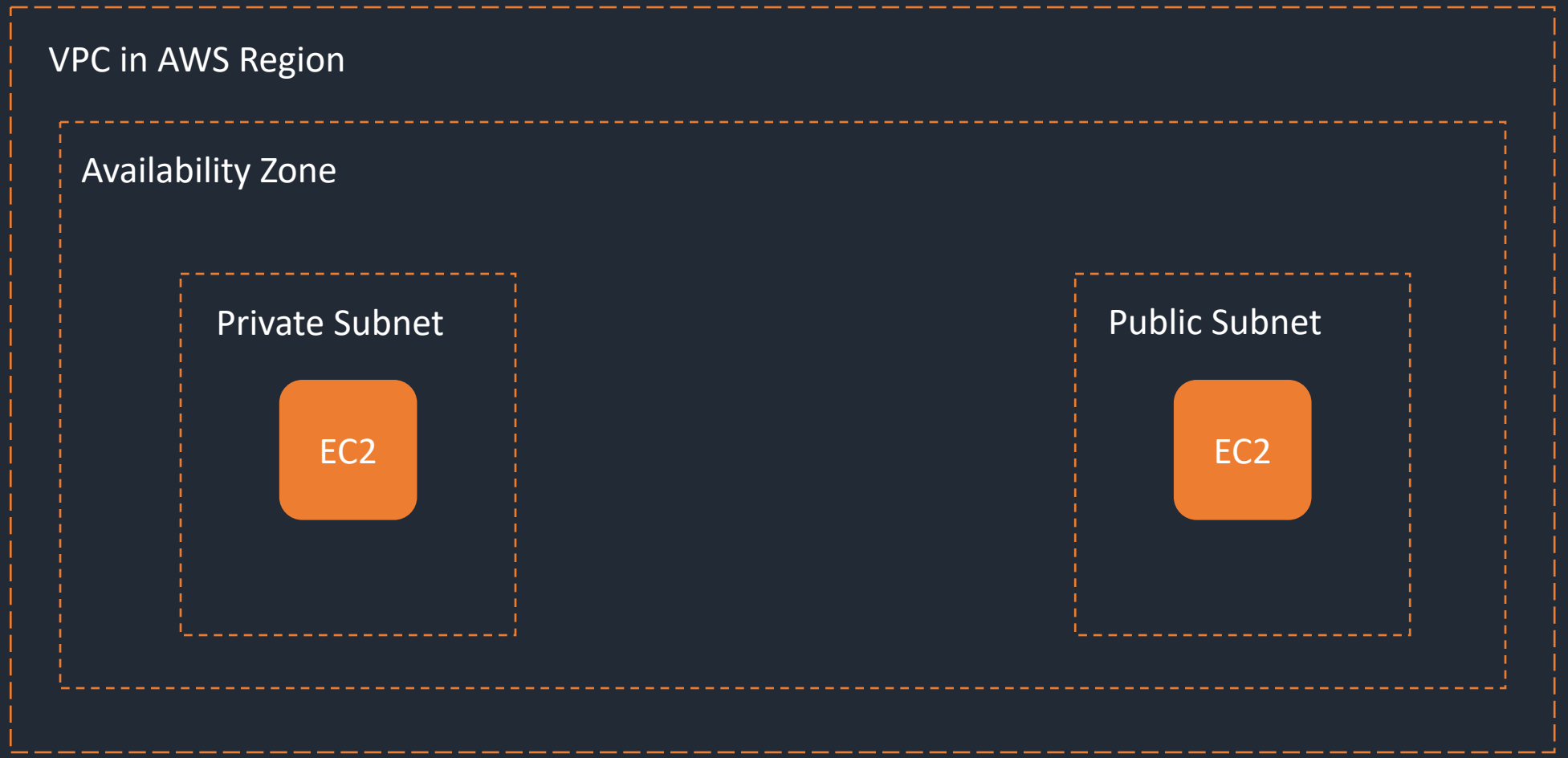


EC2 – Networking & Security

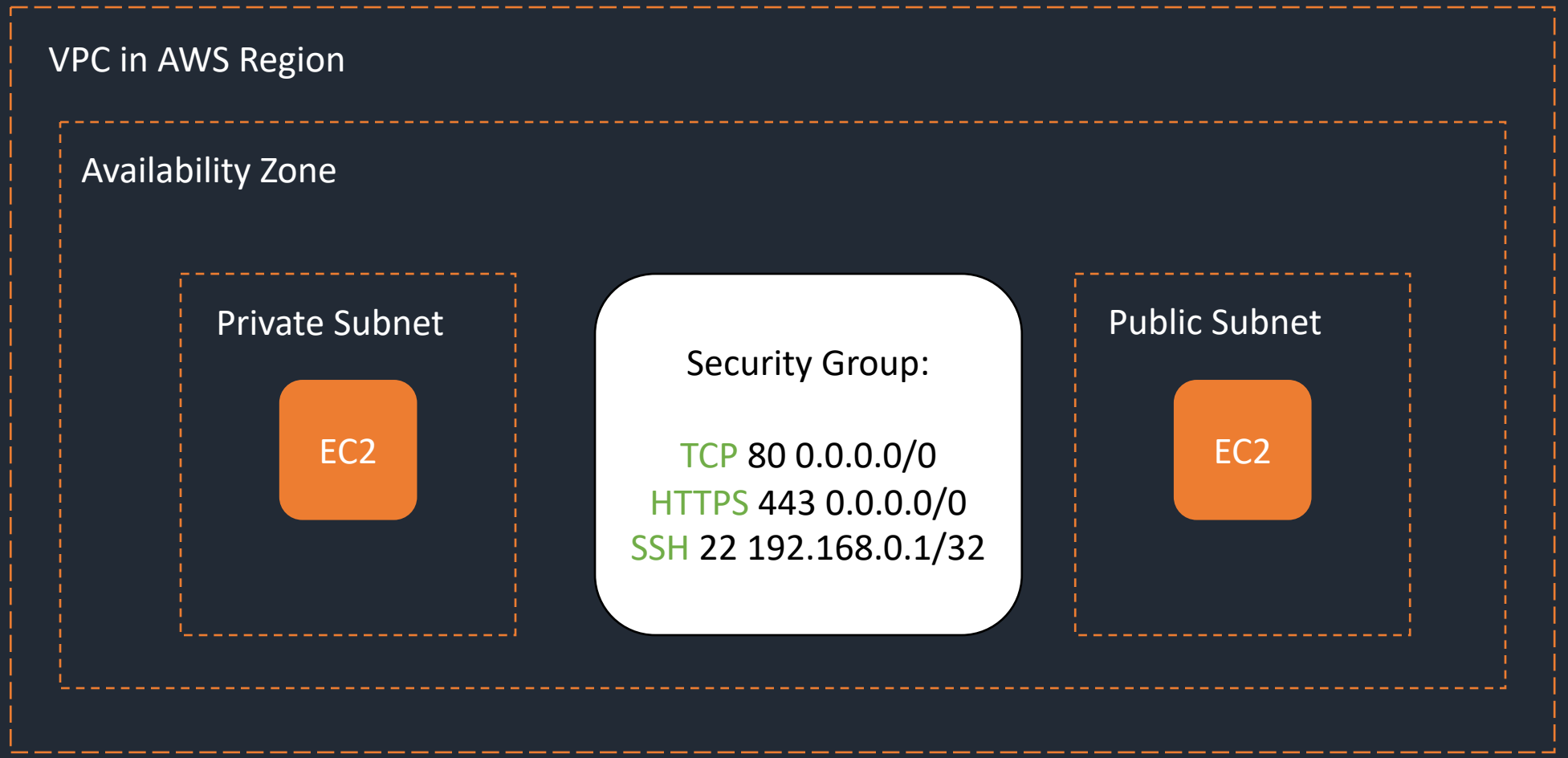


EC2 – Networking & Security

EC2 – Networking & Security



EC2 – Networking & Security



Thank you !

Find me on Linkedin:
Thomas LE MOULLEC