AWSOME DAY
ONLINE CONFERENCE

Module 2 :
Core Services

Amazon Virtual Private Cloud (VPC)

Introduction

Private, virtual network in the AWS Cloud

Similar constructs as on-premises network

Customizable network configurations to meet your needs

Features

Characteristics

- Allows you to provision virtual networks

Logically isolated

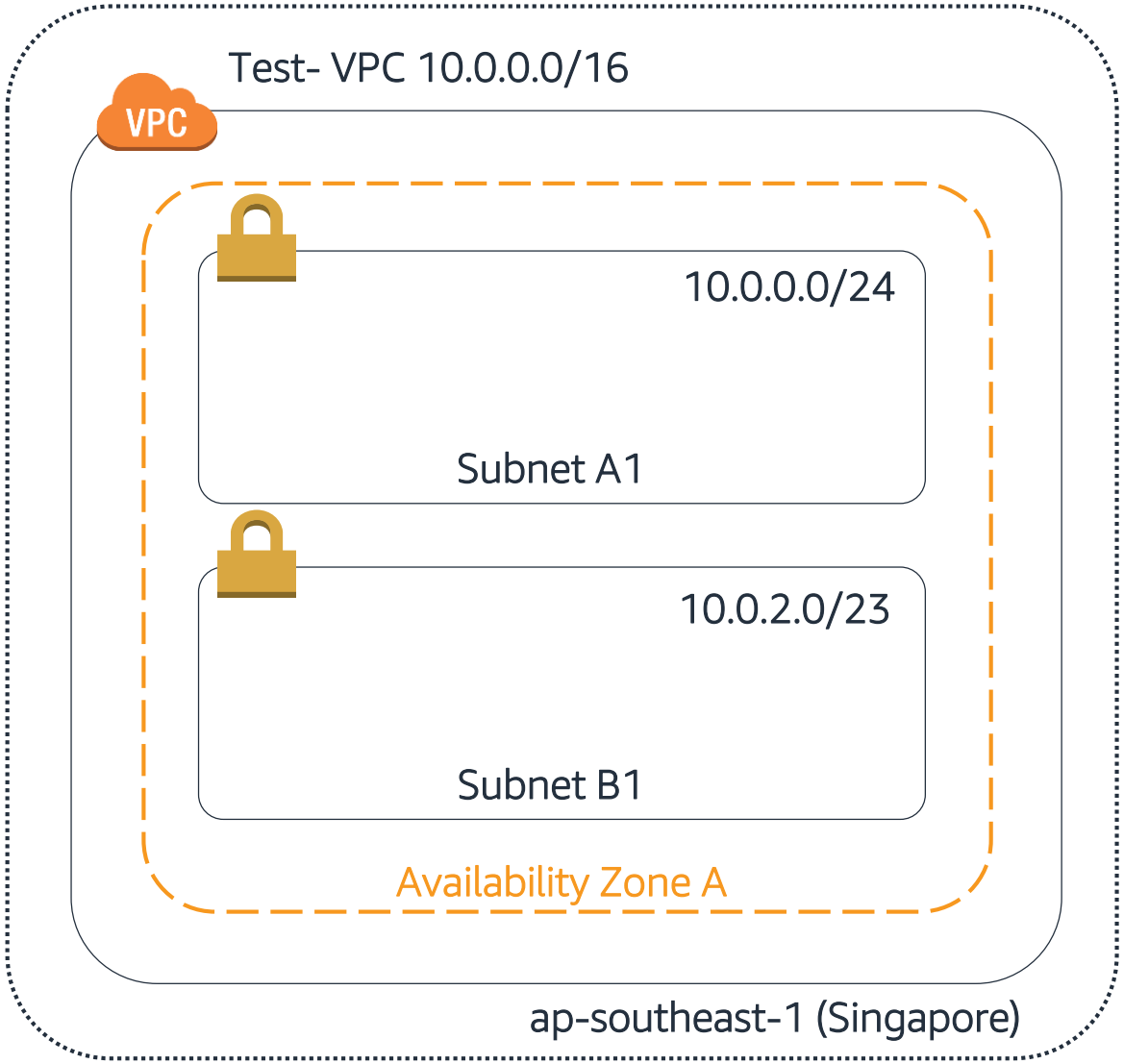
Configurable key features

- IP ranges
- Routing
- Network gateways
- Security settings

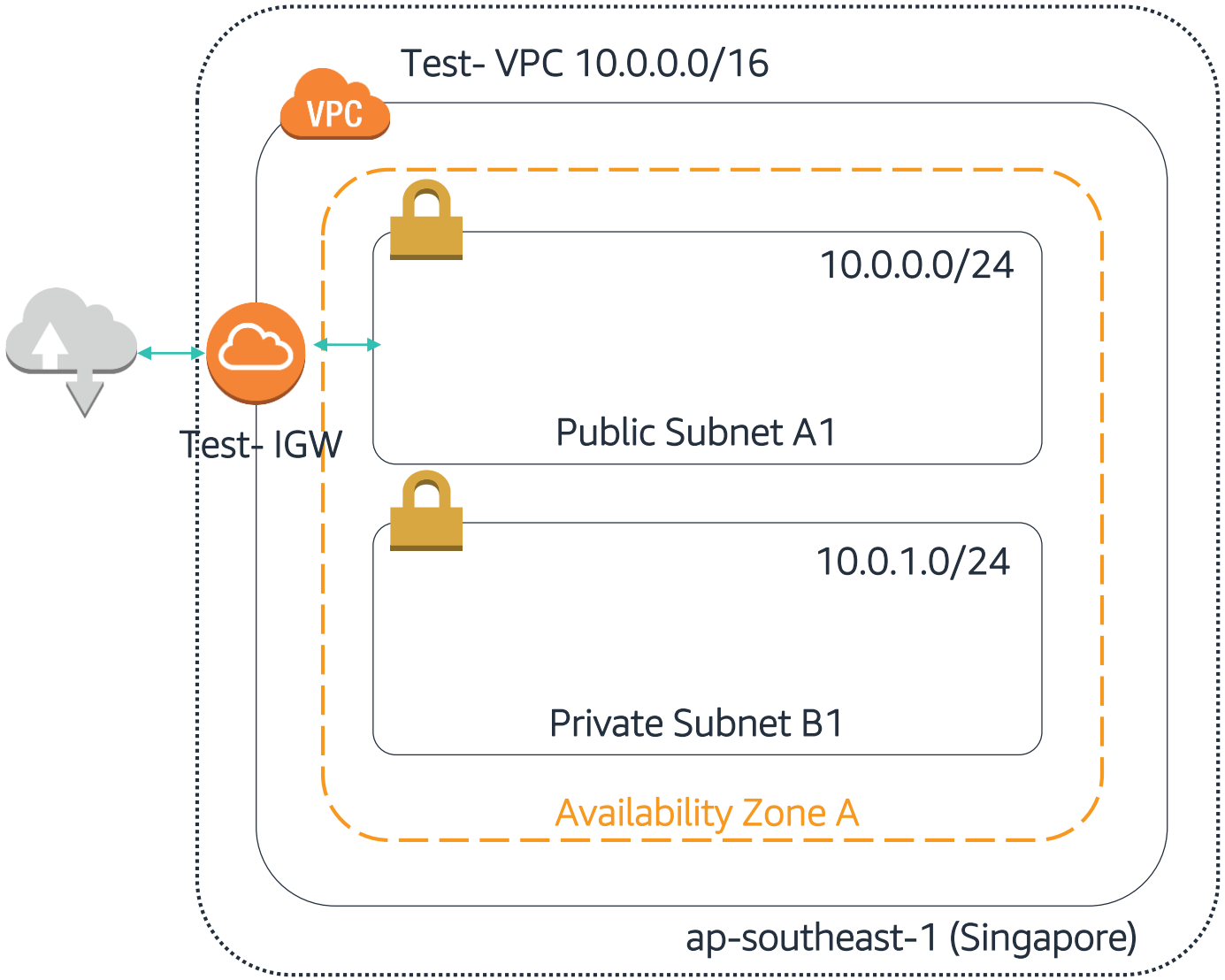
Route Tables

- Control traffic going out of the subnets

Example



Example



Summary

VPC concepts:

- Region Scope
- An internet gateway
- Public subnet
- Private subnet

Learn More

- Route tables and isolation methods
- Other Amazon VPC features (e.g., VPC endpoints and peering connections)

AWS Security Groups

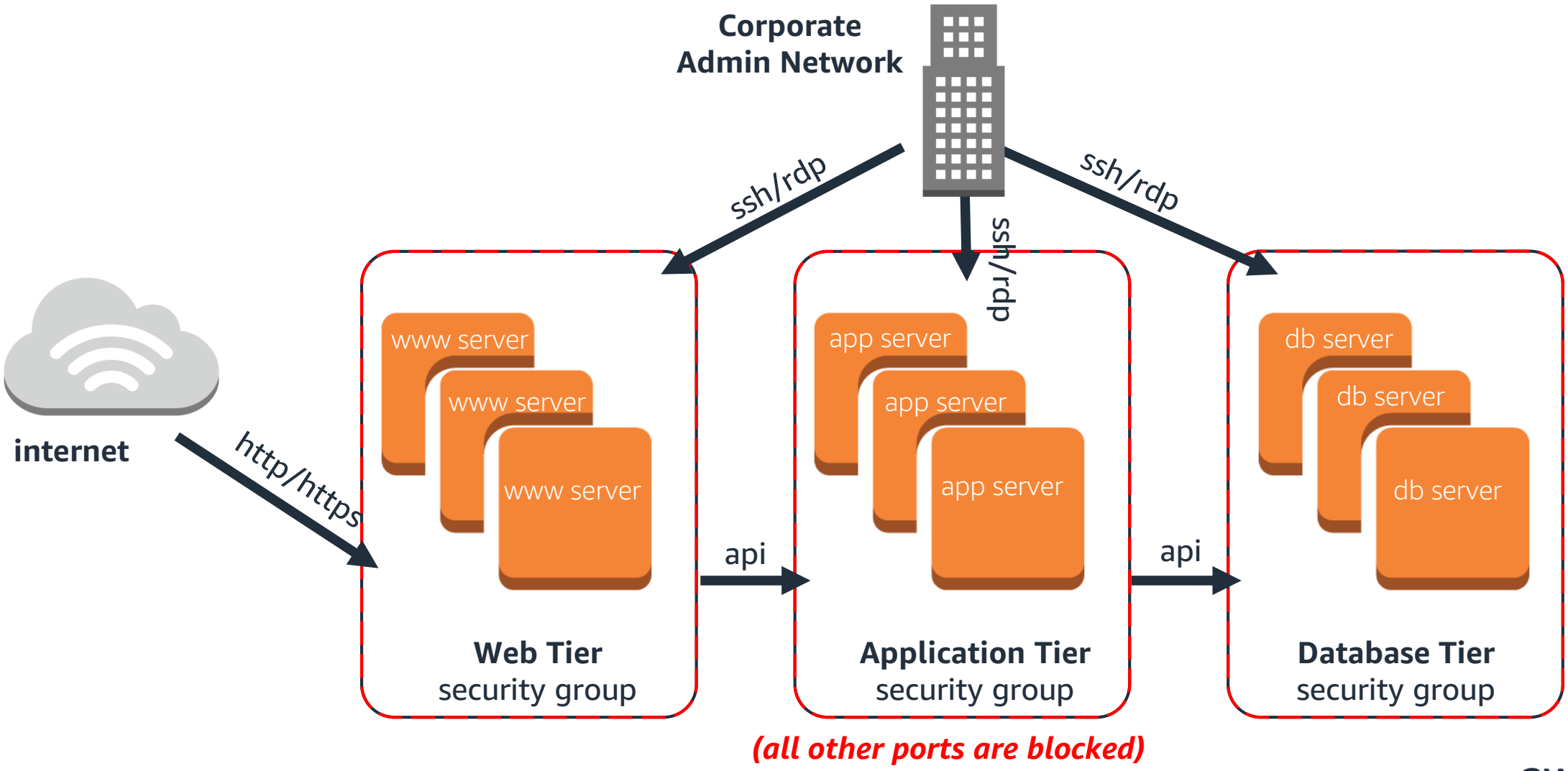
AWS Security

Is the highest priorities

Security groups

- Act as built-in firewalls
- Control accessibility to instances

AWS Security



Compute Services

Compute Services

AWS

- Flexible
- Cost effective

Amazon EC2

- Flexible configuration and control

AWS Lambda

- Pay only for what you use
- No administration

Compute Services

Amazon Lightsail

- Launch virtual private server
- Manage simple web and application servers

Amazon ECS

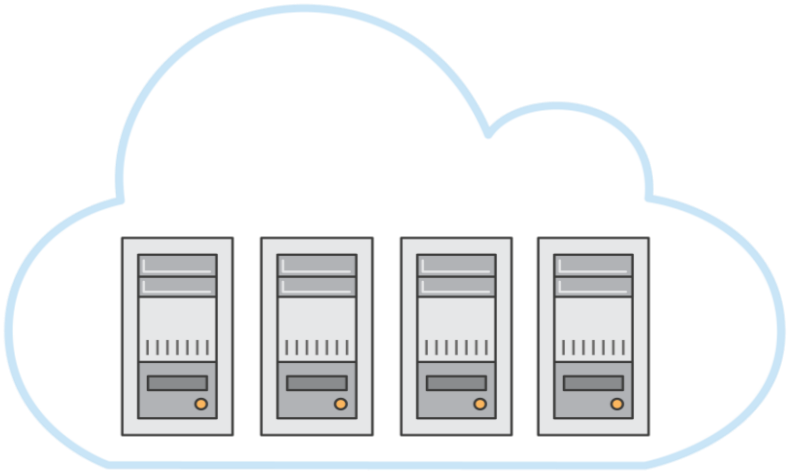
- Managed containers
- Highly scalable, high performance

AWS Fargate

Amazon EKS

Amazon Elastic Compute Cloud (EC2)

Elastic Compute Cloud



- ✓ Application Server
- ✓ Web Server
- ✓ Database Server
- ✓ Game Server
- ✓ Mail Server
- ✓ Media Server
- ✓ Catalog Server
- ✓ File Server
- ✓ Computing Server
- ✓ Proxy Server

What is Amazon EC2?

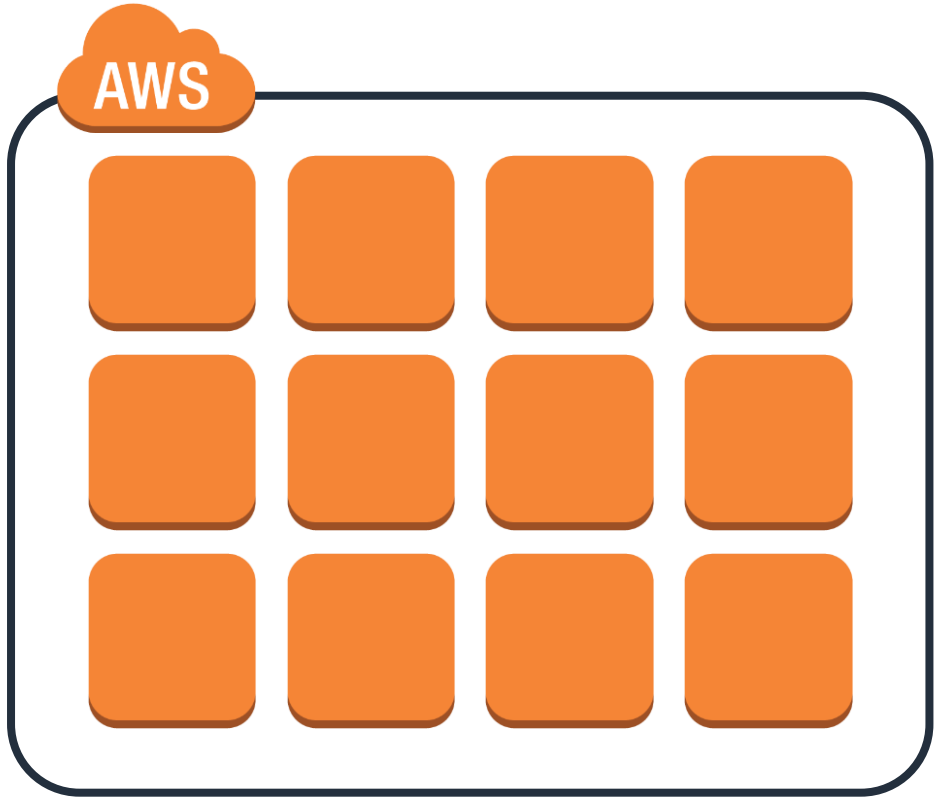
Amazon EC2 Instances

Pay as you go

Broad selection of HW/SW

Global hosting

Much more (aws.amazon.com/ec2)



Instance Types

Families	Description	Example Use Cases
t2, m4, m3	**General Purpose** Balanced Performance	Websites, web applications, Dev, code repos, micro services, business apps
c3, c4, cc2	**Compute Optimized** High CPU Performance	Front-end fleets, web-servers, batch processing, distributed analytics, science and engineering apps, ad serving, MMO gaming, video-encoding
g2, p2	**GPU Optimized** High-end GPU	Amazon AppStream 2.0, video encoding, machine learning, high perf databases, science
r3, r4, x1, cr1	**Memory Optimized** Large RAM footprint	In-memory databases, data mining
d2, i2, i3, hi1, hs1	**Storage Optimized** High I/O, High density	NAS, data warehousing, NoSQL

Choosing the Right Amazon EC2 Instances

EC2 Instance types are optimized for different use cases, workloads & come in multiple sizes. This allows you to optimally scale resources to your workload requirements.

AWS utilizes Intel® Xeon® processors for EC2 Instances providing customers with high performance and value.

Consider the following when choosing your instances: core count, memory size, storage size & type, network performance, I/O requirements & CPU technologies.

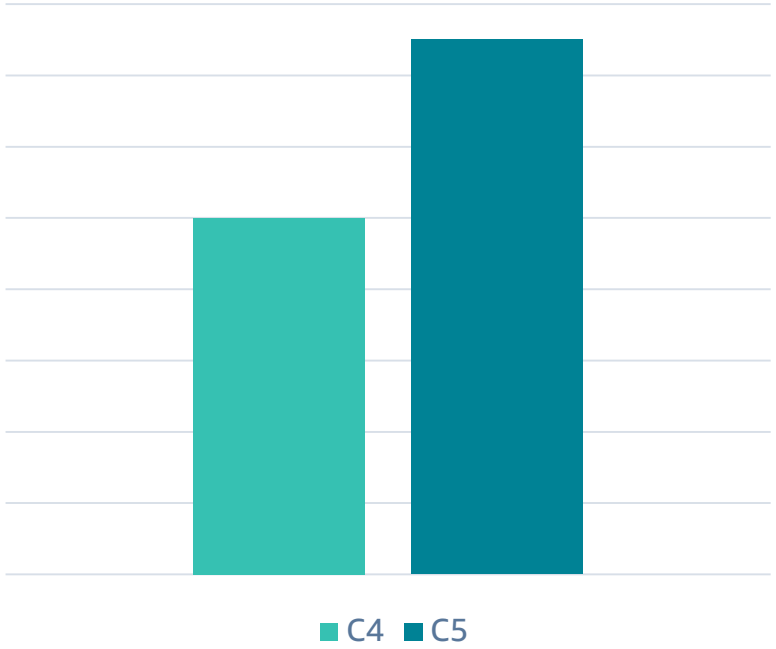
Hurry Up & Go Idle - A larger compute instance can save you time and money, therefore paying more per hour for a shorter amount of time can be less expensive.

EC2 Instances Powered by Intel Technologies

EC2 Instance Type	Compute Optimized		General Purpose			Memory Optimized			Storage Optimized		
	C5	C4	M5	M4	T2	X1	X1e	R4	H1	I3	D2
Intel Processor	Xeon Platinum 8175M	Xeon E5 2666 v3	Xeon Platinum 8175M	Xeon E5 2686 v4 2676 v3	Xeon Family	Xeon E7 8880 v3	Xeon E7 8880 v3	Xeon E5 2686 v4	Xeon E5 2686 v4	Xeon E5 2686 v4	Xeon E5 2676 v3
Intel Processor Technology	Skylake	Haswell	Skylake	Broadwell Haswell	Yes	Haswell	Haswell	Broadwell	Broadwell	Broadwell	Haswell
Intel AVX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intel AVX2	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	Yes
Intel AVX-512	Yes	-	Yes	-	-	-	-	-	-	-	-
Intel Turbo Boost	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Storage	EBS-only	EBS-only	EBS-only	EBS-only	EBS-only	SSD EBS-Opt	SSD EBS-Opt	-	HDD	SSD	HDD

C5: Compute Optimized Instances

25% price/performance improvement over C4



Based on 3.0 GHz Intel Xeon Scalable Processors (Skylake)

Up to 72 vCPUs and 144 GiB of memory (2:1 Memory:vCPU ratio)

25 Gbps NW bandwidth

Support for Intel AVX-512

NETFLIX

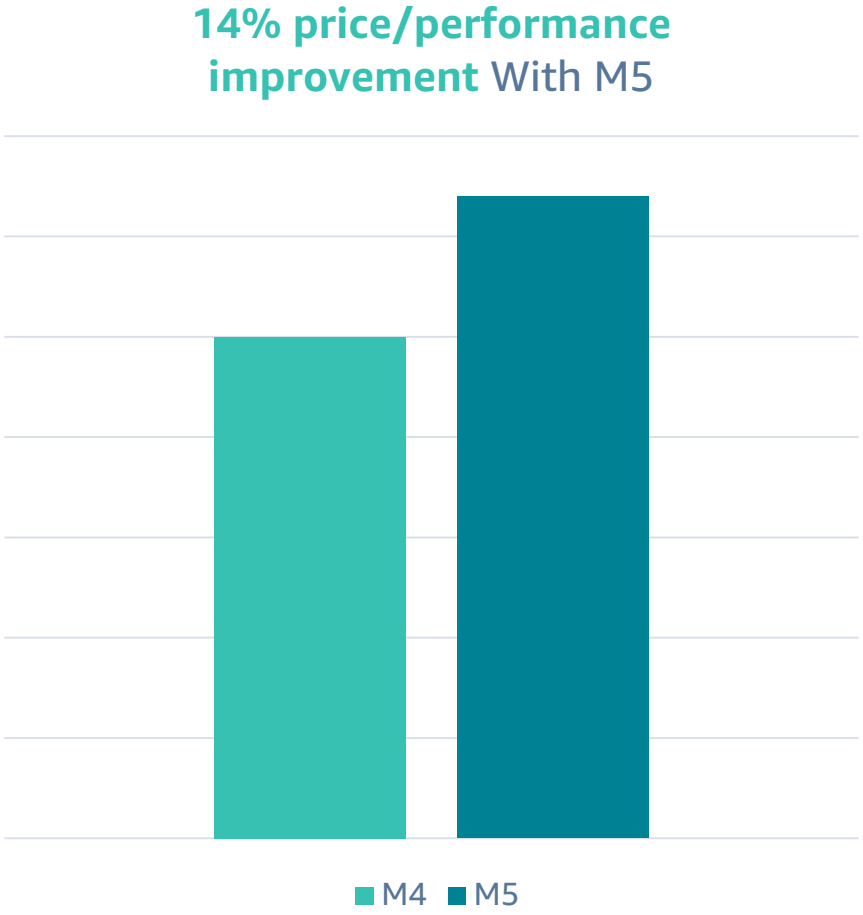
"We saw significant performance improvement on Amazon EC2 C5, with up to a 140% performance improvement in industry standard CPU benchmarks over C4."

GRAIL

"We are eager to migrate onto the AVX-512 enabled c5.18xlarge instance size… . We expect to decrease the processing time of some of our key workloads by more than 30%."

aws | intel

M5: Next-Gen General Purpose instance



Powered by 2.5 GHz Intel Xeon Scalable Processors (Skylake)

New larger instance size—m5.24xlarge with 96 vCPUs and 384 GiB of memory (4:1 Memory:vCPU ratio)

Improved network and EBS performance on smaller instance sizes

Support for Intel AVX-512 offering up to twice the performance for vector and floating point workloads

Elastic Load Balancing (ELB)

Introduction to Elastic Load Balancing

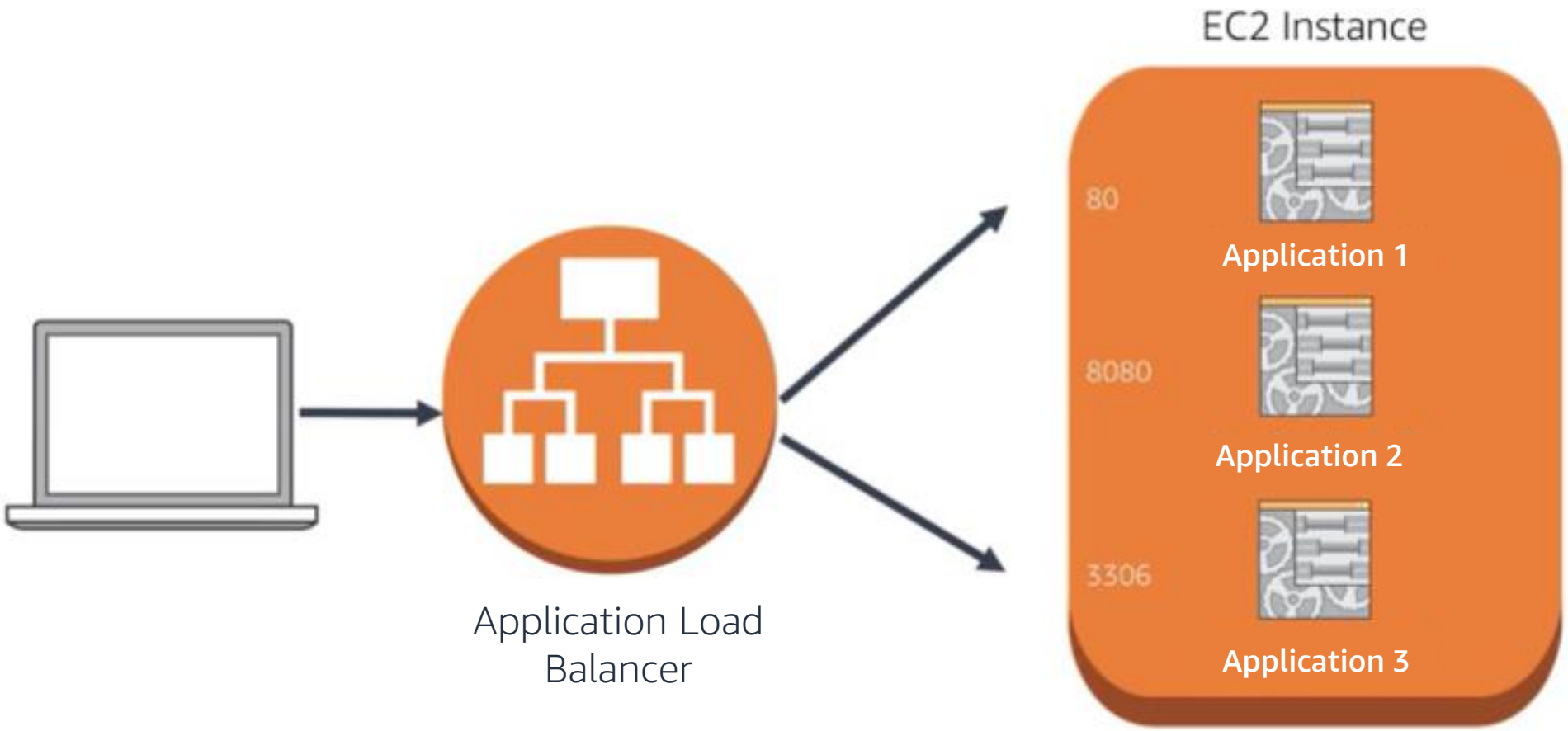
Managed load balancing service

Distributes loads between instances

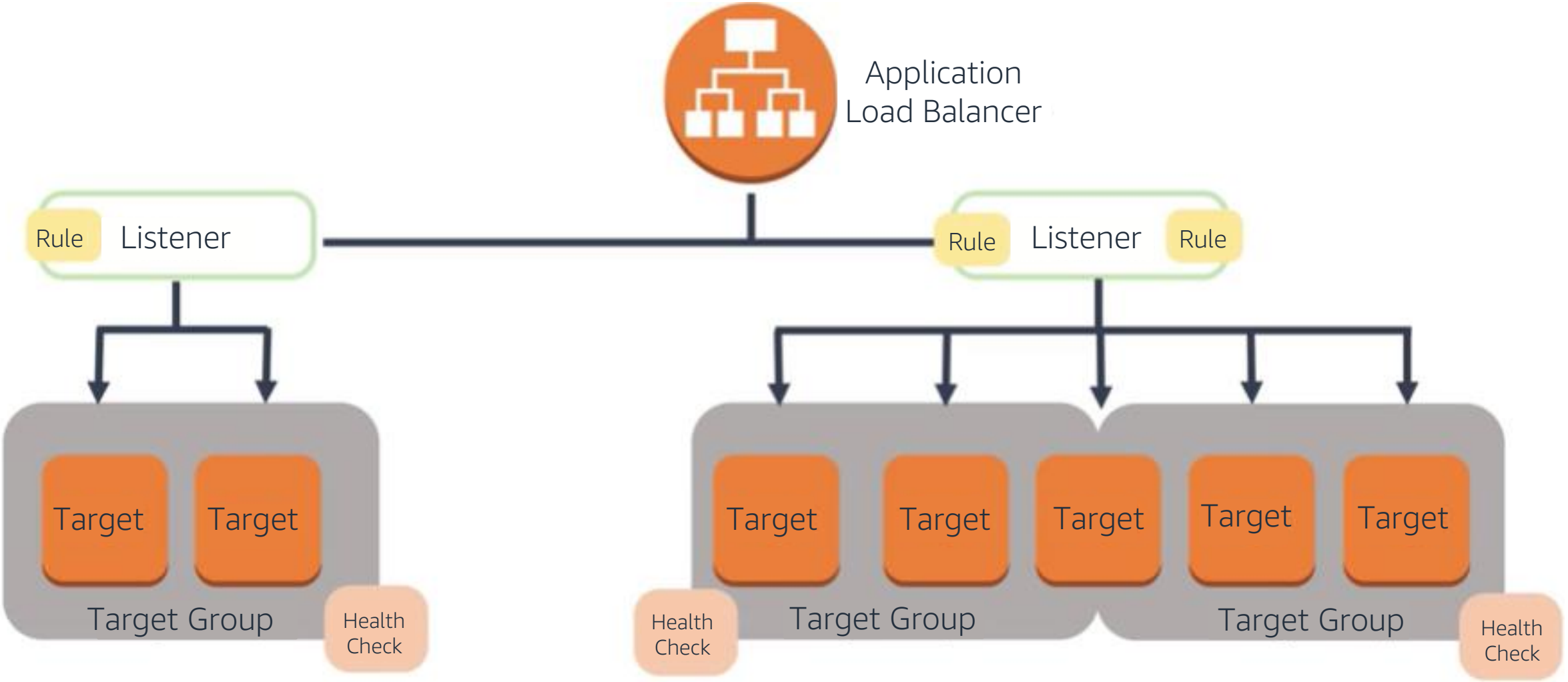
Elastic Load Balancing Products

Application Load Balancer (ALB)	Network Load Balancer (NLB)	Classic Load Balancer (CLB)
HTTP HTTPS	TCP	PREVIOUS GENERATION for HTTP, HTTPS, and TCP
• Flexible application management • Advanced load balancing of HTTP and HTTPS traffic • Operates at the request level (Layer 7)	• Extreme performance and static IP for your application • Load balancing of TCP traffic • Operates at the connection level (Layer 4)	• Existing application that was built within the EC2-Classic network • Operates at both the request level and connection level

Application Load Balancer Use Cases



Application Load Balancer Use Cases



Network Load Balancer Use Cases

Sudden and volatile traffic patterns

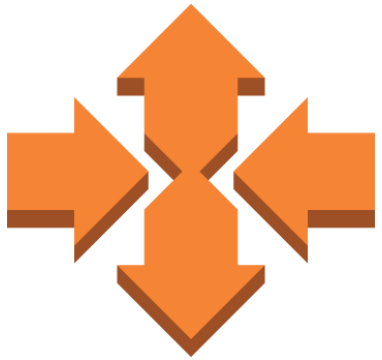
Single static IP address per Availability Zone

Ideal for applications that require extreme performance

Auto Scaling

What Is Auto Scaling?

Helps you verify that you have the desired number of Amazon EC2 instances available to handle the load for your application

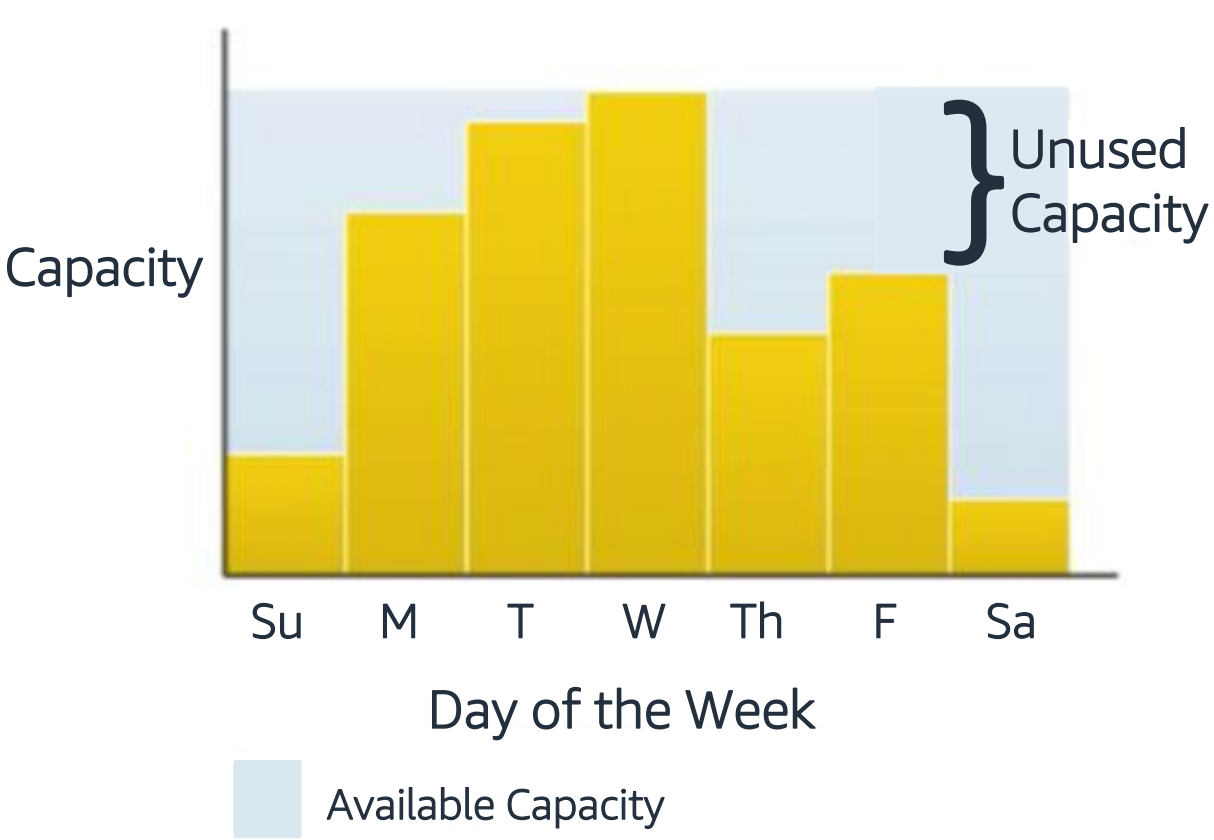


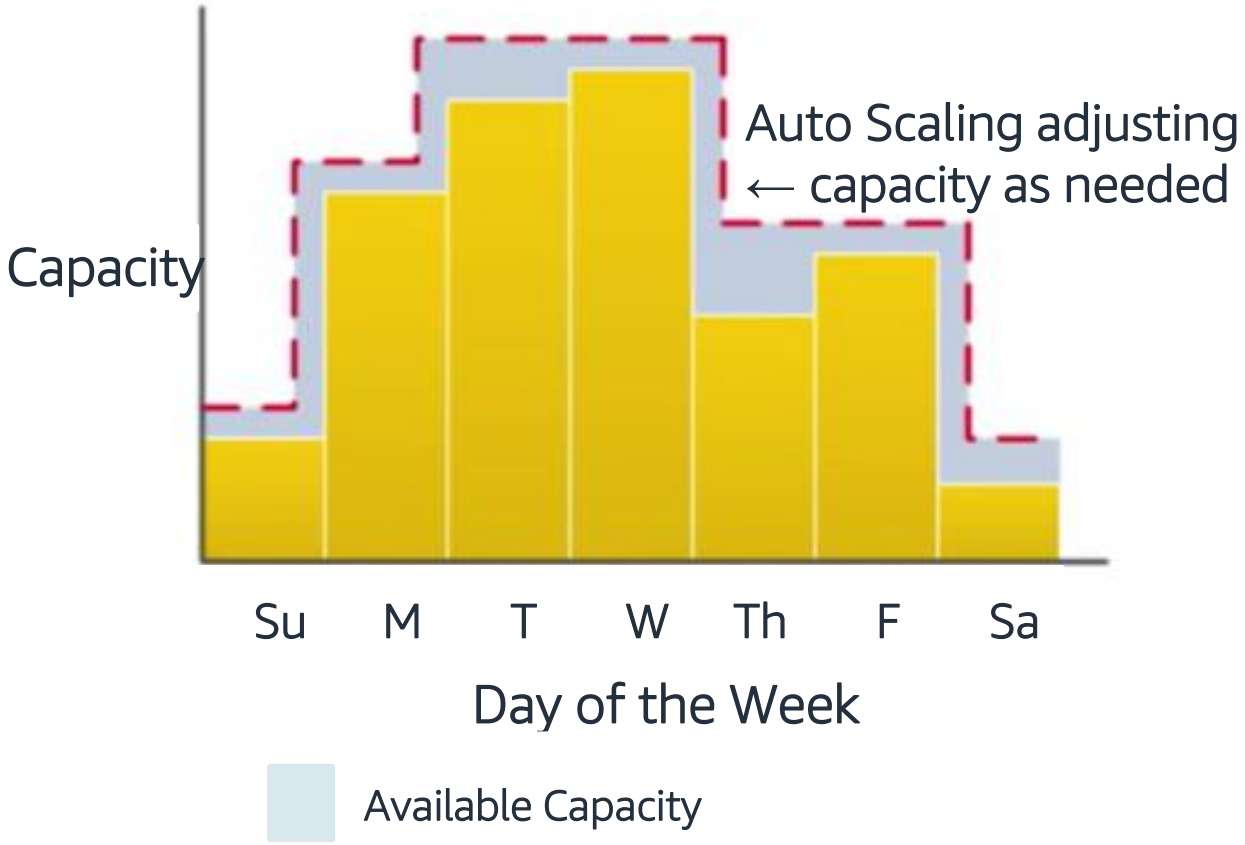
Monitoring Resource Performance

Amazon CloudWatch to monitor performance

Auto Scaling to add or remove EC2 instances

Capacity Management





Critical Questions

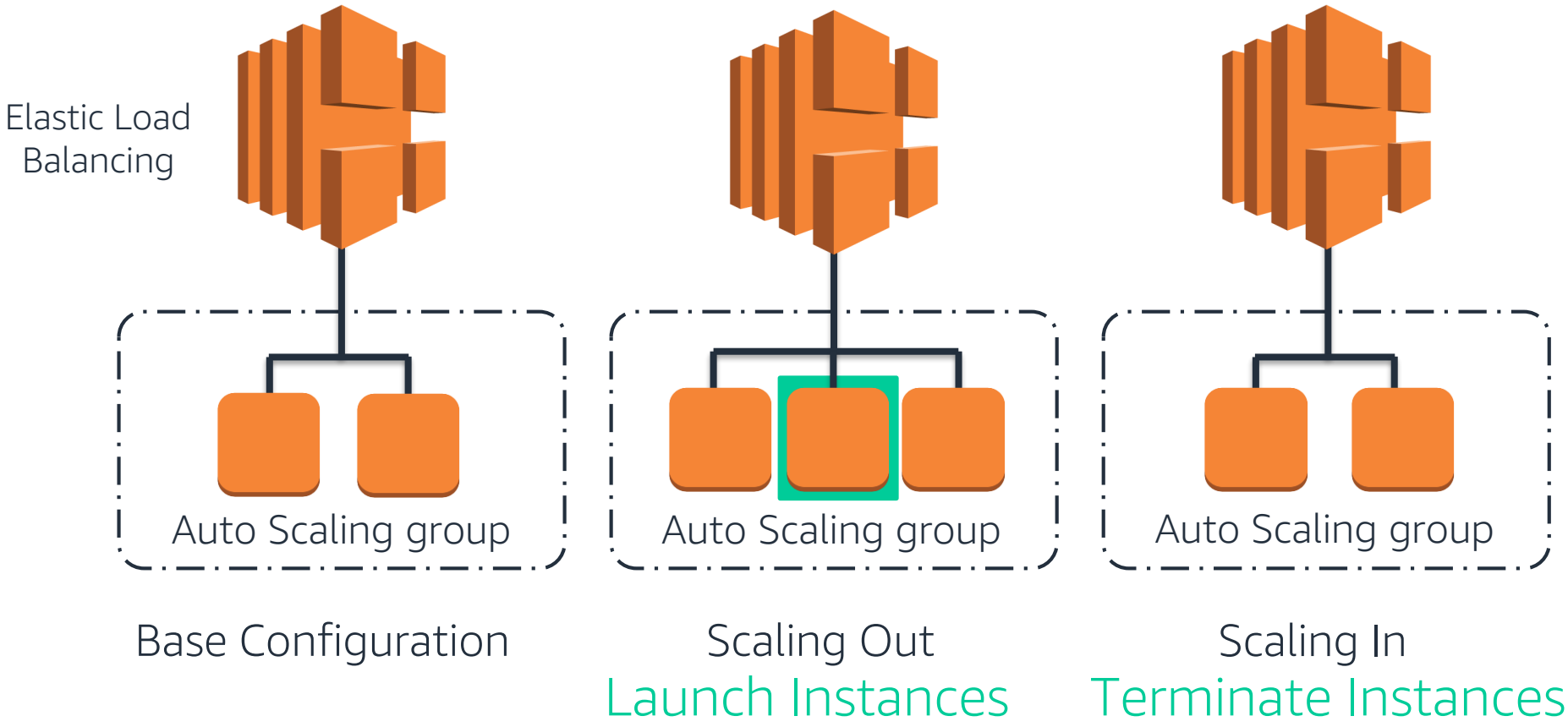
How can I make sure that my workload has enough EC2 resources to meet fluctuating performance requirements?

Scalability

How can EC2 resource provisioning occur on-demand?

Automation

Scaling Out and Scaling In



Auto Scaling Components

Launch Configuration

Auto Scaling groups

Auto Scaling Policy

Auto Scaling Components

Launch Configuration: *What will be scaled?*

Launch settings

- AMI
- Instance type
- Security groups
- Roles

Auto Scaling Components

Auto Scaling Group: *Where will it take place?*

Deployment settings

- VPC and subnets
- Load balancer
- Minimum instances
- Maximum instances
- Desired capacity

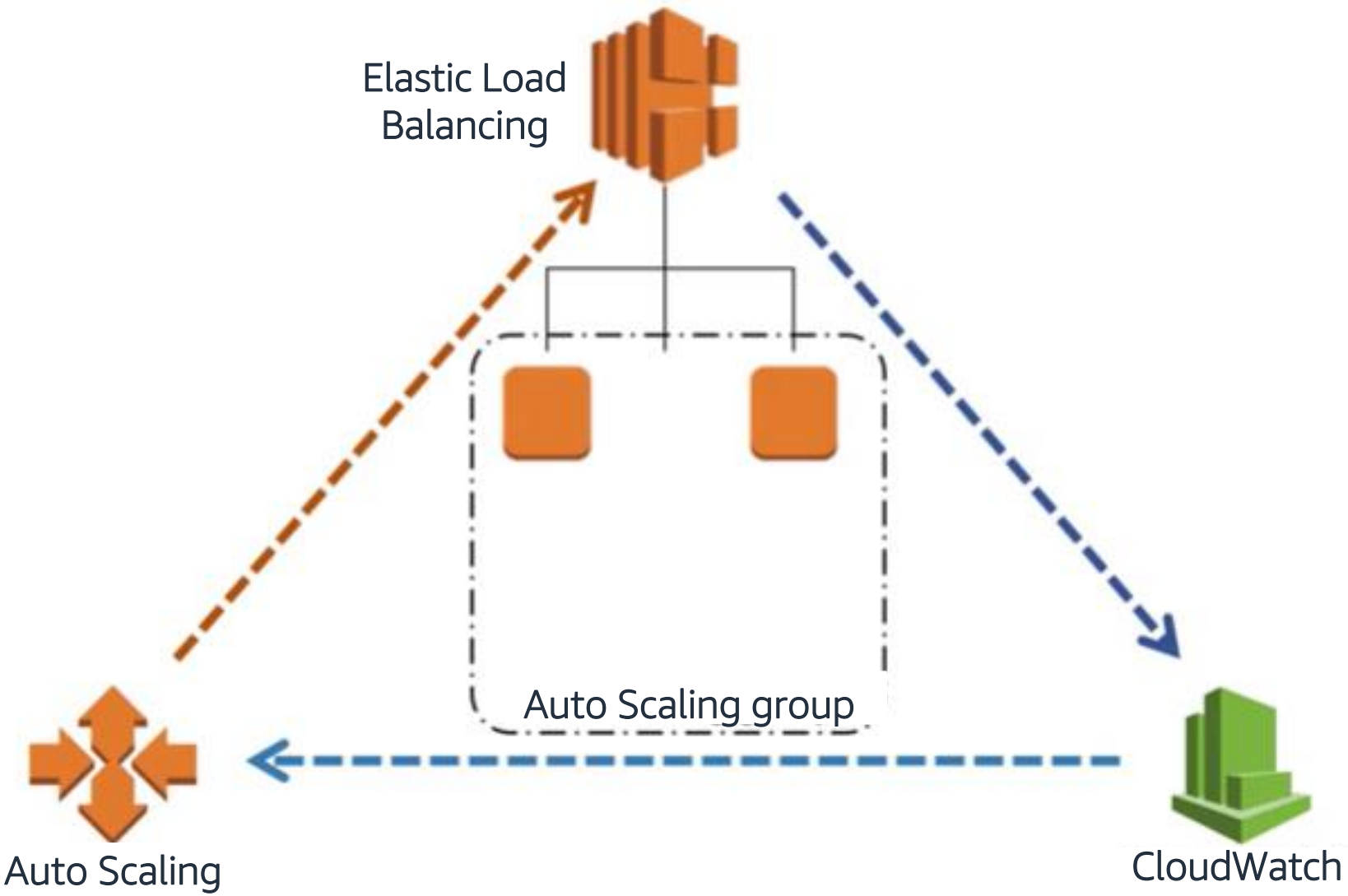
Auto Scaling Components

Auto Scaling Policy: *When will it take place?*

Policy settings

- Scheduled
- On-demand
- Scale-out policy
- Scale-in policy

Dynamic Auto Scaling



CloudWatch Alarm for Auto Scaling

Whenever: CPUUtilization

is : >= 80

for: 1 consecutive period(s)

AutoScaling Action Delete

Whenever this alarm: State is ALARM

From resource type: AutoScaling

From the: IREASG

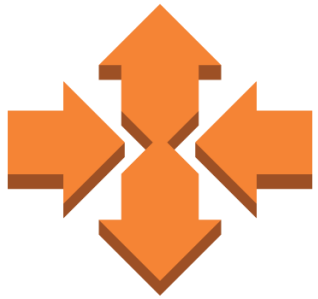
Take this action: Increase Group Size – Add 2 instances

Summary

Created

- A launch configuration
- Auto Scaling group
- Auto Scaling policy

Triggered Auto Scaling



Amazon Elastic Block Store (EBS)

EBS Volumes

Characteristics

- Persistent and customizable block storage for EC2 instances
- HDD and SSD types
- Use Snapshots for backups
- Easy and transparent encryption
- Elastic

EBS Volumes

Availability

- Durable and automatically replicated

Drive Types

- Storage that best fits your needs
- Magnetic or SSD
- Performance and price requirements

Amazon EBS

Snapshots

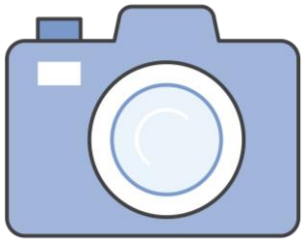
- Point-in-time snapshots
- Recreate a new volume at any time

Encryption

- Encrypted EBS volumes
- No additional cost

Elasticity

- Increase capacity
- Change to different types



Summary

Features

- Persistent and customizable block storage for EC2 instances
- HDD and SSD types
- Replicated in the same Availability Zone
- Easy and transparent encryption
- Elastic volumes
- Back up using snapshots

Amazon Simple Storage Service (S3)

Amazon S3

Features

- Fully managed cloud storage service
- Rich security controls

Functionality

- Store virtually unlimited number of objects
- Access any time, from anywhere

Getting Started with S3

welcome.mp4

my-bucket-name

med welcome.mp4

Key

Object

my-bucket-name

Example - http://**my-bucket-name**.s3.amazonaws.com/**welcome.mp4**

Access the Data Anywhere

AWS Management Console

AWS command line interface

AWS software development kits

Common Use Cases

Storing application assets

Static web hosting

Backup and disaster recovery (DR)

Staging area for big data

Summary

Fully managed cloud storage service

Store virtually unlimited number of objects

Access any time, from anywhere

Rich security controls

Common use cases

Let’s take a look at an Amazon S3 Demo

Amazon S3 Demo

Amazon Relational Database Service (RDS)

Challenges of Relational Databases

Server maintenance and energy footprint

Software installation and patches

Database backups and high availability

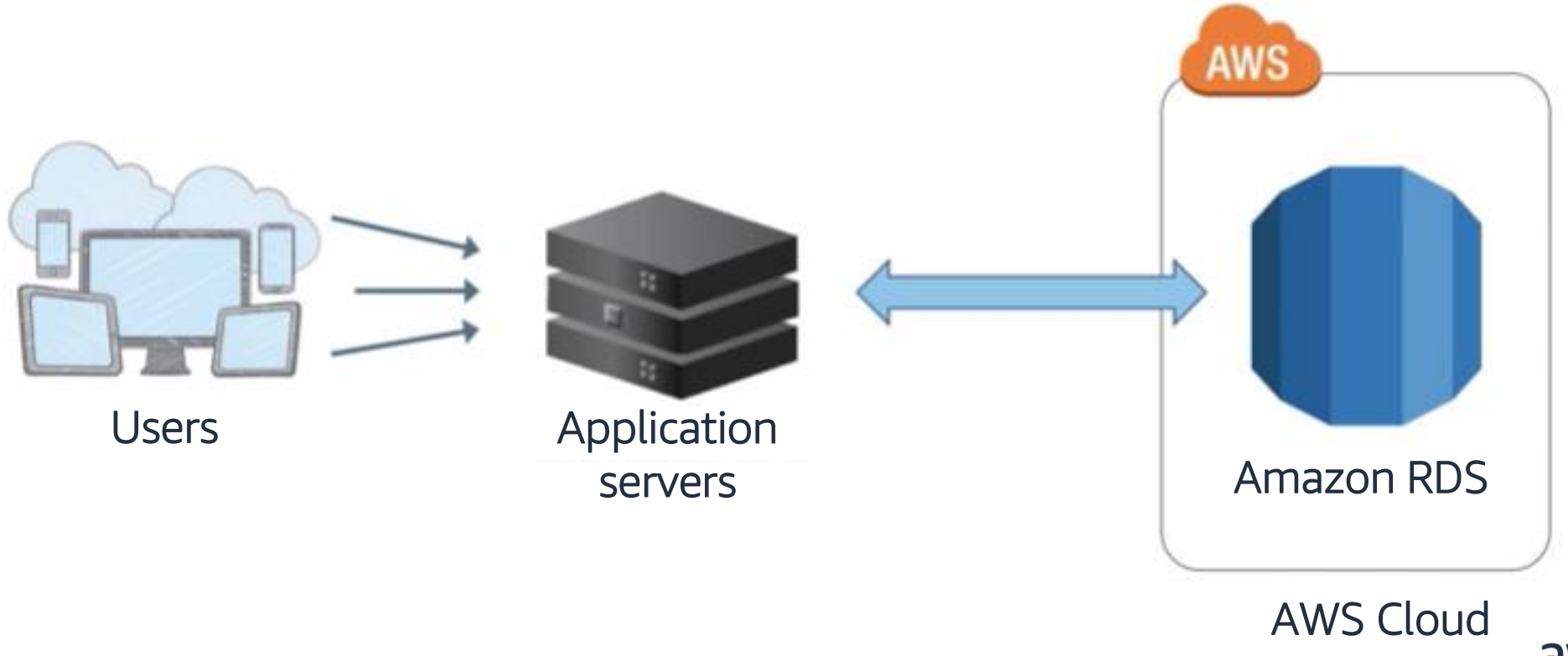
Limits on scalability

Data security

OS install and patches

Amazon RDS

Managed service that sets up and operates a relational database in the Cloud



Amazon RDS

Customer manages:

- Application Optimization
- Database schema
- Data

AWS manages:

- OS installation and patches
- Database software installation and patches
- Database backups
- High availability
- Scaling
- Power, rack and stack
- Server maintenance

Amazon RDS DB Instances

Amazon RDS

RDS DB master instance

DB Instance Class

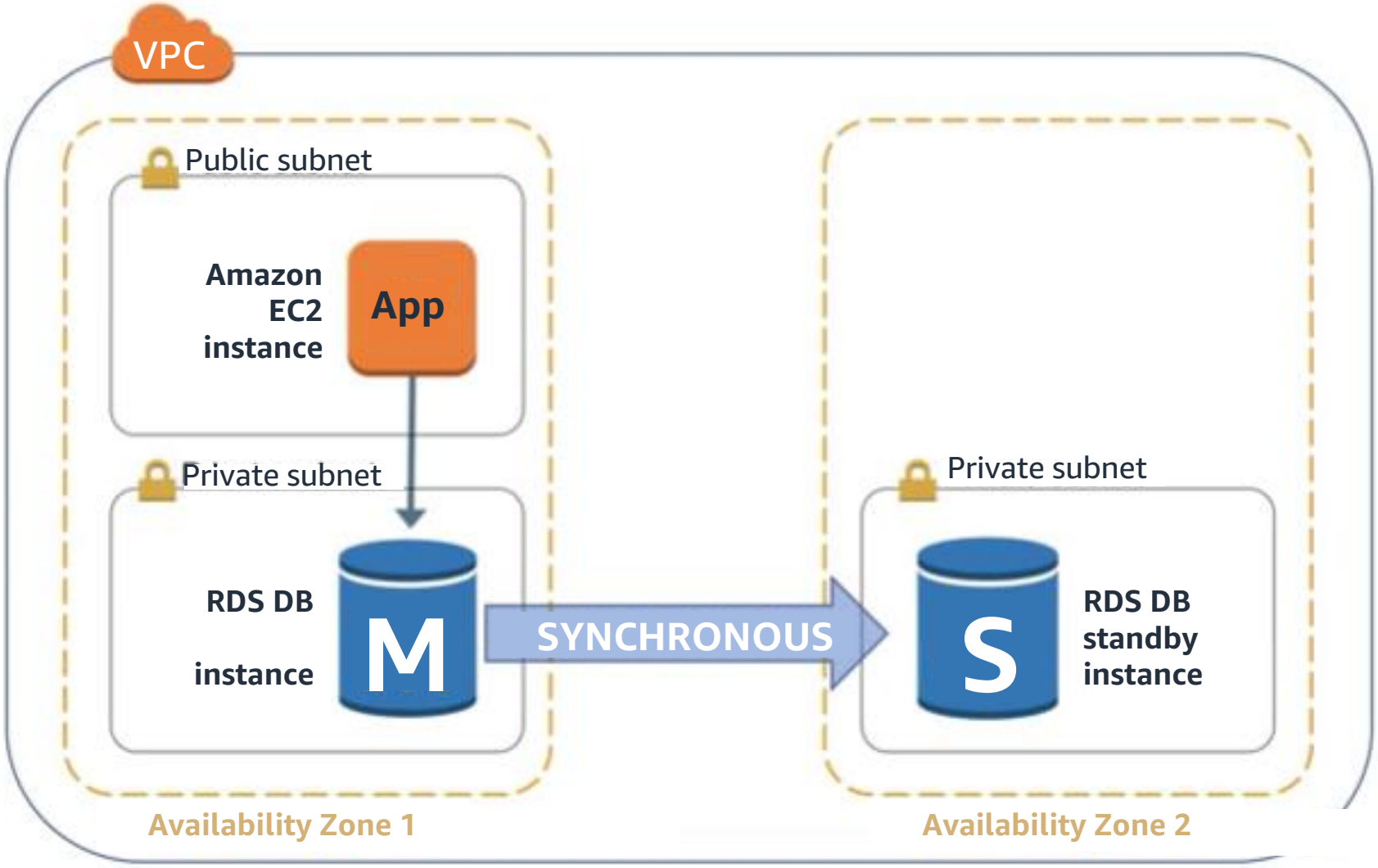
- CPU
- Memory
- Network Performance

DB Instance Storage

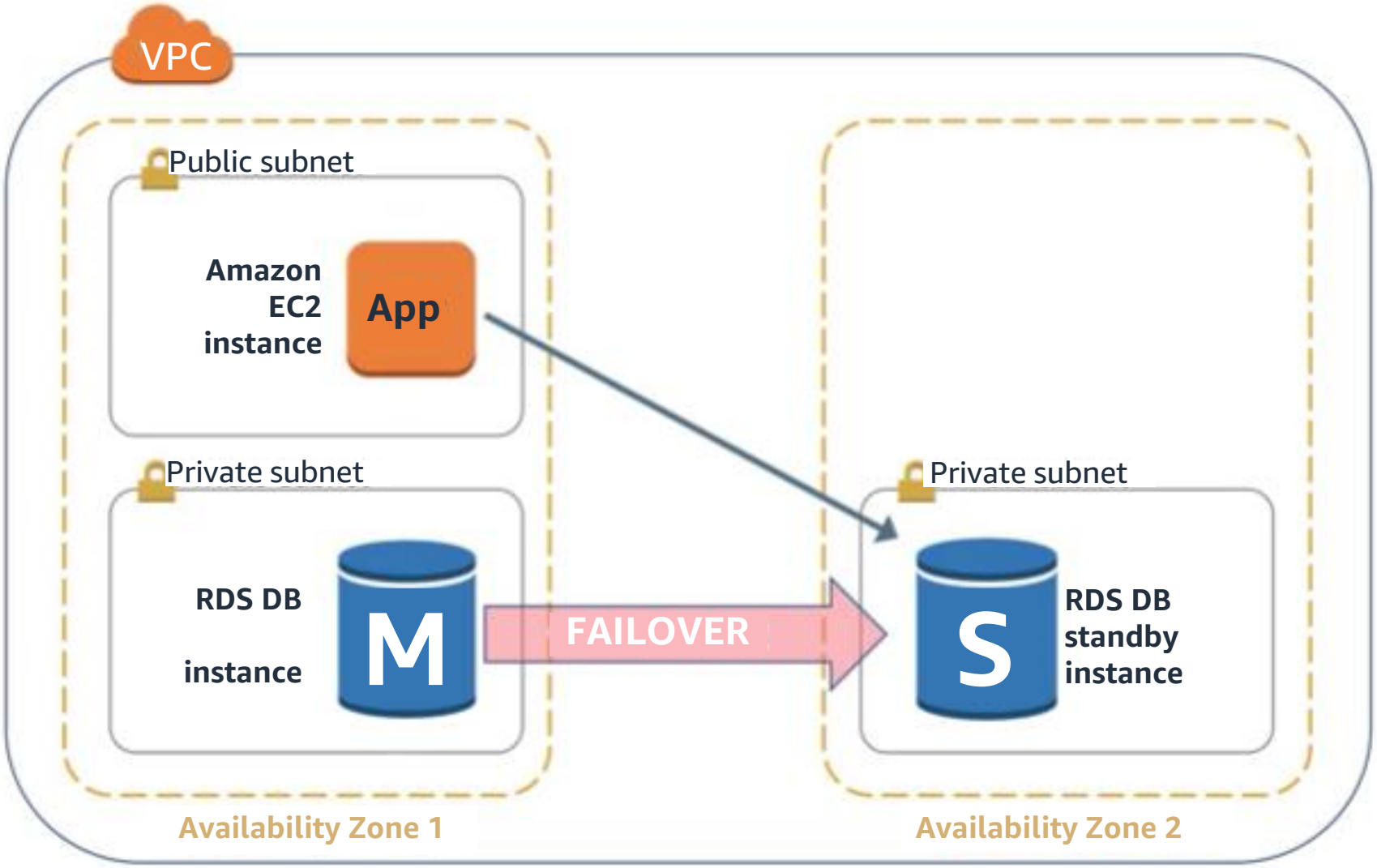
- Magnetic
- General Purpose (SSD)
- Provisioned IOPS

DB Engines

High Availability with Multi-AZ



High Availability with Multi-AZ



Amazon RDS Read Replicas

Features

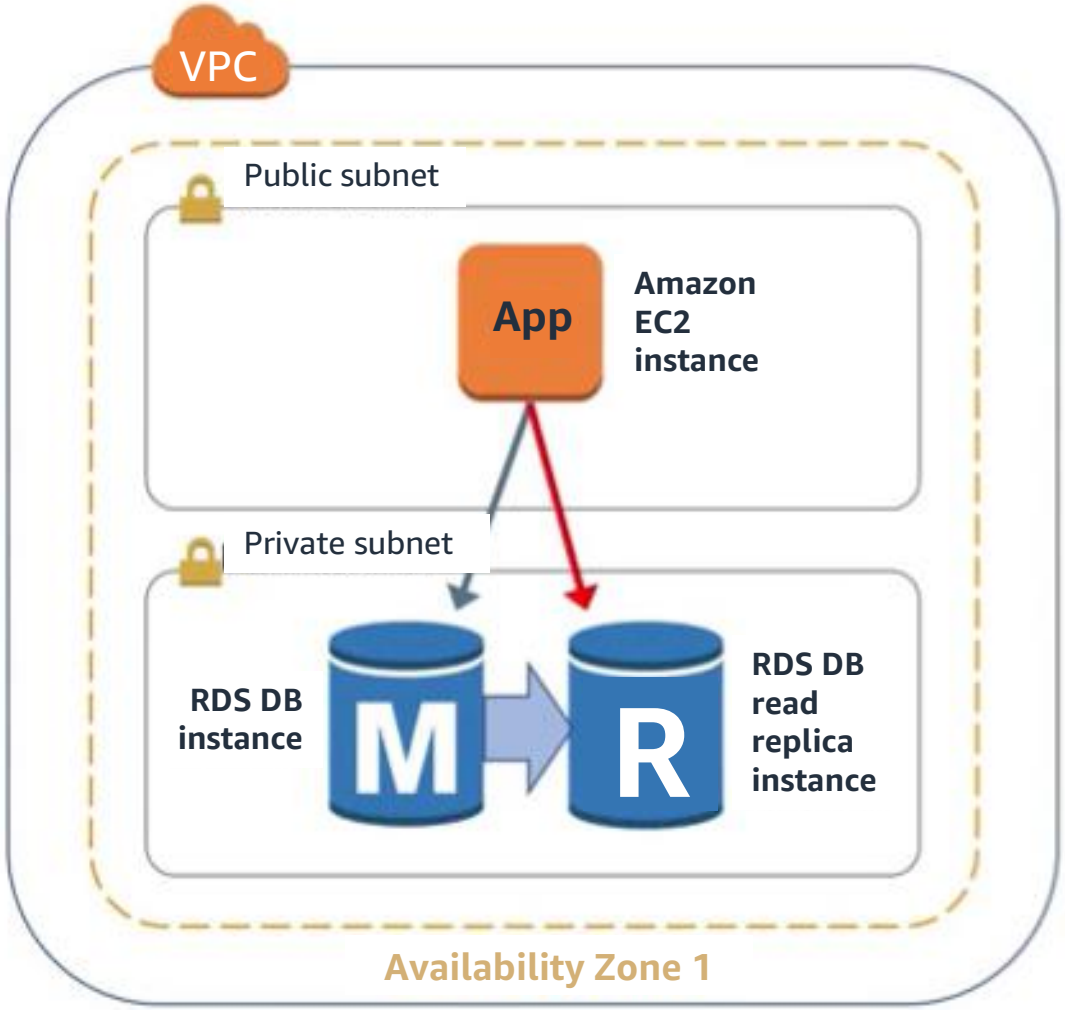
Asynchronous replication

Promote to master if necessary

Functionality

Read-heavy database workloac

Offload read queries



Summary

Highly scalable

High performance

Easy to administer

Available and durable

Secure and compliant

Amazon DynamoDB

What Is Amazon DynamoDB?

NoSQL database tables

Virtually unlimited storage

Items may have differing attributes

Low-latency queries

Scalable read/write throughput

Common Use Cases

Web

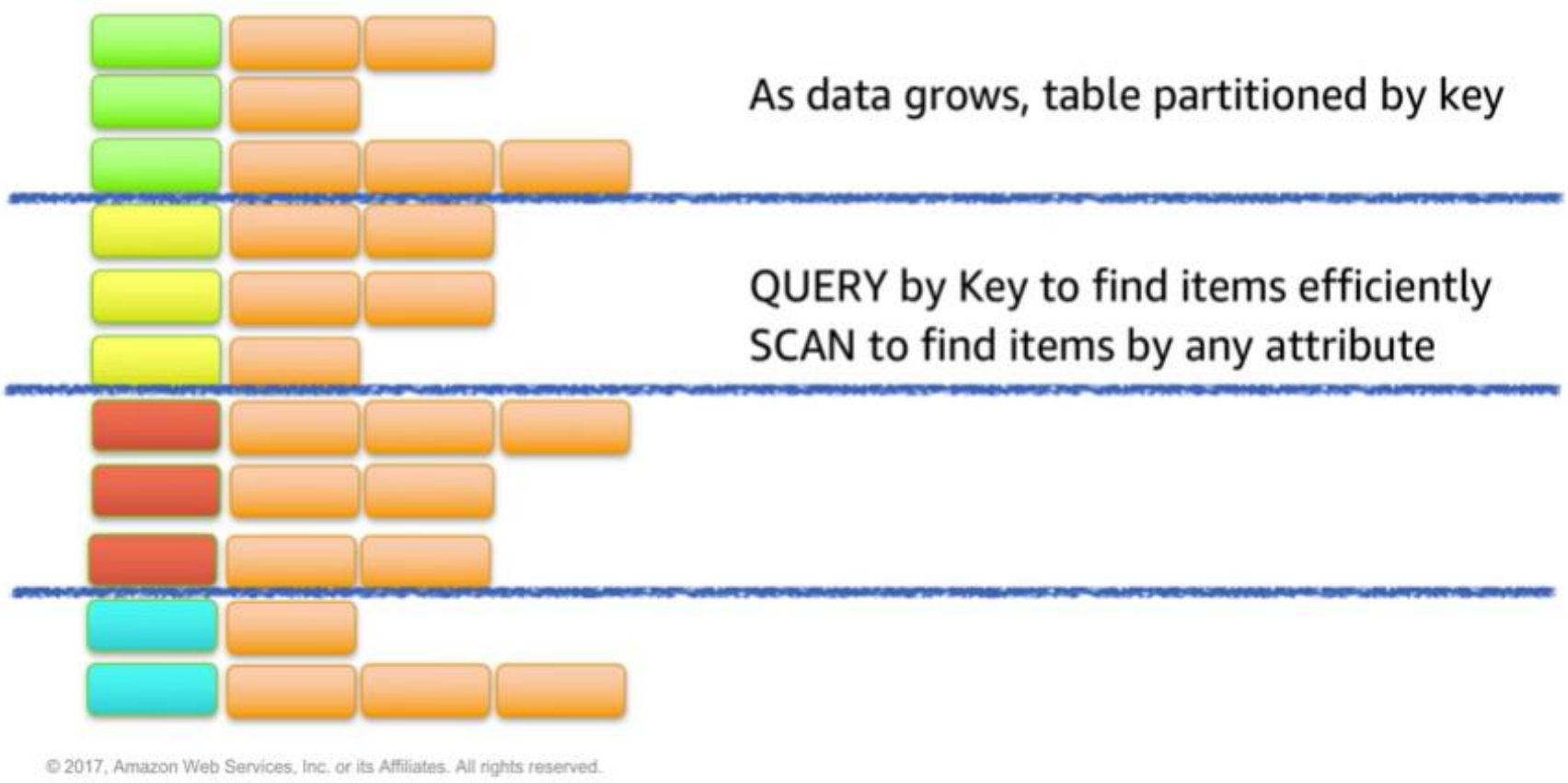
Mobile apps

Internet of Things

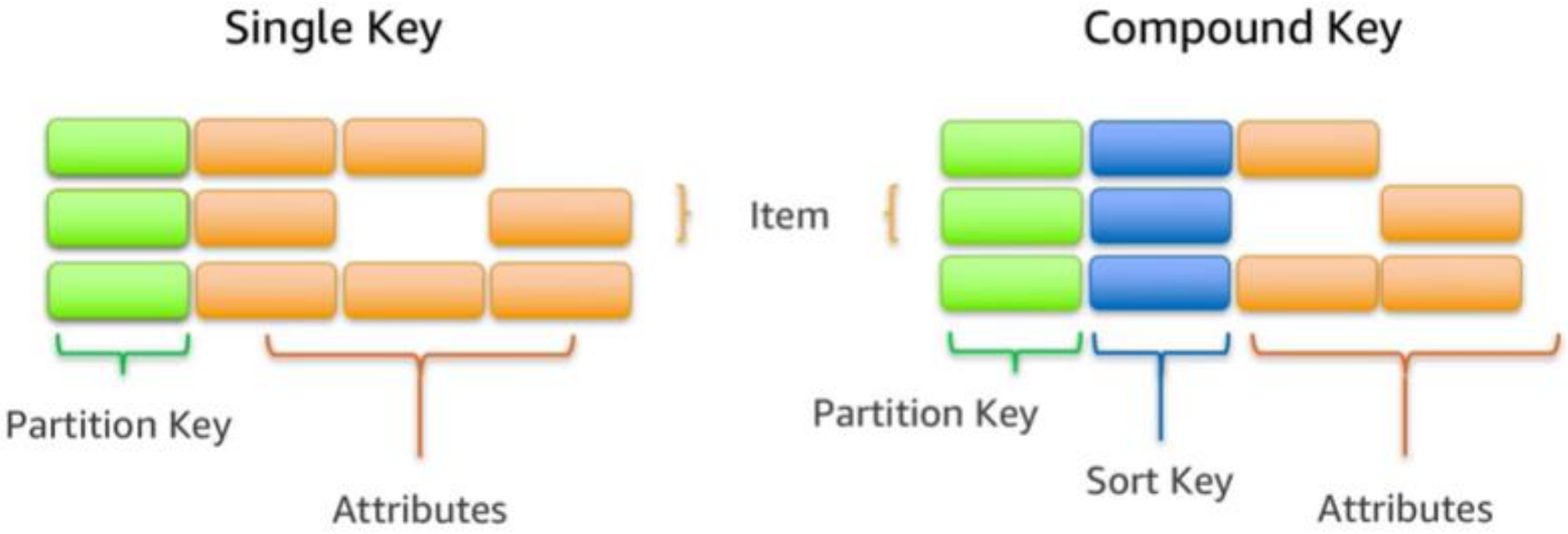
Ad tech

Gaming

Partitioning



Items in a Table Must Have a Key



Summary

Managed NoSQL database service

Data store for applications

- Store large amounts of data
- Support high request volume
- Require low-latency query performance

End of Module 2
Test Your Knowledge