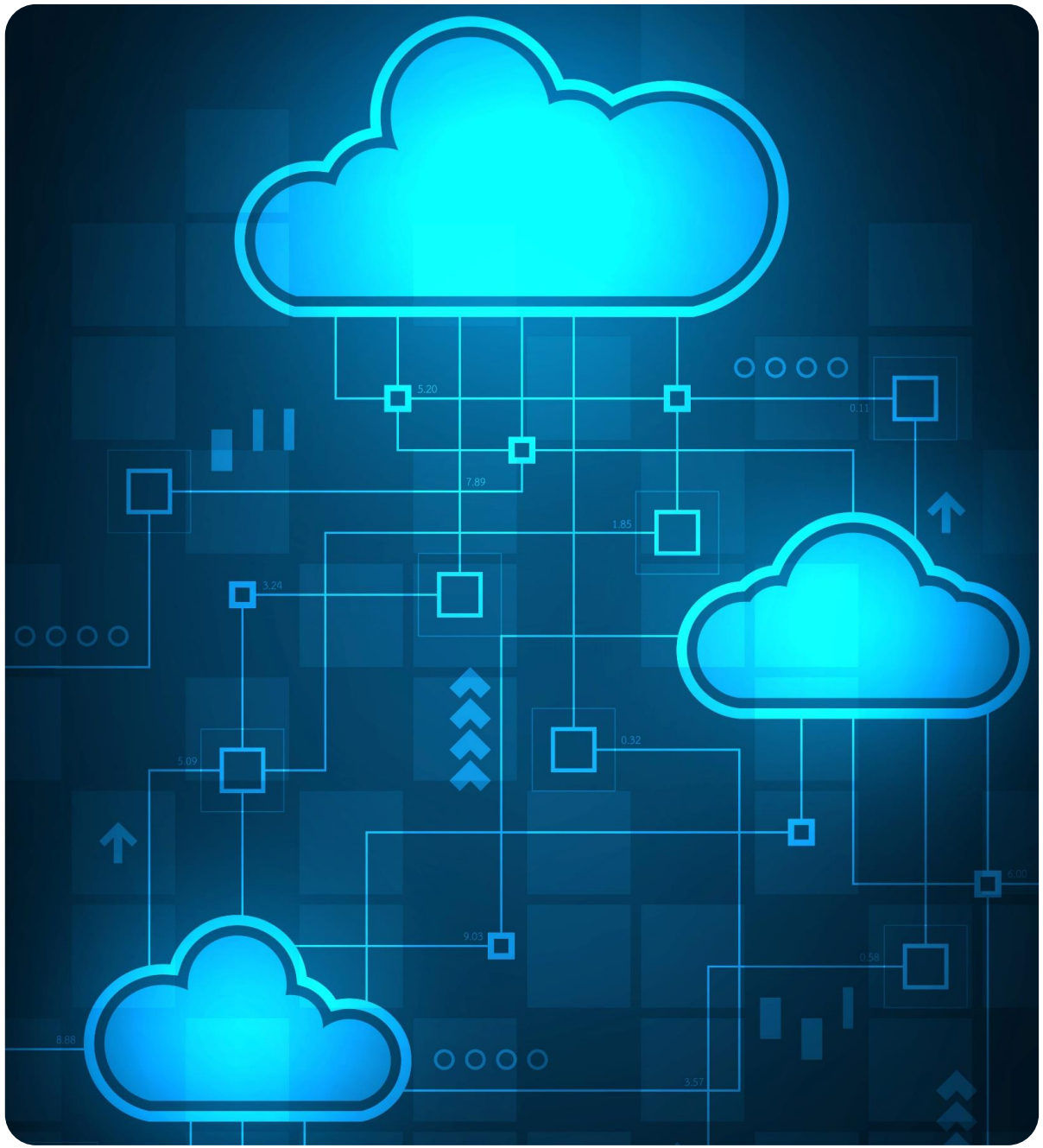
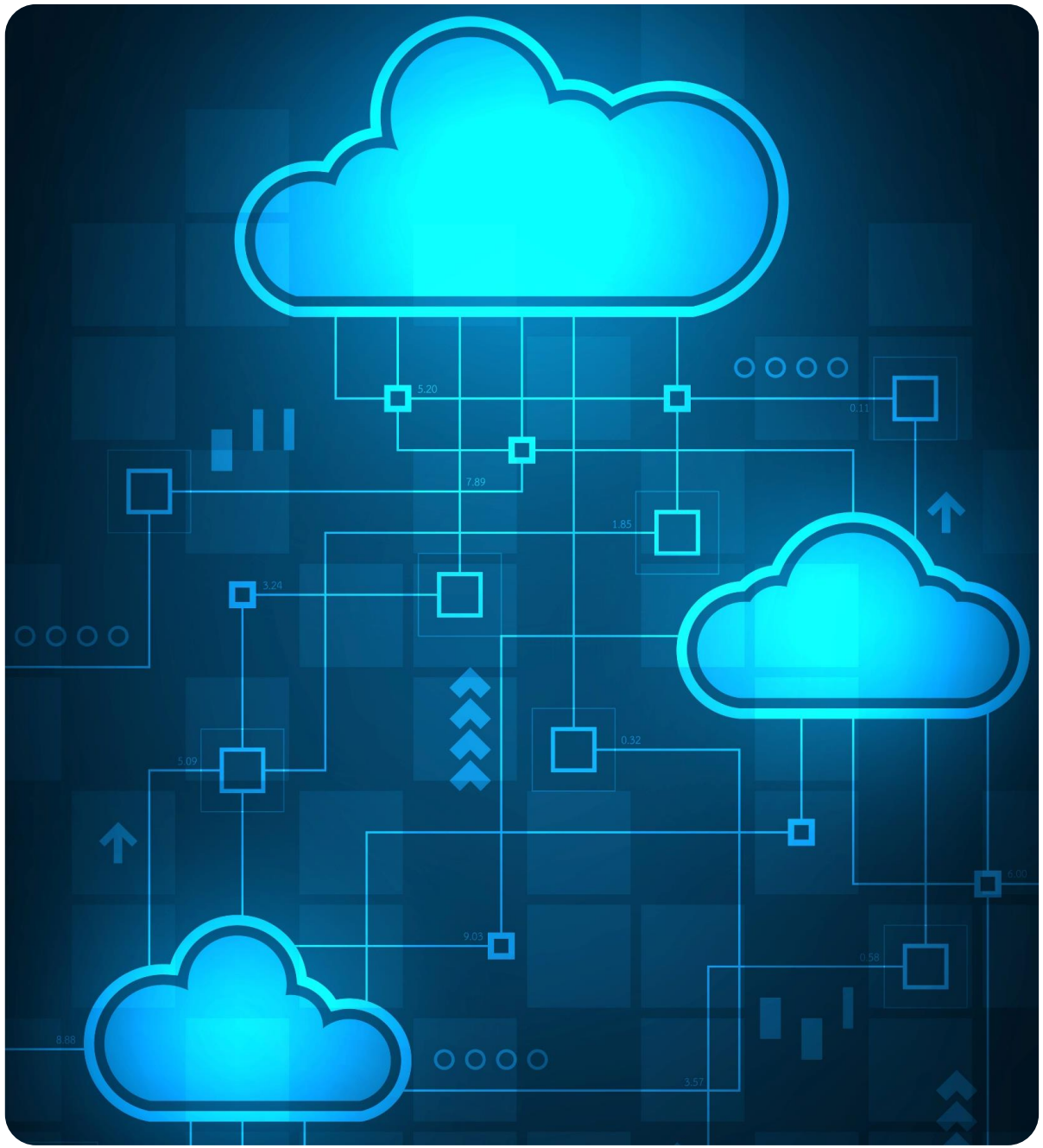
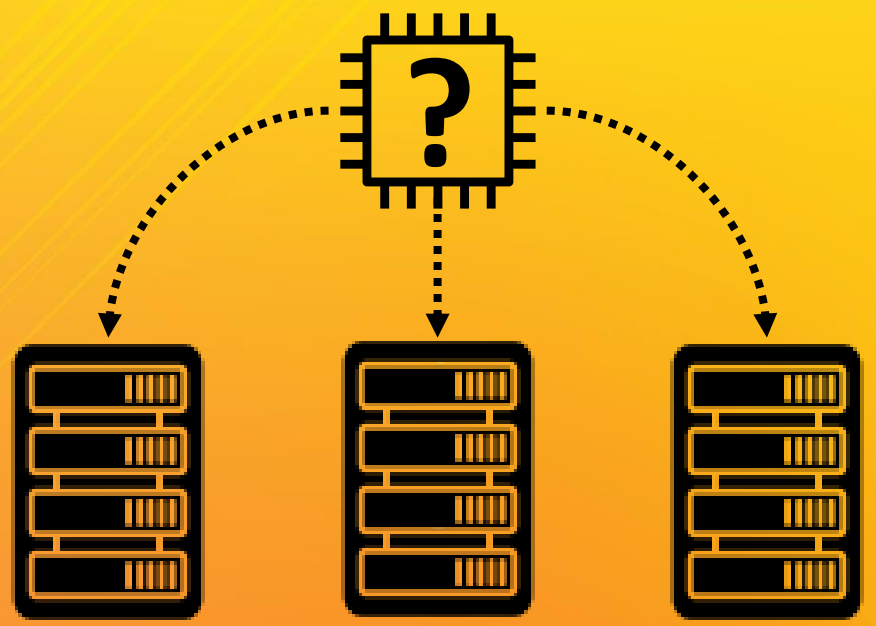
Compute in AWS – Part 2



Agenda

- Placement Group
- Dedicated Instance
- ASG Lifecycle Hooks
- Amazon Data Lifecycle Manager (Amazon DLM)



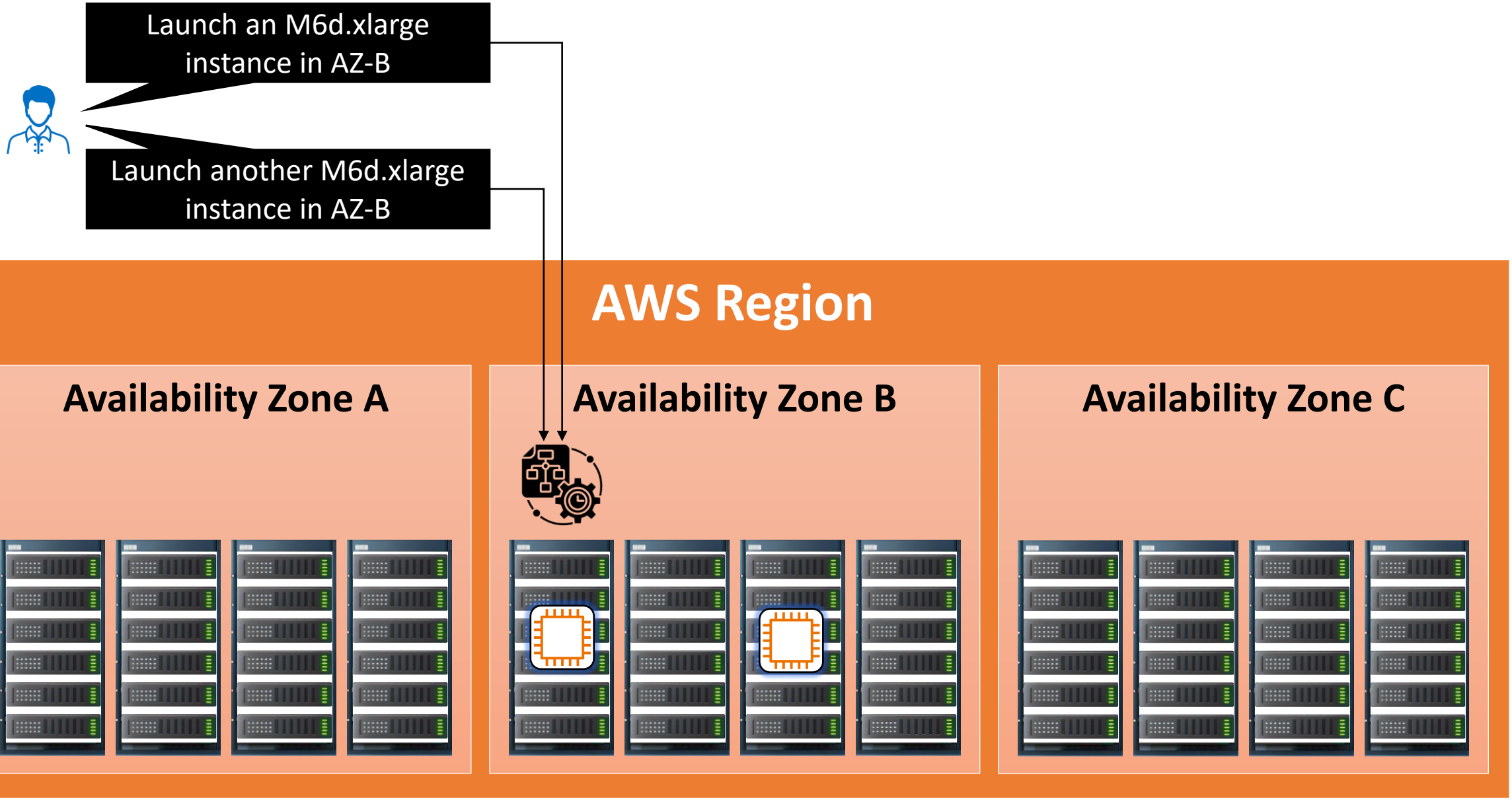


Amazon EC2 Placement Group

AWS Global Infrastructure

Availability Zone

Launching an Amazon EC2 Instance



Amazon EC2 Placement Groups

- You can influence placement of Amazon EC2 instance through placement groups, to suit your use case.

- There is no charge for creating a placement group.

- Depending on the type of workload, you can create three types of placement group :

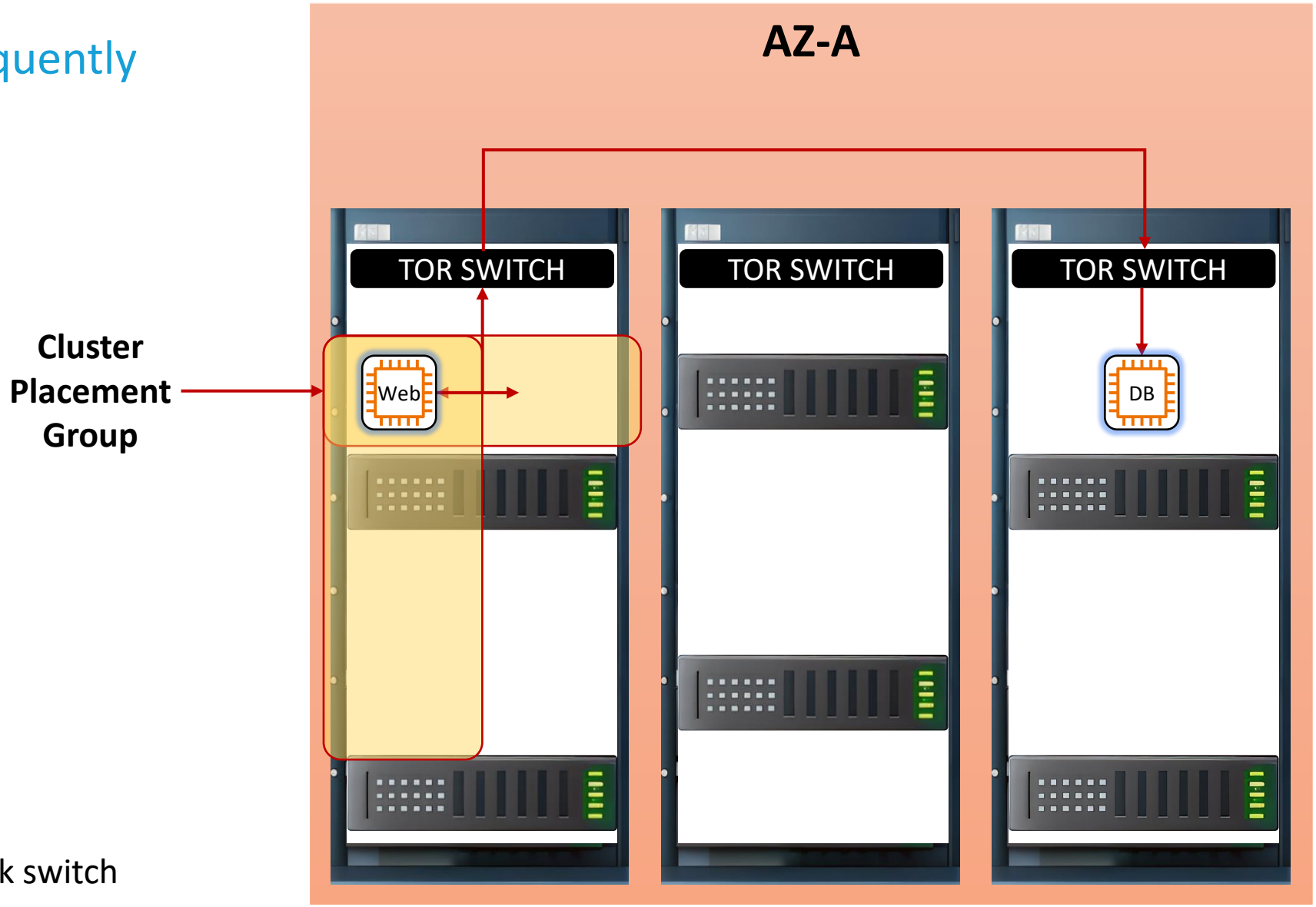
Cluster Placement Group

Spread Placement Group

Partition Placement Group

Cluster Placement Group

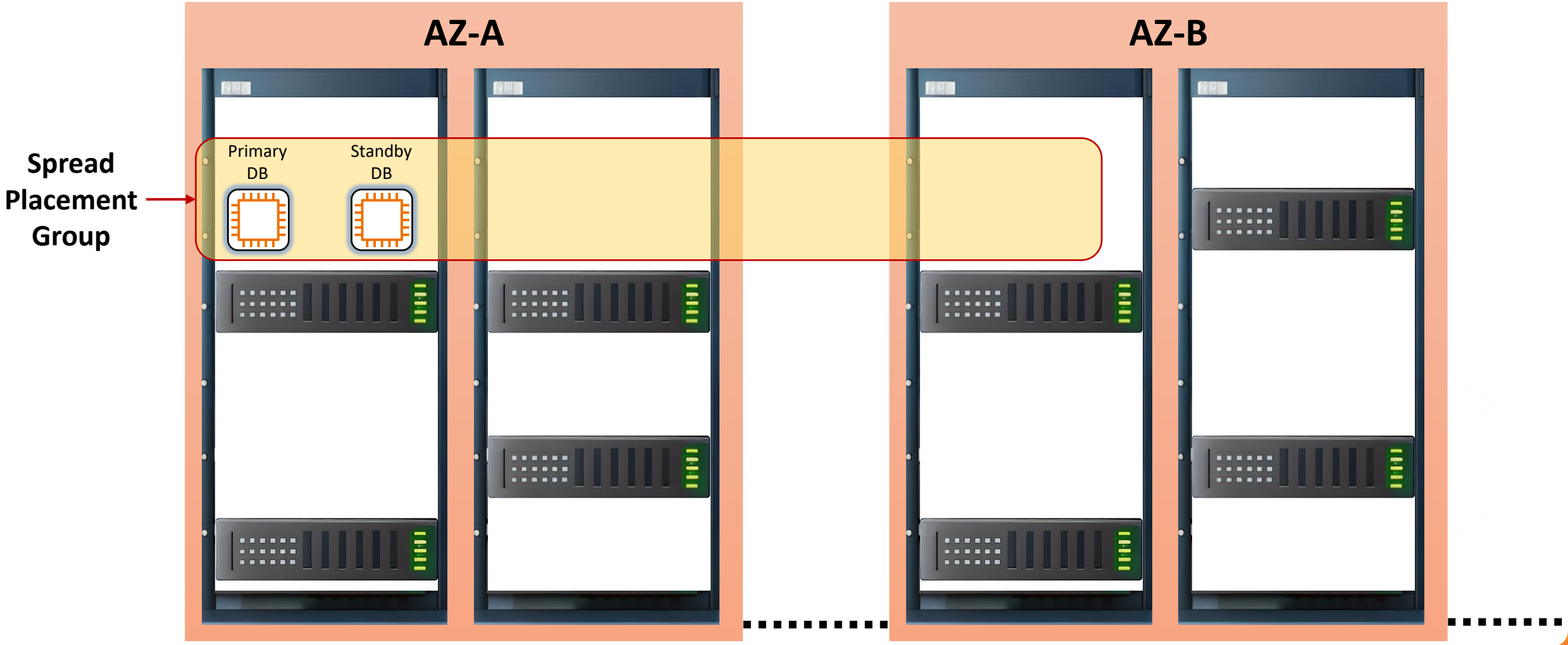
- Application Profile
  - Individual servers frequently communicate



TOR = Top of the rack switch

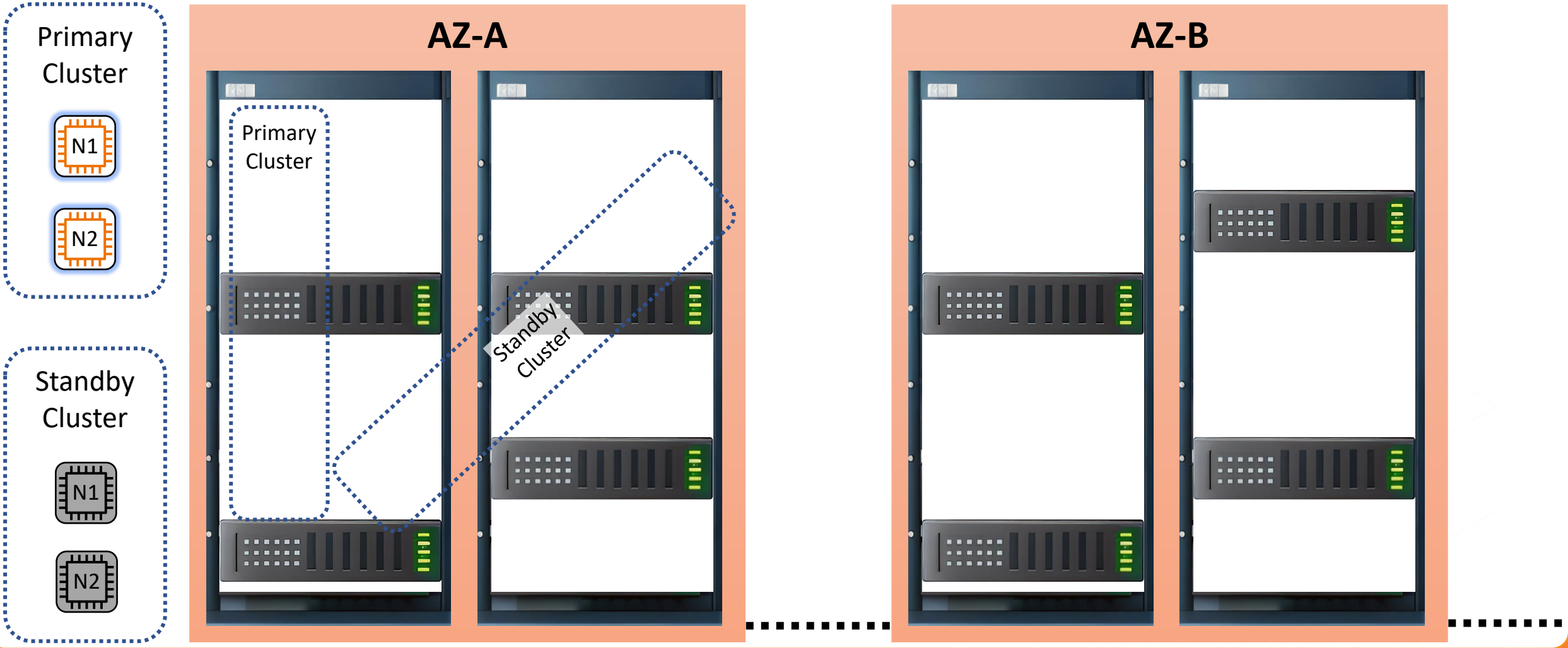
Spread Placement Group

- Application Profile
  - Critical Servers which are related but should not fail together



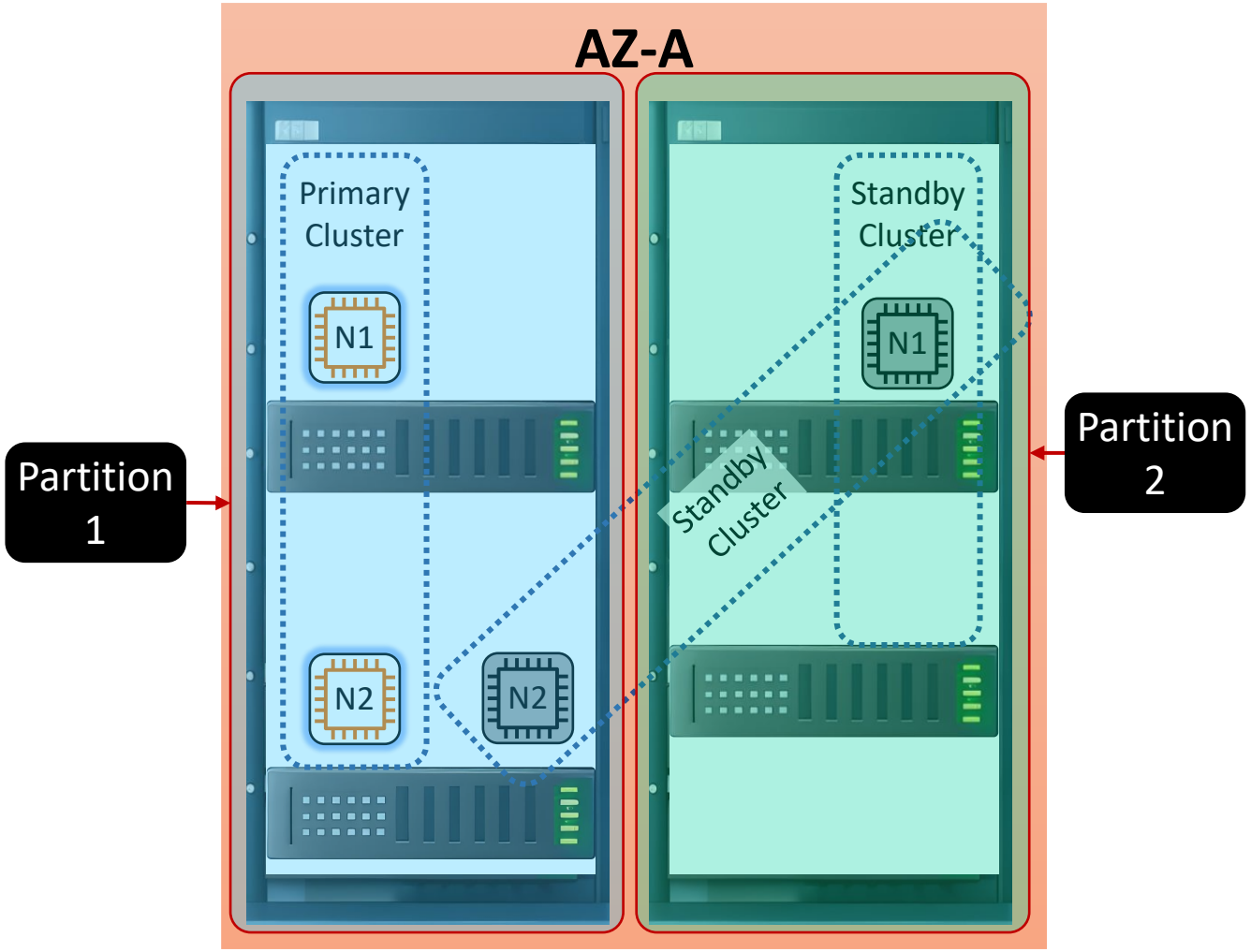
Partition Placement Group

- Application Profile
  - Group of critical servers which are related but should not fail together



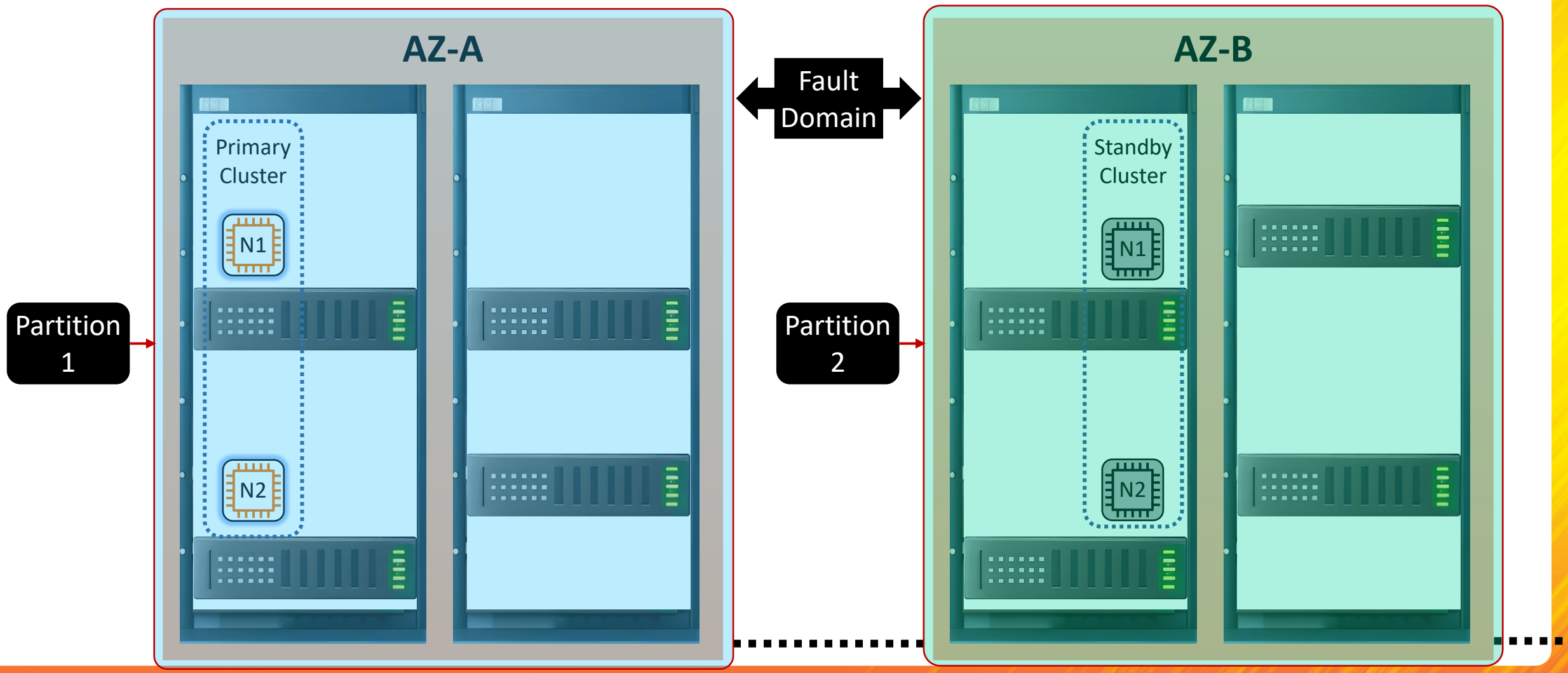
Partition Placement Group

- Application Profile
  - Group of critical servers which are related but should not fail together



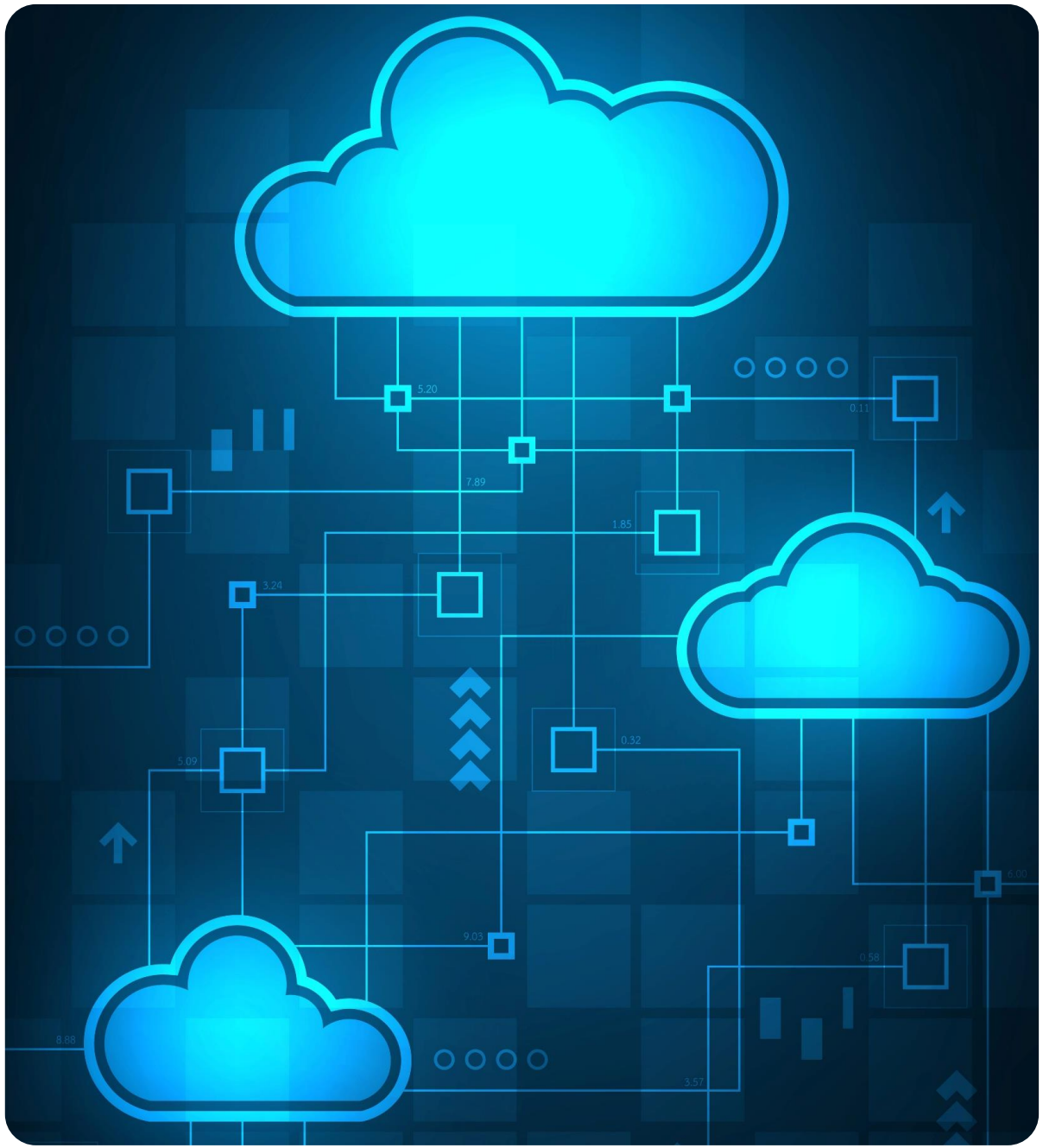
Partition Placement Group

- Application Profile
  - Group of critical servers which are related but should not fail together



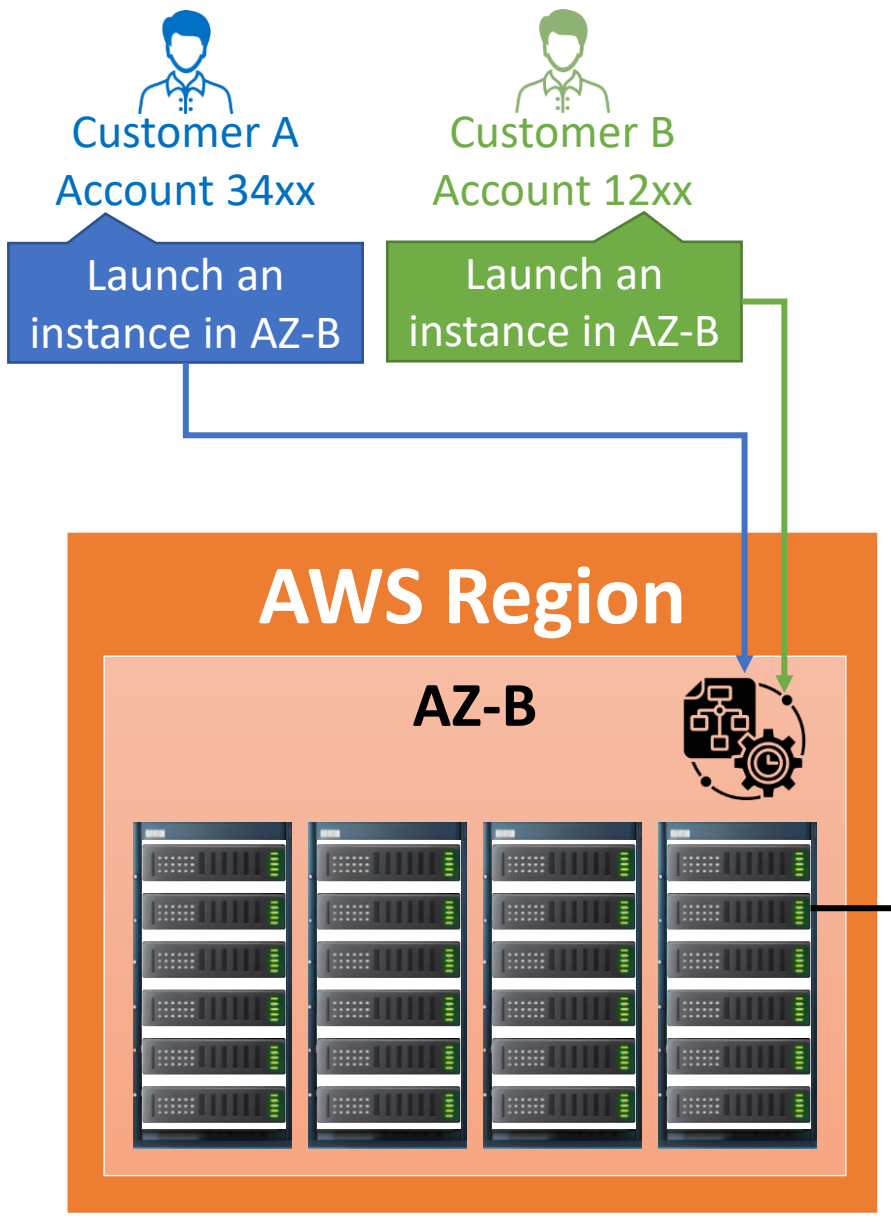
Comparison : Amazon EC2 Placement Groups

	Cluster Placement Group	Spread Placement Group	Partition Placement Group
Definition	Offers low-latency networking by placing instances as near as possible (possibly same server or same rack).	Spreads instances across underlying hardware to reduce the risk of simultaneous failures.	Similar to Spread, but groups instances into partitions, each set on distinct racks with their own network and power source.
Benefit	Network Performance	Fault Isolation	Fault Isolation and Network Performance
Can it span across multiple AZs?	No	Yes	Yes
Suitable for	HPC clusters, tightly-coupled distributed applications, weather modeling, computational fluid dynamics, etc.	Ideal for applications needing isolated hardware, critical applications that have a small number of critical instances.	Large distributed and replicated workloads like Hadoop, Cassandra, and other replicated workloads which supports multiple instances per partition.

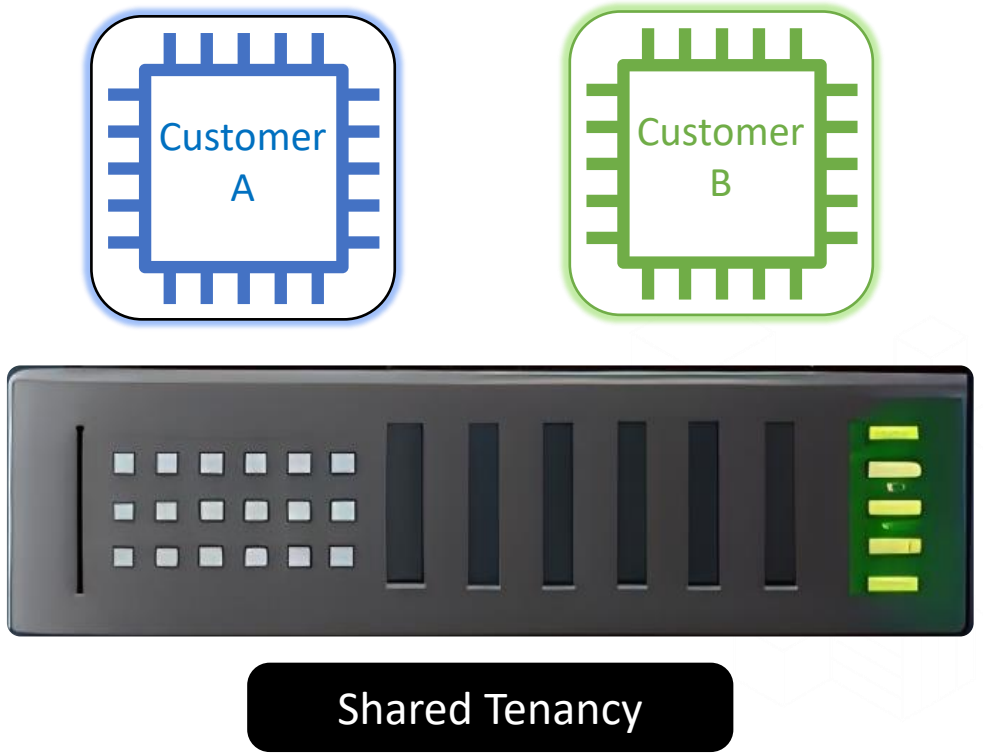


Dedicated Tenancy

Launching an Amazon EC2 Instance

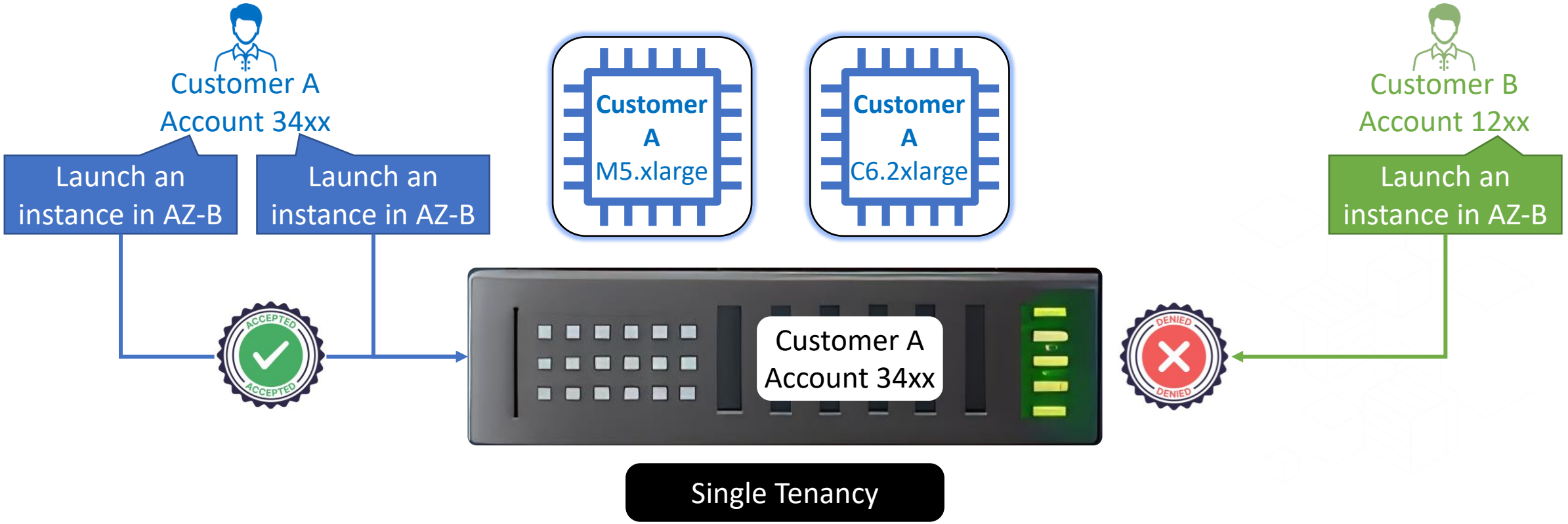


- By default, EC2 instances run on shared tenancy hardware. This means that multiple AWS accounts might share the same physical hardware.



Single Tenancy

- Single Tenancy (Dedicated Instance and Dedicated Host) allows you run your instance on a hardware which is dedicated for your AWS account.
- It can host only your instances, and no instances from other customers will be hosted on it.



Dedicated Instance vs. Dedicated Host

Dedicated Instance

Allocate me a host to run my instances.

A randomly selected host is allocated to run your instance.

If that host fails health checks your instances will be restarted on a similar (or higher) capacity host.

Dedicated Host

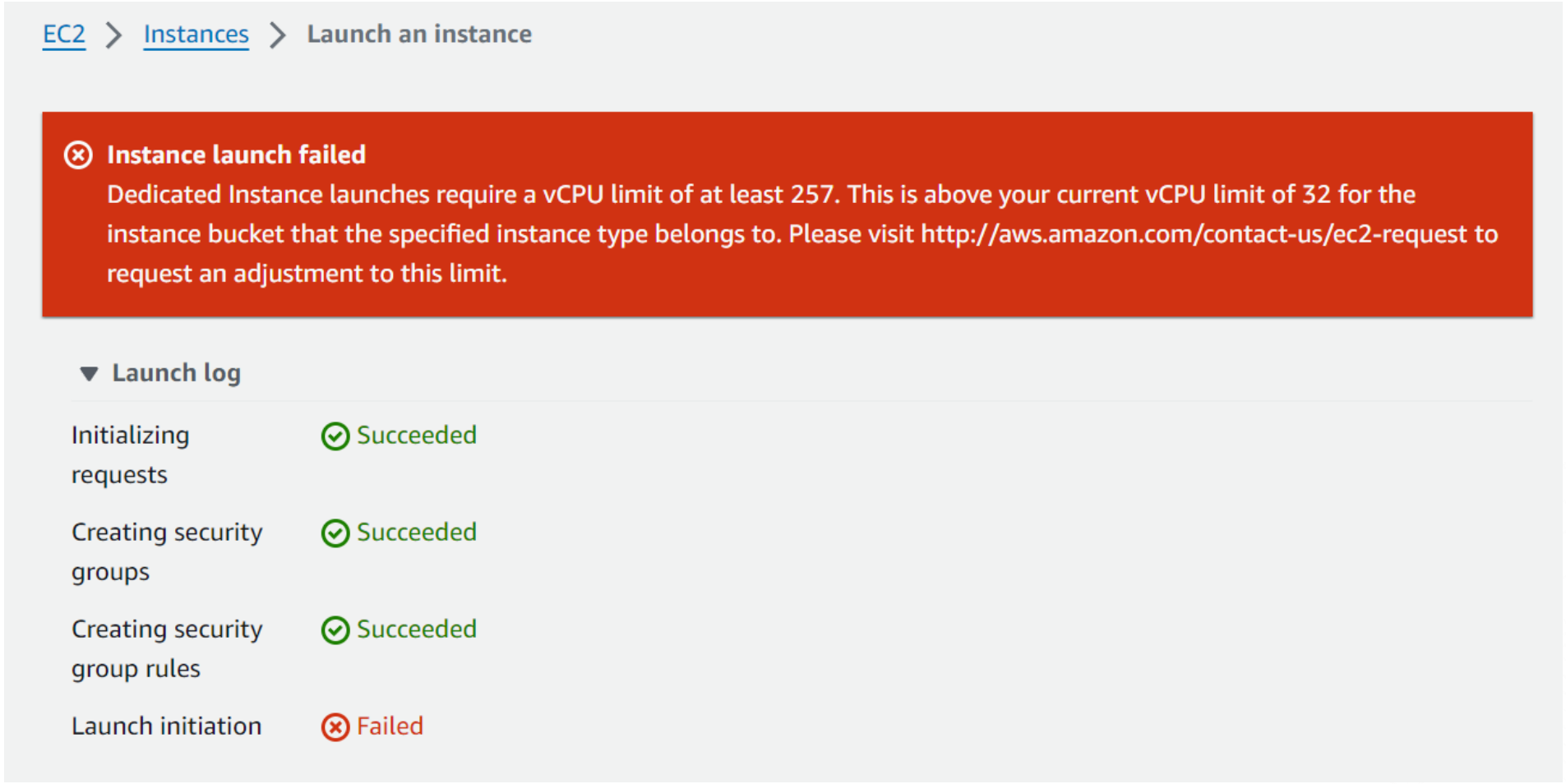
Allocate me the host to run my instances.

You select a host (sockets / core / host id) to run your instance.

If that host fails health checks, your instances will be restarted on another exact same host.

Account Limits

- You may have to get your vCPU limit increased in order to launch dedicated instance.



Dedicated Instance vs. Dedicated Host

		Dedicated Instance	Dedicated Host
	Use Case	To meet compliance and regulatory requirements	To address CPU / Socket / VM bound software licenses
Similarity	Dedicated physical server	Yes	Yes
	Automatic instance recovery	Yes	Yes
Difference	Charges	Only for the instance	For the entire host
	Bring Your Own License (BYOL)	Partial	Yes
	Visibility of sockets, cores, and host ID	No	Yes
	Targeted instance placement	No	Yes
	Consistently deploy your instances to the same physical server over time	No	Yes
	Can other accounts in the same billing family share the host where instances run	No	Yes

Analogy

Shared Tenancy

Staying in a dormitory

Room type

- 1 Bed in a 6-Bed Mixed Shared Dorm
  1 bunk bed
- Bed in 10 + Mixed Dormitory Room
  1 bunk bed

Dedicated Instance

Staying in a room in a hotel

Room type

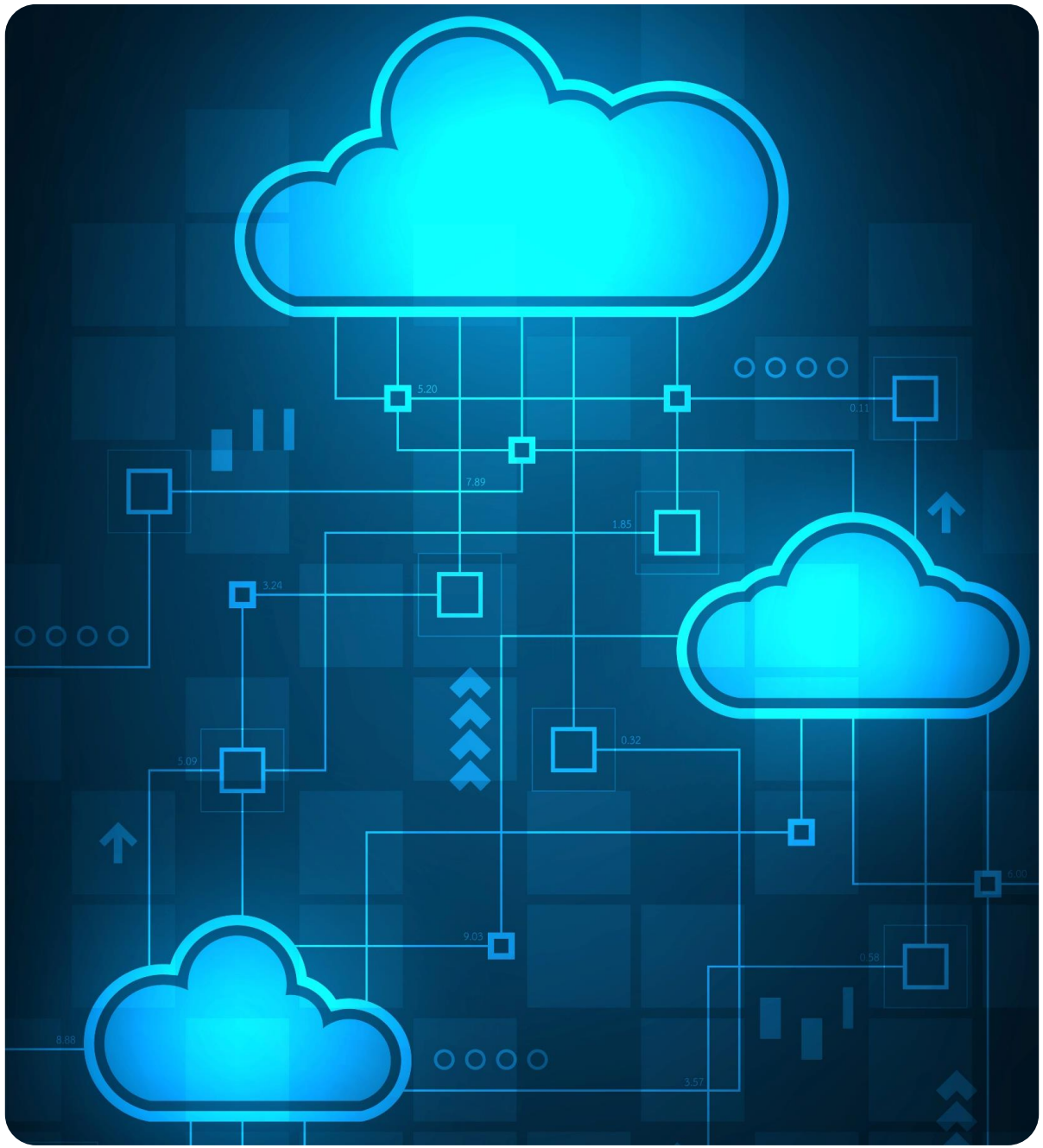
- Family Suite - Non-Smoking
  2 single beds and 1 large double bed
- Deluxe King Room - Non-Smoking
  1 extra-large double bed

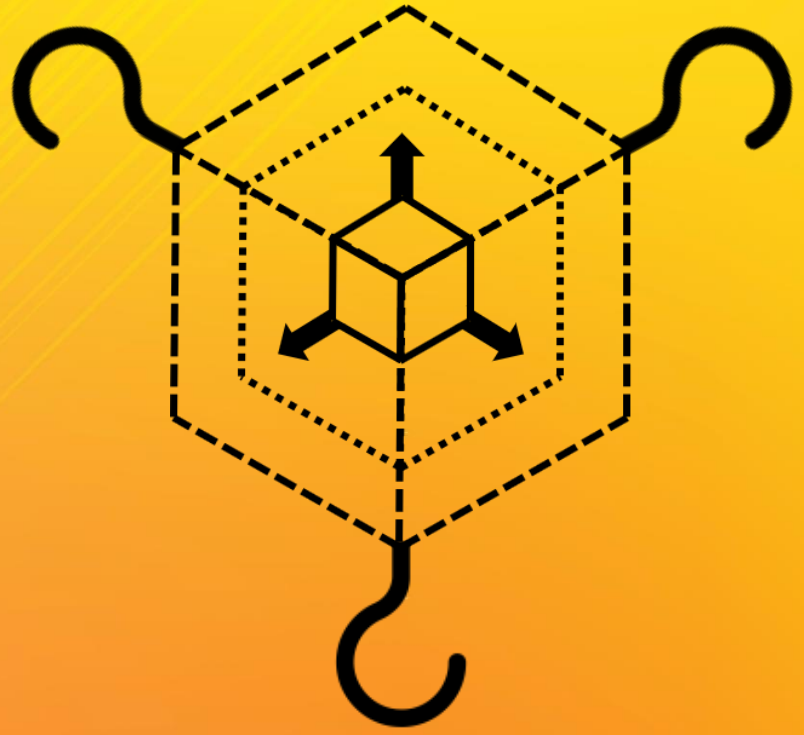
Dedicated Host

Staying in that specific holiday home

Accommodation Type

- Holiday Home
  Bedroom 1: 1 extra-large double bed
  Bedroom 2: 1 extra-large double bed
  Bedroom 3: 1 bunk bed
  Living room: 1 sofa bed



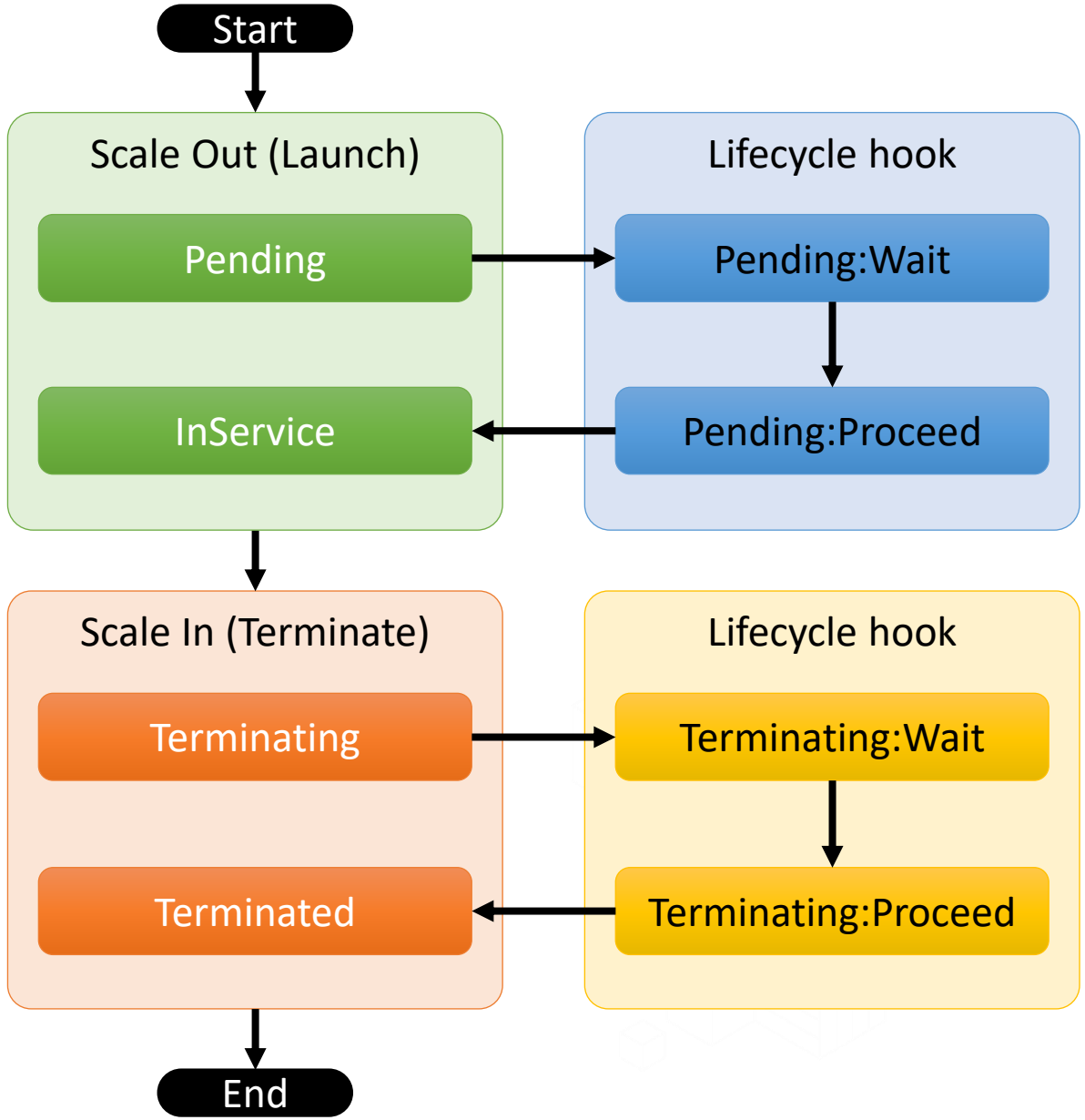


Amazon EC2

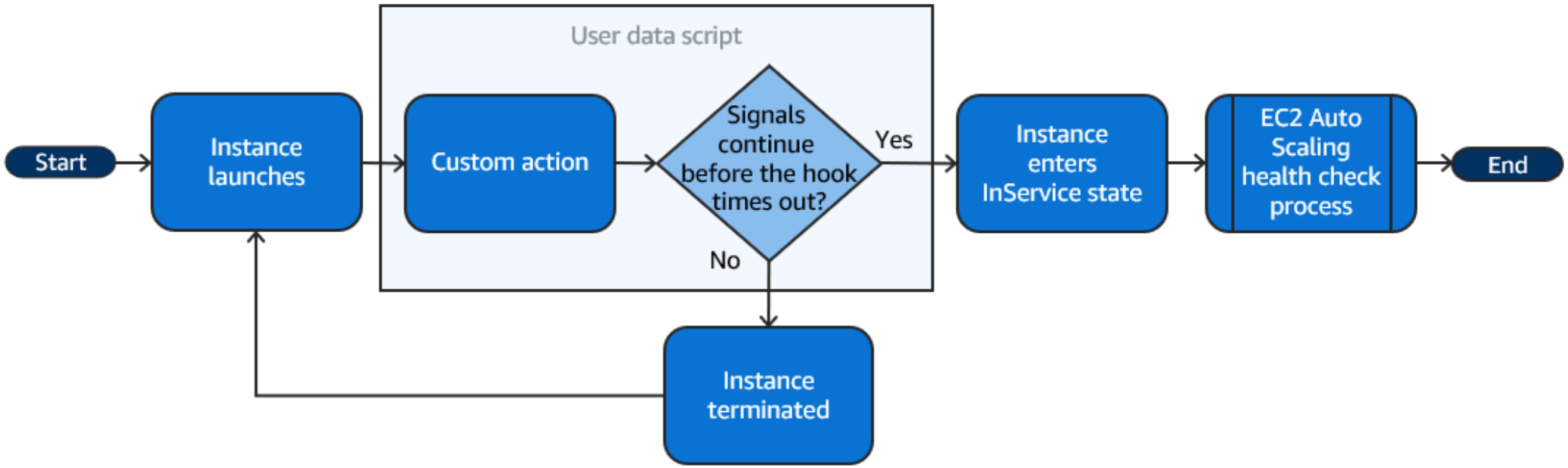
Auto Scaling lifecycle hooks

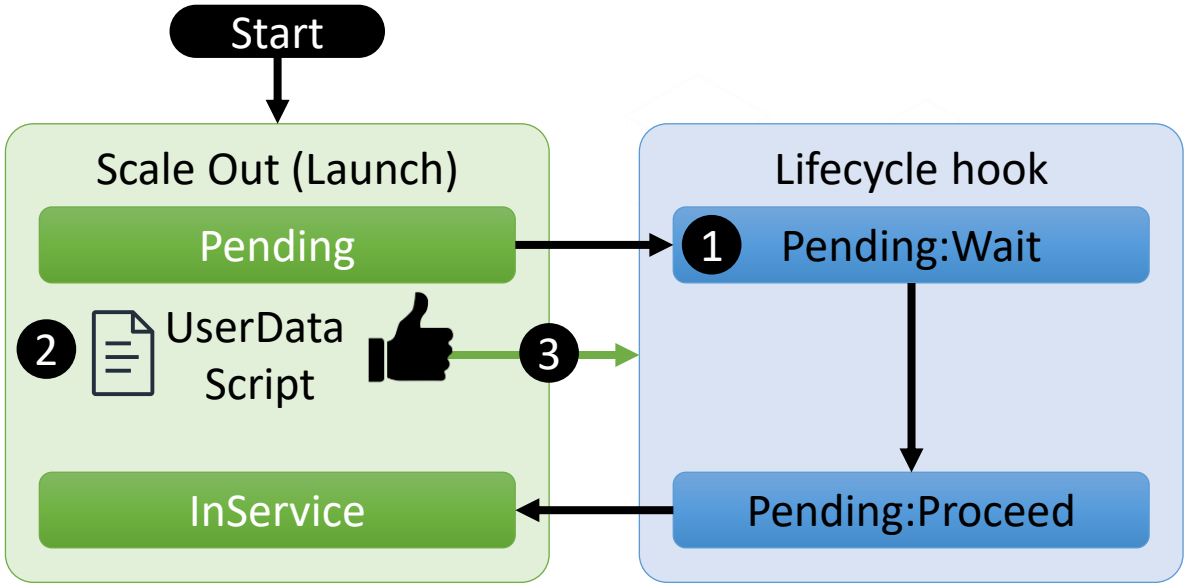
Amazon EC2 Auto Scaling - Lifecycle Hooks

- An Amazon EC2 instance transitions through different states from the time it launches until it is terminated.
- You can create custom actions for your Auto Scaling group to act when an instance transitions into a wait state due to a lifecycle hook.
- A lifecycle hook provides a specified amount of time (one hour by default) to wait for the action to complete before the instance transitions to the next state.

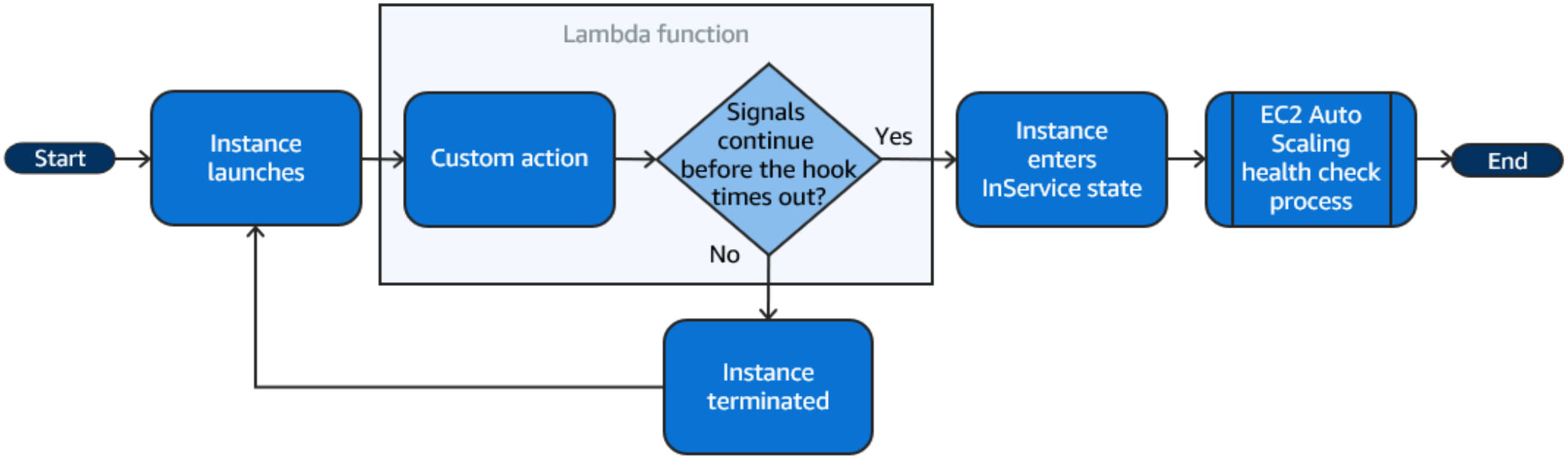


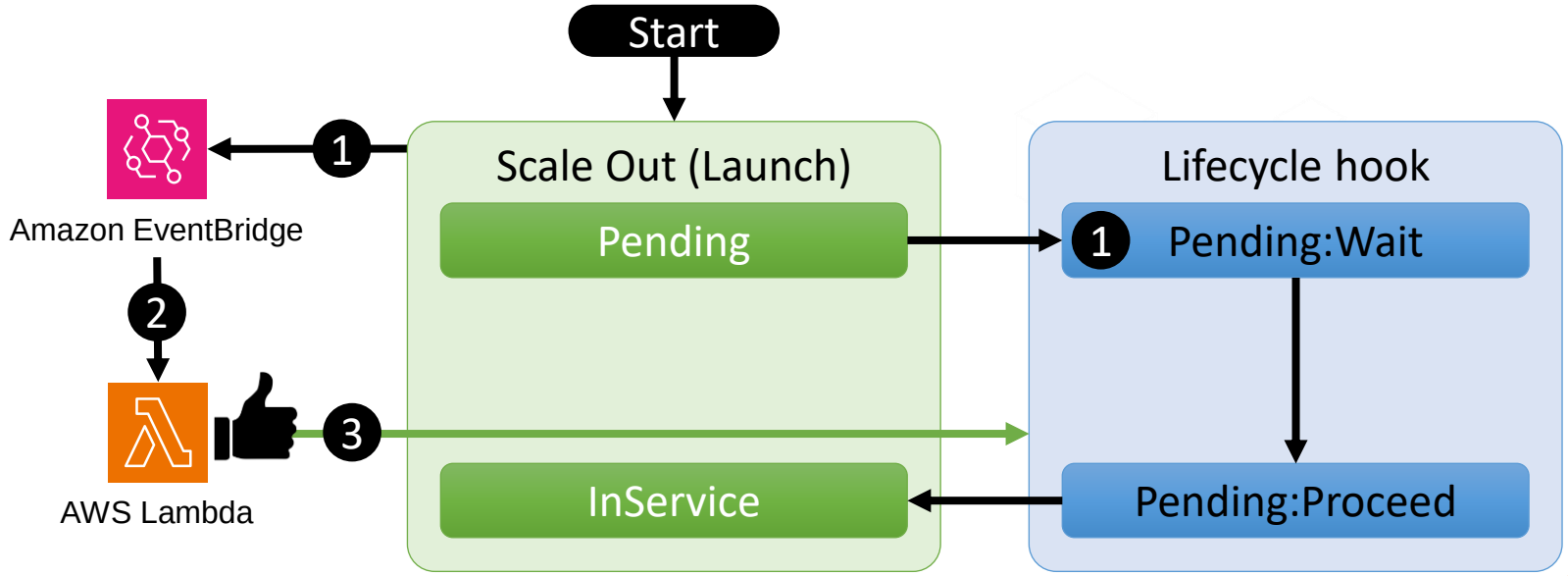
Amazon EC2 Auto Scaling - Lifecycle Hooks

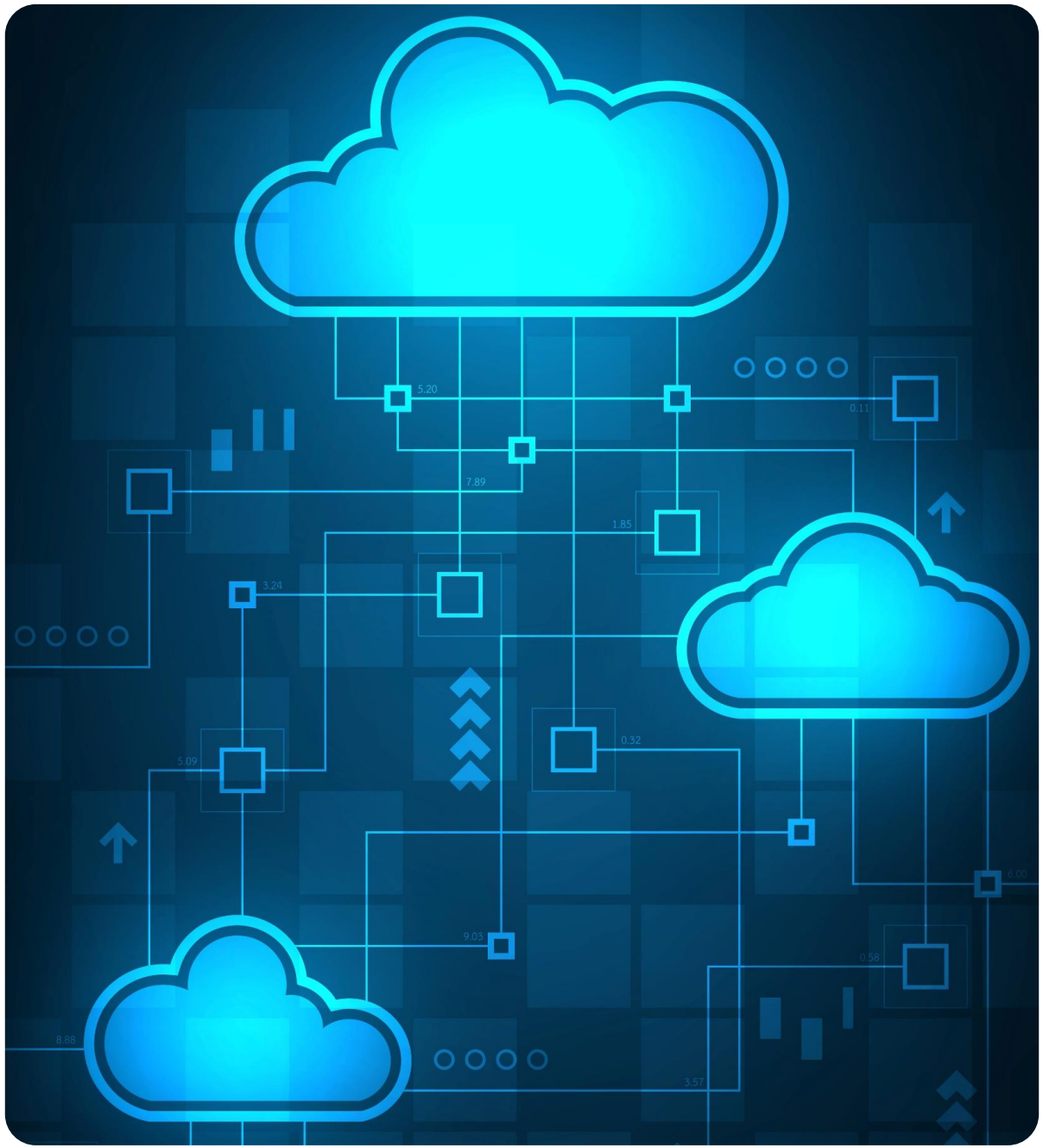




Amazon EC2 Auto Scaling - Lifecycle Hooks



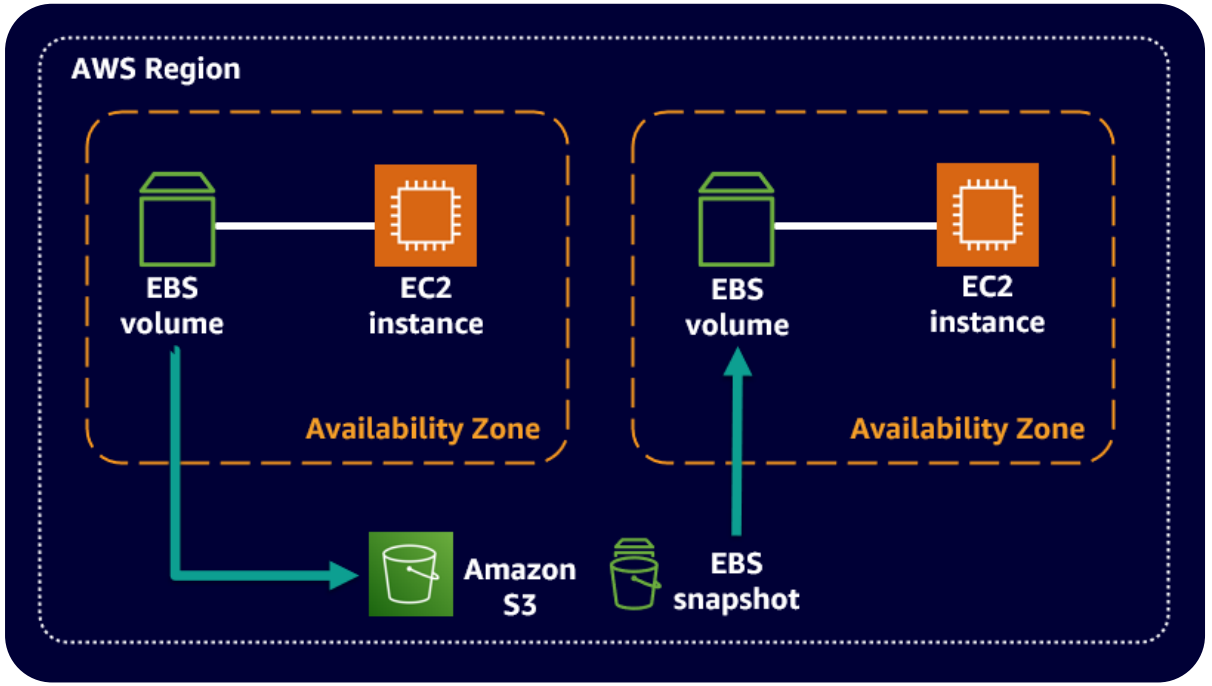




Amazon Data Lifecycle Manager (Amazon DLM)

Amazon Data Lifecycle Manager

- Amazon EBS snapshots
  - Point-in-time backup of modified volume blocks
  - Subsequent snapshots are incremental
  - Deleting snapshot only removes data exclusive to that snapshot



- Amazon Data Lifecycle Manager provides a simple, automated way to create, retain, copy and delete snapshots and AMIs.

Amazon Data Lifecycle Manager - Use cases

Automate snapshot of individual EBS Volumes	Automate snapshot of EBS Volumes connected to an Instance	Automate AMI creation from an Instance