
AWS
re:Invent

CMP311

AWS Wavelength: Run apps with ultra-low latency at 5G edge

Chris Barclay

Principal Product Manager, AWS Wavelength

AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Introduction to AWS Wavelength

Use cases

Developer experience

Q&A



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Introduction



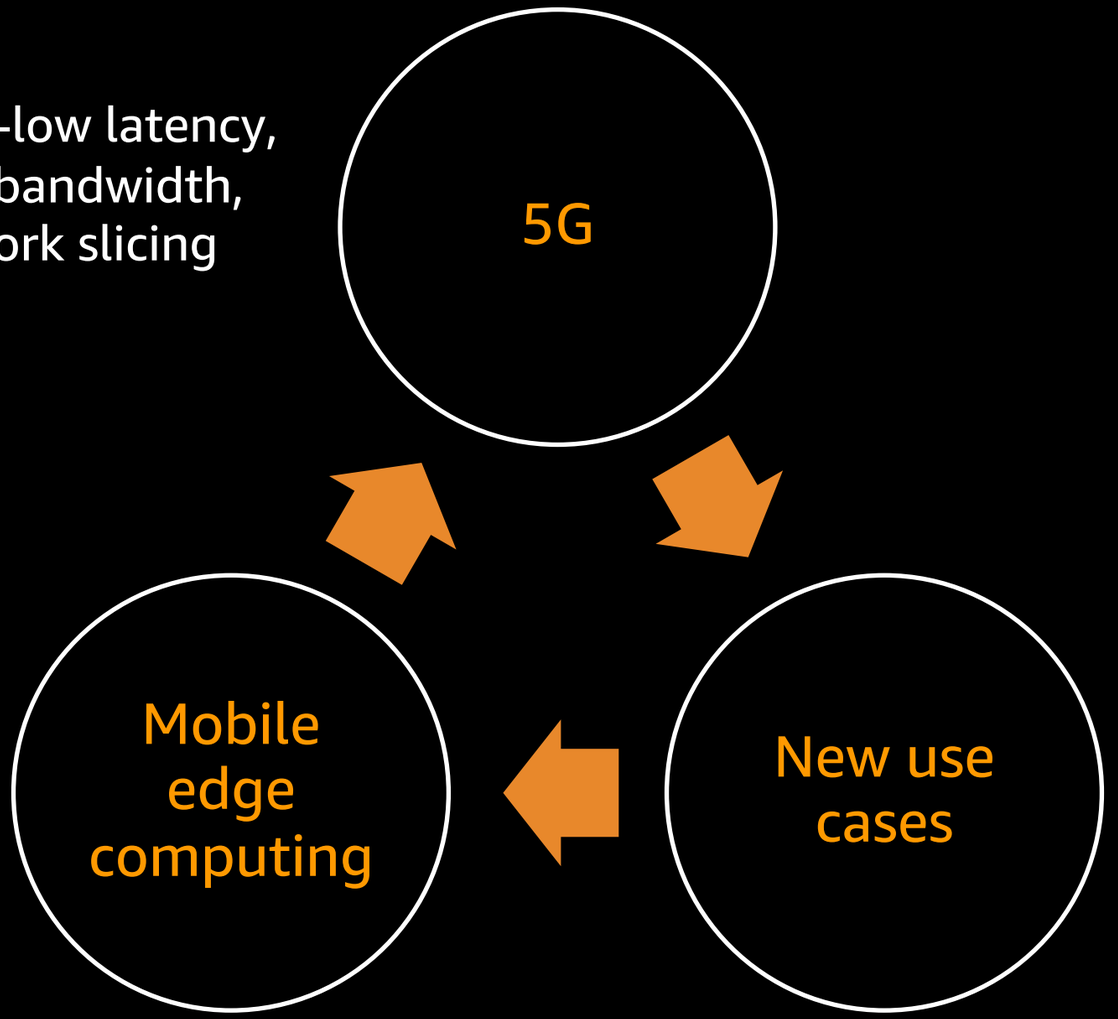
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

5G and mobile edge computing

MOBILE SERVICE DELIVERY MODEL THAT IS CONSUMER AND BUSINESS FOCUSED

Ultra-low latency, high bandwidth, network slicing



Mobile gaming, video streaming, augmented reality / virtual reality (AR / VR), Internet of Things (IoT), video recognition, Cellular Vehicle-to-Everything (C-V2X), etc.

Reduced network hops to deliver full value of 5G

Mobile edge computing: Characteristics

Massive number of devices such as sensors, video cameras, and IoT devices generating large volumes of data

Processing of data needs to happen close to data generation

In real time with very low latency

5G network will carry 35% of mobile traffic globally by 2024

Total global mobile traffic expected to reach 131 EB / month by 2024

Customer expectations for mobile edge infrastructure

- Ultra-low latency
- Cloud model
- Consistent programming interface
- Consistent operations
- New features

AWS for the edge



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS for the Edge

AWS OFFERS THE MOST COMPREHENSIVE PLATFORM FOR BUILDING AND DEPLOYING EDGE APPLICATIONS

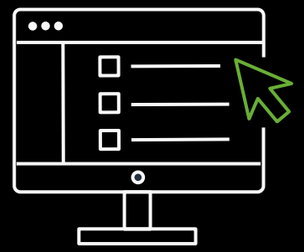
	AWS Outposts	AWS local zones	AWS Wavelength
			
Overview	AWS infrastructure and services on premises	AWS infrastructure and services in large metro centers	AWS infrastructure and services in CSP 5G networks
Use cases	Migration, local critical applications, data residency	Migration, low latency, local data processing	Ultra-low latency, local data processing
Service model	Expandable capacity in customer's data center, collocated, on-premises location	Scalable capacity in AWS-managed and operated facility	Scalable capacity in CSP data center managed and supported by AWS

AWS Wavelength: Built for the 5G mobile edge

AWS services from inside the communications service provider (CSP) 5G mobile network



AWS compute and storage infrastructure embedded inside CSP mobile network



Single pane of management, across Wavelength Zone and AWS Regions



Access to services in the AWS Region



Develop applications once and deploy for use with 5G network globally

AWS Wavelength availability

Growing list of CSP collaborators

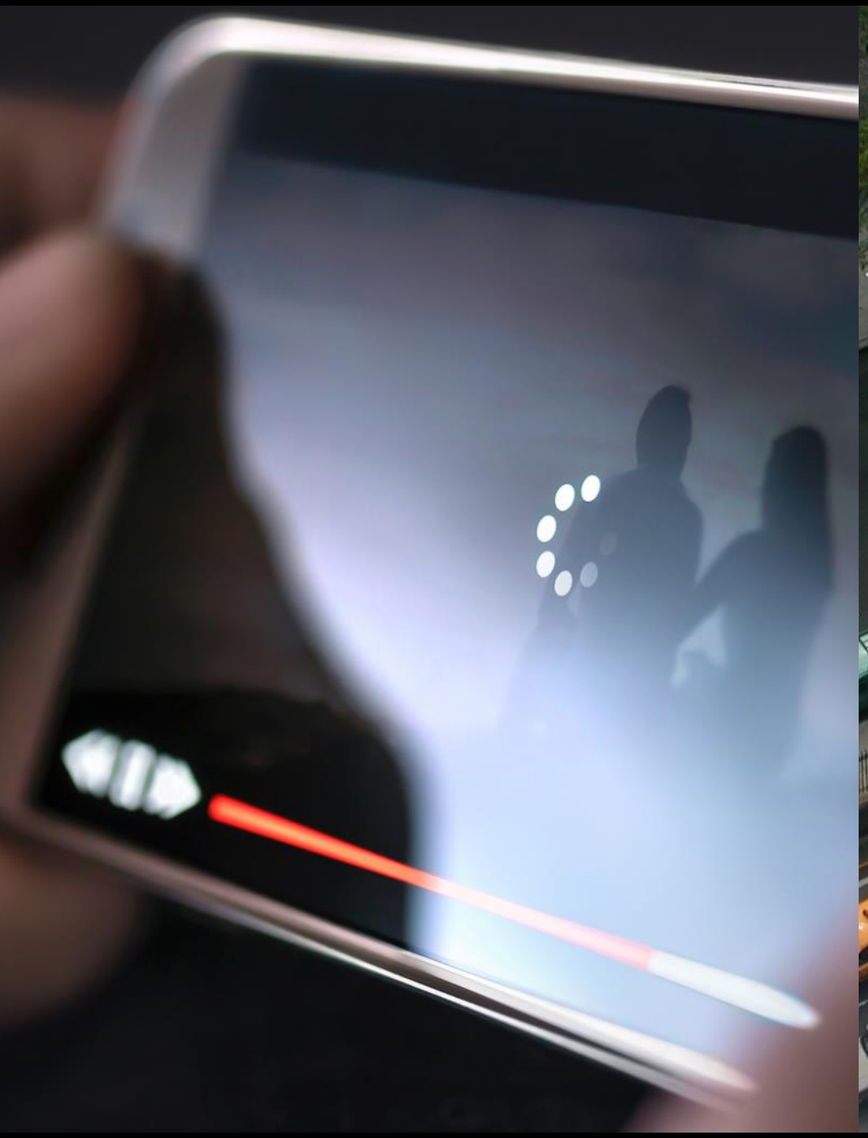
verizon vodafone Tomorrow, Together KDDI SK telecom

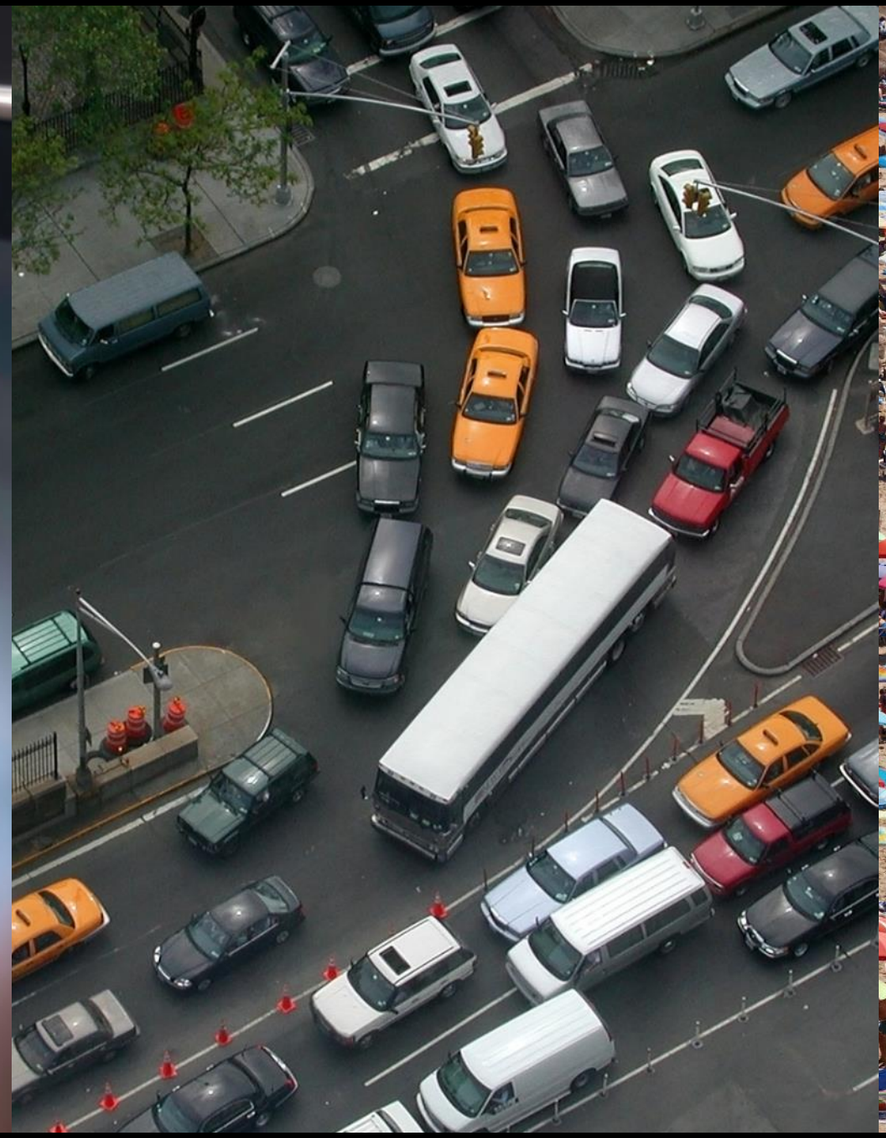


GA at BOS, NYC, ATL, WAS, SFO with Verizon

Wavelength expansions to new locations across US and worldwide in 2020 and 2021

Why AWS Wavelength: Latency







Ultra-low latency

"AWS Wavelength and the 5G network provide Sony the ultra-low latency needed to live stream high-resolution video and high-fidelity audio, as well as to embed interactive experiences into live video streams."

Hisashi Tamai
Senior Vice President of System and Platform Technology Development, Fundamental Technology Research and Development, R&D Center
Sony



Edge data processing and inference

"With critical GPU workloads rendered from AWS Wavelength coupled with 5G's bandwidth, we can achieve the low latency and increased connectivity to deliver a pathway to truly connected factories, where data flows seamlessly between machines and cloud to enable AI powered real-time decisioning at the edge."

Shanky Viswanathan
CTO Communication,
Media and Information Service Business Group
Tata Consultancy Services



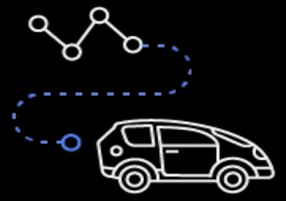
AWS Wavelength use cases

ML-assisted diagnostics for Healthcare



Inference at the edge for processing and analyzing video, images, and data for real-time diagnosis

Connected vehicles



Real-time monitoring of data from sensors for road safety, secure connectivity, and in-car telematics

Smart factories



Accelerating the industrial edge with AI / ML, video analytics for software-defined manufacturing

Real-time gaming



Cloud gaming streaming for high-quality, interactive gaming with very low latency

AR / VR / XR



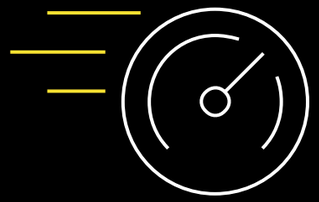
Reduce MTP latencies to offer realistic experiences across a wide range of applications

Interactive live video streams



Embed interactive experiences into live video streams

AWS Wavelength benefits



Low latency / high bandwidth

Wavelength brings services to the edge of the 5G network



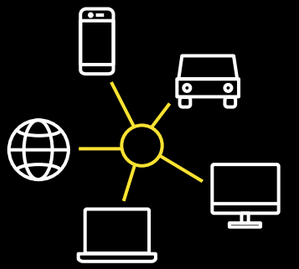
Consistent development experience

Delivering a consistent developer experience across multiple 5G networks around the world



Same AWS benefits

Including no capital costs, pay-as-you-go pricing, and the ability to scale up and down on demand

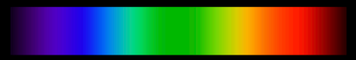


Ubiquity

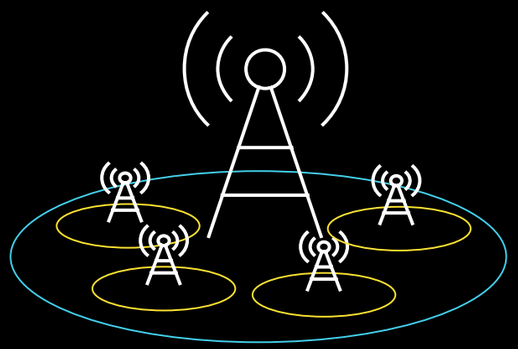
The same benefits whether end users connect over the internet or mobile networks, in the Americas, Europe, Japan, Korea, and more

5G advancements

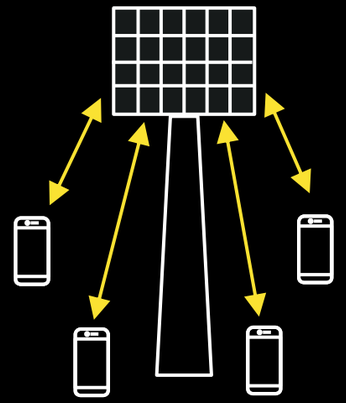
1. Advanced wireless technologies



mm Wave Spectrum



Small cell topology



MIMO antennas

2. Virtualization and control of network components

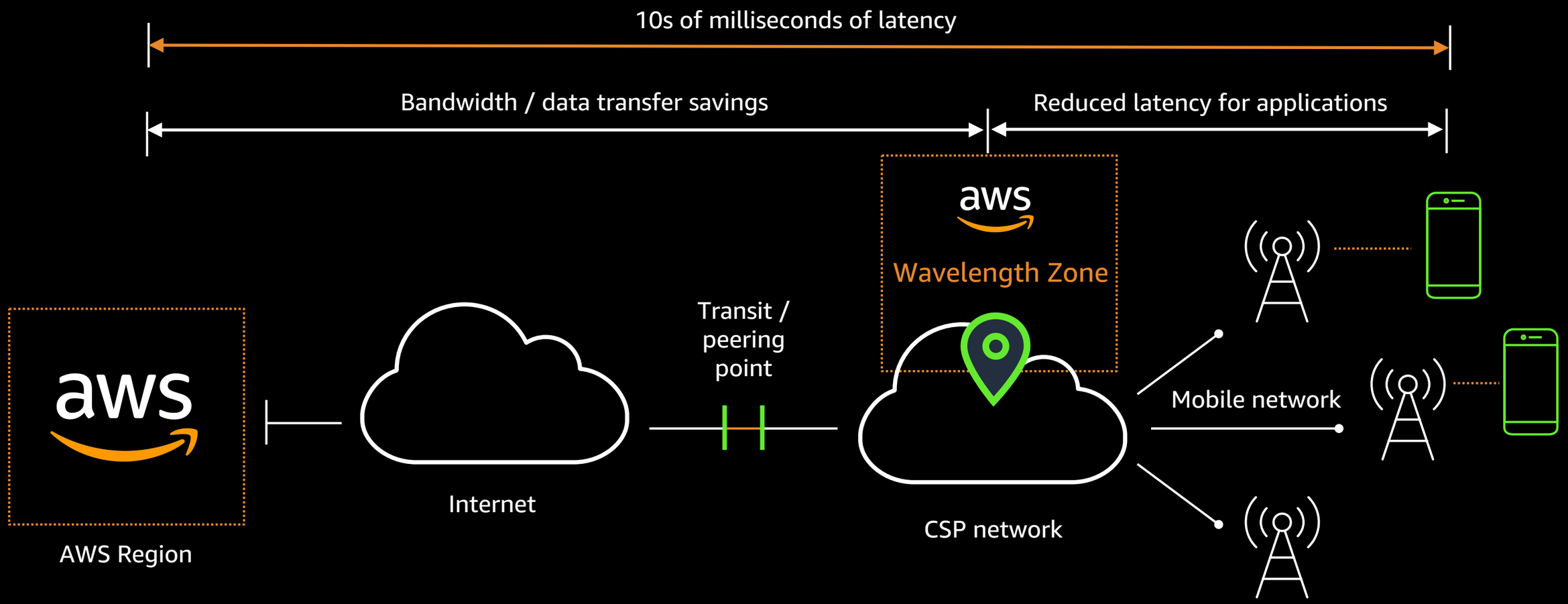


Dedicated HW appliances



Applications on servers

AWS Wavelength in a 5G network



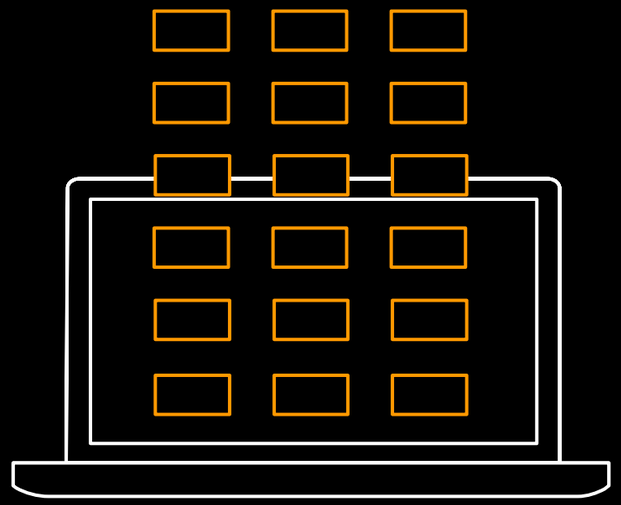
Developer experience



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Built on familiar AWS Services

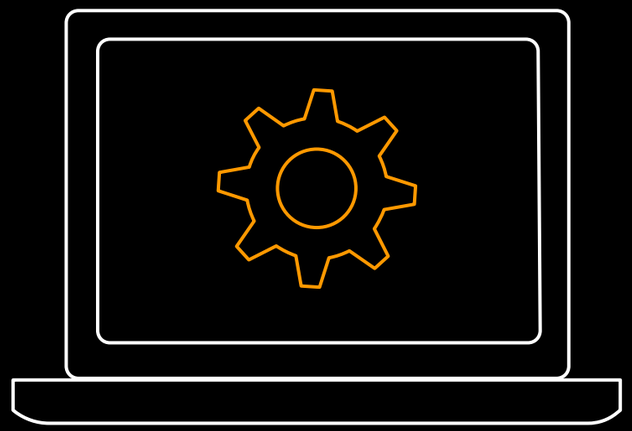


Compute and storage – Amazon Elastic Compute Cloud (Amazon EC2) instances and Amazon Elastic Block Store (Amazon EBS) volumes

Networking – Amazon Virtual Private Cloud (Amazon VPC)

Containers – Amazon Elastic Container Service (ECS) and Amazon Elastic Kubernetes Service (EKS)

With the same AWS APIs and tools as in the AWS Region



Amazon EC2 Auto Scaling groups

AWS CloudFormation

Amazon CloudWatch

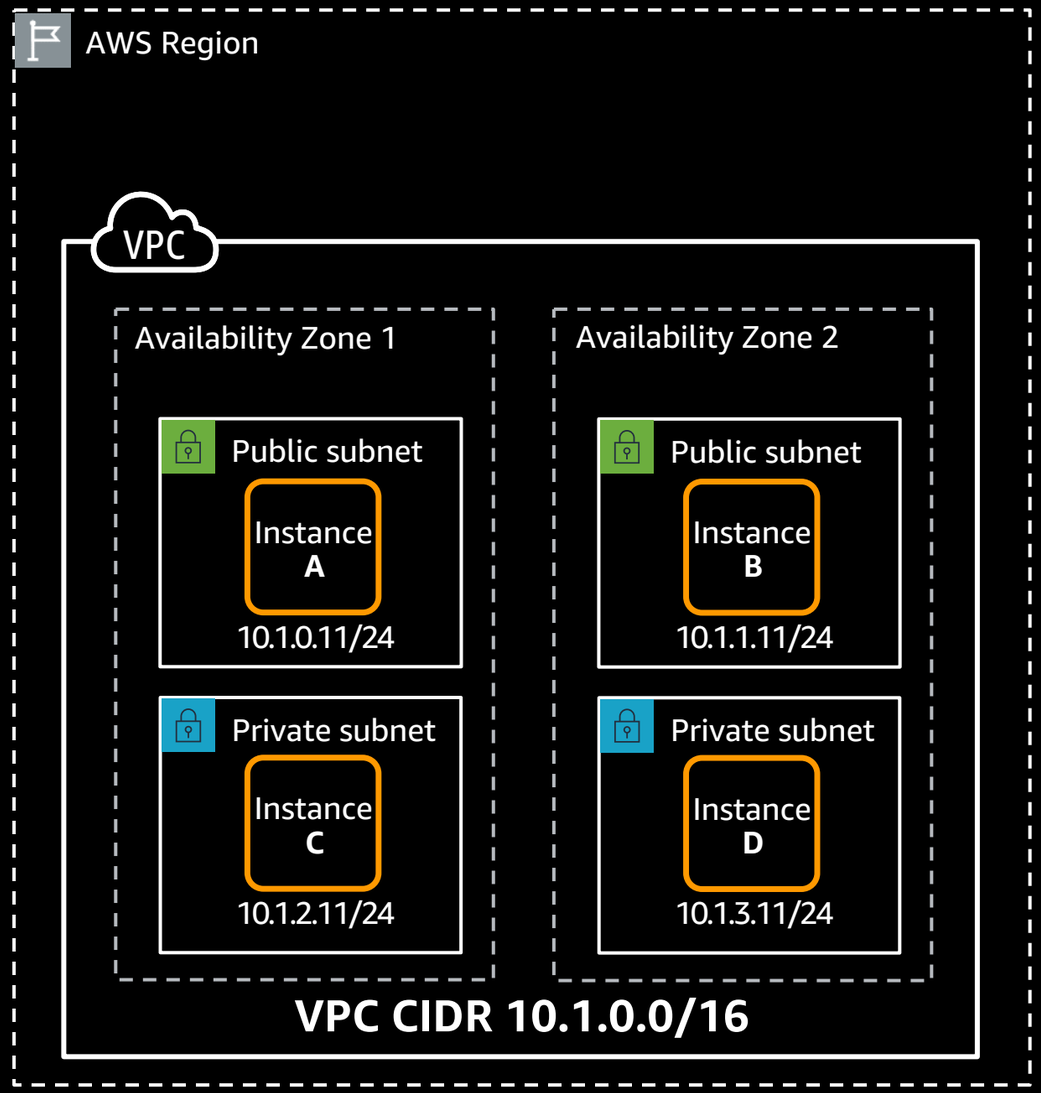
AWS CloudTrail

And more

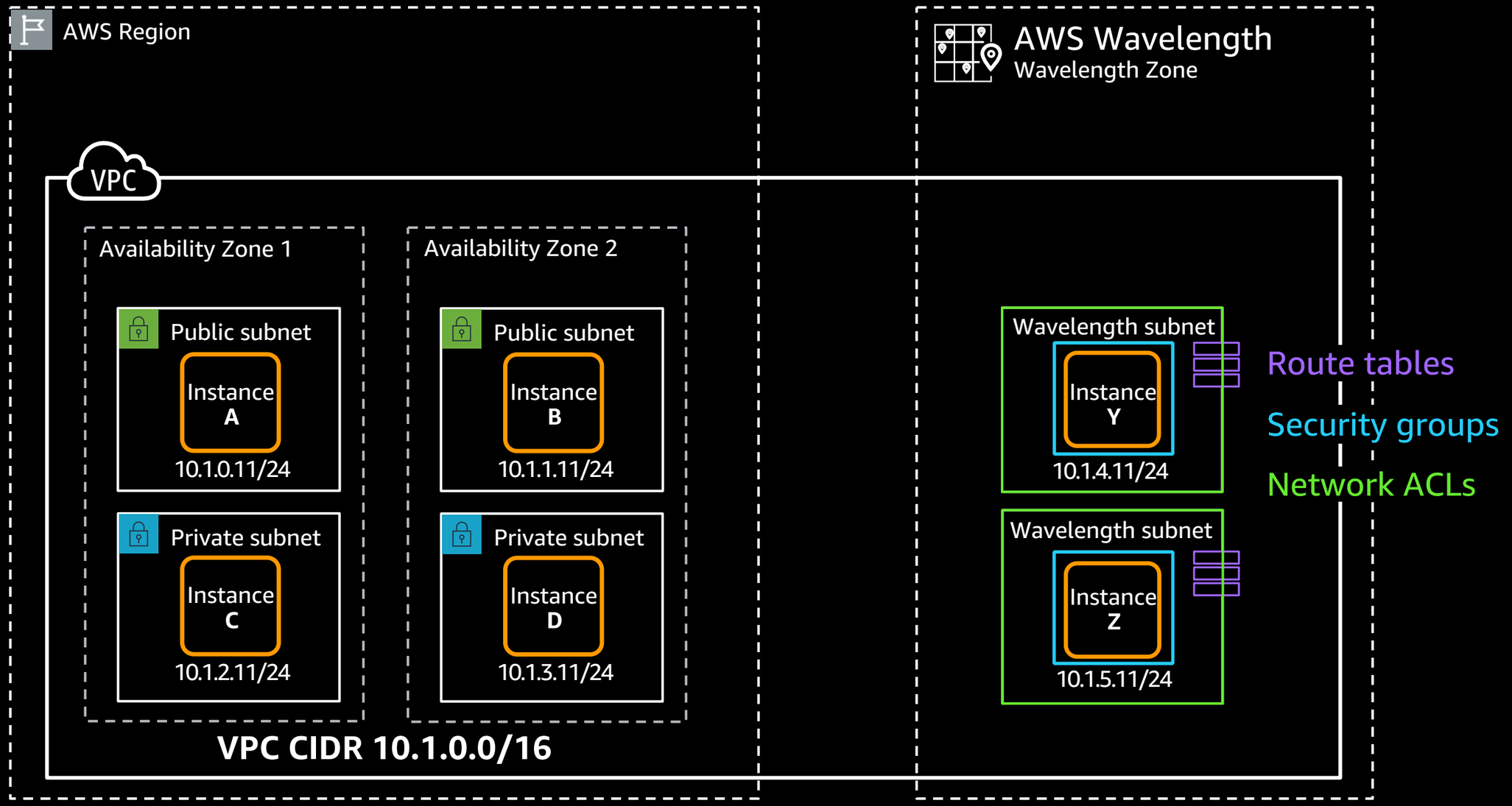
AWS Wavelength concepts

```
$ aws ec2 --region us-east-1 describe-availability-Zones
{
    "AvailabilityZones": [
--snip--
        {
            "State": "available",
            "OptInStatus": "opted-in",
            "RegionName": "us-east-1",
            "ZoneName": "us-east-1-wl1-bos-wlz-1",
            "ZoneId": "use1-wl1-bos-wlz1",
            "GroupName": "us-east-1-wl1",
            "NetworkBorderGroup": "us-east-1-wl1-bos-wlz-1",
            "ZoneType": "wavelength-Zone",
            "ParentZoneName": "us-east-1a",
            "ParentZoneId": "use1-az4"
        }
    ]
}
```

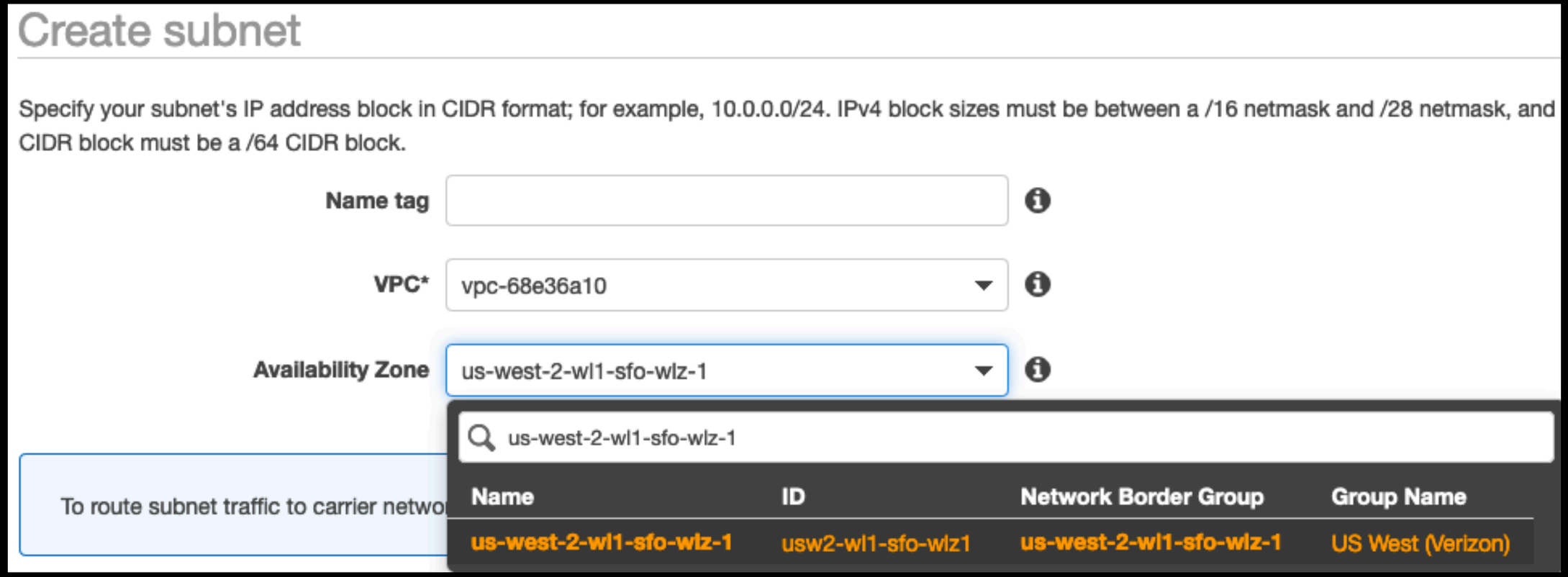
AWS Wavelength networks



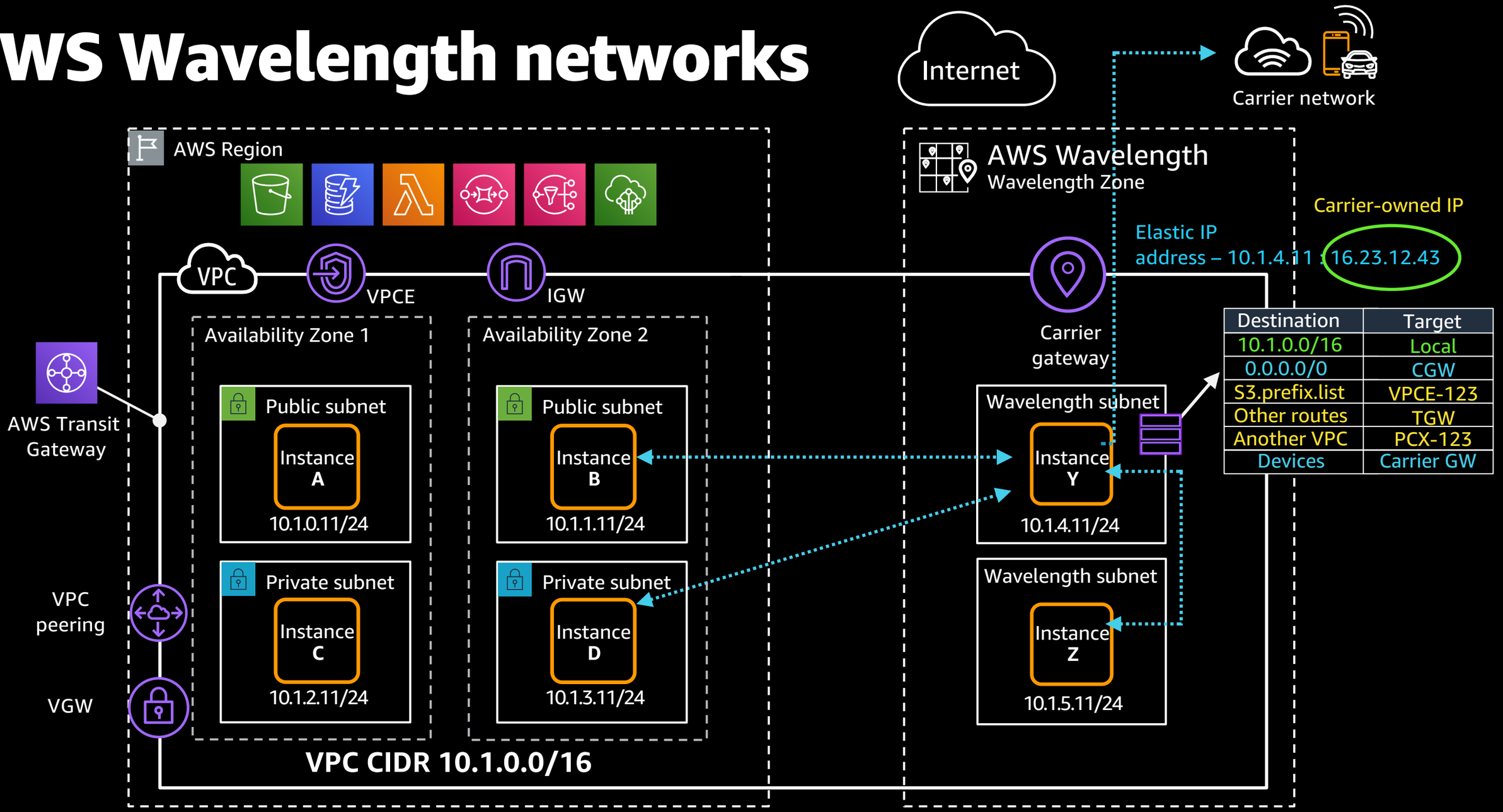
AWS Wavelength networks



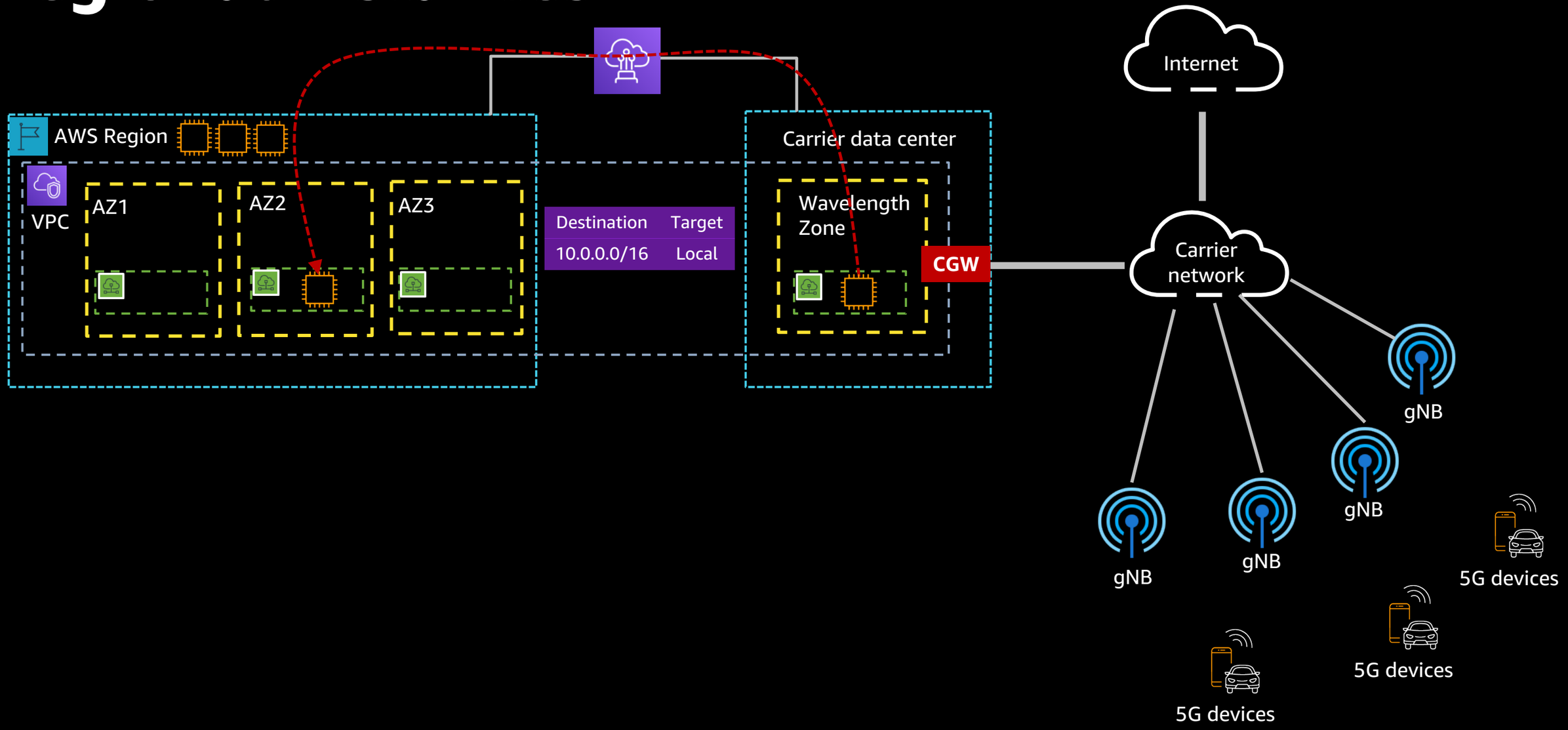
AWS Wavelength networks



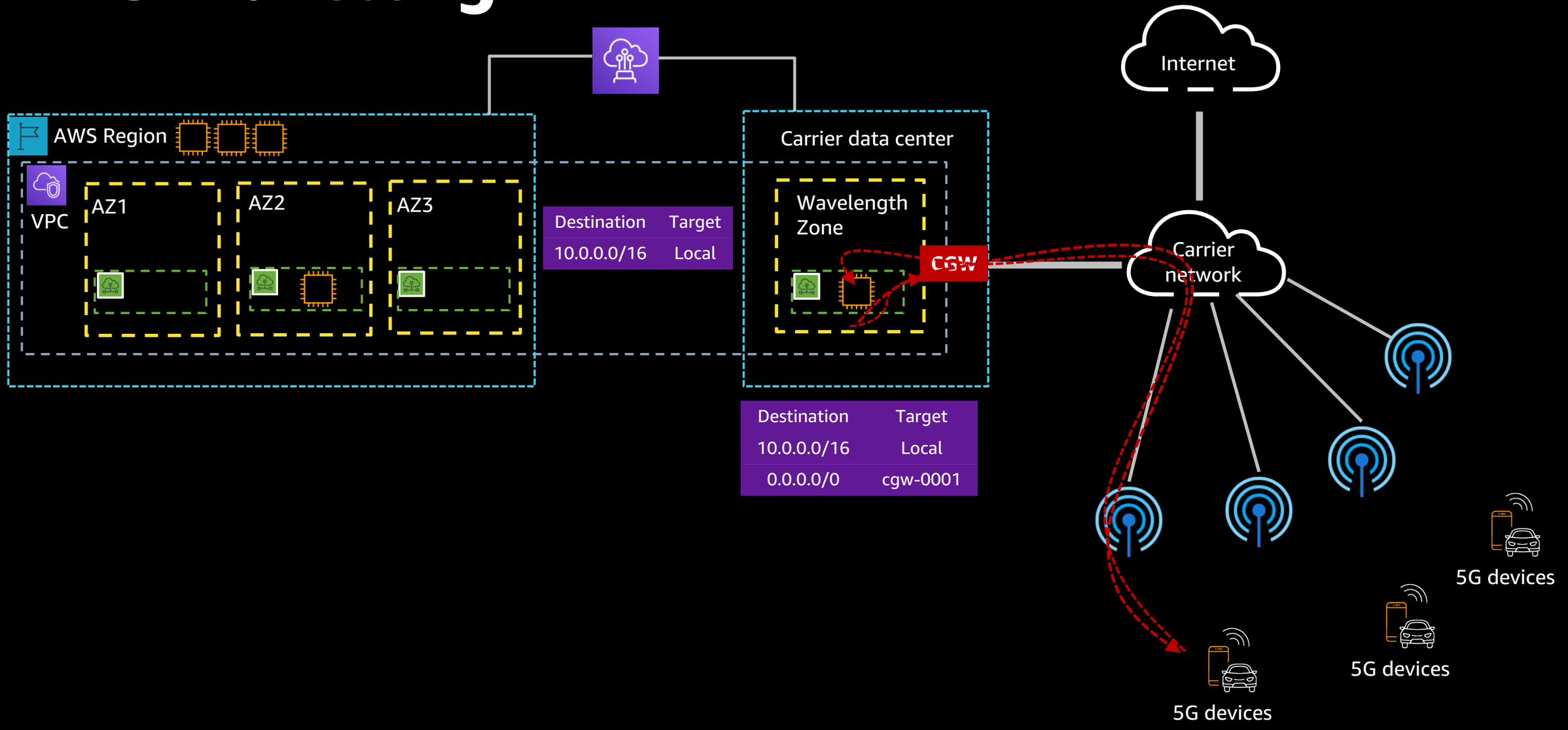
AWS Wavelength networks



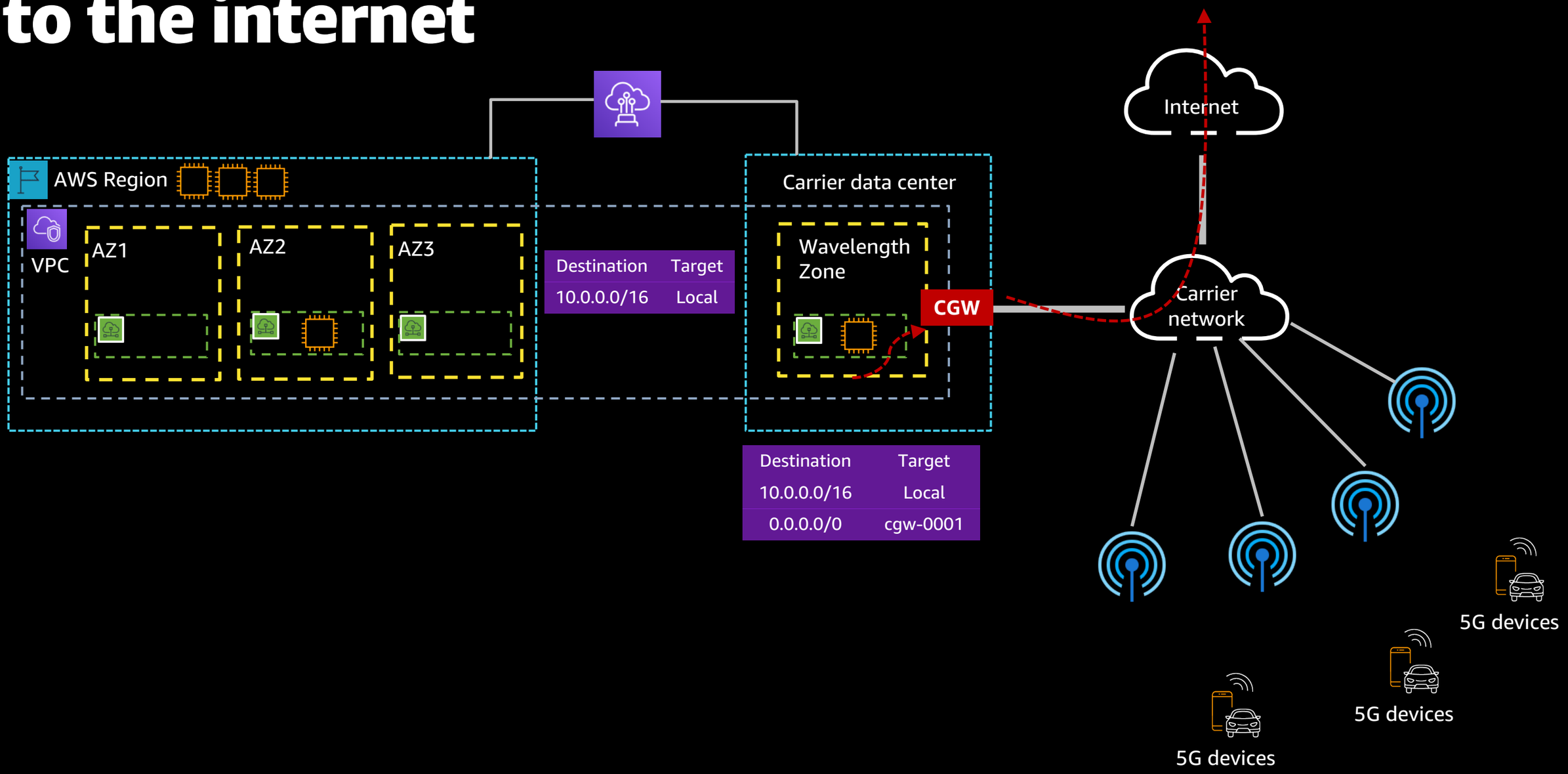
Packet flow between AWS Wavelength and regional instances



Packet flow between devices and AWS Wavelength



Packet flow from Wavelength Zone to the internet



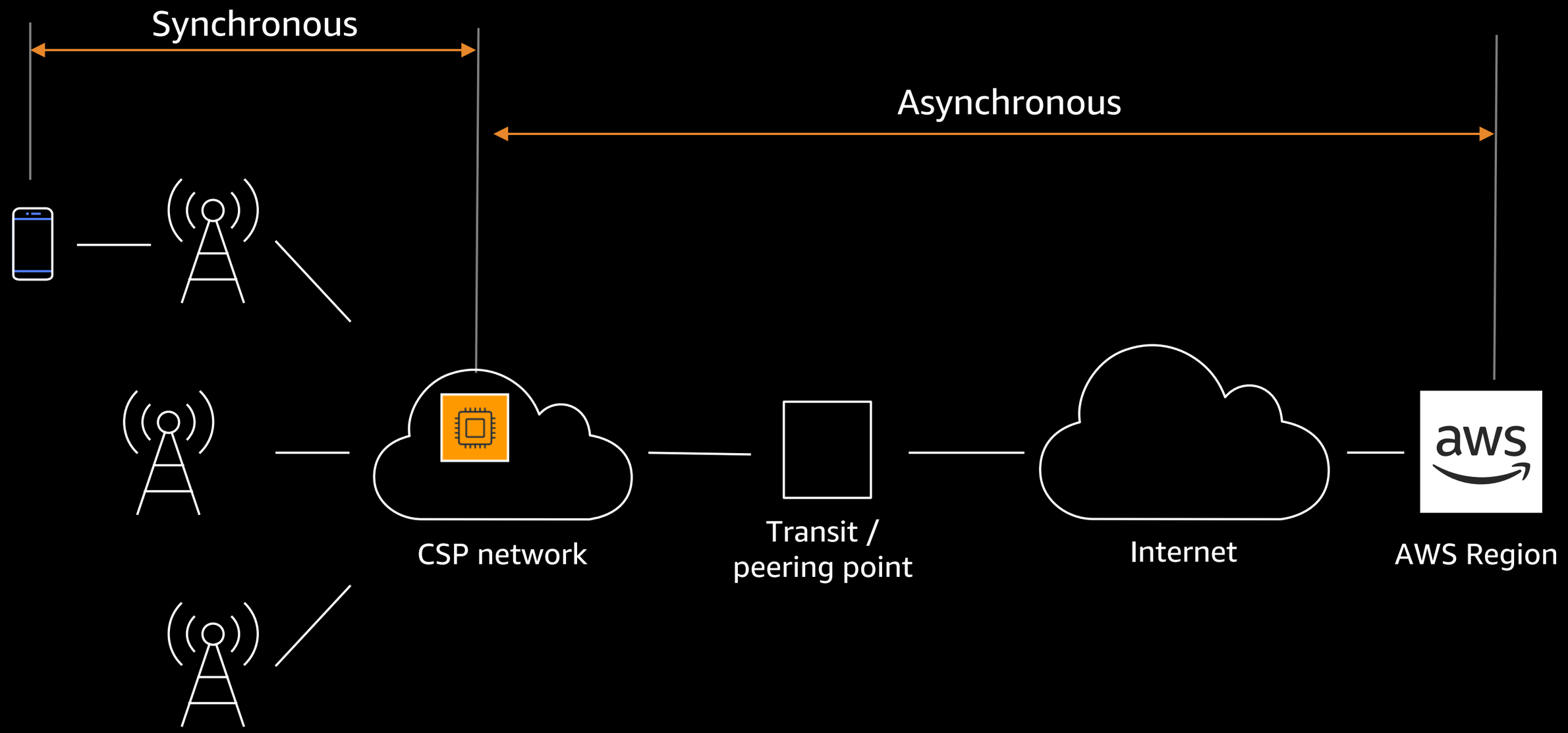
Architecture and provisioning



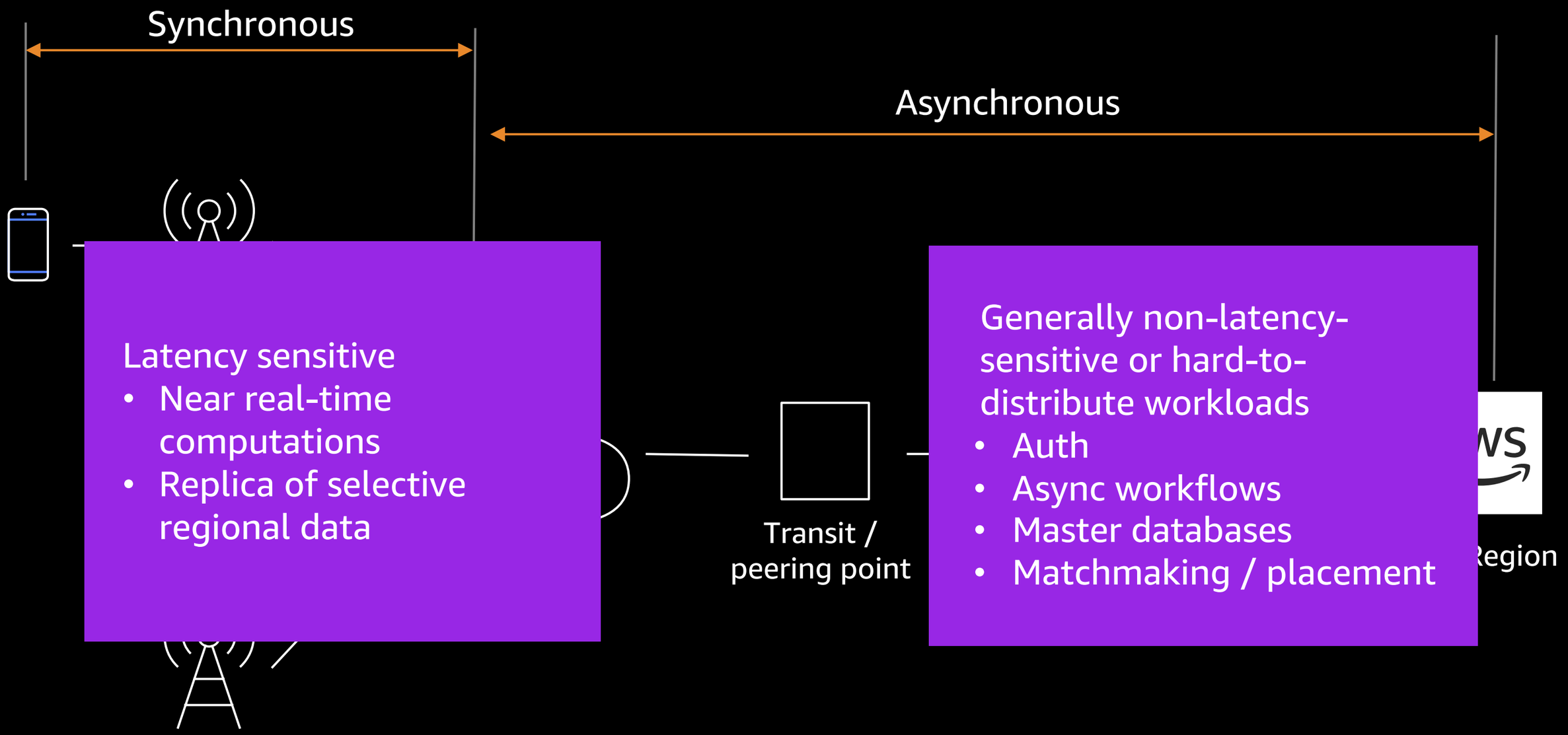
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

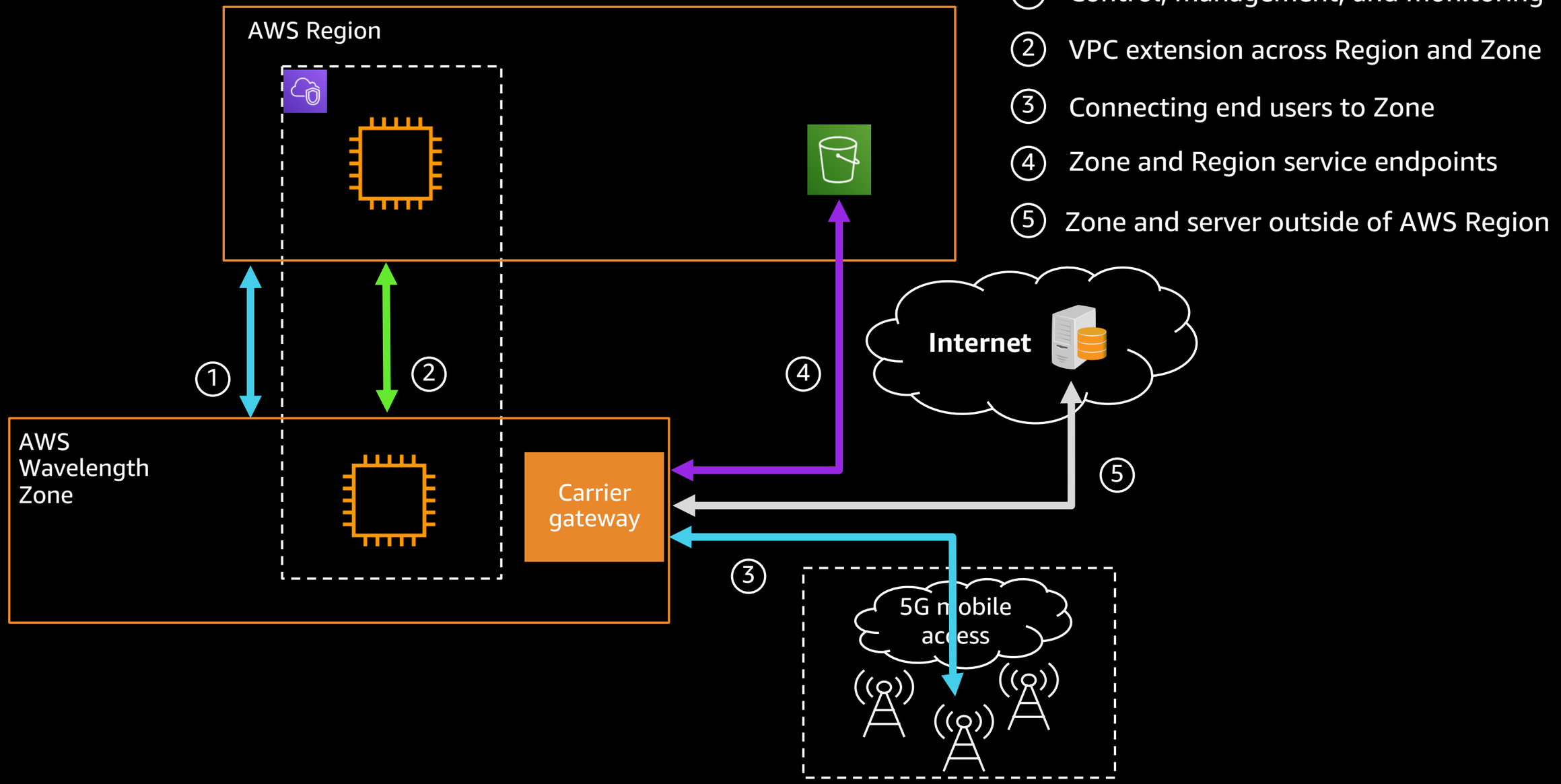
Application architecture considerations



Application architecture considerations

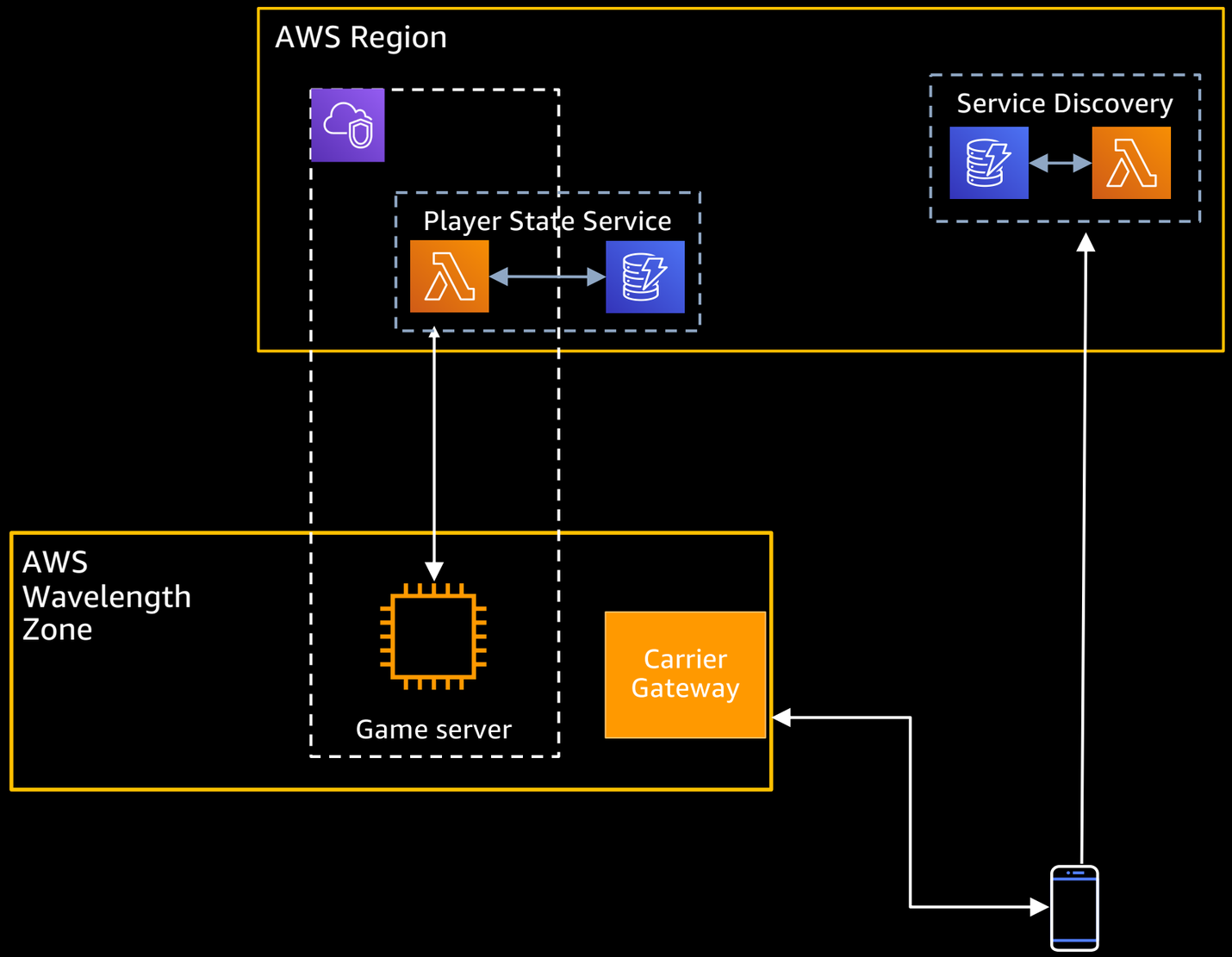


Deploying applications – Wavelength

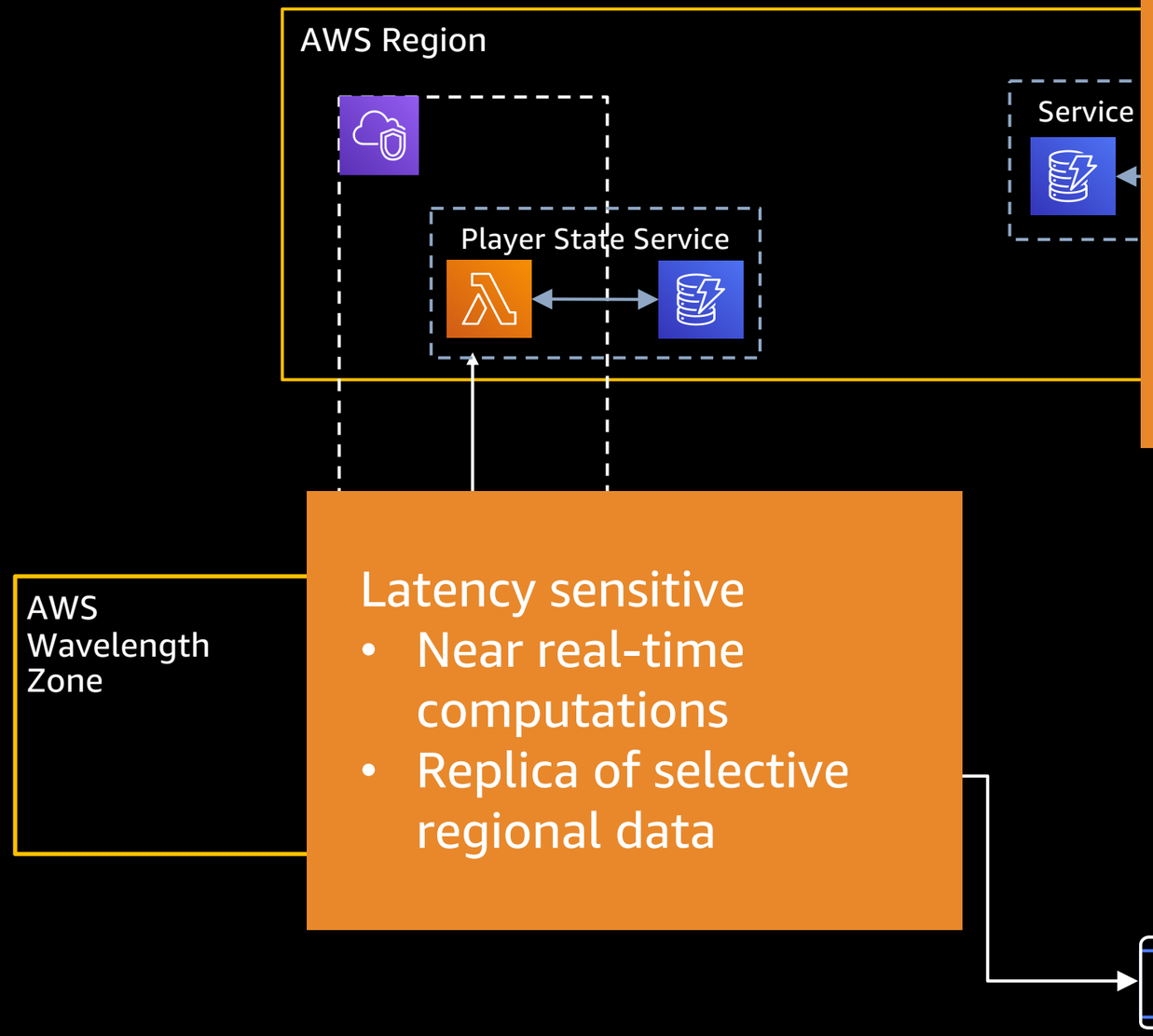


1. Control, management, and monitoring
2. VPC extension across Region and Zone
3. Connecting end users to Zone
4. Zone and Region service endpoints
5. Zone and server outside of AWS Region

Deploying a gaming application



Deploying a gaming application



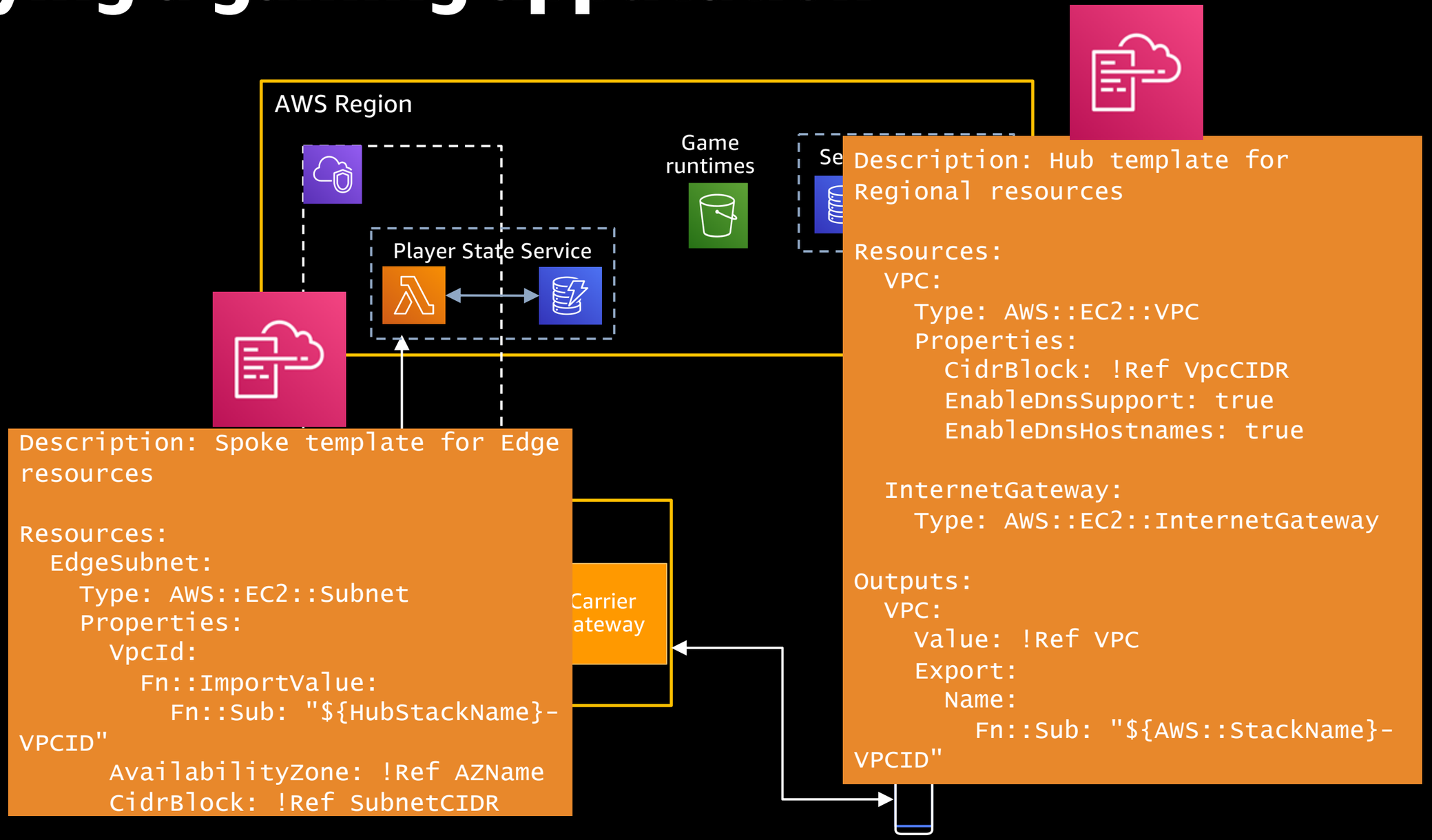
Generally non-latency-sensitive or hard-to-distribute workloads

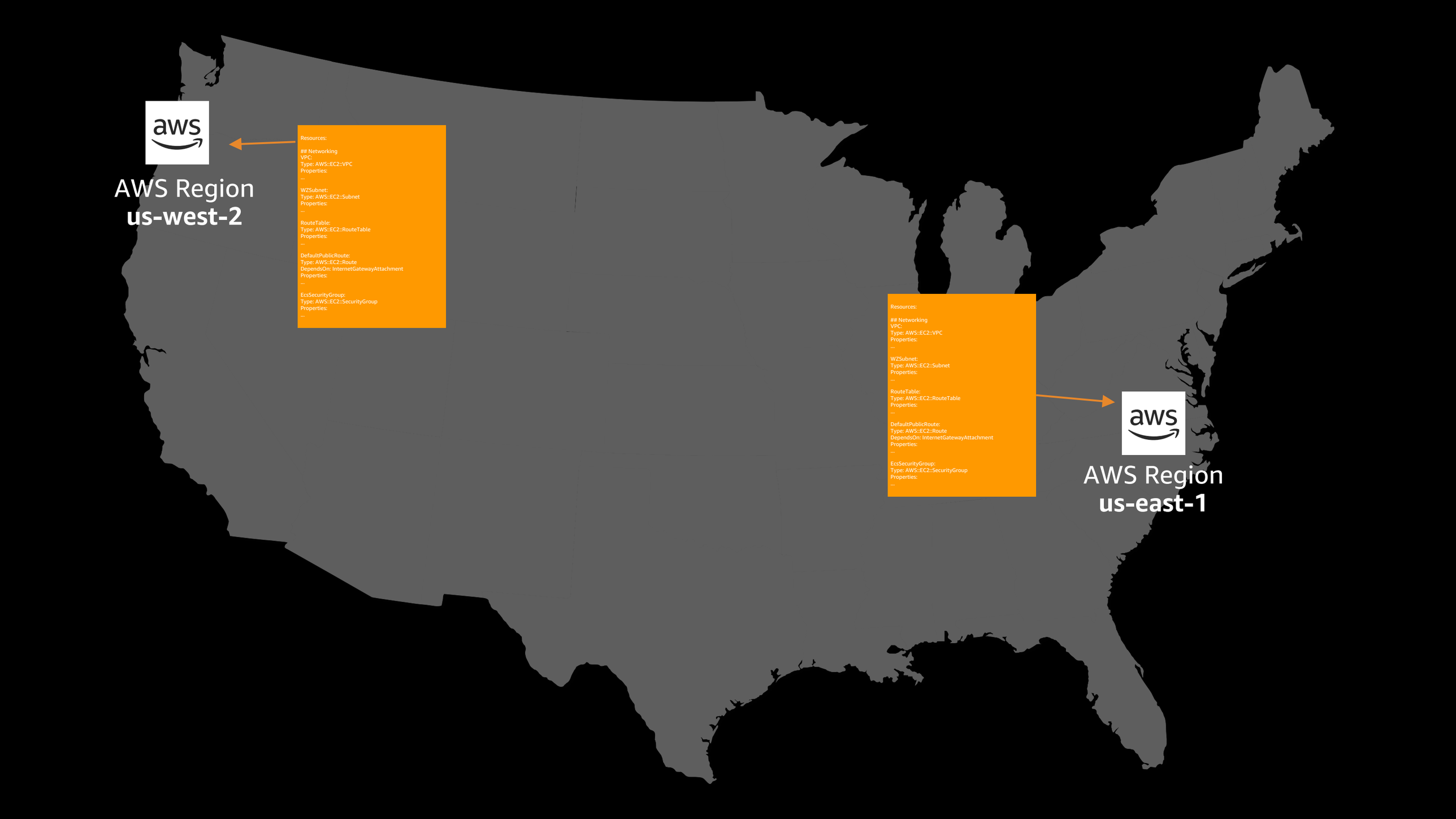
- Auth
- Async workflows
- Master databases
- Matchmaking / placement

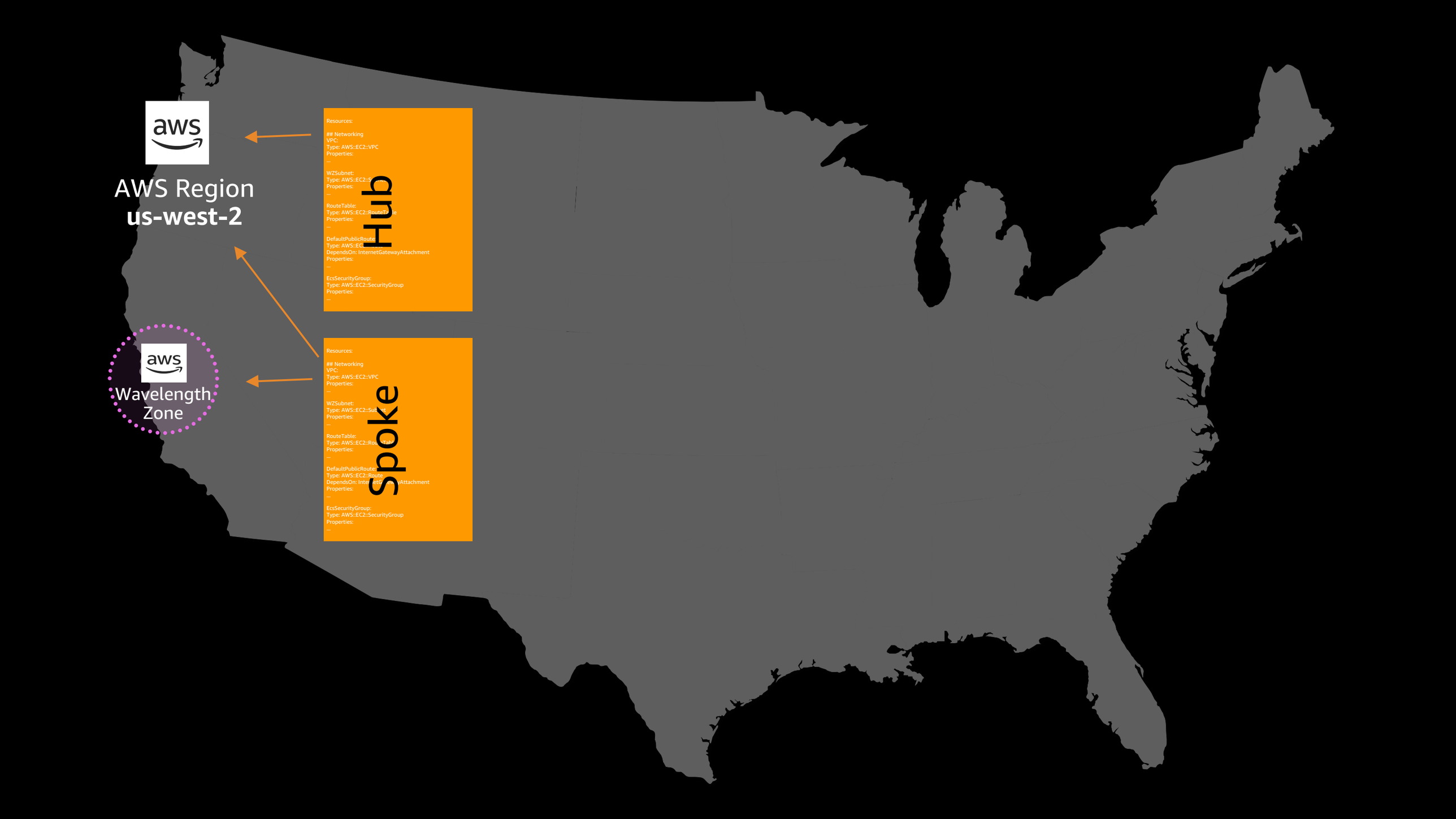
Latency sensitive

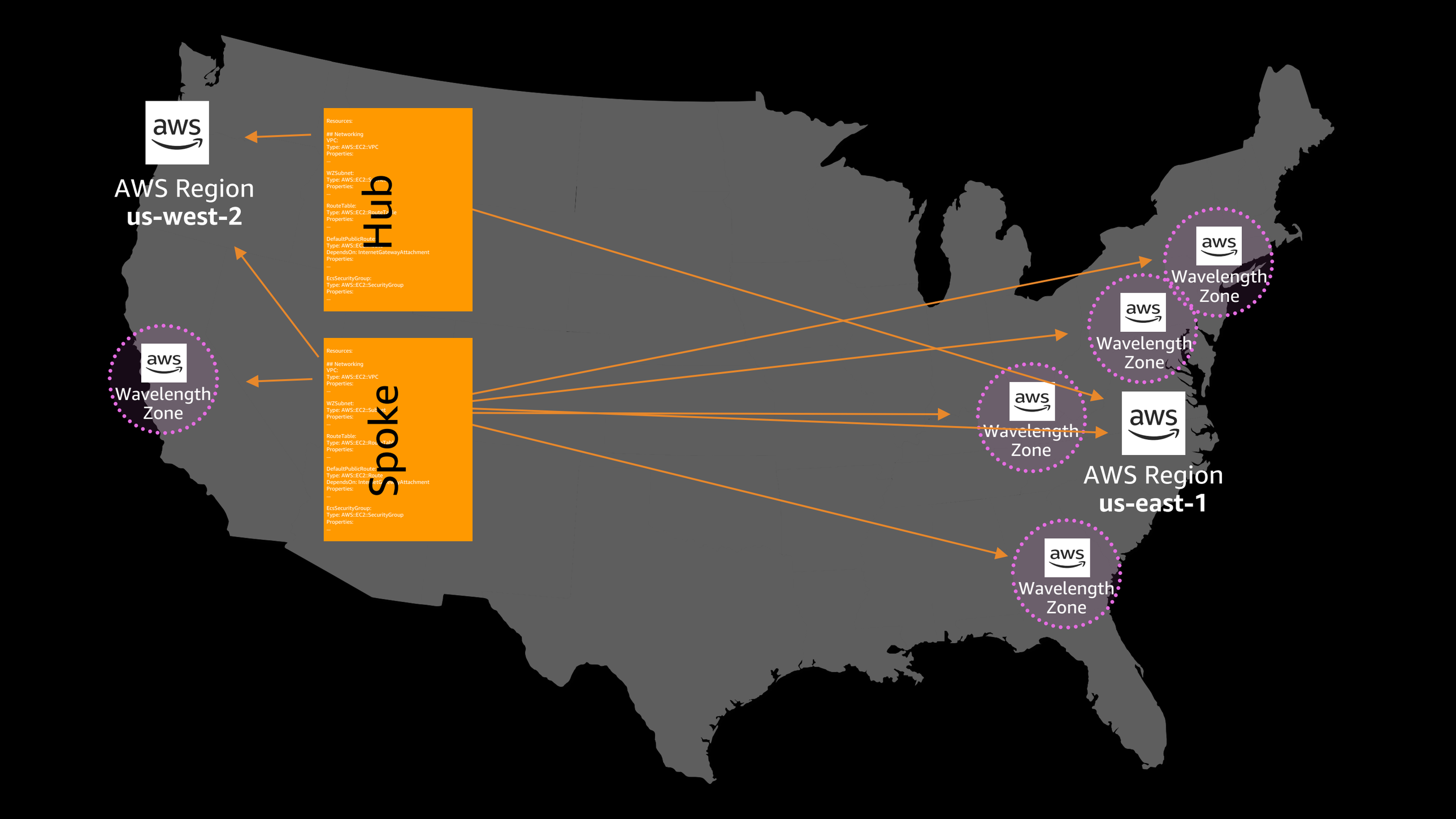
- Near real-time computations
- Replica of selective regional data

Deploying a gaming application

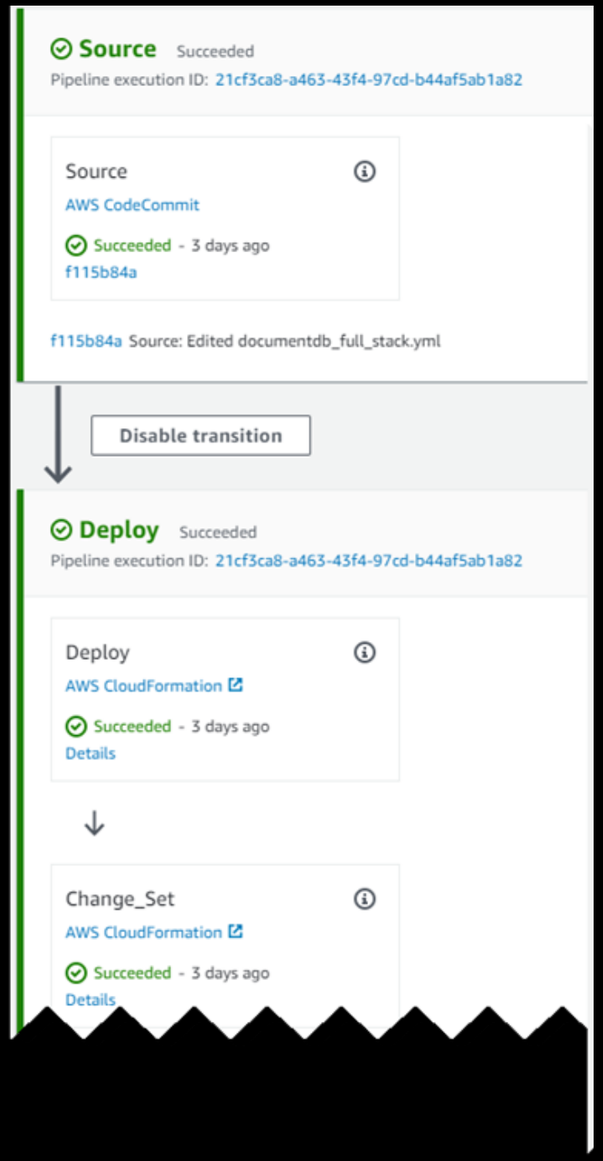



AWS Region
us-west-2
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Type: AWS::EC2::Subnet
Properties:
...
RouteTable:
Type: AWS::EC2::RouteTable
Properties:
...
DefaultPublicRoute:
Type: AWS::EC2::Route
DependsOn: InternetGatewayAttachment
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Type: AWS::EC2::Subnet
Properties:
...
RouteTable:
Type: AWS::EC2::RouteTable
Properties:
...
DefaultPublicRoute:
Type: AWS::EC2::Route
DependsOn: InternetGatewayAttachment
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
AWS Region
us-east-1


AWS Region
us-west-2
Wavelength
Zone
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Properties:
...
RouteTable:
Properties:
...
DependsOn: InternetGatewayAttachment
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
Hub
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Properties:
...
RouteTable:
Properties:
...
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
Spoke


AWS Region
us-west-2
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Type: AWS::EC2::S
Properties:
...
RouteTable:
Type: AWS::EC2::RouteTa
Properties:
...
DefaultPublicRoute:
Type: AWS::EC
DependsOn: InternetGatewayAttachment
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
Hub
Wavelength
Zone
Resources:
Networking
VPC:
Type: AWS::EC2::VPC
Properties:
...
WZSubnet:
Type: AWS::EC2::Subnet
Properties:
...
RouteTable:
Type: AWS::EC2::Route
Properties:
...
DefaultPublicRoute:
Type: AWS::EC2::Route
DependsOn: InternetGatewayAttachment
Properties:
...
EcsSecurityGroup:
Type: AWS::EC2::SecurityGroup
Properties:
...
Spoke
Wavelength
Zone
Wavelength
Zone
Wavelength
Zone
Wavelength
Zone
AWS Region
us-east-1

Deployment pipeline



Stages to deploy hub templates to each Region

Stages to deploy edge templates with actions for each Zone

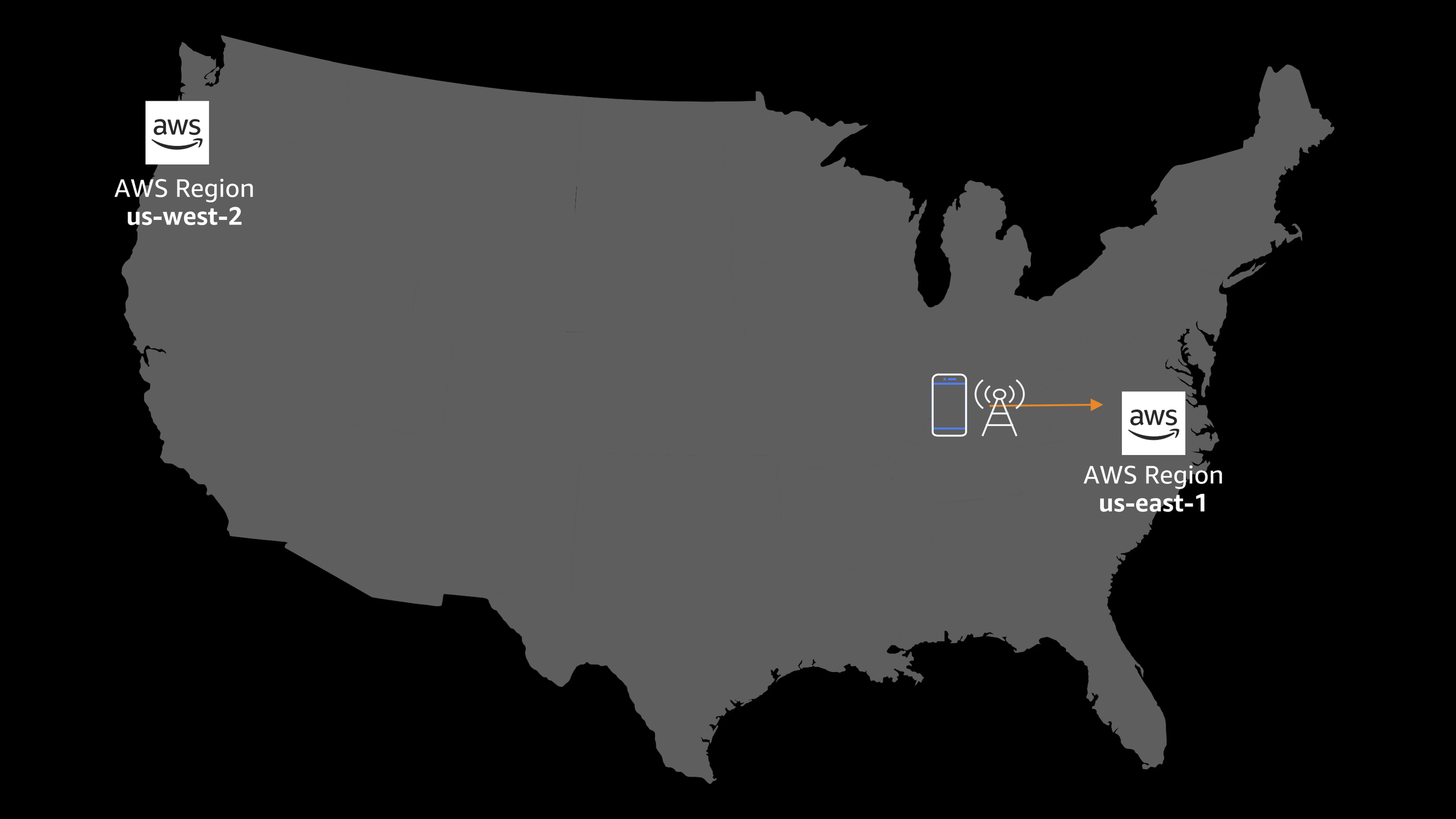
Input and output parameters

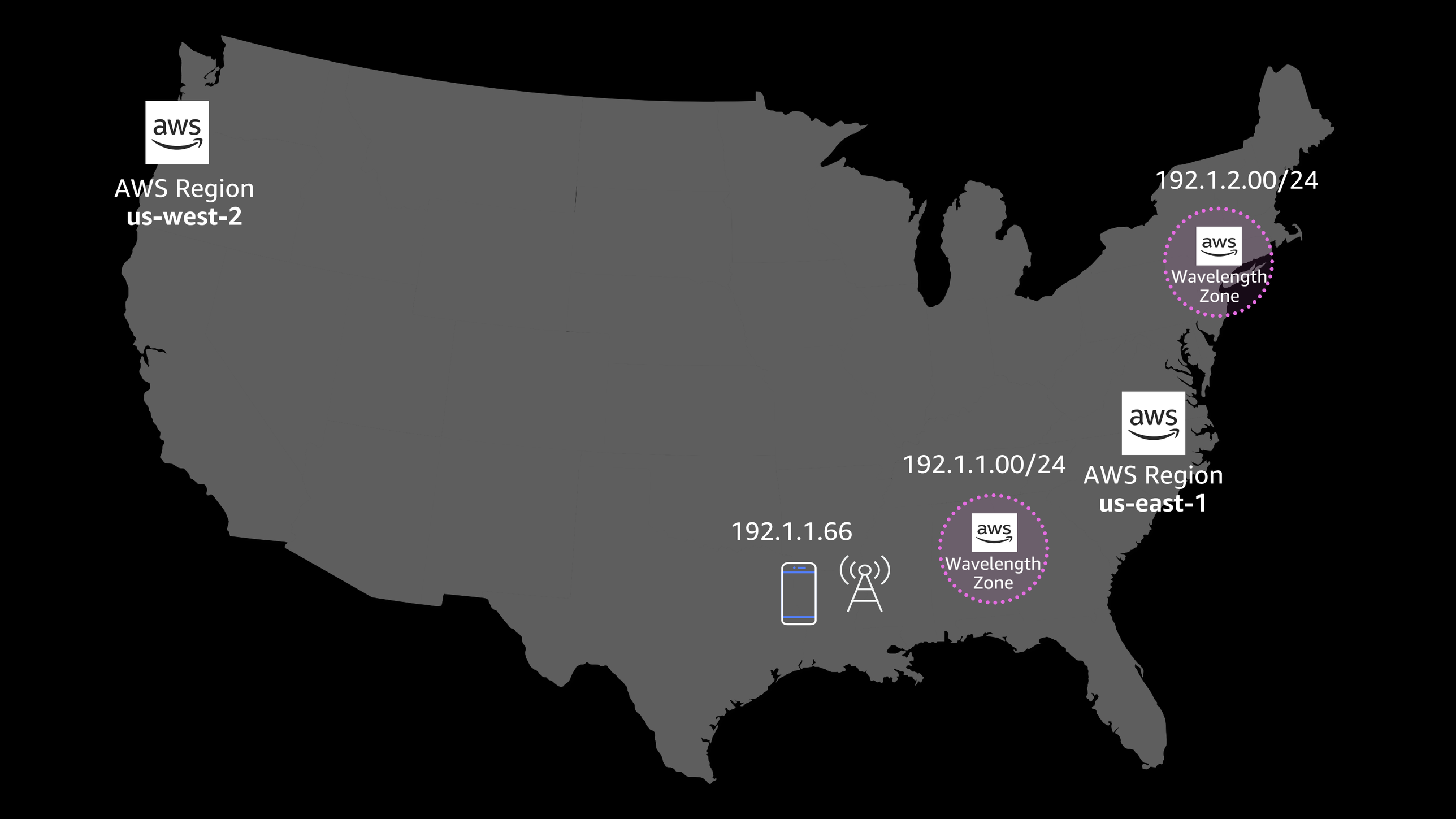
Endpoint discovery

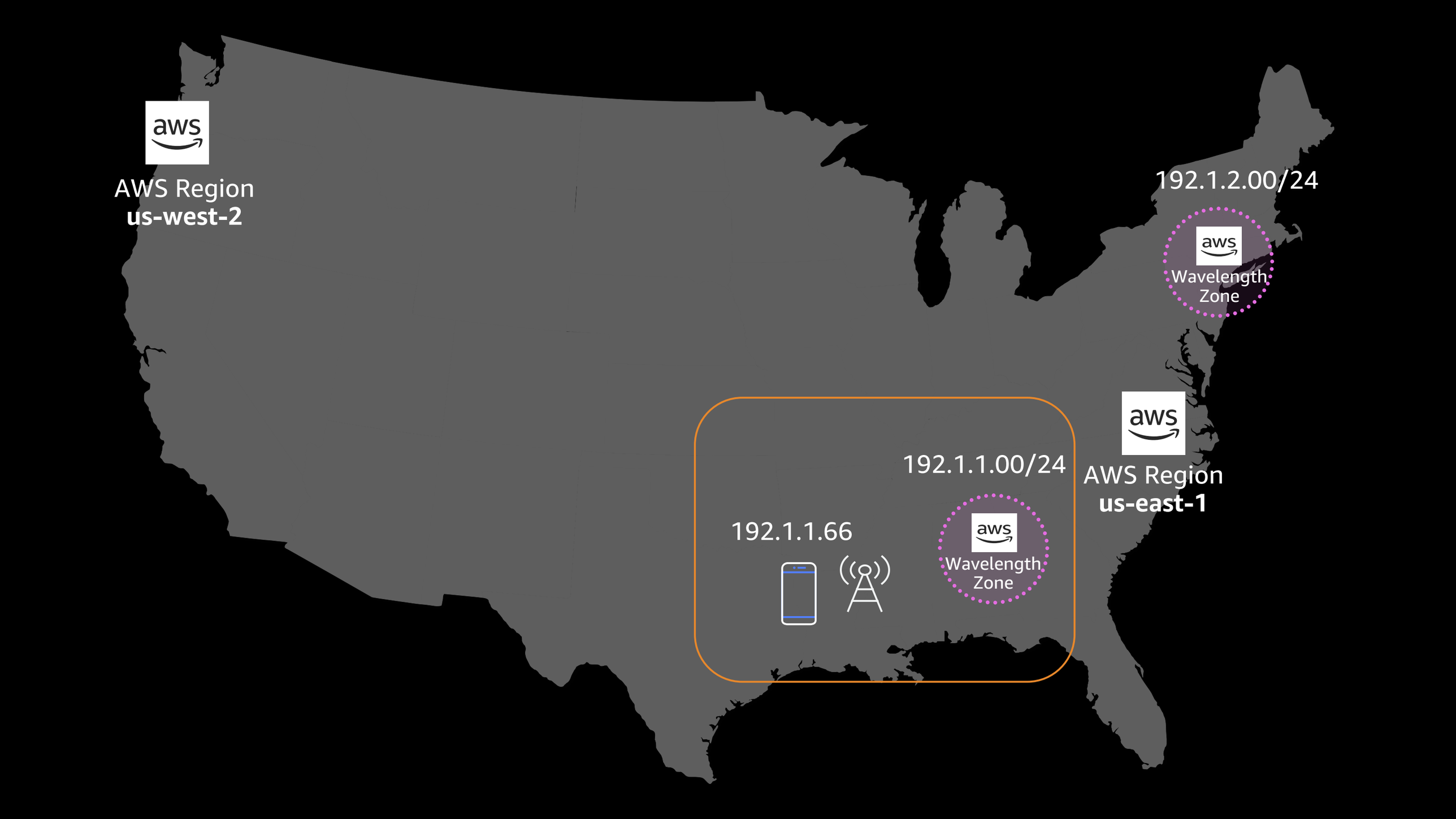


AWS re:Invent

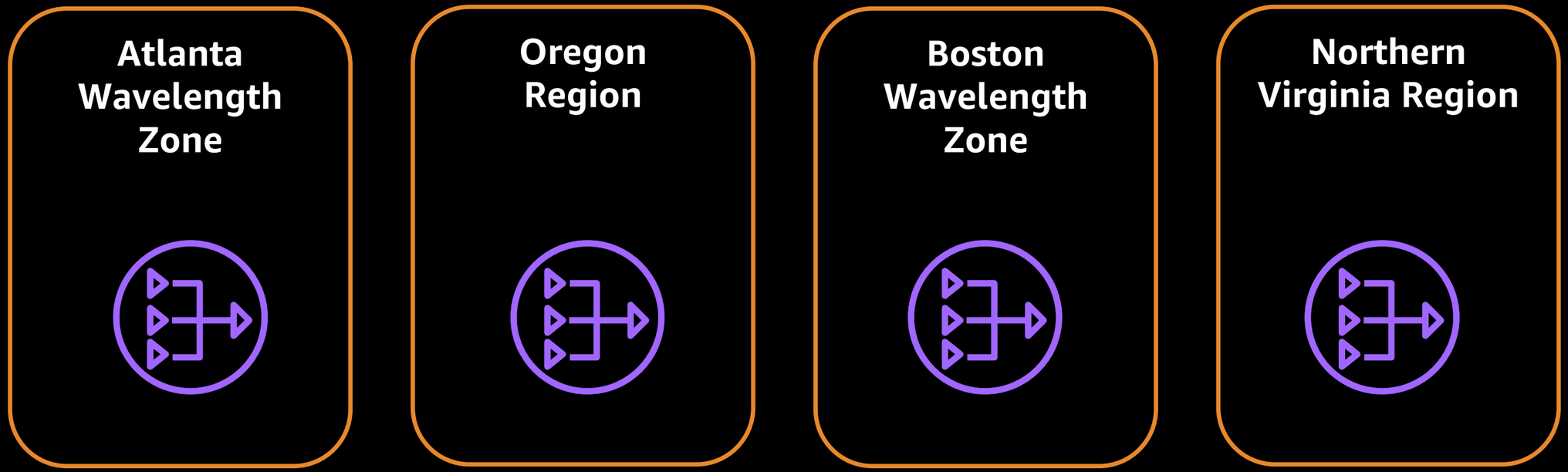
© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.


AWS Region
us-west-2
AWS Region
us-east-1


AWS Region
us-west-2
192.1.2.00/24
Wavelength Zone
AWS Region
us-east-1
192.1.1.00/24
Wavelength Zone
192.1.1.66


AWS Region
us-west-2
192.1.2.00/24
Wavelength Zone
192.1.1.00/24
192.1.1.66
Wavelength Zone
AWS Region
us-east-1

Register application resources



Endpoint registration

```
$ aws servicediscovery create-service --name game-service
$ aws servicediscovery register-instance --service-id game-service \
--instance-id i-0ab1cdefghij23 --attributes
AWS_INSTANCE_IPV4=54.20.10.1,stage=prod,version=1.0,zone=us-east-1-wl1-atl-wlz-1,
lat=33.753746,lon=-84.386330
```

Endpoint discovery

```
# In the client application, determine the geolocation

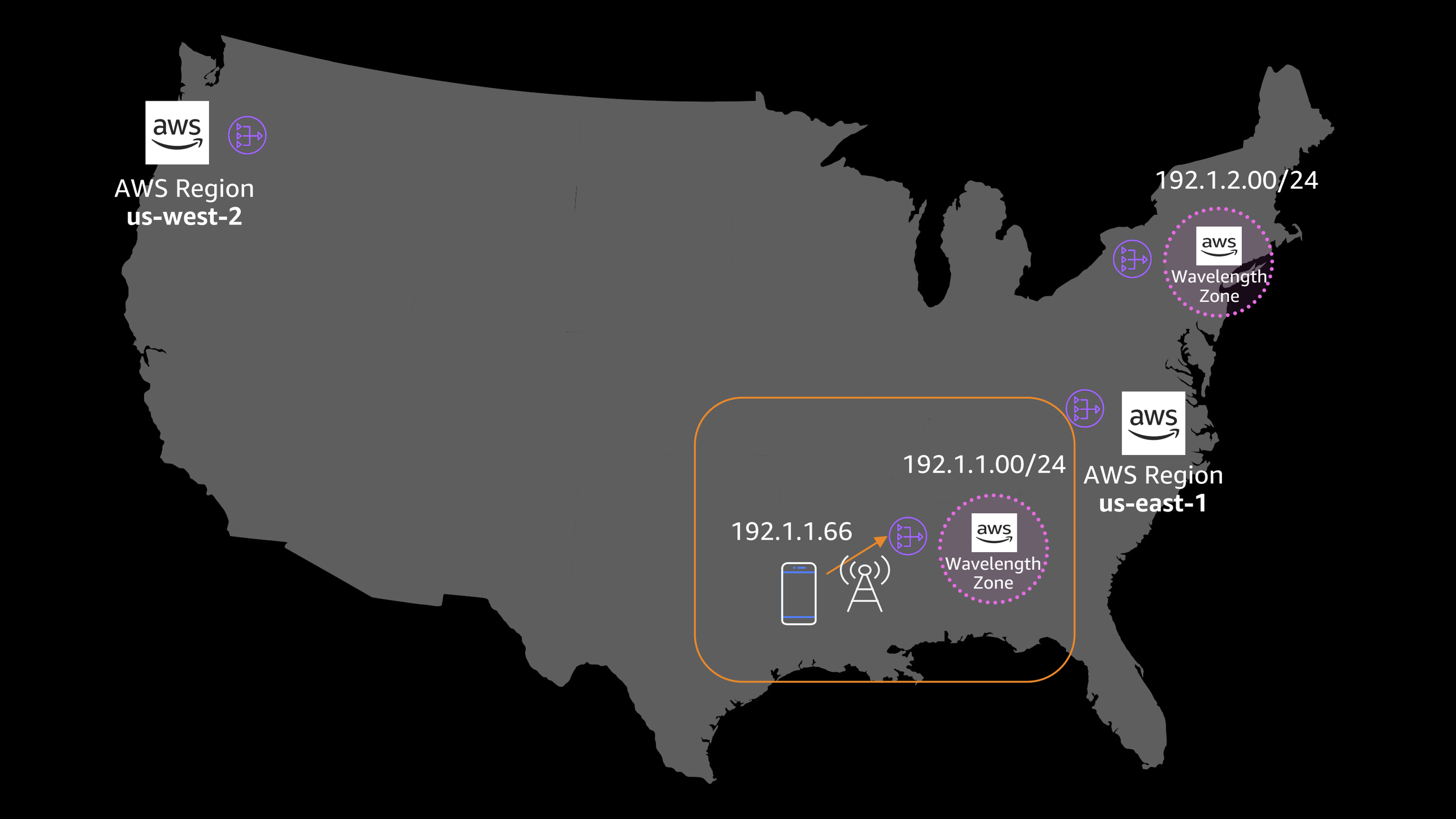
navigator.geolocation.getCurrentPosition(success, error, geoOptions);

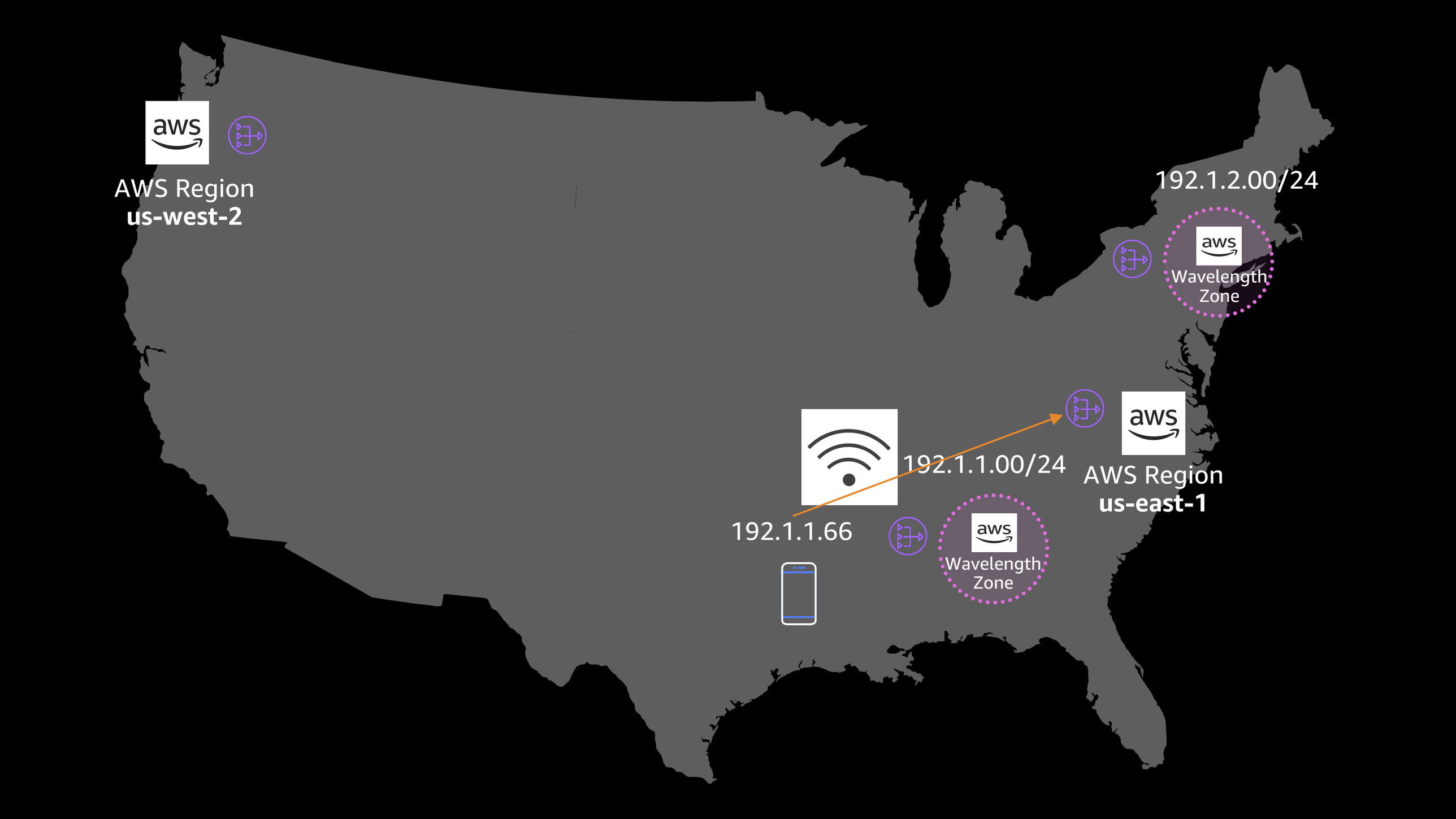
const latitude = position.coords.latitude;
const longitude = position.coords.longitude;

# Determine closest airport based on lat/lon

# Determine whether on LTE, 5G or WiFi

$ aws servicediscovery discover-instances --service-name game-service \
--query-parameters zone=us-east-1-wl1-atl-wlz-1
```


AWS Region
us-west-2
192.1.2.00/24
Wavelength Zone
192.1.1.00/24
192.1.1.66
Wavelength Zone
AWS Region
us-east-1


AWS Region
us-west-2
192.1.2.00/24
Wavelength Zone
192.1.1.00/24
AWS Region
us-east-1
192.1.1.66
Wavelength Zone

Thank you!



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.