

AWS re:Invent

N E T 3 1 8 - R

Building highly available applications using AWS Global Accelerator

James Wenzel

Sr Partner SA, Networking
Amazon Web Services, Inc.

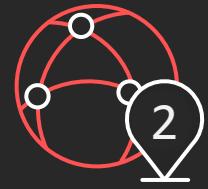
Harvo Jones

Principal SDE
Amazon Web Services, Inc.

Agenda

- Building on AWS
 - Challenges of the internet
 - Creating and using an accelerator
 - Traffic control
- The power of AWS Global Accelerator
 - Security, availability, performance
- Building with AWS Global Accelerator

Key features we'll learn



Single entry point
with global static
anycast IPs



Intelligent
distribution of TCP &
UDP traffic



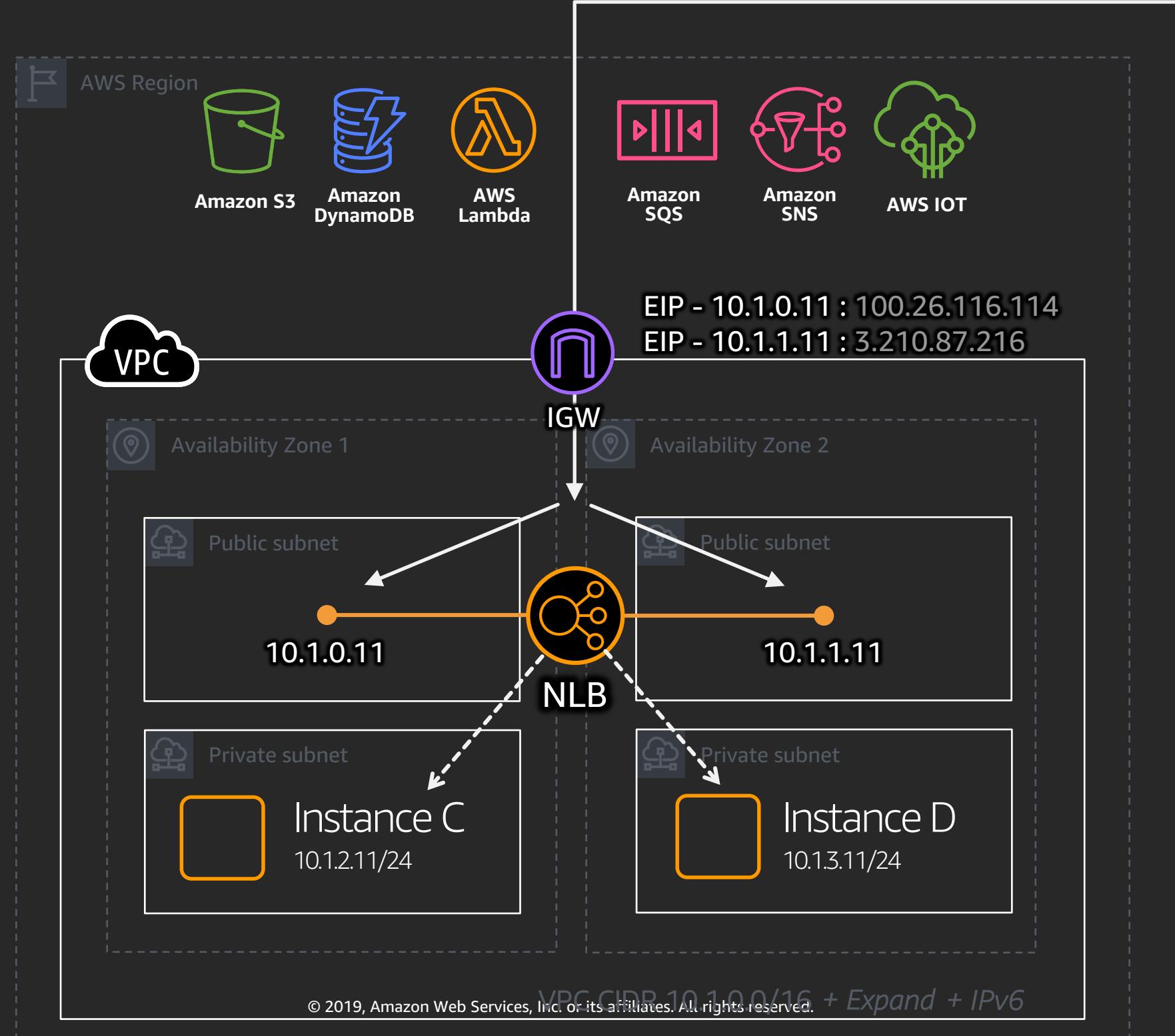
Target EC2 instances
and Elastic Load
Balancers (ALBs & NLBs)
Preserve client IPs

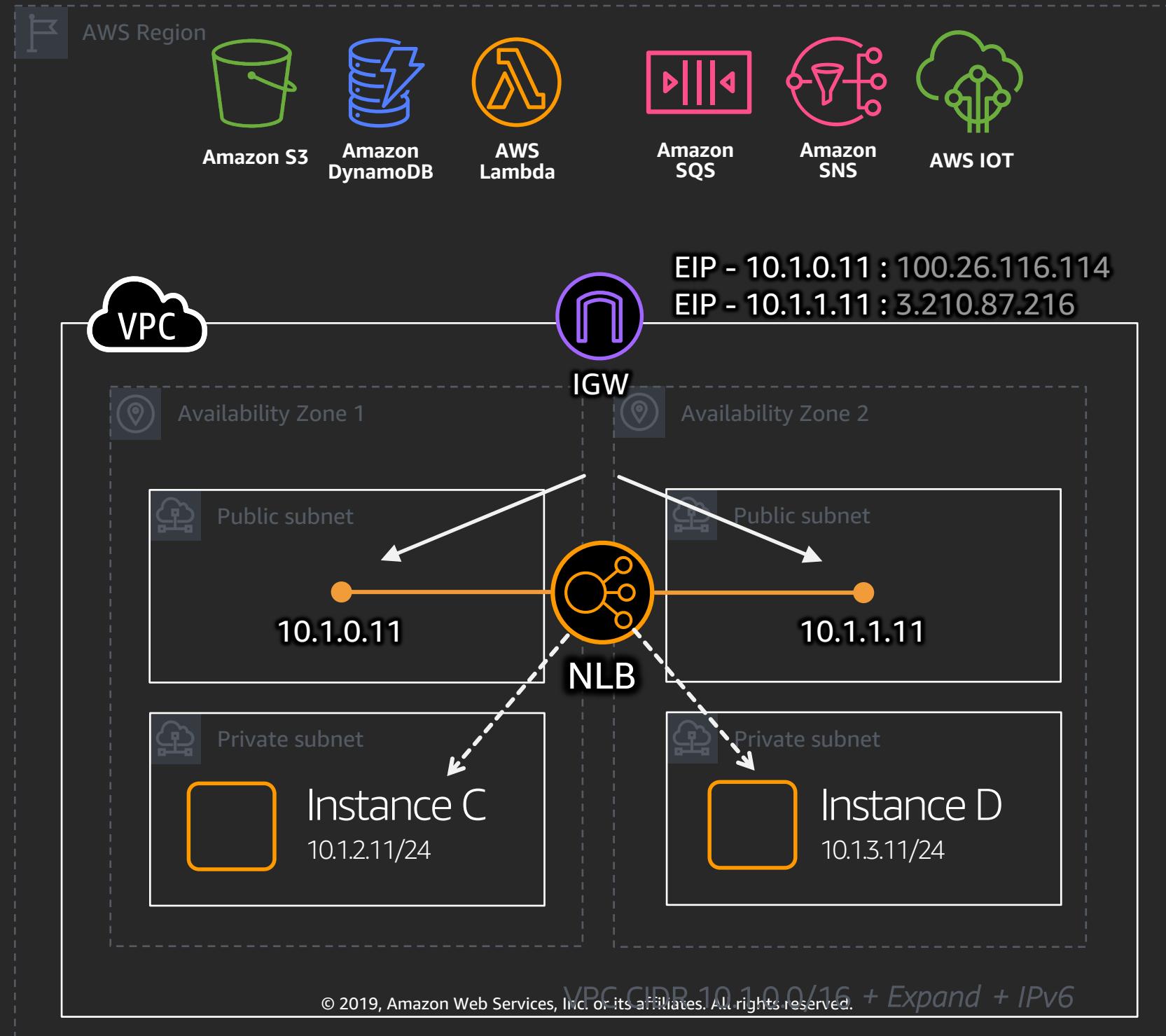


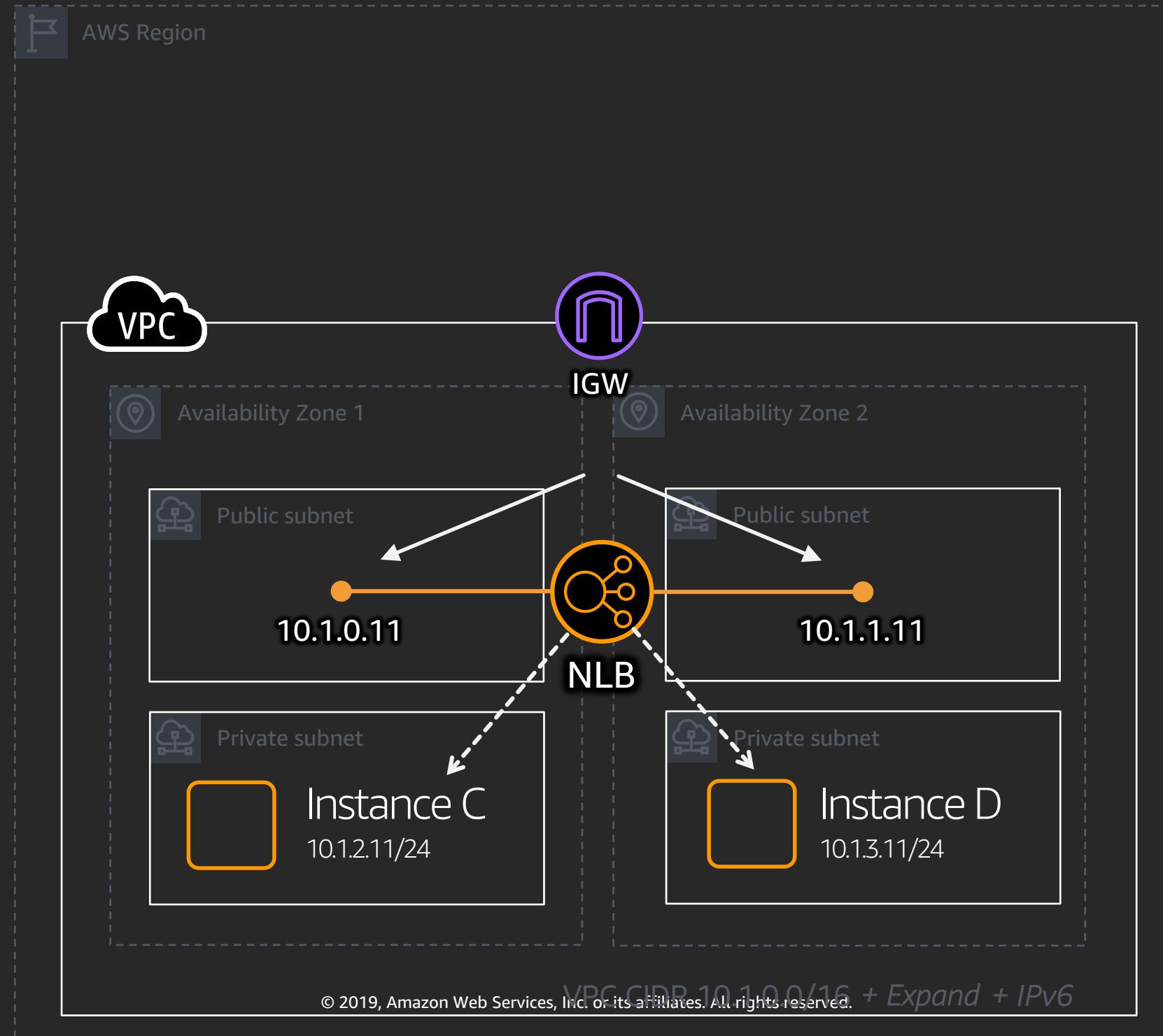
Accelerate both
single and multi-
region workloads

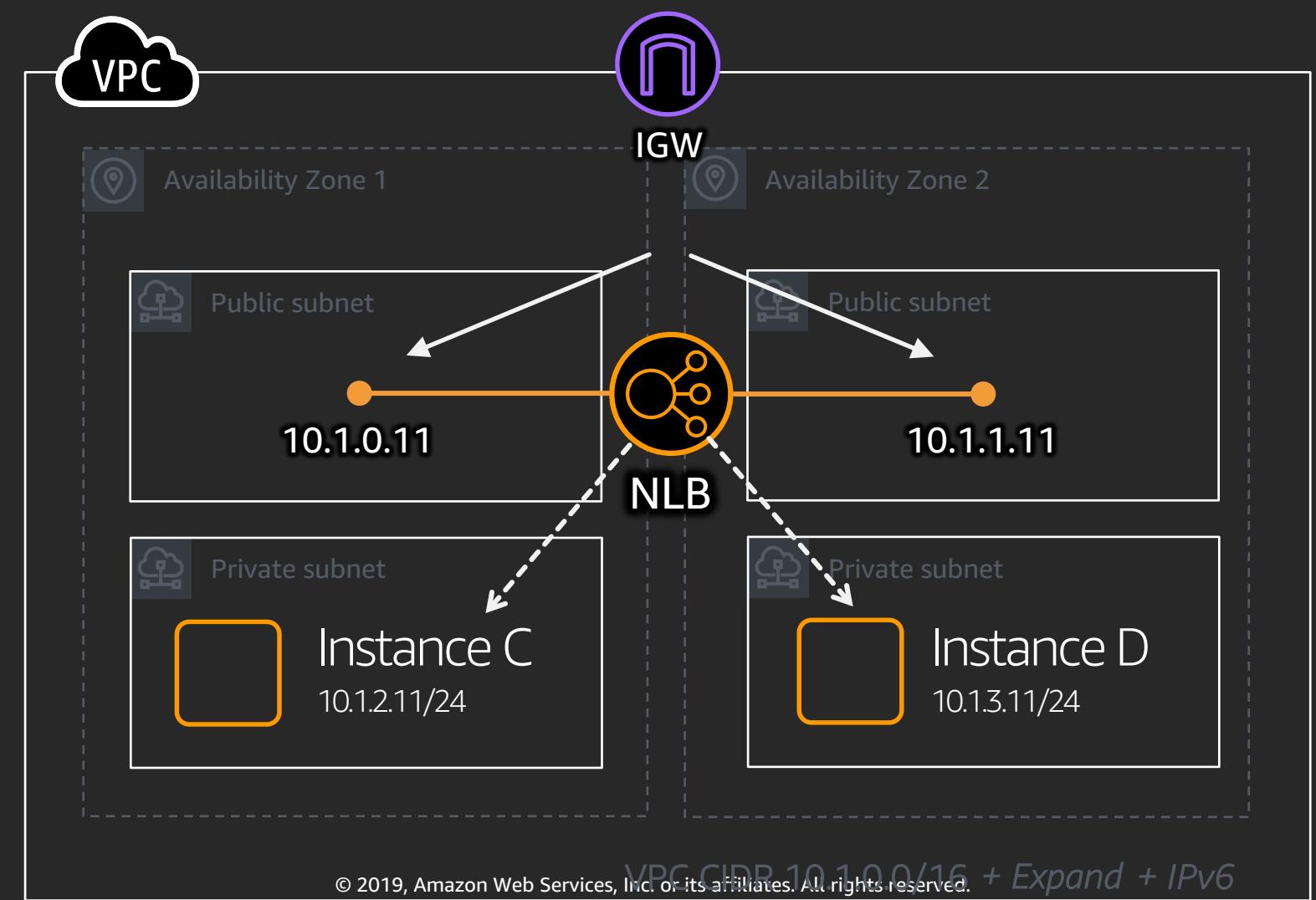
All traffic traverses the backbone* and is
protected from DDoS attacks

* Except within the People's Republic of China

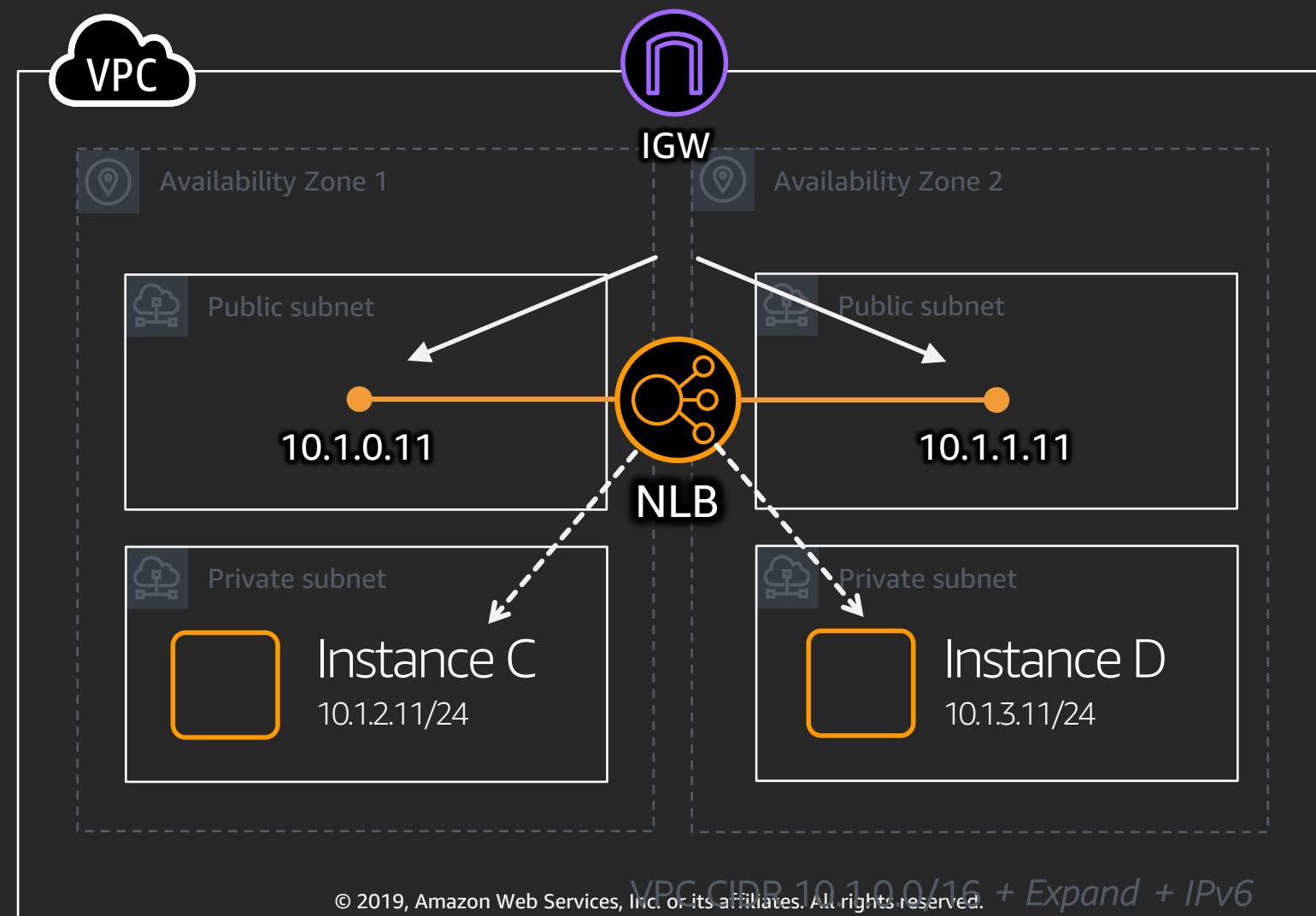


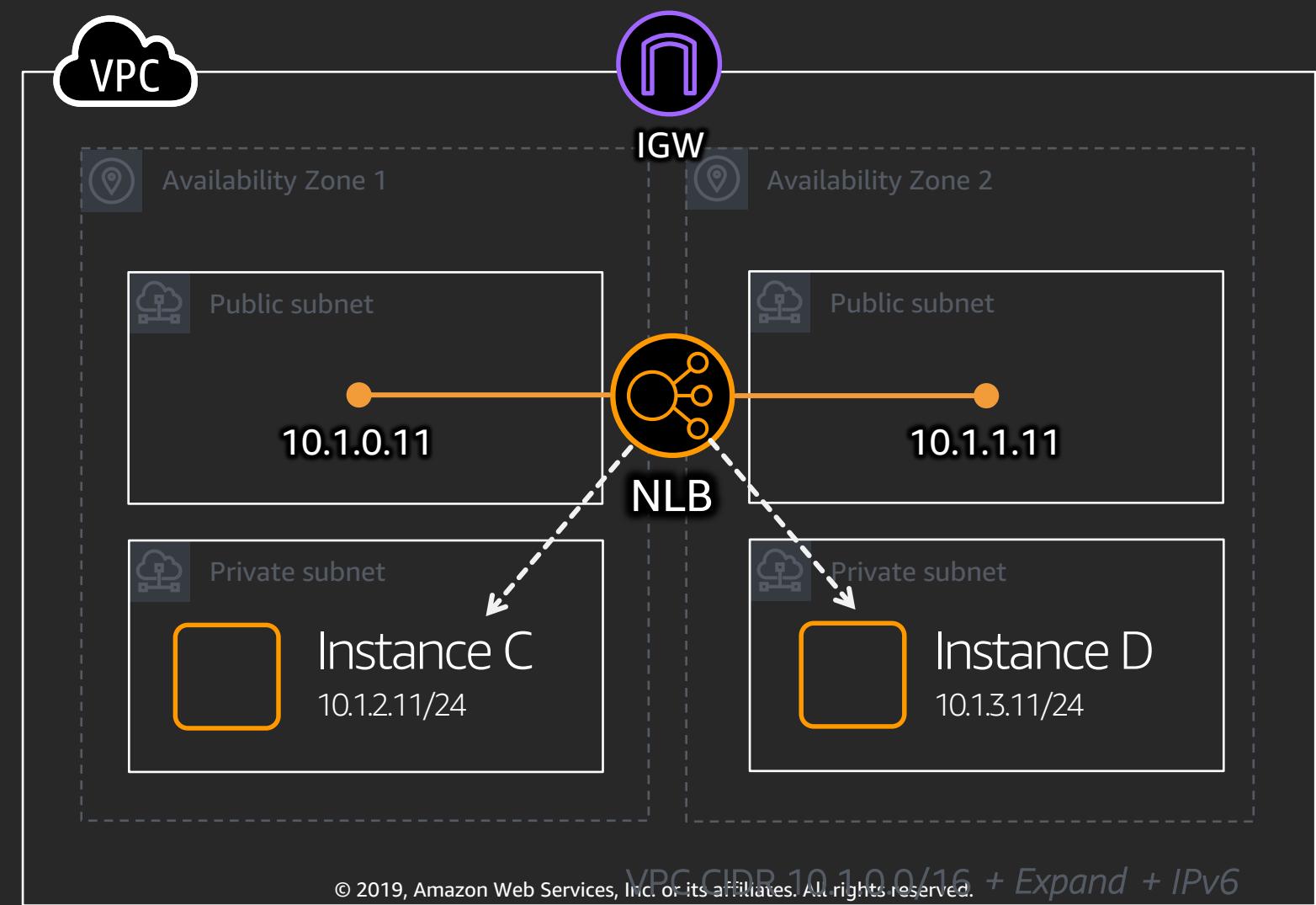




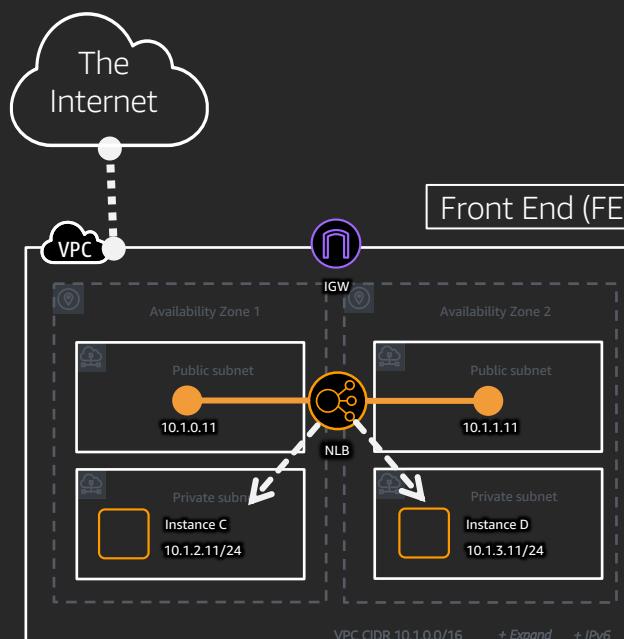


Let's start with a common work load...



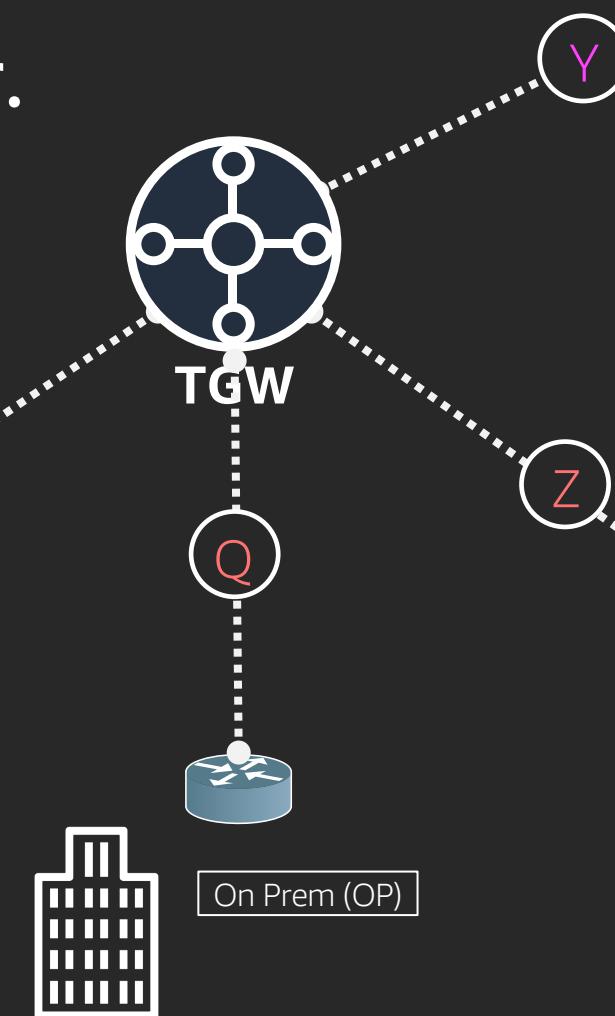


Let's build our
Great Everything is
workload!
Working the way we want
in a well-architected manner.

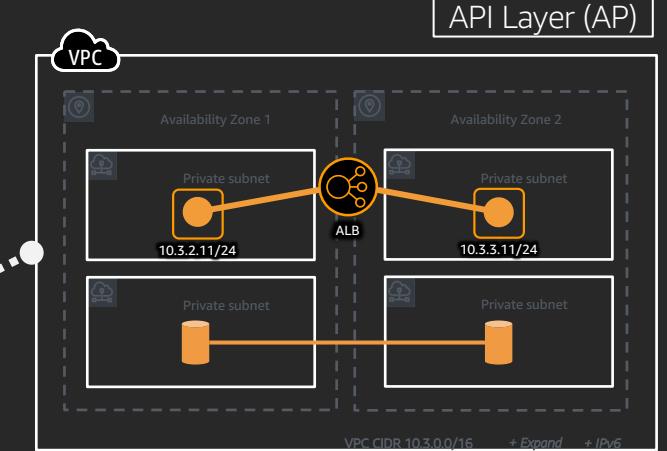


Destination	Target
10.1.0.0/16	Local
0.0.0.0/0	IGW
10.0.0.0/8	TGW

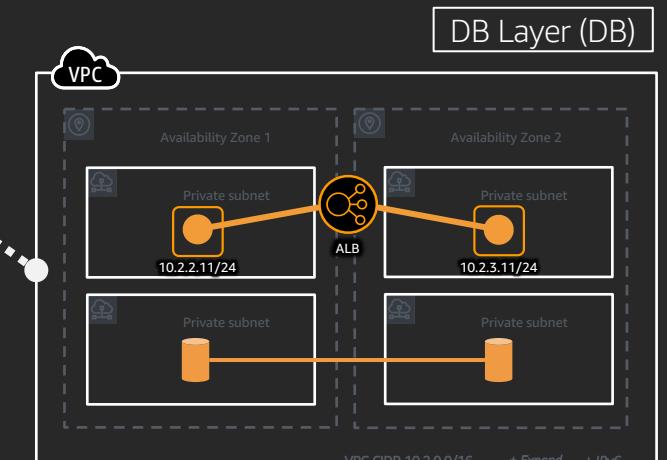
Associations	Propagations	Routes
FE from X	FE from X	10.1.0.0/16 >X
DB from Z	DB from Z	10.2.0.0/16 >Z
API from Y	API from Y	10.3.0.0/16 >Y
OP from Q	OP from Q	10.4.0.0/16 >Q



Destination	Target
10.3.0.0/16	Local
0.0.0.0/0	TGW



Destination	Target
10.2.0.0/16	Local
0.0.0.0/0	TGW



Now we open up our workload to the world



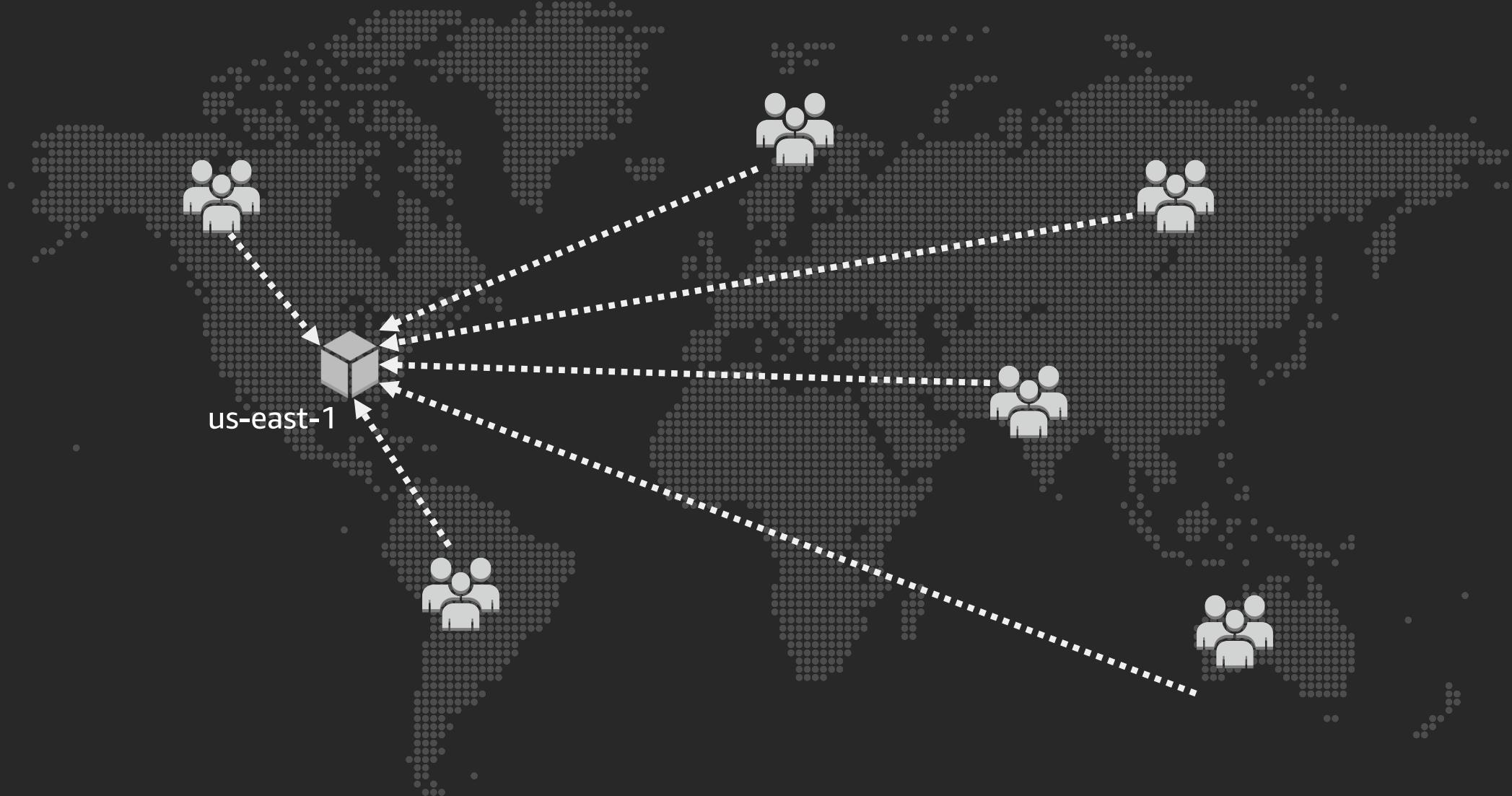
Now we open up our workload to the world



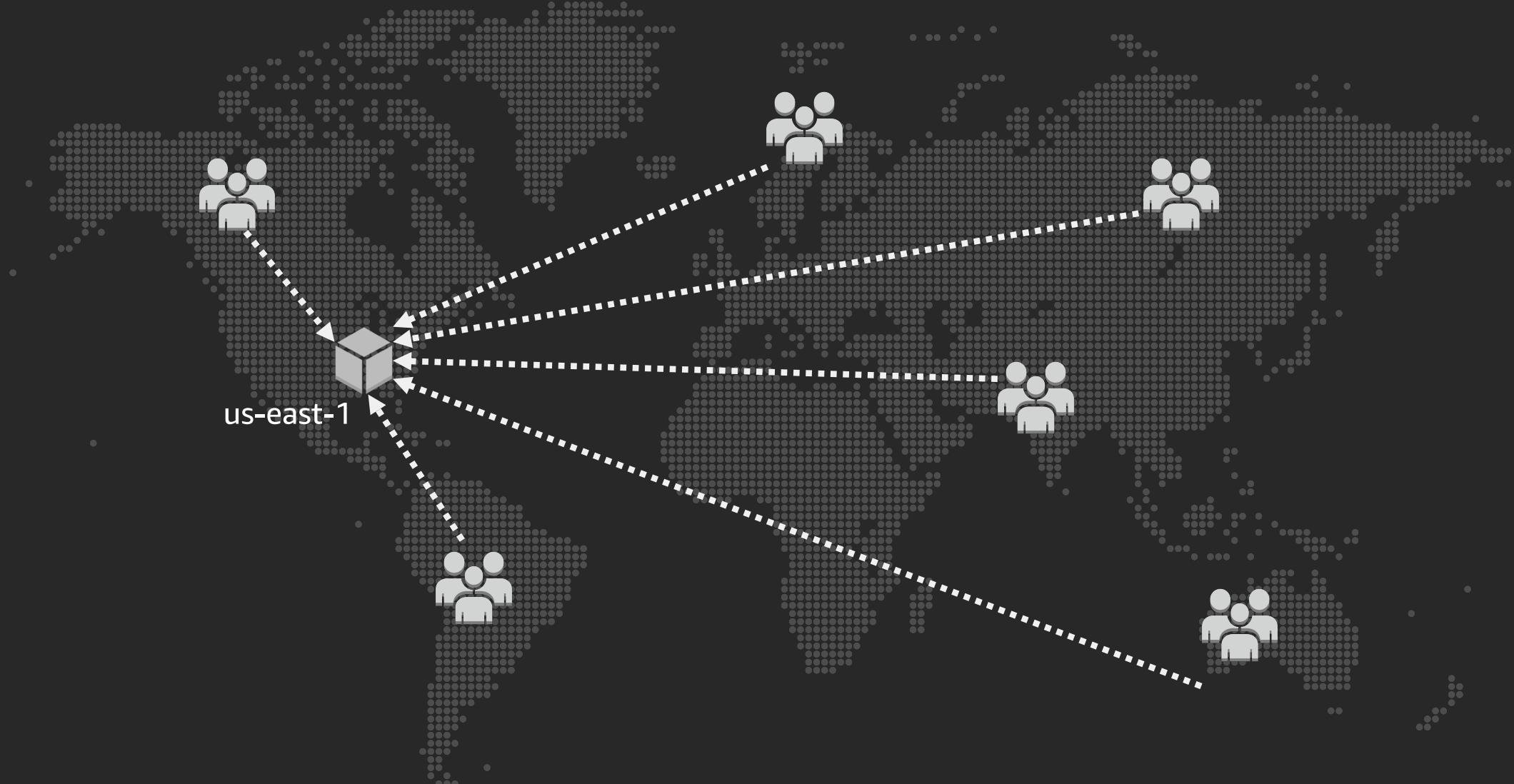
Because we are on the internet, it's accessible from everywhere.



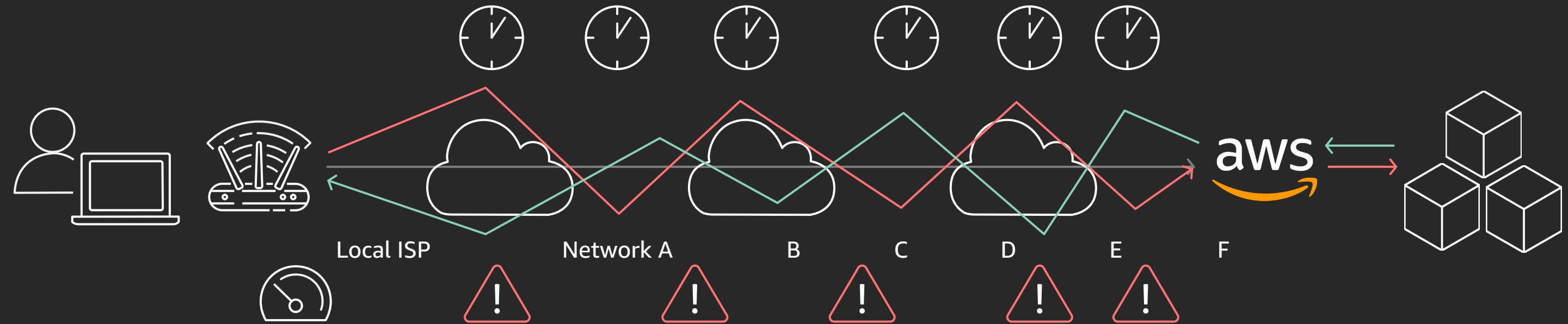
Because we are on the internet, it's accessible from everywhere.



Not all of our customers will have the same experience due to internet weather...



Internet weather



Accessing your application is not this straightforward
It can take many networks to reach the application
Paths to and from the application may differ
Each hop impacts performance and can introduce risk

There are a few ways to handle this...



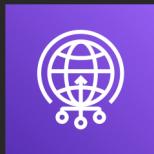
Amazon Route 53



Amazon CloudFront

Our global DNS solution
with routing capabilities

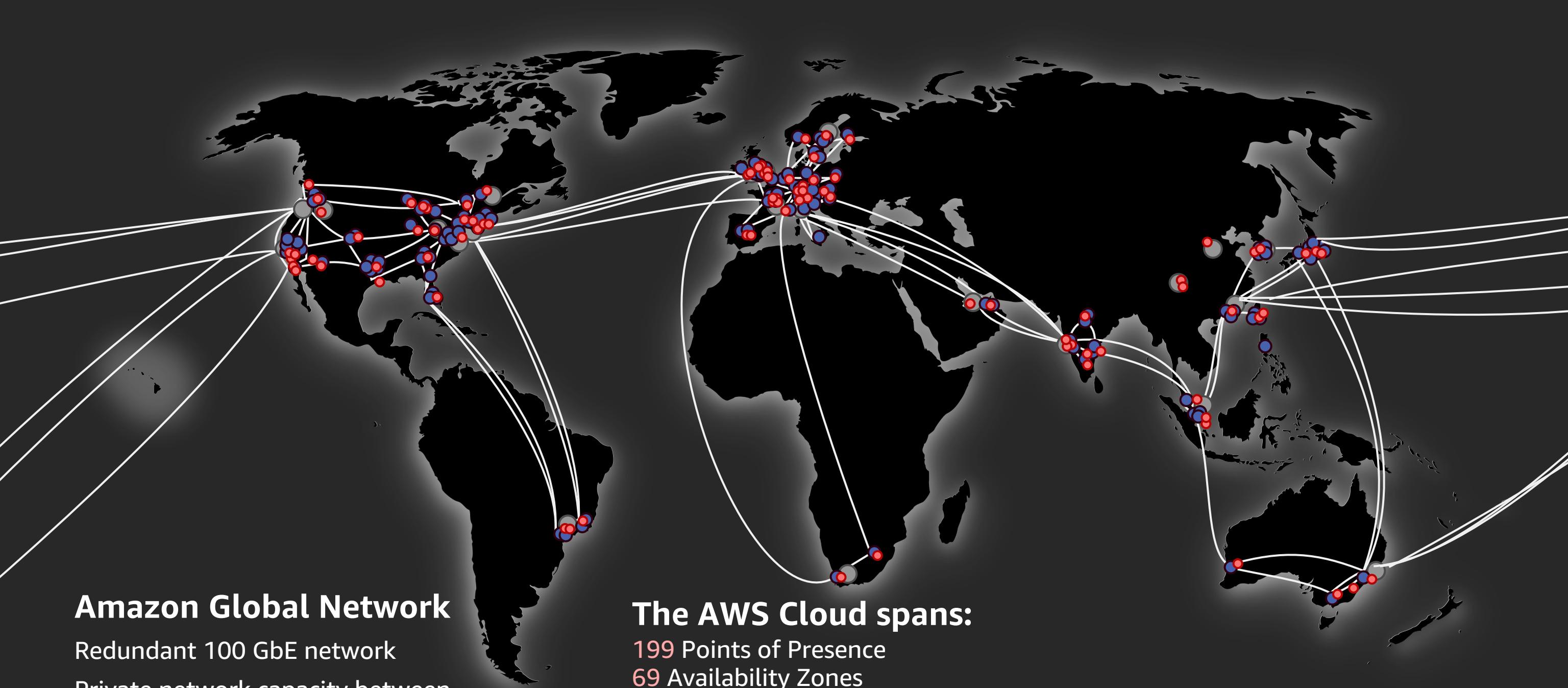
Our HTTP/S CDN solution
with WAF and Edge Compute



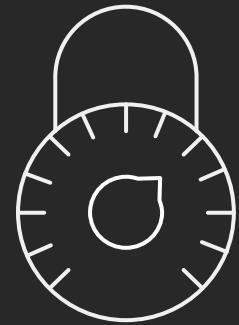
AWS Global Accelerator

Accelerators that direct TCP & UDP traffic to optimal endpoints
over the AWS global network

What makes a global accelerator?



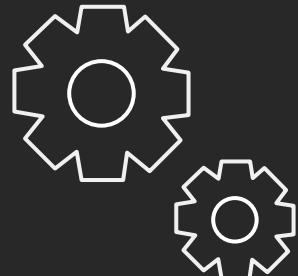
What does having a global backbone do for our customers?



Security

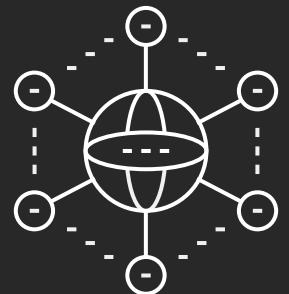
Traffic traverses our infrastructure rather than the internet

*End-to-end backbone encryption



Reliable performance

Controlling paths customer traffic traverses



Availability

Controlling scaling and redundancy



Connecting closer to customers

Avoiding internet hot spots or suboptimal external connectivity

What does this look like for our workloads with AGA?

This lets us reduce jitter and latency



Traffic enters the AWS global network at edge locations
Leverages the Global AWS network
Resulting in improved performance

Accessibility with a globally accessible Anycast IP set

What is Anycast?

Anycast is a network addressing and routing methodology in which a single destination address has multiple routing paths to $N+$ endpoint destinations.

Anycast helps us

Whitelisting only 2 IPs

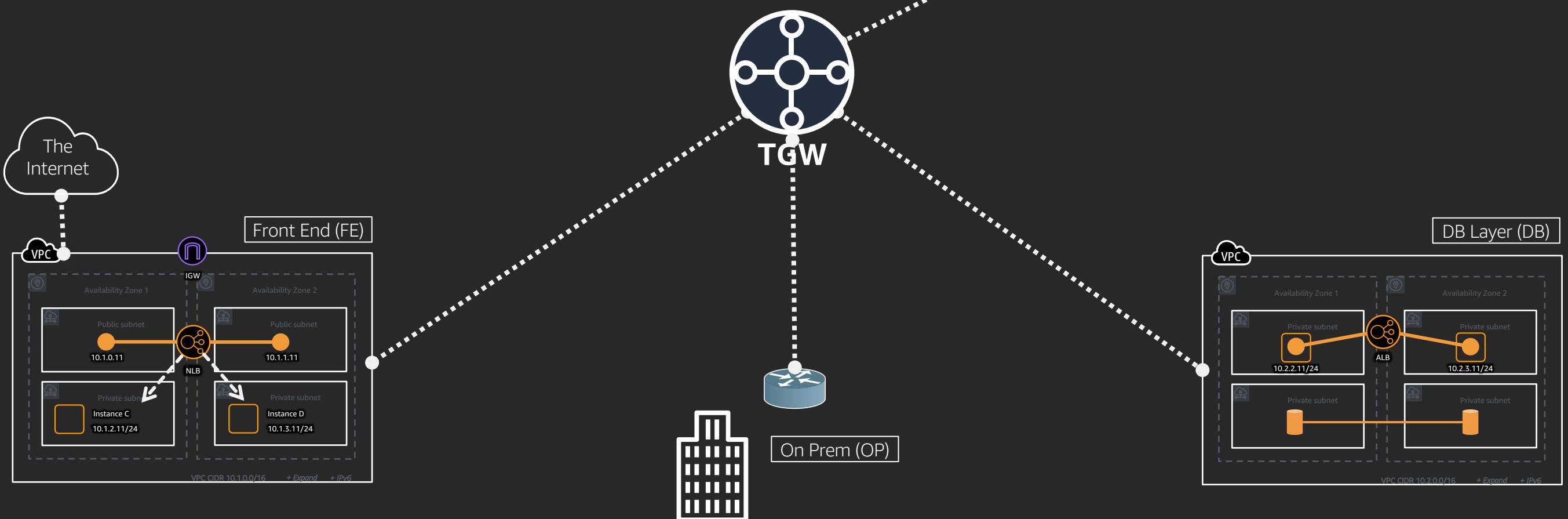
Reduce global address complexity

Source obfuscation

(your endpoints can be private)

Gives us a global presence

Let's replicate our workload from
us-east1 to ap-east 1 and use
Global Accelerator





3.10.3.125

Amazon edge location

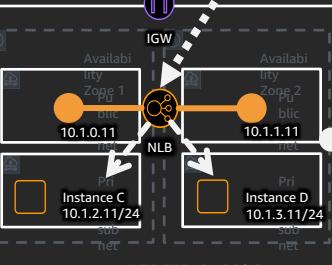
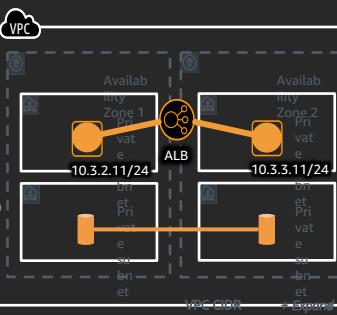


3.10.3.125

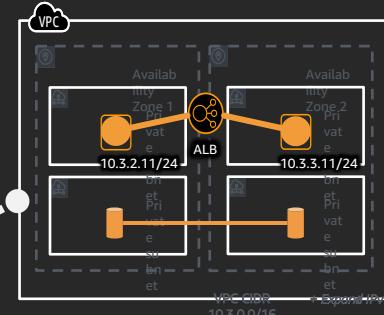
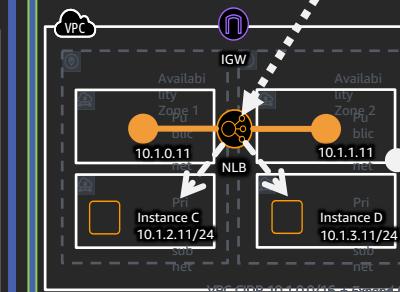
Amazon edge location

AWS Global Accelerator

us-east-1

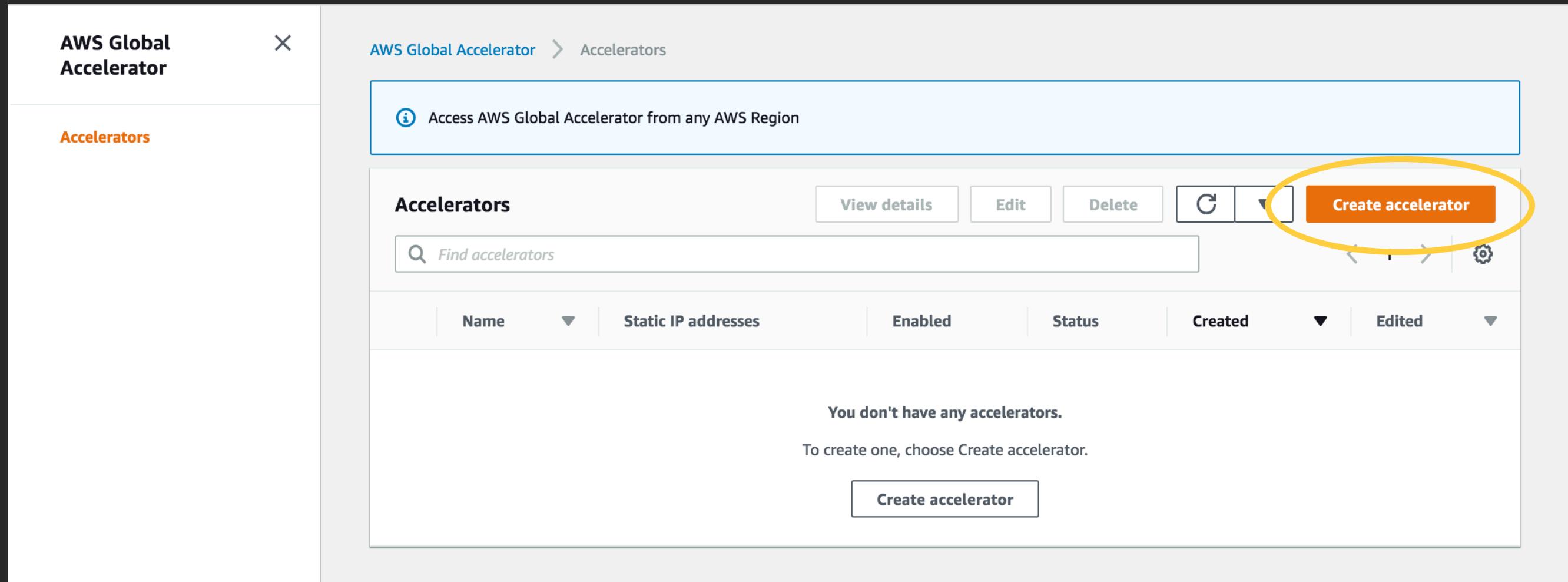


ap-east-1



Creating your accelerator

Navigate to Global Accelerator in the console in the networking section



AWS Global Accelerator X

AWS Global Accelerator > Accelerators > Create accelerator

Step 1 Enter name

Step 2 Add listeners

Step 3 Add endpoint groups

Step 4 Add endpoints

Enter name

An accelerator includes one or more listeners that direct traffic to one or more endpoint groups. An endpoint group includes endpoints, such as load balancers.

Basic configuration

To get started with creating your accelerator, provide a name for it.

Accelerator name
Provide a name to associate with your accelerator.
 Use only letters or numbers, with no spaces.

IP address type

Choose SourceIP if you have stateful applications requiring affinity

The screenshot shows the AWS Global Accelerator 'Create accelerator' wizard at Step 2: Add listeners. The left sidebar lists steps: Step 1 Enter name, Step 2 Add listeners (highlighted), Step 3 Add endpoint groups, and Step 4 Add endpoints. The main area has a heading 'Add listeners' and a sub-instruction: 'A listener is a process that checks for connection requests that arrive to an assigned set of static IP addresses on a port or port range that you specify.' Below this is a 'Listeners' section with the instruction: 'You designate a listener by choosing a specific port or port range to listen on'. It shows two rows of configuration: Row 1 has ports 80, 443, 8080-8081, protocol TCP, client affinity None, and remove buttons. Row 2 has port 500, protocol UDP, client affinity None, and remove buttons. A yellow arrow points to the 'Client affinity Info' dropdown menu in the second row.

Traffic dial is from 0 – 100% of traffic you want to send to this workload

Listener: 80, 443, 8080-8081 TCP

Each listener can have multiple endpoint groups. Each endpoint group can only include endpoints that are in one Region. You aren't required to add an endpoint group, but until you do, traffic to this listener won't reach any endpoints.

Region Info	Traffic dial Info
us-west-2	100
▶ Configure health checks	
us-east-2	50
A number from 0 to 100.	
▶ Configure health checks	
Add endpoint group	

Listener: 500 UDP

Each listener can have multiple endpoint groups. Each endpoint group can only include endpoints that are in one Region. You aren't required to add an endpoint group, but until you do, traffic to this listener won't reach any endpoints.

Region Info	Traffic dial Info
us-west-2	100
▶ Configure health checks	
us-east-2	100
A number from 0 to 100.	
▶ Configure health checks	
Add endpoint group	

Create an endpoint group for every AWS Region where your application is hosted.

Health checks are for Elastic IPs and EC2 instances

ELB health checks are inherited from the load balancer

Health checks for EIPs can be TCP, HTTP or HTTPS

Remember UDP Health Checks are not Supported at this time.

Health check port
The port to use when Global Accelerator performs health checks on endpoints that are part of this endpoint group. The default is the port for the listener that the endpoint group is associated with. If listener port is a list, the first specified port in the list is used.

A number from 1 to 65535.

Health check protocol
The protocol to use when Global Accelerator performs health checks on endpoints that are part of this endpoint group.

▼

Health check path
If the protocol is HTTP/S, provide the ping path for the endpoints to be checked.

A path name from 1 and 1024 characters.

Health check interval
The interval, in seconds, between health checks for each endpoint.

▼

Threshold count
The number of consecutive health checks required before considering an unhealthy target healthy or a healthy target unhealthy.

A number from 1 to 10.

Available resource ARNs will appear in the endpoint info dropdown

Listener: 80, 443 TCP
AWS Global Accelerator routes traffic that arrives on these ports to endpoints in regional endpoint groups. All endpoints for an endpoint group must be in the same Region.

▼ Endpoint group: us-west-2
Traffic dial: 0%

Endpoint type	Info	Endpoint Info	Weight Info
Application Load ...	▼	arn:aws:elasticclo...	128
Application Load Balancer	info		A number from 0 to 255.
Network Load Balancer			es the client IP address for internet-facing Application Load Balancers unless you clear the feature. All internal Application Load Balancers and EC2 instances automatically preserve the client IP address. Make sure that your endpoints are configured to accept traffic from the preserved client IP addresses.
EC2 instance			
Elastic IP address			
<input checked="" type="checkbox"/> Preserve client IP address			

Weight lets you portion your traffic to a resource in the endpoint group

Listener: 500 UDP
AWS Global Accelerator routes traffic that arrives on these ports to endpoints in regional endpoint groups. All endpoints for an endpoint group must be in the same Region.

▼ Endpoint group: us-west-2
Traffic dial: 100%

Endpoint type	Info	Endpoint Info	Weight Info
EC2 instance	▼	i-000a8bd9aa6aa...	128
EC2 instance			A number from 0 to 255.
Elastic IP address	info		Global Accelerator preserves the client IP address for internet-facing Application Load Balancers unless you clear the check box to disable the feature. All internal Application Load Balancers and EC2 instances automatically preserve the client IP address. Make sure that your endpoints are configured to accept traffic from the preserved client IP addresses.
<input checked="" type="checkbox"/> Preserve client IP address			
Add endpoint			

All of the Region's endpoint options will show in the dropdown

Now the Global Accelerator is up and deployed.

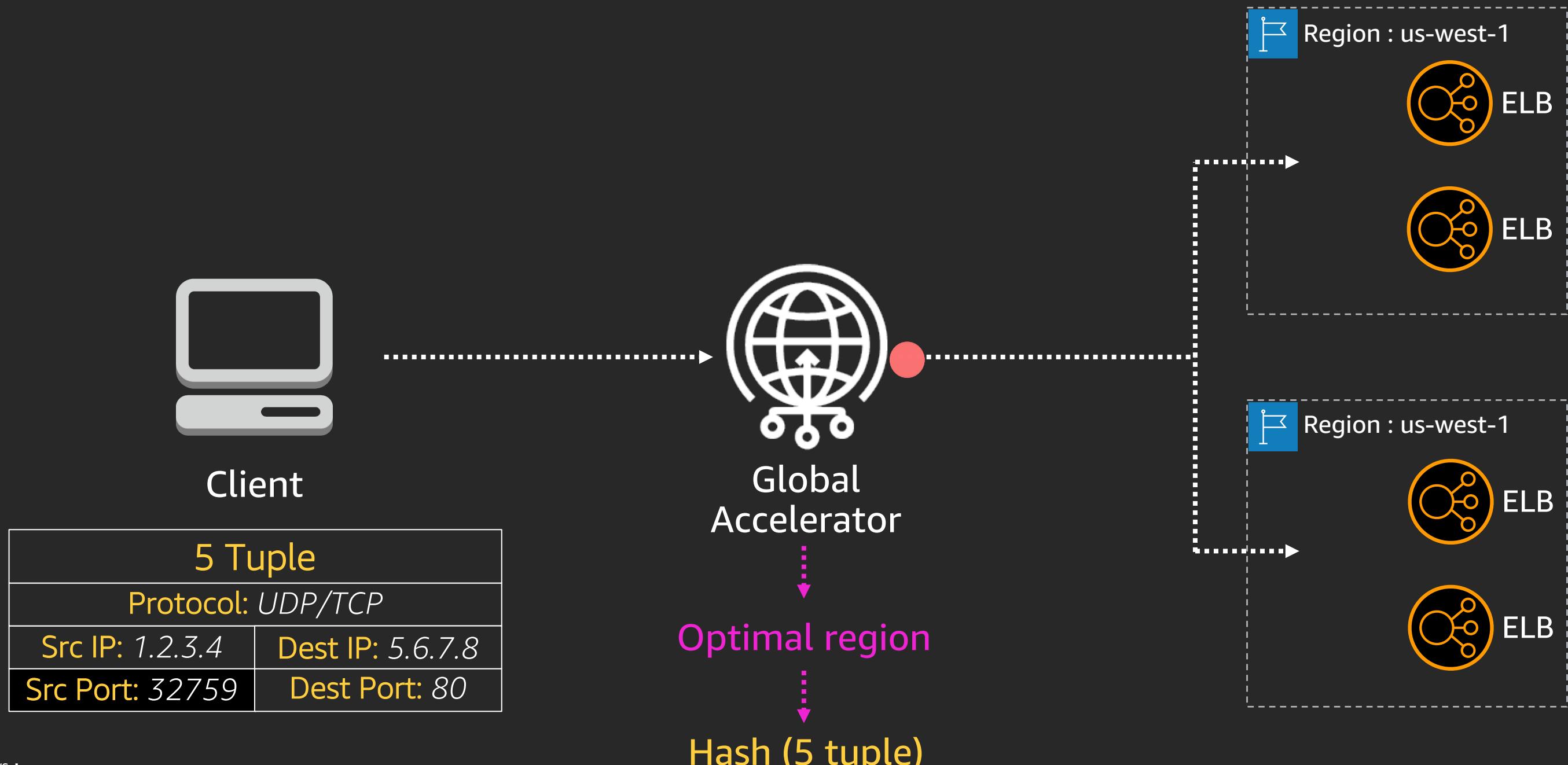
Accelerators (1)						View details	Edit	Delete			Create accelerator		
						<input type="text"/> Find accelerators							
Name	Static IP addresses	Enabled	Status	Created	Edited								
	DemoAccelerator	13.248.157.111, 76.223.21.30	On	Deployed	Thursday, May 23, 2019 2:41 PM GMT	Thursday, May 23, 2019 2:41 PM GMT							



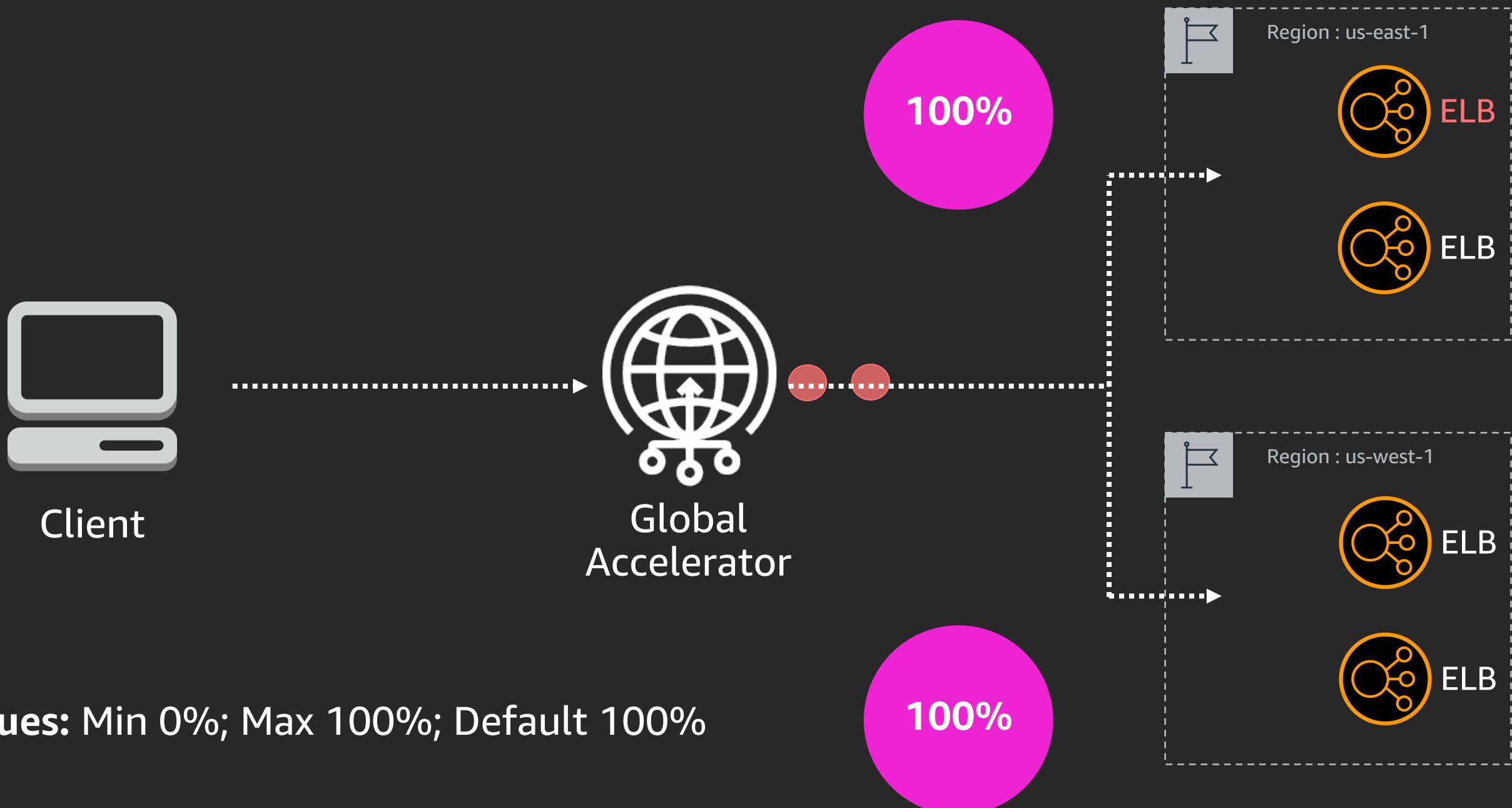
Allocated a set of IP addresses. Each is in an isolated network zone. If one zone fails, the other zone takes over.

Traffic control

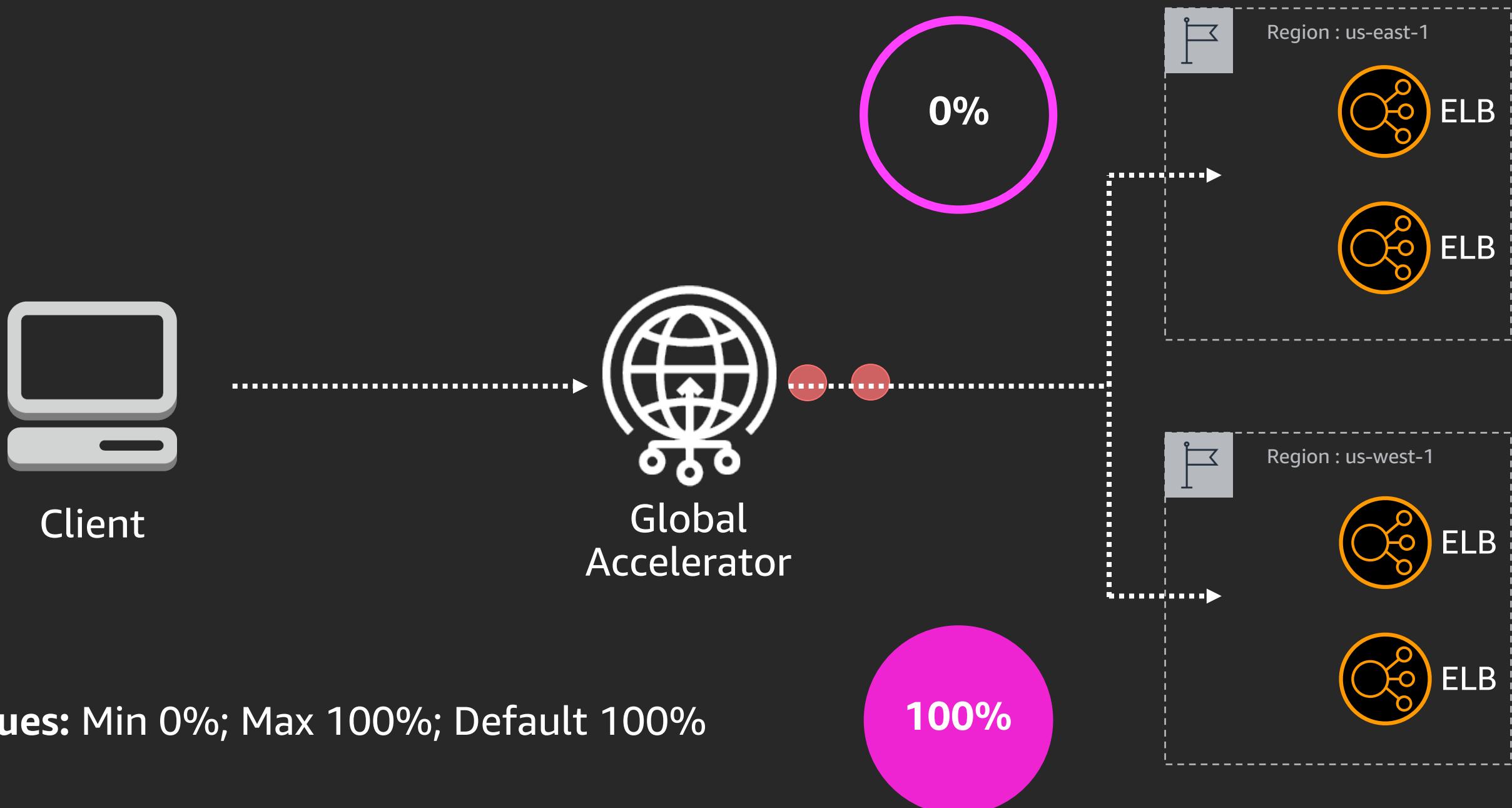
Easy traffic control – Optimal endpoint selection



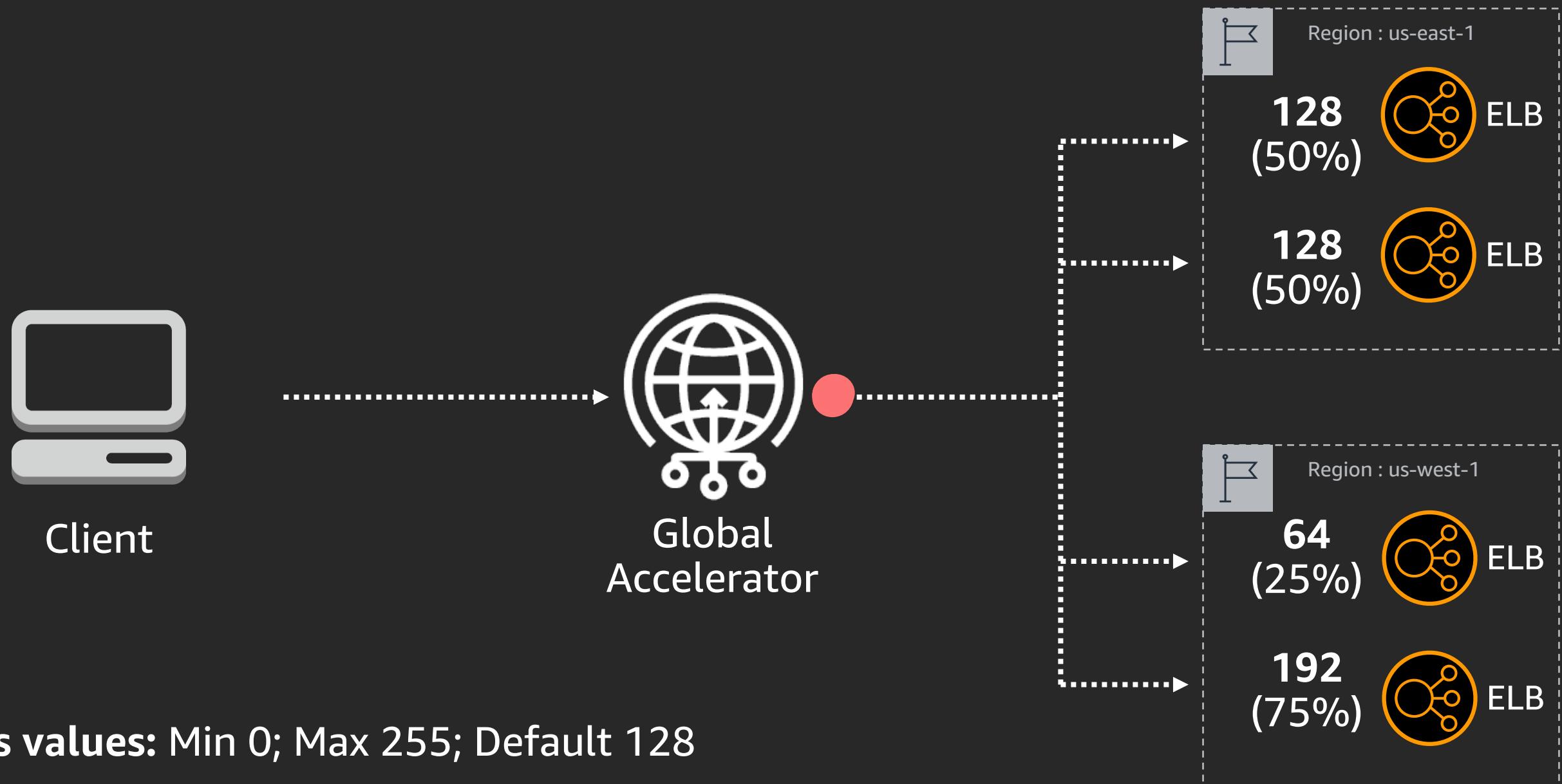
Easy traffic control – Regional traffic dials



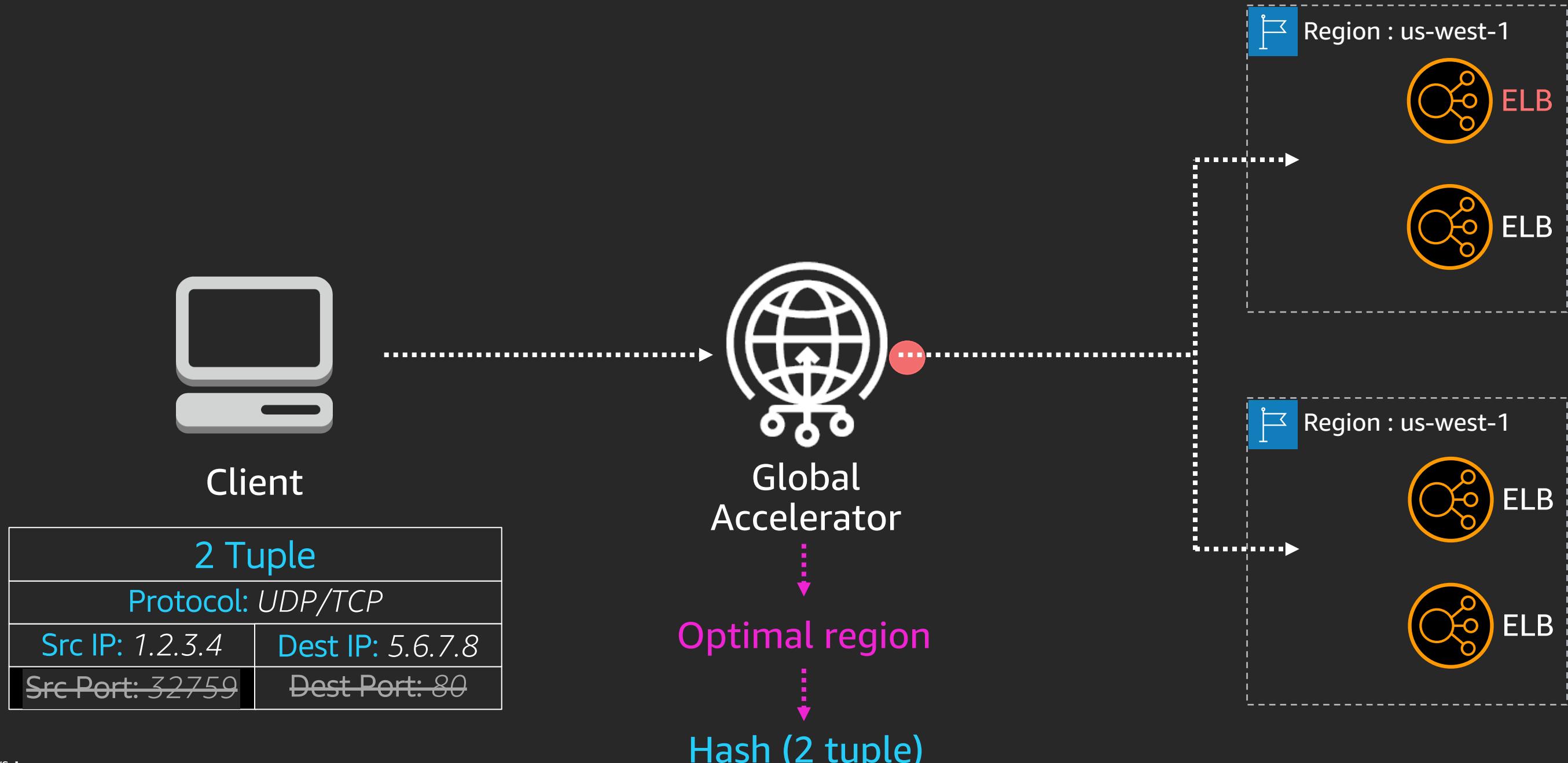
Easy traffic control – Regional traffic dials



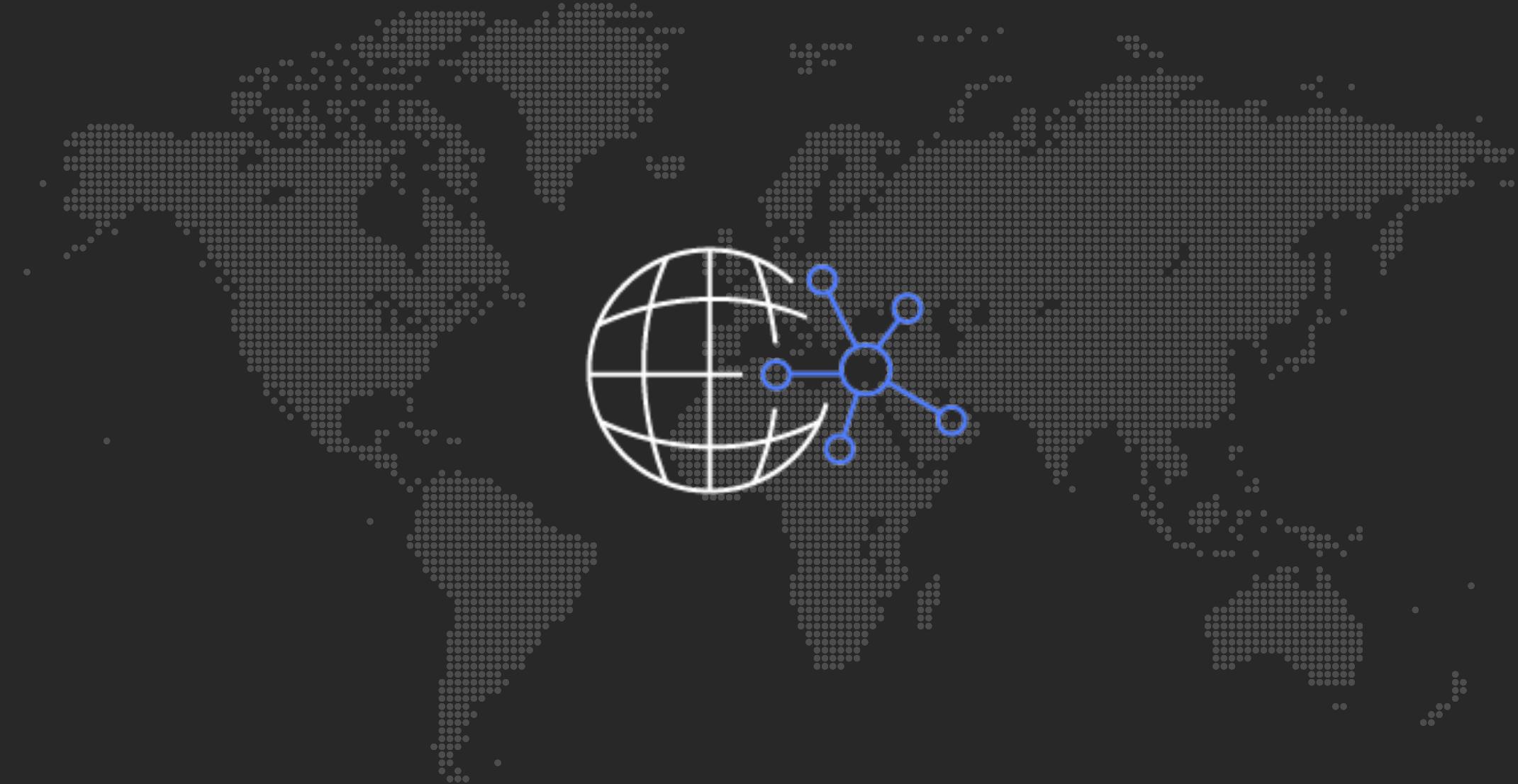
Easy traffic control – Endpoint weights



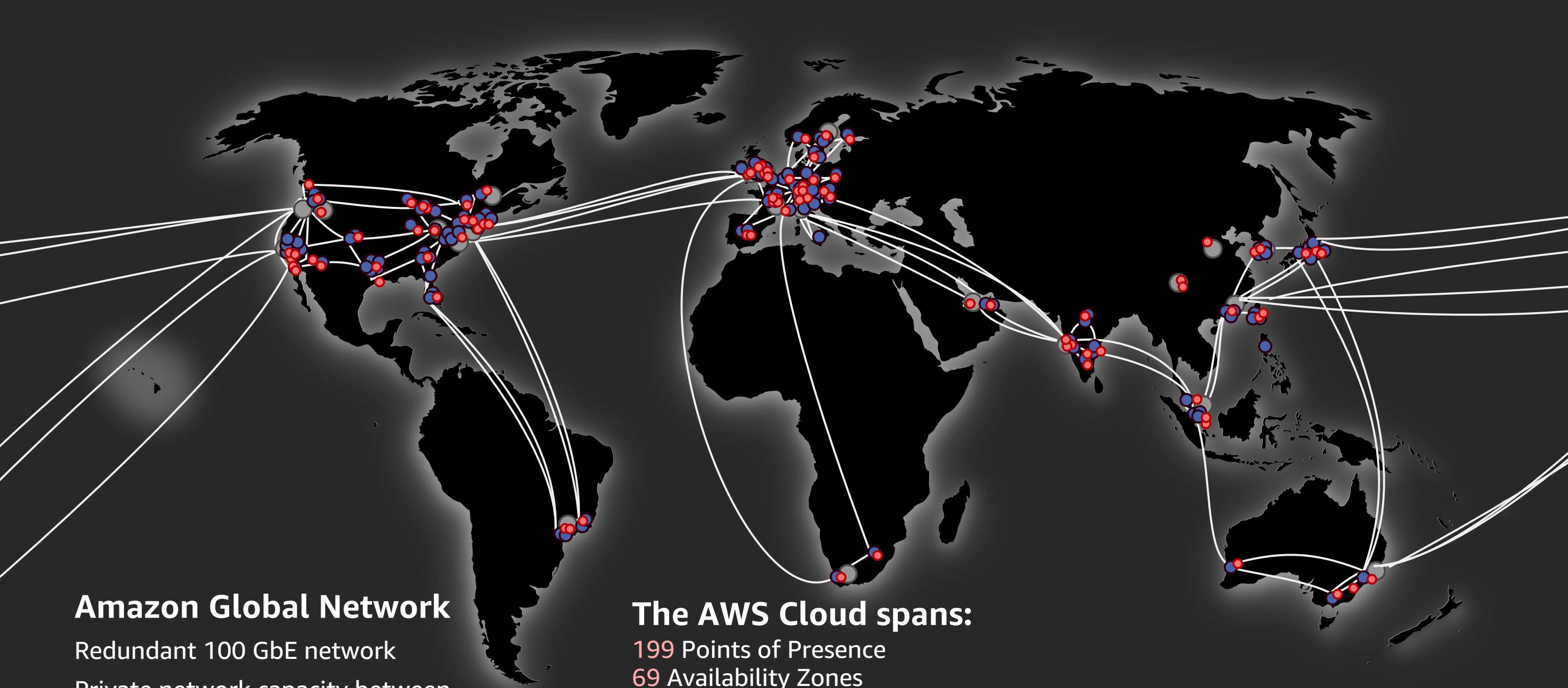
Easy traffic control – Client affinity



Propagation worldwide takes about 30 seconds



Security



Amazon Global Network

Redundant 100 GbE network

Private network capacity between
all AWS Regions, except China

The AWS Cloud spans:

199 Points of Presence

69 Availability Zones

22 Geographic AWS Regions around the world*

*With announced plans for 13 more Availability Zones and four more Regions in
[Cape Town](#), [Jakarta](#), [Milan](#), and [Spain](#).

Security



Data
encryption key

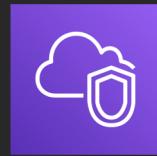
- Network level encryption

Security



Data
encryption key

- Network level encryption



Amazon VPC

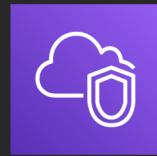
- Security groups

Security



Data
encryption key

- Network level encryption



Amazon VPC

- Security groups



AWS WAF

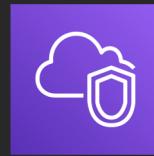
- AWS WAF on ALB

Security



Data
encryption key

- Network level encryption



Amazon VPC

- Security groups



AWS WAF

- AWS WAF on ALB

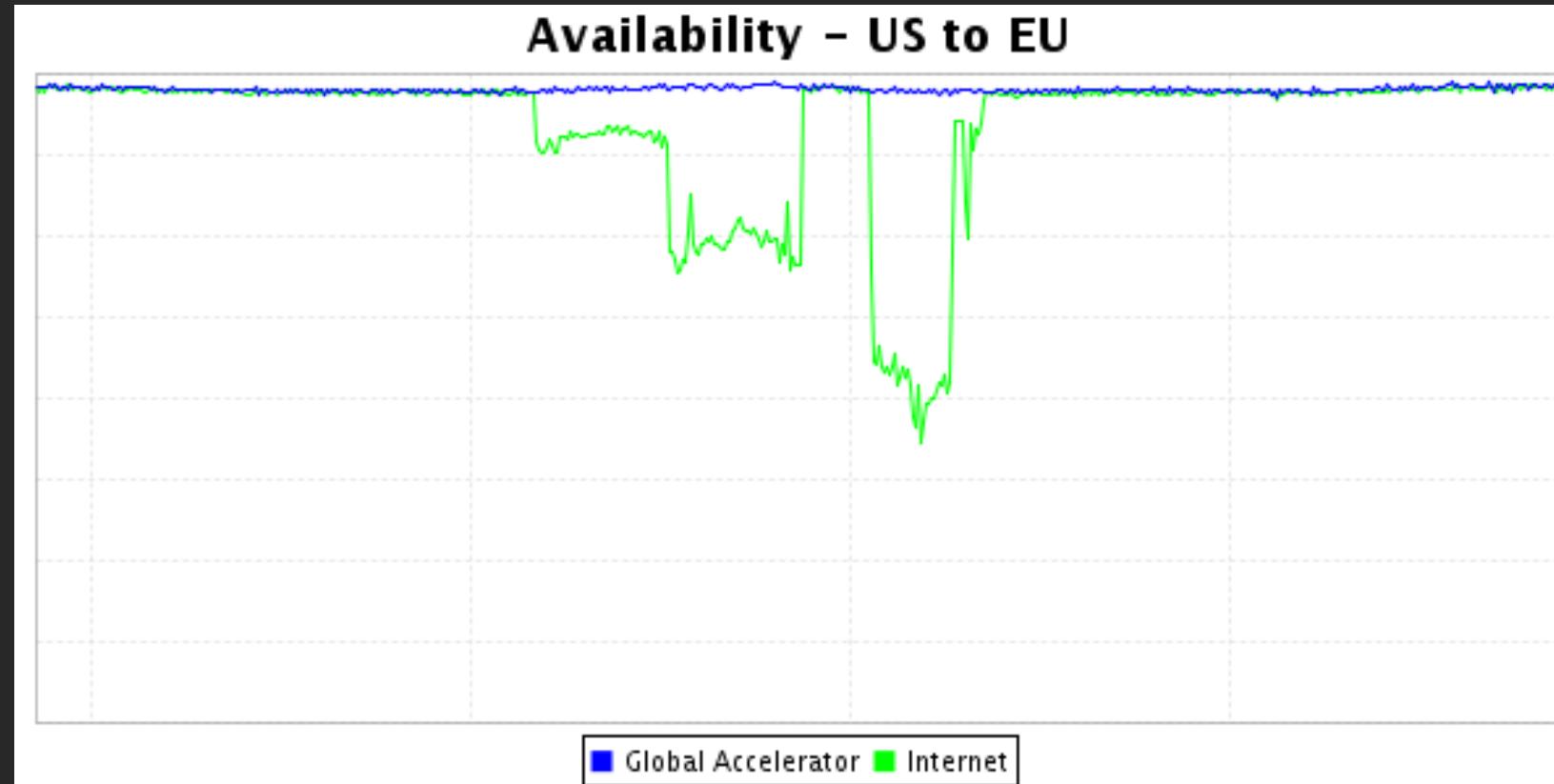


Amazon EC2

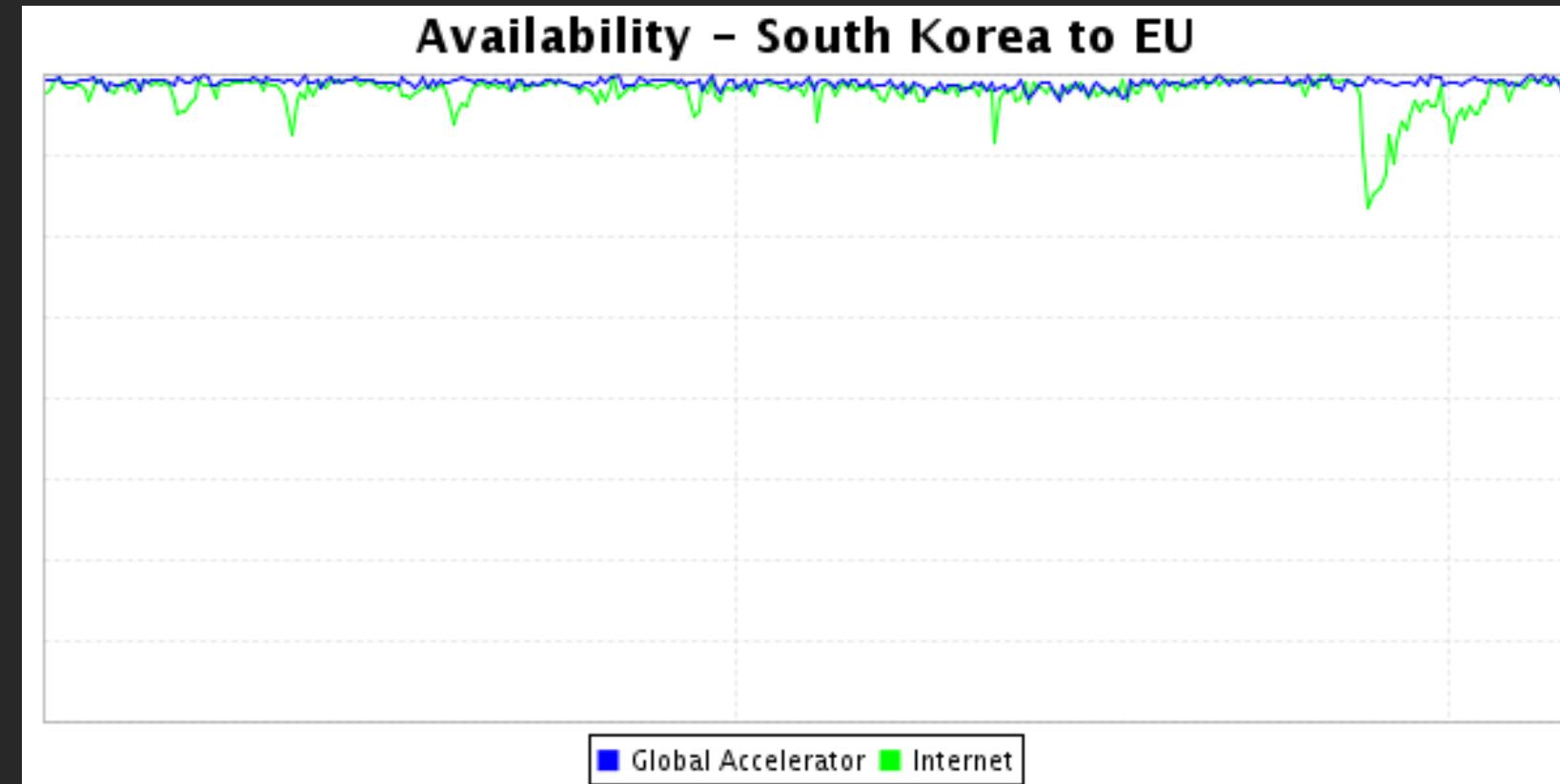
- Internal EC2 instances and ALBs

Availability

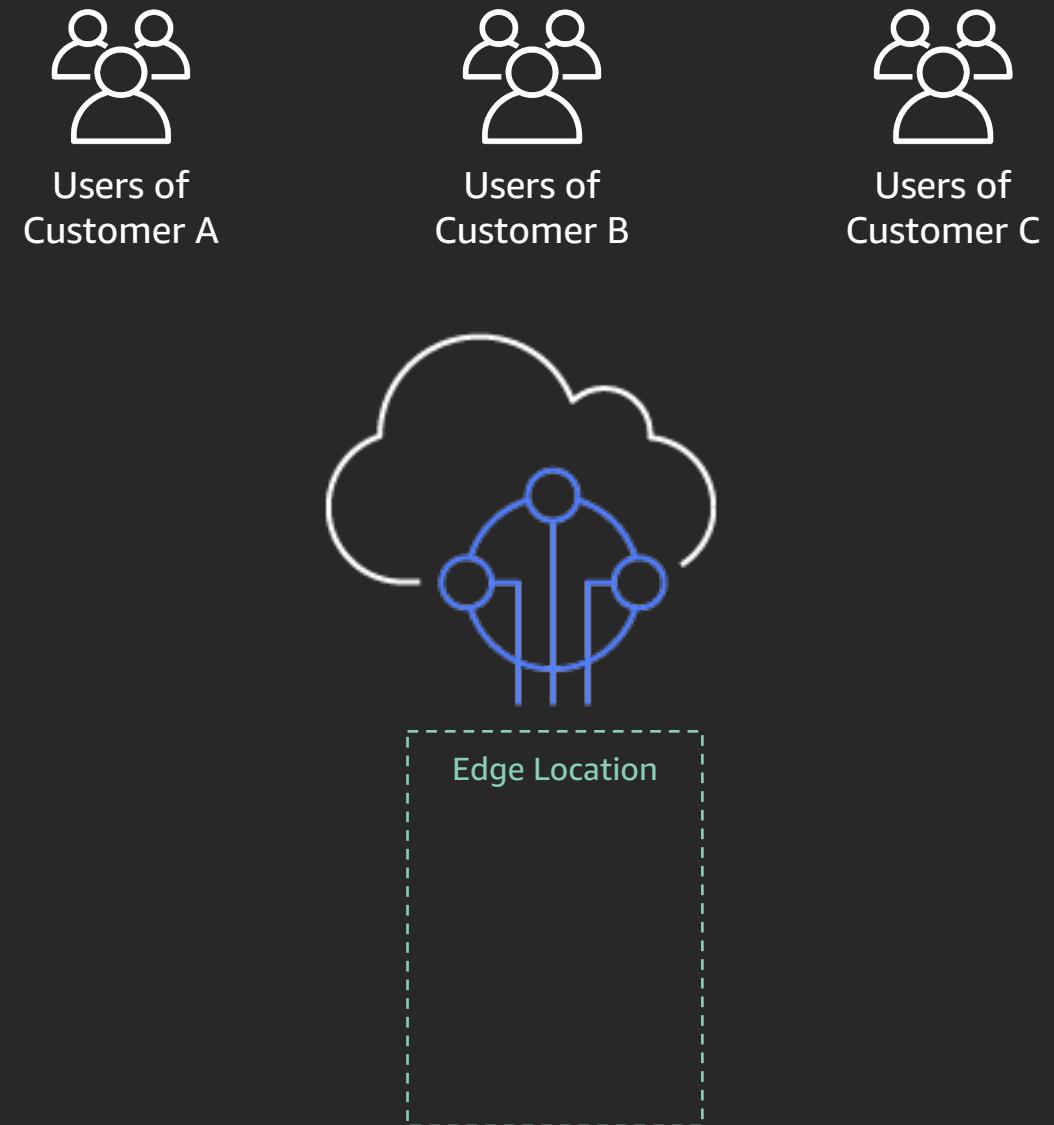
Availability



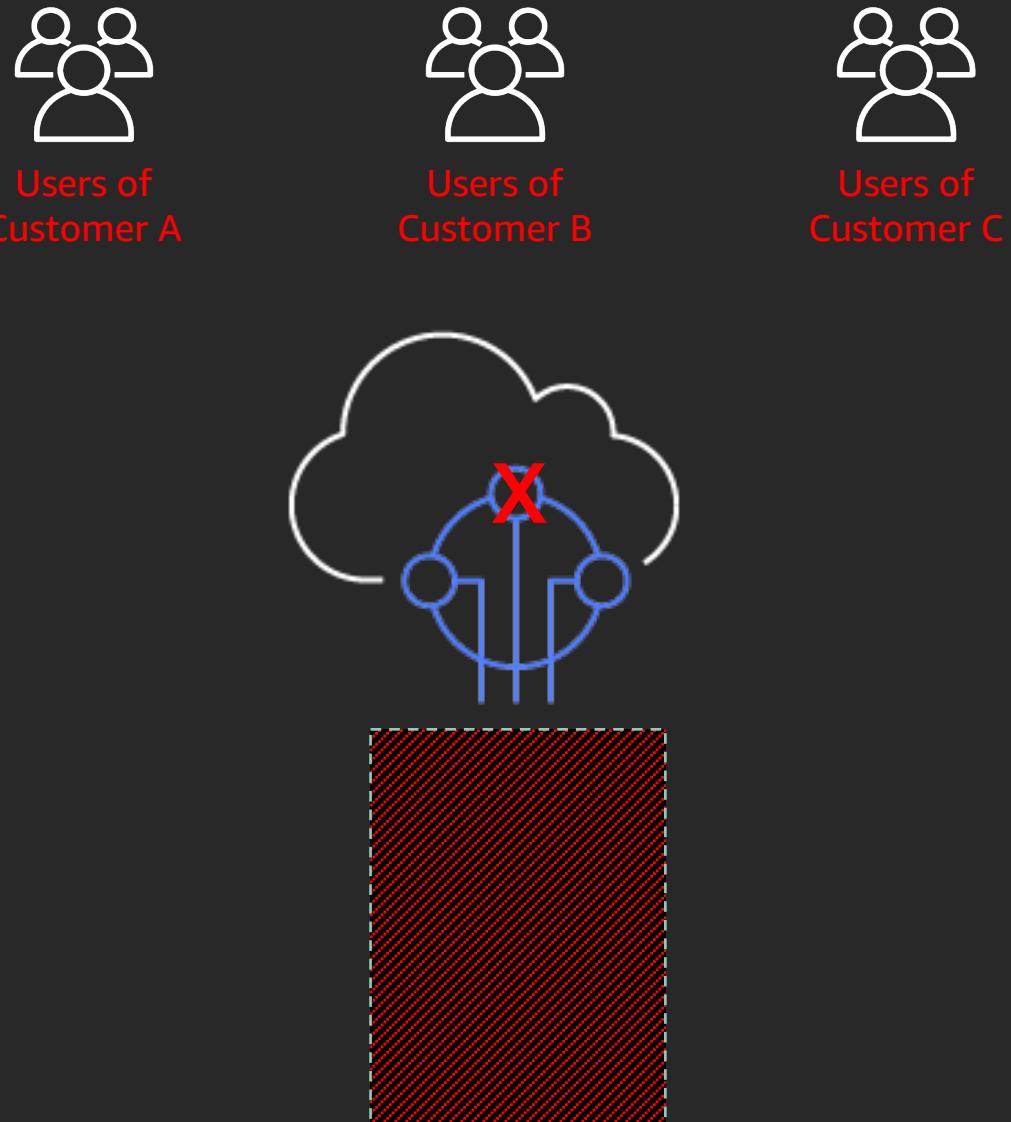
Availability



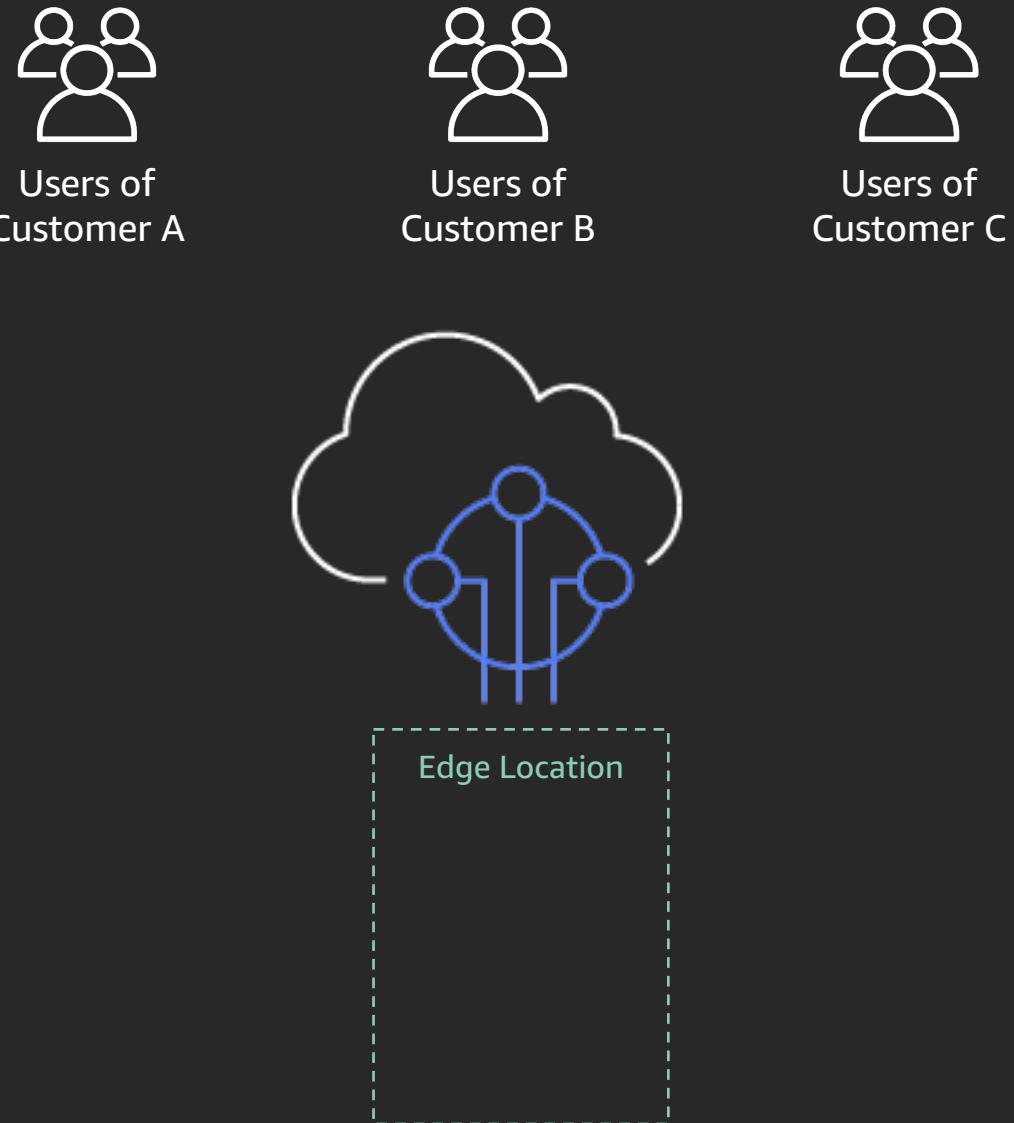
Edge location – BGP announcements



Network events



Heavy workloads



Heavy workloads – noisy neighbors



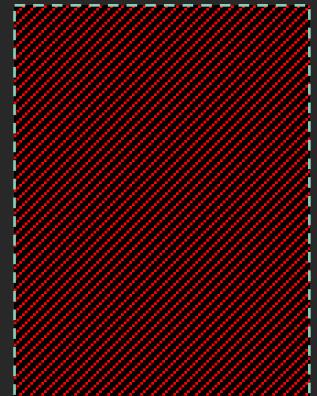
Users of
Customer A



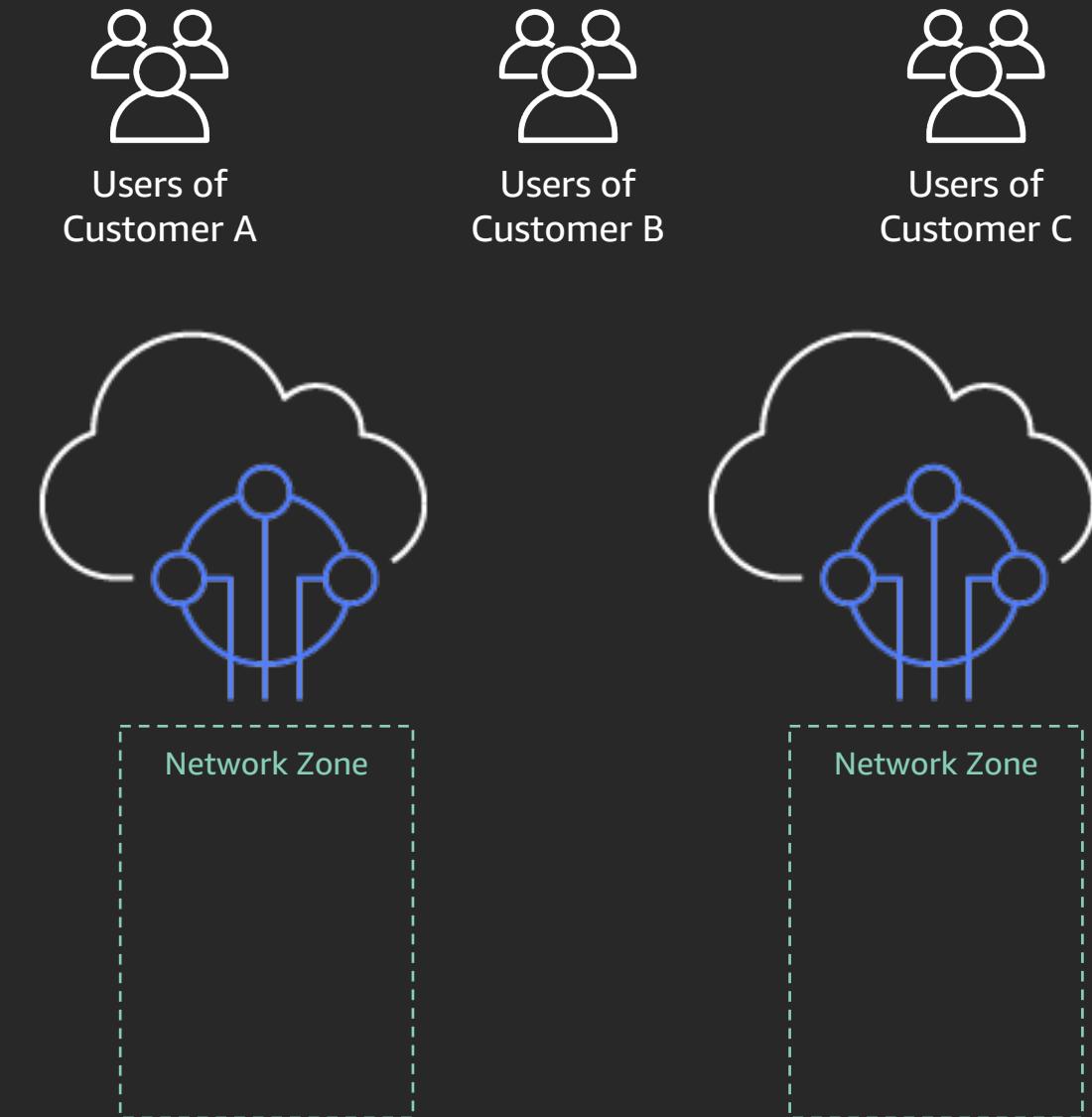
Users of
Customer B



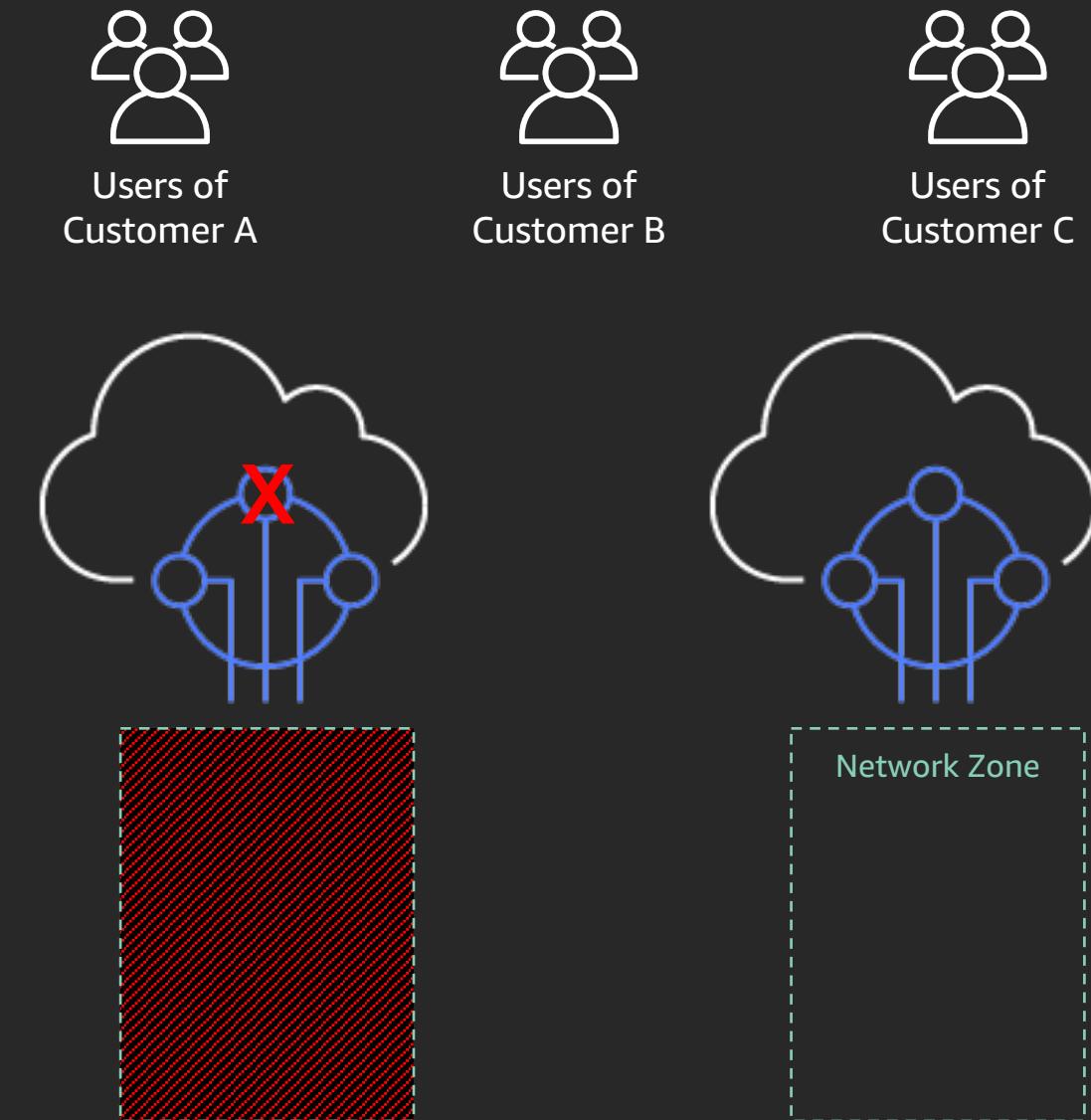
Users of
Customer C



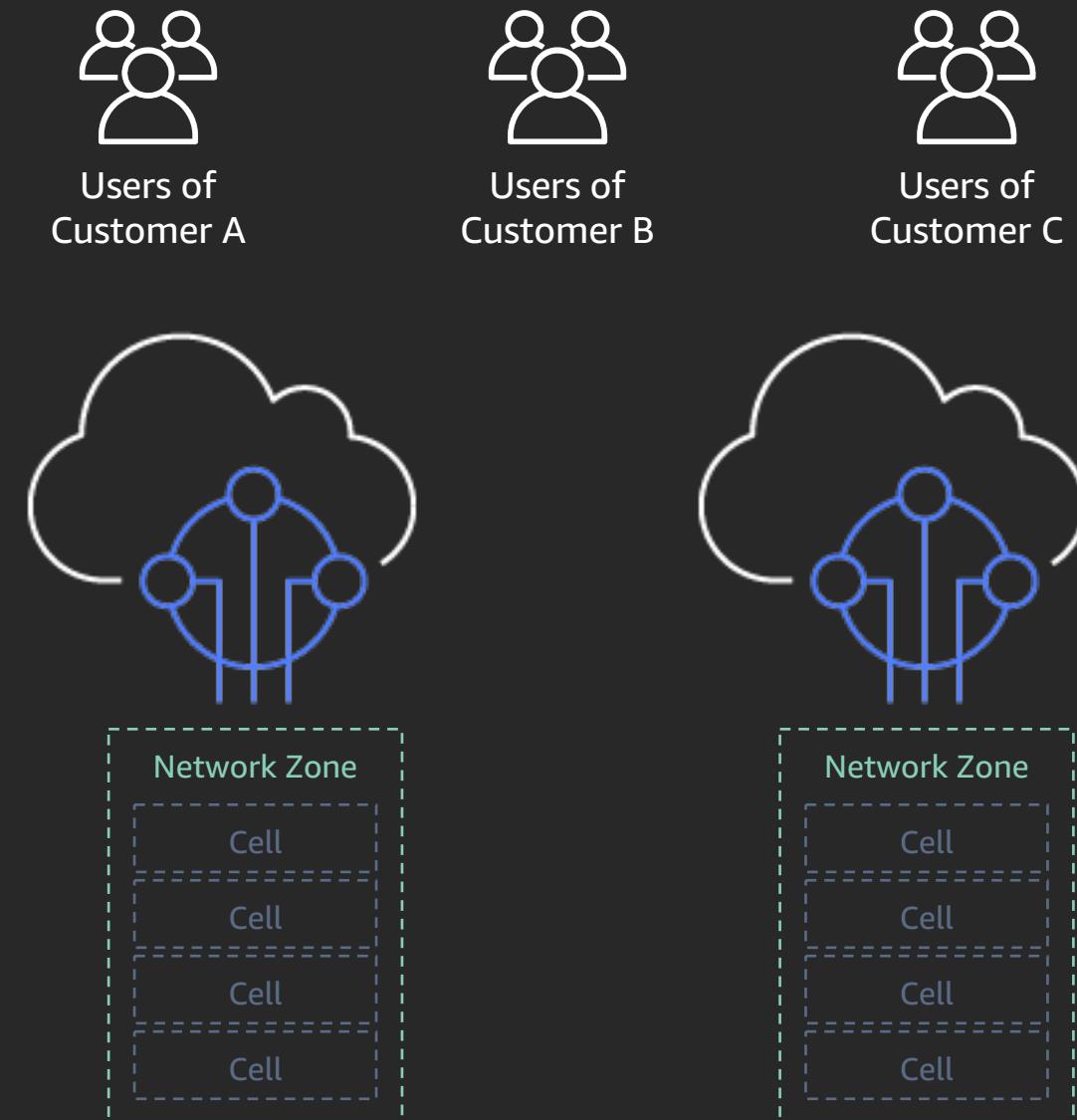
Network zones



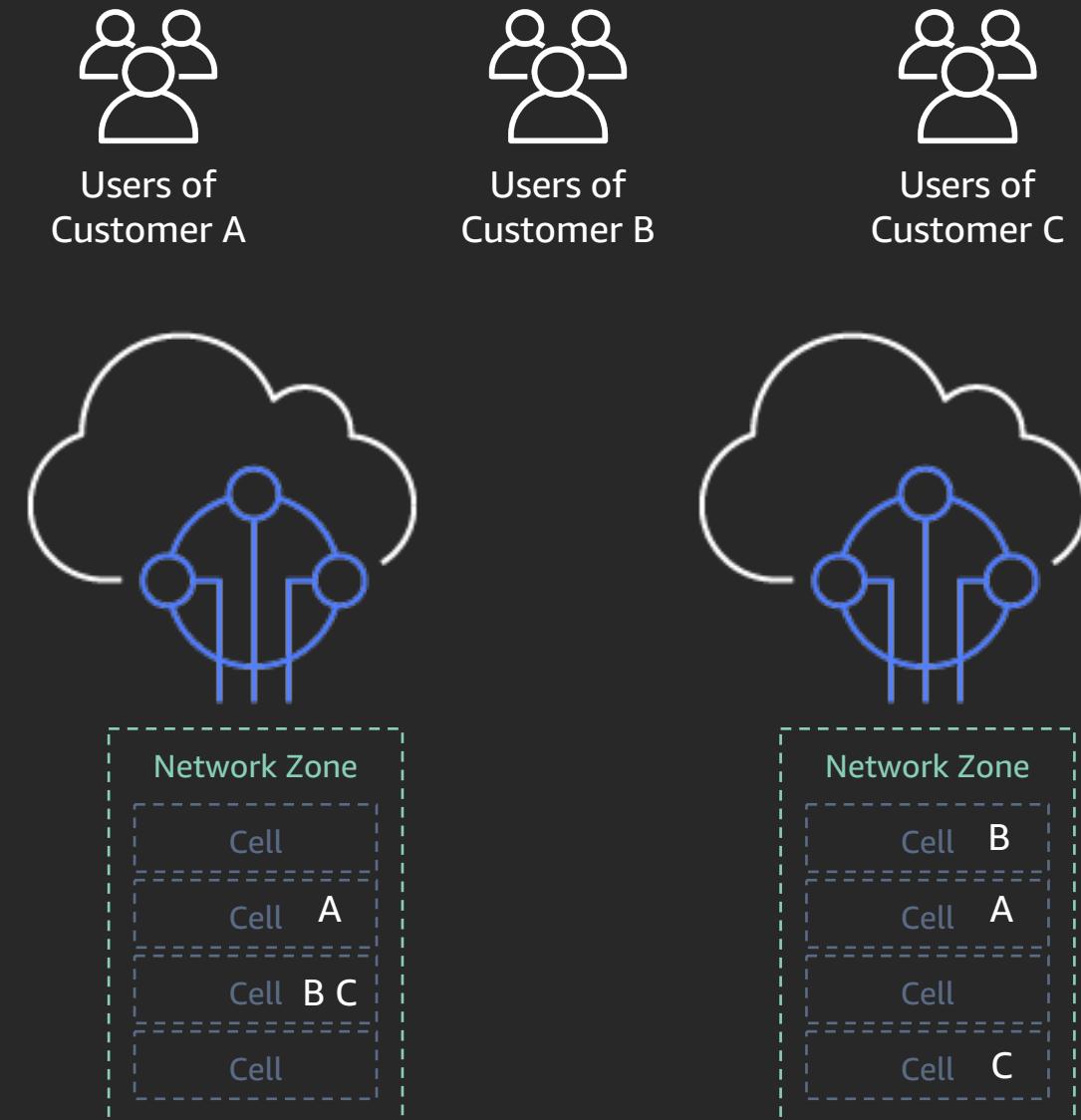
Network event affecting one zone



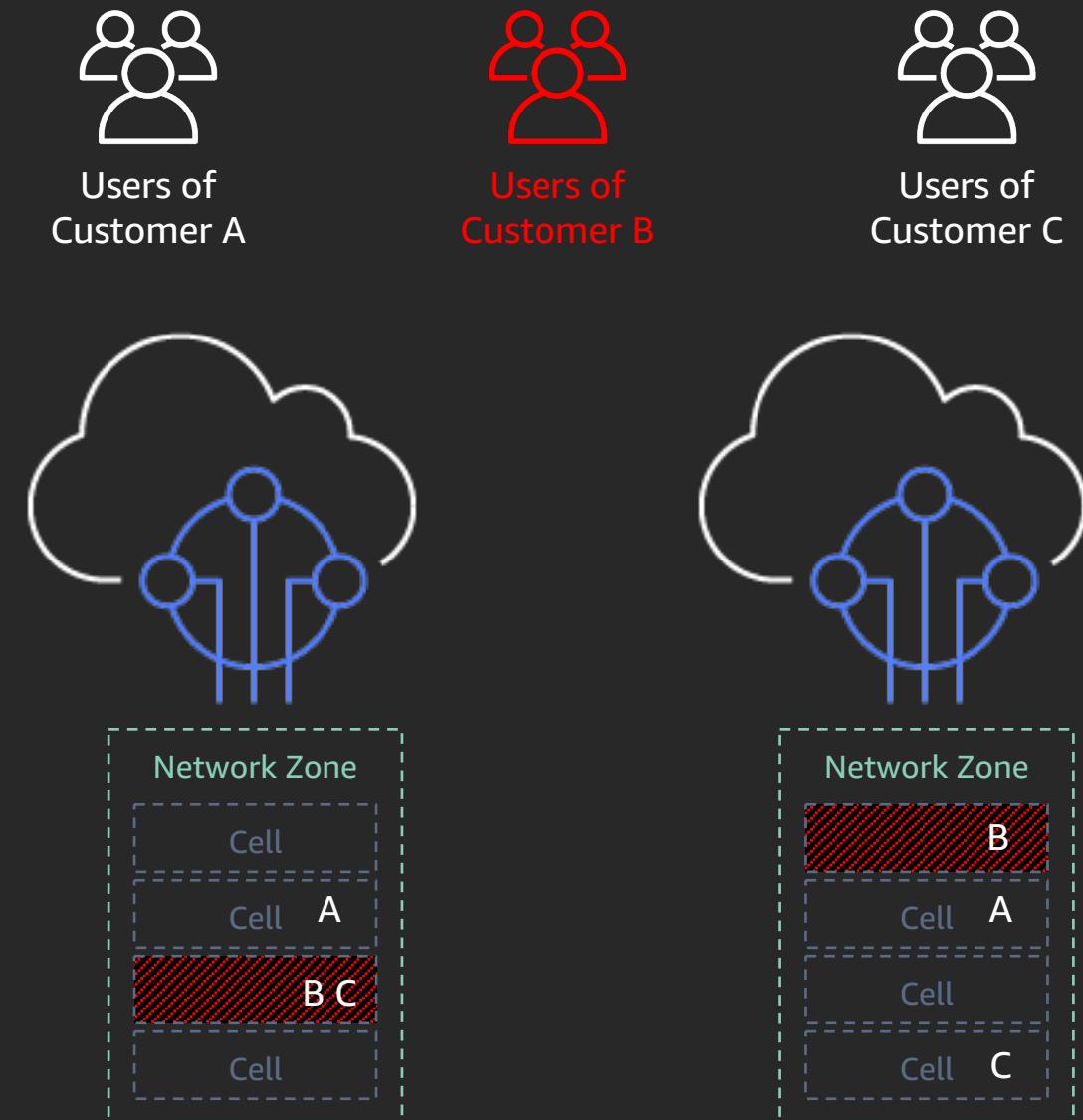
Cells per zone



Cells per zone



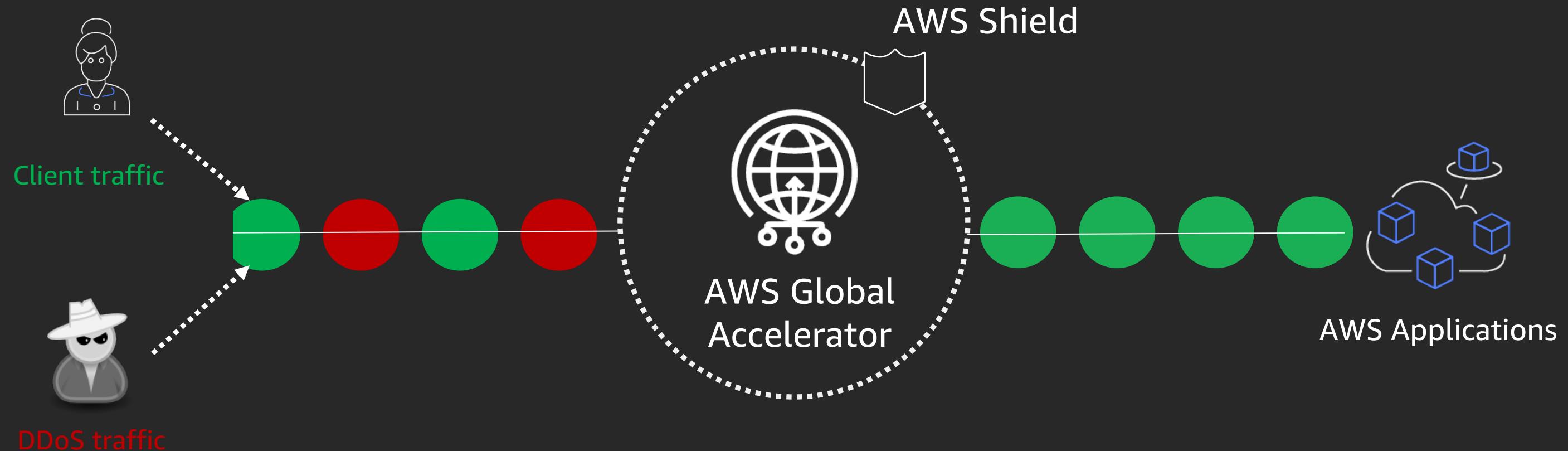
Noisy neighbor affecting one cell per zone



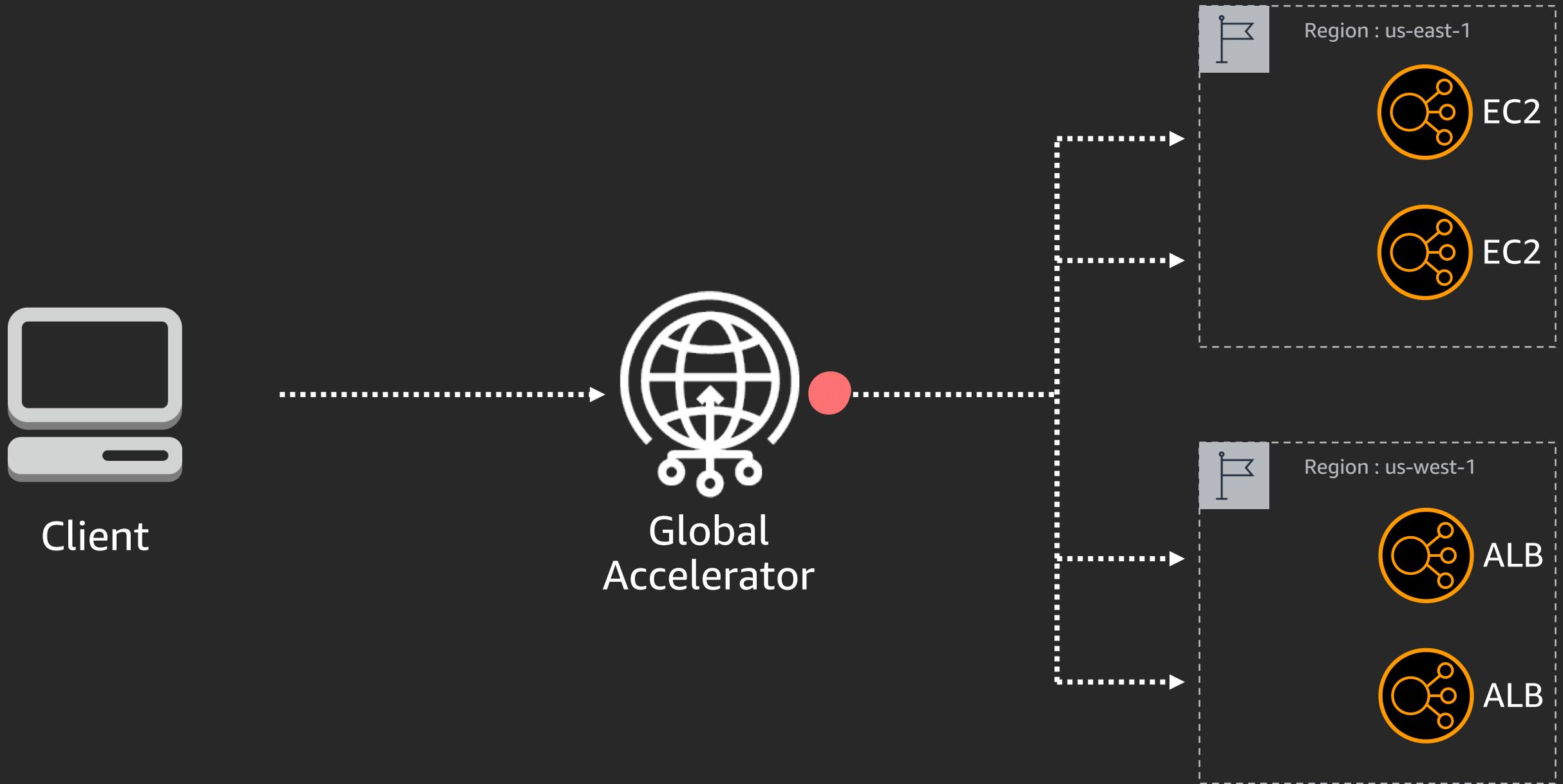
Blast radius reduction

	1 Network Zone	2 Network Zones	4 Network Zones
2 Cell Per NZ	$1/2^1 = 50\%$	$1/2^2 = 25\%$	$1/2^4 = 6.25\%$
4 Cells Per NZ	$1/4^1 = 25\%$	$1/4^2 = 6.25\%$	$1/4^4 = 0.4\%$
8 Cells Per NZ	$1/4^1 = 12.5\%$	$1/8^2 = 1.56\%$	$1/8^4 = 0.02\%$

AWS Shield will safeguard against L3/L4 DDoS attacks



Availability across failures



Network Address Translation



	Source	Destination
IP	1.2.3.4	5.6.7.8
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

Network Address Translation



	Source	Destination
IP	1.2.3.4	5.6.7.8
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

	Source	Destination
IP	9.8.7.6	10.2.3.4
TCP	45678	443
Payload	GET /fun.jpg HTTP/1.1	

Client IP preservation



	Source	Destination
IP	1.2.3.4	5.6.7.8
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

Client IP preservation



	Source	Destination
IP	1.2.3.4	5.6.7.8
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

Encapsulation

IP
Transport
Metadata
IP
TCP
Payload

Client IP preservation



	Source	Destination
IP	1.2.3.4	5.6.7.8
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

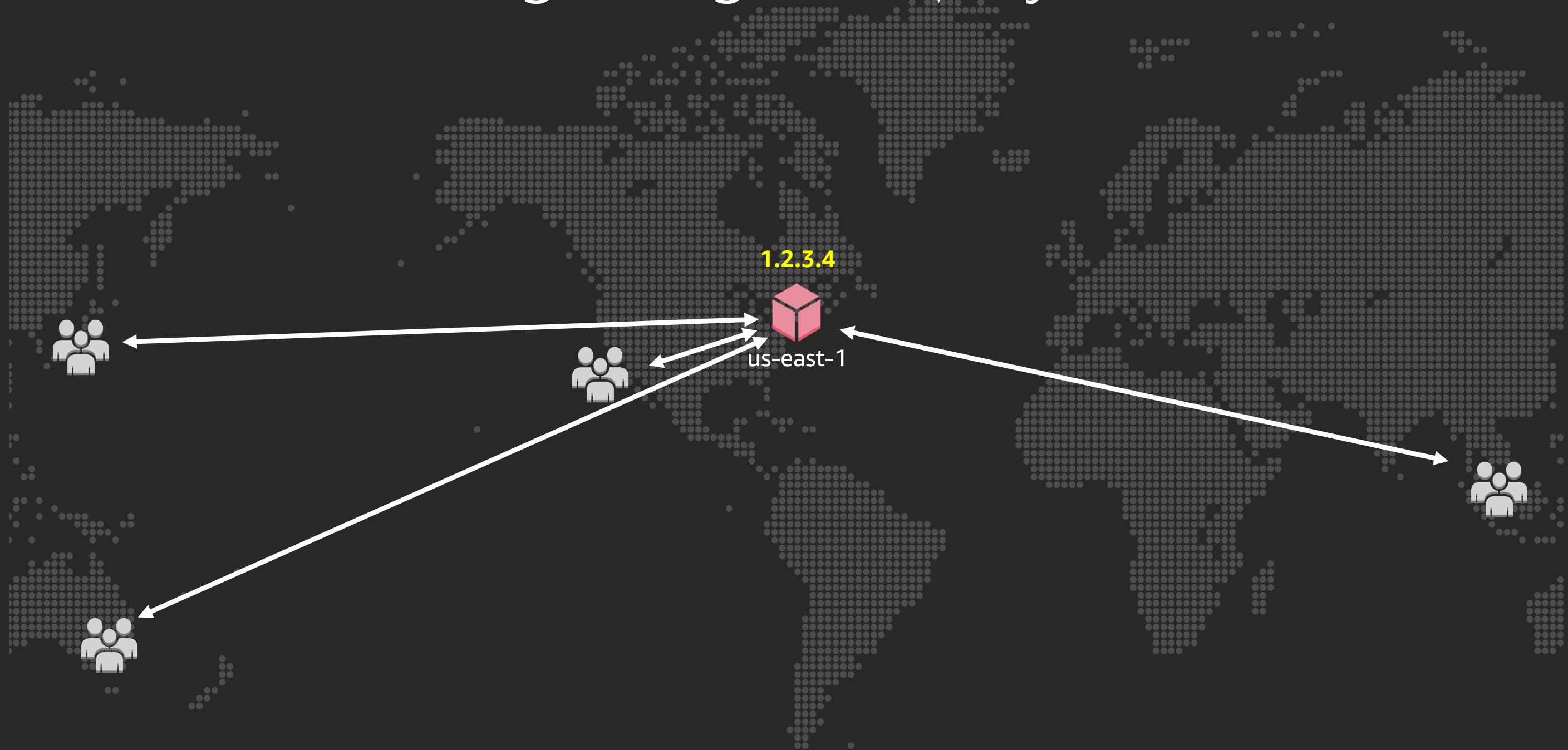
Encapsulation

IP
Transport
Metadata
IP
TCP
Payload

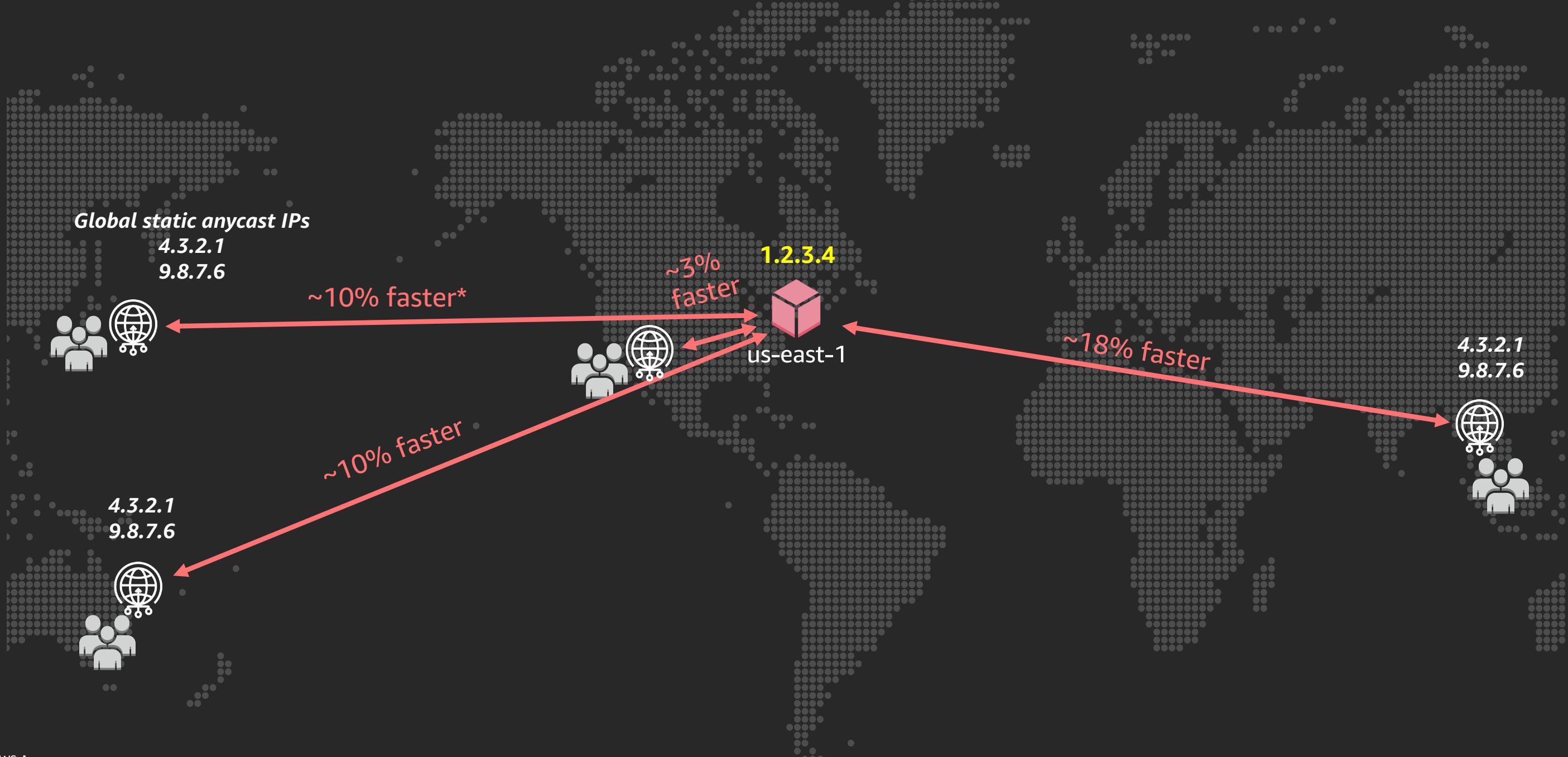
	Source	Destination
IP	1.2.3.4	10.2.3.4
TCP	32456	443
Payload	GET /fun.jpg HTTP/1.1	

Performance

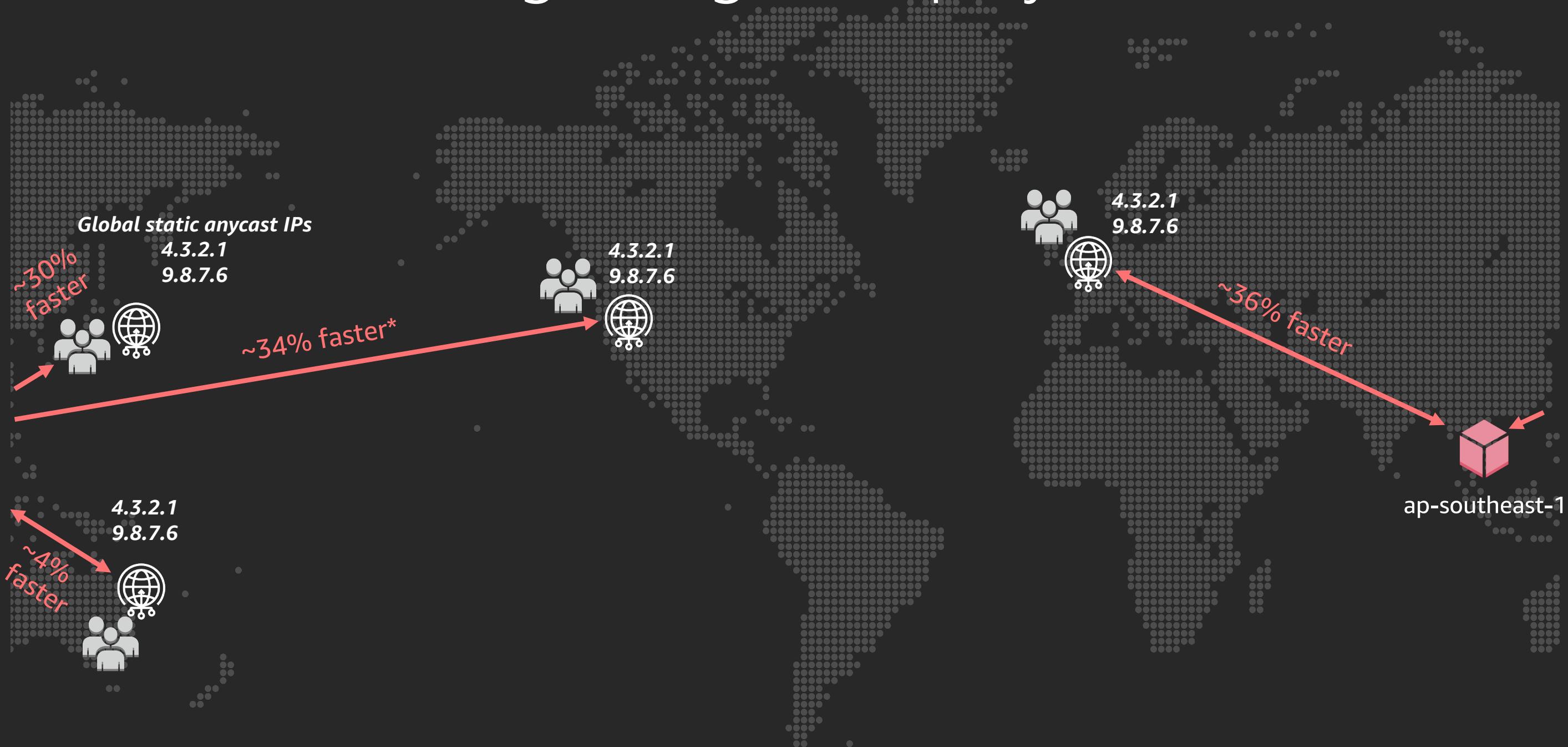
Performance – Single-Region deployments



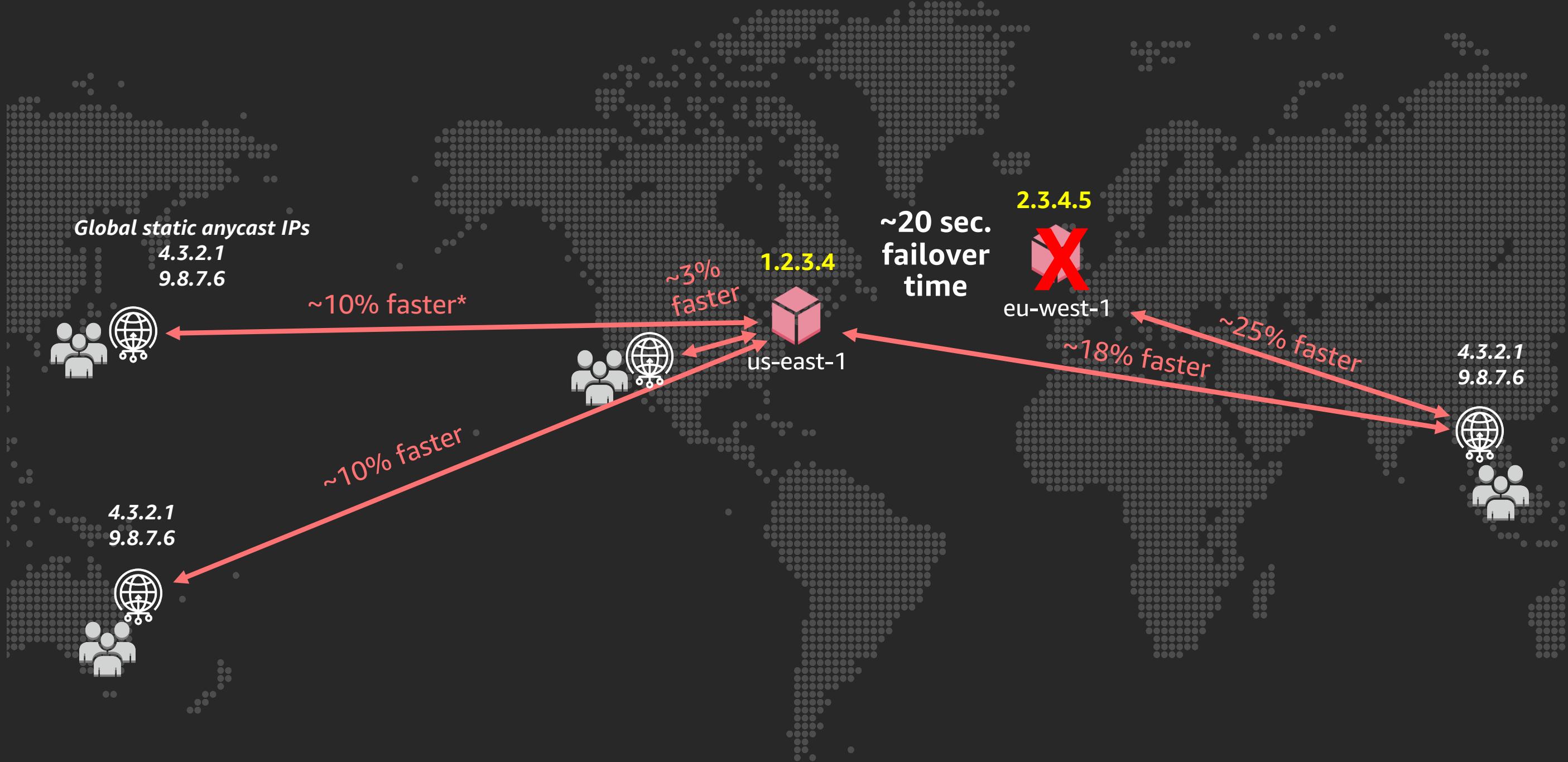
Performance – Single-Region deployments



Performance – Single-Region deployments



Performance – Plus failover



Building with Global Accelerator

Build an application



AWS Management
Console



Developer Tools

- CodeStar
- CodeCommit
- CodeBuild
- CodeDeploy
- CodePipeline
- Cloud9
- X-Ray

HTTP front end



AWS CodeCommit

Developer Tools > CodeCommit > Repositories > reysrv

reysrv

reysrv / app / webroot / todo / index.html [Info](#)

```
1 <!DOCTYPE html>
2 <html>
3 <head>
4   <title>todo</title>
5   <meta name="viewport" content="width=device-width,
```

HTTP front end



AWS CodeCommit

```
250   <table>
251     <tr>
252       <td>Categories:</td>
253       <td><span ng-repeat="viewName in viewNames | orderBy">
254     </td>
255     <tr>
256       <td align="right">Upcoming:</td>
257       <td>
258         <a href="" ng-click="viewToday()">Today</a>
259         <a href="" ng-click="viewTomorrow()">Tomorrow</a>
260         <a href="" ng-click="viewBacklog()">Backlog</a>
261       </td>
262     </tr>
263   </table>
```

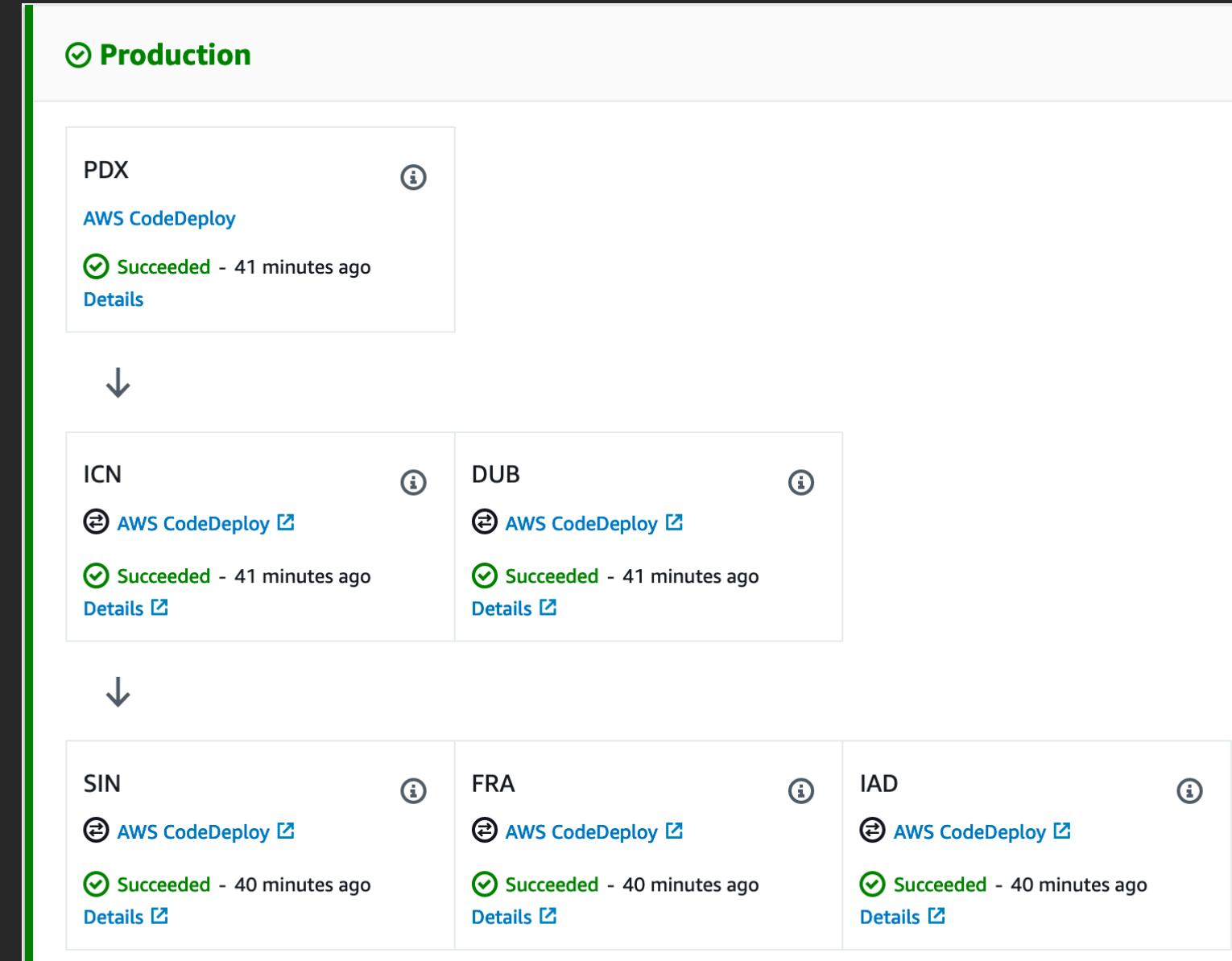
HTTP backend



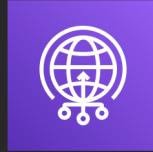
AWS CodeCommit

```
690     newVersion := *task.Version + 1
691     params := &dynamodb.UpdateItemInput{
692         TableName: aws.String("Task"),
693         Key: map[string]*dynamodb.AttributeValue{
694             "ViewID": {S: task.View},
695             "TaskID": {S: task.TaskID},
696         },
697         ConditionExpression: aws.String("Version = :version OR attribute_not_exists"),
698         UpdateExpression:   aws.String("SET Description = :desc, DoDate = :doDate,
699         ExpressionAttributeValues: map[string]*dynamodb.AttributeValue{
700             ":desc":      {S: task.Desc},
701             ":doDate":    {S: task.DoDate},
702             ":version":   {N: aws.String(strconv.Itoa(*task.Version))},
703             ":newVersion": {N: aws.String(strconv.Itoa(newVersion))},
704         },
705     }
```

Deploy worldwide



Set up your accelerator



AWS Global Accelerator

Listener: 80, 443 TCP

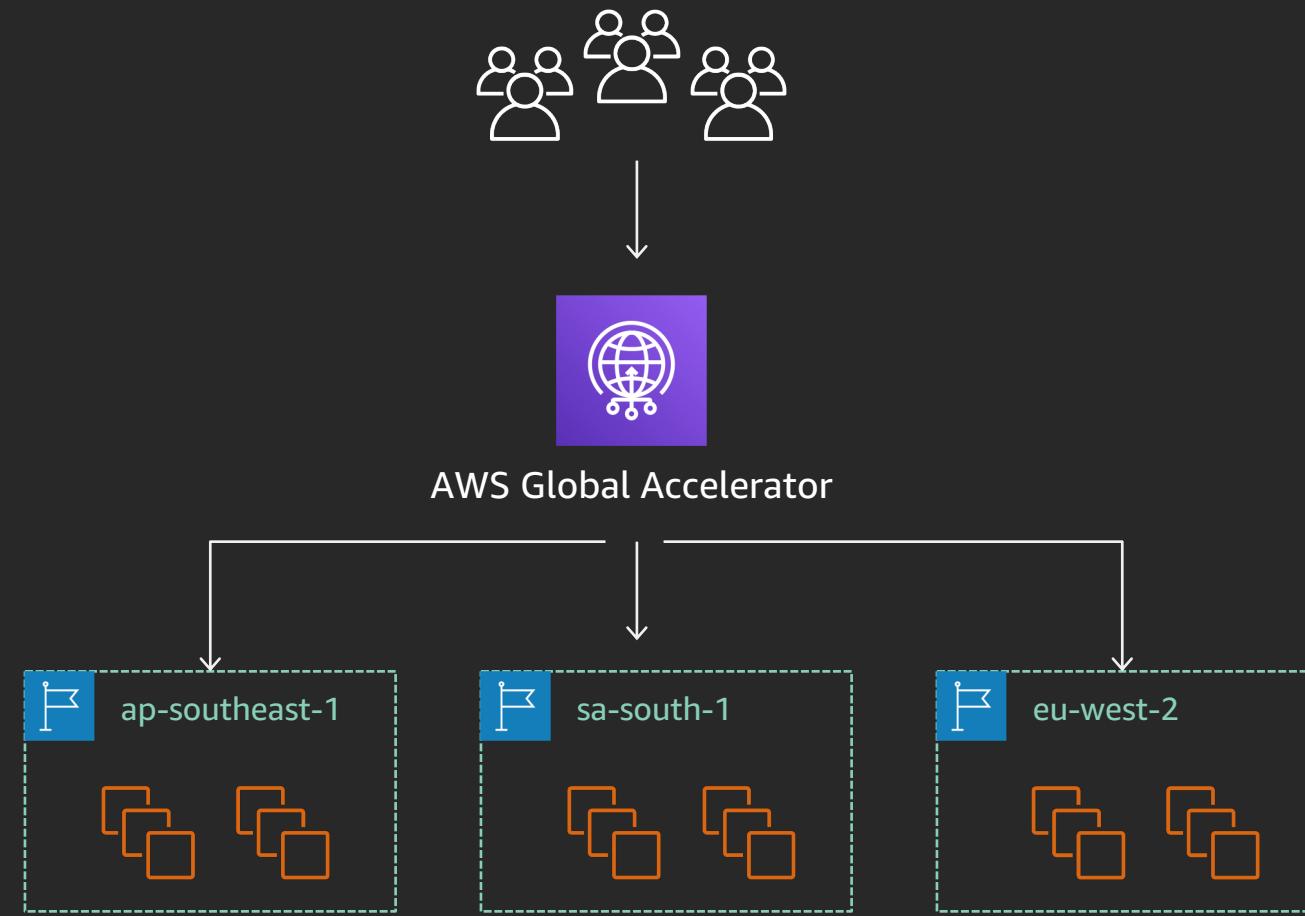
AWS Global Accelerator routes traffic that arrives on these ports the same Region.

▼ Endpoint group: ap-southeast-1

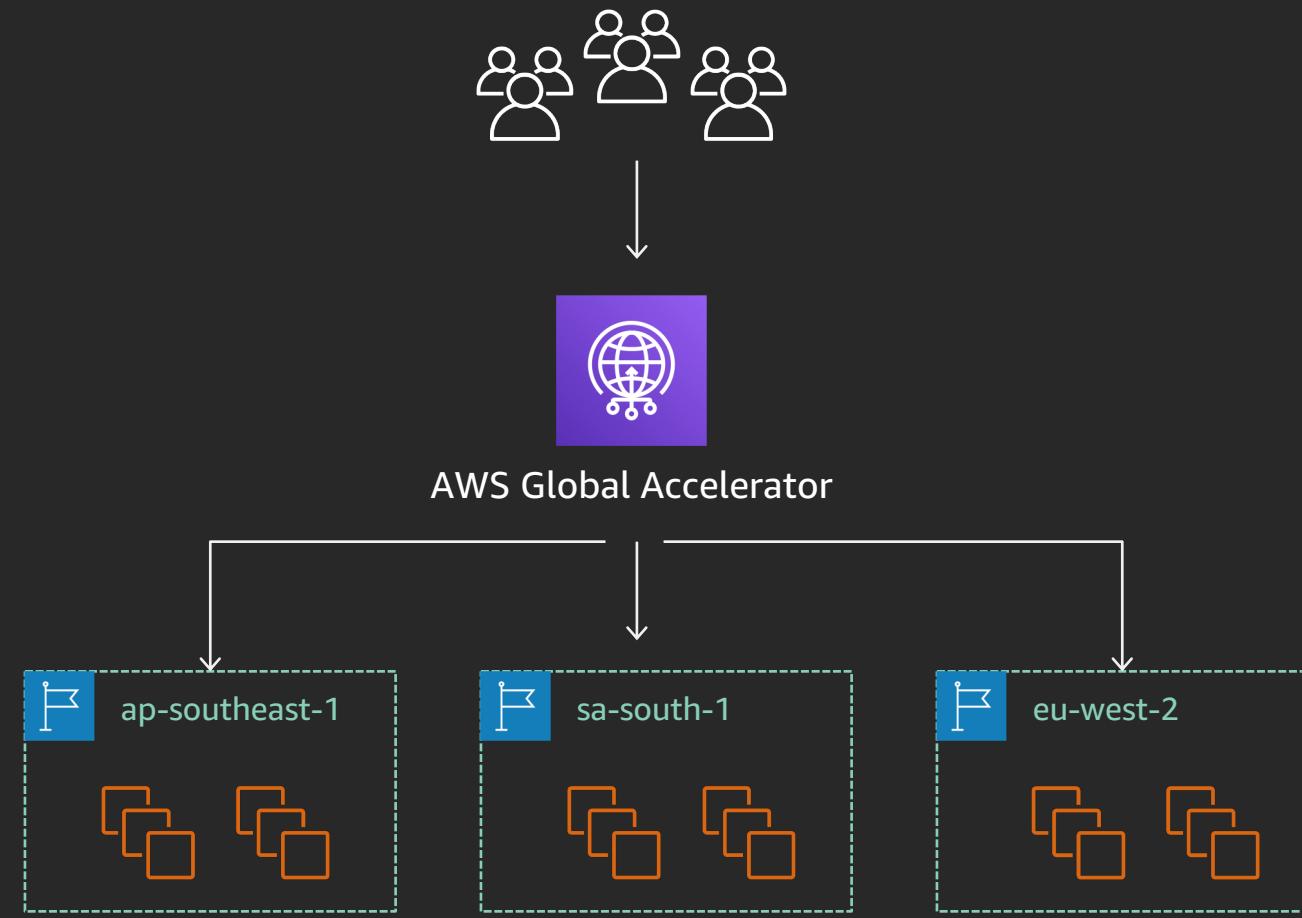
Traffic dial: 100%

Endpoint type Info	Endpoint Info
Choose endpoint	Choose endpoint
Application Load Balancer	
Network Load Balancer	
EC2 instance	
Elastic IP address	

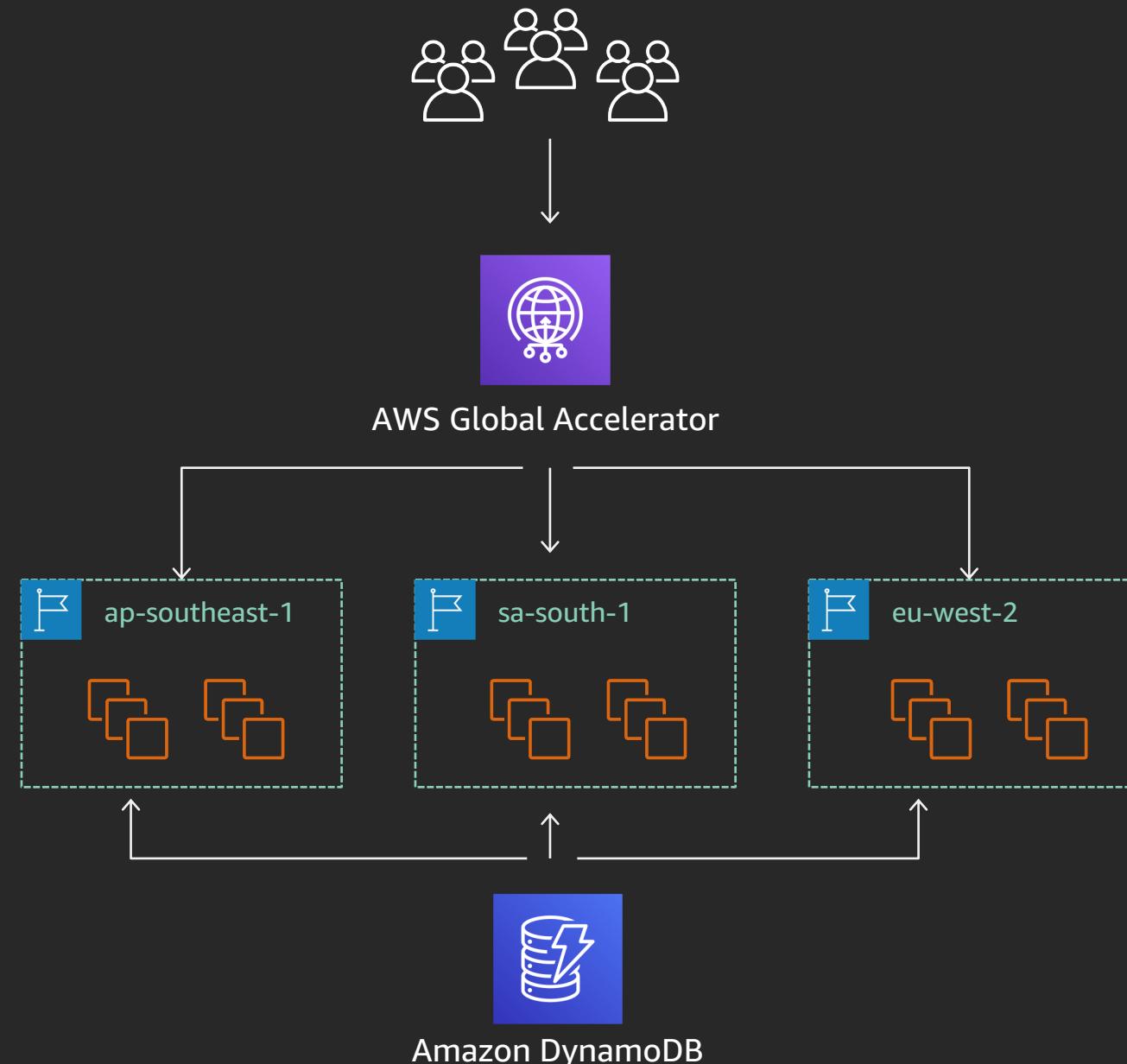
Multi-Region application



One thing missing...



Distributed state



Multi-Region application

Categories: [Reading List](#) [Weekend Plans](#)
Upcoming: [Today](#) [Tomorrow](#) [Backlog](#)

Weekend Plans

Sat Dec 7

- ✗ ↓ ↘ 1. Morning kayak
- ✗ ↓ ↘ 2. Roller rink!
- ✗ ↓ ↘ 3. Evening bonfire

Add task:

Desc:

Category: [Weekend Plans](#)

2019/12/07 ▾ [Add Item](#)

When to use Global Accelerator

theTradeDesk®



LEVER

Team Internet
Ideas. Change. Markets.

New Relic®



REFINITIV



flowplayer

CrazyCall

webflow

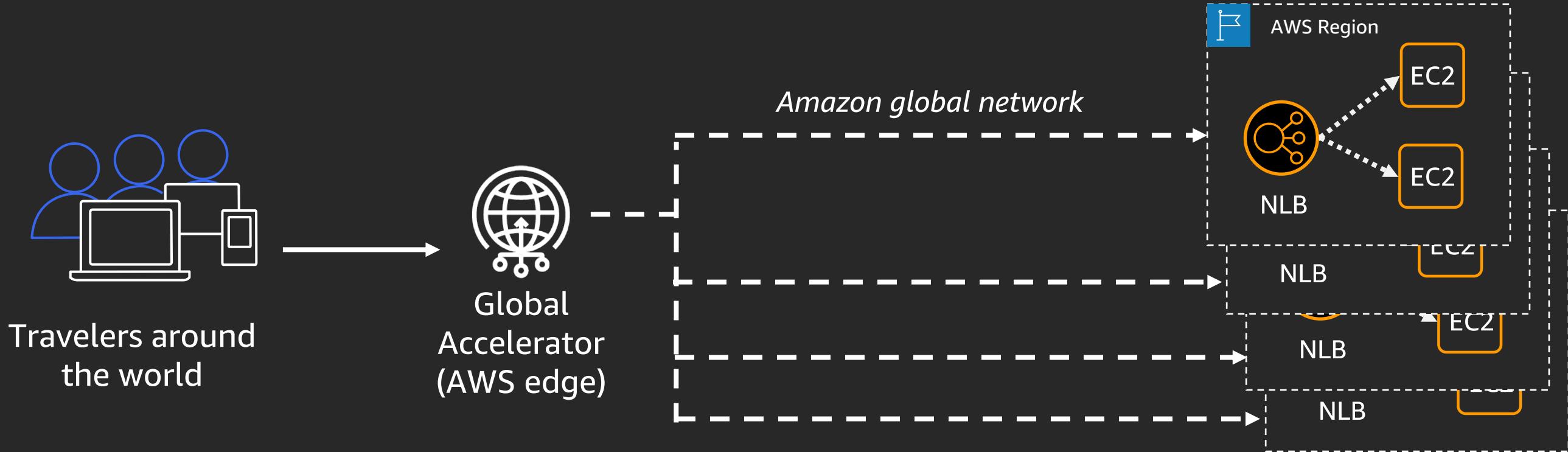
Non-HTTP applications

- Gaming (UDP)
- Real-time video (RTP)
- Voice over IP
- VPN/private connectivity
- DNS hosting (UDP)
- IoT (MQTT)
- File uploads (FTP)
- Push notifications (WebSocket)

HTTP applications

- Blue/green deployment and A+B testing
- Static IPs for IP whitelisting
- Failover resilience for multi-Region applications

Use case: DNS management



Travelers around
the world

Global
Accelerator
(AWS edge)

Amazon global network

AWS Region

EC2

"Skyscanner is a leading global travel search site, a place where people are inspired to plan and book direct from millions of travel options at the best prices. AWS Global Accelerator enables us to effortlessly build resilient, multi-region services, allowing us to focus on providing our travelers with the best experience possible."

— **Stewart Wallace, Senior Engineer, Skyscanner**

Use case: IP whitelisting + API acceleration



“All our points of sale use AWS Global Accelerator to reach our applications running on AWS with high availability and low latency.”

—Leonard Redles, System Architect, ParTech, Inc.



Use case: Live video ingest



“Our video platform is hosted on AWS and uses Global Accelerator to improve the performance and availability of video ingest for our users all around the world.”



—Erik Viklund, VP Dev Ops, Flowplayer

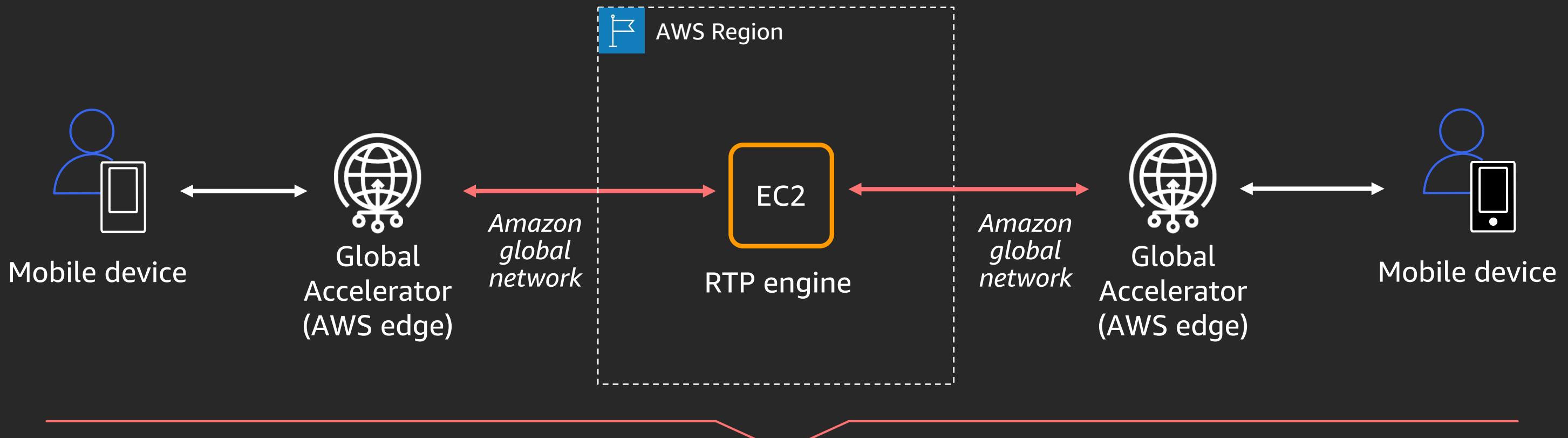
Blog post: <https://aws.amazon.com/blogs/networking-and-content-delivery/how-flowplayer-improved-live-video-ingest-with-aws-global-accelerator/>

AWS
re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Use case: Voice over IP

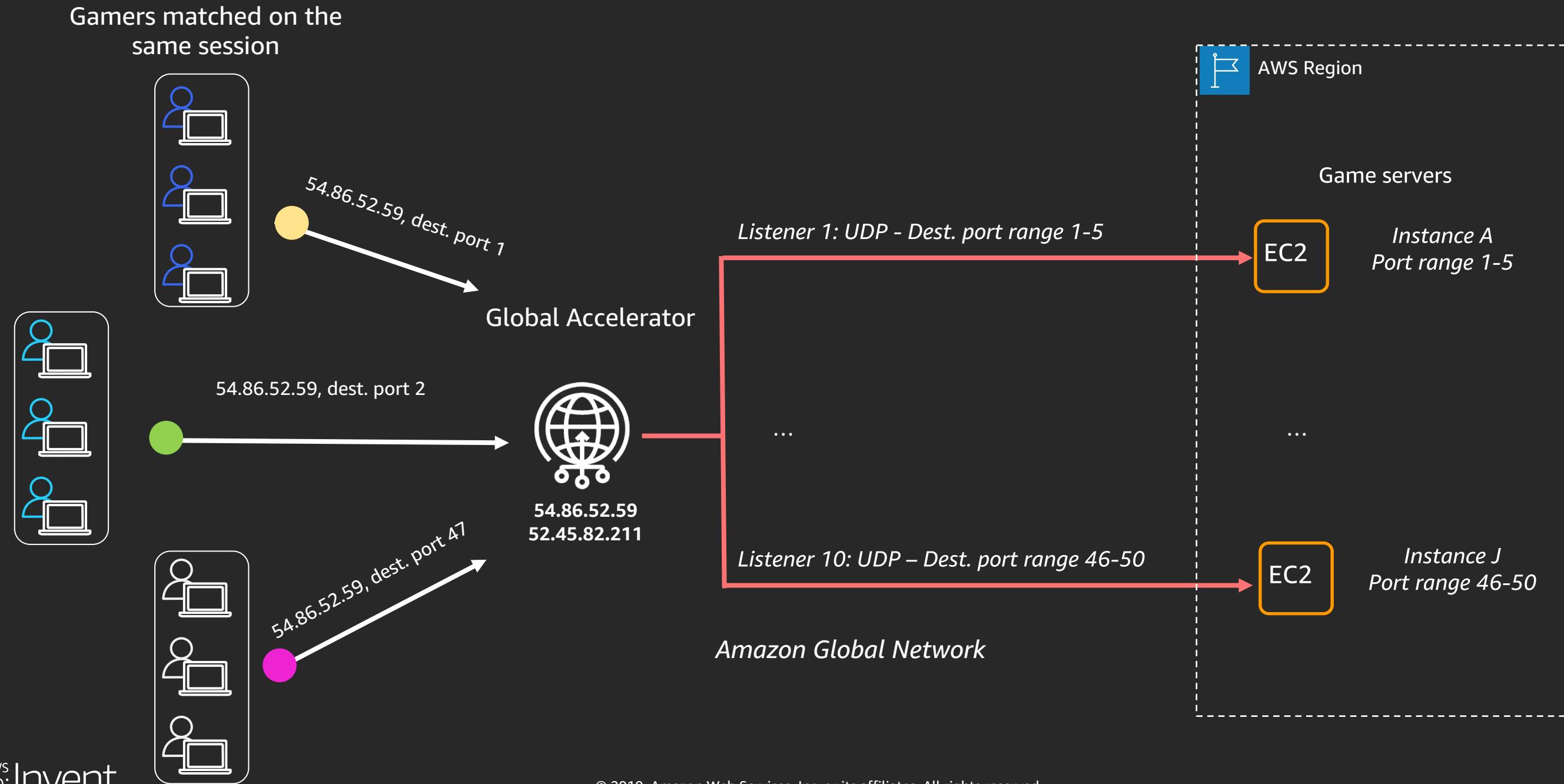


"We use AWS Global Accelerator's static anycast IPs and global network to handle voice over IP (VoIP) traffic, thus ensuring that our customers get the best quality of service."

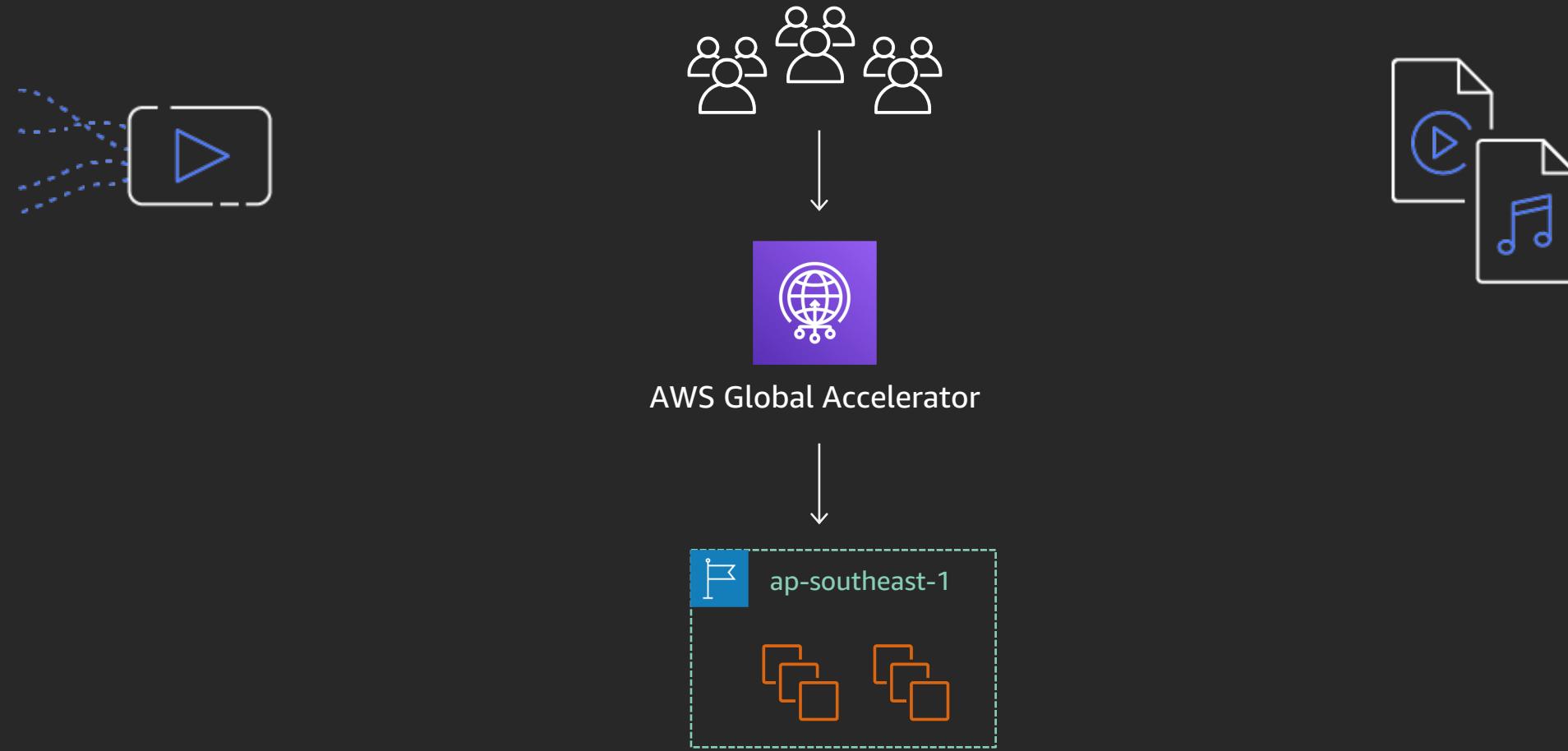
CrazyCall

—*Marcin Kowalczyk, AWS Architect, CrazyCall*

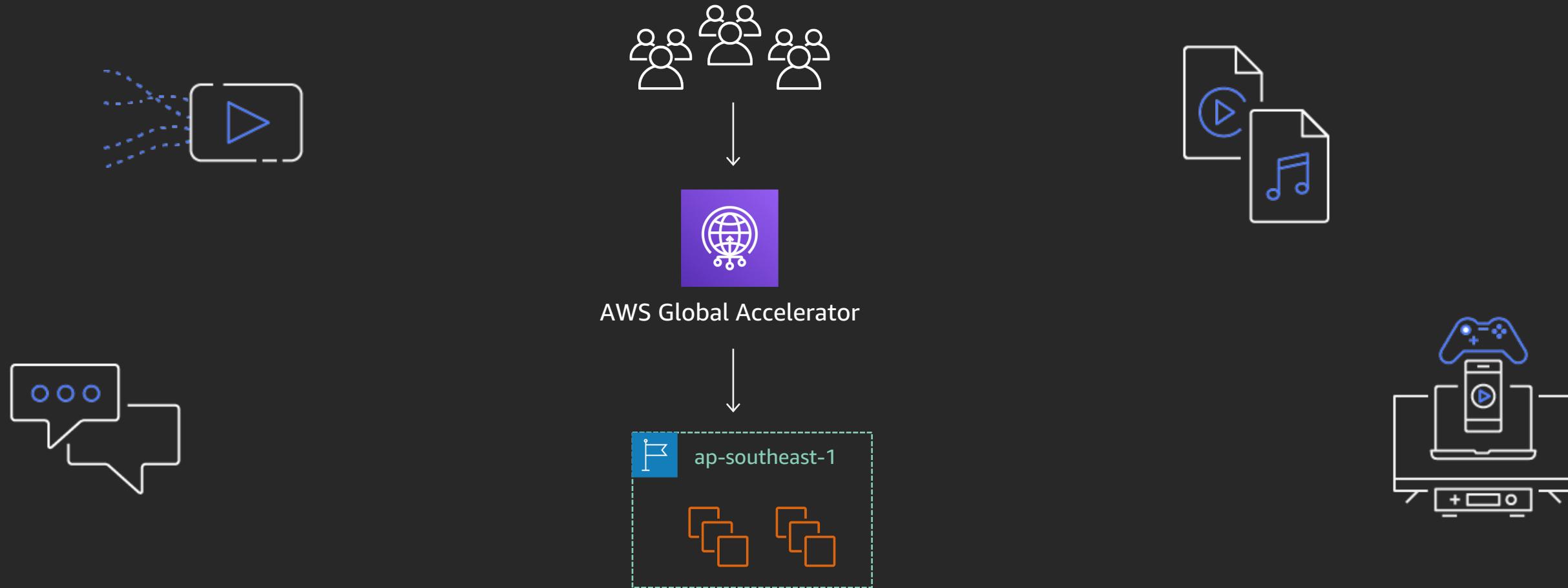
Use case: Multiplayer real-time games with Global Accelerator



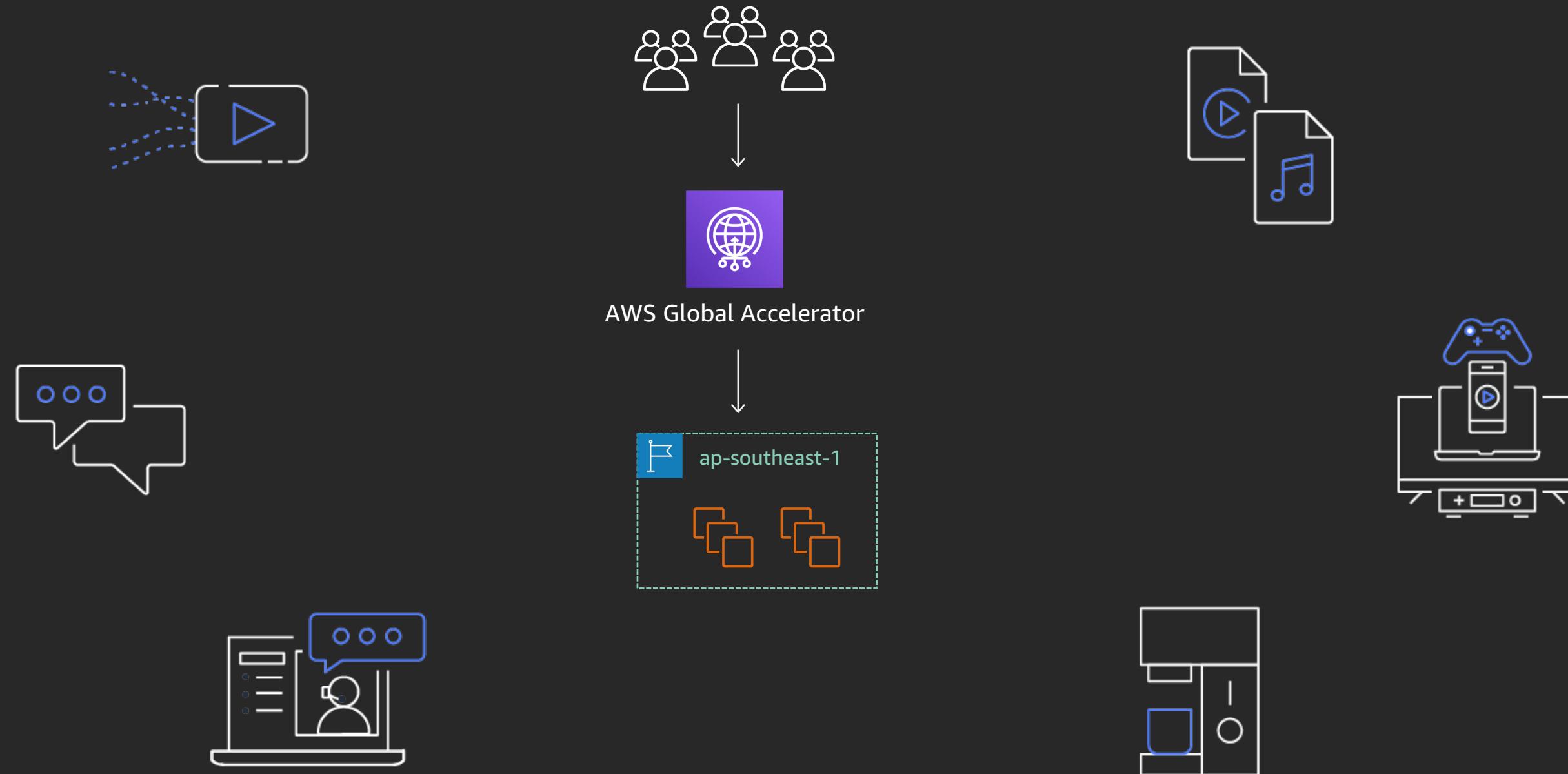
UDP galore



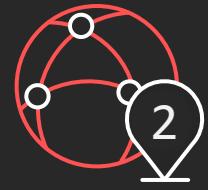
UDP galore



UDP galore



Key features of AWS Global Accelerator



Single entry point
with global static
anycast IPs



Intelligent
distribution of TCP &
UDP traffic



Target EC2 instances
and Elastic Load
Balancers (ALBs & NLBs)
Preserve client IPs



Accelerate both
single and multi-
Region workloads

All traffic traverses the backbone* and is
protected from DDoS attacks

* except within the People's Republic of China

Thank you!

James Wenzel

Sr Partner SA, Networking
Amazon Web Services, Inc.

Harvo Jones

Principal SDE
Amazon Web Services, Inc.



Please complete the session
survey in the mobile app.