Global Accelerator

Chandra Lingam

Cloud Wave LLC

Motivation

Elastic Load Balancers are regional resources

Global Accelerator – Cross-region load balancing

What is the need for cross-region load balancing?
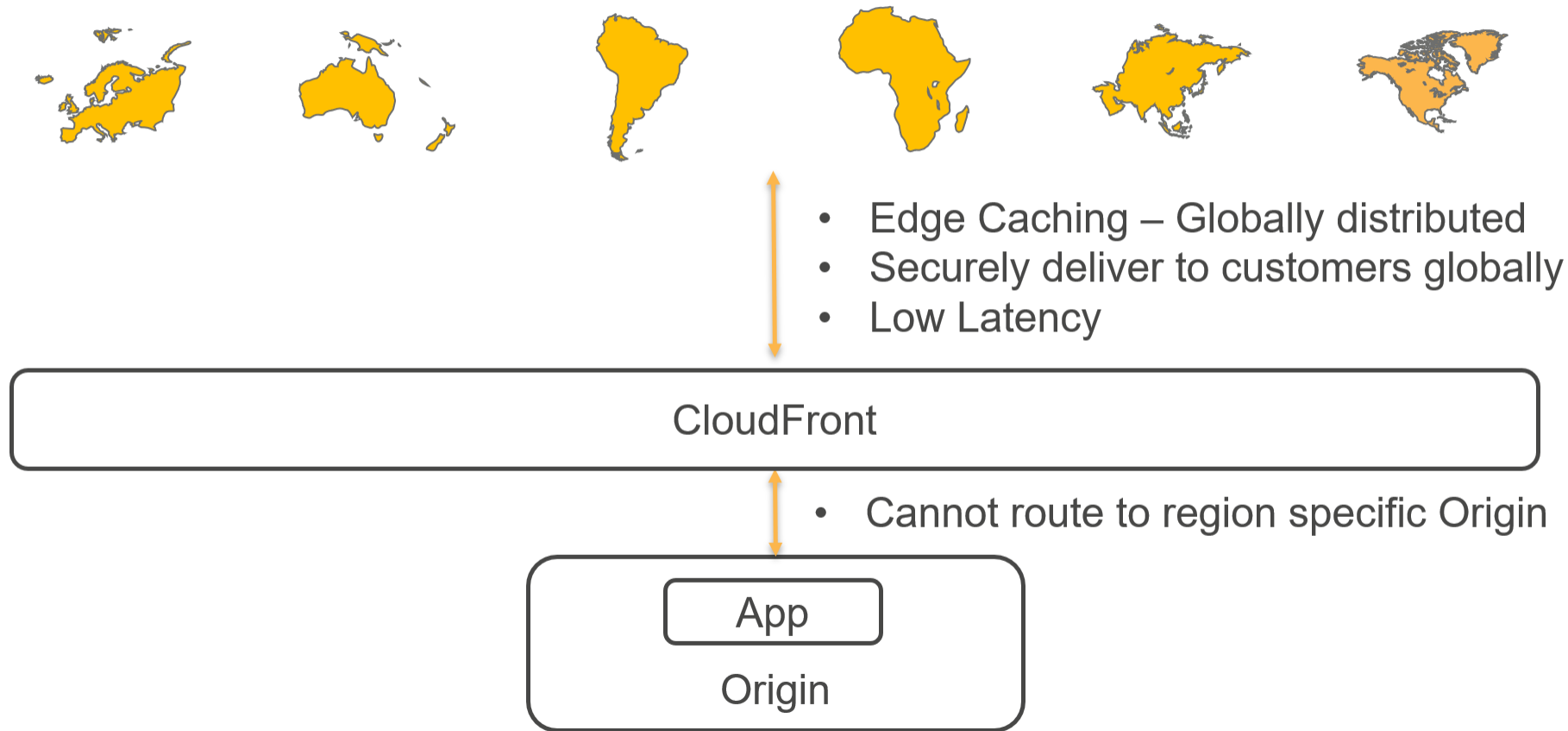
Go Global

- Global Audience
- Performance

App

Content

Data Video Image API

CloudFront CDN

- Edge Caching – Globally distributed
- Securely deliver to customers globally
- Low Latency

CloudFront

- Cannot route to region specific Origin

App

Origin

Keep your app in several regions

- Route traffic to appropriate region
- Complex setup

App Region 1 | App Region 2 | App Region 3 | App Region 4 | App Region 5
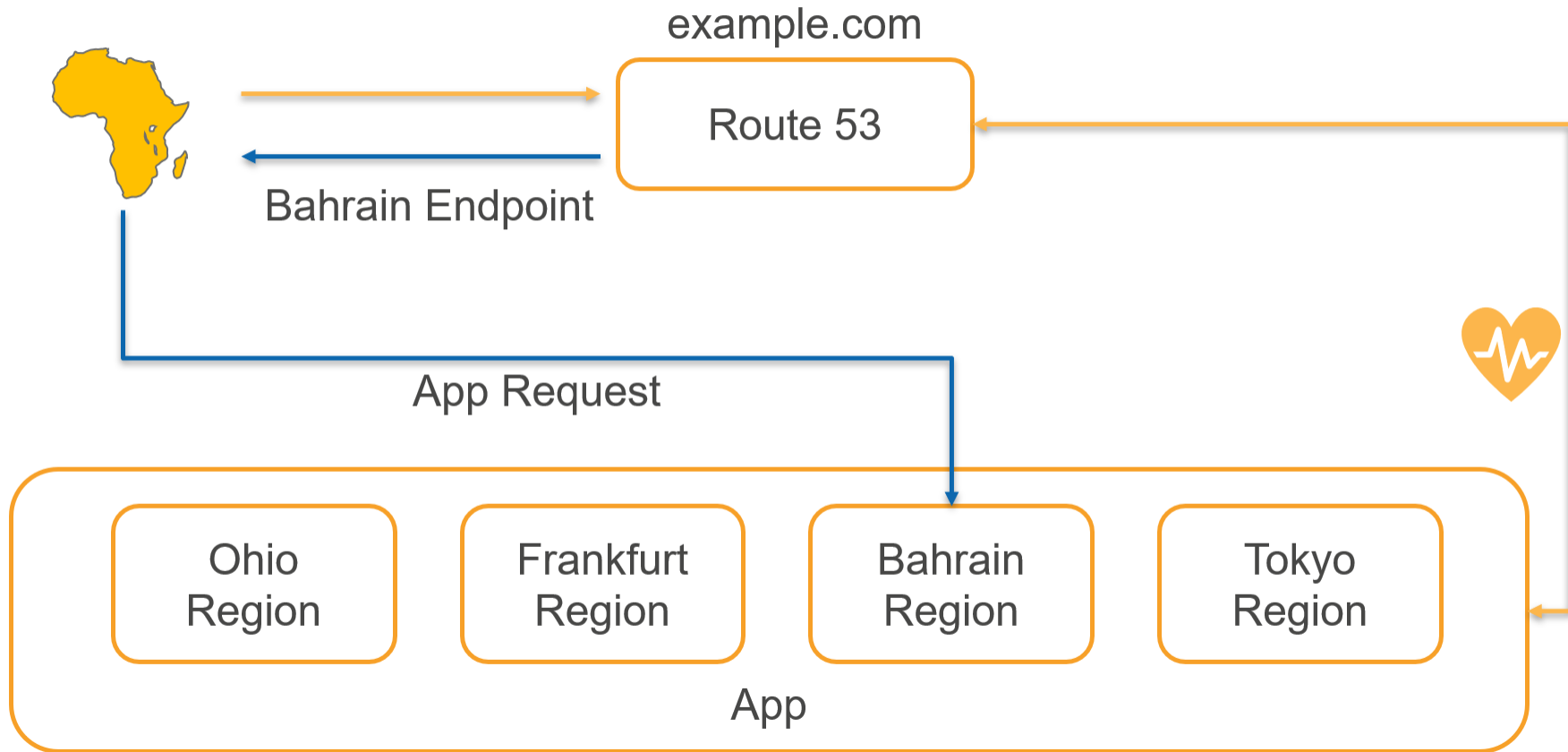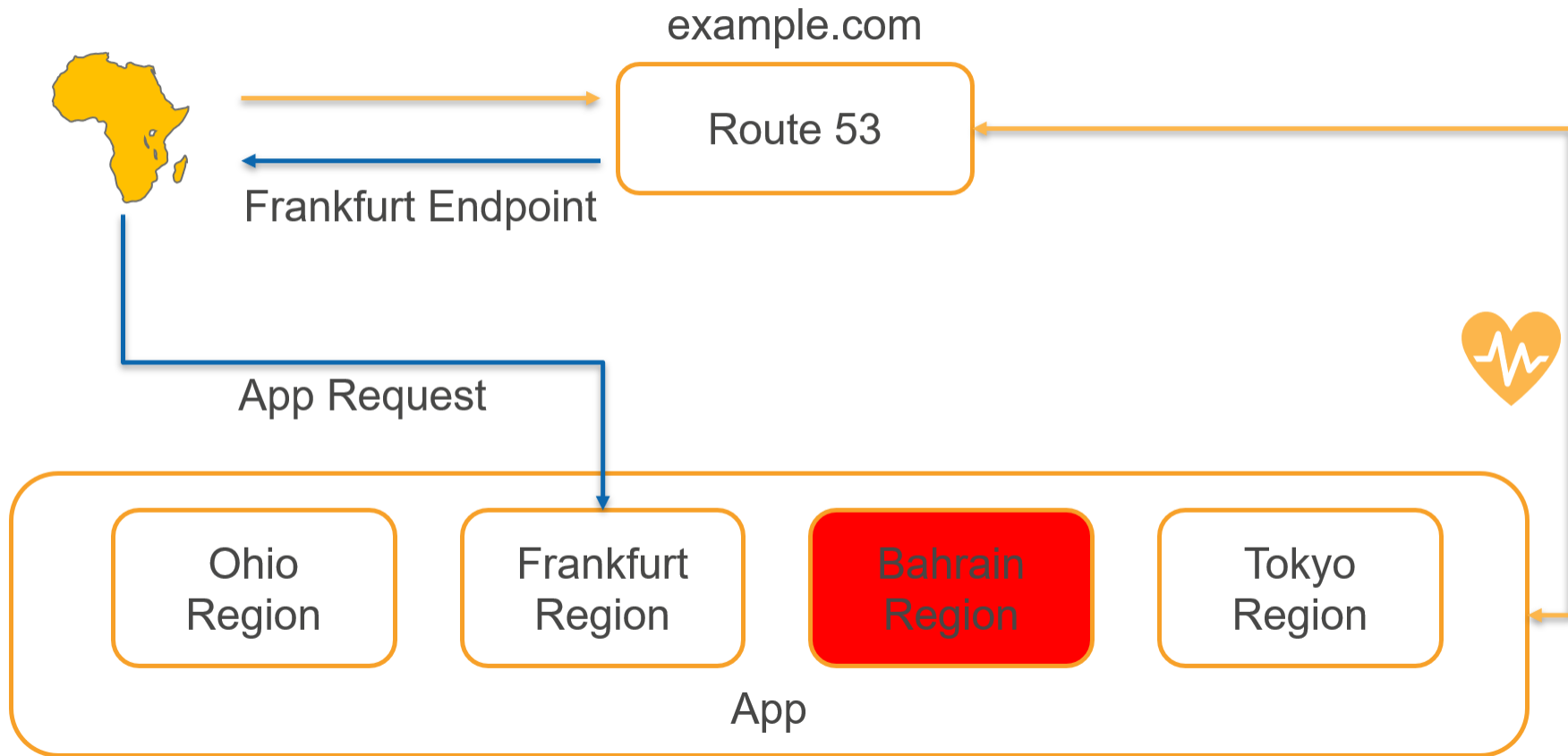
App in major regions

Route 53 DNS Resolution

example.com

Route 53

Bahrain Endpoint

App Request

| Ohio Region | Frankfurt Region | Bahrain Region | Tokyo Region |

App

Route 53 DNS Resolution

example.com

Route 53

Frankfurt Endpoint

App Request

| Ohio Region | Frankfurt Region | Bahrain Region | Tokyo Region |

App

Route 53 based routing issues

DNS Records are cached in client and DNS resolvers for duration specified in TTL

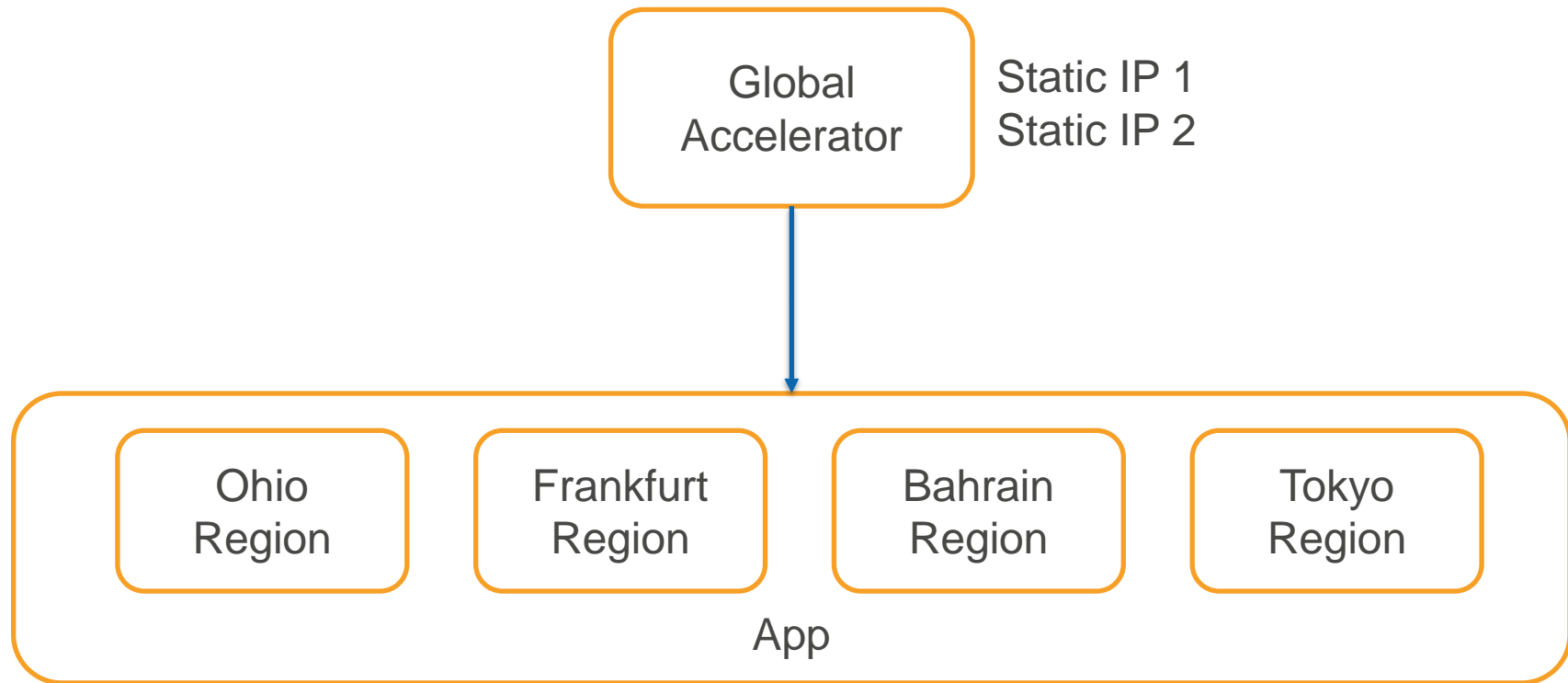Delay in responding and rerouting traffic to new endpoint

Infrastructure scaling actions results in IP address changes
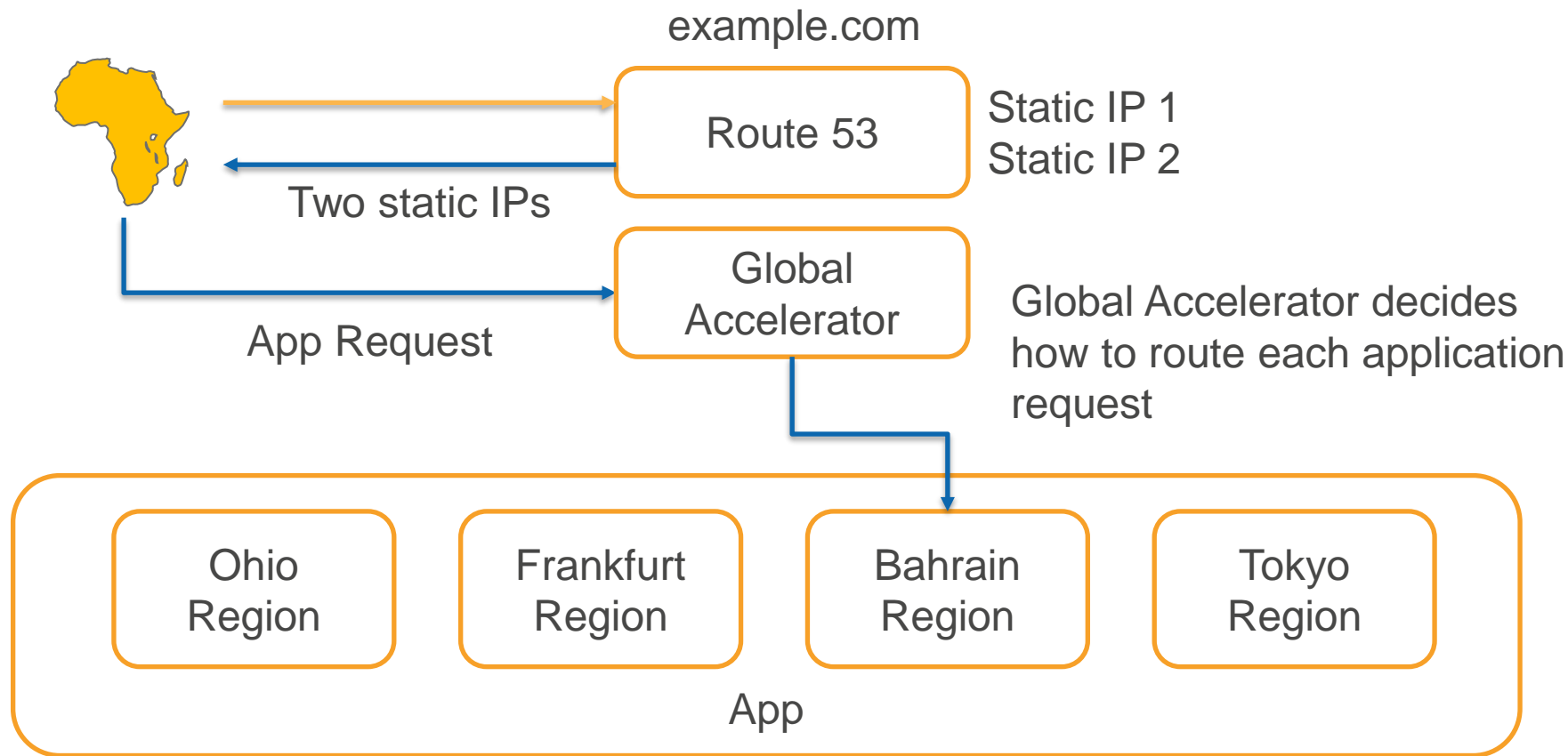
ALB adds and removes load balancer nodes

List of load balancer nodes is maintained by ALB in Route 53

Polling based

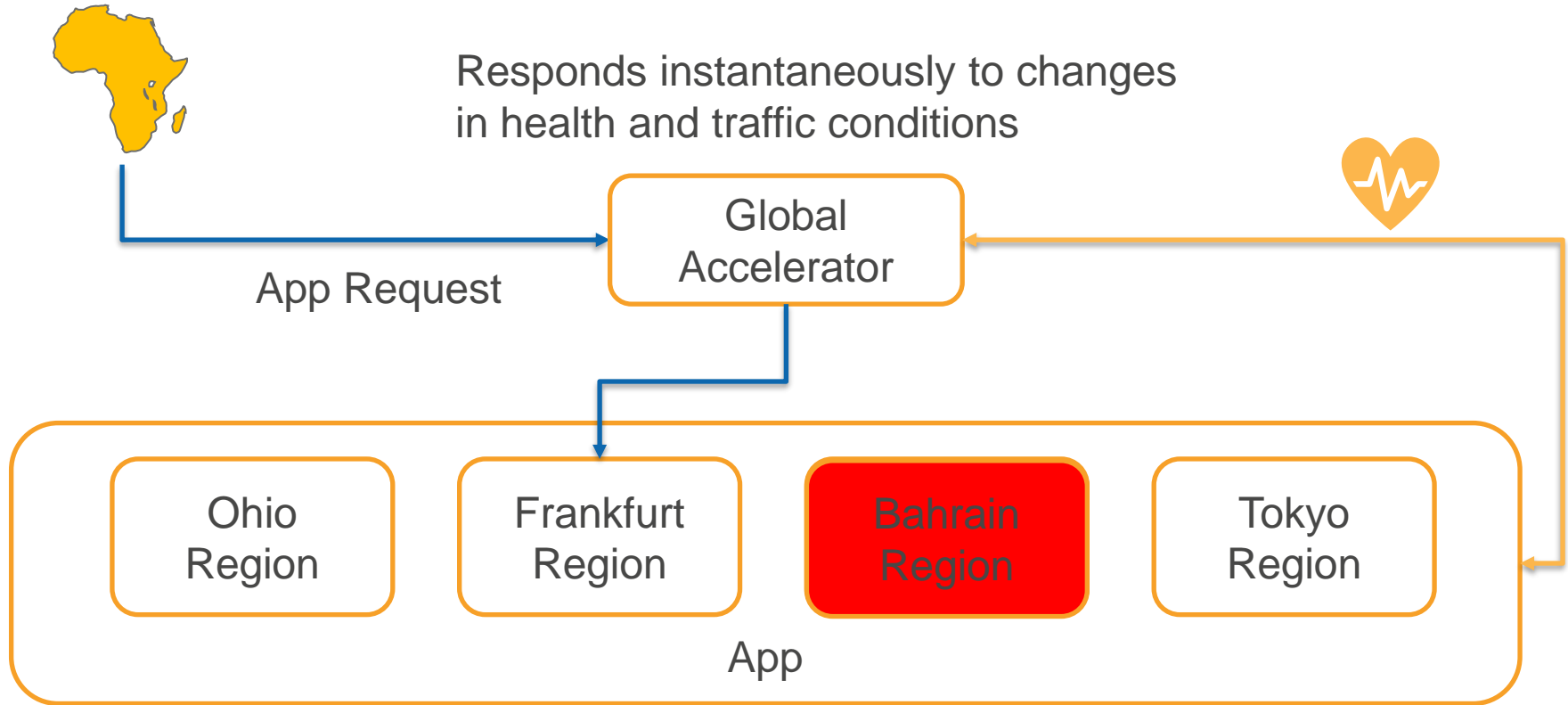# Global Accelerator

Global
Accelerator

Static IP 1
Static IP 2

Ohio
Region

Frankfurt
Region

Bahrain
Region

Tokyo
Region

App

Global Accelerator

example.com

Route 53

Static IP 1
Static IP 2

Two static IPs

Global Accelerator

App Request

Global Accelerator decides how to route each application request

Ohio Region

Frankfurt Region

Bahrain Region

Tokyo Region

App

Global Accelerator

Responds instantaneously to changes in health and traffic conditions

Global Accelerator

App Request

Ohio Region

Frankfurt Region

Bahrain Region

Tokyo Region

App

Global Accelerator Speed Test Tool

AWS Global Accelerator
Speed Comparison

About this tool

AWS Global Accelerator is a service that improves the availability and performance of your applications. This tool compares Global Accelerator to the public internet. Choose a file size to see the time to download a file from application endpoints in different AWS Regions to your browser.

Files are downloaded over HTTPS/TCP from Application Load Balancers (ALBs) in different AWS Regions to your browser. Learn more

Choose a file size and click "Start" to start the tests: [100KB ▼] **Reset**

We welcome suggestions for how to improve this tool. Provide feedback

ⓘ Results may differ when you run the test multiple times. Download times can vary based on factors that are external to Global Accelerator, such as the quality, capacity, and distance of the connection in the last-mile network that you're using.

N. Virginia (us-east-1)

Direct over internet | Total time
146ms

AWS Global Accelerator | 101ms

31% faster with AWS Global Accelerator

Oregon (us-west-2)

Direct over internet | Total time
645ms

Ireland (eu-west-1)

Direct over internet | Total time
686ms

Global Accelerator

Clients connect to Edge location

Request travels in AWS Global network (and not public internet)

Endpoints can be private (Global Accelerator creates a peering connection)

Client-affinity – when enabled, request from a client is sent to the same endpoint

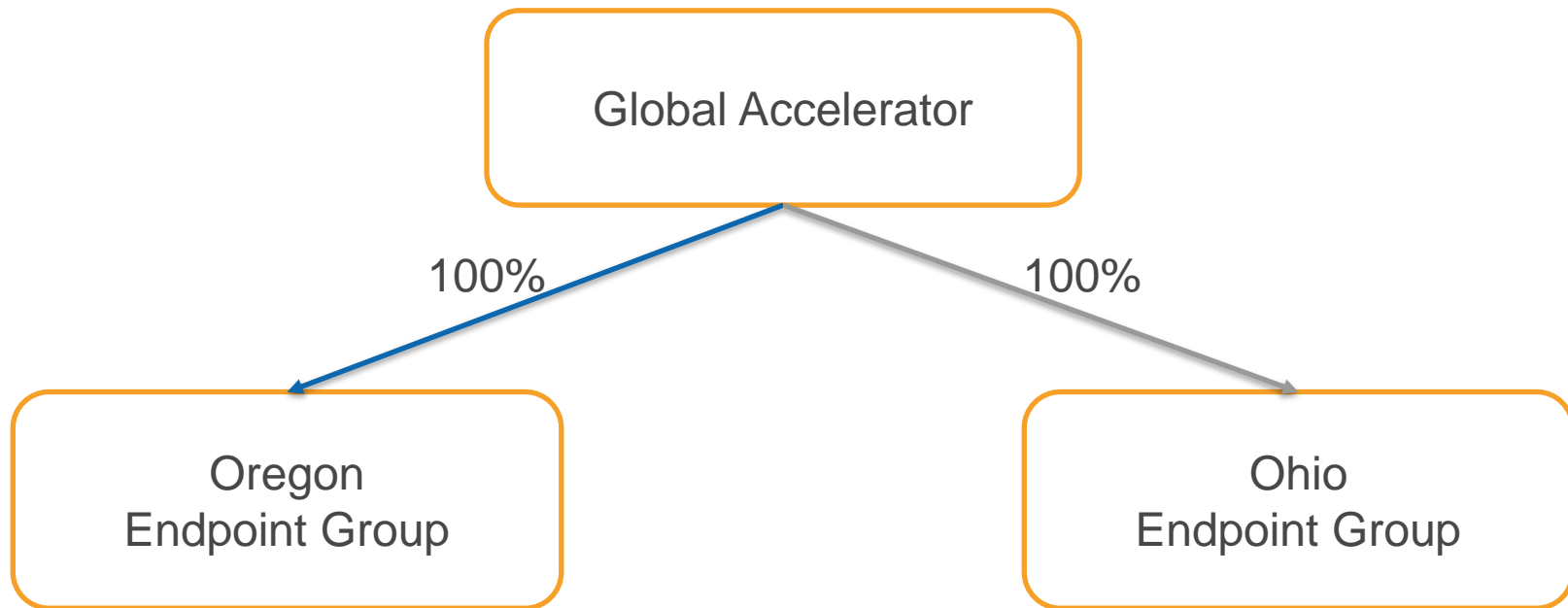# Global Accelerator

Hides infrastructure

DDoS Protection

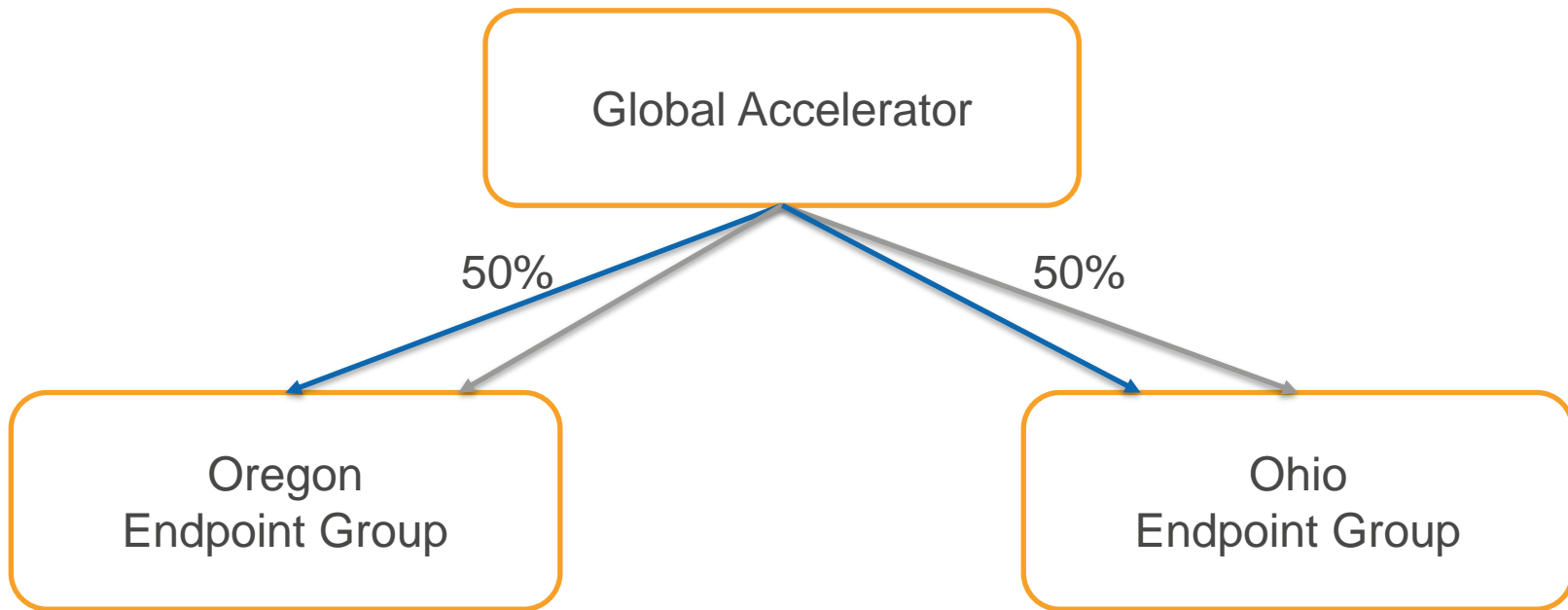AWS Shield Standard and Advanced Protection

Bring your own IP

Traffic Dial

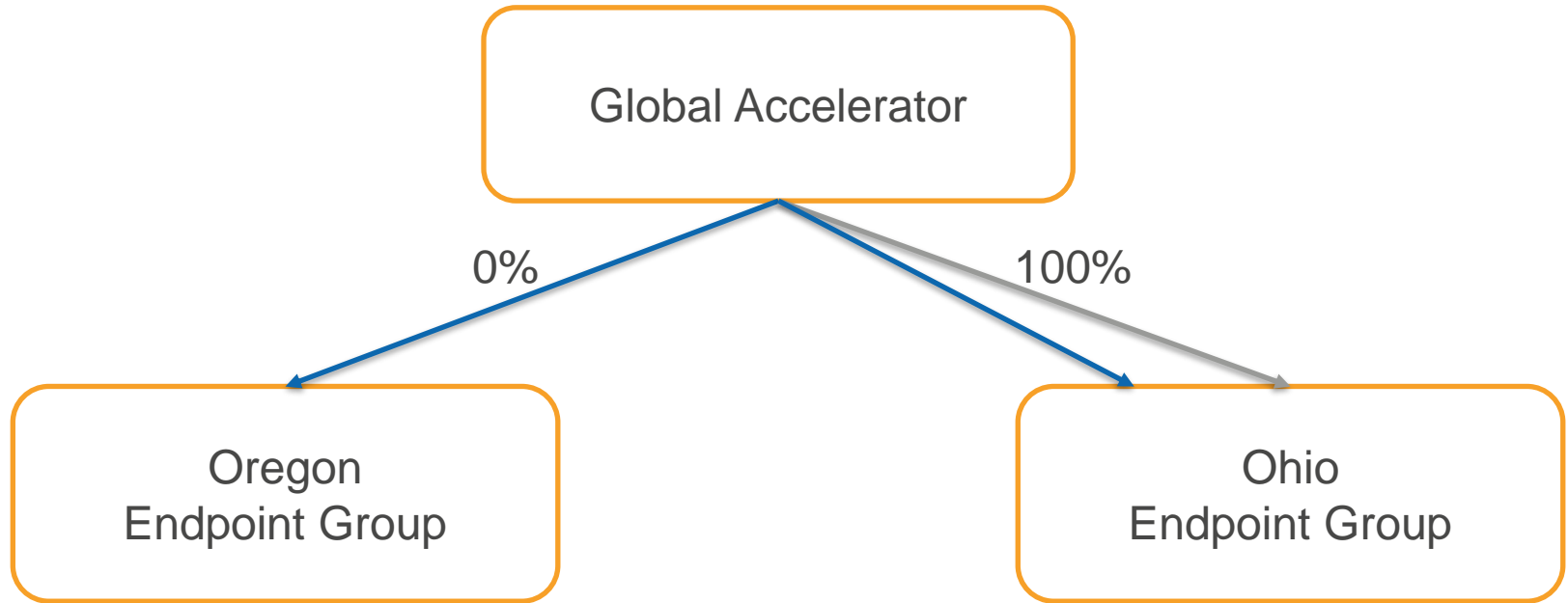Global Accelerator selects a healthy endpoint group closest to the requester

Traffic Dial – Shape traffic flow between regions

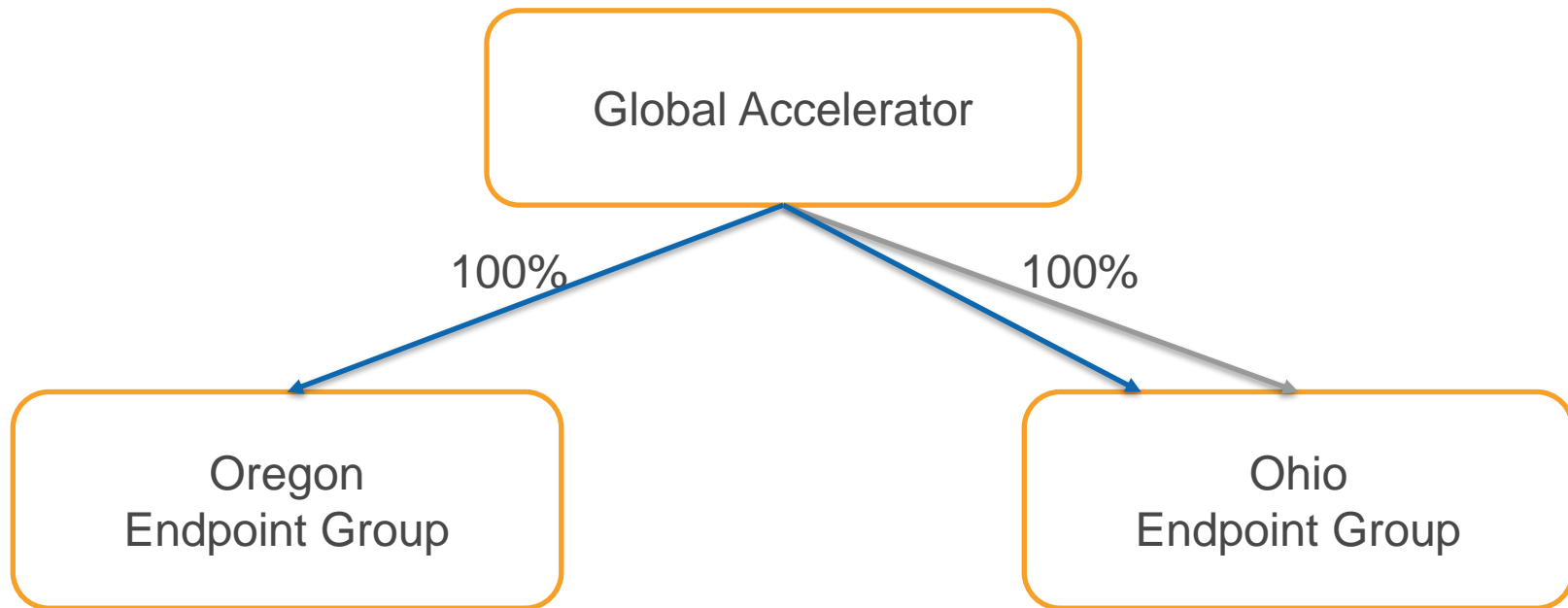50% of Oregon traffic is routed to Ohio and 50% of Ohio traffic is routed to Oregon
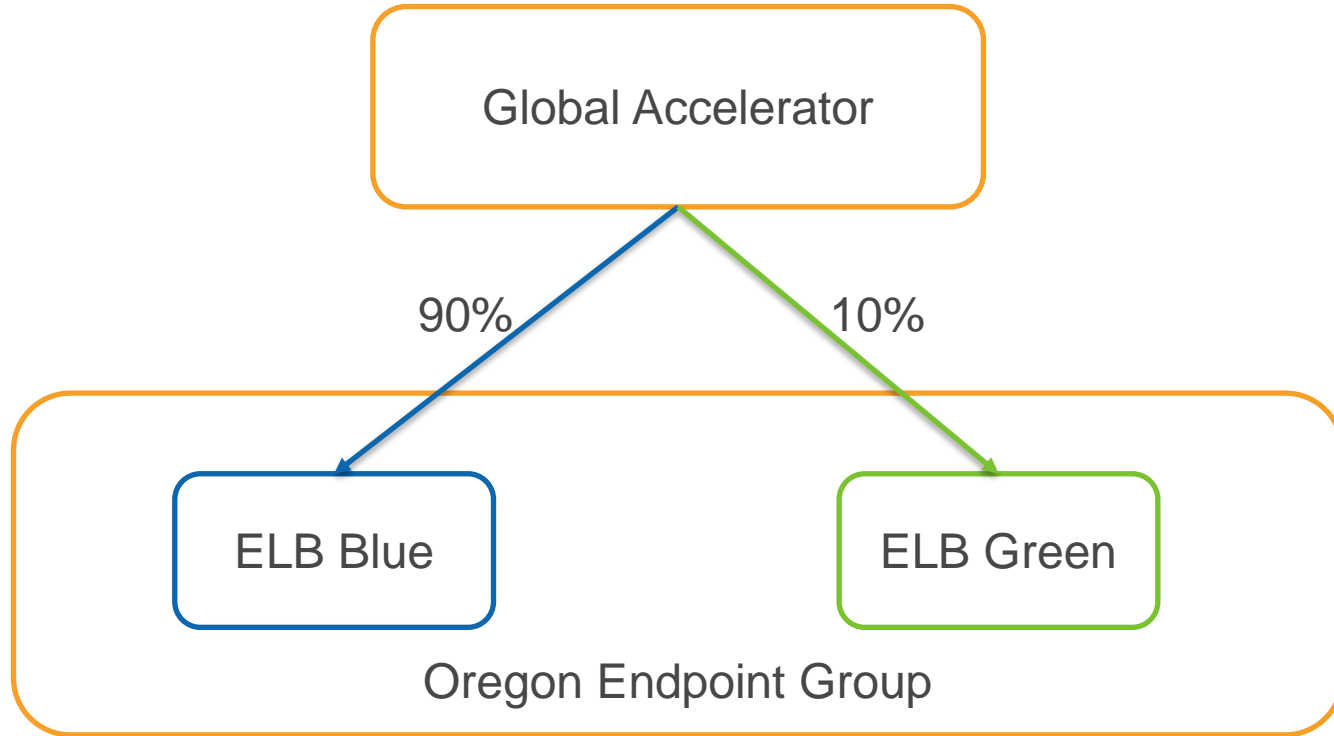
Traffic Dial – During Deployment

Global Accelerator

0%

100%

Oregon
Endpoint Group

Ohio
Endpoint Group

All traffic routed to Ohio

Traffic Dial – Restore traffic after deployment

```
                    ┌─────────────────────────┐
                    │    Global Accelerator   │
                    └─────────────────────────┘
              100%                          100%
      ┌──────────────────┐         ┌──────────────────┐
      │      Oregon      │         │       Ohio       │
      │  Endpoint Group  │         │  Endpoint Group  │
      └──────────────────┘         └──────────────────┘
```

Weight – Shape traffic flow inside a region

Summary

Reacts instantaneously to changes in traffic and health

Request routed through Edge network and AWS global network

Add regions, move regions without impact to client
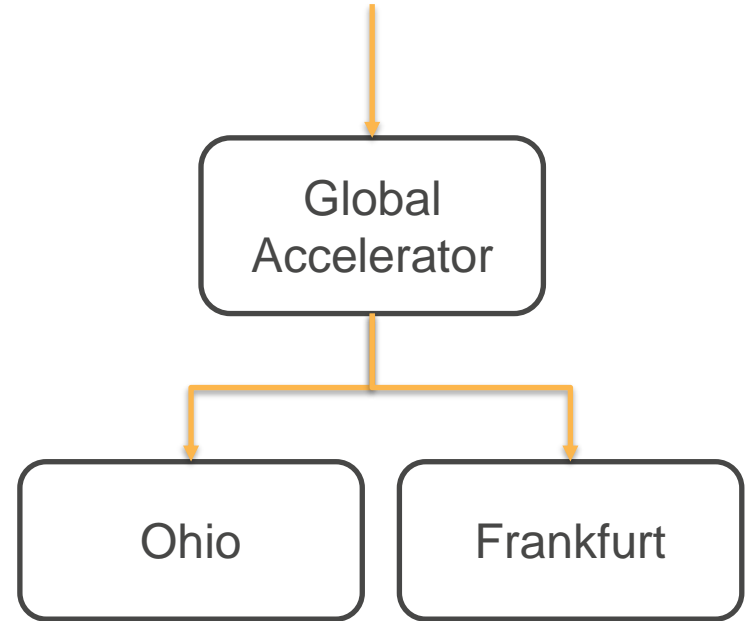
DDoS protection

Lab – Global Accelerator

Setup ELB in multiple regions
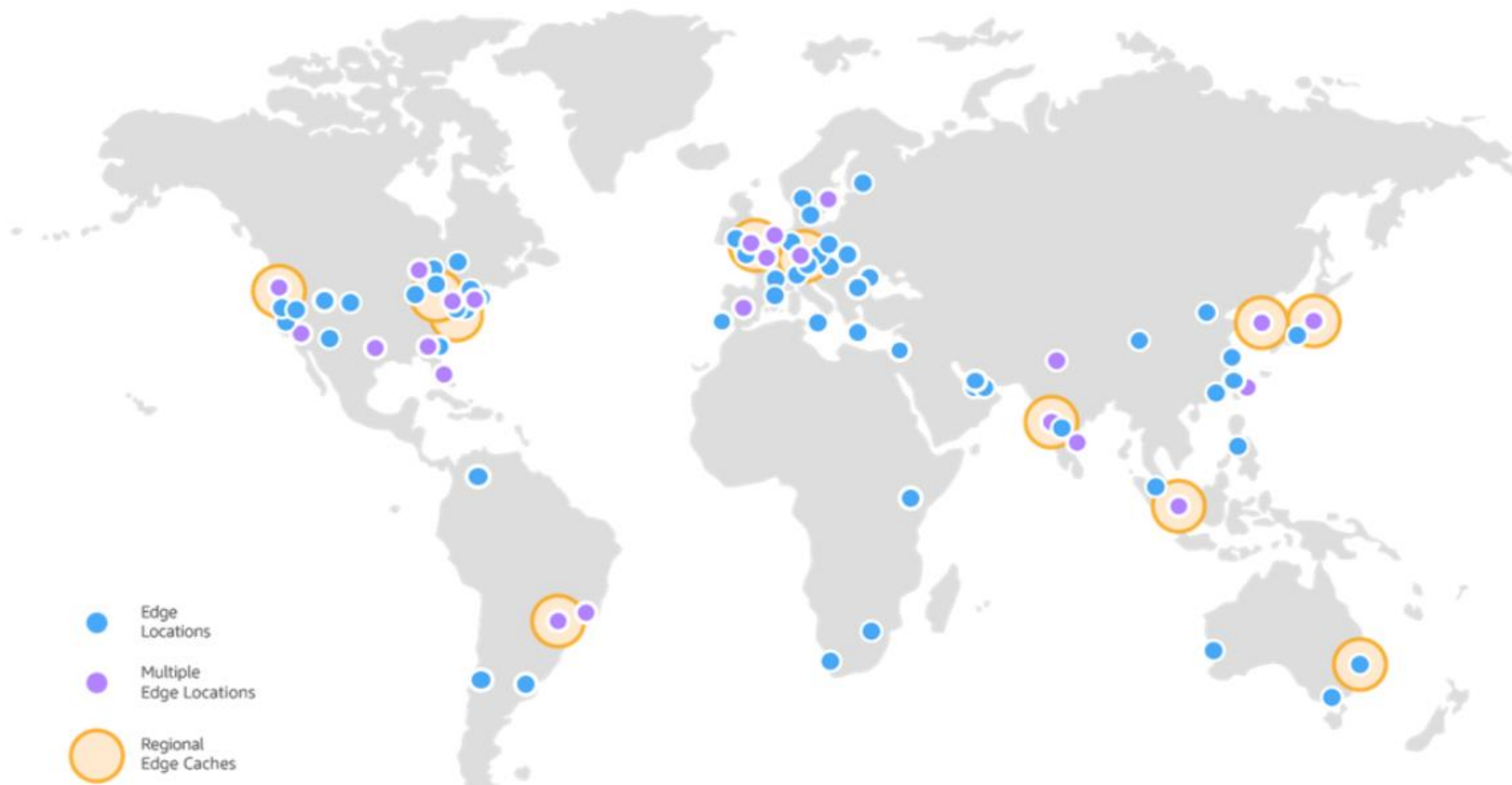
Setup Global Accelerator

Adjust traffic dial to reroute traffic

Simulate Failure

```
                    │
                    ▼
            ┌───────────────┐
            │    Global     │
            │  Accelerator  │
            └───────────────┘
                    │
          ┌─────────┴─────────┐
          ▼                   ▼
    ┌──────────┐        ┌───────────┐
    │   Ohio   │        │ Frankfurt │
    └──────────┘        └───────────┘
```

Edge Locations: https://aws.amazon.com/cloudfront/features/

Locations and 11 Regional Edge Caches) in 84 cities across 42 countries. Amazon CloudFront Edge locations are located in:

Chandra Lingam

57,000+ Students

For AWS self-paced video courses, visit:

https://www.cloudwavetraining.com/

Cloud Wave LLC