Behind the Scenes of the AWS Edge

Giorgio Bonfiglio
(he/him)
Principal TAM
AWS Enterprise Support

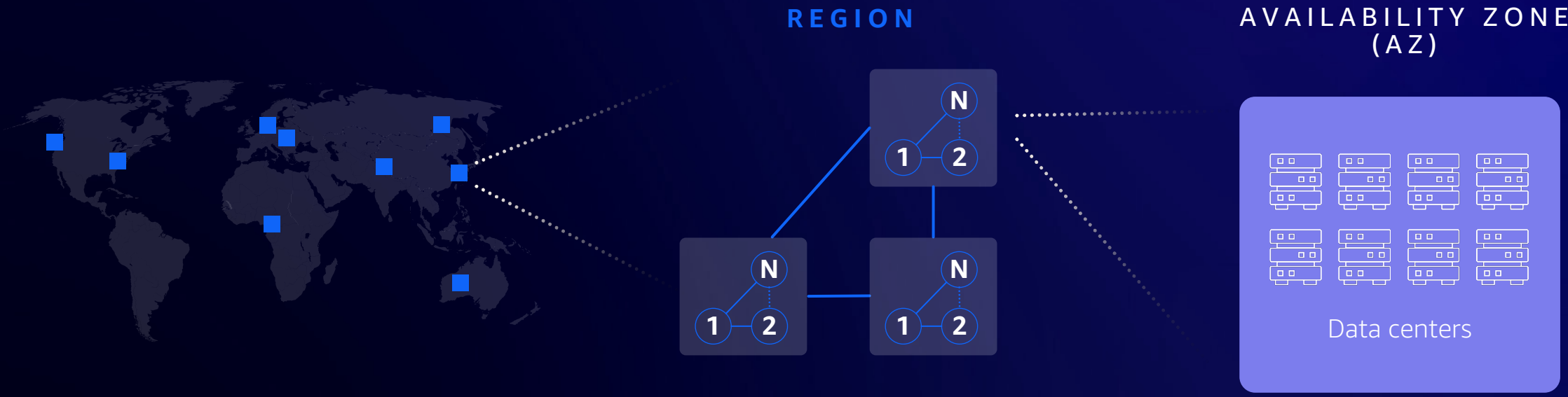
Agenda

- AWS Infrastructure
- Amazon CloudFront
- AWS Global Accelerator
- Q&A

AWS Infrastructure

AWS Regions

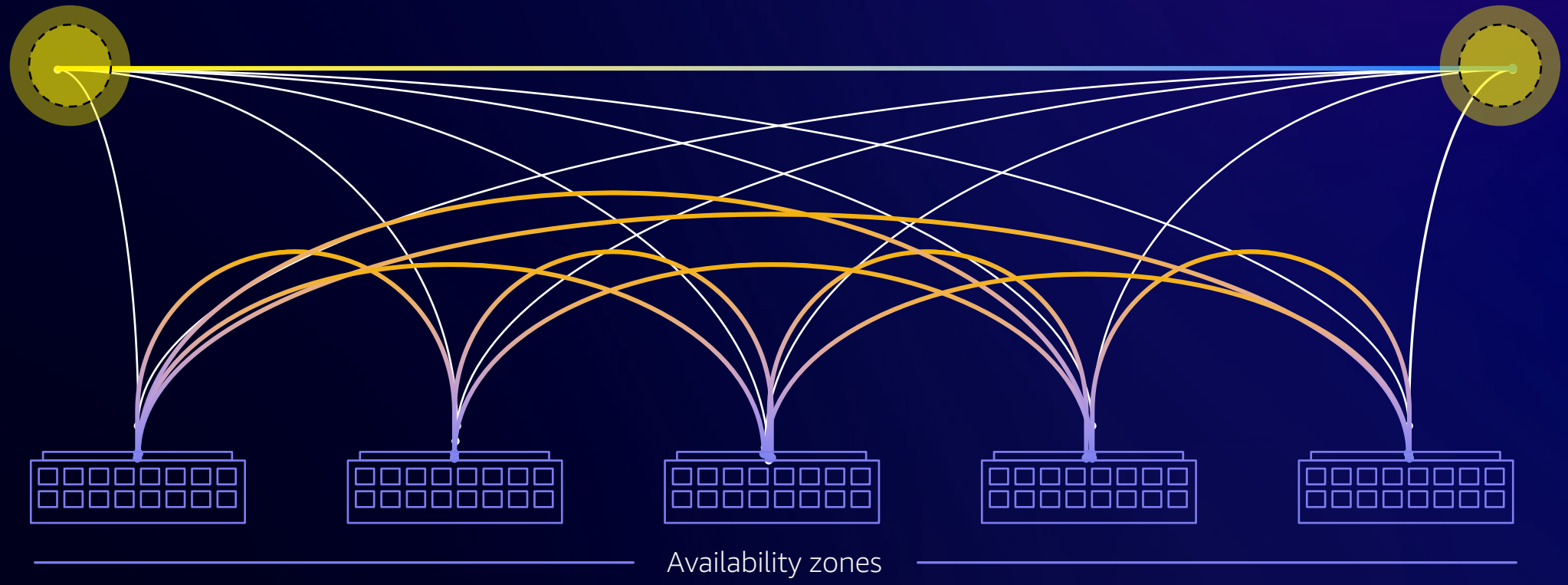
AWS Regions are comprised of multiple AZs for **high availability, high scalability,** and **high fault tolerance.** Applications and data are replicated in real time and consistent in the different AZs.



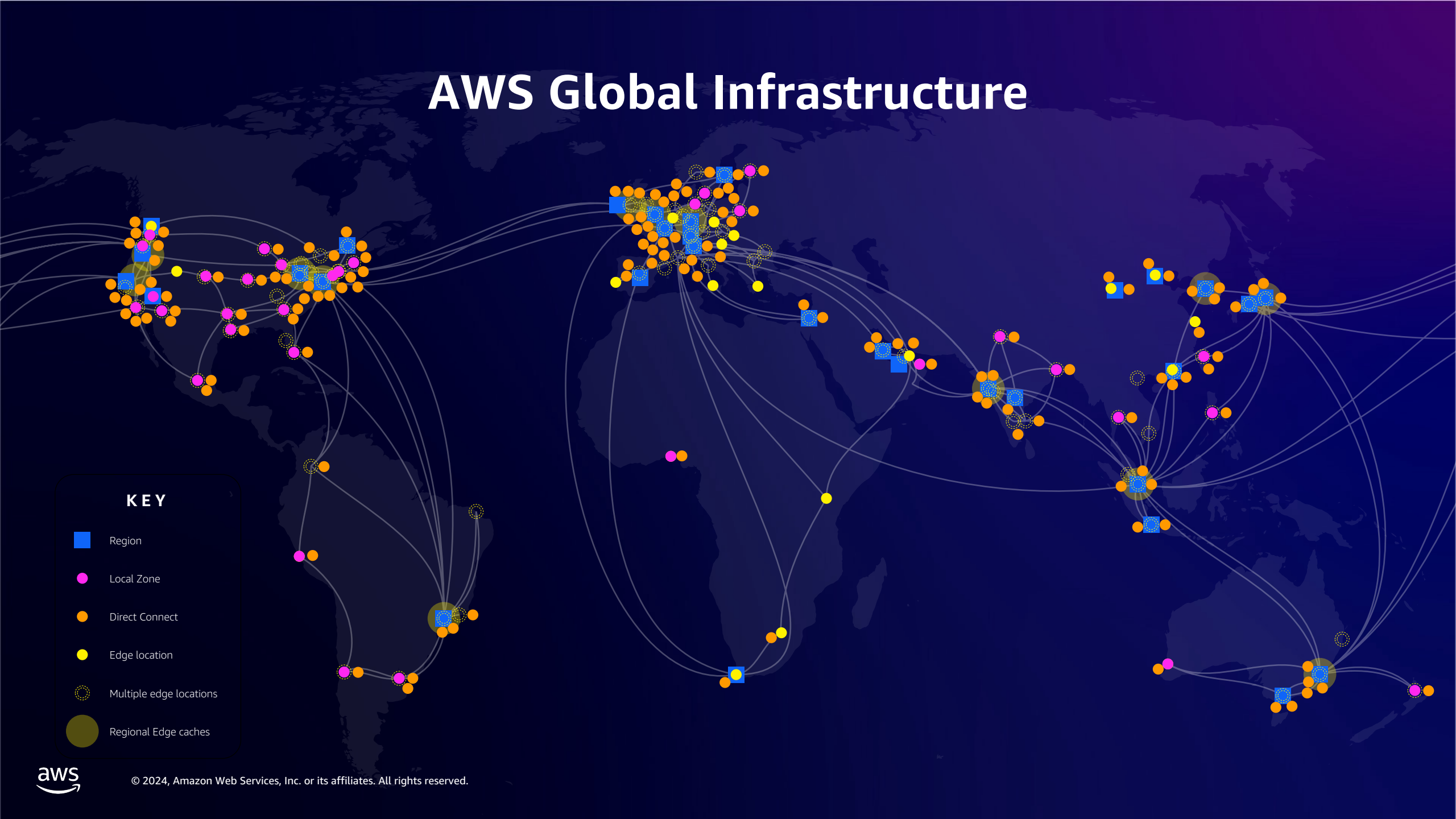
Regional network

At least 2 redundant transit centers

Highly peered & connected



AWS Global Infrastructure



Amazon CloudFront

Amazon Global Edge Network

GLOBAL NETWORK

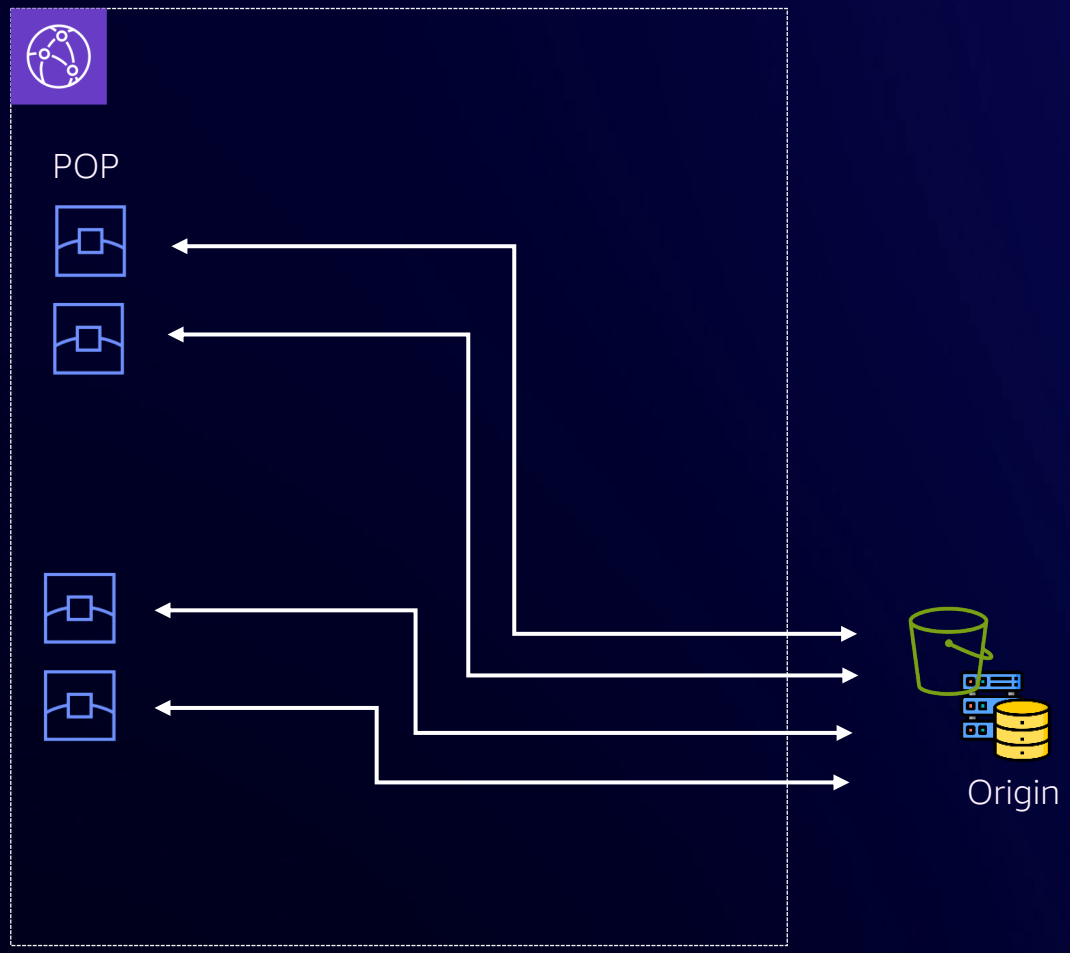
Redundant 400 GbE network and private capacity between all regions except for the AWS China*

EDGE NETWORKING

600+ PoPs in 50 countries and 100+ cities, with direct peering to all major ISPs

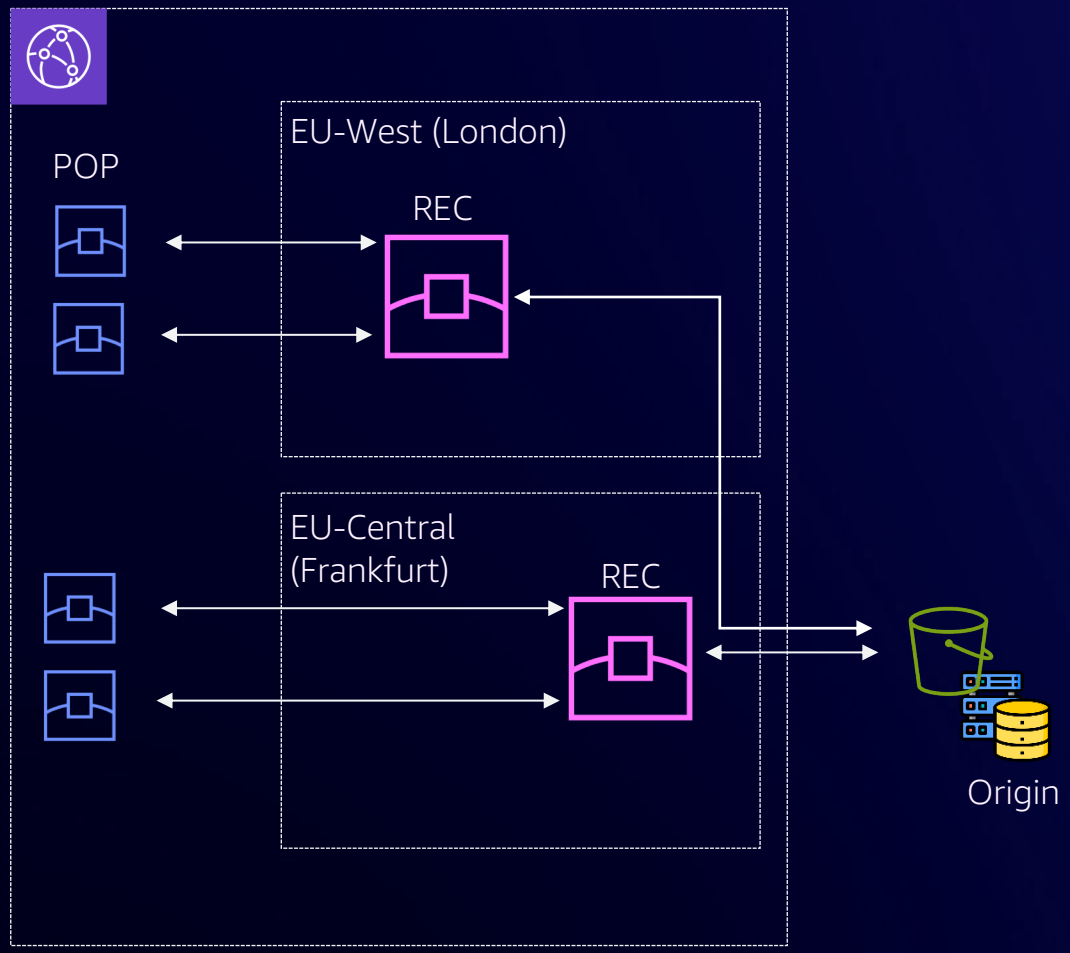
aws

Caching Layers v1



Point of Presence (POP)

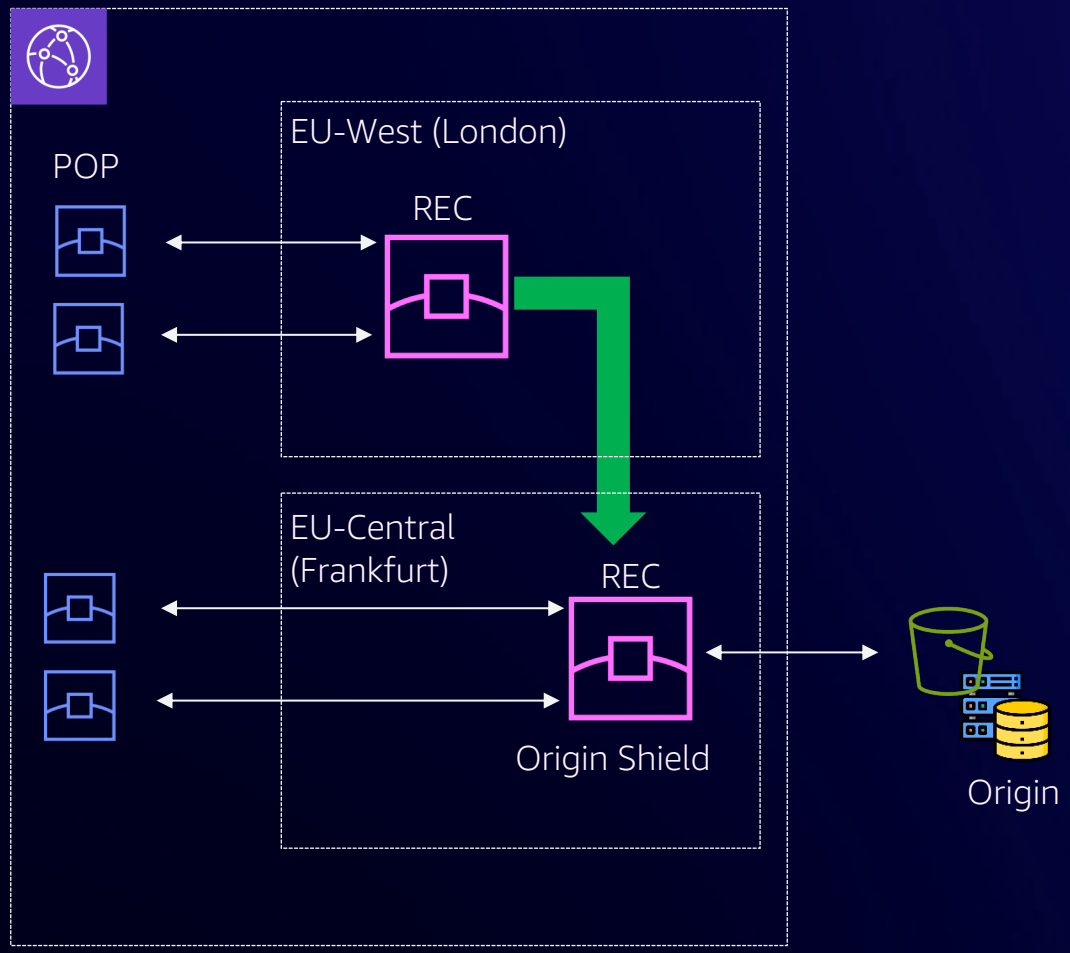
Caching Layers v2



Point of Presence (POP)

Regional Edge Cache (REC)

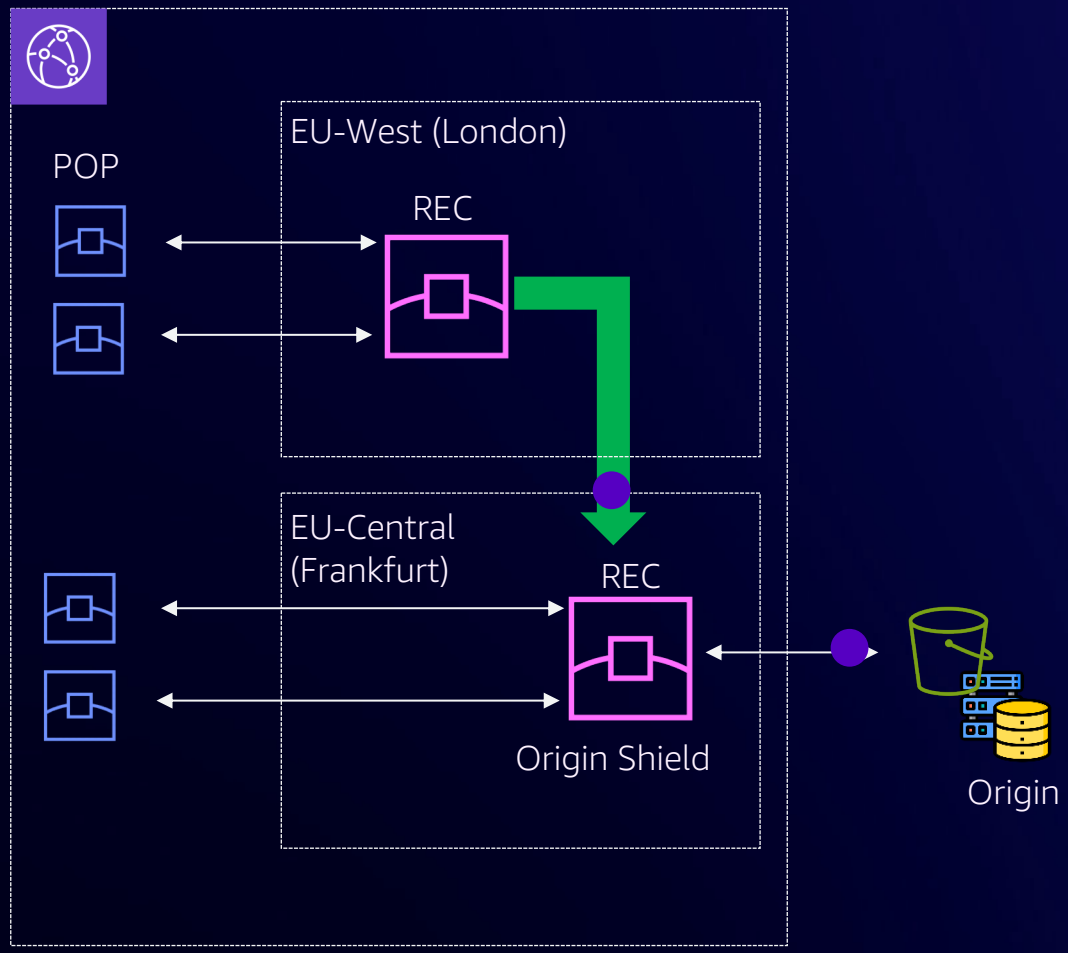
Caching Layers v3



Point of Presence (POP)

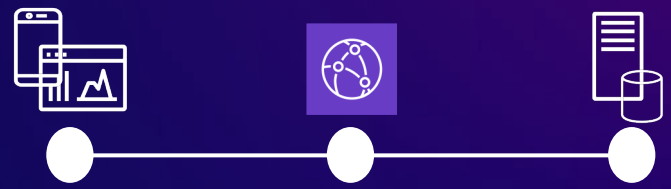
Regional Edge Cache (REC)

Caching Strategy



Point of Presence (POP)

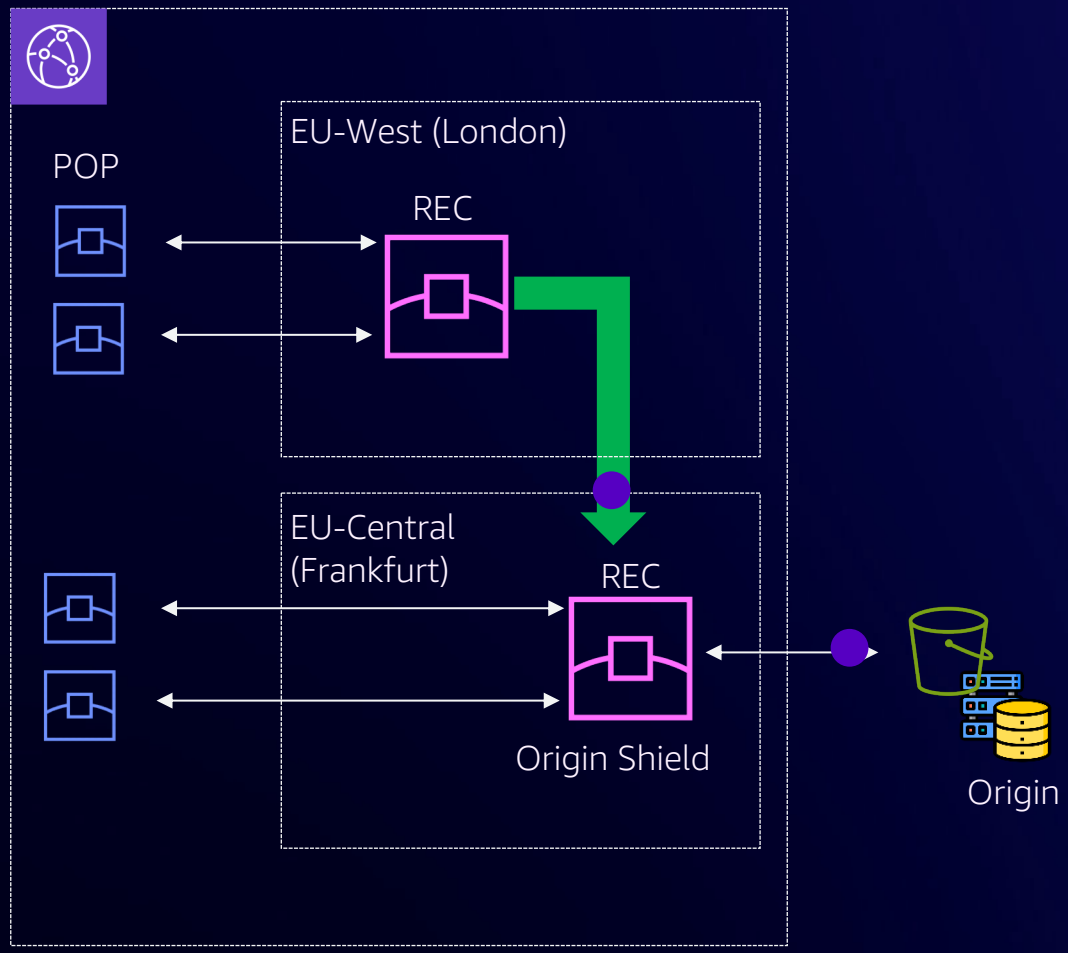
Regional Edge Cache (REC)



JavaScript and CSS files -
Cache-Control: public,max-age=31536000, immutable

index.html - Cache-Control: public,max-age=60

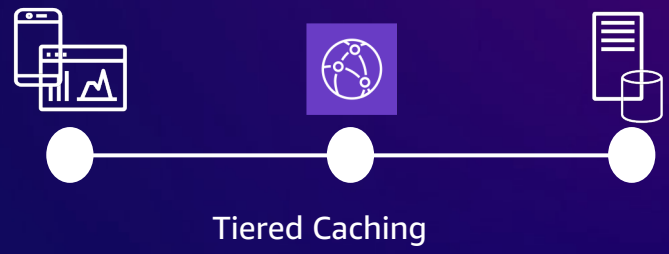
Caching Strategy



Point of Presence (POP)

Regional Edge Cache (REC)

Popularity



Mutability

JavaScript and CSS files -
Cache-Control: public,max-age=31536000, immutable

index.html - Cache-Control: public,max-age=60

Shareable

Cache-Control: Private, max-age=3600

Performance & Efficiency

ETag: "1234"
Cache-Control: stale-while-revalidate

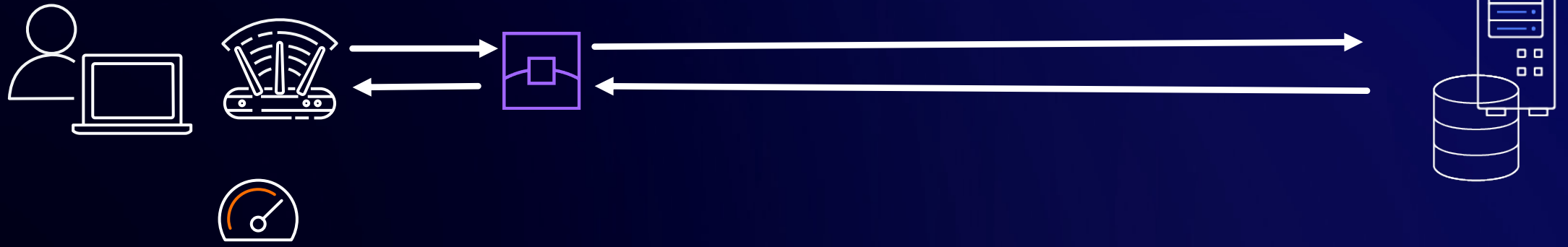
Dynamic Content Acceleration

20KB Object
393 ms

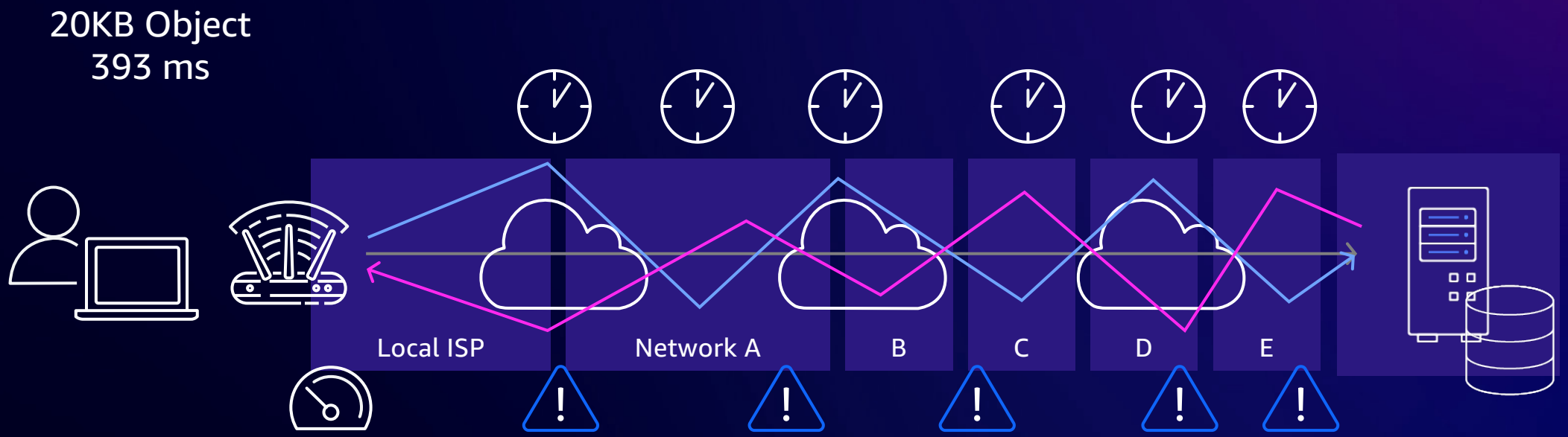
20KB Object
121 ms

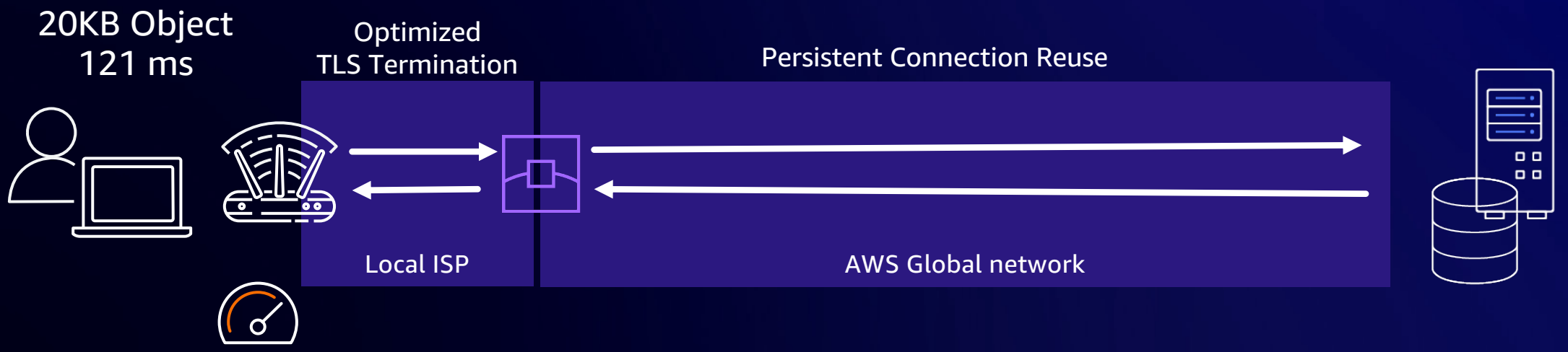
Optimized
TLS Termination

Persistent Connection Reuse

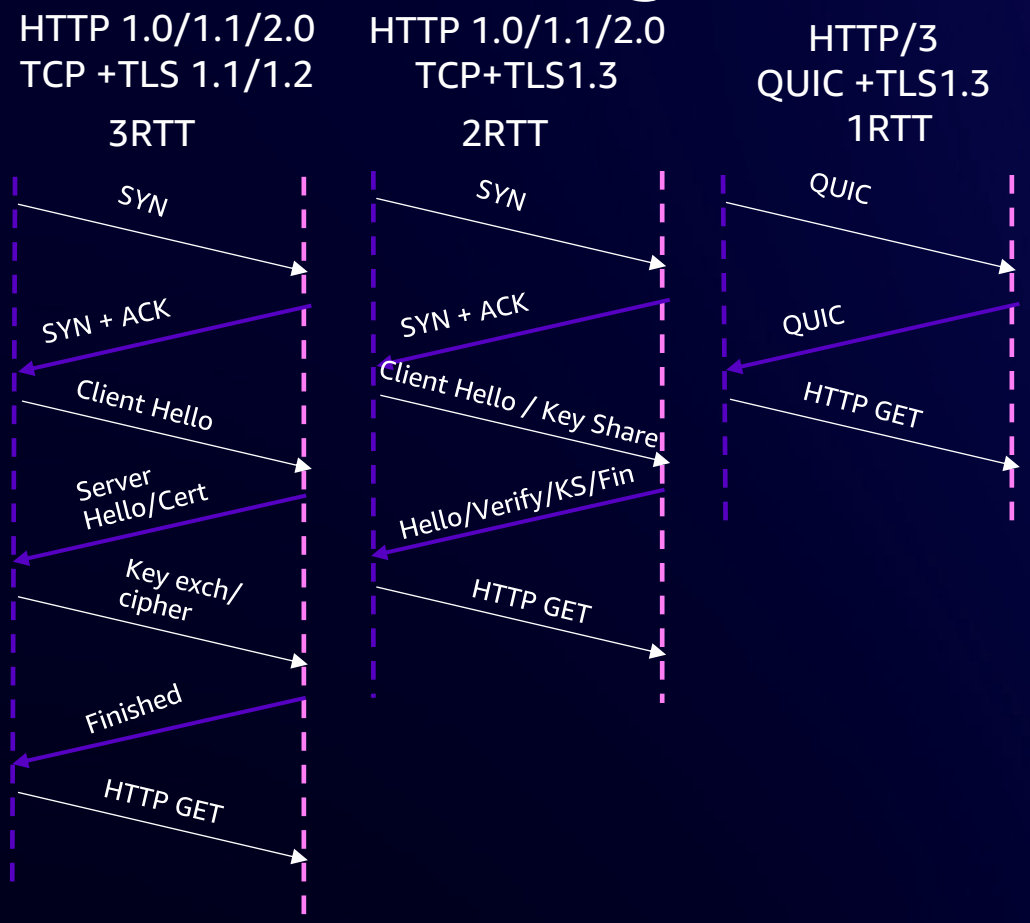


Dynamic Content Acceleration



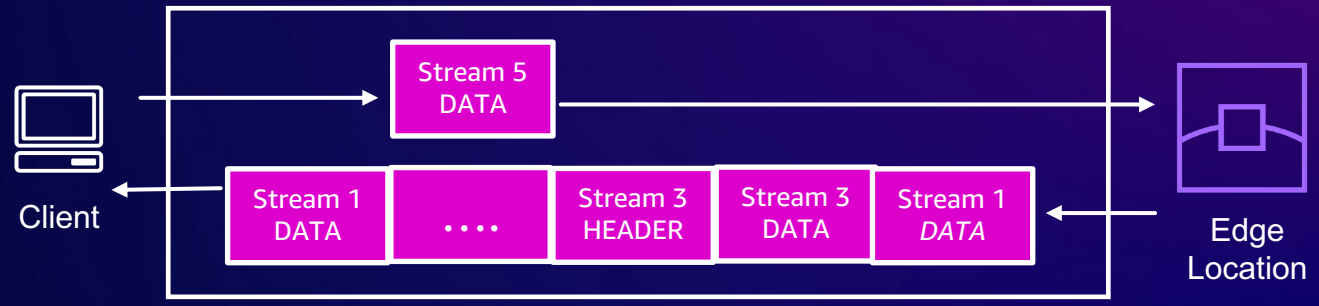


Protocol Progression



*Additional optimizations (not pictured)
- TCP Fast Open
- TLS Session Resumption

HTTP 2.0 Connection



WebSocket Protocol

```
GET /chat HTTP/1.1
Host: server.example.com
Upgrade: websocket Connection:
Upgrade Sec-WebSocket-Key: bsZSBub25jZQ==
Origin: https://example.com
Sec-WebSocket-Protocol: chat, superchat
Sec-WebSocket-Version: 13
```

gRPC / HTTP2

```
message HelloRequest {
        string firstName = 1;
        string lastName = 2;
}
```

Granular Configuration with Edge Functions

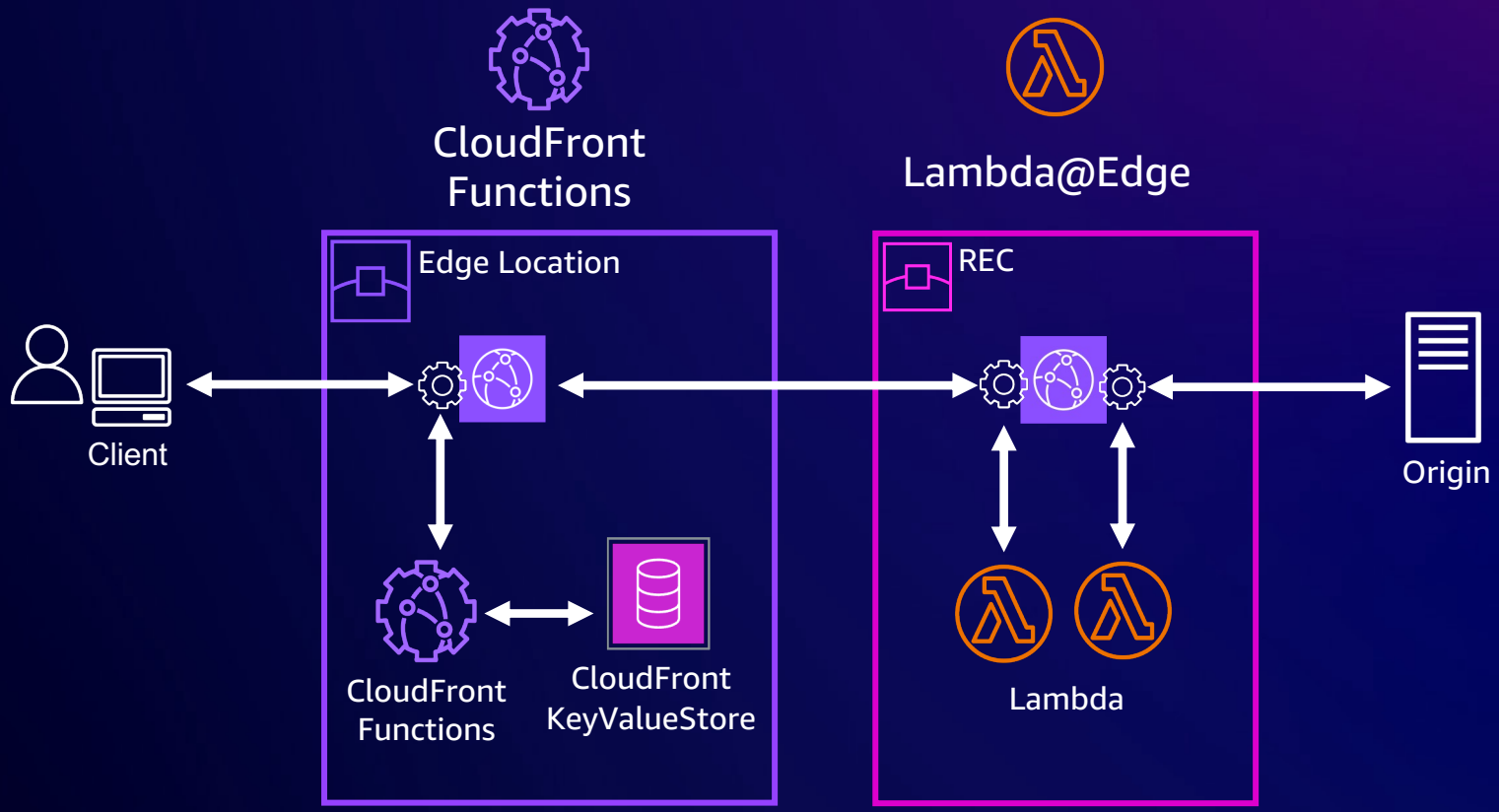
Request / Response Manipulation

Cache Key Normalization

URI Re-write / Redirect

Custom Response Generation

Software Integration



High volume, latency sensitive functions

Longer running functions, capable of network calls

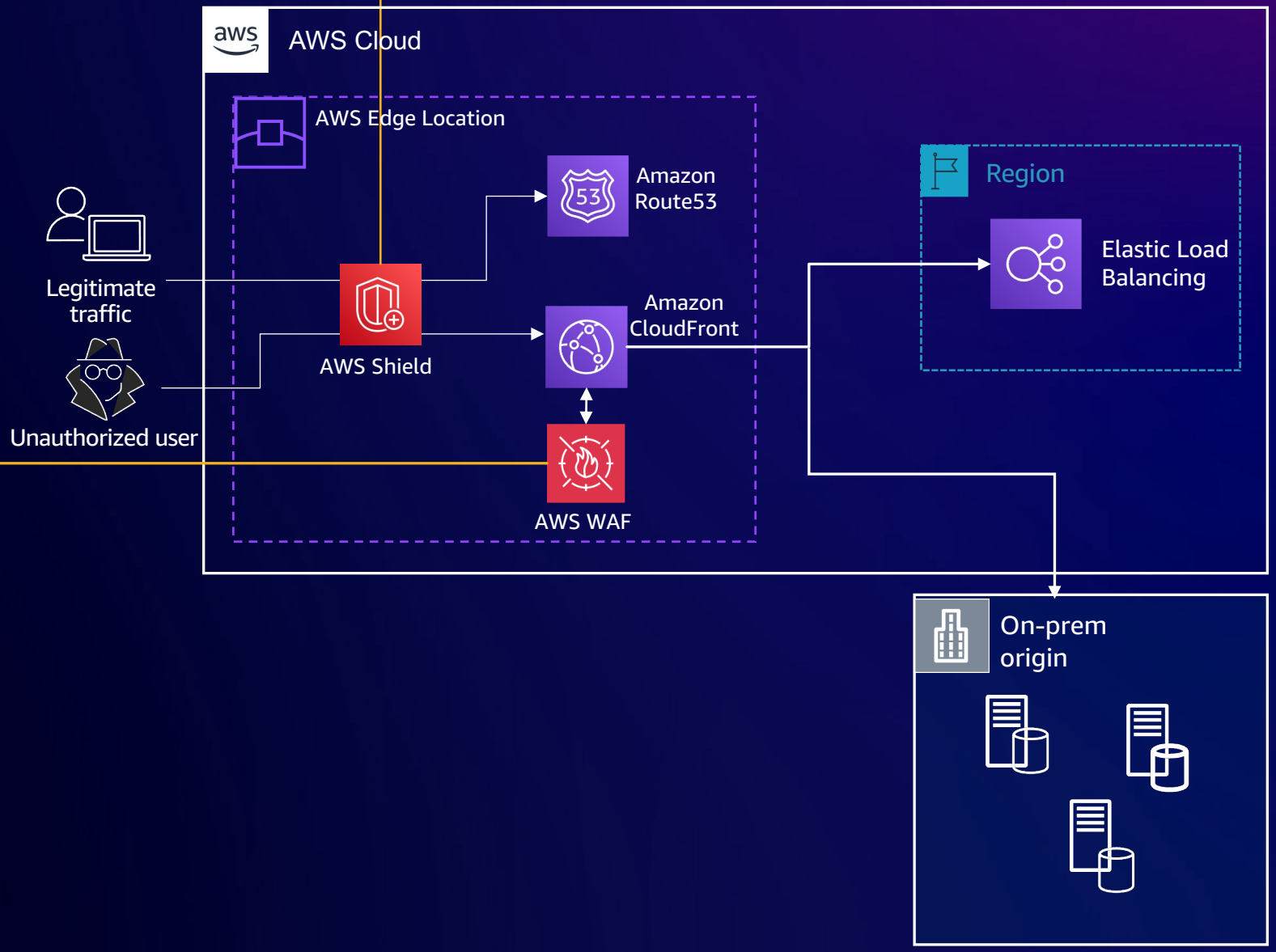
Secure Perimeter

Layer 3/4

- Black Watch
- SYN Proxy
- Continuous inspection (inline)
- Packet validation
- Distributed scrubbing capacity
- Automated routing policies to absorb large attacks

Layer 7

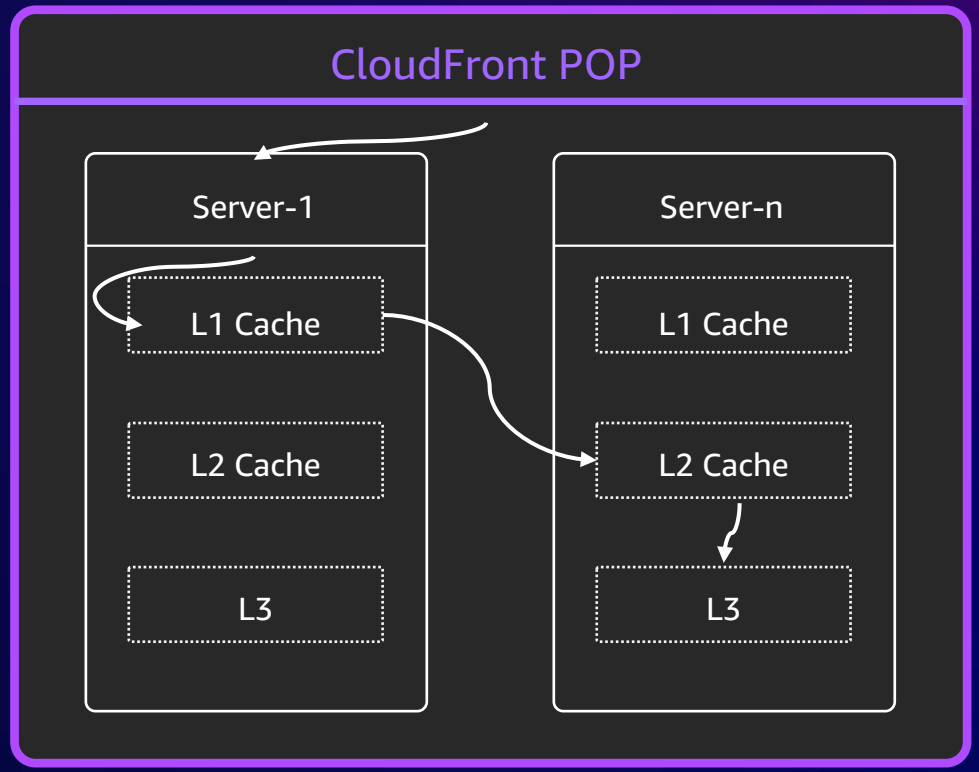
- Bot Managment
- RateLimiting
- Managed Rules (OWASP)
- Custom Rules
- Security Automation



POP Architecture

Technology Stack

NGINX SQUID



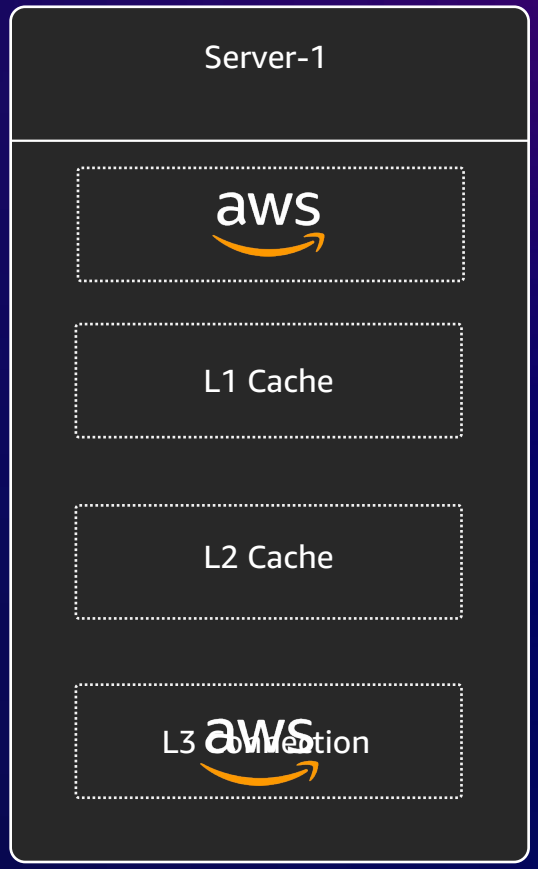
Challenges

- Designed for outdated protocols
- Suboptimal for CPU expensive requests
- Contains capabilities not relevant to CDN
- Difficult to update with latest security features

Re:Inventing Amazon CloudFront

aws
web server

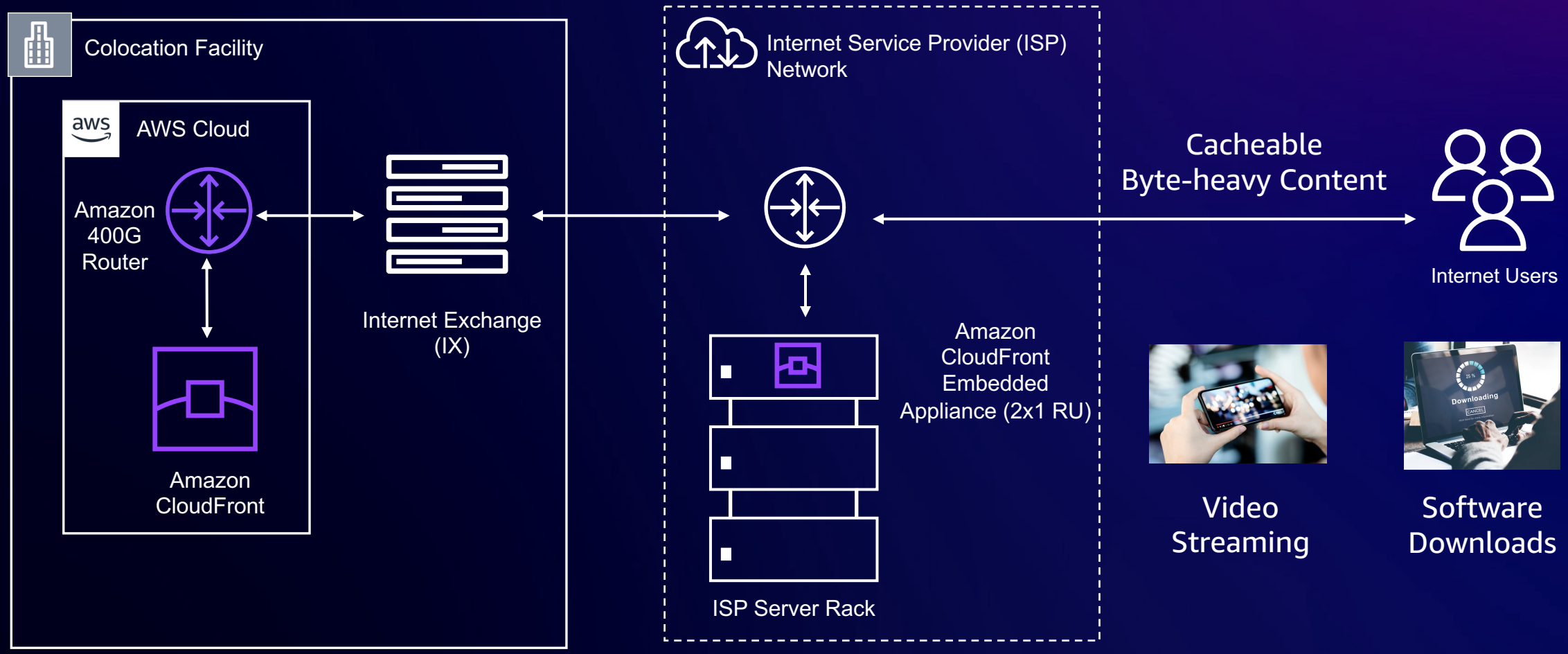
- AWS Built server
- Built on Tokio Runtime maintained by AWS engineers
- Built to enable QUIC HTTP/3
- Multi-threaded, work stealing scheduler
- Written Rust



100ms improvement for HTTP2/3 connections

Embedded POP

95% Offload
65% FBL reduction



AWS Global Accelerator

AWS Global Accelerator

- Global static anycast IP addresses for applications
- Route to Elastic Load Balancer or direct to Instances
- Accelerate TCP and UDP traffic over the AWS global backbone network

Why does a redundant and available backbone matter?

An availability story:

- Third-party ISP had a fiber cable event in Southeast Asia: 48+ hours of impact
- End users could not connect to endpoints in AWS Singapore Region
- Customer onboarded to AWS Global Accelerator: recovery in minutes

Designed for high availability

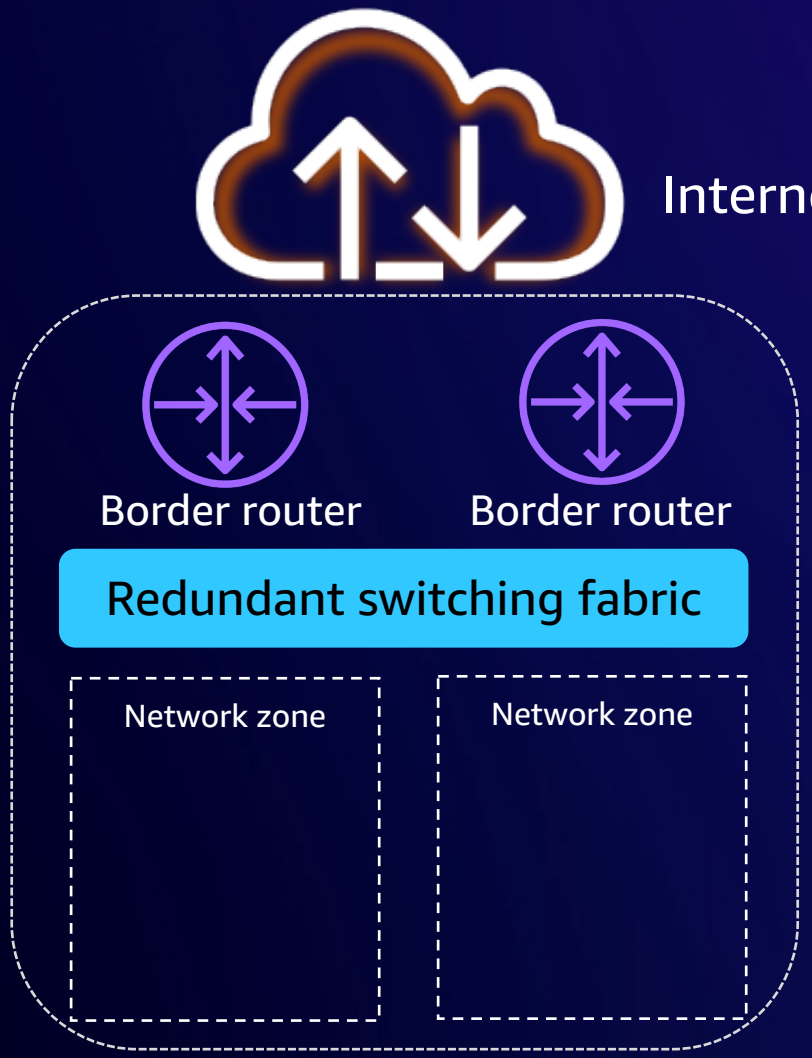
Resilient architecture

- Network zones
- Cellular architecture
- Shuffle sharding

Handling impairments

- Route around transit failures
- Monitor endpoint health
- Fast failover

Built-in redundancy with network zones

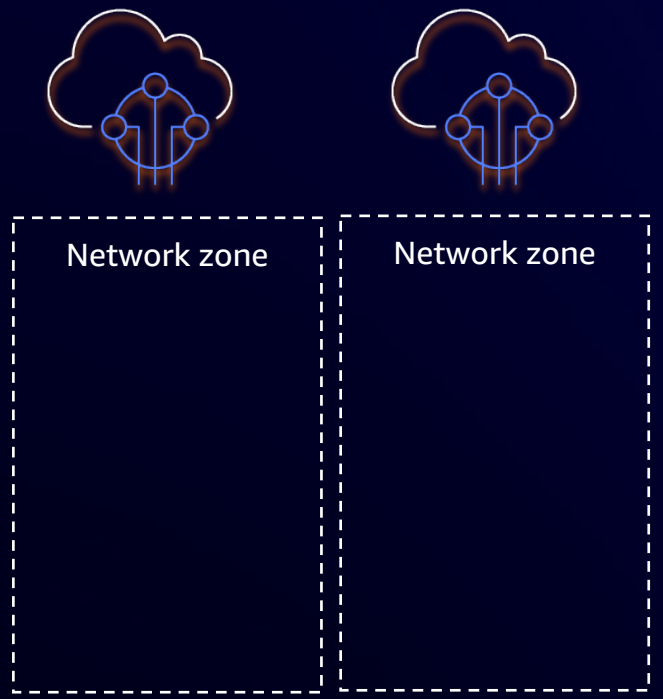


Connect to external networks (e.g., via peering)

AWS Global Accelerator Points of Presence (PoPs)

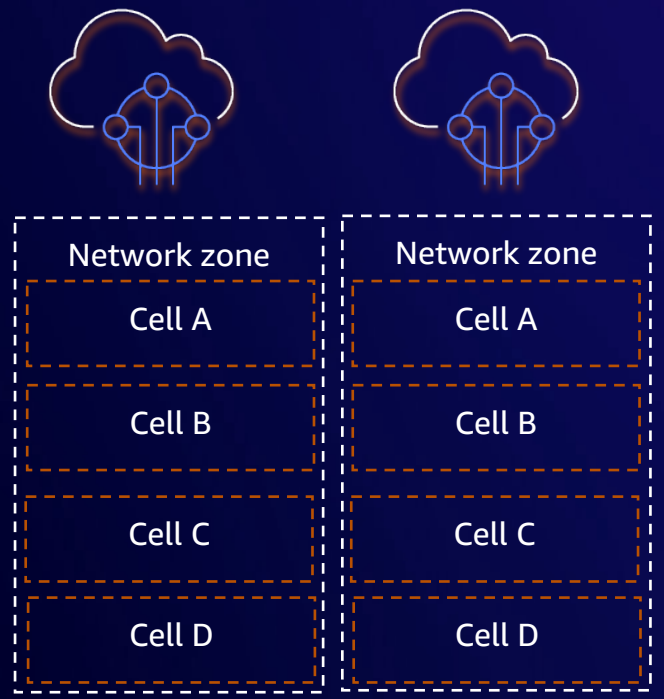
High availability: Cellular architecture

Network zones



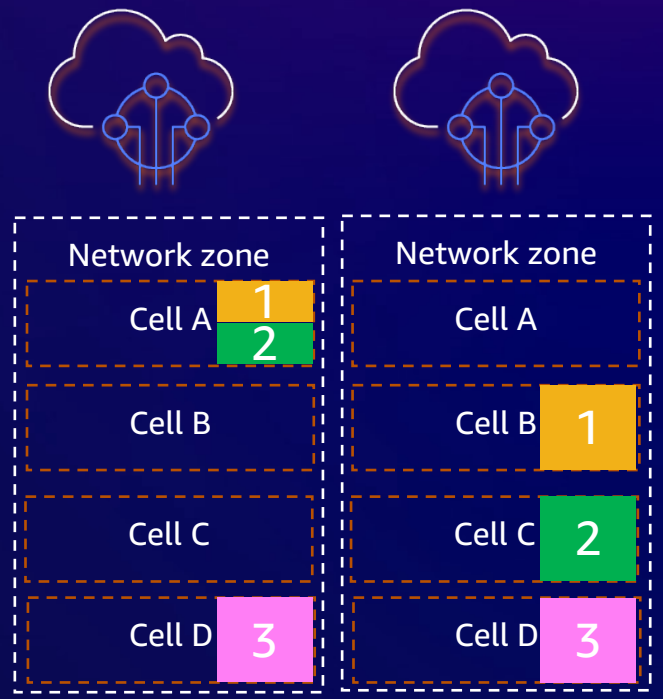
- Two network zones per accelerator
- Each anycast static IP address comes from separate network zones

Cellular architecture



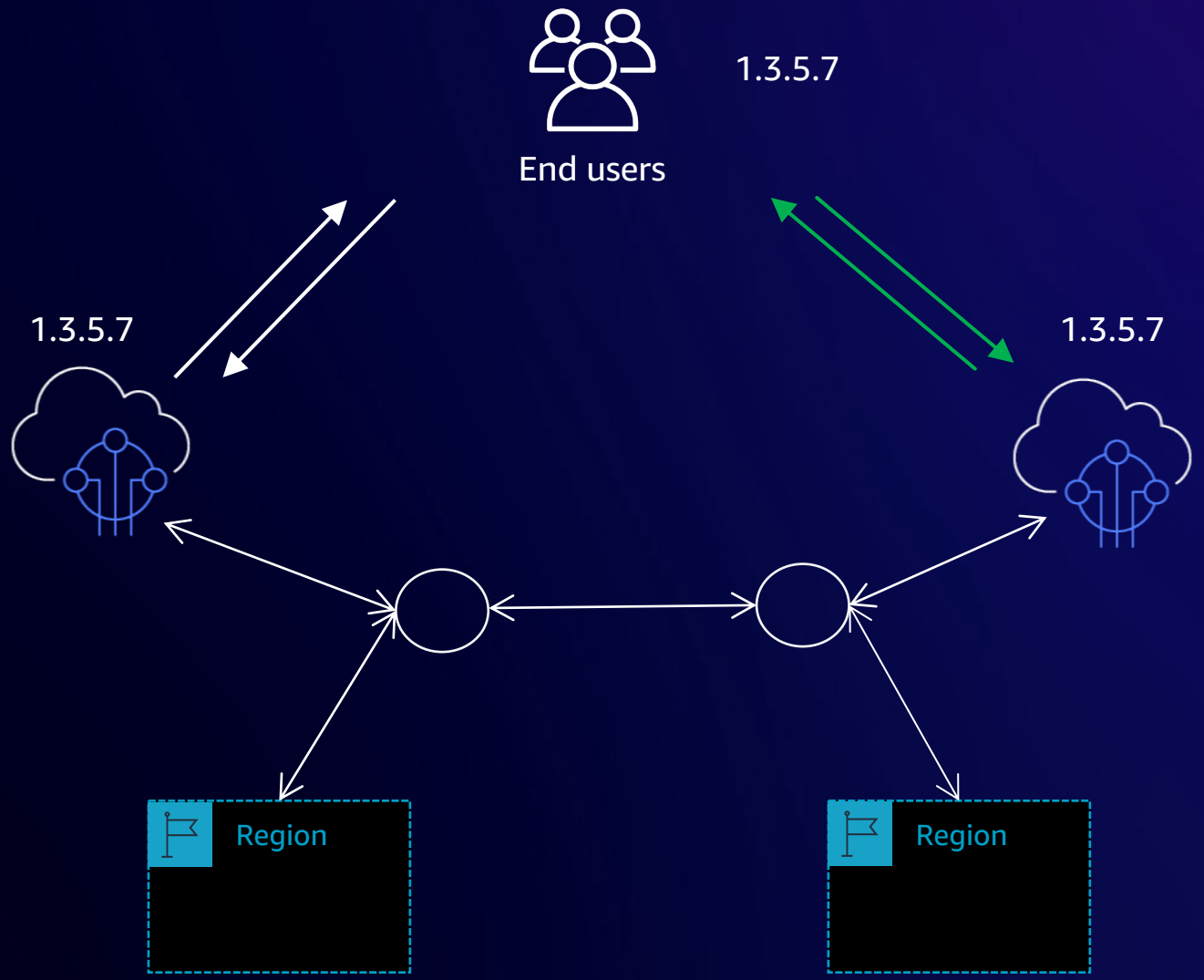
- Each network zone has four cells
- Each cell has multiple servers

Shuffle sharding

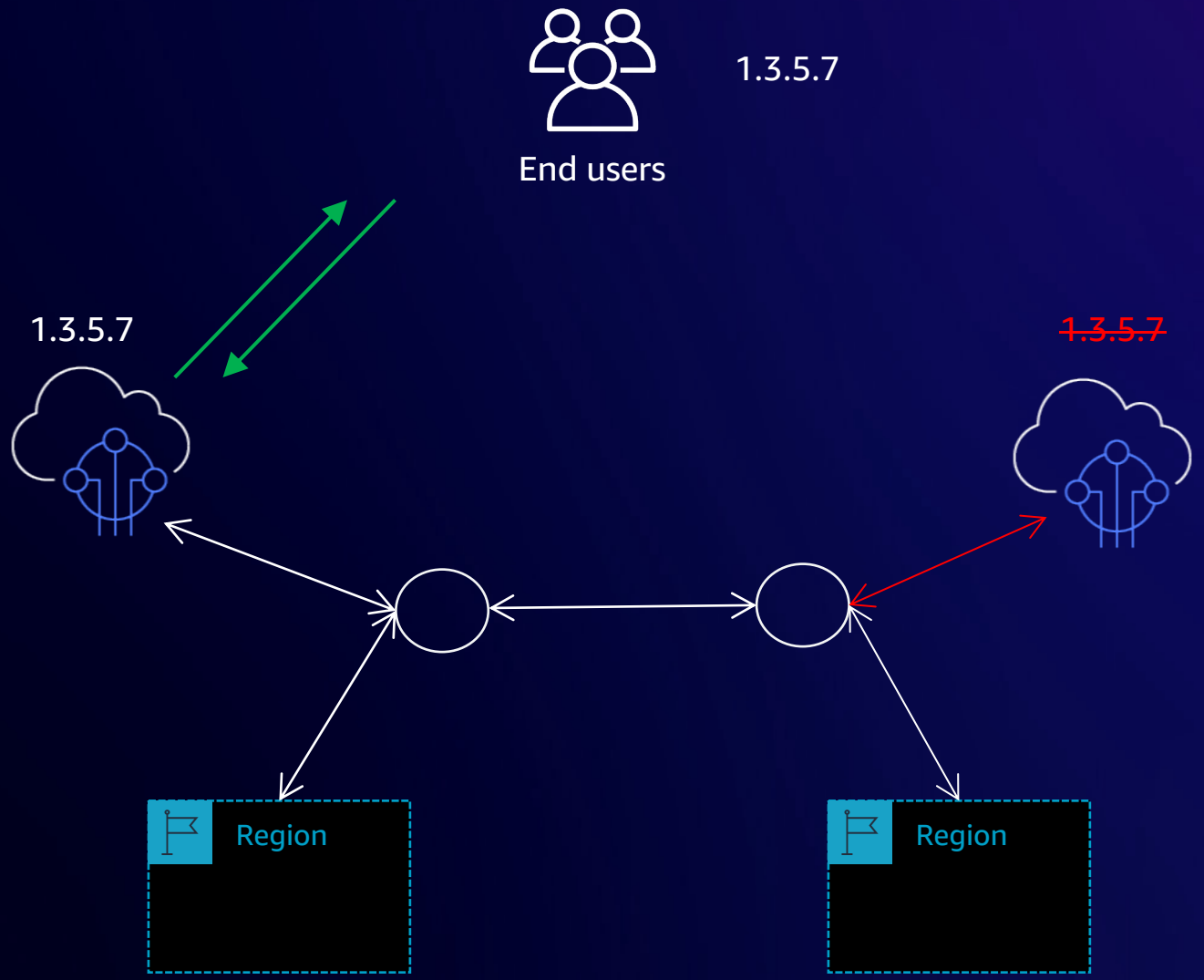


- Customers shuffled across cells to reduce "noisy neighbor" issues

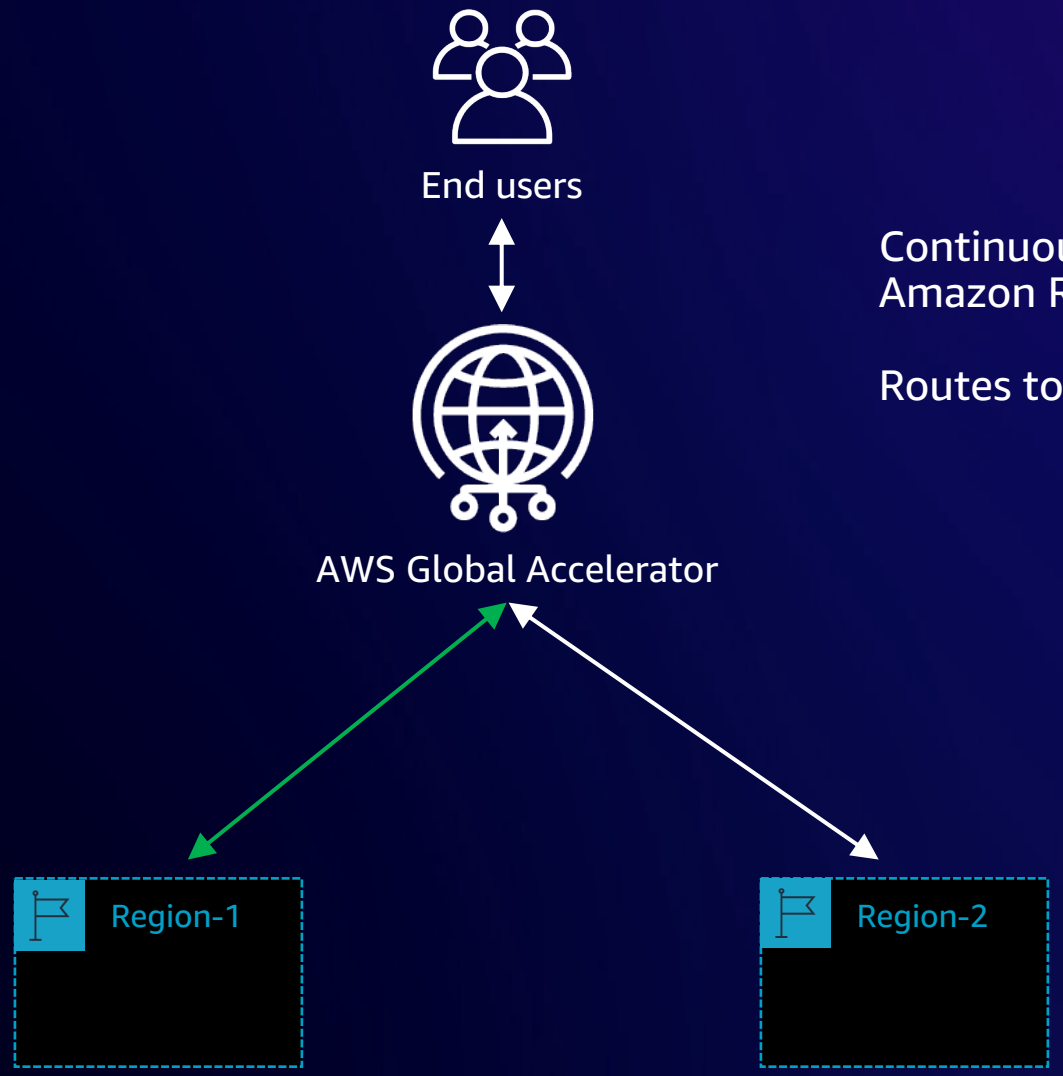
Self-healing network



Self-healing network



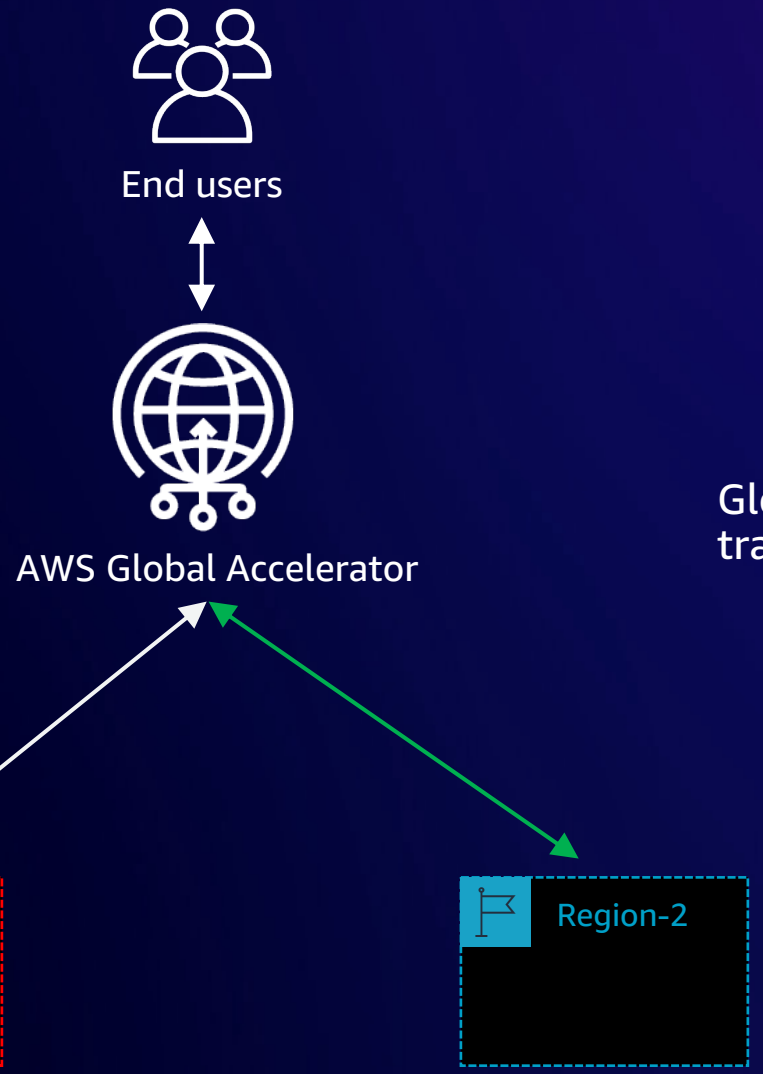
Healthy endpoints and fast failover



Continuously checks health status of endpoints
Amazon Route 53 health checks

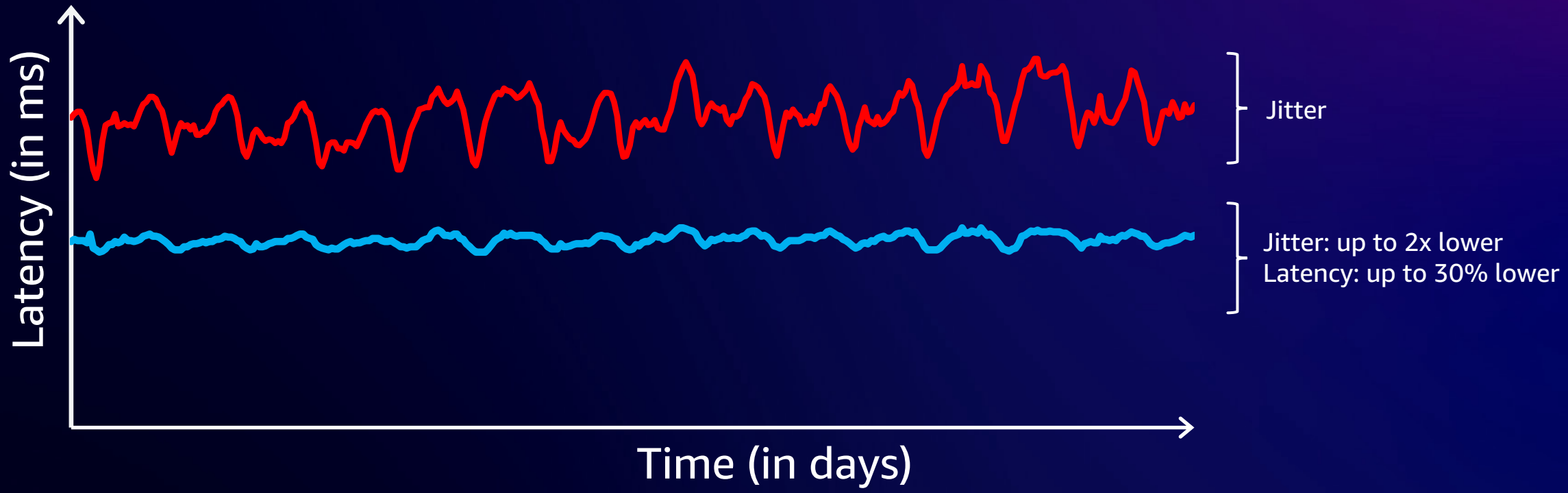
Routes to healthy endpoints

Healthy endpoints and fast failover



Global Accelerator automatically shifts traffic to next healthy available endpoint

Lower latency and jitter vs. public internet*



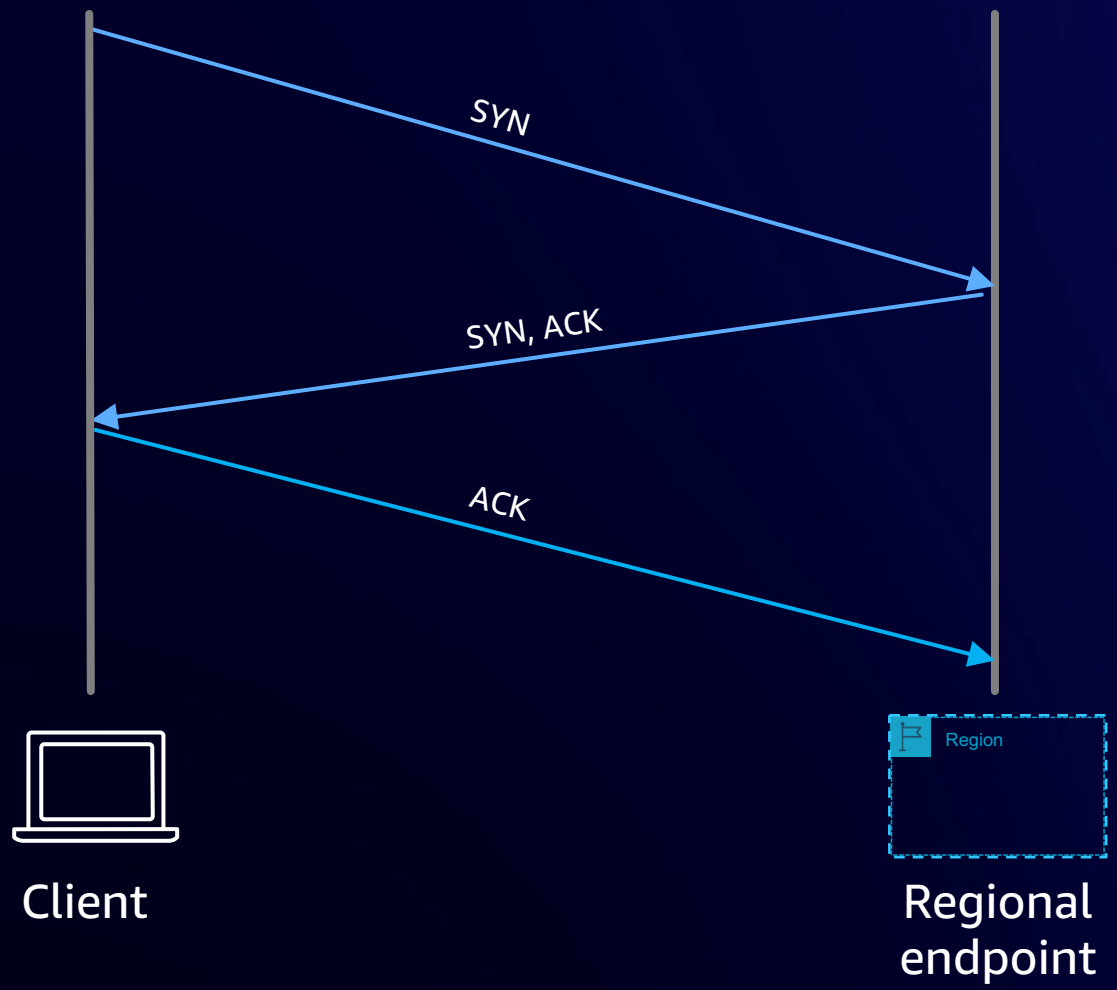
AWS Global Accelerator

Public internet

P90 first byte latency from client locations in US to endpoints in EU-West-1 with TCP termination

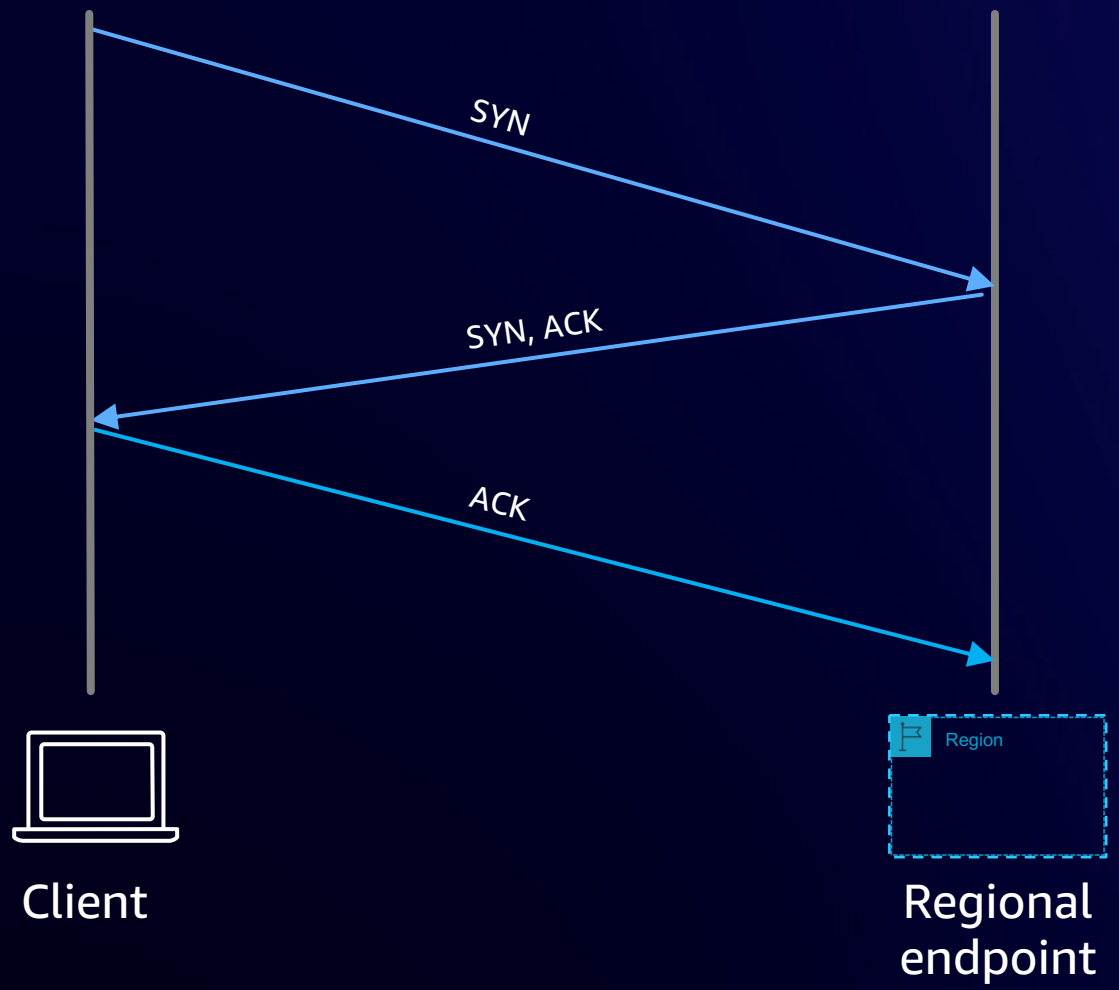
* Actual numbers are obscured on graph

TCP termination at edge

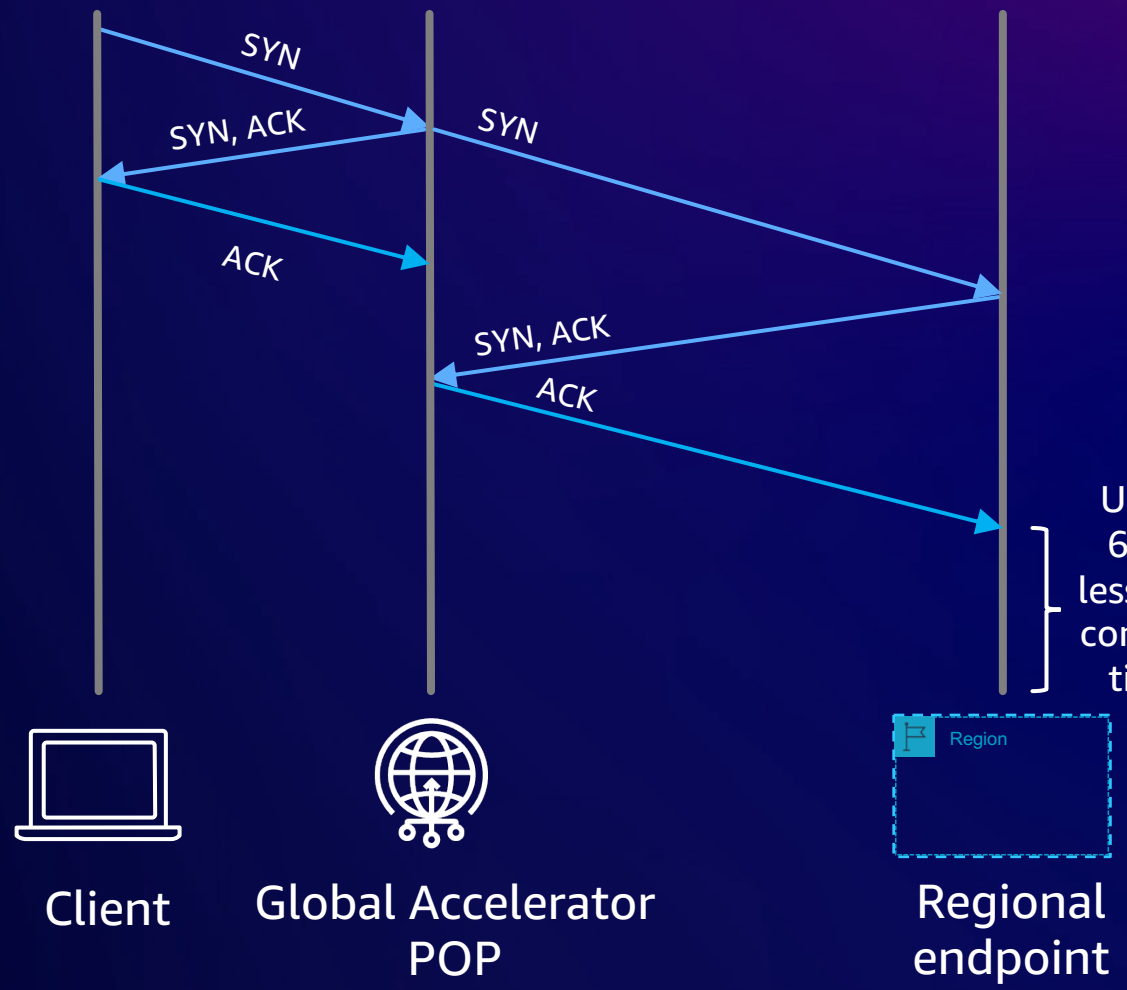


Without Global Accelerator

TCP termination at edge

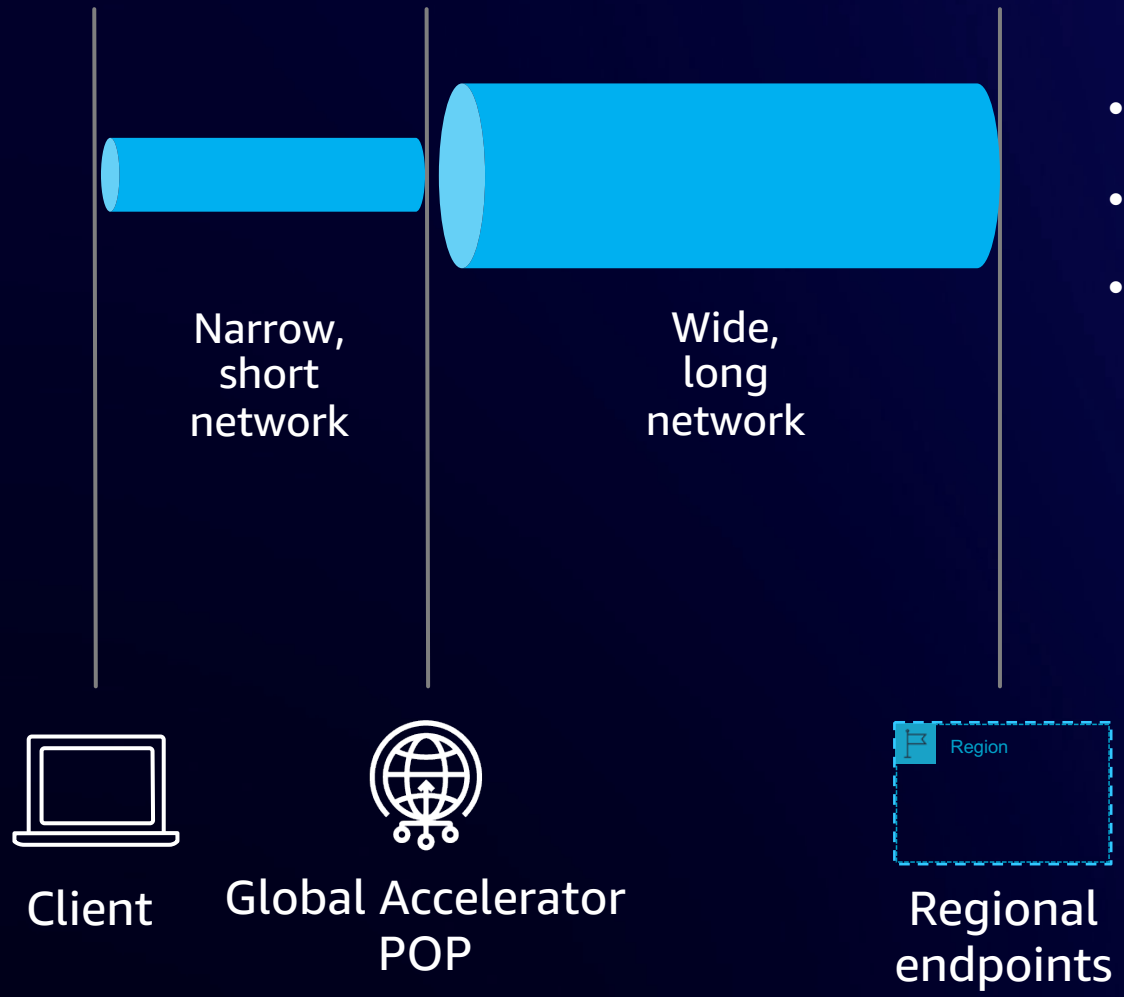


Without Global Accelerator



With Global Accelerator

Up to 60% improvement in TCP throughput



- Enables fast handshake between clients and endpoints
- Jumbo frames between edge and AWS Region
- Uses TCP buffers and larger TCP window to achieve higher throughput

IPv6 support for all endpoints

- Dual-stack accelerator for routing to dual-stack endpoints
- Two static anycast IPv6 addresses in addition to two IPv4 addresses
- Unique dual-stack DNS name, for controlled migration
- Optionally, upgrade your IPv4 accelerator to dual-stack

Q&A

Thank you!

Giorgio Bonfiglio

Email bonfigg@amazon.com
LinkedIn giorgiobonfiglio
Twitter @g_bonfiglio