# AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

NET314-R2

# Elastic Load Balancing best practices


Jon Zobrist (he/him)
SysDev Manager
Elastic Load Balancing
AWS

Priya Patel (she/her)
Product Manager
Elastic Load Balancing
AWS


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Agenda

Elastic Load Balancing – ALB, NLB

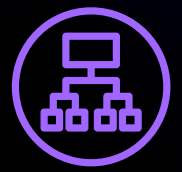
New features, highlighted use cases, customer stories

Availability on Elastic Load Balancers – deep dive

Operations / Planning for production

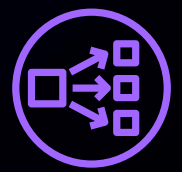
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Elastic Load Balancing Portfolio



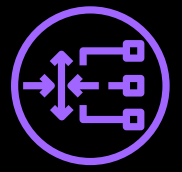
## Application Load Balancer

- Request level balancing
- HTTP, HTTPS, and gRPC support Advanced rule-based routing



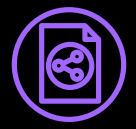
## Network Load Balancer

- Connection level balancing
- TCP, TLS, and UDP support
- Built for scale, low latency and high performance



## Gateway Load Balancer

- Transparent network gateway and load balancer
- Single entry and exit point for all IP traffic
- Improve appliance availability
- Elastically scale and reduce costs



## Classic Load Balancer

- L4-7 load balancing
- Built for the EC2-Classic environment

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Benefits of using Elastic Load Balancers



Scalability



Security



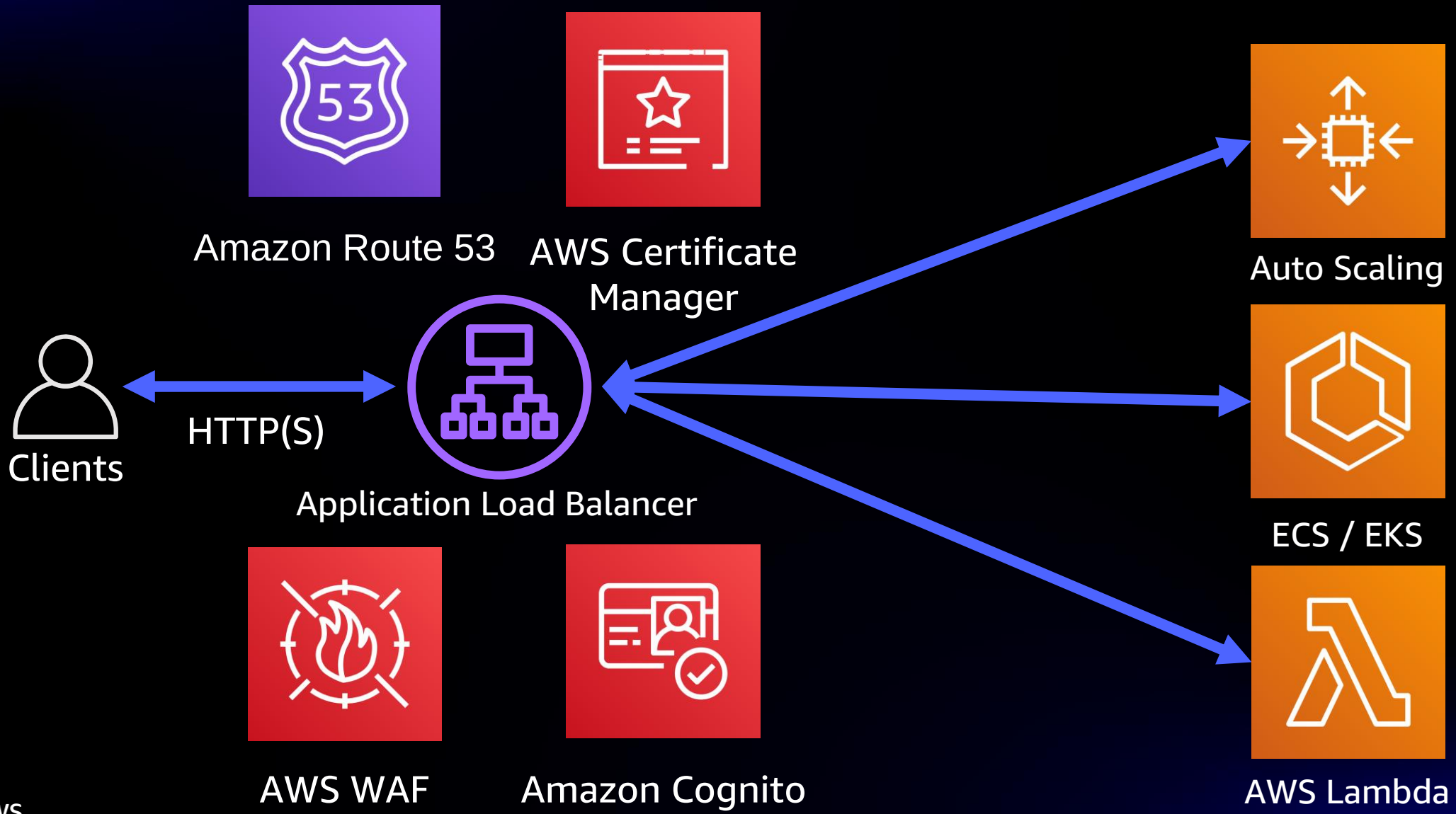
Availability

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Application Load Balancer – Overview, Features, Customer stories

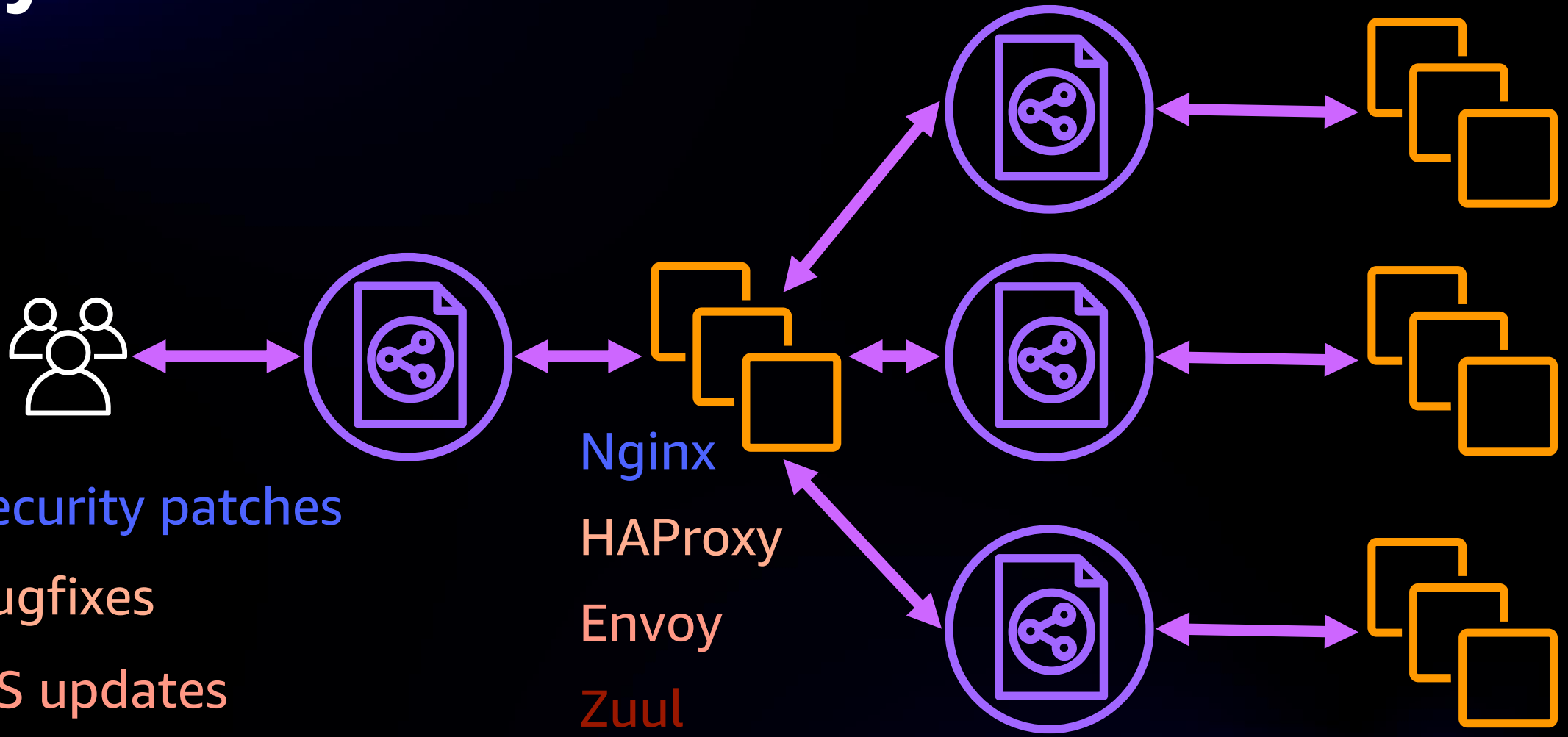
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Application Load Balancer (ALB)




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Why we built ALB – The ELB Sandwich




Security patches

Bugfixes

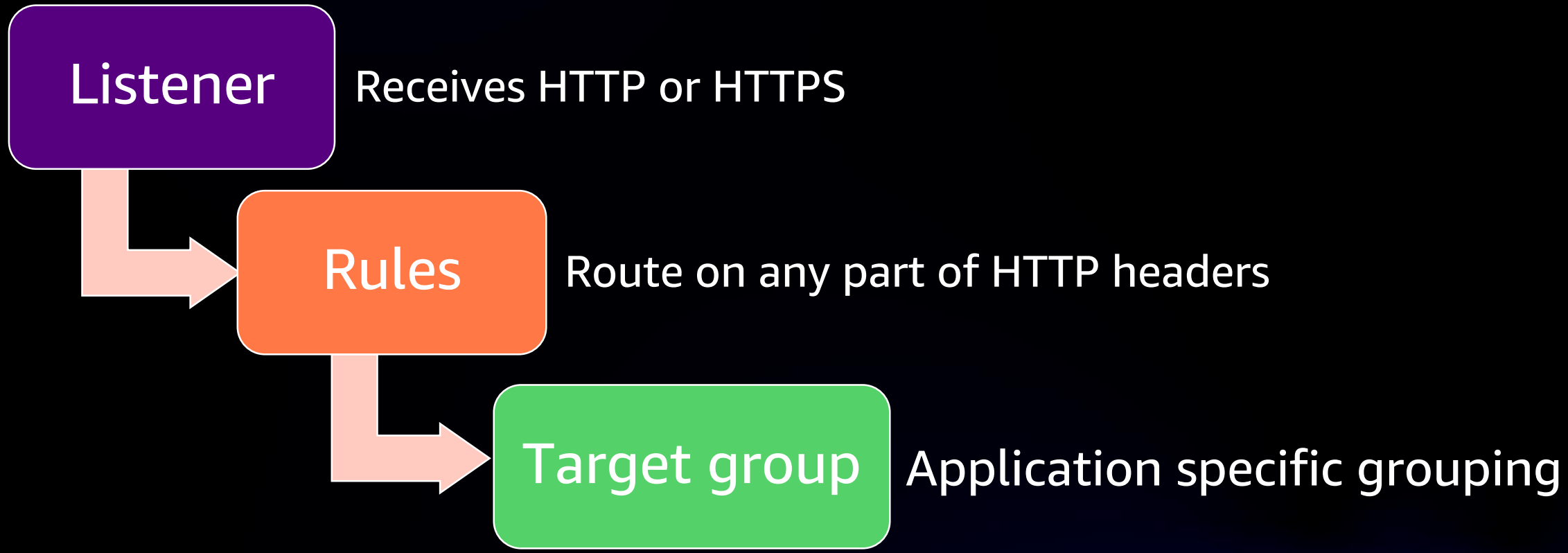
OS updates

Configuration changes

Scaling

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Logical overview of ALB




aws

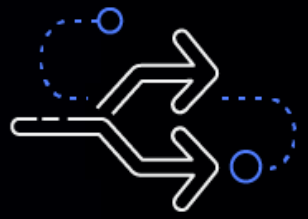
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB: Recently Launched Features



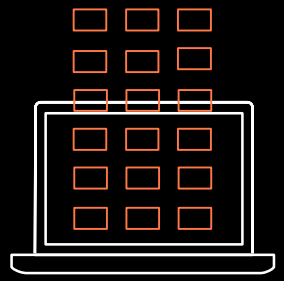
## Security and Compliance

- Desync Mitigation Mode
- Non-CBC Security Policy



## Routing

- WAF Fail Open
- Application Cookie Stickiness



## Protocols

- gRPC
- XFF Header for Source Port

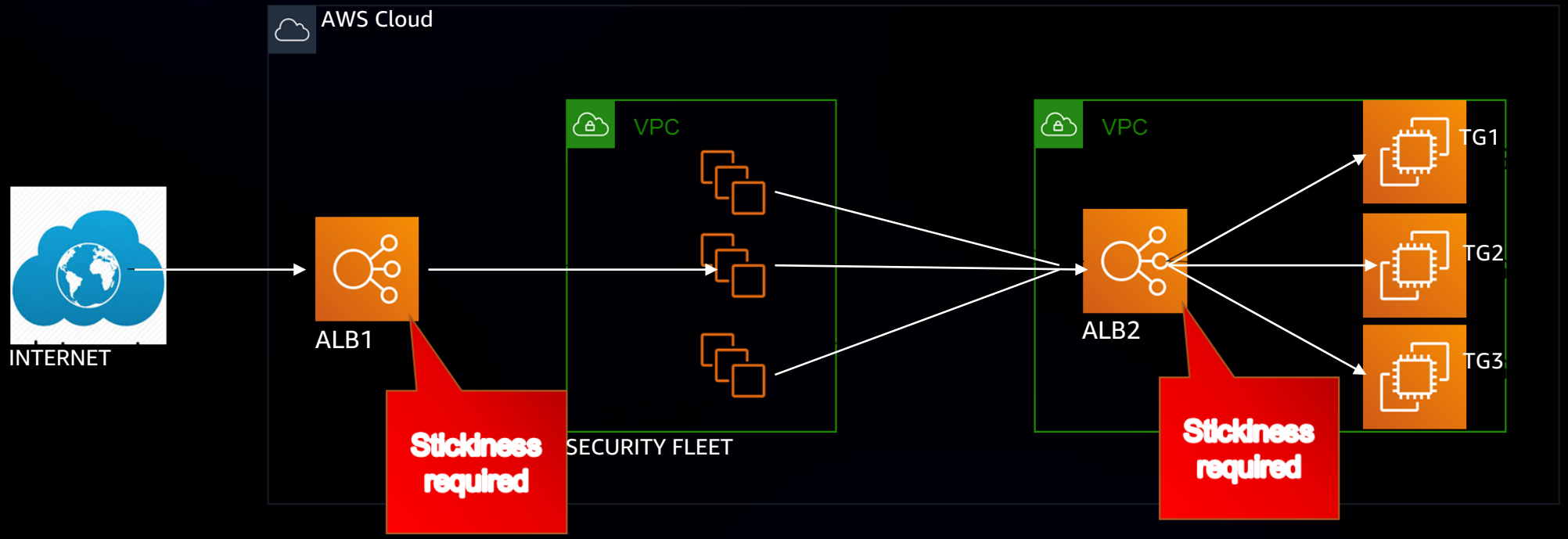


## Additional Capabilities

- ALB on Outposts
- Kubernetes Integration (AWS LB Controller)
- ALB-as-a-target of NLB

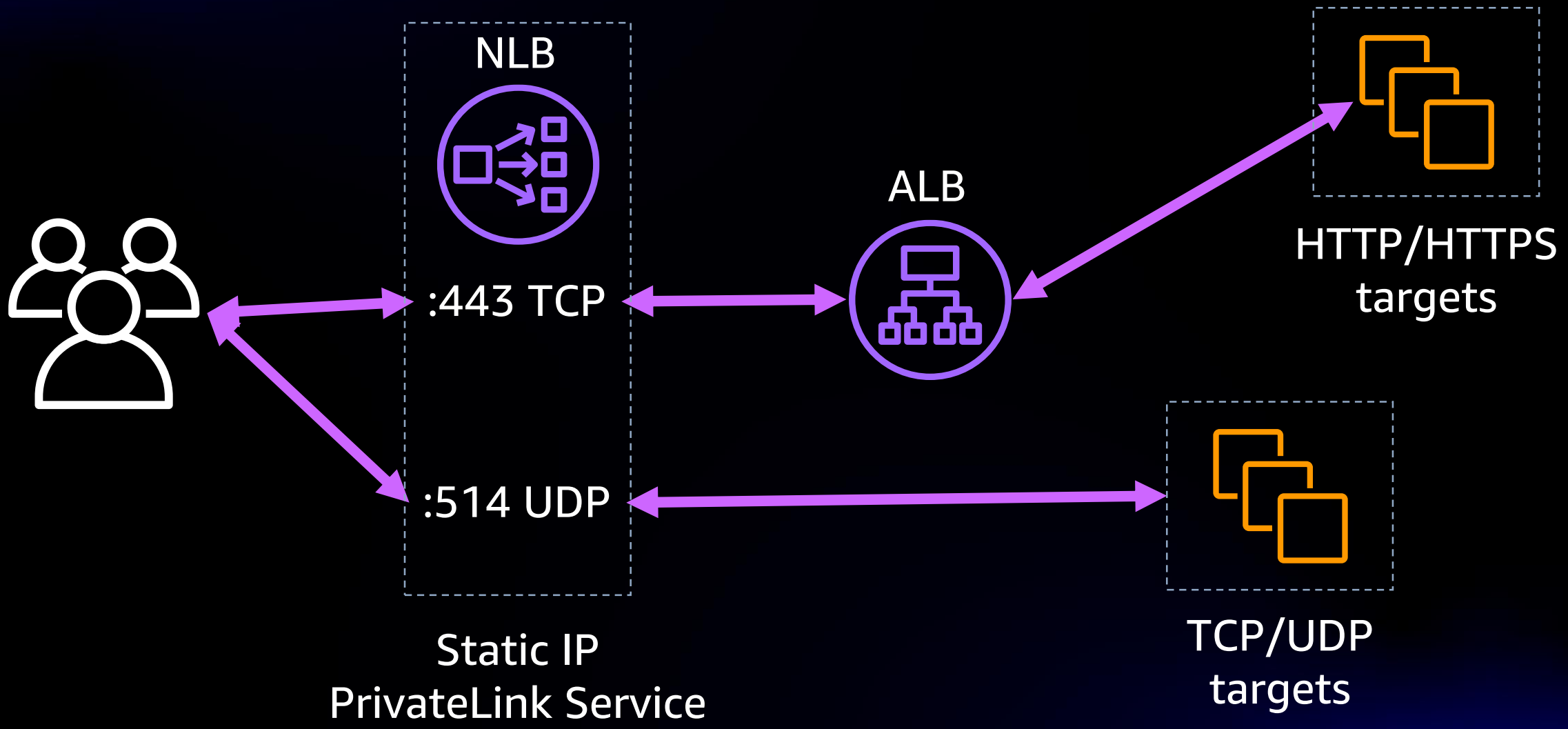
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB: Application Cookie Stickiness




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

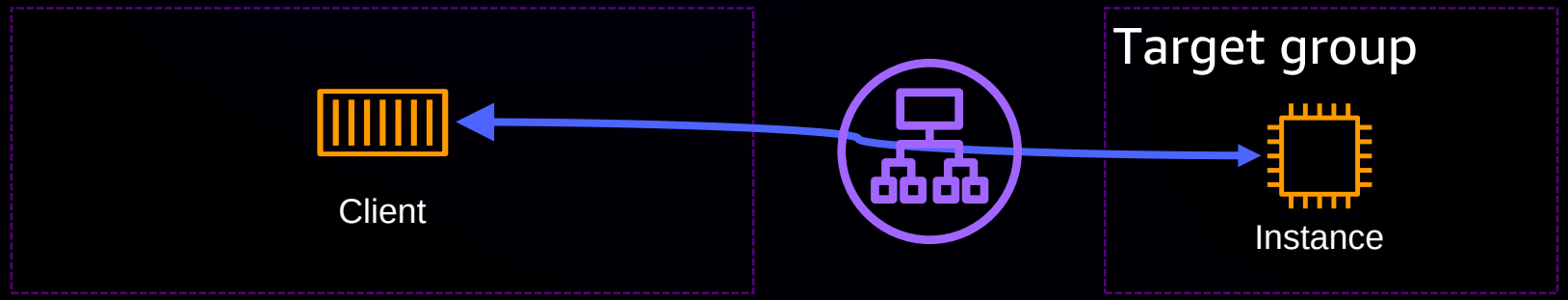
# ALB: ALB as a Target of NLB



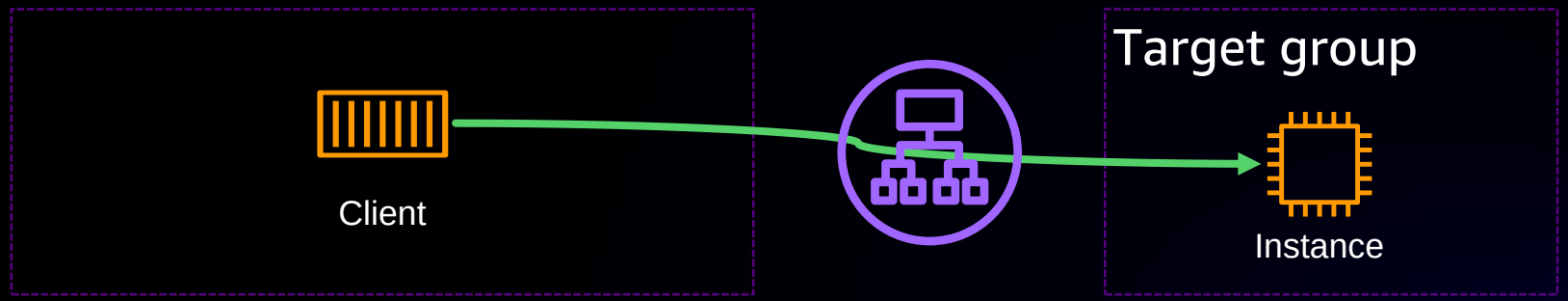

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB: gRPC and HTTP/2 Application Load Balancer

## gRPC listener, and target groups




## HTTP/2 target groups




gRPC

Unary calls

Streaming:

Client

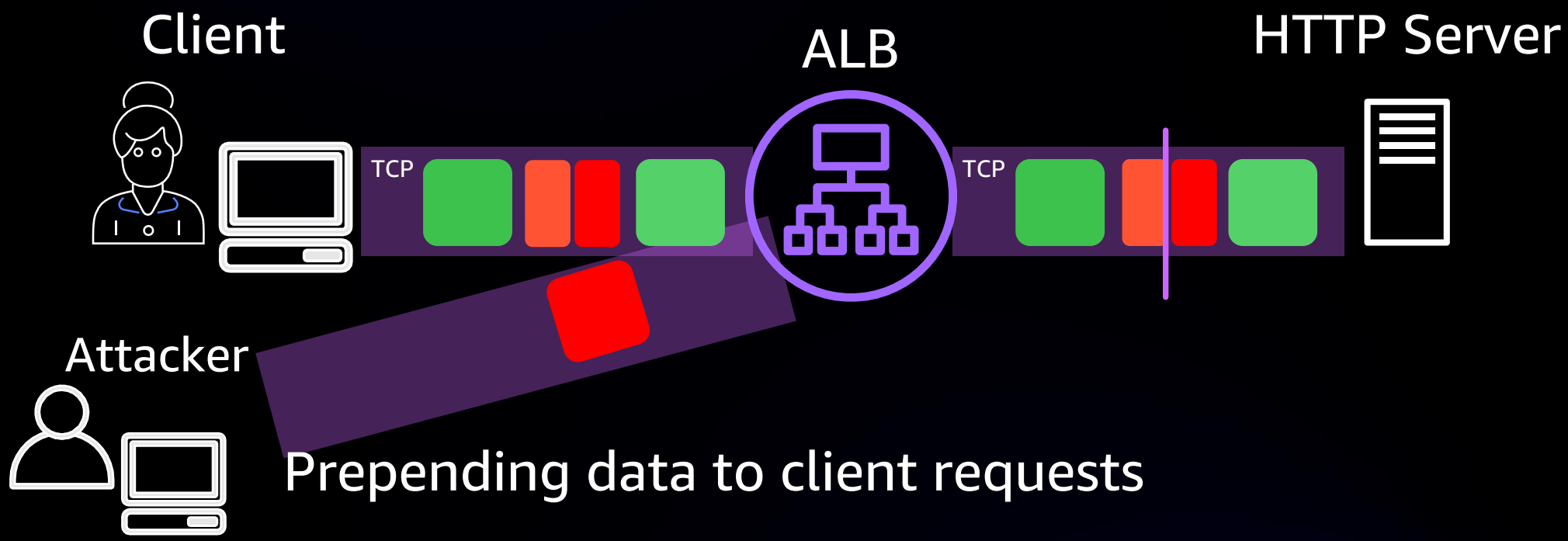
Server

Bi-directional

HTTP/2

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB: HTTP DeSync mitigation




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Awesome applications





















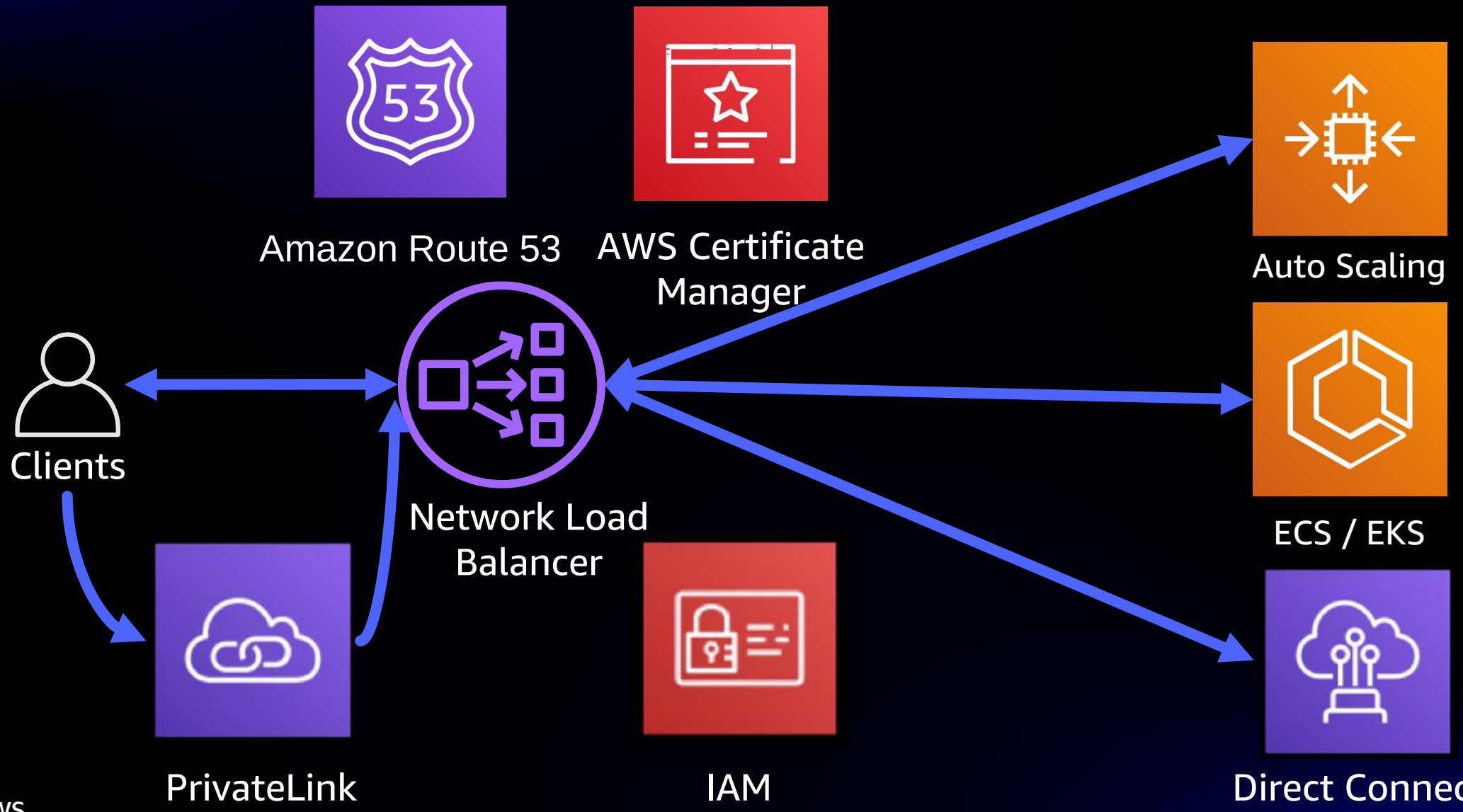



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Network Load Balancer – Overview, Features

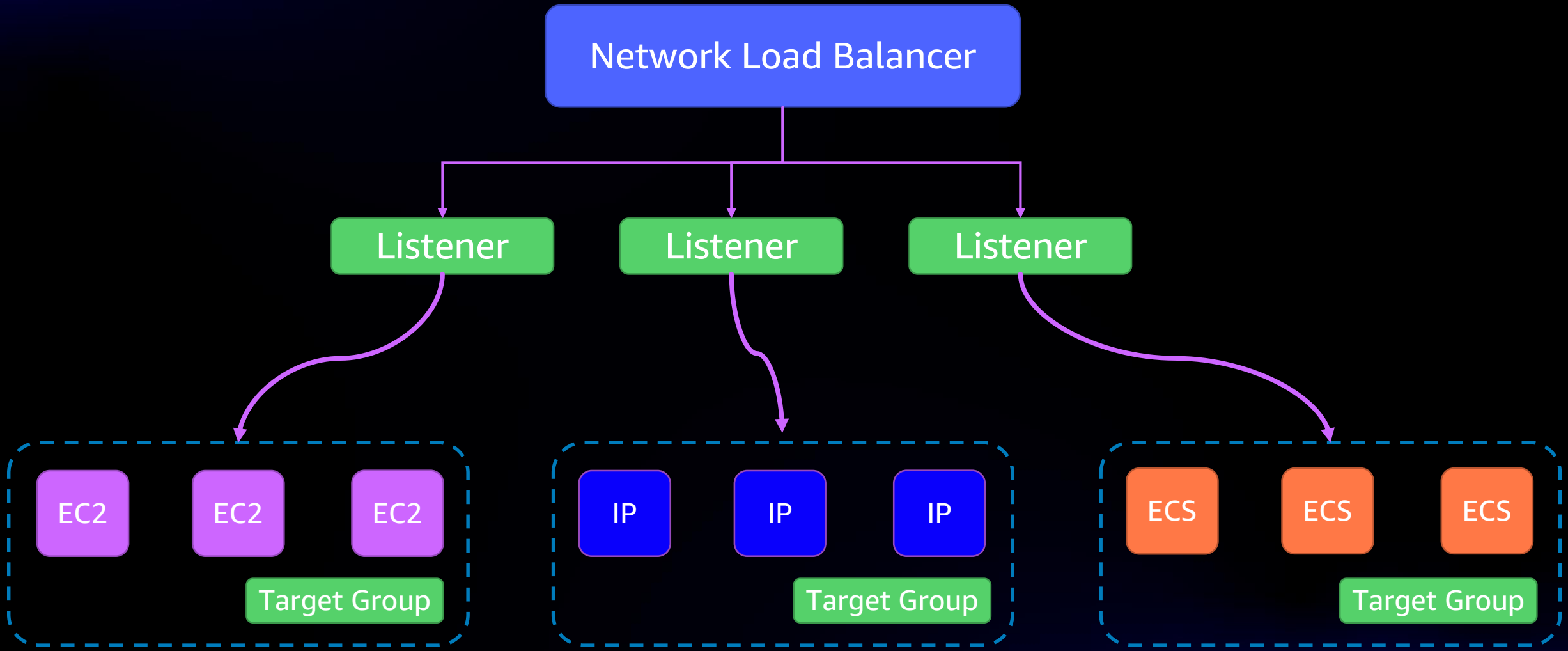
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Network Load Balancer (NLB)




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Logical overview of NLB



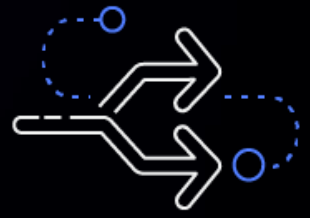

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# NLB: Recently Launched Features



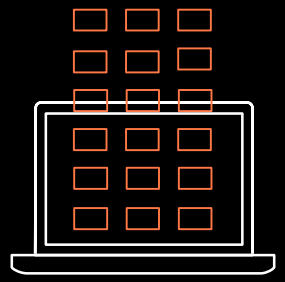
## Security and Compliance

- ALPN Support
- Non-CBC Security Policy
- TLS 1.3



## Routing

- IP Session Affinity
- Terminate Connections to Deregistered Targets



## Protocols

- UDP for IP Targets
- IPv6 Dual Stack



## Additional Capabilities

- Kubernetes Integration (AWS LB Controller)

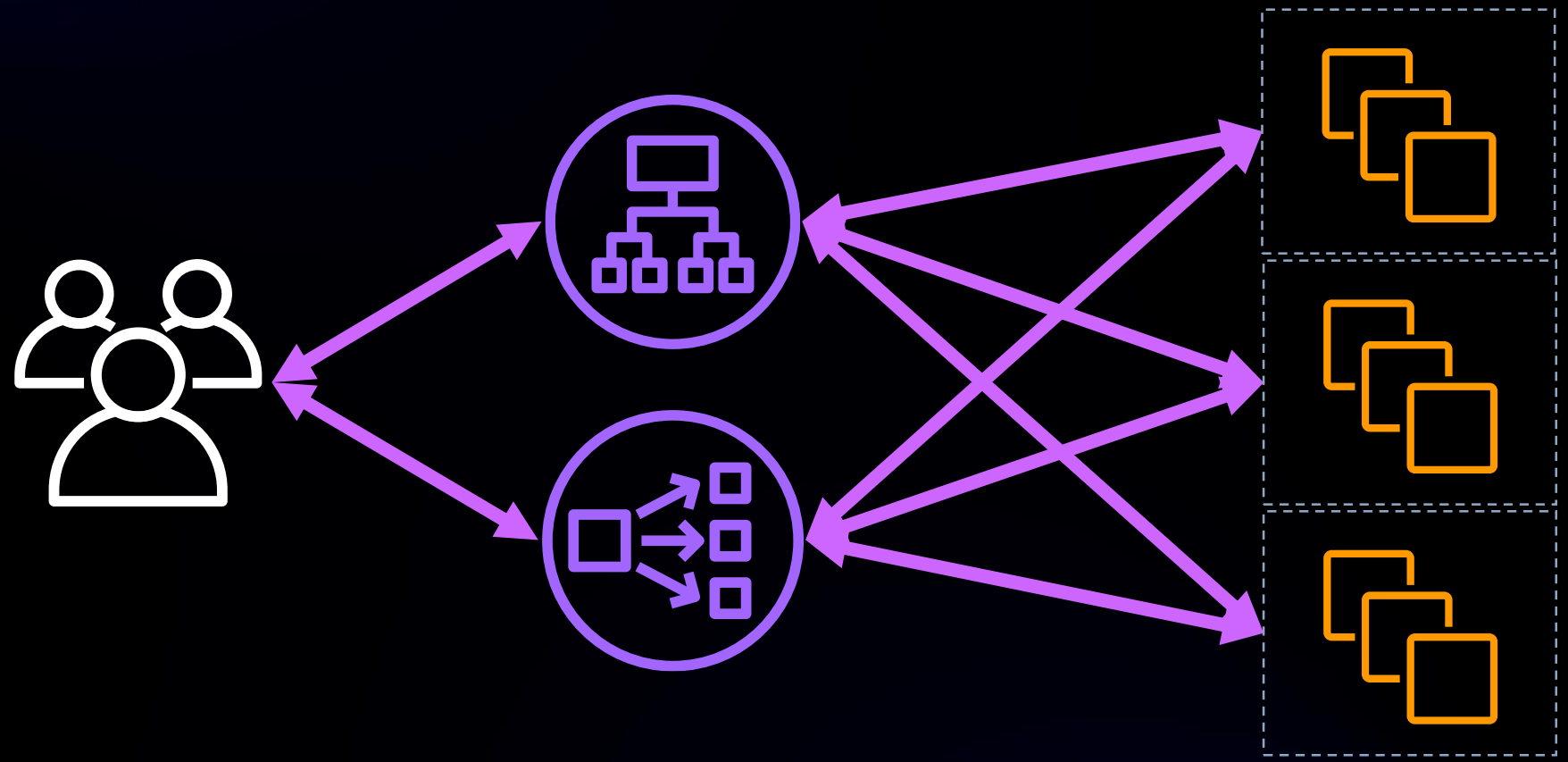
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELBs improve availability

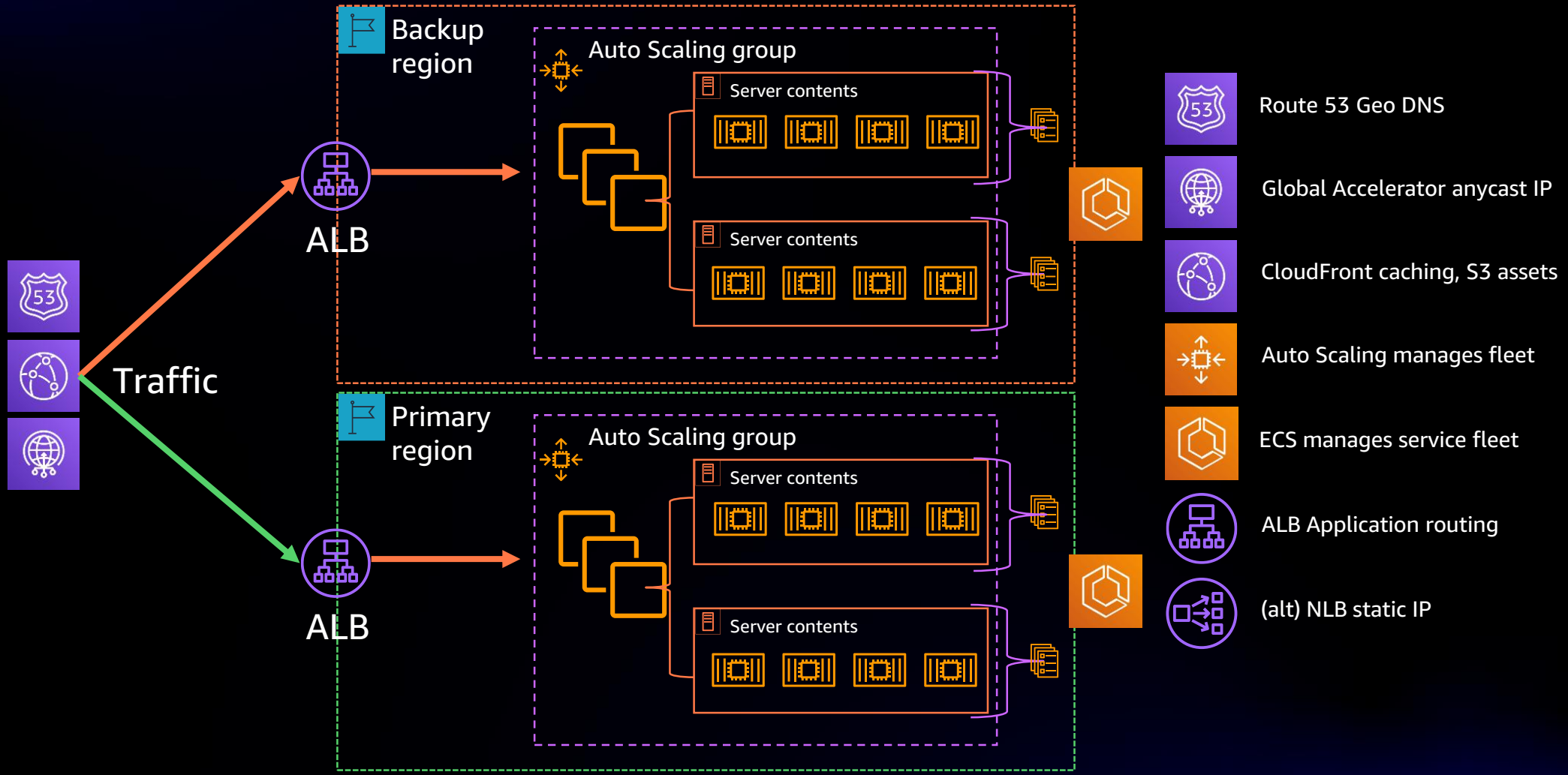
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Elastic Load Balancers improve availability



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB integrations improve availability




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Services in front of ELB improving availability




Amazon Route 53



Amazon CloudFront

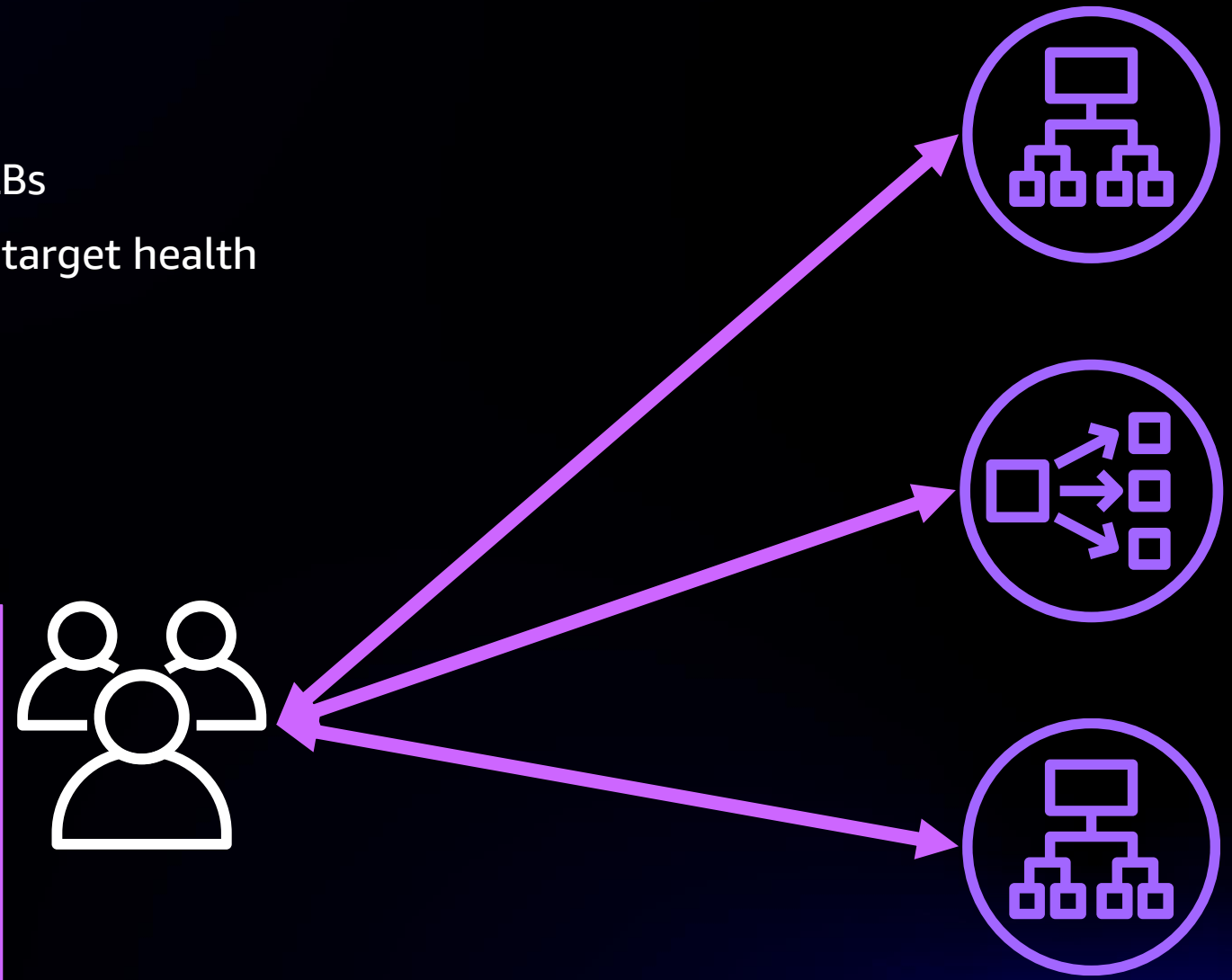


AWS Global Accelerator

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

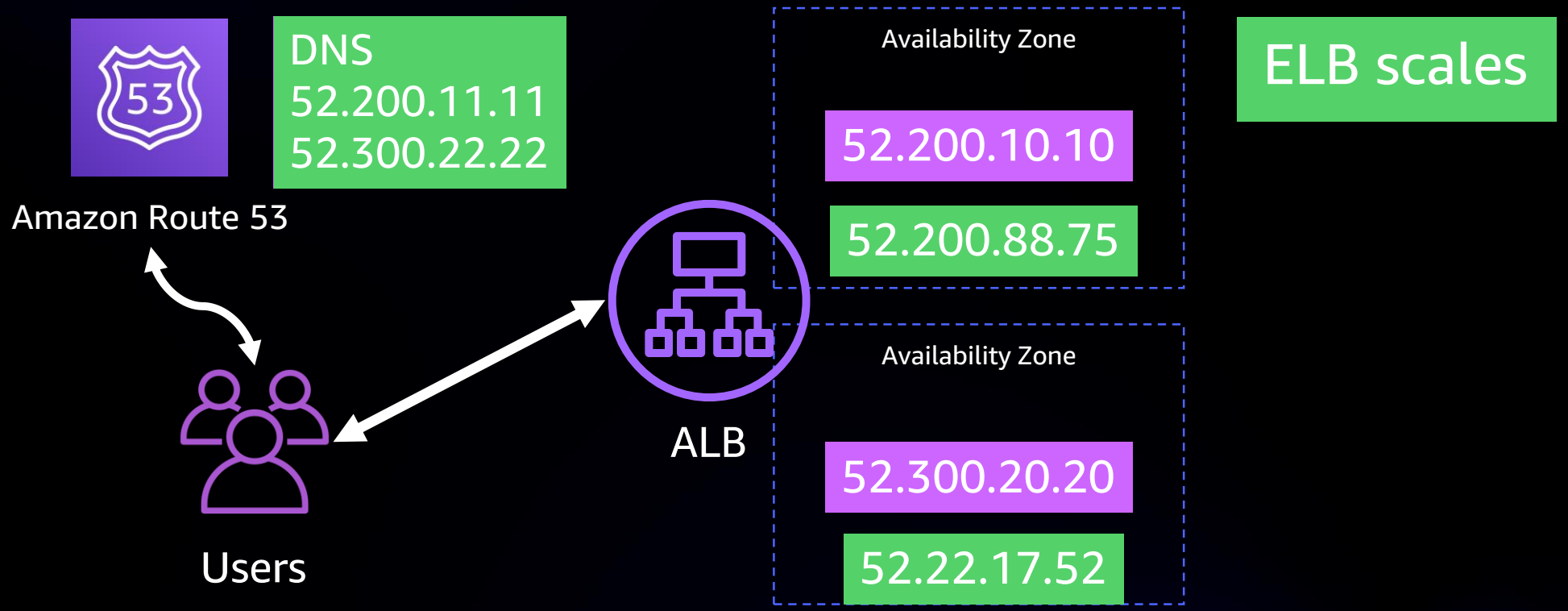
# ELB and Amazon Route 53

- Use DNS records with multiple ALBs
- Use Route 53 and select evaluate target health
- Set TTL to 60 to match




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB DNS – healthy and scaled endpoints




aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Client connectivity best practices

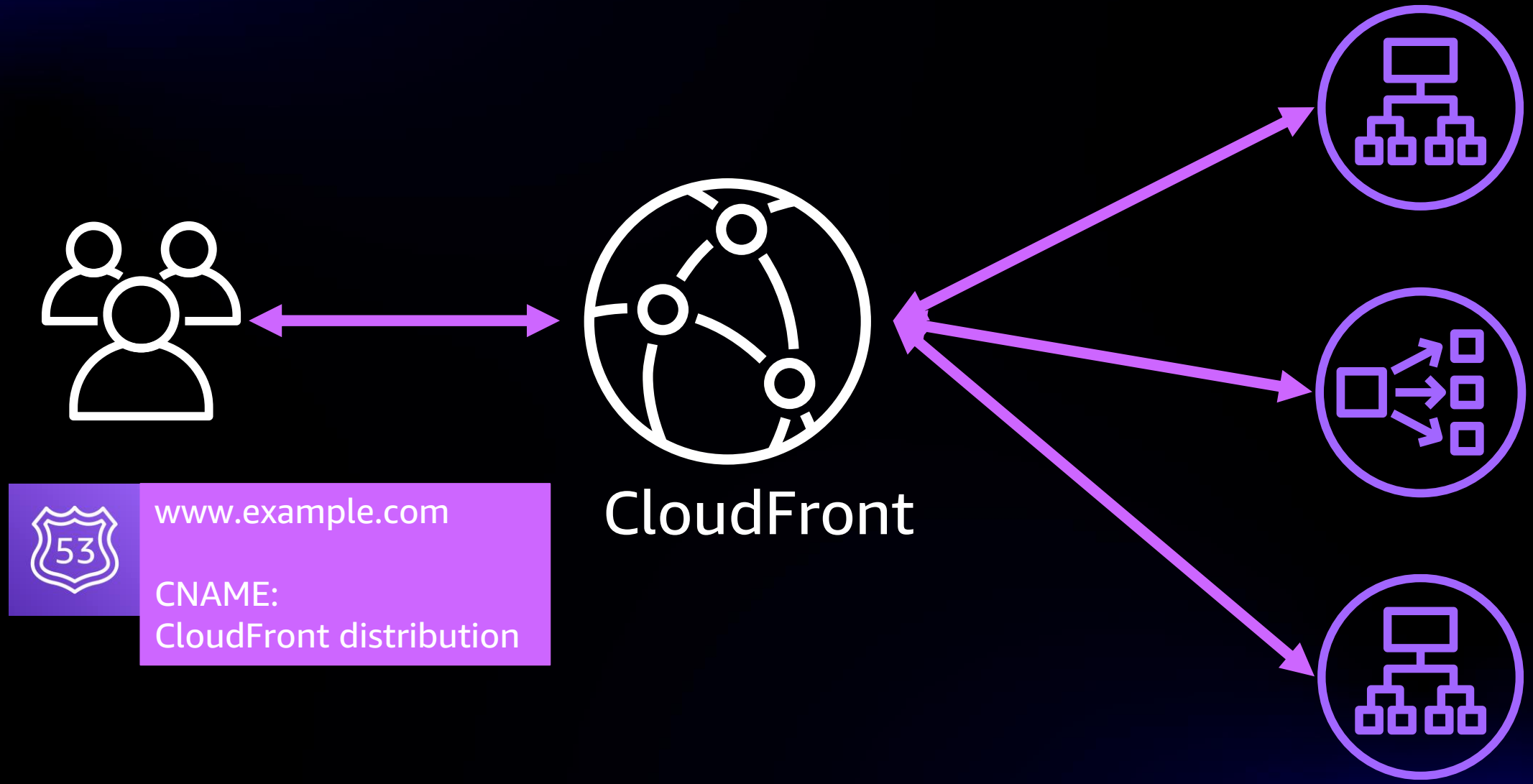
1. Use & respect DNS
2. Refresh DNS on reconnect
3. Use exponential backoff with jitter




aws

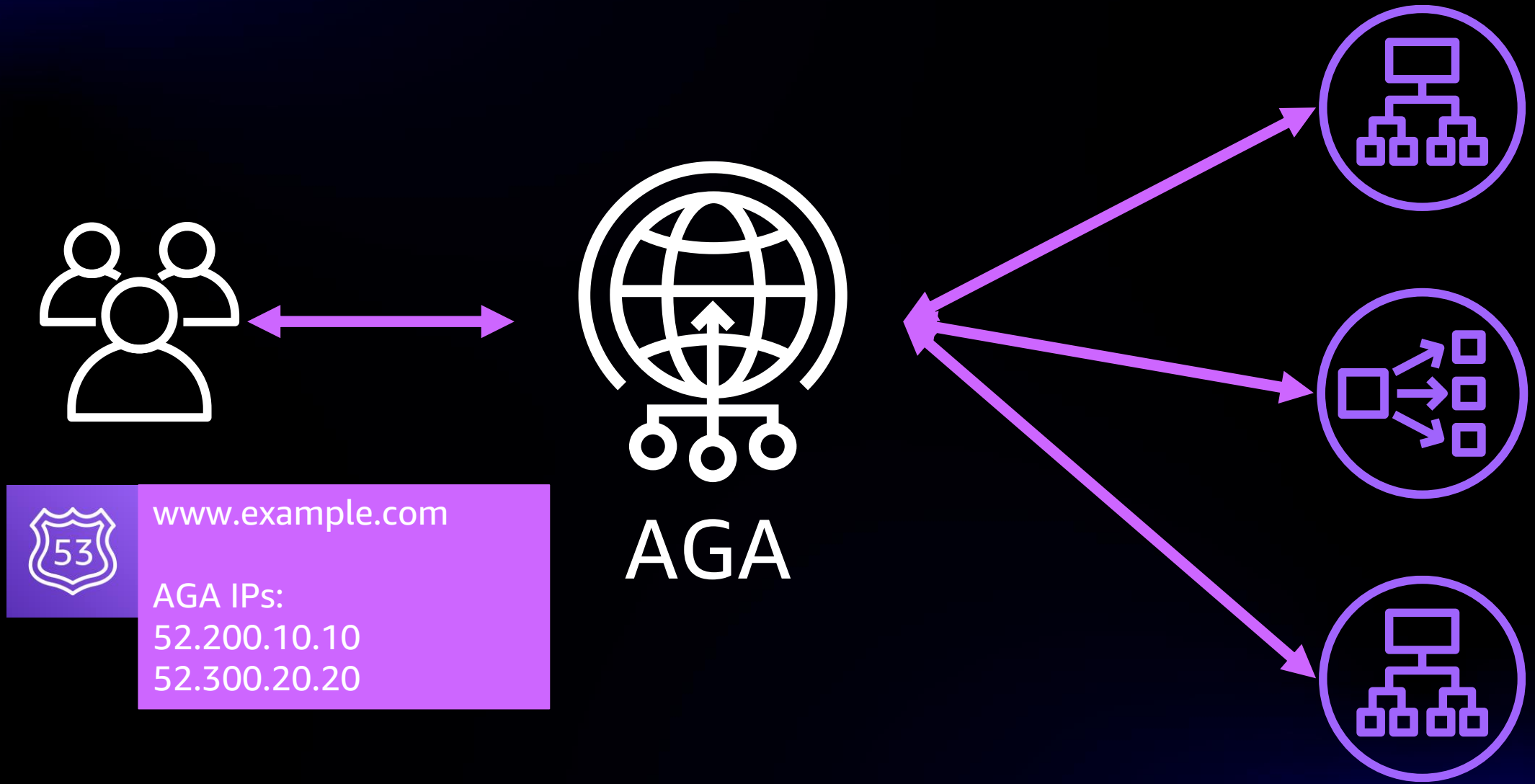
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB with Amazon CloudFront



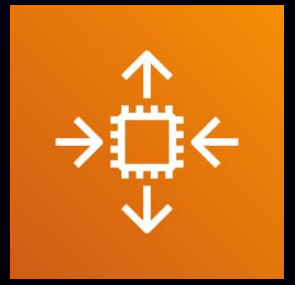

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELBs with AWS Global Accelerator




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Services behind ELB improving availability



AWS Auto Scaling



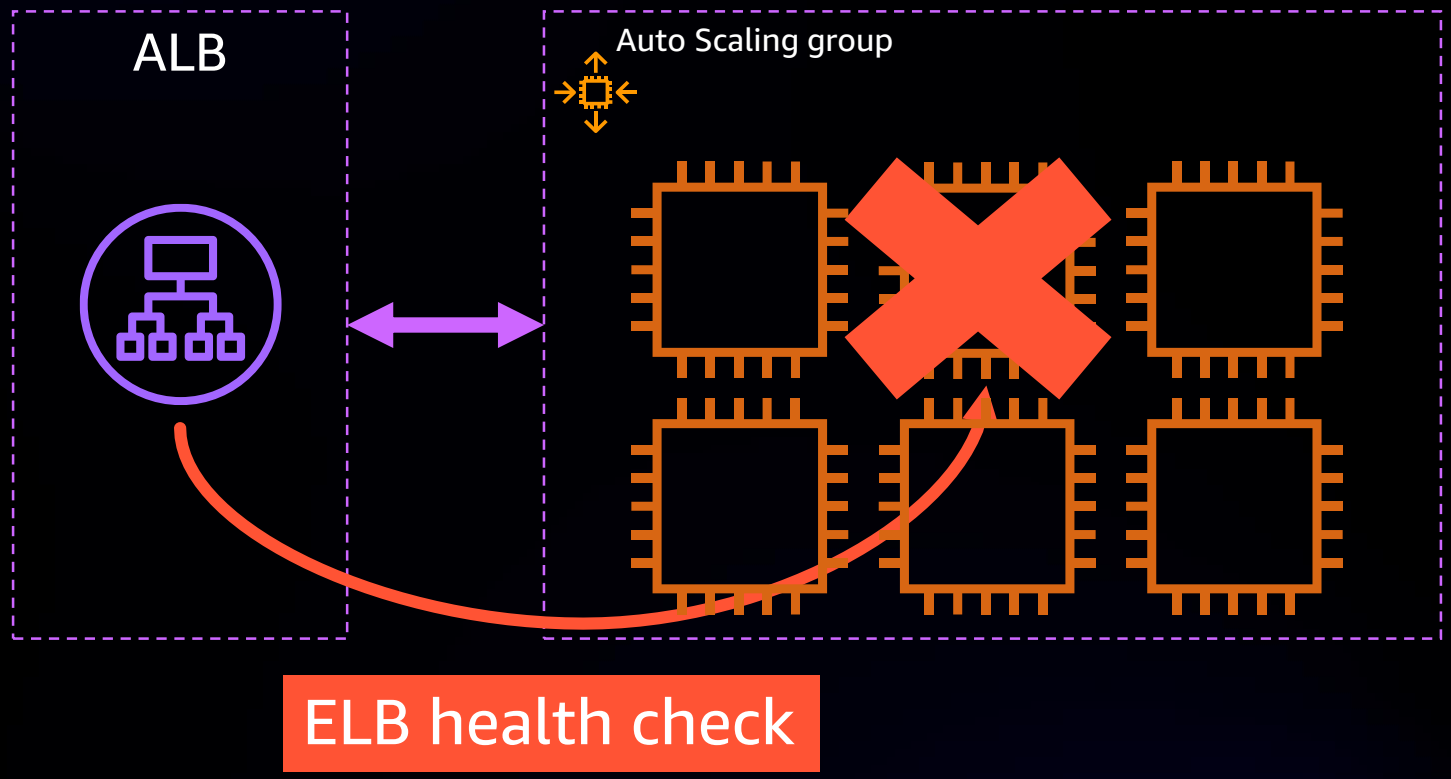
ECS / EKS



Lambda, Fargate

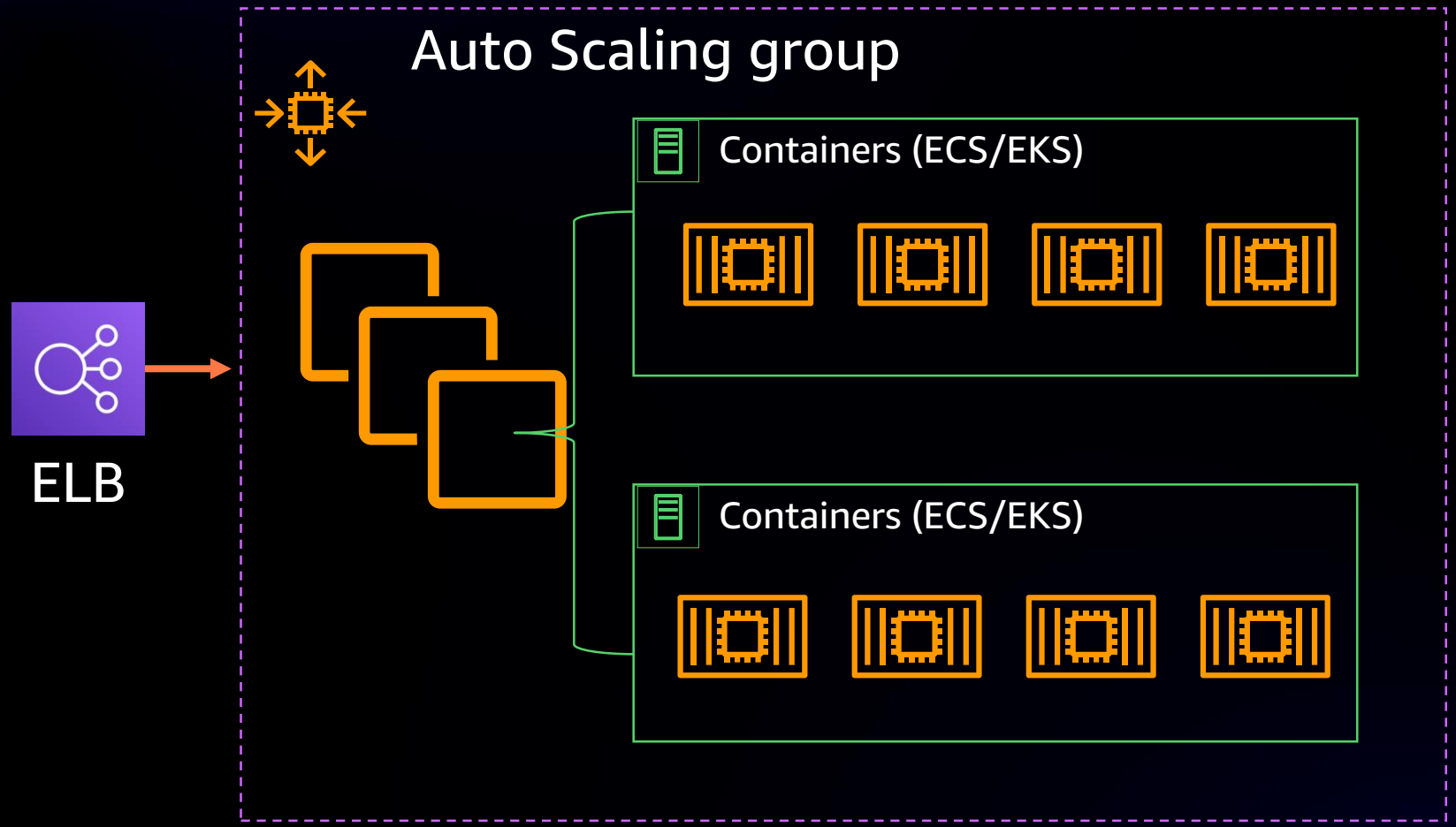
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELBs with AWS Auto Scaling




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Scaling containers – ASG + containers

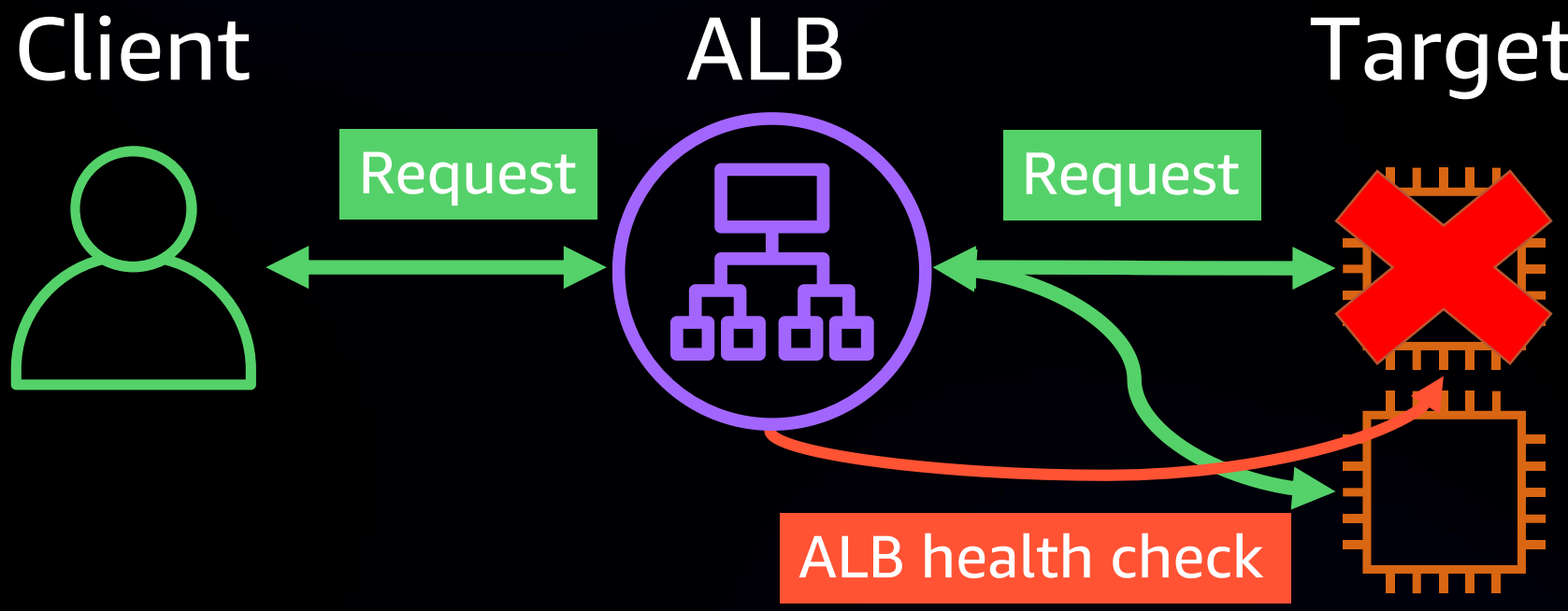



Auto Scaling
Launch instance
= minutes

ECS/EKS
Lunch container
= seconds

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ALB target failure lifecycle




Client connected to ALB IP

Each request routed to target

Target fails ALB health check

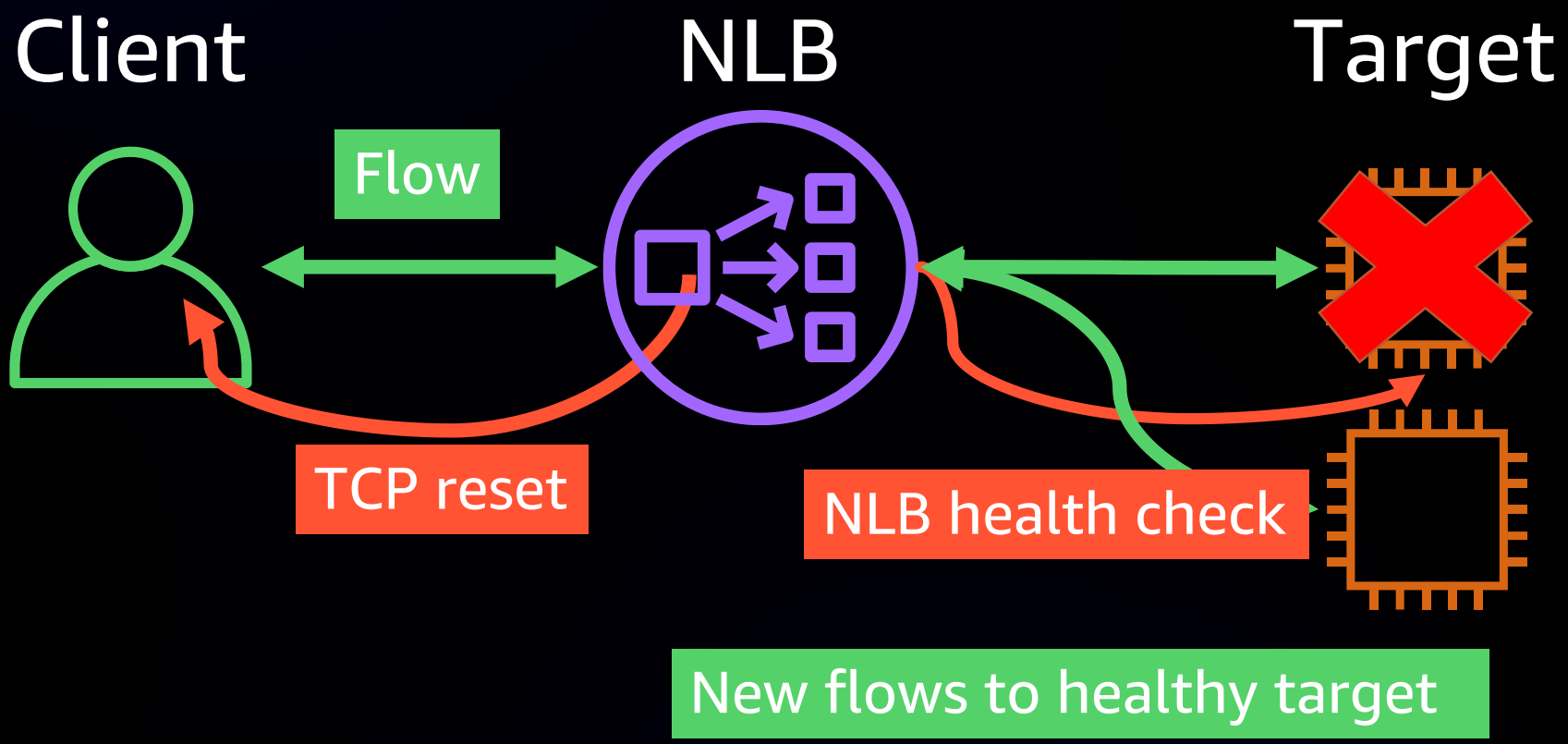
In-flight requests allowed to complete

Future requests routed to healthy targets

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# NLB target failure lifecycle

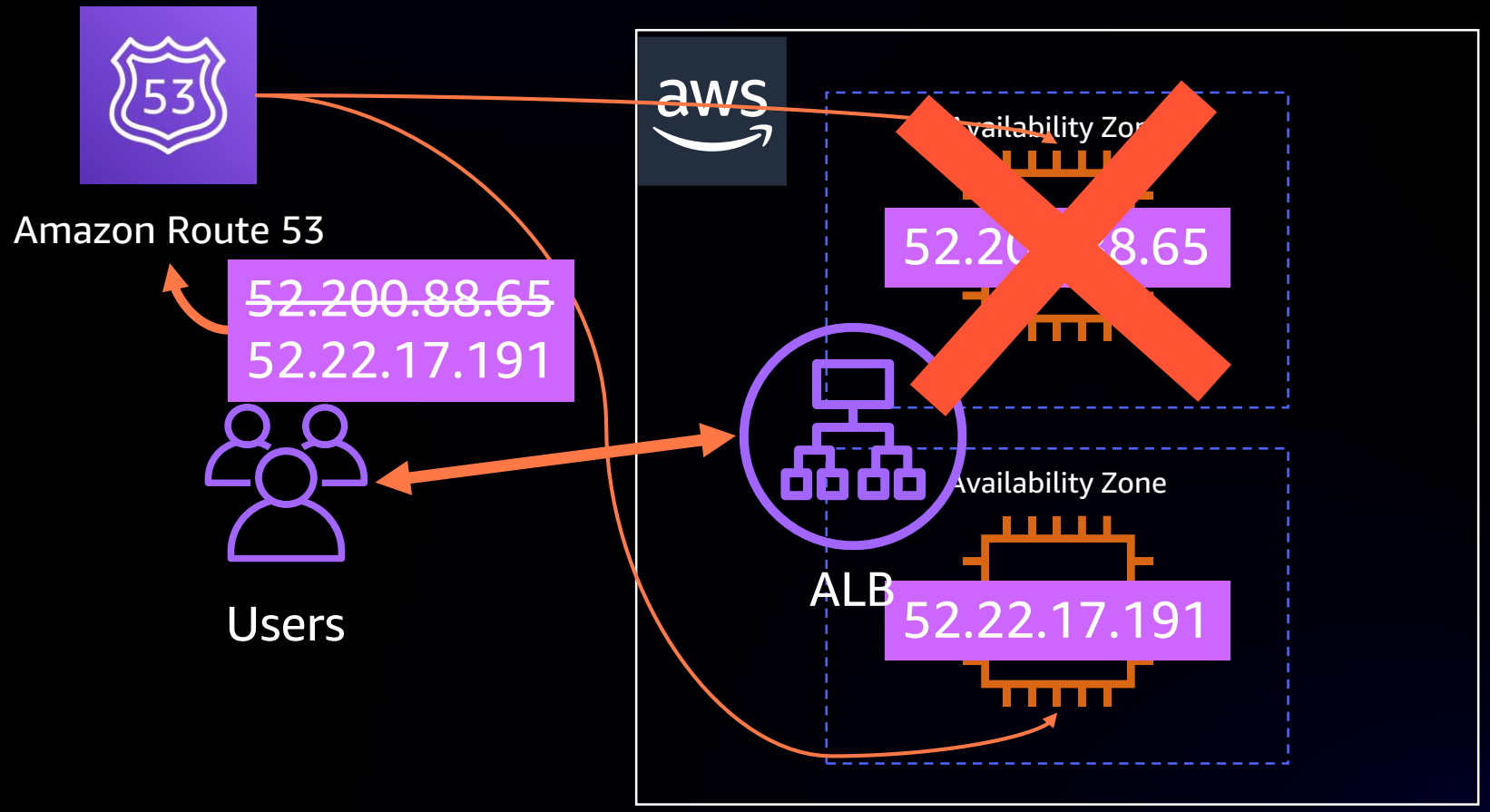



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB health

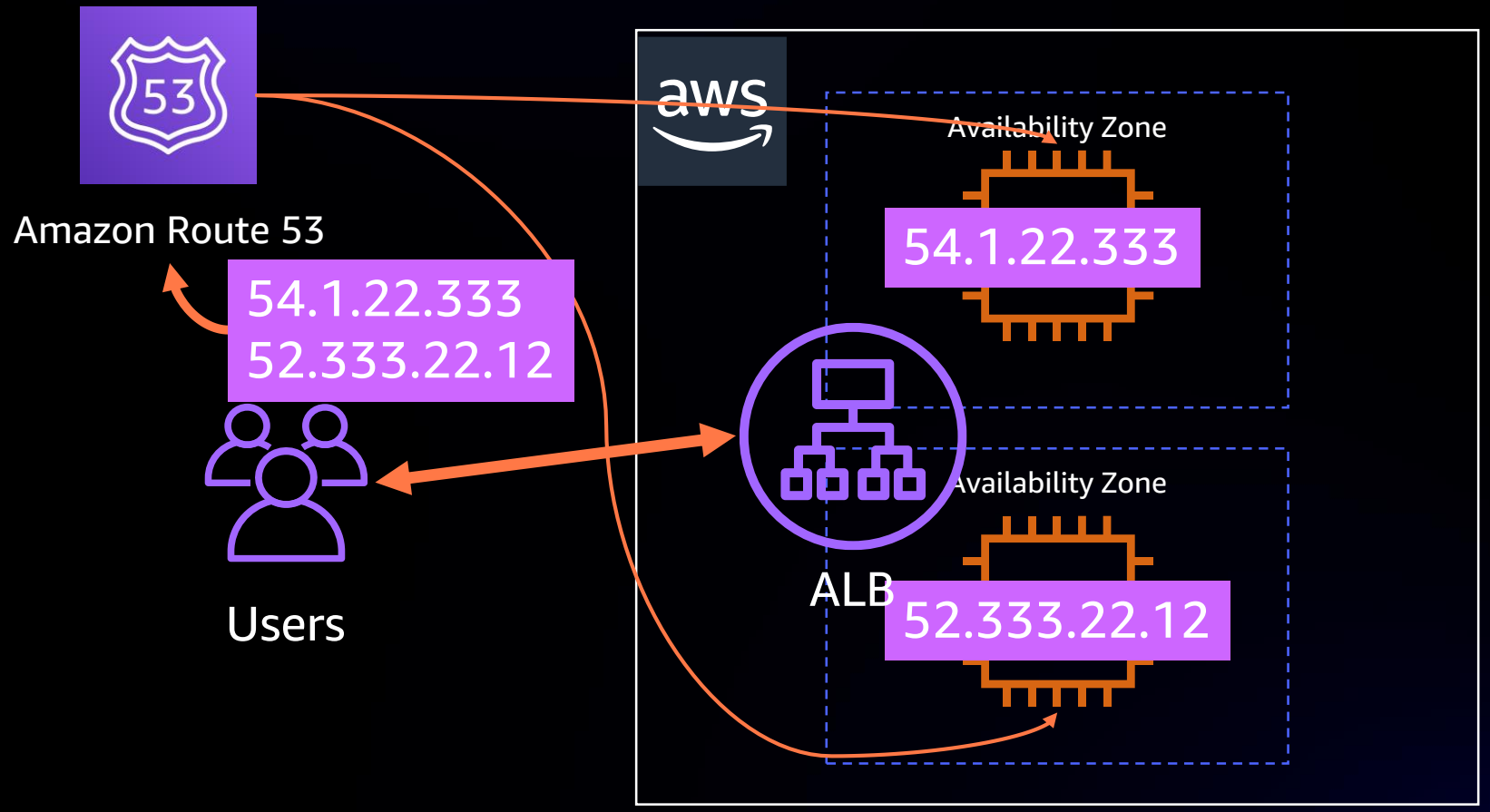
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB IPs are health-checked by Route 53



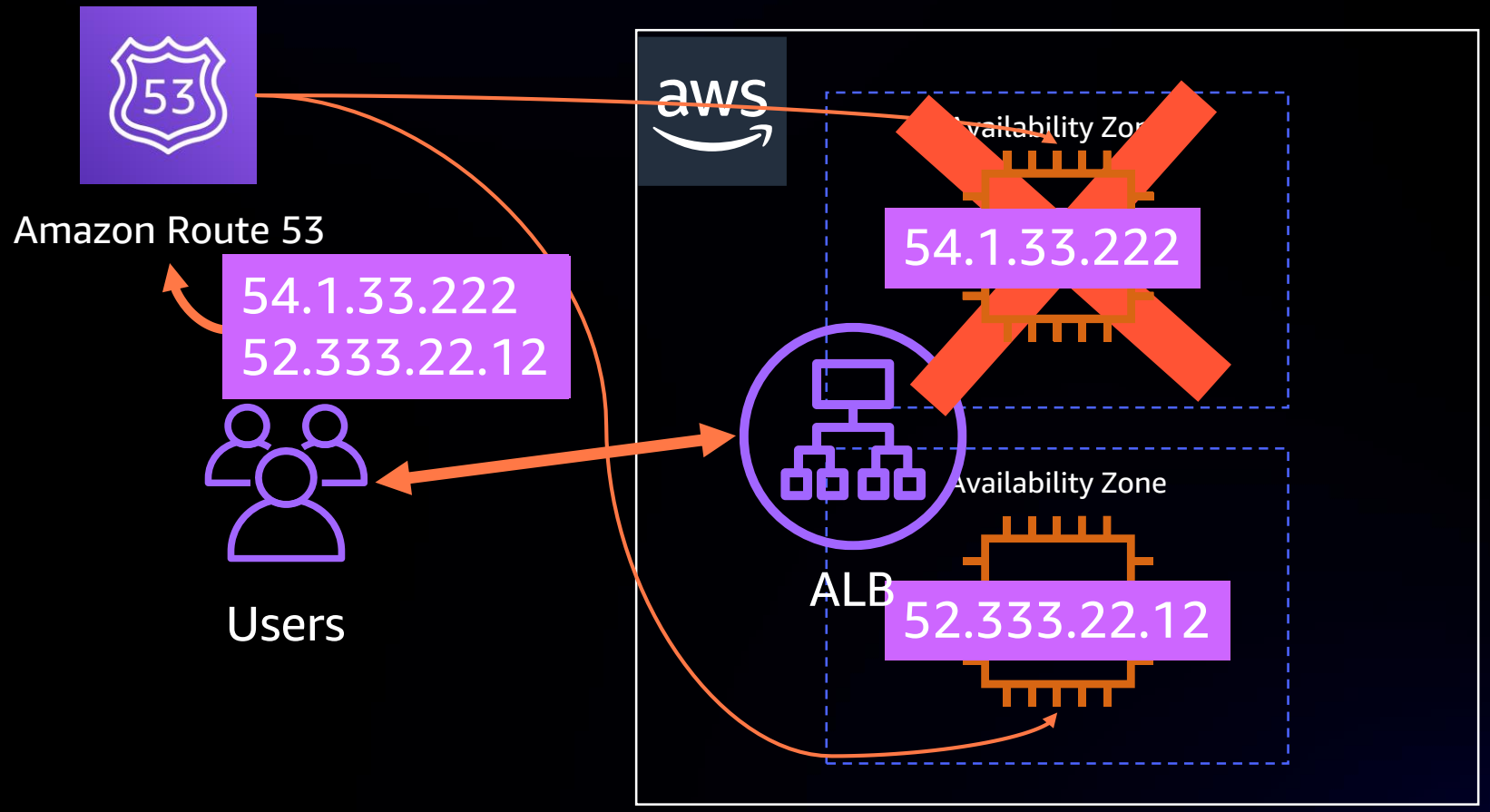

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB availability systems – health checking




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# ELB availability systems – health checking




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Serverless

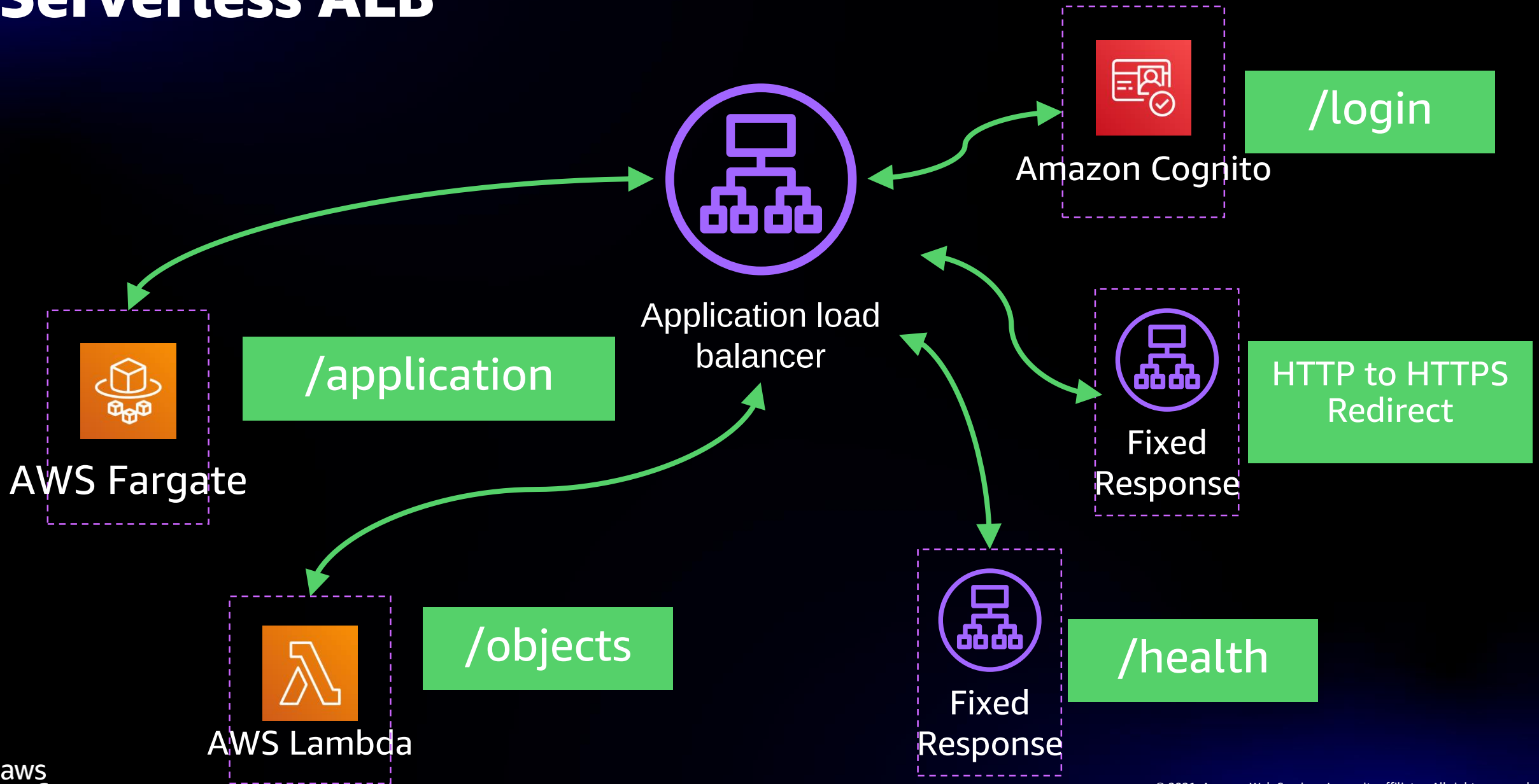
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Serverless ALB

| | Serverless | Servers |
|---|---|---|
| Non-Fargate containers | | ✓ |
| Custom application servers | | ✓ |
| Simple HTML fixed responses | ✓ | |
| Fargate containers | ✓ | |
| HTTP to HTTPS redirects | ✓ | |
| Lambda | ✓ | |
| AWS WAF | ✓ | |
| Authentication | ✓ | |

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Serverless ALB

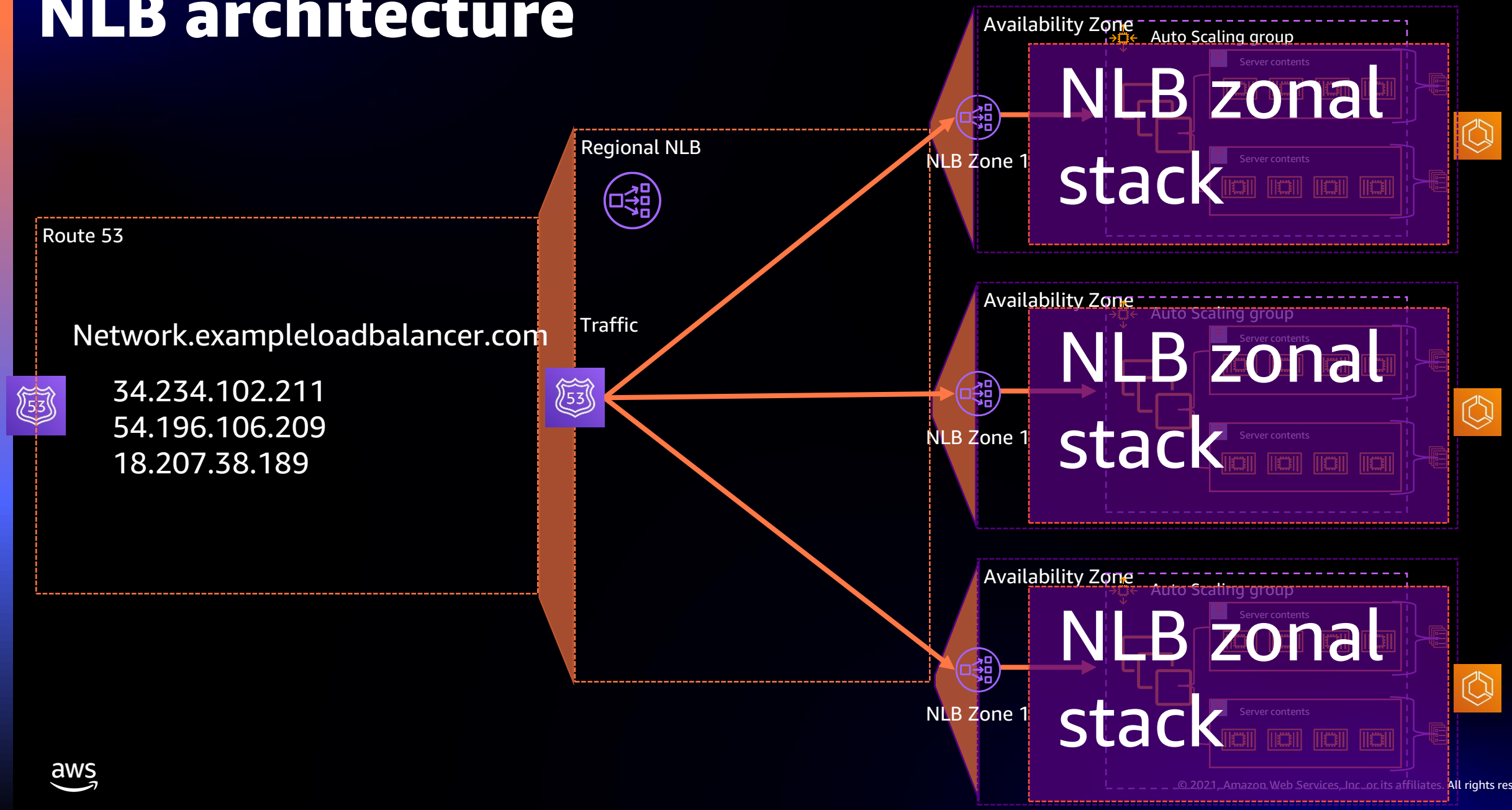



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Network Load Balancer DNS

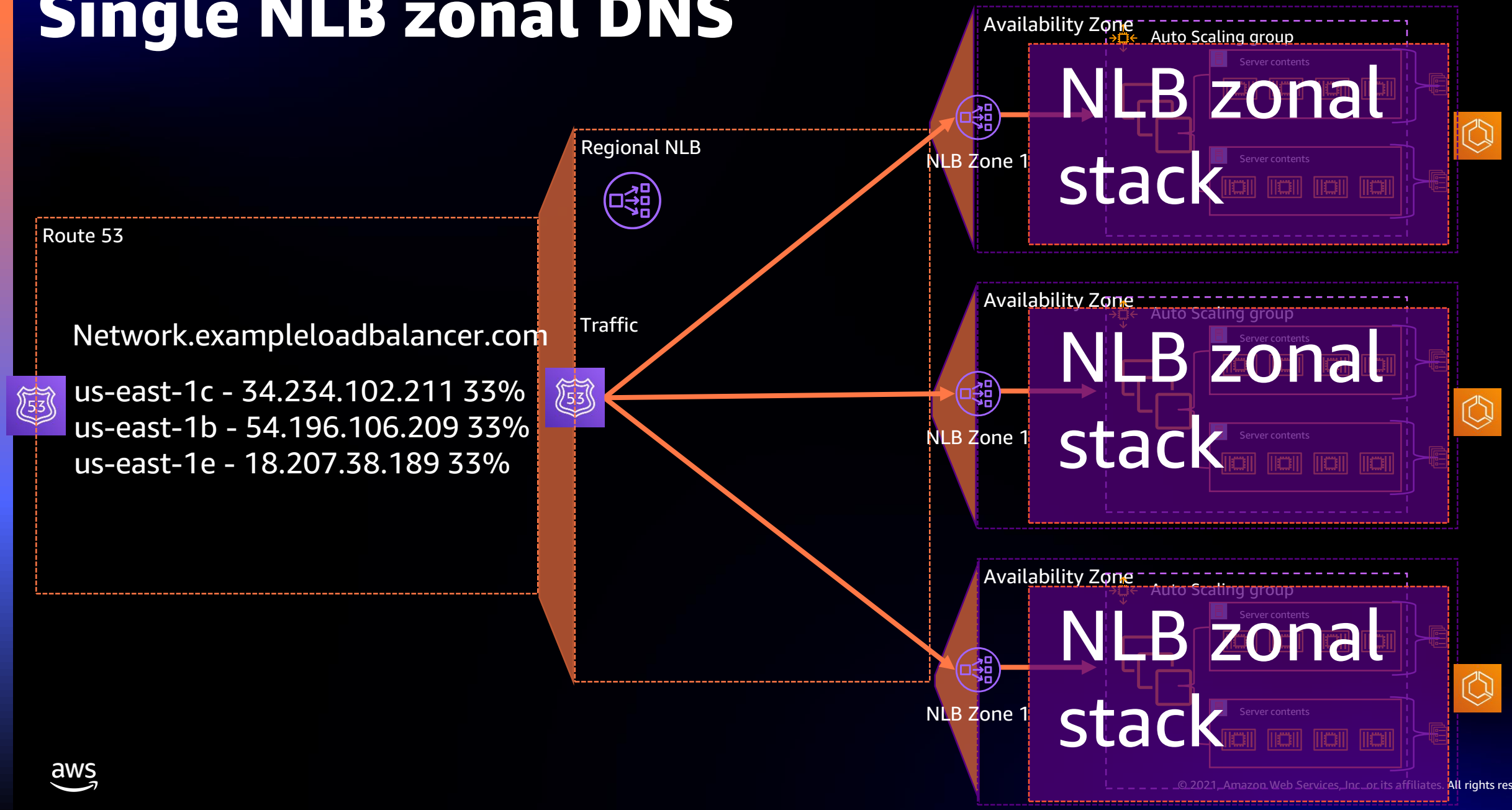
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# NLB architecture



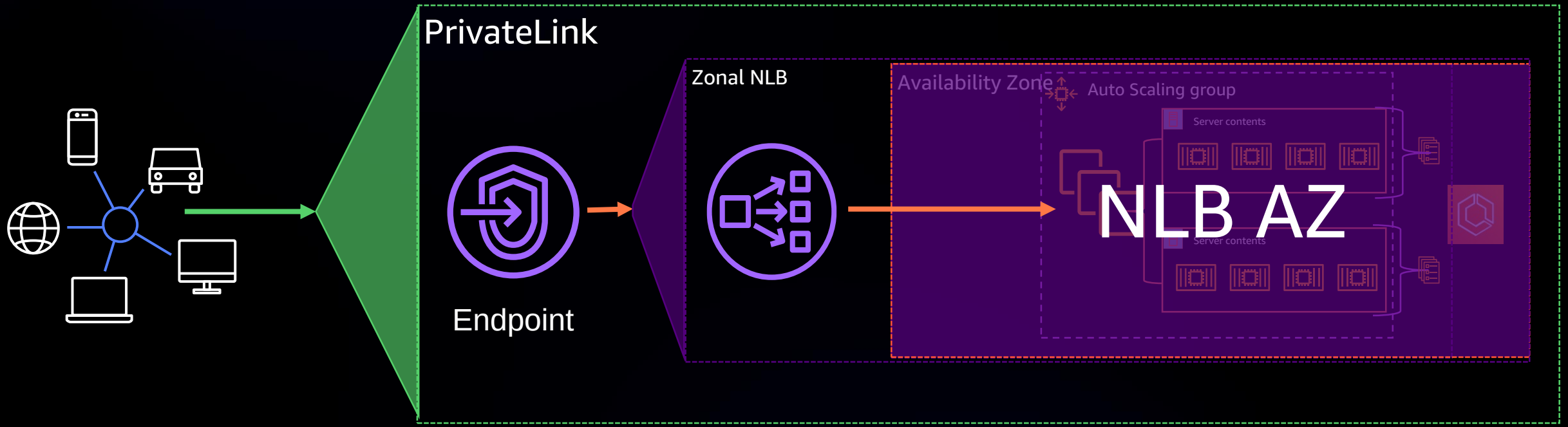

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Single NLB zonal DNS




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# PrivateLink with NLBs




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Operations / Planning for production

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

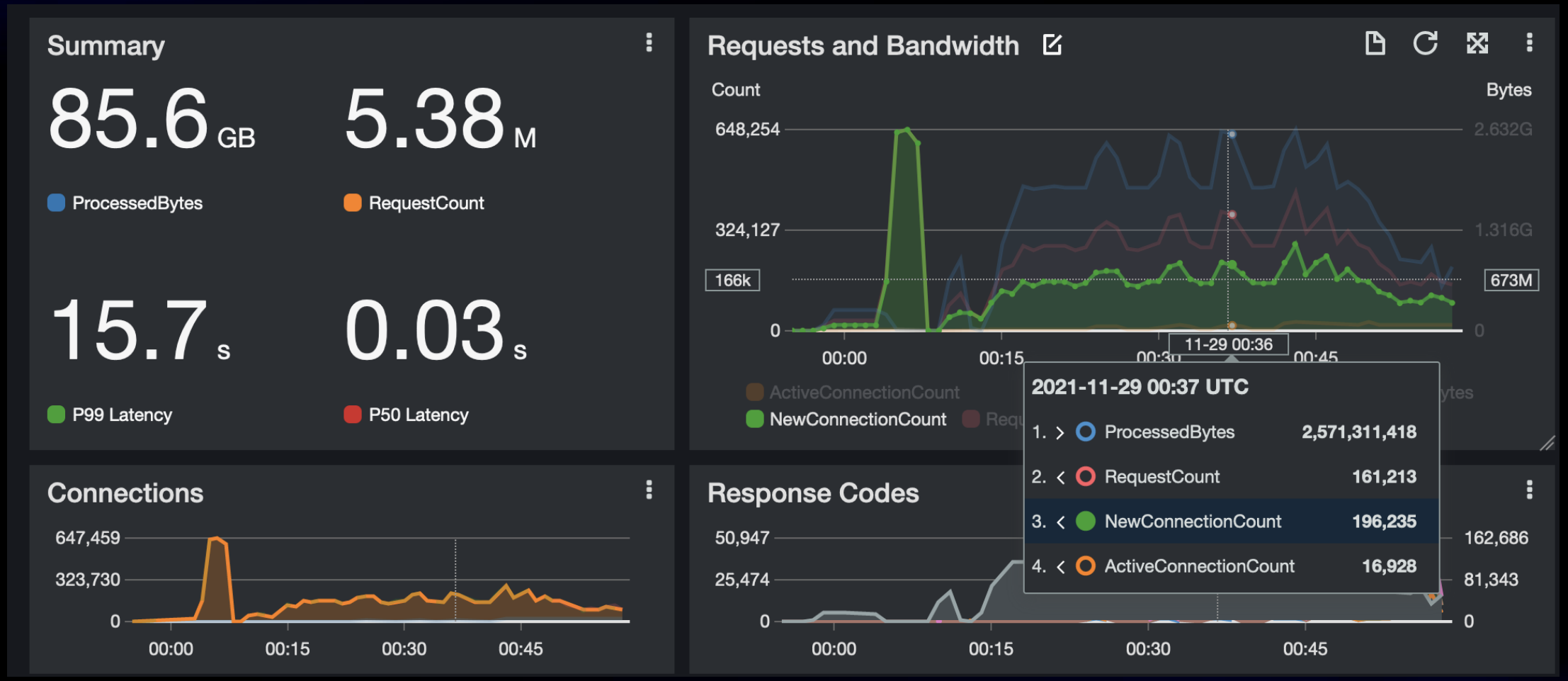
# Operational visibility



- Metrics & dashboards
- Alarms
- Operational hygiene

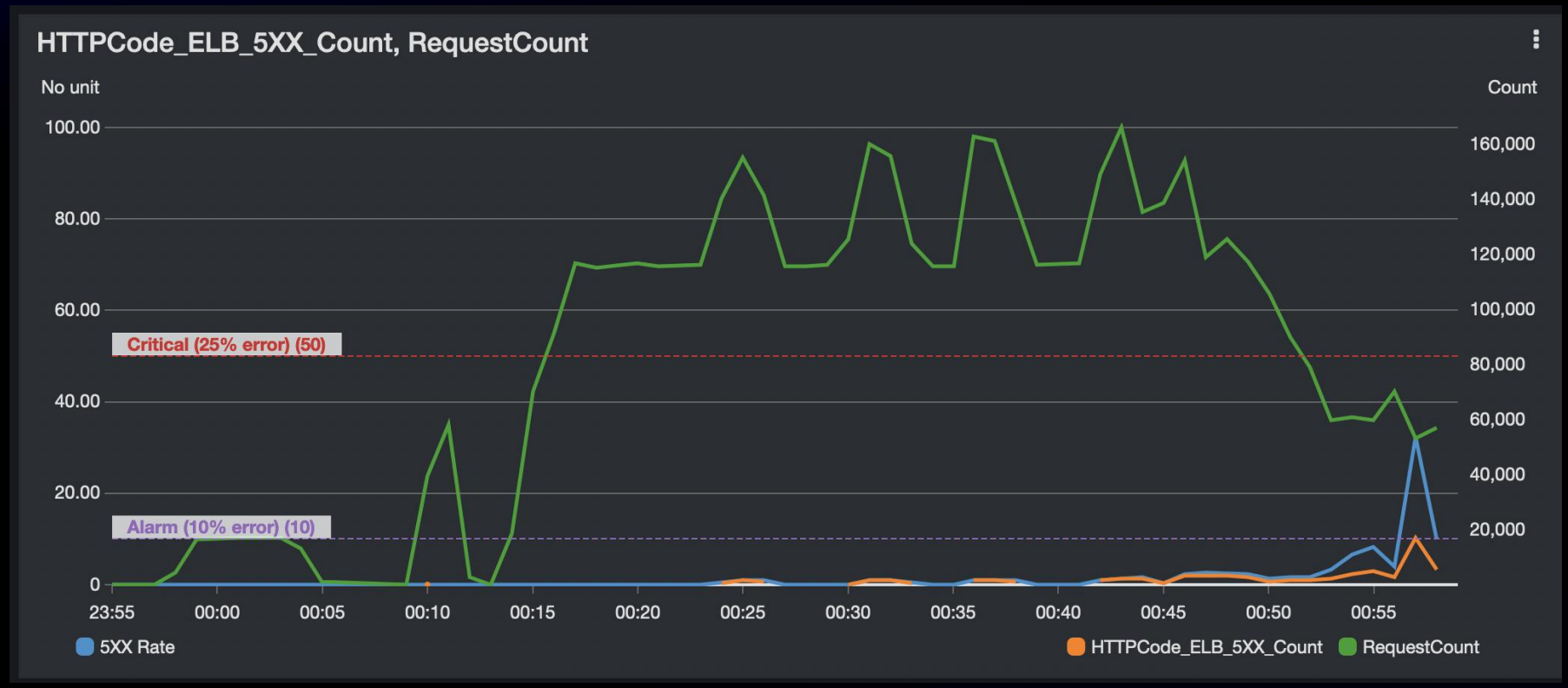
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Metrics and dashboards




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Alerting & operational paranoia



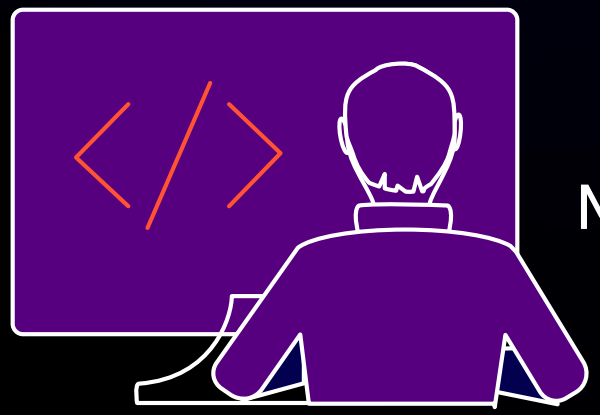



Alarms on key metrics
Thresholds > 1
Multiple periods
Multiple alarms w/different thresholds



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Operational priorities



Mitigating impact – always priority number 1



Finding root cause – not the most important thing to do

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Thank you!

Jon Zobrist
zob@amazon.com

Priya Patel
ppripa@amazon.com

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.