

re:Invent 2023 re:cap Session 3 : Analytics, Database, Storage and Architecture

Rohit Singh
Solutions Architect

AWS RE:INVENT

re:Cap



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Amazon S3

Express One Zone

The fastest cloud object storage class
for performance-critical applications

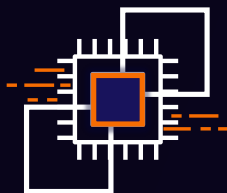
Millions of access requests per
minute with single-digit millisecond
latency

New bucket type, pinned to the
Availability Zone of your choice

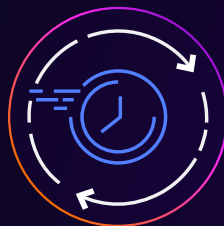


Available now

Amazon S3 Express One Zone storage class



**For compute-
intensive workloads**



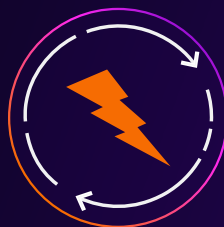
Single-digit millisecond latency

One zone S3 storage class that delivers the fastest data access speed and highest performance of any cloud object storage for customers' most latency-sensitive applications



Most frequently accessed data

Designed for **request intensive applications** – ML training and inference, interactive analytics, media content creation



10x faster + 50% lower request costs

Data access speeds up to **10x faster**, and request costs up to **50% lower** than S3 Standard. Fully elastic with no storage provisioning



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The newest Amazon S3 storage class

New



S3 Express One Zone

Single Availability Zone

Re-creatable, latency sensitive, frequently accessed data

- Single-digit millisecond access
- Uses directory buckets
- Co-locate storage and compute resources to further reduce latency



S3 Intelligent-Tiering

Changing access patterns

- Milliseconds access
- No retrieval charge
- Archive Instant Access tier



S3 Standard

Frequently accessed data

- Milliseconds access
- No retrieval charge



S3 Standard-IA

Infrequently accessed data

- Milliseconds access
- Per-GB retrieval charge



S3 Glacier Instant Retrieval

Rarely accessed data

- Milliseconds access
- Per-GB retrieval charge



S3 Glacier Flexible Retrieval (formerly S3 Glacier)

Archive data

- Retrieval options from minutes to hours
- Free bulk retrievals



S3 Glacier Deep Archive

Long term archive data

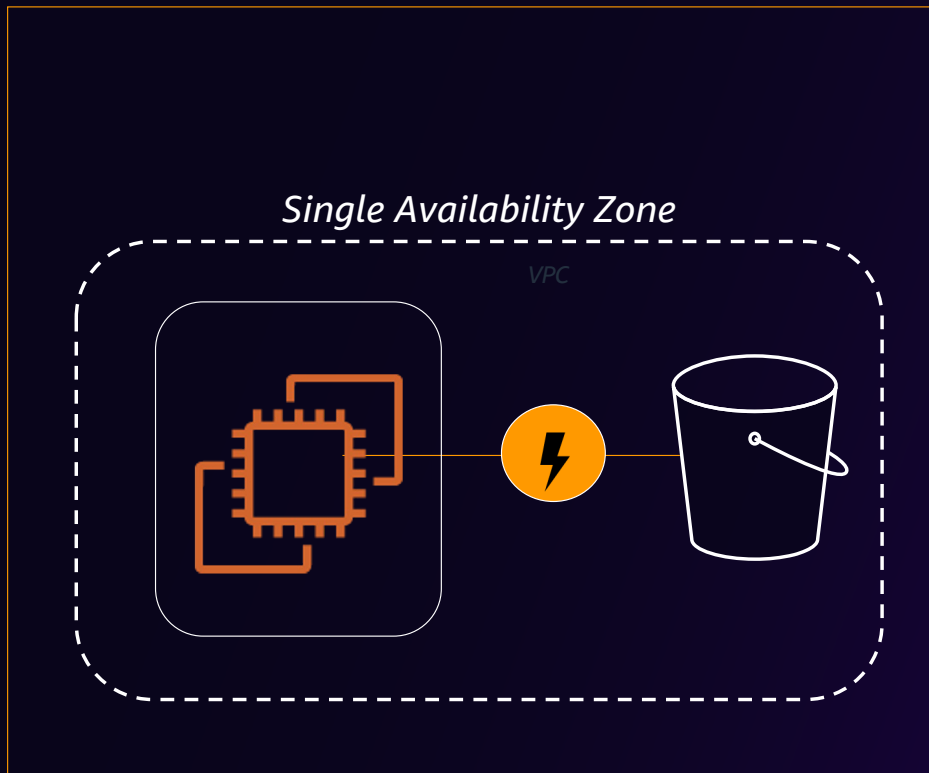
- Retrieval in hours

AWS Region \geq 3 Availability Zones

Purpose built for fastest data processing for performance critical applications

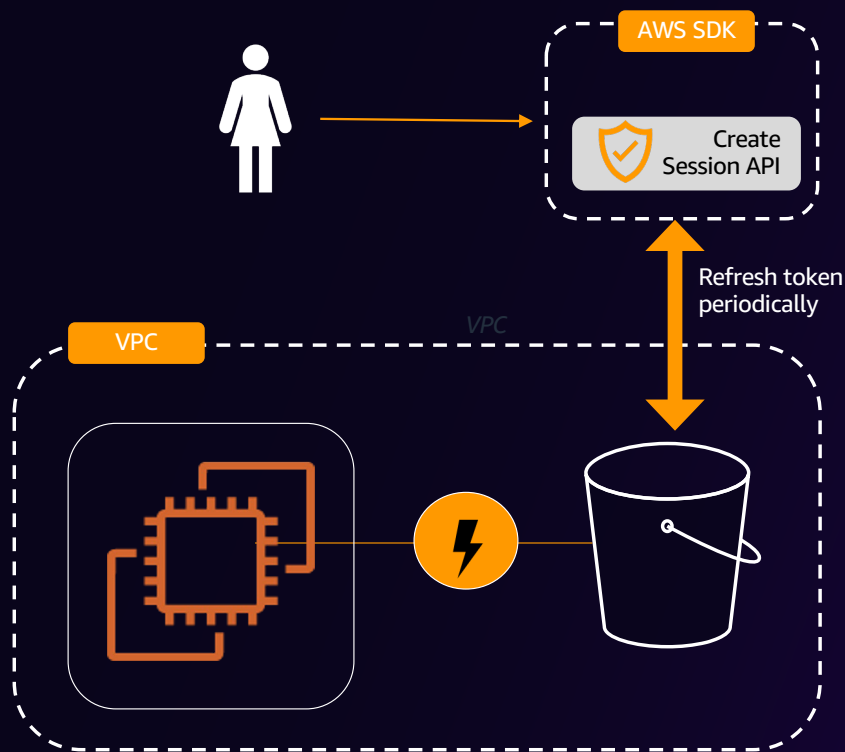


Single-Availability Zone storage class enables low latency



- Single-Availability Zone storage class enables **co-location of storage and compute resources** for highest levels of performance
- **Single-digit millisecond** first-byte P99 latency

A faster authentication model to support lower latency



- Call **CreateSession API** to initialize session between client and S3
- CreateSession API returns a session **token**
- **Perform data operations** with token for authorization
- **Refresh token** after set period of time. SDKs perform operation automatically

Performance gains for AWS services

REQUEST-HEAVY WORKLOADS SEE THE MOST MEANINGFUL BENEFITS

Amazon EMR on EC2 (Spark)



Up to **4.0x** performance
Improvement

Amazon Athena (SQL)



Up to **2.1x** performance
Improvement

Amazon SageMaker (File Mode)



Up to **1.64x** performance
Improvement



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Mountpoint for **Amazon S3**

Container Storage Interface (CSI)
driver for Kubernetes now available

Added support for the new S3
Express One Zone storage class



Demo



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Amazon EFS

New Archive storage class optimized for long-lived file data that is accessed a few times a year or less

Infrequent Access storage class price reduction by 36%

Now supports up to 250,000 read and up to 50,000 write IOPS per file system





AWS Backup

Now supports automatic archival
of EBS Snapshots into the EBS
Snapshot Archive

Restore testing helps perform
automated and periodic restore
tests of backed up resources



GENERALLY AVAILABLE

AWS Secrets Manager

batch-retrieval of secrets

BatchGetSecretValue identifies and retrieves secret groups with a single API call

Use with list of secret names or ARNs, along with filter criteria, such as tags



GENERALLY AVAILABLE

Amazon Inspector

Lambda code scanning
with generative AI

Assess function code for security issues like injection risks, data leaks, and weak or missing encryption

Actionable security findings including affected code snippets and remediation suggestions



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon Inspector container image security

Scans container images for security issues from within the CI/CD pipeline

Works with CI/CD tools hosted in AWS, on-premises, or hybrid clouds



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ECS: GA / EC2: IN PREVIEW

AWS Guard Duty

ECS and EC2 runtime monitoring

[GuardDuty](#) > Runtime Monitoring

Runtime Monitoring [Info](#)

Monitor runtime activity to detect threats to your compute workloads on Amazon ECS (including AWS Fargate), Amazon EKS, and Amazon EC2. To use need to manage a GuardDuty security agent that will ingest runtime activity for your resource.

Configuration | **Runtime coverage**

Coverage statistics [Info](#)

Healthy EKS clusters

0/0 (100%)

Healthy ECS clusters

1/1 (100.0%)

Healthy EC2 instances

0/0 (100%)

EKS clusters runtime coverage

ECS clusters runtime coverage

EC2 instance runtime coverage

Clusters list (1)

Cluster name	Account ID	Agent management type (Fargate)	Coverage status
SpaceTwin-Infra-SpaceTwinCluster3D77526A-5D6rTXjgdSb2	636183562745	Auto-managed	Healthy



GENERALLY AVAILABLE

Amazon CloudWatch

Pattern Analysis and Anomaly Detection

Patterns view visualizes recurring patterns while querying your logs

Compare mode helps you quickly find changes over time

Anomaly detection evaluates incoming logs against historical baselines



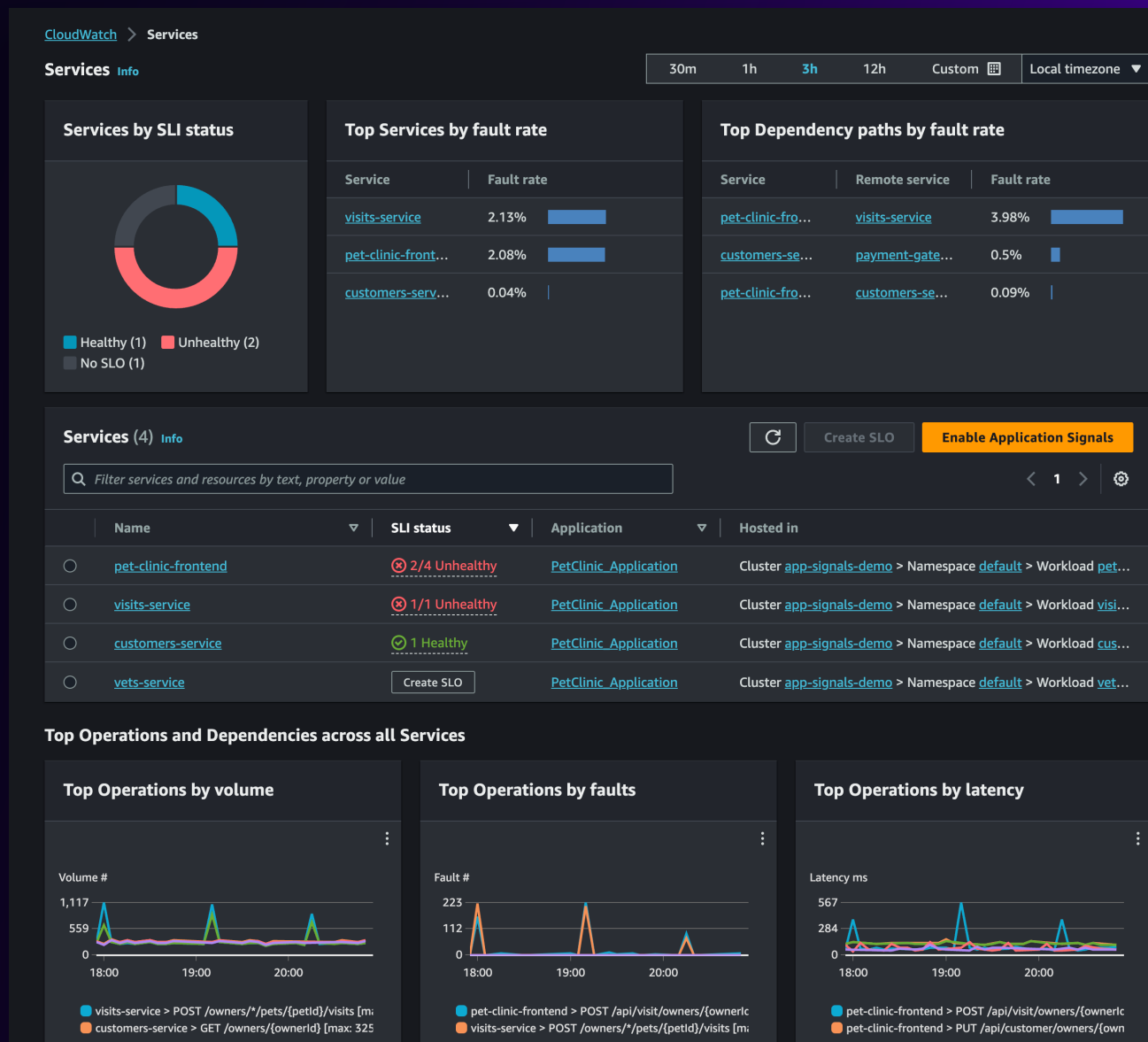
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

IN PREVIEW

Amazon CloudWatch Application Signals



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



GENERALLY AVAILABLE

Amazon CloudWatch

Infrequent Access class

A new log class for cost-effective log consolidation

Best suited for logs that require infrequent querying, e.g. for forensic analysis



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

AWS Systems Manager Automation low-code Runbook design

Drag and drop automation actions and AWS APIs and connect them as steps

Features a guided flow, best practice recommendations, and error highlighting

Integrates with Amazon CodeGuru Security for embedded Python scripts



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



IN PREVIEW

Amazon Aurora Limitless Database

Managed horizontal scale-out beyond
the limits of a single database instance



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon ElastiCache Serverless

Serverless Amazon ElastiCache
configuration for Redis and
Memcached



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon ElastiCache Serverless

Instantly scale without managing
infrastructure



Create a highly available cache in < 1 minute without managing infrastructure



Instantly scales in response to changes in application access patterns



99.99% high availability for both Redis and Memcached engines



Pay for what you use economics



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon ElastiCache Serverless features & benefits



**Create a cache in
under a minute**



**No capacity
management**



**700 microseconds at
p50, 1.2ms at p99**



**Up to 5 TB
of storage**



**Pay-per-use
pricing**



**99.99%
availability SLA**



**Single
endpoint experience**



**PCI-DSS, SOC, HIPAA
compliances**



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon Q generative SQL in **Amazon Redshift**

Simplify query authoring by expressing
queries in natural language



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Zero-ETL integrations with **Amazon Redshift**

**Amazon Aurora
PostgreSQL**

**Amazon RDS
for MySQL**

**Amazon
DynamoDB**

AVAILABLE IN PREVIEW



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Zero-ETL integrations

**Amazon
OpenSearch**
with
Amazon S3

**Amazon
DynamoDB**
with
Amazon OpenSearch

Amazon Athena
with
**AWS CloudTrail
DataLake**

AVAILABLE IN PREVIEW



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Fault Injection Service

A FULLY MANAGED SERVICE FOR RUNNING FAULT INJECTION EXPERIMENTS.

RESILIENCE LIFECYCLE: TEST & EVALUATE

Easy to
get started



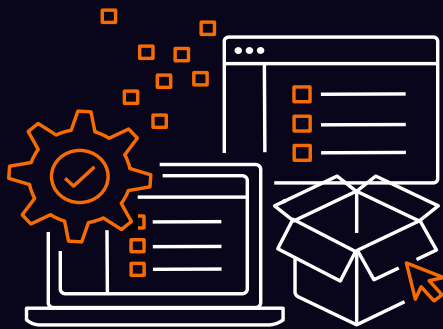
Real-world
conditions



Safeguards



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Easy to
Get Started



Verify Resilience using FIS through its native integrations with AWS Services.



Use the AWS Management Console or the AWS CLI



Use pre-existing experiment templates and get started in minutes





Real-World Conditions



Experiments can include multiple fault actions running in sequence or in parallel



Target all levels of the system (API, host, infrastructure, network, etc.)



Unique actions injected at the service control plane level (e.g., api-internal-error)

Scenario Library

- List of scenarios
- Detailed description
- Create Experiment Templates
- Bulk edit parameters

The screenshot displays the AWS FIS Scenario Library interface. At the top, the AWS logo and navigation bar are visible, including the 'Services' menu, a search bar, and the current region 'N. Virginia'. The main heading is 'Scenario Library (1/12)' with an 'Info' link. Below this is a search bar with the placeholder text 'Find by resource types, services, or actions'. Three scenario cards are shown: 'EC2 Stress - CPU' (selected), 'EC2 Stress - Disk', and 'EC2 Stress - Instance Failure'. Each card provides a brief description and target types. The 'EC2 Stress - CPU' card is expanded, showing tabs for 'Overview', 'Content', and 'Details'. The 'Overview' tab is active, displaying a description: 'Explore impact of increased CPU on your EC2 based application.' and a 'Prerequisites' section: 'EC2 instances: you will need one or more EC2 instances to target'.

Scenario Library

- List of scenarios
- Detailed description
- Create Experiment Templates
- Bulk edit parameters

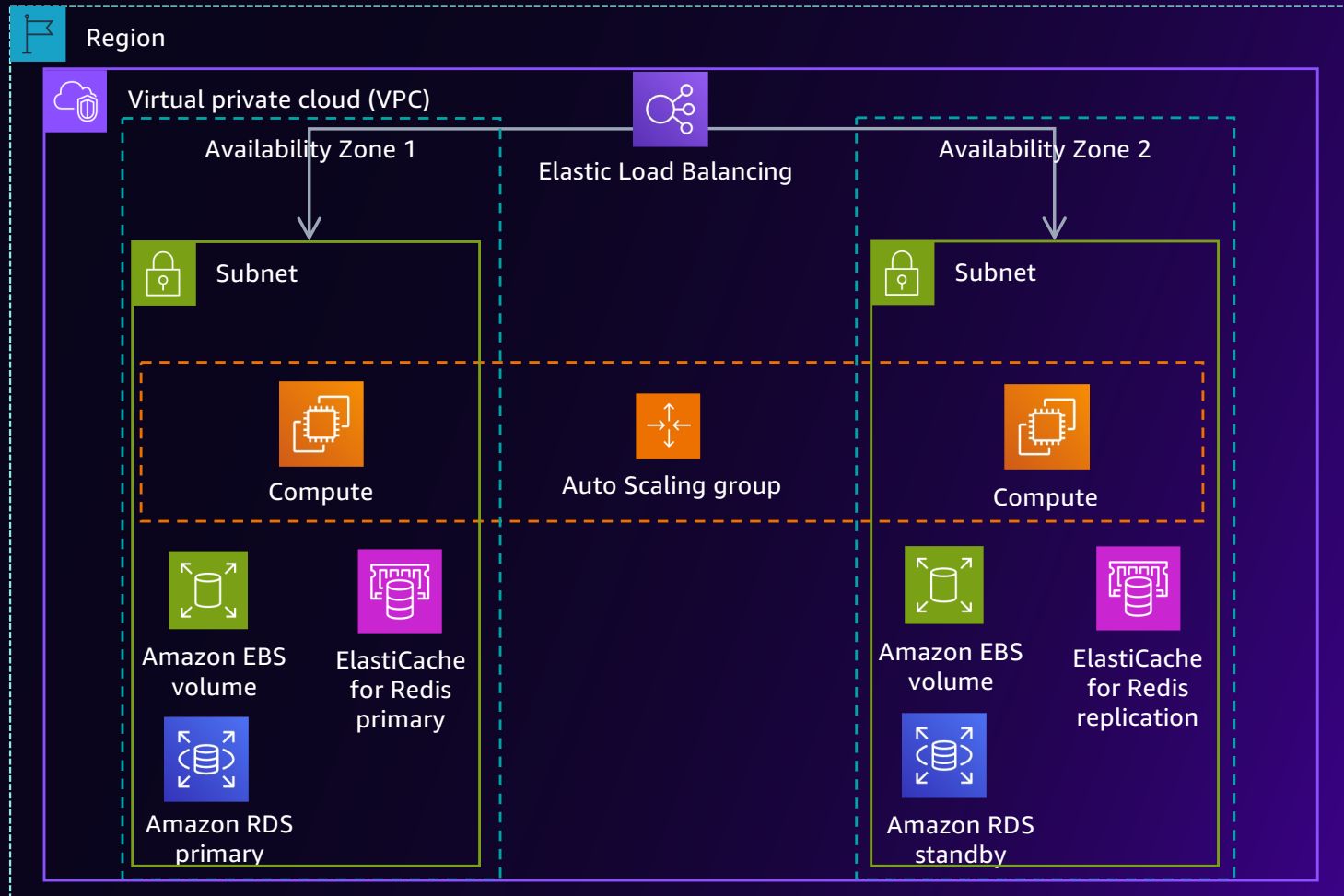
The screenshot shows the AWS FIS console interface for creating an experiment template. The breadcrumb navigation is 'AWS FIS > Experiment templates > Create experiment template'. The main heading is 'Create experiment template' with an 'Info' link. The 'Description and name' section contains a 'Description' field with the text 'Inject increasing CPU load on one ore more instances, targeting based on instance tags' and a note that the description must have 1 to 512 characters. Below it is a 'Name - optional' field with the text 'EC2 Stress - CPU' and a note that the name must have 1 to 256 characters. A toggle for 'New targets and actions experience' is set to 'On'. A blue information box states: 'Edit shared parameters of the EC2 Stress - CPU scenario. Changes made to the experiment template directly or through use of the shared parameter editing dialog will disable shared parameter editing features. [Learn more about scenarios](#)'. An 'Edit shared parameters' button is located to the right. The 'Actions (3)' section lists two actions: 'Ec2StressCpuAction-010-CpuStress / aws:ssm:send-command/AWSFIS-Run-CPU-Stress (5 min)' and 'Ec2StressCpuAction-020-CpuStress / aws:ssm:send-command/AWSFIS-Run-CPU-Stress (5 min)'. Each action has 'Edit' and 'Remove' buttons. Below the first action, it says 'Start: At beginning of experiment / Target: Ec2StressCpuTarget'.

AZ Availability: Power Interruption Scenario

Power interruption scenario to verify application is resilient to an AZ power interruption

Test Multi-AZ application against the most common symptoms of a complete AZ impairment

Multi-AZ architectures



AZ Availability: Power Interruption Scenario

Power interruption scenario to verify application is resilient to an AZ power interruption



Amazon EC2

Stop instances, insufficient capacity errors for provisioning APIs



Auto Scaling Group (ASG)

Impair provisioning to AZ via insufficient capacity errors



Amazon EKS*

Stop EC2 instances (delete pod), prevent EKS provisioning (Via ASG, Karpenter)



Amazon ECS*

Stop EC2 instances (stop task), prevent ECS provisioning (via ASG)



Amazon ElastiCache**

AZ unavailability



Amazon RDS

AZ failover



Amazon EBS

Pause IO (unresponsive volume) post-recovery

* No AWS Fargate support

** Redis only



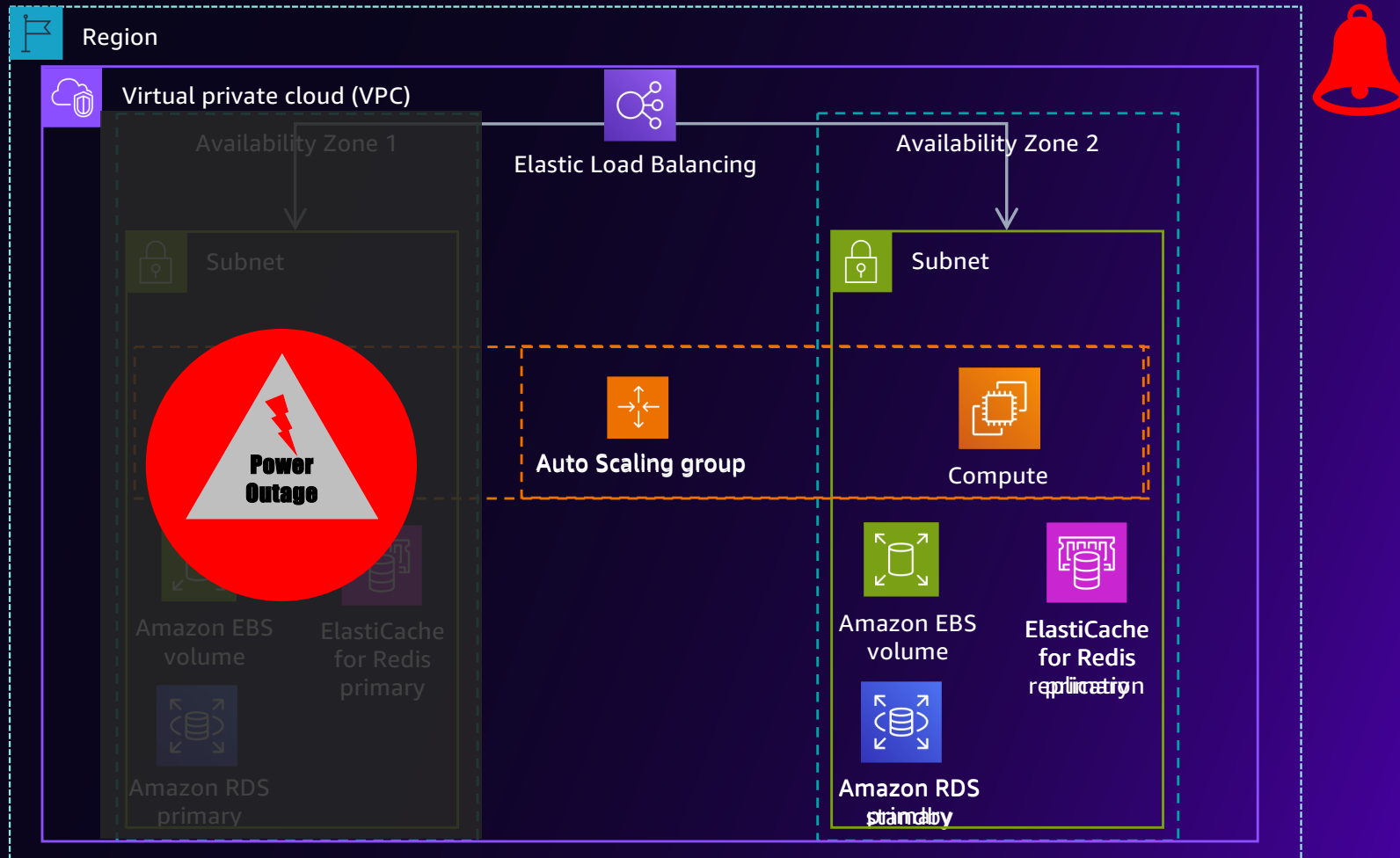
© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AZ Availability: Power Interruption Scenario



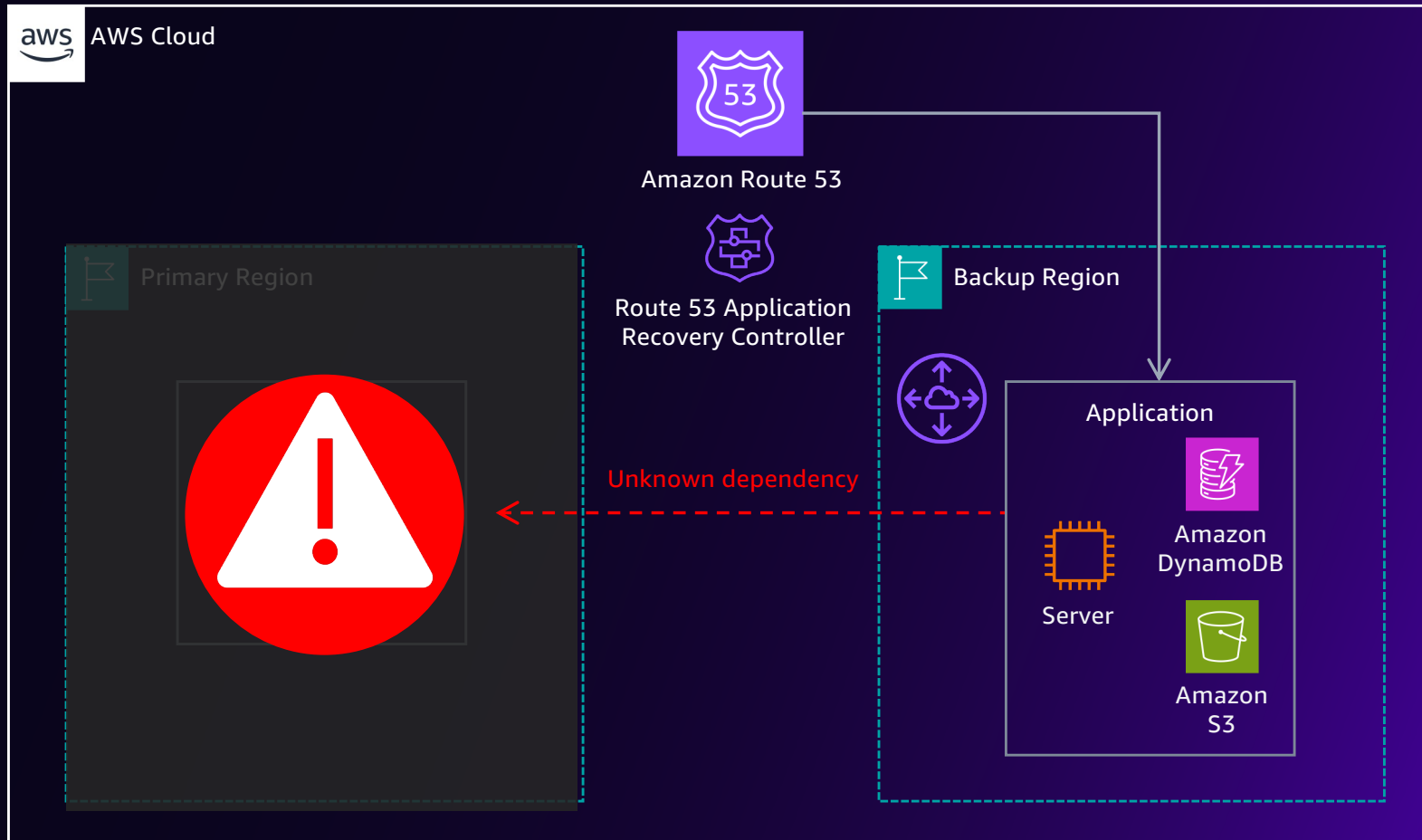
Cross-Region: Connectivity Scenario

Test for unknown dependencies from one region to another

Test multi-Region applications can operate when their primary region is inaccessible

Cross-Region: Connectivity Scenario

Test for unknown dependencies from one region to another



Cross-Region: Connectivity Scenario

Cross-region connectivity disruption to test for hidden dependencies from one region to another



Network* – block cross-Region traffic originating from VPC (internet gateway, VPC peering, AWS Transit Gateway)



Amazon DynamoDB** – pause global tables cross-Region replication

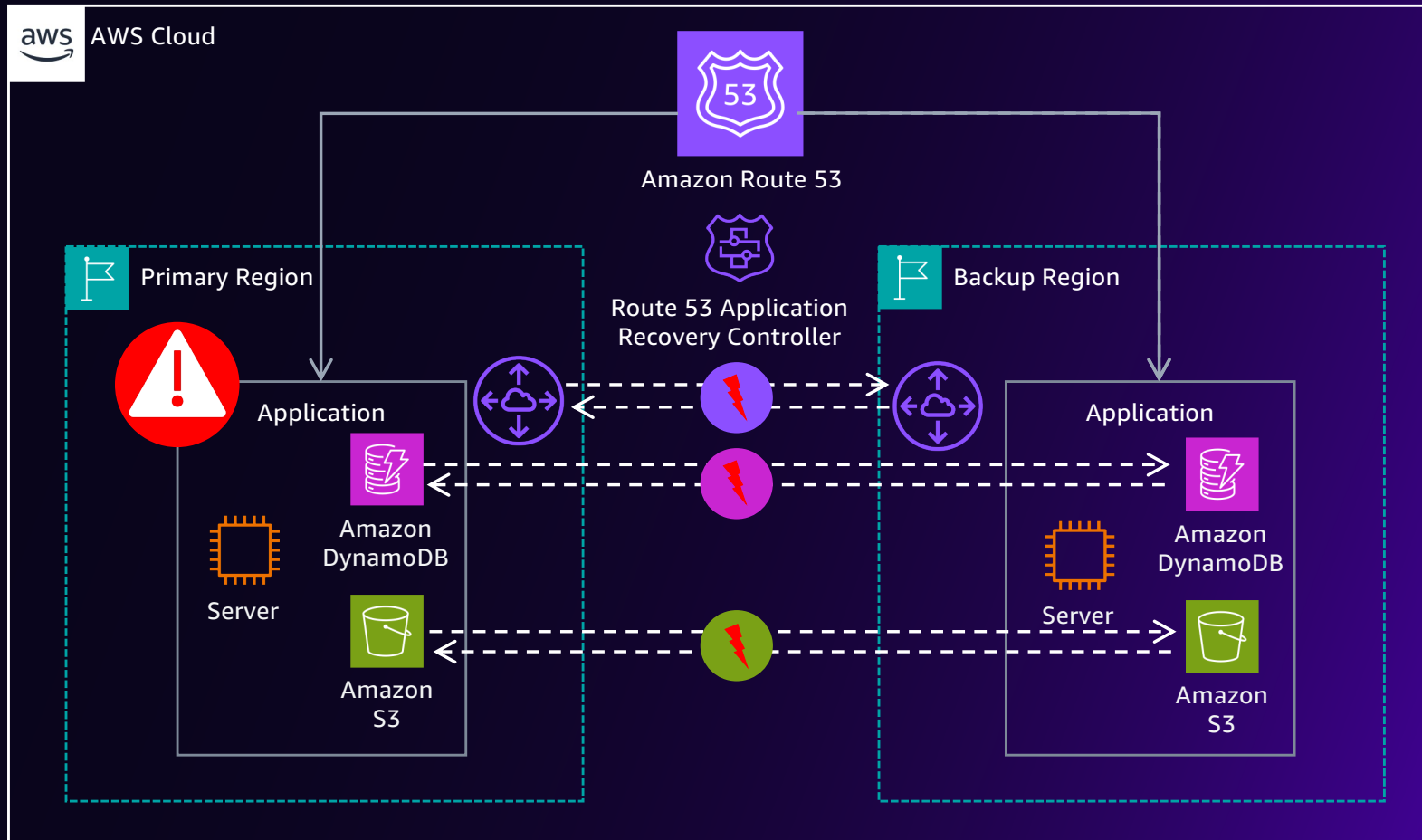


Amazon S3 – pause replication from source to destination buckets

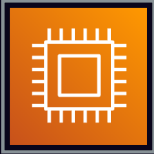
*AWS Direct Connect not supported (hybrid)

** Table encrypted with customer managed keys only

Multi-Region Architectures



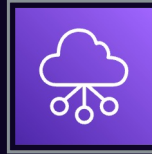
Native Targets And Actions



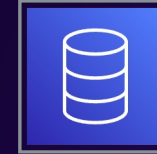
Compute



Storage



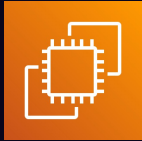
Networking



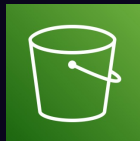
Database



Management



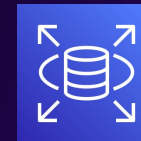
Amazon Elastic Compute Cloud (Amazon EC2)



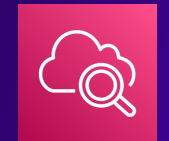
Amazon Simple Storage Service (Amazon S3)



Amazon Virtual Private Cloud (Amazon VPC)



Amazon Relational Database Service (Amazon RDS)



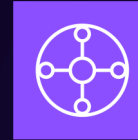
Amazon CloudWatch



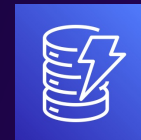
Amazon Elastic Kubernetes Service (Amazon EKS)



Amazon Elastic Block Store (Amazon EBS)



AWS Transit Gateway



Amazon DynamoDB



Amazon Elastic Container Service (Amazon ECS)

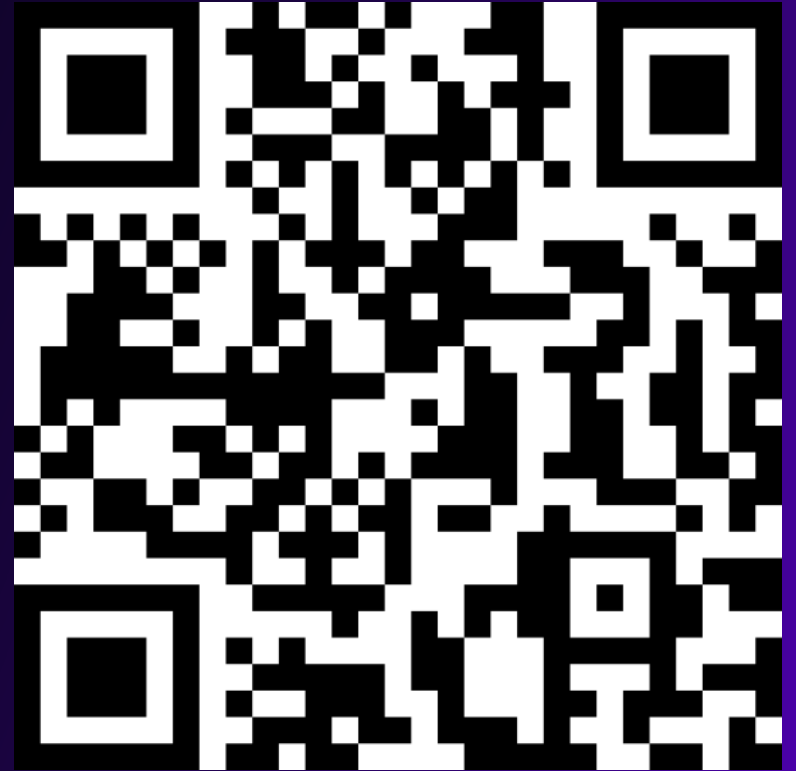


Amazon E3, C2 Auto Scaling



Amazon ElastiCache

Thank you!



Please complete the session survey



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.