Persistent Storage for Amazon ECS and AWS Fargate Using Amazon EFS

Will Ochandarena
Principal Product Manager, AWS
April 2020

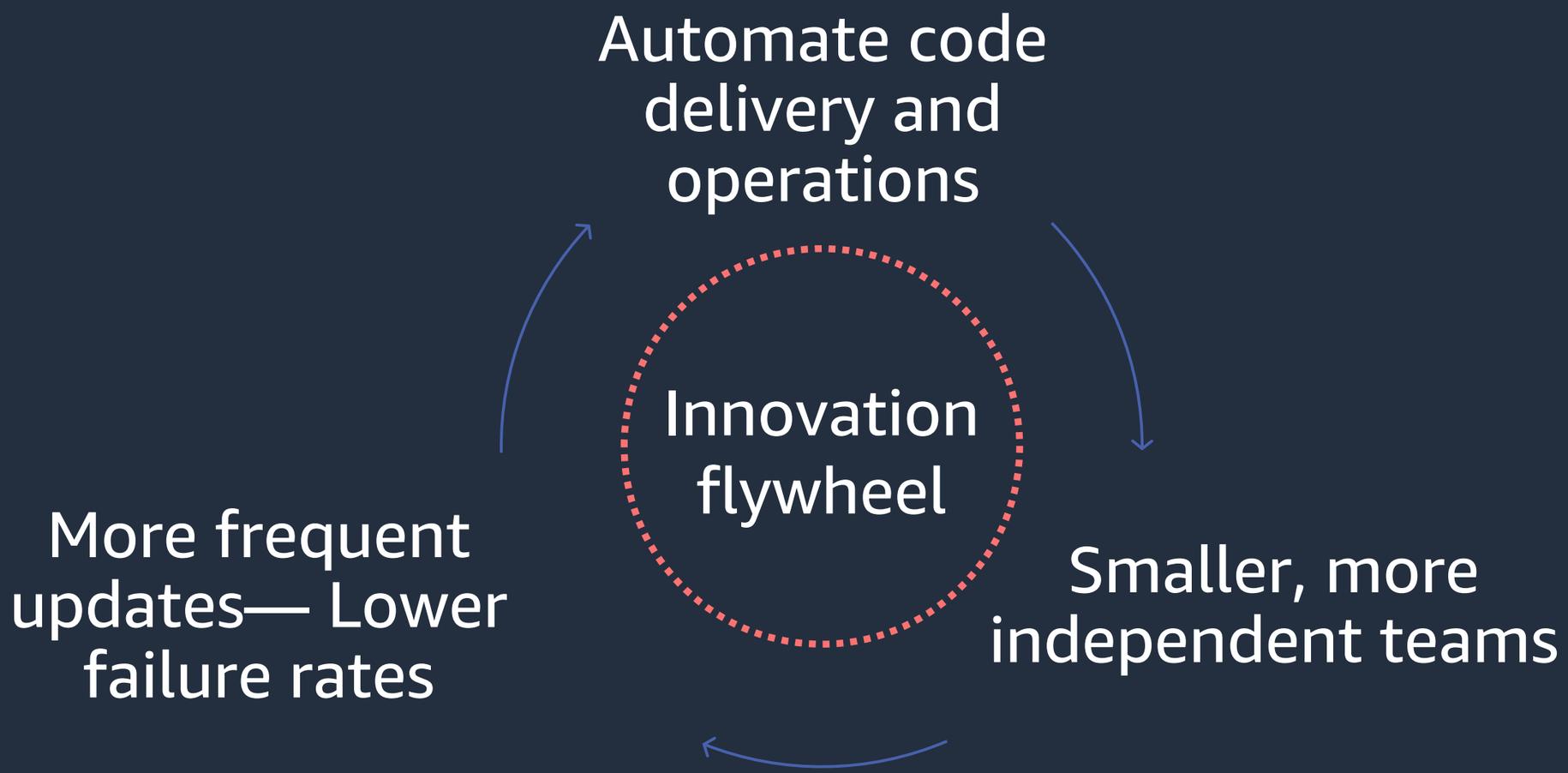
Agenda

- Intro: Why Persistent Storage for Containers?
- What's New: ECS & Fargate Support for EFS
- Best Practices: Security, Performance, & Cost
- Demo: Securely Connecting to EFS from a Fargate App

aws

Intro:
Why Persistent Storage for Containers?

Business value for container adoption



Why customers love running containers on AWS

Security, reliability, scalability, & integration

Hundreds of thousands of customers from all industries run mission-critical apps on AWS

Application focus

Serverless compute means developers don’t need to manage infrastructure

Choice: Picking the best tool for the job

Broadest choice of orchestrators and tools to support all use cases

Many containerized applications need persistent storage

Long-running
Stateful Applications

Shared Data Sets

Developer
Tools

JIRA
Git
Artifactory

Content
Management

WordPress
Drupal

Machine
Learning

MXNet
TensorFlow

Shared
Notebooks

Jupyter
Jupyterhub

aws

Traditional storage is not designed for modern applications

Lack of scalability

Lack of Agility

Administrative overhead

aws

Amazon Elastic File System (Amazon EFS)

Amazon EFS
is a fully managed file system that is…

Accelerated web serving

"We set up a new website in a few hours to announce an important donation. That task would have taken at least several days using our previous environment. That wasn't acceptable, because the announcement was very time-sensitive," says McKinnon. "With our small staff, we wouldn't have been able to create a site like this without the agility and scalability of Amazon EFS and Amazon ECS."

Aaron McKinnon, Associate director of academic computing at Caltech

Stateful microservices on-demand

"The move by T-Mobile to Amazon EFS was prompted by a need for more flexibility for our application development teams," says Amreth Chandrasehar, senior enterprise architect at T-Mobile. "By using Amazon EFS, we can address storage needs on demand. This flexibility also allows us to control costs and even shift storage requests to be managed as self-service by our application teams."

Amreth Chandrasehar, Senior Enterprise Architect, T-Mobile US

ML and AI that run consistently for Data Scientists

"We are saving deployment engineers days of effort every time we deploy the solution." Faculty has also improved collaboration between data scientists. "There is more consistency in the data science development process because everyone is working with the most updated code base thanks to Amazon EFS," says Stevenson. "This ultimately helps us deliver a better product to our customers."

Scott Stevenson, Data Engineer, Faculty

aws

What's New:
ECS & Fargate Support for EFS

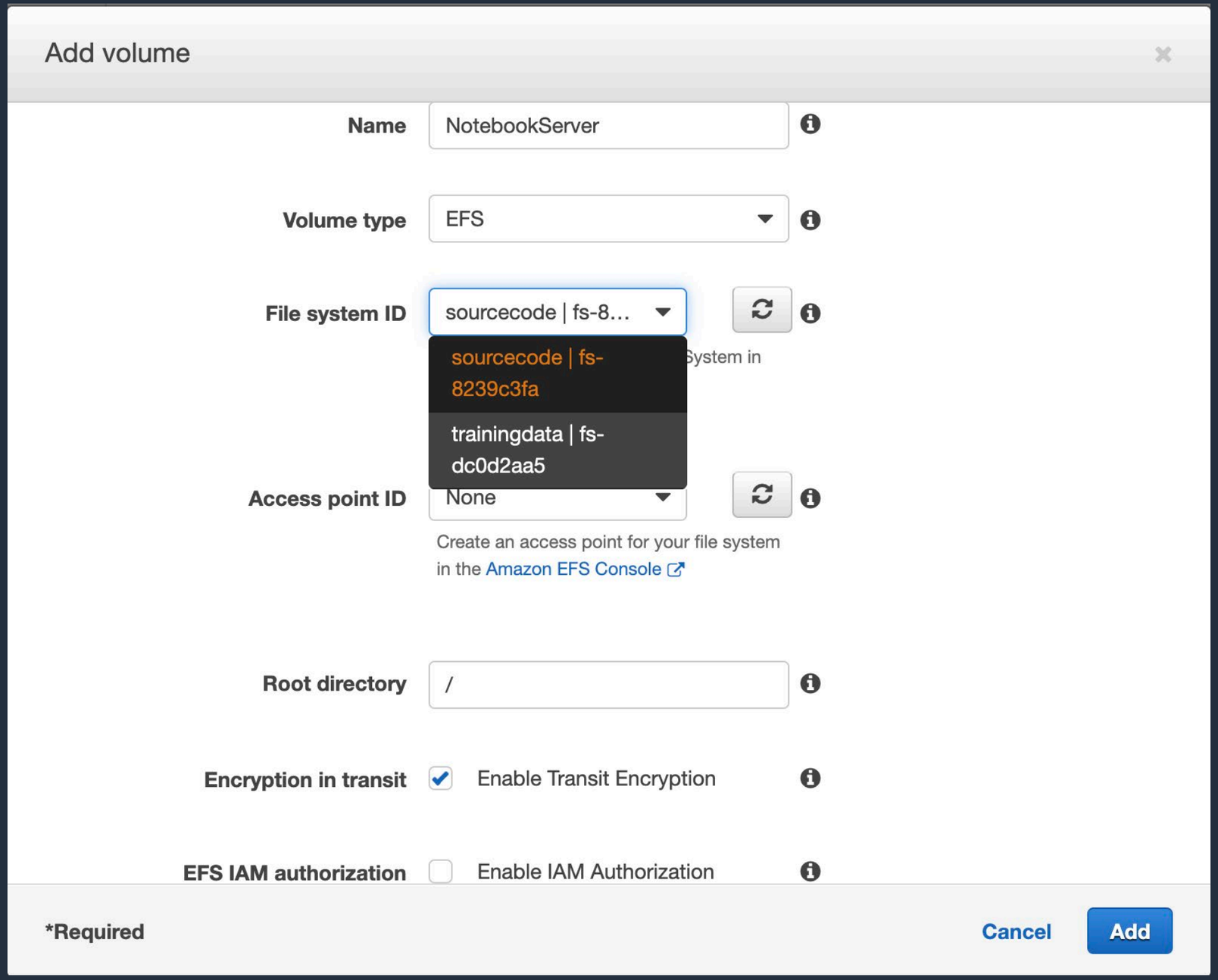
New:
Amazon ECS and AWS Fargate Support for Amazon EFS

Simple: All EFS configuration is inside the ECS task definition, and connectivity is handled behind the scenes.

Serverless: AWS Fargate tasks can now leverage shared persistent storage.

Secure: Access to file systems can be authorized by IAM, and access to data controlled by EFS Access Points.

aws



Simplify Persistent Storage for Modern Applications with Amazon EFS for Amazon ECS and AWS Fargate

Simple

Amazon EFS configuration is inside Amazon ECS task definition, including connectivity, so developers can focus on their applications, not infrastructure.

Available and Durable

Amazon ECS, AWS Fargate, and Amazon EFS are regional services. Customers can build applications that span multiple availability zones, with automatic failover.

Elastic

Amazon ECS, AWS Fargate, and Amazon EFS are fully elastic, scale up and down rapidly based on demand. Customers pay only for what they use.

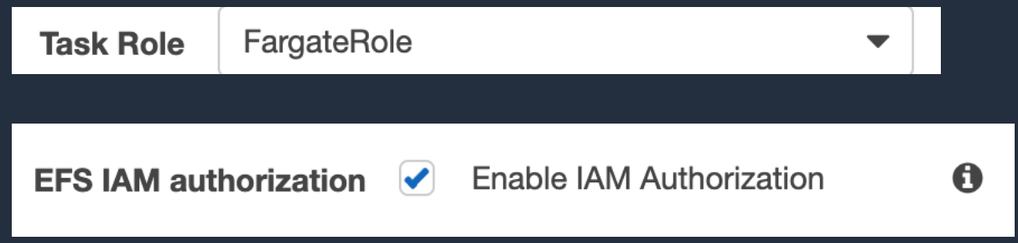
Secure

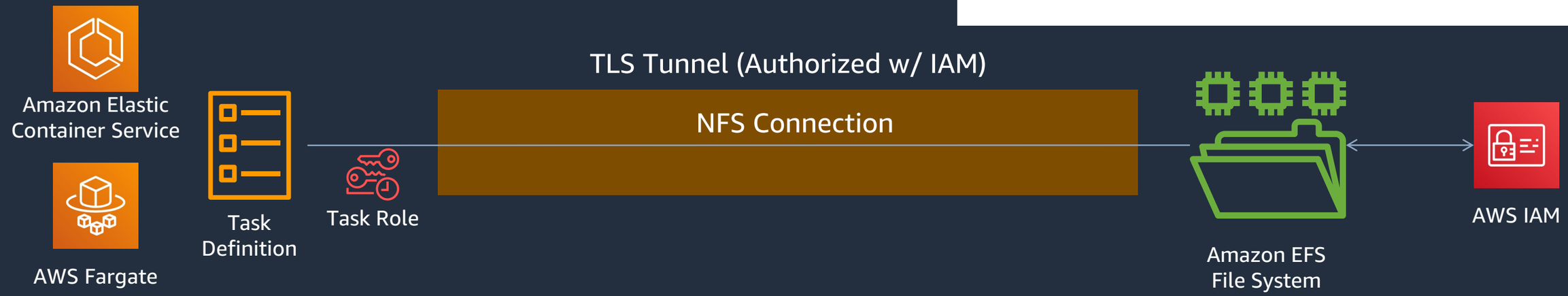
Access to Amazon EFS can be restricted based on the IAM role of the Amazon ECS task.

Amazon EFS Access Points can enforce file system permissions when multiple apps share a file system.

Best Practices: Security, Performance, & Cost

Using IAM for File System Access





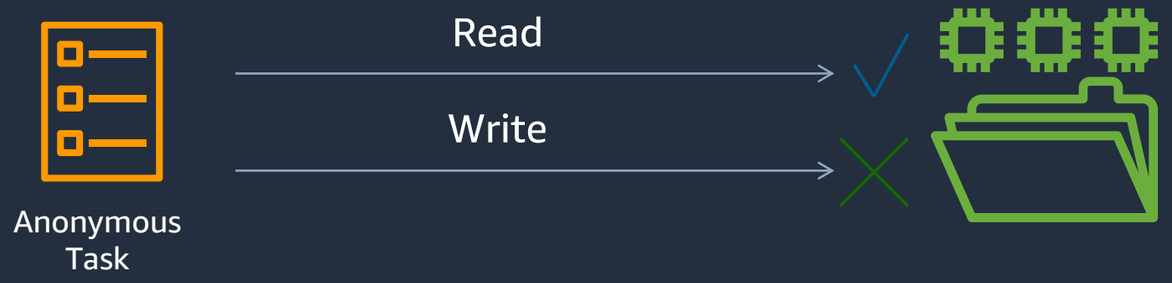
Identity Policy Attached to IAM Role

```
{
  “Statement” : {
    “Effect” : “allow”,
    “Action” : “elasticfilesystem:Client*”,
    "Resource": ”fs-deadbeef"
  }
}
```

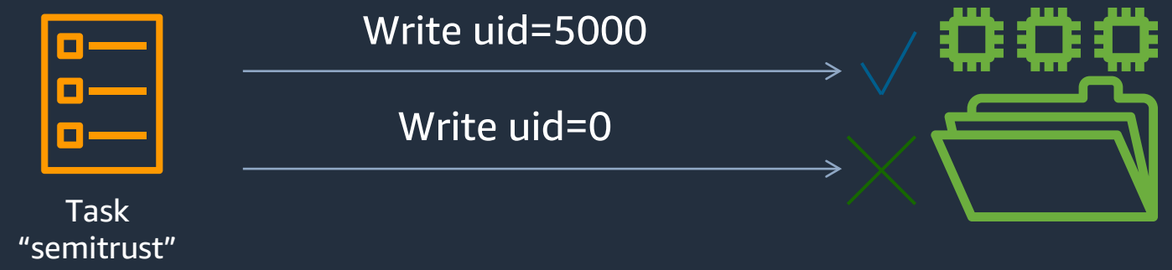
File System Resource Policy

```
{
  “Statement” : {
    “Effect” : “allow”,
    “Action” : “elasticfilesystem:Client*”,
    “Principal” : { “AWS”: “FargateRole” }
  }
}
```

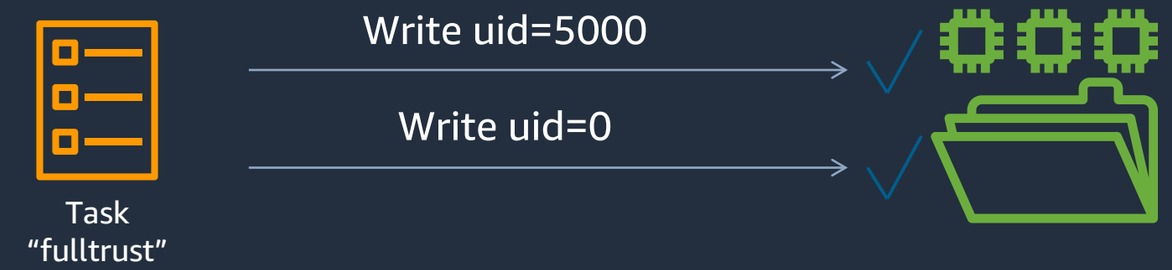
Handling EFS Authorization Using IAM




```
"Effect" : "allow",
"Action" : "elasticfilesystem:ClientMount",
"Principal" : "*"
```




```
"Effect" : "allow",
"Action" : ["elasticfilesystem:ClientMount",
            "elasticfilesystem:ClientWrite"],
"Principal" : { "AWS": "semitrust" }
```




```
"Effect" : "allow",
"Action" : ["elasticfilesystem:ClientMount",
            "elasticfilesystem:ClientWrite",
            "elasticfilesystem:ClientRootAccess],
"Principal" : { "AWS": "fulltrust" }
```

Application-specific Access with EFS Access Points

Creates App-specific Directory & Permissions

No EC2 instance required!

Apps only see data they need

Enforces File System Identity

Root containers can’t escalate access

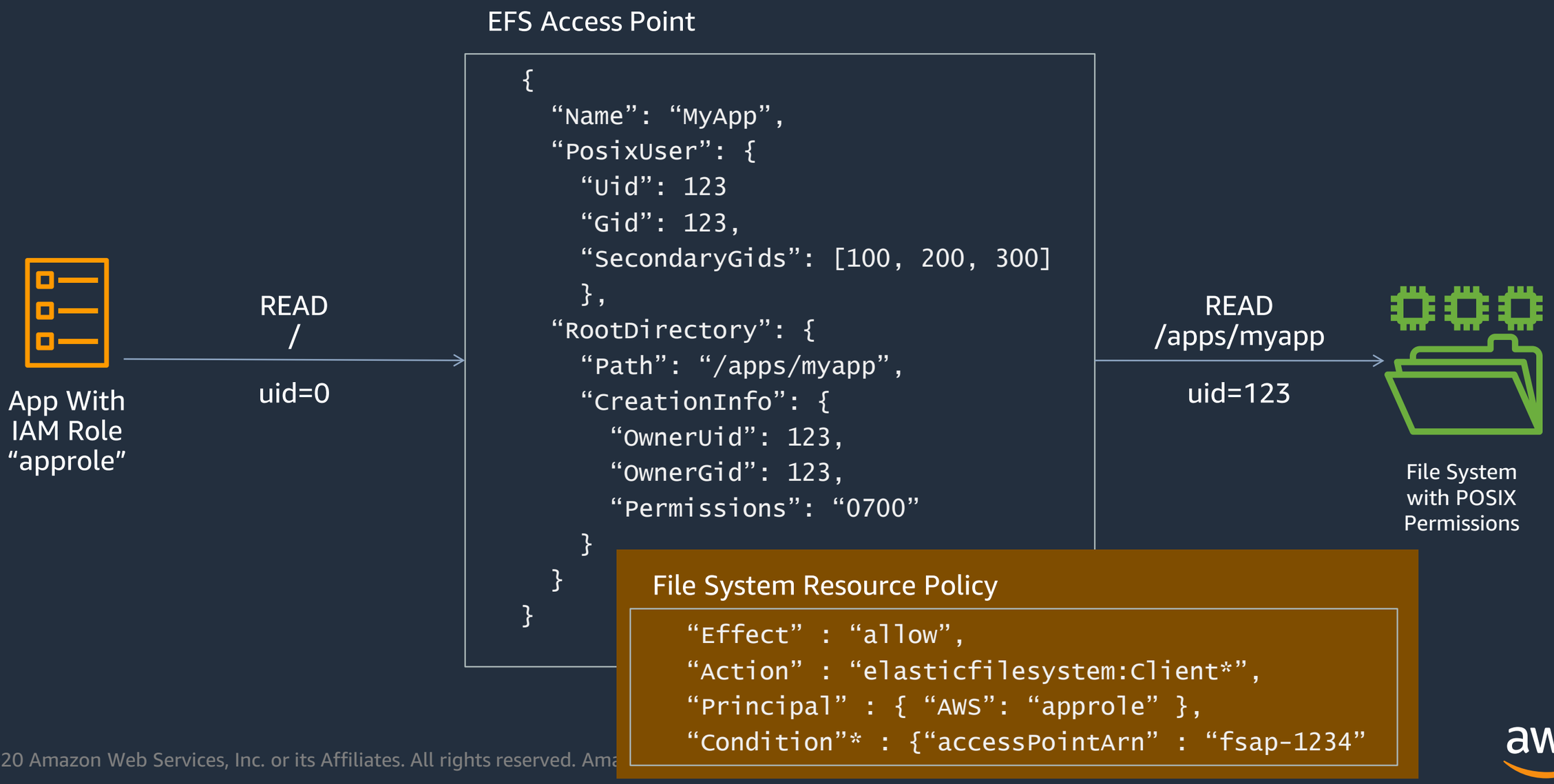
Arbitrary users aren’t locked out

EFS Access Point

```
{
  “Name”: “MyApp”,
  "FileSystemId": "string",
  “PosixUser”: {
    “Uid”: 123
    “Gid”: 123,
    “SecondaryGids”: [100, 200, 300]
  },
  “RootDirectory”: {
    “Path”: “/apps/myapp”,
    “CreationInfo”: {
      “OwnerUid”: 123,
      “OwnerGid”: 123,
      “Permissions”: “0700”
    }
  }
}
```


aws

How EFS Access Points Work



Best Practices for IAM and Access Points, & Security

- Use EFS Access Points, even if single app per file system!
  - Simplifies directory permission setup
  - Consistent experience regardless of user/group setup in container
  - Future-proof for adding apps to share data
- Use IAM Authorization
  - Use Resource Policies to restrict IAM roles to Access Points
  - Use Identity Policies to give single role “admin” access to file systems
- Enable Encryption @ Rest and Encryption in Motion
  - Simple setup, no performance penalty

aws

Best Practices - Performance

Amazon Elastic File System announces 400% increase in read operations for General Purpose mode file systems

Posted On: Apr 1, 2020

- **Use General Purpose** for most apps
  - GP lower latency, now supports up to 35K read IOPS
  - MaxIO for scale-out analytics/ML that need 100k+ IOPS
- **Configure provisioned throughput** for initial need
  - As your file system grows you'll eventually be given higher throughput
- **Set up Amazon CloudWatch**, monitor throughput, IOPS, and burst credits*

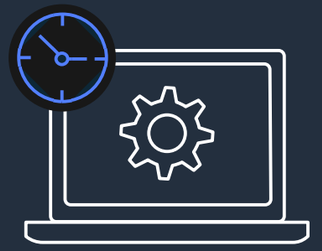
* https://github.com/aws-samples/amazon-efs-tutorial/tree/master/monitoring

Optimize cost with Amazon EFS Infrequent Access

Amazon EFS IA storage class for infrequently accessed files for $0.025/GB per month*

No changes to existing applications using Amazon EFS

Cost savings up to 92%

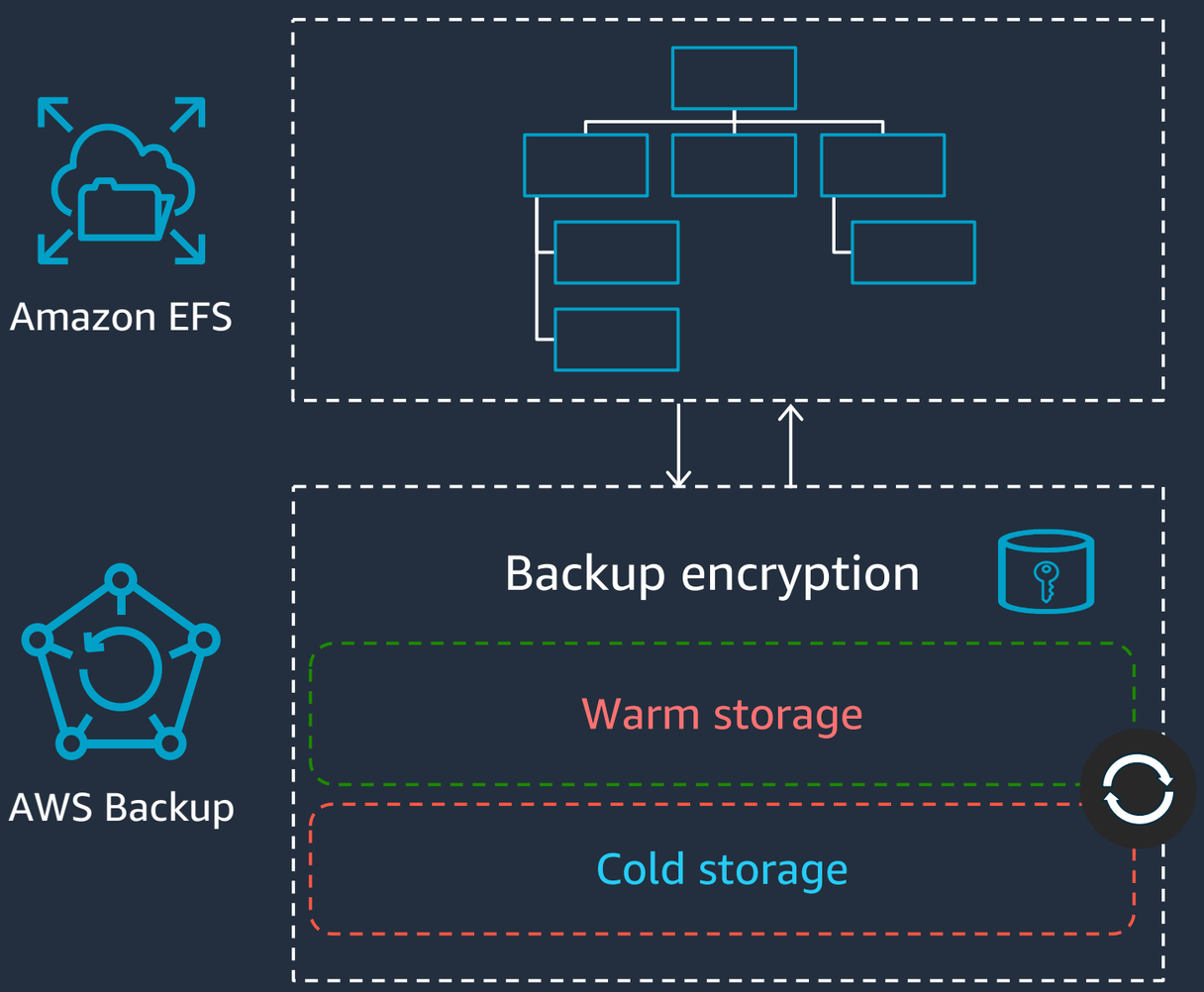


Automated lifecycle management

* Pricing in the US East (N. Virginia) region

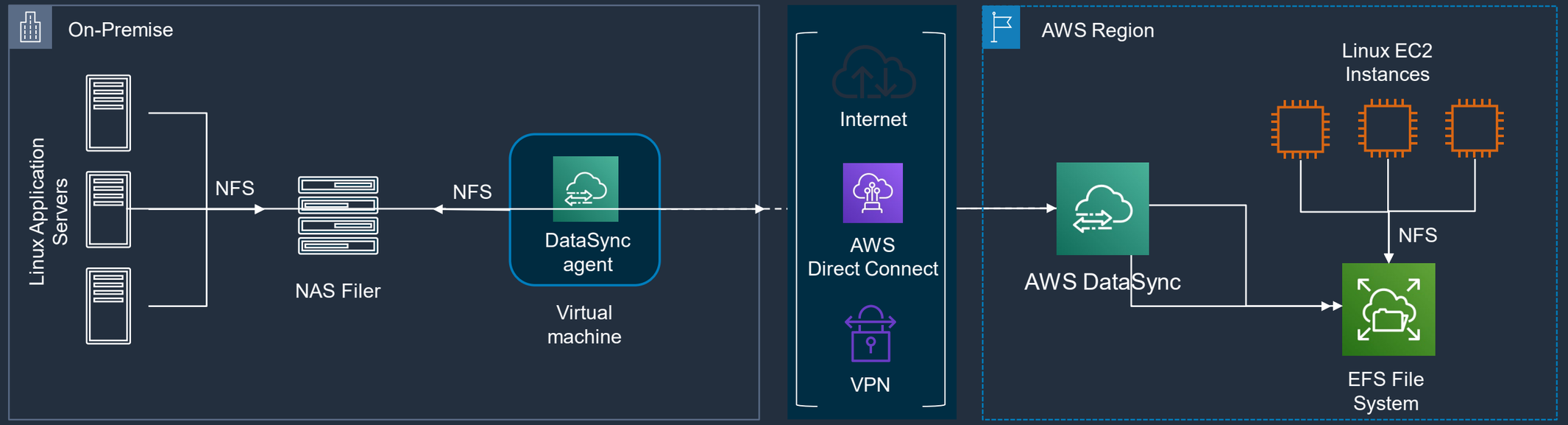
aws

Backup for Amazon EFS



- EFS file systems can be backed up and restored using AWS Backup
- AWS Backup provides automated backup scheduling and retention per user defined policy
- AWS Backup offers two classes of service backup storage with the ability to lifecycle to cold storage
- Restore individual files and directories

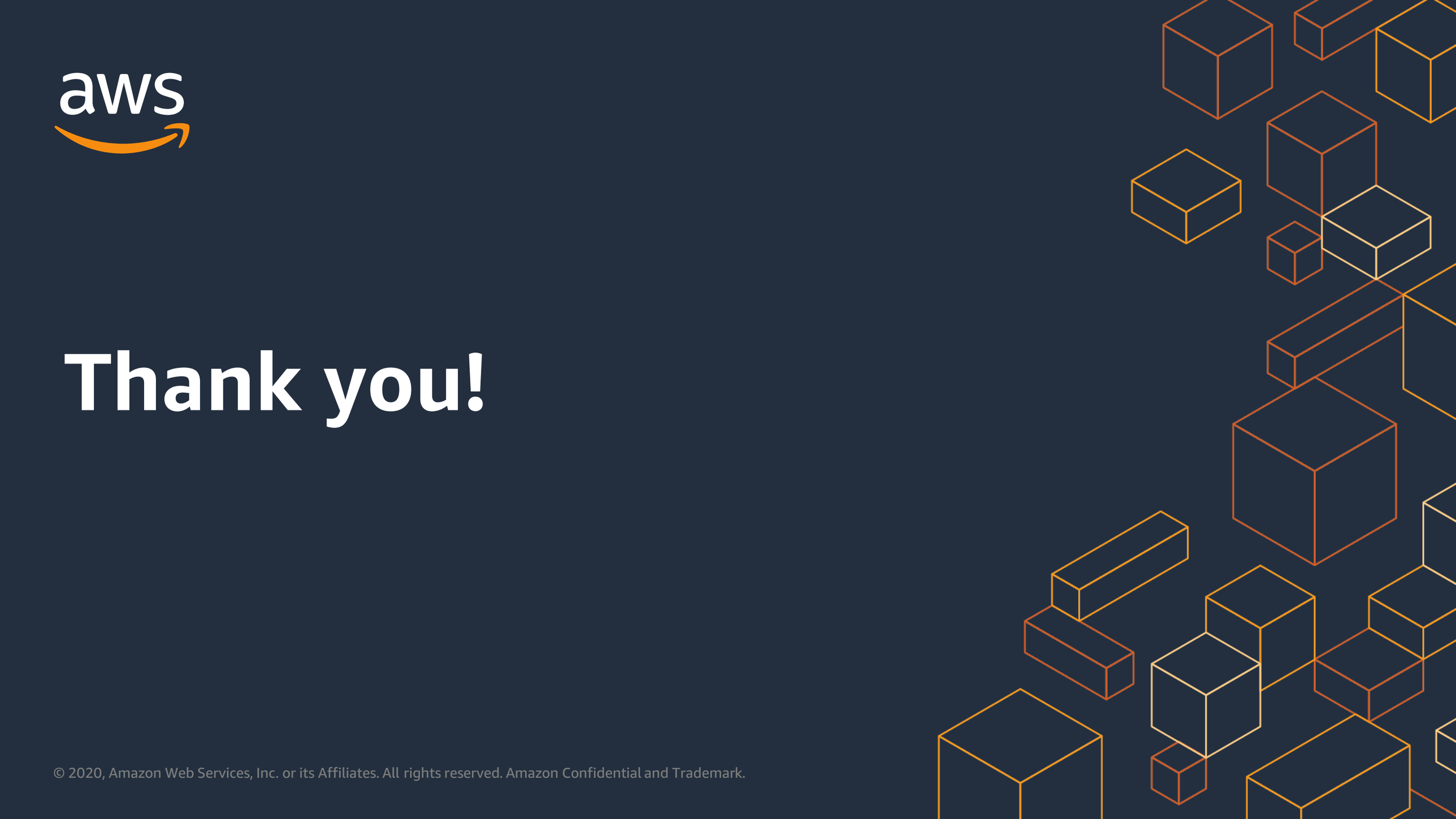
Migrating NFS workloads to EFS



AWS DataSync: Online transfer service that simplifies, automates, and accelerates moving data between on-premises storage and AWS

Demo:
Securely Connecting to EFS from a Fargate App

Q&A



Thank you!

Cloud native | Amazon EFS

Elastic

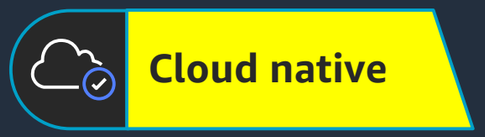
- Grow & shrink on demand
- No need to provision and manage infrastructure & capacity
- Pay as you go, pay only for what you use
- Simple to use, create a file system in seconds

Scalable

- Grow up to petabytes
- Performance modes for low latencies and maximum I/O
- Throughput that scales with storage
- Provisioned throughput available

Integrated

- Shared access from on-premises, inter region, and cloud-native applications
- Integrated with various AWS computing models
- Access concurrently from thousands of Amazon EC2 instances
- Attach to containers launched by both Amazon ECS and EKS
- Use with Amazon SageMaker notebooks



Highly reliable | Amazon EFS

Highly available, durable

- Stores data across three availability zones for high availability and durability
- Access your file system from multiple AZs
- Strong consistency for concurrent access

Secure

- Control network traffic
- Control file and directory access
- Control administrative (API) access with AWS IAM
- Encrypt data at rest and in transit
- Control NFS client access with AWS IAM

Global footprint

- Available in all AWS commercial regions
- Feature parity in every region

Cost optimized | Amazon EFS

No minimum commitments or upfront fees

No need to provision storage

Use with Spot Instances

Automatic lifecycle management to lower cost storage