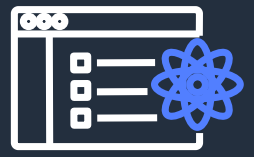
Achieve faster time to insights for data analytics with Amazon EFS

Delwin Olivan (he/him)

Senior Product Manager, Amazon EFS

Put your data to work

Machine learning and AI



Data science

Financial analytics

Genomics research

AWS Builders are continuing to look for ways to get more value out of their data

Developers

Artists and editors

Data scientists and ML engineers

A modern approach to data analytics

Monolith vs.
Microservices

Traditional servers vs.
Containers and serverless compute

Amazon EKS

Amazon ECS

AWS Fargate

AWS Lambda

Containers

Serverless

Deliver insights faster with a modern approach to data analytics

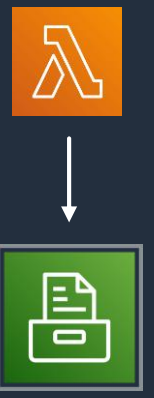
Increase agility by instantly scaling up according to demand

Maximize value from your limited resources

Develop and deploy applications with greater efficiency

Faster time to insights

Modern data analytics applications need scalable, persistent, shared storage

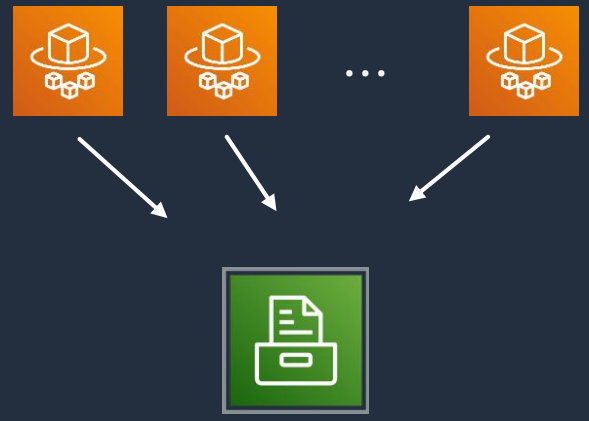


Scalable

Fast access to large volumes of data

Persistent

Long-lived durable storage for data sets, trained models, …



Shared

Natively accessible from distributed jobs, end-users, …

Builders don’t want to think about storage

Limited storage expertise

Used to storage that “just works”

File storage for AWS Builders

“Just works”

Fully elastic, scalable storage
Highly durable (11x9s)
Highly available

Modern integrations

Fully integrated with services like Amazon ECS, Amazon EKS, AWS Lambda to support modern applications

Amazon EFS

Serverless, elastic file system

High performance

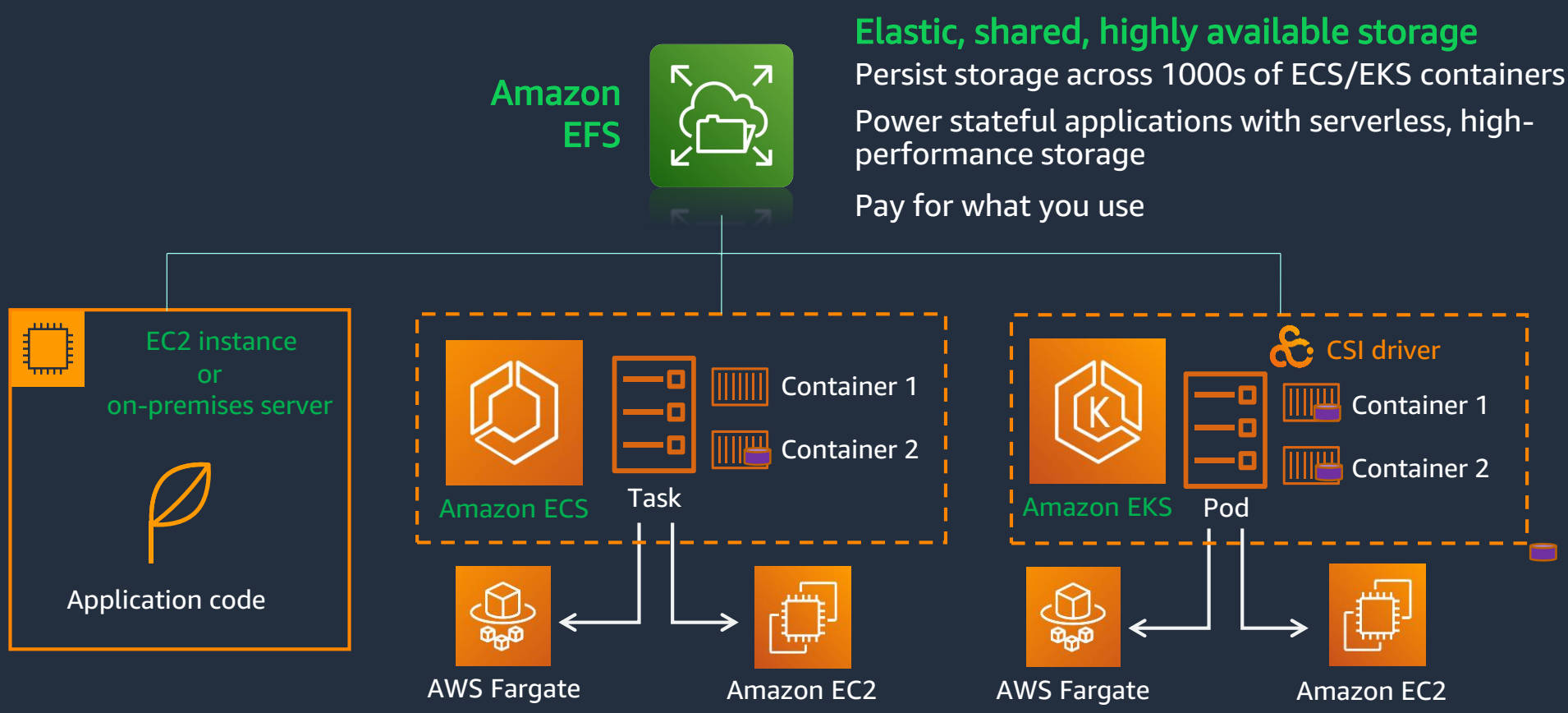
Latencies as low as 0.6 ms
Up to 35 K IOPS
Up to 3 GB/s of throughput

Cost optimized

Pay for what you use model
Automatically tier infrequently accessed data to colder storage

Containers: EFS for AWS container services (ECS/EKS)

PERSISTENT SCALABLE STORAGE FOR CONTAINERIZED APPLICATIONS



J&J builds data science platform using Amazon EKS and Amazon EFS

Challenge

J&J needed storage to share and perform analytics on their data science workbench for genomics neuroscience, R&D, and drug discovery; a highly scalable solution was required to share and analyze hundreds of TiB of data from external and internal studies

Solution

Amazon EFS provides performant analytics storage with shared file access to data scientists using open-source genomics and Shiny, as well as Domino Data Lab, running on an Amazon EKS cluster

Benefits

- Faster time to insights: reduced analysis time by 35%
- Scales elastically to 500 TiB and growing
- Reduced costs by 37% using EFS lifecycle management and EFS One Zone lower cost storage classes

---

Data scientists need access to massive data from different sources for genomics, R&D, and drug discovery to perform advanced analytics. **With Amazon EFS, we were able to share data on a massive scale, increase innovation by accelerating analytics, and save 37% in costs.**

Greg Rusin, IT Manager, Johnson & Johnson

aws

Serverless: EFS for AWS Lambda

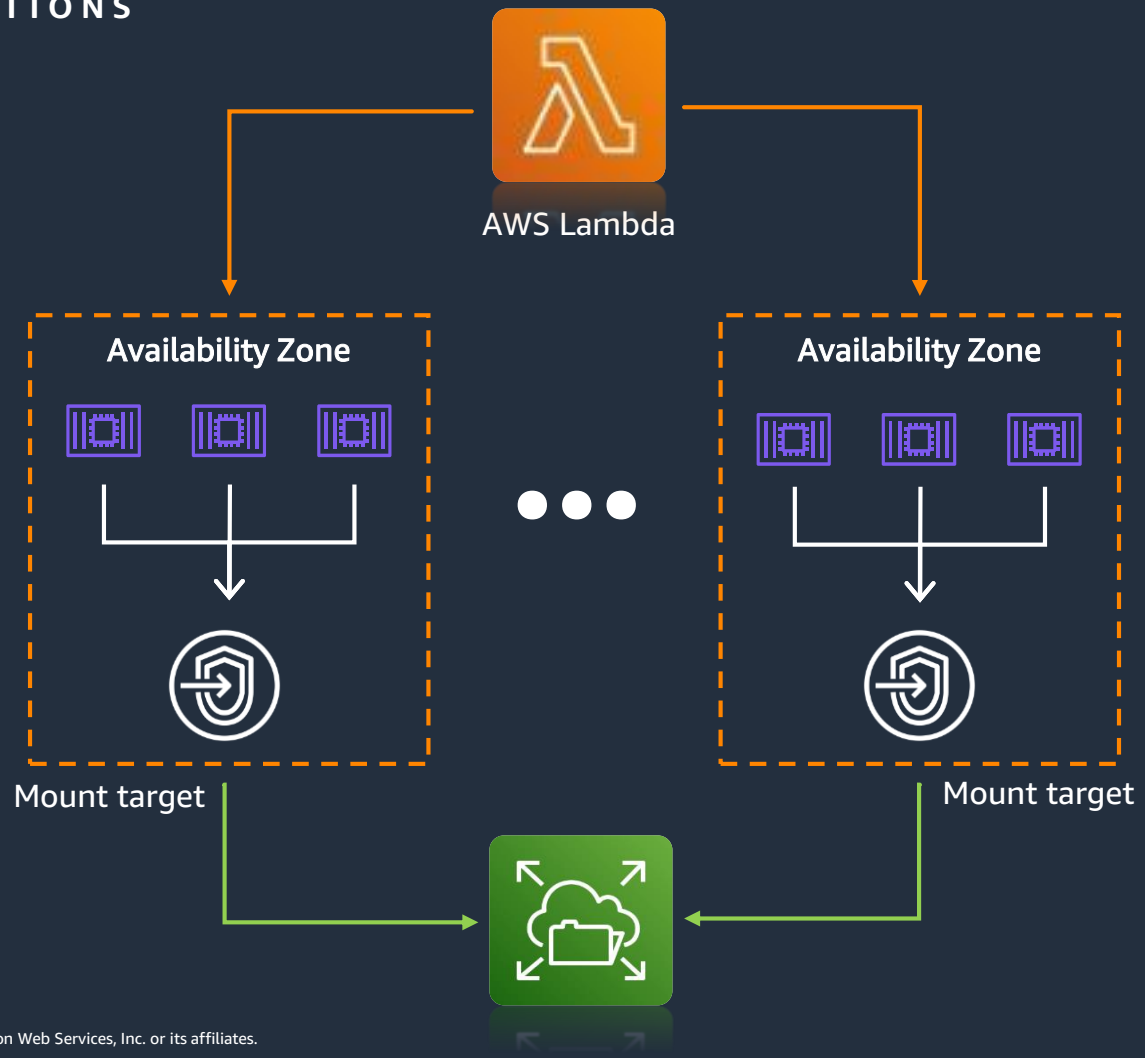
SERVERLESS STORAGE FOR SERVERLESS APPLICATIONS

Elastic, shared, serverless storage

Provide petabyte-scale storage for 1000s of Lambda functions

Power stateful applications with serverless, high-performance storage

Pay for what you use



Asurion powers real-time machine learning insights using Amazon EFS

Challenge

Asurion needed to perform real-time analysis of customer experience during support calls, leveraging their machine learning models. Recordings didn't fit in the available storage on AWS Lambda

Solution

Amazon EFS provides shared, persistent storage for Asurion's trained machine learning models, powering their real-time inference workflows running on AWS Lambda

Benefits

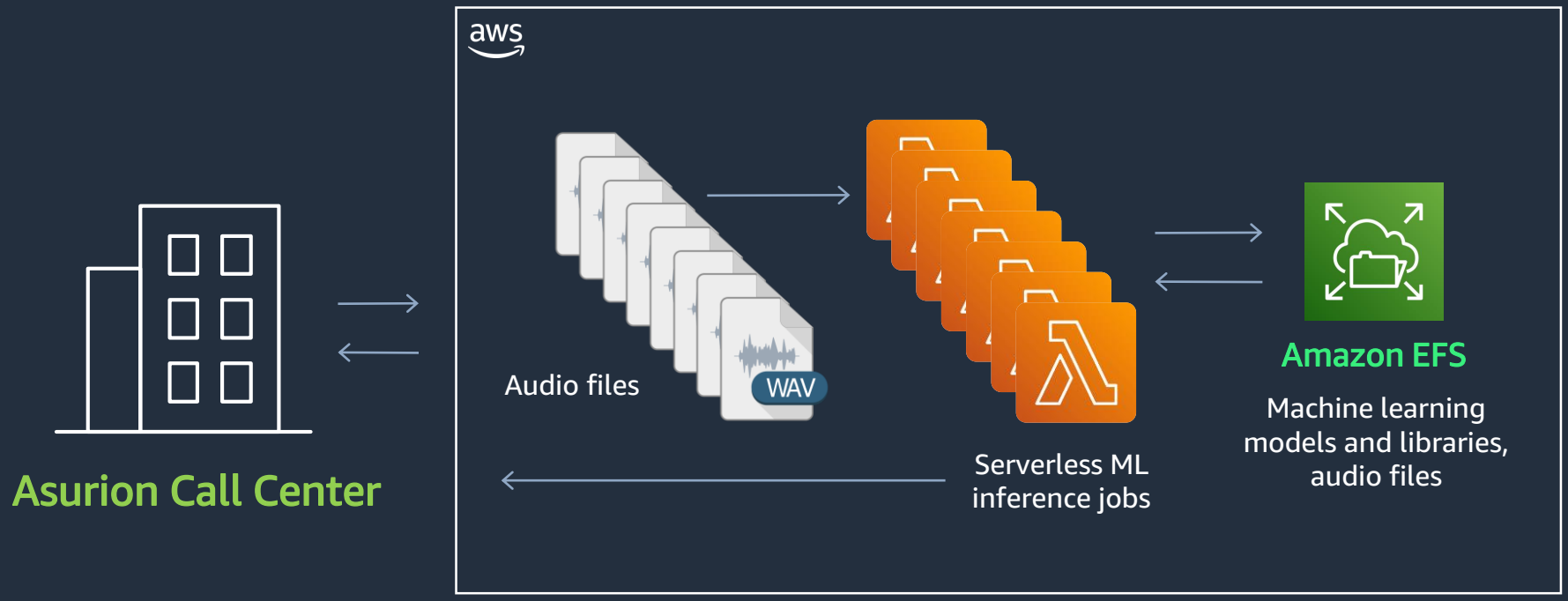
- ML inference infrastructure scales elastically with call volume
- Reduced operational overhead compared to maintaining instances and auto-scaling

---

We really wanted to use AWS Lambda to make our ML inference elastic but thought we wouldn't be able to because of the size of data the process required.
With Amazon EFS, we were easily able to give our [Lambda] function all of the storage space it needs.

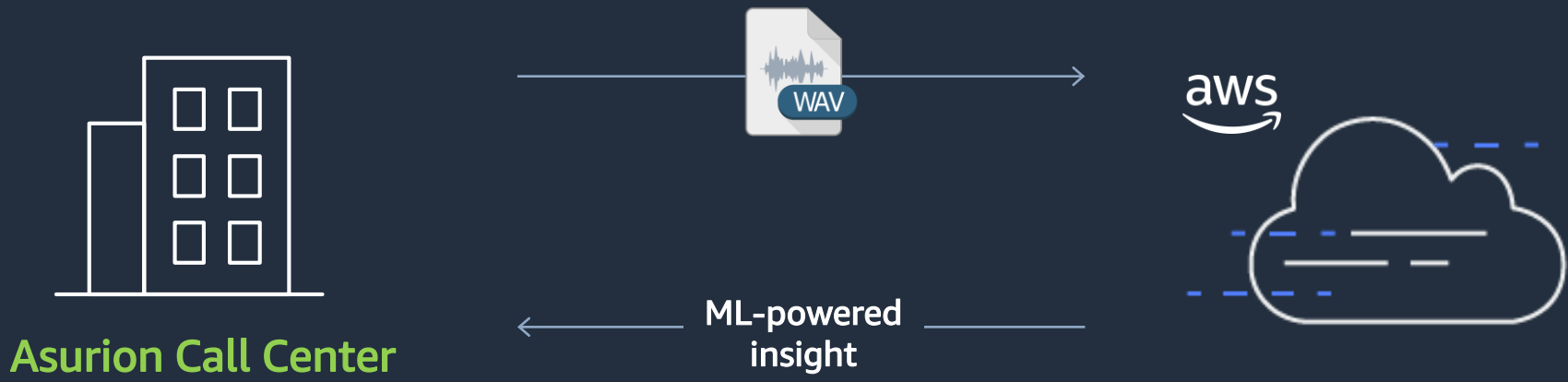
Jeff Tougas, Senior Principal Software Engineer, Asurion

How does this look in practice?



Delivering real-time analysis of customer experience during support calls

How does this look in practice?



Delivering real-time analysis of customer experience during support calls

Modern storage for modern data analytics

Amazon
EFS

Amazon
S3

"Just works"

Modern integrations

High performance

Cost optimized

Thank you!