STG315

AWS re:INVENT

Case Study: Learn How Thomson Reuters Uses Amazon EFS to Deliver Billions of Pieces of Content to Hundreds of Millions of Visitors Every Year

Edward Zhitnitsky - Production Infrastructure Manager, Thomson Reuters
Andy Crudge – Amazon EFS Product Manager

November 30, 2017

In the fast-paced world of news media, having a high performance, reliable storage service is a critical component for delivering content at scale. Learn how Thomson Reuters leverages Amazon EFS to host their rich media content, saving them time and money while delivering unmatched reliability. Come hear about their journey, their approach, their architecture, and the considerations they took when building out their environment. Best practices and tips for success will be shared throughout.

Agenda:

Amazon Storage and EFS Overview
Case Study: Thomson Reuters
Reference and Other Sessions
Q&A

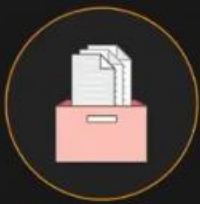# What do you think about when choosing a storage solution?

Storage type	Features and performance	Economics

Three types of storage

File	Block	Object

Three types of storage

Data stored as files in a directory hierarchy

Shared over a network

File

Three types of storage

Block

Data stored as blocks on a disk or disks

Locally attached

Three types of storage

Object

Data is stored as an object that's identified by a key in a flat space
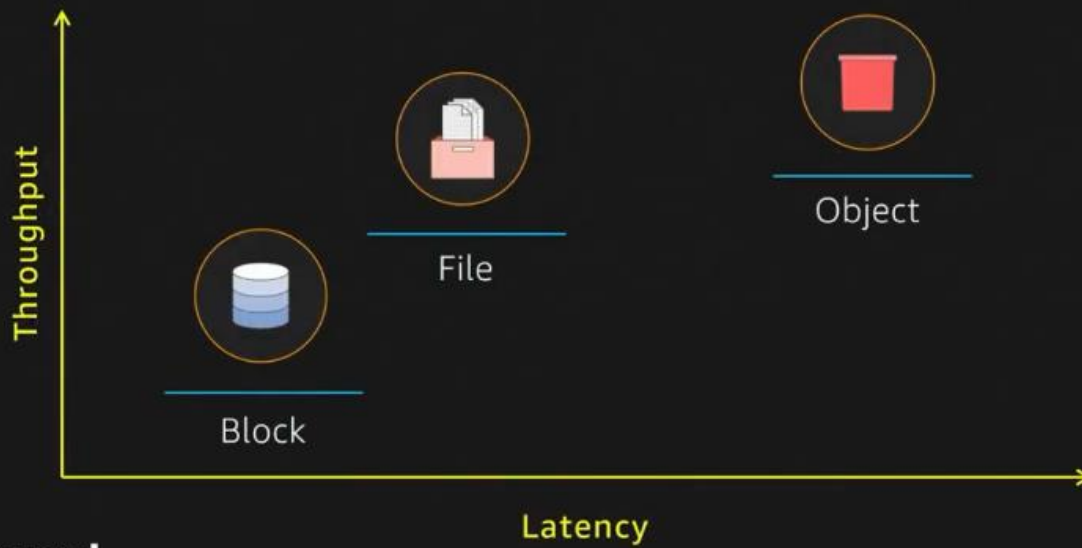
Simple API to get and put data based on key

Why is file storage so popular?

Works natively with operating systems

Provides shared access while providing consistency guarantees and locking functionality

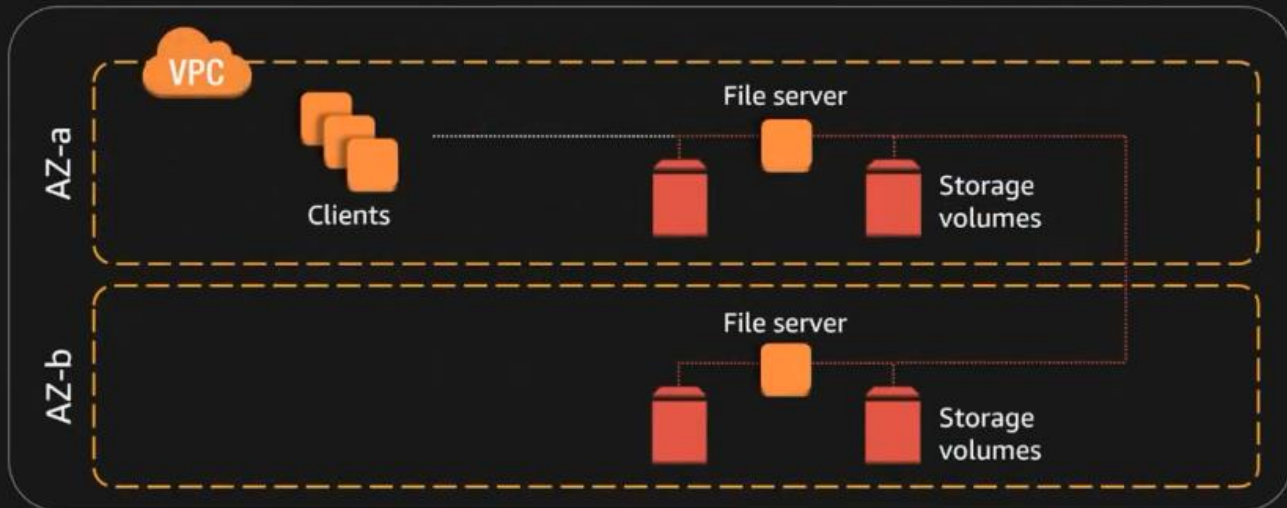**Provides hierarchical namespace**

How does performance compare?

Throughput vs Latency chart showing Block, File, and Object storage

Before EFS... DIY file storage

Amazon EFS
A fully managed file service

Key features of Amazon EFS

Simple → Elastic → Scalable

Highly available and durable

aws

Designed for a wide spectrum of needs

Analytics	Web serving	Dev tooling
M&E workflows	Content management	Home directories
	Database backups	Container storage
	Enterprise apps & messaging	
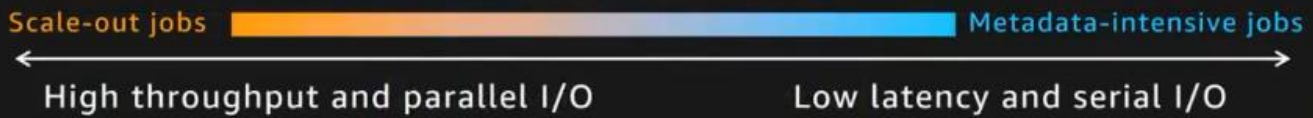
Scale-out jobs → Metadata-intensive jobs

High throughput and parallel I/O — Low latency and serial I/O

Amazon EFS customers

Database backups
J.D. POWER   Sysco

Media and entertainment workflows
bambuser   zynga

Enterprise applications
CardinalHealth   Atlassian   infor

Developer tools
zend   A

Home directories
coursera   ASI DATA SCIENCE

Container storage
NETFLIX   here

Web serving + content management
THOMSON REUTERS   workfront   PlanGrid   FINRA
SPOKEO   BeSoftware   Seeking Alpha   DigitalReLab
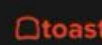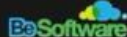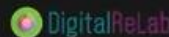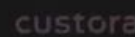
Big data and analytics
toast   ZS   MONSANTO
custora   kissmetrics   Celgene

Security model and security features

Control network traffic using VPC security groups and network ACLs

Control file and directory access using POSIX permissions

Control administrative access (API access) using AWS Identity and Access Management (IAM) (action-level and resource-level permissions)
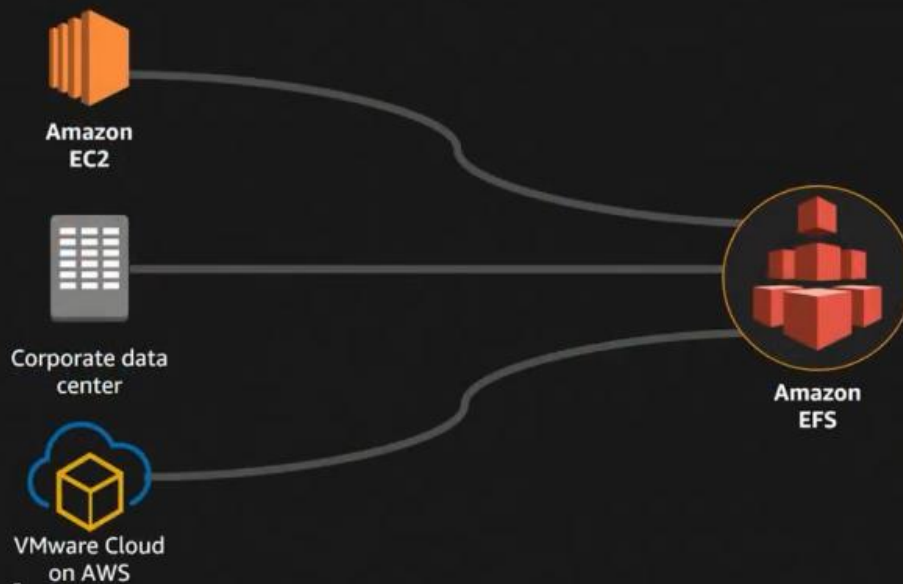
Encrypt data at rest using keys managed in AWS KMS

Access EFS from various environments

Amazon EC2

Corporate data center

VMware Cloud on AWS

Amazon EFS

EFS economics

No minimum commitments or up-front fees

No need to provision storage in advance

No other fees, charges, or billing dimensions
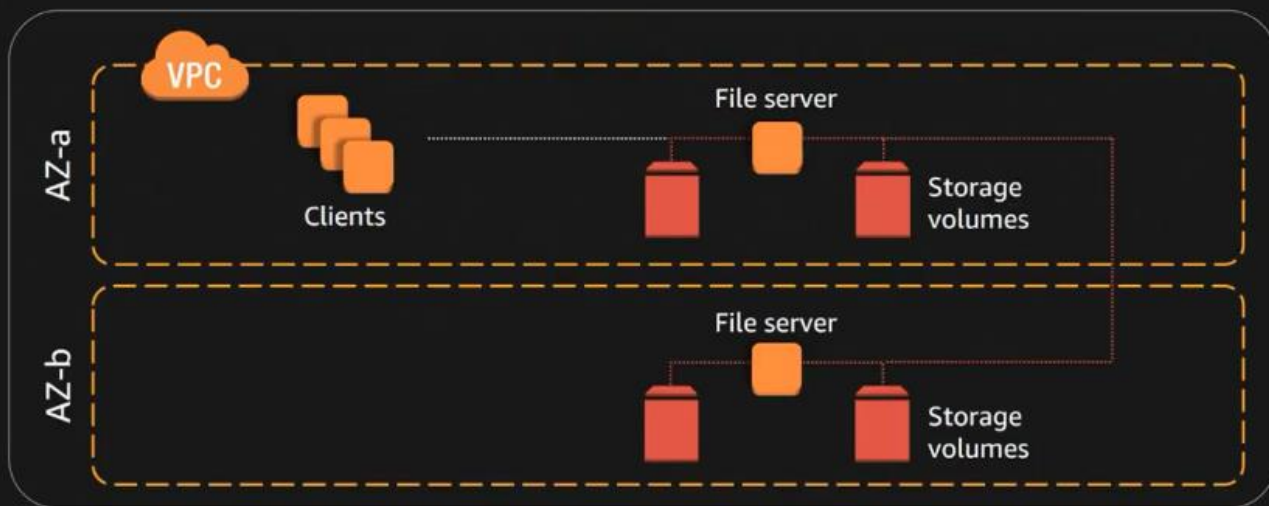
$0.30/GB-Month (US Regions)
$0.33/GB-Month (EU Ireland)
$0.36/GB-Month (EU Frankfurt)
$0.36/GB-Month (AP Sydney)

Before EFS... Costs of DIY file storage

VPC

AZ-a

Clients

File server

Storage volumes

AZ-b

File server

Storage volumes

Before EFS... Costs of DIY file storage
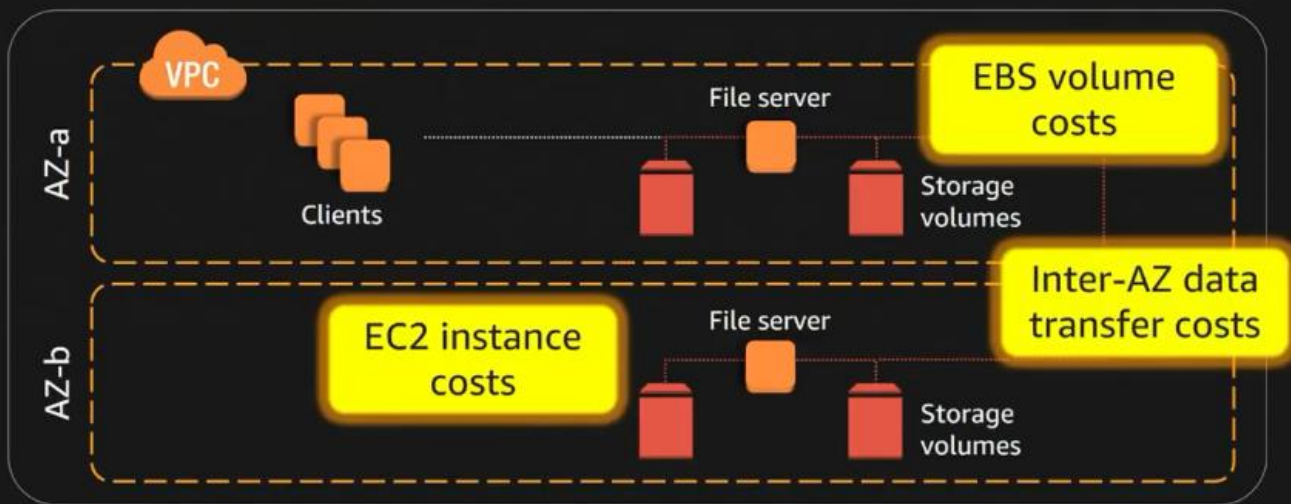
TCO example: For storing 500 GB, Amazon EFS is 70% less than DIY

EFS cost: (500GB * $0.30/GB-month) = <u>$150 per month</u>

For DIY, you might provision 600 GB of EBS (i.e., ~85% utilization):

Storage (2x 600 GB EBS gp2 volumes):	$120 per month
Compute (2x m4.xlarge instances):	$290 per month
Inter-AZ data transfer costs (est.):	$130 per month
Total	<u>$540 per month</u>

Where is EFS available today?

- US West (Oregon)
- US East (N. Virginia)
- US East (Ohio)
- EU (Ireland)
- EU (Frankfurt)
- Asia Pacific (Sydney)

More coming soon!

Complete set of data building blocks
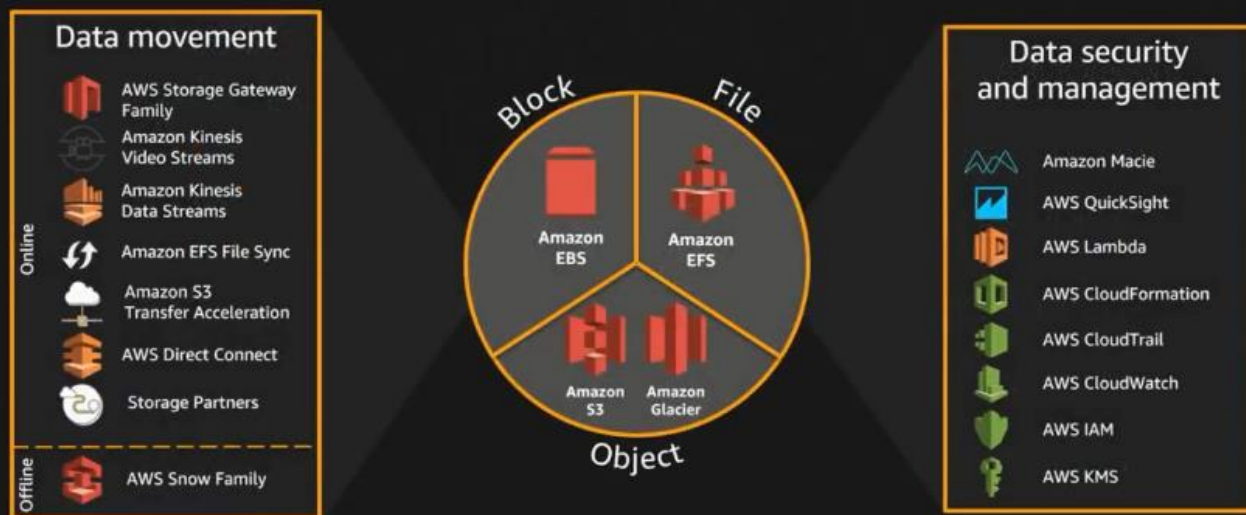
Data movement

Online
- AWS Storage Gateway Family
- Amazon Kinesis Video Streams
- Amazon Kinesis Data Streams
- Amazon EFS File Sync
- Amazon S3 Transfer Acceleration
- AWS Direct Connect
- Storage Partners

Offline
- AWS Snow Family

Block — Amazon EBS
File — Amazon EFS
Object — Amazon S3, Amazon Glacier

Data security and management
- Amazon Macie
- AWS QuickSight
- AWS Lambda
- AWS CloudFormation
- AWS CloudTrail
- AWS CloudWatch
- AWS IAM
- AWS KMS

Introducing EFS File Sync

Sync data from existing file systems into Amazon EFS file systems

Simple
Set up and manage easily
from the AWS Console

Fast
Up to 5x faster than
standard Linux copy tools

Secure
Encrypted parallel data
transfer to AWS

Use EFS File Sync to copy...

file systems from on-premises to EFS

DIY in-cloud file systems to EFS

EFS file systems between AWS Regions

aws

AWS re:Invent

Thomson Reuters Journey to EFS
Reuters.com Re-Hosting – Edward Zhitnitsky

aws

Reuters.com News, Financial Info, Video
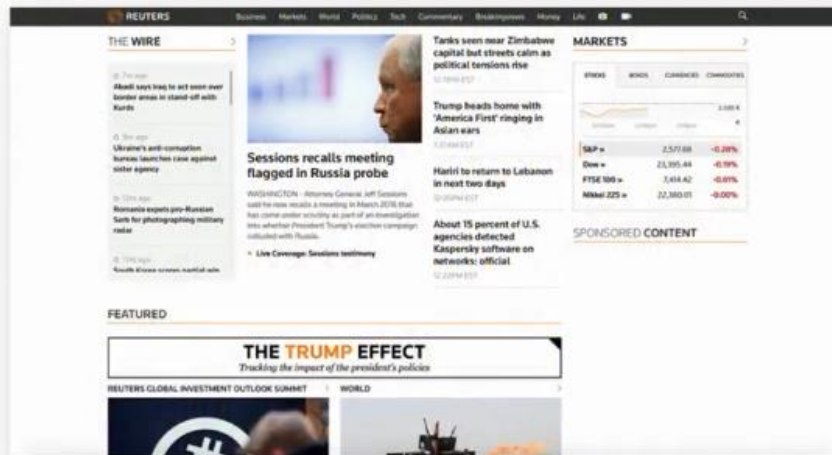
Reuters.com – Usage Metrics

- 221,710,480 total views/month
- 48,546,796 unique visitors/month
- 17 desktop editions
- 4 mobile editions
- News
    - Articles
    - Video
    - Pictures
- Financial data
    - Stocks
    - Company Look-Ups
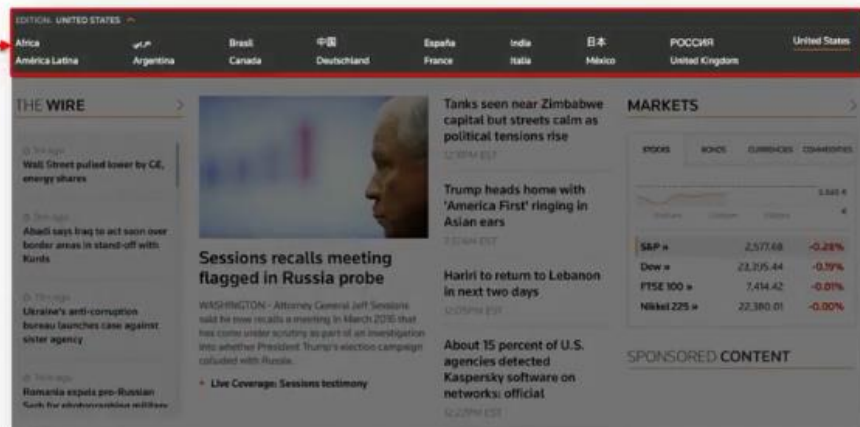
Previous Implementations

Previous Implementations
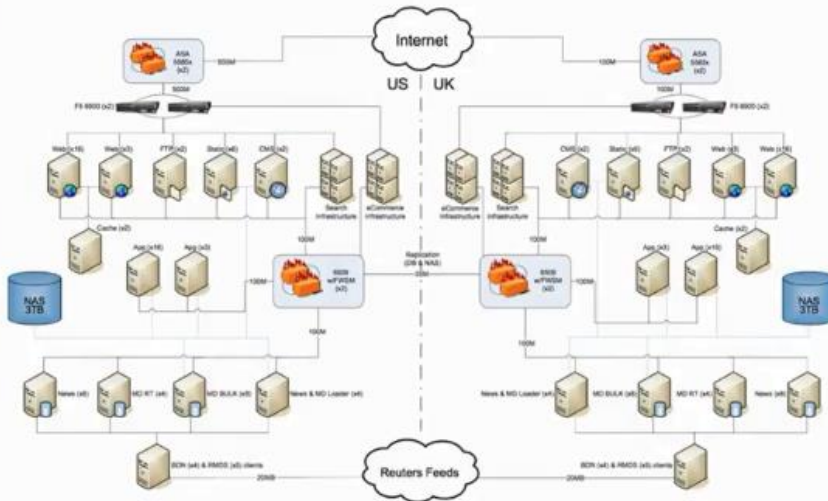
- Classic 3-tier architectures
- 120+ servers per environment
- GTMs
- Perimeter & Internal Firewalls
- 10 VLANs
- Load Balancers
- Edge Caching
- Master/replica news DB
- NAS/NFS for static content
- CMS
- Fixed assets require over provisioning

aws

Previous Implementations

- Aging architecture – 7+ years old

- Bandwidth commitments/constraints

- Performance

- Supportability

- Cost

The Amazon Journey

The Amazon Journey – 3 possible solutions

Existing Provider Proposal

- Higher costs, particularly to design and commission service

- Unproven tools

- Lock-in (sizing was locked into approximating existing footprint)

Internal Data Centers

- Data centers location

- Shared resource and support limitations

- Difficulty quantifying true costs

- Lead times to scale, augment, or limits on system access

Amazon

- Learning curve

- Application migration

- Most cost effective

- Capacity on demand

- Full system control

- Clarity in costs

Decision: We're moving to AWS!!!

Technologies Used
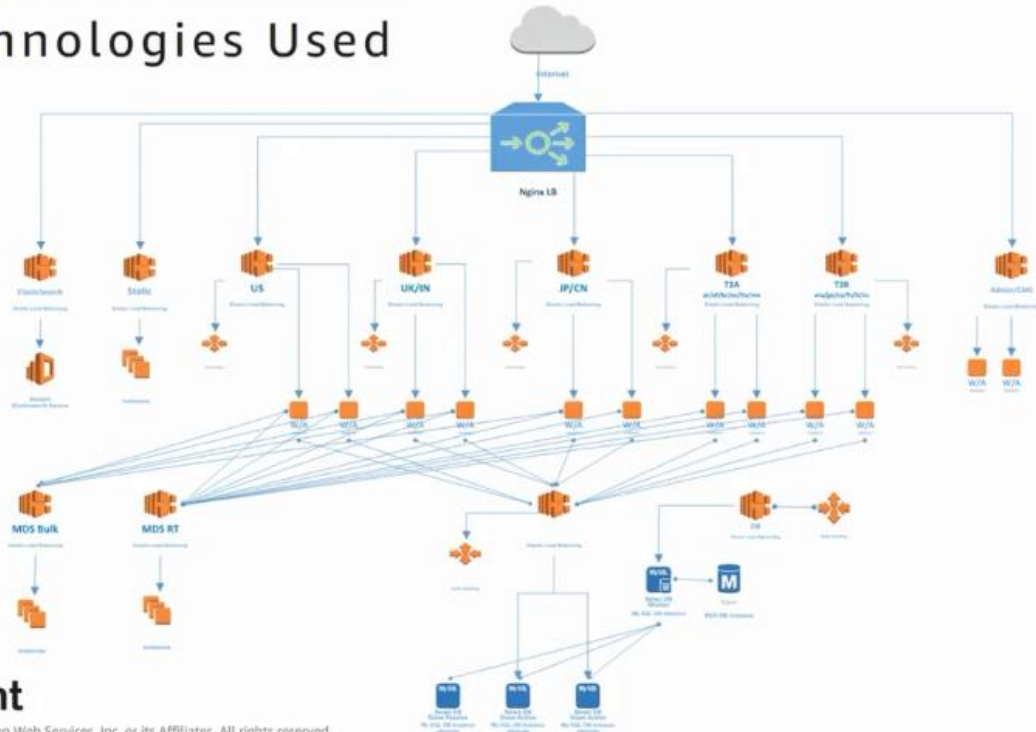
Technologies Used

Technologies Used

Amazon Route 53

Amazon CloudFront

AWS Certificate Manager

AWS WAF

Amazon S3

Amazon SQS

Amazon EFS

AWS CloudFormation

Amazon CloudWatch

Amazon ES

Elastic Load Balancing

AWS CloudTrail

Amazon RDS

openswan + AWS cloud
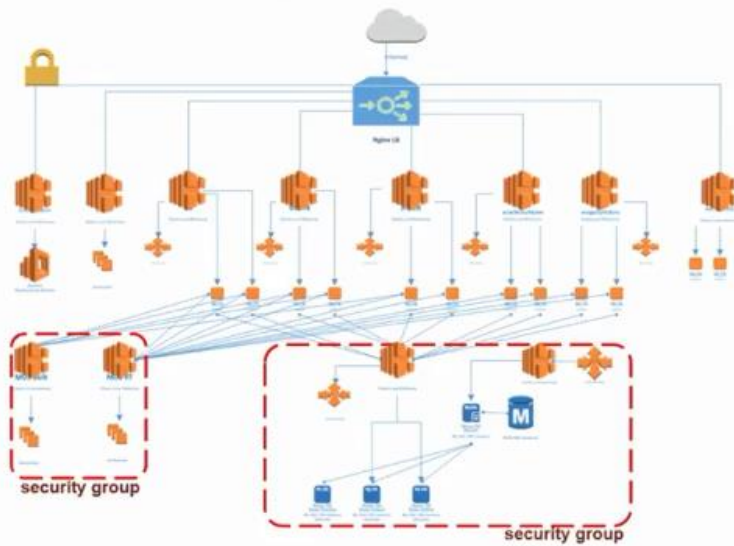
Technologies Used

Migration to Amazon EFS

- Great performance
- Redundancy
- Metadata protection
- No capacity limitations
- Easily mounted
- Cost savings

Technologies Used

Challenge: How to migrate existing content to EFS?

AWS Snowball

- Prompt delivery
- Data security
- Metadata protection

The Bottom Line

The Bottom Line
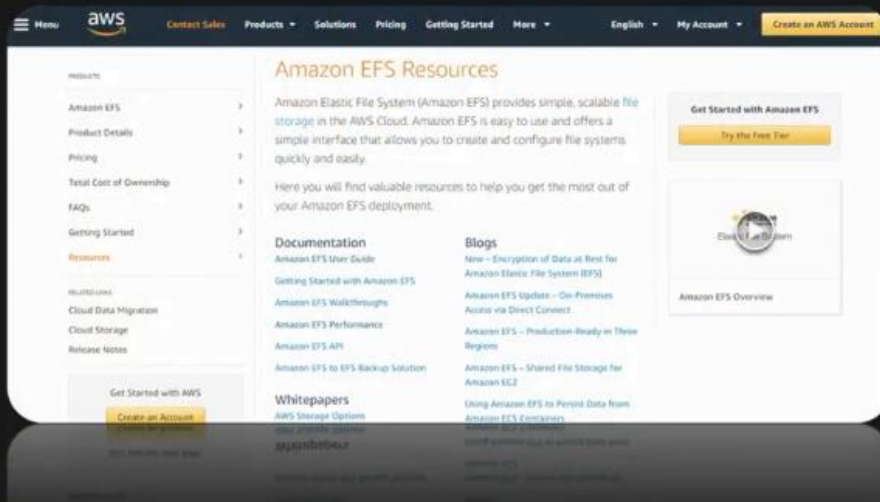
- Embraced new technology
- Using capacity on demand
- No more NAS Volume expansion
- Current hardware
- Bandwidth capacity eliminated
- Additional security added
  - DDoS protection
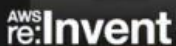 - DNS scaling protection
- 67% cost savings

How to learn more: aws.amazon.com/efs

Feature blogs

Whitepapers

Ref architectures

TCO calculator

10-minute tutorials

Documentation

re:Invent EFS Sessions

Monday, 11:30AM	STG314: Case Study: Learn How HERE Uses JFrog Artifactory with Amazon EFS
Tuesday, 12:15PM	STG401: This Is My Architecture - Lightning Round: Most Innovative Storage Solutions
Tuesday, 2:00PM	STG327: Chalk Talk: Learn How to Maximize Throughput to EFS by Using Parallel Data Access Methods
Tuesday, 2:30PM	STG324: Learn to Build a Cloud-Scale WordPress Site That Can Keep Up with Unpredictable Changes
Tuesday, 4:45PM	STG307: Deep Dive on Amazon Elastic File System (Amazon EFS)
Wednesday, 12:15PM	STG325: Case Study: Learn How SiriusXM & Digital ReLab Leveraged Amazon EFS for their Media Workflows
Wednesday, 4:00PM	STG310: Files in AWS: Overcoming Storage Challenges for Common File Use Cases - Special Guest Celgene
Thursday, 12:15PM	STG307-R: Deep Dive on Amazon Elastic File System (Amazon EFS) [REPEAT]
Thursday, 3:15PM	STG315: Case Study: Learn How Thomson Reuters Uses Amazon EFS to Deliver Billions of Pieces of Content
Thursday, 3:15PM	STG321: Amazon EFS: Leverage the Power of a Distributed Shared File System in the Cloud [WORKSHOP]
Thursday, 5:30PM	STG326: How to Build a Scalable & Highly Available Messaging Service with IBM & TIBCO Using Amazon EFS

New Storage Training

aws training and certification

For Enterprise Storage Engineers	Online at www.aws.training	At re:Invent
• Learn how to architect and manage highly available solutions on AWS storage services	• Access 100+ new digital training courses including advanced training on storage	• Visit Hands-on Labs at the Venetian
• Advance toward AWS certifications	• Deep Dives on S3, EFS, and EBS	• Attend a proctored "Introduction to EFS" Spotlight Lab on Thursday at 3pm at the Venetian
• Help your organization migrate to the cloud faster	• Migrating and Tiering Storage to AWS (Hybrid Solutions)	• Meet Storage experts at the Ask the Experts in Hands-on Labs room at the Venetian

AWS re:Invent

THANK YOU!