Accelerate Analytics & Data Science with Cloud-Native File Storage

Data science tools are critical to the wide array of analytics and machine learning services. Data science teams using frameworks such as Python, MXNet, Tensorflow, and Jupyter, can benefit from shared file storage to accelerate their analytics workflows.

- Gain faster data insights
- Eliminate capacity planning and improve utilization for better ROI
- Increase collaboration to iterate quickly

Why run data science workloads on AWS?

Amazon Elastic File System (Amazon EFS) enables greater collaboration with secure access to data across teams and geographies. Amazon EFS works seamlessly with AWS container services to provide persistent shared storage for your modern applications.

Elastic

- Grow and shrink file system capacity on demand
- No need to manage and provision infrastructure capacity
- Pay as you go and pay only for what you use

Highly Available and Durable

- Amazon EFS stores data across multiple Availability Zones (AZs)
- Available in every AWS Region
- 99.9% availability SLA
- Designed for 11 9's of durability

Scalable

- Scale to petabytes
- Performance modes for low latencies and maximum I/O
- Throughput that scales with storage

Secure

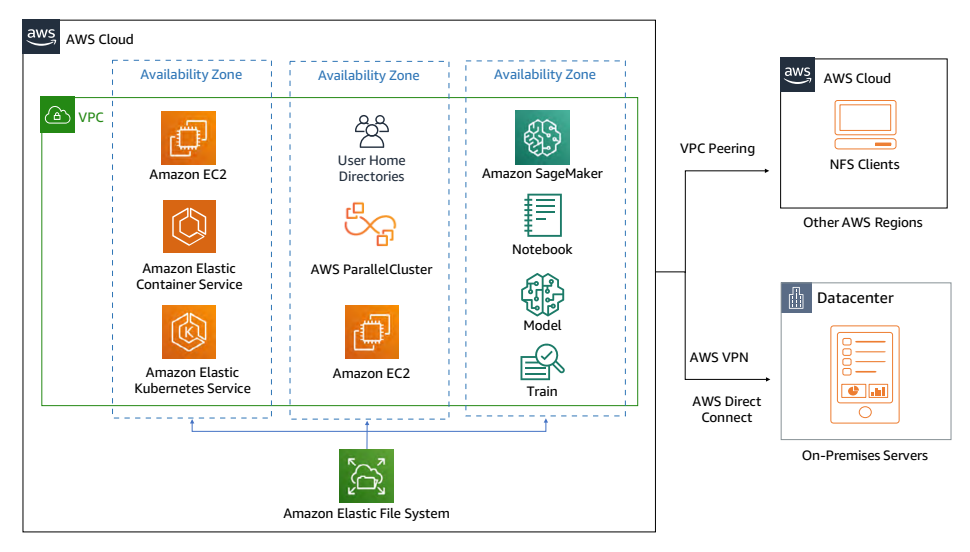
- Native cloud authentication and authorization with AWS IAM
- Manage multi-tenant access at scale with Amazon EFS Access Points
- Encryption at rest and in transit

Amazon EFS is architected for your data science requirements

Amazon EFS is a fully managed cloud native file service that delivers the simplicity and performance needed for your data science workloads. It is designed to provide massively parallel shared access to thousands of Amazon Elastic Compute Cloud (Amazon EC2) instances, enabling your applications to achieve high levels of aggregate throughput and IOPS with consistent low latencies. It provides a POSIX compliant interface that integrates with existing workflows.

Integration with native AWS services

Data science teams can use Amazon EFS to host user home directories, store and share notebook files, shared data sets, and model artifacts, and enable access from multiple AZs. Amazon SageMaker integrates with Amazon EFS allowing data scientists to iterate quickly and speed up machine learning training models.



Customer highlight

Bayer Crop Science, a leading sustainable-agriculture company, continuously strives to give its customers expert advice on where and when to plant its products. To do that, the organization supports a large geospatial-data platform containing sets of vector and raster data, soil and elevation layers, maps, and digital aerial imagery of farms. The company performs analytics on these data sets for internal data scientists and external business partners.

*"We can now run analytics at scale for our geospatial data platform using Amazon EFS. We can **provision compute and storage in seconds**, which means we can quickly expand the platform to meet the growing demand for analytics. Amazon EFS gives us a shared file system in the cloud, with **minimal setup and no management required**. We can also use the elasticity of the cloud to automatically scale up and down in support of our seasonal workloads."*

Vishnu Alavur Kannan, Analytical Technical Platform Lead, Bayer Crop Science

Additional resources

Learn more on our website: https://aws.amazon.com/efs/