**Lecture Notes for AWS Solution Architect**

**MODULE-6**

**Amazon Elastic File System (Amazon EFS)**

Amazon Elastic File System (Amazon EFS) provides a simple, serverless, set-and-forget, elastic file system that lets you share file data without provisioning or managing storage. It can be used with AWS Cloud services and on-premises resources, and is built to scale on demand to petabytes without disrupting applications. With Amazon EFS, you can grow and shrink your file systems automatically as you add and remove files, eliminating the need to provision and manage capacity to accommodate growth.

Amazon EFS offers you the choice of creating file systems using Standard or One Zone storage classes. Standard storage classes store data within and across multiple availability zones (AZ). One Zone storage classes store data redundantly within a single AZ, at a 47% lower price compared to file systems using Standard storage classes, for workloads that don’t require multi-AZ resilience.

Amazon EFS offers four storage classes: two Standard storage classes, Amazon EFS Standard and Amazon EFS Standard-Infrequent Access (EFS Standard-IA), and two One Zone storage classes, Amazon EFS One Zone, and Amazon EFS One Zone-Infrequent Access (EFS One Zone-IA).

While workload patterns vary, customers typically find that 80% of files are infrequently accessed (and suitable for infrequent access storage classes), and 20% are actively used (suitable for EFS Standard and EFS One Zone storage classes), resulting in an effective storage cost as low as $0.043/GB-month1 for file systems with One Zone storage classes, and $0.08/GB-month2 for file systems with Standard storage classes. Amazon EFS transparently serves files from both frequently accessed and infrequent accessed storage classes in a common file system namespace.

Amazon EFS is designed to provide massively parallel shared access to thousands of Amazon EC2 instances, and AWS containers and serverless compute services including Amazon Elastic Container Service (ECS), Amazon Elastic Kubernetes Service (EKS), AWS Fargate, and AWS Lambda, enabling your applications to achieve high levels of aggregate throughput and IOPS with consistent low latencies.

**Benefits**

1. **POSIX-compliant shared file storage**

Amazon EFS provides secure access for thousands of connections for Amazon EC2 instances and on-premises servers, as well as AWS compute services, including ECS, EKS, AWS Fargate, and AWS Lambda, simultaneously using a traditional file permissions model, file locking capabilities, and hierarchical directory structure via the NFSv4 protocol. Amazon EC2 instances can access your file system across AZs, regions, and VPCs, while on-premises servers can access using AWS Direct Connect or AWS VPN.

1. **Scalable performance**

Amazon EFS is designed to provide the throughput, IOPS, and low latency needed for Linux workloads. Throughput and IOPS scale as a file system grows and can burst to higher throughput levels for short periods of time to support the unpredictable performance needs of file workloads. For the most demanding workloads, Amazon EFS can support performance over 10 GB/sec and more than 500,000 IOPS.

1. **Dynamic elasticity**

Amazon EFS automatically and instantly scales your file system storage capacity up or down as you add or remove files without disrupting your applications, dynamically providing the storage capacity you need as you need it. You simply create your file system and start adding files with no need to provision storage in advance.

1. **Fully managed**

Amazon EFS is a fully managed service providing shared file system storage for Linux workloads. It provides a simple interface allowing you to create and configure file systems quickly and manages the file storage infrastructure for you, removing the complexity of deploying, patching, and maintaining the underpinnings of a file system.

1. **Cost-effective**

With Amazon EFS storage, you pay only for what you use. There is no need to provision storage in advance and there are no minimum commitments or up-front fees. With EFS Lifecycle Management, you can automatically move files accessed less frequently to a cost-optimized storage class, reducing file storage costs by up to 92%. With Amazon EFS One Zone storage classes, customers with workloads that do not require the multi-AZ resilience of Standard storage classes can choose additional savings of 47% by storing their data in a single-AZ. You can also use AWS Budgets to monitor your file system costs.

1. **Security and compliance**

Amazon EFS allows you to securely access your files using your existing security infrastructure. Control access to your Amazon EFS file systems with POSIX permissions, Amazon VPC, and AWS IAM. Secure your data by encrypting your data at rest and in transit. Amazon EFS also meets many eligibility and compliance requirements to help you meet your regulatory needs.



**Use cases**

1. **Containers and serverless persistent file storage**

Amazon EFS enables customers to persist data and state from their containers and serverless functions, providing fully managed, elastic, highly-available, scalable, and high-performance, cloud-native shared file systems. These same attributes are shared by Amazon Elastic Container Service (Amazon ECS), Amazon Kubernetes Service (Amazon EKS), AWS Fargate, and AWS Lambda, so developers don’t need to design for these traits, the services are simply ready for modern application development with data persistence. Amazon EFS allows data to be persisted separately from compute, and enables applications to have cross-AZ availability and durability. Amazon EFS provides a shared persistence layer that allows stateful applications to elastically scale up and down, such as for DevOps, web serving, web content systems, media processing, machine learning, analytics, search index, and stateful microservices applications.

1. **Move to managed file systems**

Amazon EFS provides the scalability, elasticity, availability, and durability to be the file store for enterprise applications and for applications delivered as a service. Its standard file system interface, file system permissions, and directory hierarchy make it easy to migrate enterprise applications from on-premises to the AWS cloud, and to build new ones. Move your business critical, Linux-based applications to managed file systems with Amazon EFS, while lowering your total cost of ownership (TCO).

1. **Analytics & machine learning**

Amazon EFS provides the ease of use, scale, performance, and consistency needed for machine learning and big data analytics workloads. Data scientists can use EFS to create personalized environments, with home directories storing notebook files, training data, and model artifacts. Amazon SageMaker integrates with EFS for training jobs, allowing data scientists to iterate quickly.

1. **Web serving & content management**

Amazon EFS provides a durable, high throughput file system for content management systems and web serving applications that store and serve information for a range of applications like websites, online publications, and archives. Since Amazon EFS adheres to the expected file system directory structure, file naming conventions, and permissions that web developers are accustomed to, it can easily integrate with web applications.

1. **Application testing & development**

Amazon EFS provides your development environments a common storage repository that gives you the ability to share code and other files in a secure and organized way. You can provision, duplicate, scale, or archive your test, development, and production environments with a few clicks, enabling your organization to be more agile and responsive to customer needs. Amazon EFS delivers a scalable and highly available solution that is ideal for testing and development workloads.

1. **Media & entertainment**

Media workflows like video editing, studio production, broadcast processing, sound design, and rendering often depend on shared storage to manipulate large files. Amazon EFS provides a strong data consistency model with high throughput and shared file access which can cut the time it takes to perform these jobs and consolidate multiple local file repositories into a single location for all users.

1. **Database backups**

Amazon EFS presents a standard file system that can be easily mounted with NFSv4 from database servers. This provides an ideal platform to create portable database backups using native application tools or enterprise backup applications. Many businesses want to take advantage of the flexibility of storing database backups in the cloud either for temporary protection during updates or for development and test.

**Amazon FSx**

Amazon FSx makes it easy and cost effective to launch and run popular file systems that are fully managed by AWS. With Amazon FSx, you can leverage the rich feature sets and fast performance of widely-used open source and commercially-licensed file systems, while avoiding time-consuming administrative tasks such as hardware provisioning, software configuration, patching, and backups. It provides cost-efficient capacity with high levels of reliability, and integrates with a broad portfolio of AWS services to enable faster innovation.

Amazon FSx provides two file systems to choose from: Amazon FSx for Windows File Server for business applications and Amazon FSx for Lustre for compute-intensive workloads.

**Amazon FSx for Windows File Server**

Amazon FSx for Windows File Server provides fully managed file storage that is accessible over the industry-standard Server Message Block (SMB) protocol. Built on Windows Server, Amazon FSx delivers a wide range of administrative features such as data deduplication, end-user file restore, and Microsoft Active Directory (AD) integration. It offers single-AZ and multi-AZ deployment options, fully managed backups, and encryption of data at rest and in transit. With the HDD storage option, Amazon FSx for Windows File Server offers the lowest-cost file storage in the cloud for Windows applications and workloads. You can access Amazon FSx from Windows, Linux, or MacOS compute instances and devices running on AWS or on premises. You can optimize cost and performance with multiple storage options, and you can grow your file system storage and scale your performance at any time.

**Amazon FSx for Lustre**

Amazon FSx for Lustre is a fully managed service that provides cost-effective, high-performance, scalable storage for compute workloads. Many workloads such as machine learning, high-performance computing (HPC), video rendering, and financial simulations depend on compute instances accessing the same set of data through high-performance shared storage. Powered by Lustre, the world’s most popular high-performance file system, FSx for Lustre offers shared storage with low latencies, up to hundreds of gigabytes per second of throughput, and millions of IOPS. FSx for Lustre offers multiple deployment types, storage types, and throughput performance levels to optimize cost and performance for your workload requirements. FSx for Lustre file systems can also be linked to Amazon S3 buckets, allowing you to access and process data concurrently from both a high-performance file system and from the S3 API.

**Benefits**

1. **Cost-effective**

With Amazon FSx, you can spin up and spin down file systems on-demand. Amazon FSx offers a range of Solid-State Disk (SSD) and Hard Disk Drive (HDD) storage options enabling you to optimize storage price and performance for your workload requirements. Amazon FSx for Windows File Server offers data deduplication to optimize costs by removing redundant data. With FSx for Lustre, you can choose unreplicated, scratch file systems to further lower costs for shorter-term processing of data.

1. **Fast and flexible performance**

Amazon FSx supports the most demanding enterprise applications and high-performance workloads. Amazon FSx is designed to deliver fast, predictable, scalable, and consistent performance. Amazon FSx file systems deliver sustained high read and write speeds and consistent low latency data access. You can choose the storage type and throughput level to best match your application's requirements. Storage capacity can be increased at any time to respond to changing business needs.

1. **Available and durable**

Amazon FSx runs on the same highly reliable infrastructure used by other Amazon Web Services, and it offers a variety of deployment options to match your workload’s availability and durability requirements. With FSx for Windows File Server, you can choose from single-AZ or multi-AZ deployment types, depending on your applications needs. With FSx for Lustre, you can choose from scratch or persistent storage for short-term or long-term data processing.

1. **Simple and fully managed**

In minutes and with a few clicks, you can launch a fully managed, highly reliable file system. With Amazon FSx file systems, you don't have to worry about managing file servers and storage volumes, updating hardware, configuring software, monitoring storage consumption, or performing backups -- Amazon FSx automates these time-consuming administration tasks.

1. **Secure and compliant**

Amazon FSx automatically encrypts your data at-rest and in-transit. To control network access to your file system, Amazon FSx lets you run your file systems in an Amazon Virtual Private Cloud (Amazon VPC). Amazon FSx is designed to meet the highest security standards and has been assessed to comply with ISO, PCI-DSS, and SOC certifications, and is HIPAA eligible. Integration with AWS Backup enables you to meet your data protection requirements with centralized backup and compliance management for your Amazon FSx file systems.

1. **Integrated with AWS services**

Amazon FSx file systems integrate with other AWS services, including Amazon S3, Amazon CloudWatch, Amazon CloudTrail, AWS KMS, Amazon SageMaker, Amazon WorkSpaces, Amazon AppStream 2.0, Amazon Elastic Container Service (Amazon ECS), Amazon Elastic Kubernetes Service (Amazon EKS), AWS Batch, and AWS ParallelCluster. Using Amazon FSx with these services allows you to more easily develop, deploy, and run your workloads.

**Create Amazon EFS file system**

In this step, you create your Amazon EFS file system that has the service recommended settings using the Amazon EFS console.

**To create your Amazon EFS file system**

1. Open the Amazon EFS Management Console at https://console.aws.amazon.com/efs/.
2. Choose **Create file system** to open the **Create file system** dialog box.



1. (Optional) Enter a **Name** for your file system.
2. For **Virtual Private Cloud (VPC)**, choose your VPC, or keep it set to your default VPC.
3. For **Availability and Durability**, choose one of the following:
* **Regional** to create a file system that uses Standard storage classes. Standard storage classes store file system data and metadata redundantly across all Availability Zones within an AWS Region. **Regional** offers the highest levels of availability and durability.
* **One Zone** to create a file system that uses One Zone storage classes. One Zone storage classes store file sytem data and metadata redundantly within a single Availability Zone which makes it less expensive than Standard storage classes.

Because EFS One Zone storage classes store data in a single AWS Availability Zone, data stored in these storage classes may be lost in the event of a disaster or other fault that affects all copies of the data within the Availability Zone, or in the event of Availability Zone destruction resulting from disasters, such as earthquakes and floods.

1. Choose **Create** to create a file system that uses the following service recommended settings:
	* Automatic backups turned on
	* Mount targets – Amazon EFS creates mount targets with the following settings:
		+ For file systems that use Standard storage classes, a mount target is created in each Availability Zone in the AWS Region in which the file system is created. For file systems that use One Zone storage classes, a single mount target is created in the Availability Zone you specified.
		+ Located in the default subnets of the VPC you selected.
		+ Using the VPC's default security group – You can manage security groups after the file system is the created.
	* General Purpose performance mode
	* Bursting throughput mode
	* Encryption of data at rest enabled using your default key for Amazon EFS (aws/elasticfilesystem).
	* Lifecycle management enabled with a 30-day policy

After you create the file system, you can customize the file system's settings with the exception of availability and durability, encryption, and performance mode.

If you want to create a file system with a customized configuration, choose **Customize**. For more information about creating a file system with customized settings

1. The **File systems** page appears with a banner across the top showing the status of the file system you created. A link to access the file system details page appears in the banner when the file system becomes available.

