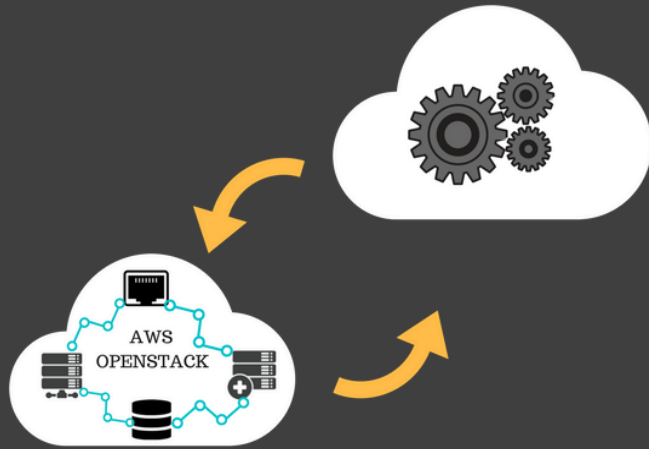


# AWS Design and Automation

## Module 1: Design Storage in AWS



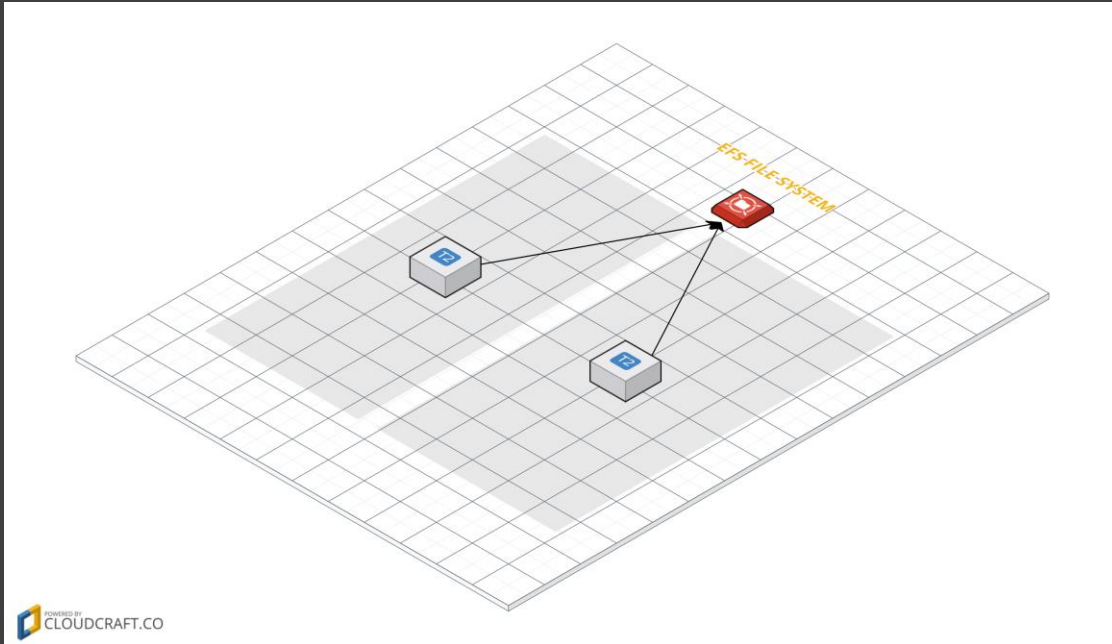
## Topic 3: Elastic File System

Mohanraj Shanmugam

# Elastic File System

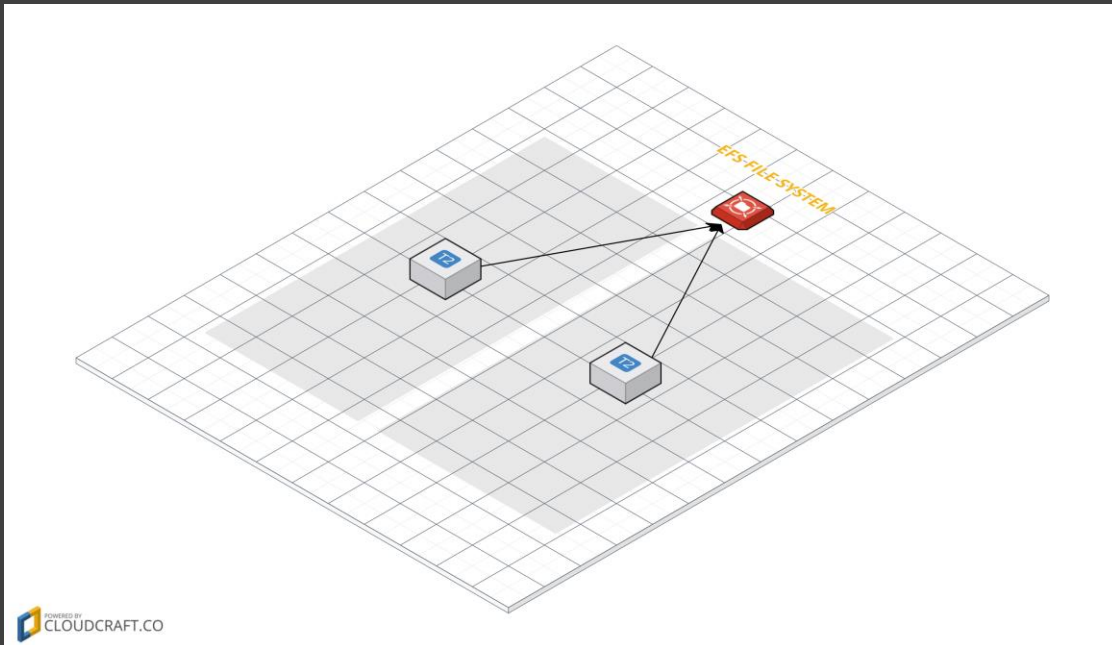
Mohanraj Shanmugam

# Elastic File System



- Amazon Elastic File System (Amazon EFS) provides simple, scalable file storage for use with Amazon EC2 instances in the AWS Cloud.
- Storage capacity is elastic, growing and shrinking automatically as you add and remove files, so your applications have the storage they need, when they need it.
- Multiple Amazon EC2 instances can access an Amazon EFS file system at the same time, allowing Amazon EFS to provide a common data source for workloads and applications running on more than one Amazon EC2 instance.

# Elastic File System



- Amazon EFS supports the Network File System version 4.1 (NFSv4.1) protocol, so the applications and tools that you use today work seamlessly with Amazon EFS.
- Multiple Amazon EC2 instances can access an Amazon EFS file system at the same time, providing a common data source for workloads and applications running on more than one instance.
- Amazon EFS file systems store data and metadata across multiple Availability Zones in a region and can grow to petabyte scale, drive high levels of throughput, and allow massively parallel access from Amazon EC2 instances to your data.

# Performance

- Amazon EFS offers two performance modes.
- The two performance modes have no additional costs, so your Amazon EFS file system is billed and metered the same, regardless of your performance mode.
  - **General Purpose Performance Mode**
  - **Max I/O Performance Mode**

# General Purpose Performance Mode

- Majority of your Amazon EFS file systems use General Purpose performance mode.
- General Purpose is ideal for latency-sensitive use cases, like web serving environments, content management systems, home directories, and general file serving.
- Amazon EFS selects the General Purpose mode for you by default.

# Max I/O Performance Mode

- File systems in the Max I/O mode can scale to higher levels of aggregate throughput and operations per second with a tradeoff of slightly higher latencies for file operations.
- Highly parallelized applications and workloads, such as big data analysis, media processing, and genomics analysis, can benefit from this mode.

# Throughput Scaling in Amazon EFS

- Throughput on Amazon EFS scales as a file system grows.
- Because file-based workloads are typically spiky—driving high levels of throughput for short periods of time, and low levels of throughput the rest of the time
- Amazon EFS is designed to burst to high throughput levels for periods of time.



# Throughput Scaling in Amazon EFS

File System Size (GiB)	Baseline Aggregate Throughput (MiB/s)	Burst Aggregate Throughput (MiB/s)	Maximum Burst Duration (Min/Day)	% of Time File System Can Burst (Per Day)
10	0.5	100	7.2	0.5%
256	12.5	100	180	12.5%
512	25.0	100	360	25.0%
1024	50.0	100	720	50.0%
1536	75.0	150	720	50.0%
2048	100.0	200	720	50.0%
3072	150.0	300	720	50.0%
4096	200.0	400	720	50.0%

# Managing Burst Credits

- When a file system has a positive burst credit balance, it can burst. You can see the burst credit balance for a file system by viewing the `BurstCreditBalance` Amazon CloudWatch metric for Amazon EFS.
- The bursting capability (both in terms of length of time and burst rate) of a file system is directly related to its size.
- Larger file systems can burst at larger rates for longer periods of time

# Amazon EFS Performance Tips

- **Average I/O Size**

- Amazon EFS's distributed nature enables high levels of availability, durability, and scalability.
- This distributed architecture results in a small latency overhead for each file operation.
- Due to this per-operation latency, overall throughput generally increases as the average I/O size increases, because the overhead is amortized over a larger amount of data.

# Amazon EFS Performance Tips

- **Simultaneous Connections –**

- Amazon EFS file systems can be mounted on up to thousands of Amazon EC2 instances concurrently.
- If your application is parallelizable across more instances, you can drive higher throughput levels on your file system in aggregate across instances.

- **Request Model –**

- By enabling asynchronous writes to your file system, pending write operations are buffered on the Amazon EC2 instance before they are written to Amazon EFS asynchronously.
- Asynchronous writes typically have lower latencies.
- When performing asynchronous writes, the kernel uses additional memory for caching. A
- file system that has enabled synchronous writes, or one that opens files using an option that bypasses the cache (for example, `O_DIRECT`), will issue synchronous requests to Amazon EFS and every operation will go through a round trip between the client and Amazon EFS.

# Amazon EFS Performance Tips

- **NFS Client Mount Settings –**

- Use the recommended mount options Recommended by amazon.
- Amazon EFS supports the Network File System version 4.1 (NFSv4.1) and NFSv4.0 protocols when mounting your file systems on Amazon EC2 instances. NFSv4.1 provides better performance.

- **Amazon EC2 Instances –**

- Applications that perform a large number of read and write operations likely need more memory or computing capacity than applications that don't.
- When launching your Amazon EC2 instances, choose instance types that have the amount of these resources that your application needs.
- Note that the performance characteristics of Amazon EFS file systems are not dependent on the use of EBS-optimized instances.

# Amazon EFS Use Cases

- **Big Data and Analytics**
- Amazon EFS provides the scale and performance required for big data applications that require high throughput to compute nodes coupled with read-after-write consistency and low-latency file operations.
- **Media Processing Workflows**
- Media workflows like video editing, studio production, broadcast processing, sound design, and rendering often depend on shared storage to manipulate large files. A strong data consistency model with high throughput and shared file access can cut the time it takes to perform these jobs and consolidate multiple local file repositories into a single location for all users.

# Amazon EFS Use Cases

- **Content Management and Web Serving**
  - Amazon EFS provides a durable, high throughput file system for content management systems that store and serve information for a range of applications like websites, online publications, and archives.
- **Home Directories**
  - Amazon EFS can provide storage for organizations that have many users that need to access and share common data sets. An administrator can use Amazon EFS to create a file system accessible to people across an organization and establish permissions for users and groups at the file or directory level.

# EFS Pricing

US East (N. Virginia)	\$0.30/GB-month
US East (Ohio)	\$0.30/GB-month
US West (Oregon)	\$0.30/GB-month
EU (Ireland)	\$0.33/GB-month