Elastic File System

25 June 2023   14:27

Amazon Elastic File System (EFS) is a scalable cloud file storage solution for use with EC2 instances. It's elastic because it will automatically **grow and shrink** as you add/remove files. It has a simple interface that enables you to create and configure file systems quickly and simply. It's similar to EBS, but with EBS you can only mount your virtual disk to one EC2 instance. You can have 2 instances sharing an EFS volume.

Key Features:
- Supports NFSv4 Protocol* (specifically v4.0 and v4.1)
- Allows for multiple EC2 instance access at the same time
- Common data source for workloads and apps running on more than one instance or server
- Offers file access semantics such as strong data consistency and file locking
- Highly scalable (can grow to PB scale), highly available, and highly durable
- Allows for high levels of throughput
- Offers two storage classes: Standard and Infrequent Access
- You only pay for storage used by your file system
- Control access to file systems using Portable Operating System Interface (POSIX) permissions
- EFS-to-EFS backup system allows you to automatically schedule incremental backups of your EFS file system
- AWS DataSync — Allows you to move and manage EFS file data that makes it faster and easier to move data between on-prem storage and EFS

How it Works

EFS provides file storage within the cloud. You create the EFS file system, mount it on an EC2 instance, and then you're able to read/write data to/from your system. It's also available in your VPC through NFSv4 protocol.

Make sure to select an AMI that supports the NFSv4 protocol when setting up your EC2 instance (Linux current gen NFSv4.1 client recommended). For select AMIs, you'll need to actually install a NFS client to mount your file system on your instance. EFS allows concurrent access from EC2 instances in your VPC allowing apps to scale beyond a single connection.

- You can mount EFS on instances in only one VPC at a time

EFS uses **mount targets** which allows you to access EFS in a VPC. This provides an IP address for a NFSv4 endpoint where you can mount EFS. You mount it by using its DNS name. This resolves to the IP address of the EFS mount target in the **same** availability zone as your EC2 instance.

You can create one mount target in each availability zone in an AWS region. If there are multiple subnets in an availability zone in your VPC, you create a mount target in one of the available subnets. All EC2 instances in that AZ will share that mount target from there.

EFS can be mounted on an EC2 instance or on-premises through an AWS Direct Connect connection.
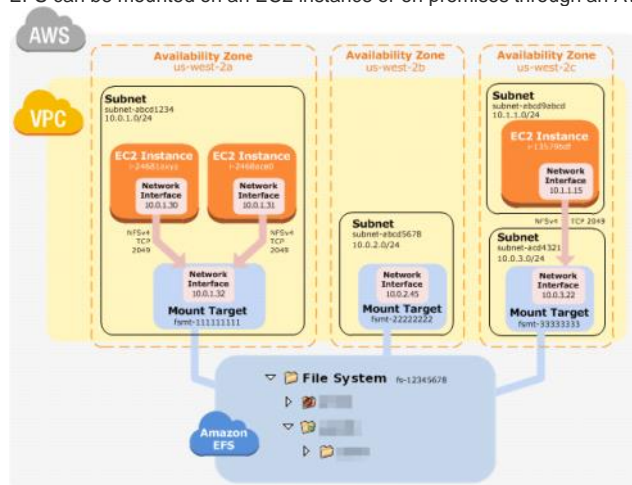
EFS Documentation

The above architecture is an example of how mount targets should be implemented in each availability zone.

EFS Performance Modes

EFS performance modes come in two different flavors. You select a performance mode when you create one. These come at no additional cost, EFS is billed the same

regardless of modes.

General Purpose Performance Mode
- This is the default mode and it's best used for latency-sensitive use cases.
- Use cases include web serving environments, content management systems, home directories, and general file serving.

Max I/O Performance Mode
- This is used to scale to higher levels of aggregate throughput and operations per second. The trade-off is slightly higher latencies for file operations.
- Use cases include highly parallelized apps and workloads like big data analysis, genomics analysis, and media processing.

Throughput Modes
You get two flavors of throughput modes for EFS:

Bursting Throughput
- This is the default mode
- Throughput on EFS scales as your file system grows
- File-based workloads are typically spiky, hence 'bursting' throughput to adequately accommodate the spikes

Provisioned Throughput
- This mode is where you specify the throughput of your file system independent of the amount of data stored

Limitations
EFS is limited to 1,000 file systems per AWS Region. You can submit a request to Amazon to increase this limit.

Resource	Limit
Number of mount targets for each file system in an Availability Zone	1
Number of mount targets for each VPC	400
Number of security groups for each mount target	5
Number of tags for each file system	50
Number of VPCs for each file system	1

Steps : -

Create Elastic File System in AWS
Create 3 ec2-linux instances and run following commands

sudo su
yum install -y amazon-efs-utils
mkdir "efs" directory in all instances

Add Security group of Ec2 instance to EFS Network Security Group

Add NFS 2049 to same security group.
Mount step (Copy step from EFS screen)
echo "While I thought that I was learning how to live, I have been learning how to die." > quote1.txt (cat test.txt and more test.txt to check file content)