Elastic File System

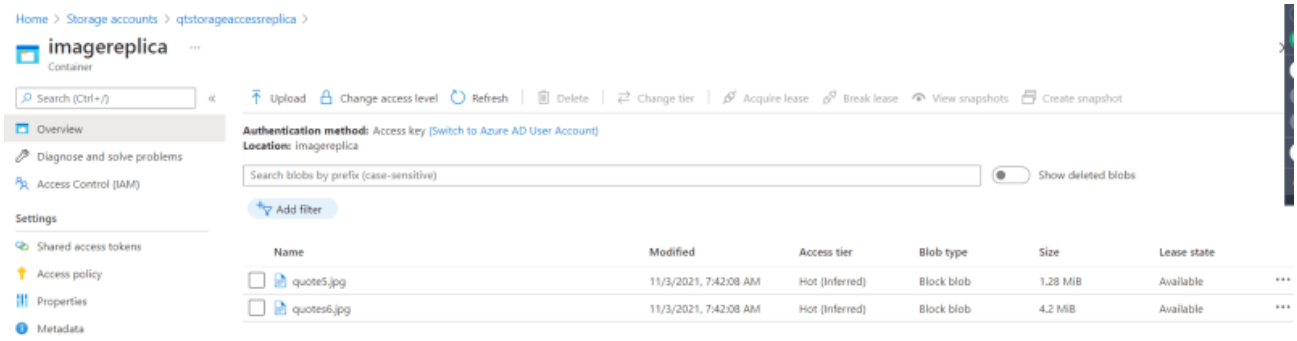
9. Elastic File System

- EFS provides simple, elastic file system for use with Cloud Services and on-premises solutions
- EFS supports NFS(Network File system) version 4 protocol.
- With EFS you pay only for the storage used by your file system
- Create an EFS in a region with selected zones.
- EFS is a network file system, so while creating EFS, We need to provide the subnet ids and security group
- Amazon Elastic File System (Amazon EFS) is a simple, server-less, set-and-forget, elastic file system.
- EFS is a user-based encryption control technique that enables users to control who can read the files on their system. The typical method of using EFS is to perform encryption at the folder level. This ensures that all files added to the encrypted folder are automatically encrypted.
- There is no minimum fee or setup charge.
- You pay only for the storage you use, for read and write access to data stored in Infrequent Access storage classes, and for any provisioned throughput.
- EBS is a high-performance per-instance block storage system designed to act as storage for a single EC2 instance (most of the time) EFS is a highly scalable file storage system designed to provide flexible storage for multiple EC2 instances.
- You can use an EFS file system as a common data source for workloads and applications running on multiple instances.
- EFS provides a file system interface and file system semantics to AWS EC2 instances. When EFS is attached to EC2 instances, it acts just like a local file system but EFS is a shared file system which means you can mount the same file system across multiple EC2 instances with low latencies.

Attributes of EFS:

- Fully managed: We don’t need to maintain any hardware or software
- File System access semantics: We get what we expect from a regular file system including read-after-write consistency locking.
- File System Interface: It exposes file system interface that works with standard os APIs. EFS appears to be like any other file system to your OS
- Shared Storage:
- Elastic and scalable: EFS elastically grows to petabyte scale. We don’t have to specify a provisioned size up front, We just create a file system, and it grows and shrinks automatically as we add/remove data
- Performance: It is built for performance across wide variety of workloads. It provides consistent, low latencies, high throughput and high IOPS
- High Available & Durable: The data in EFS is automatically replicated across AZs with in a region.
- EFS also works with NFS protocol, Using Direct connect and VPC, we can also mount EFS on your on-premises server via the NFS 4.1 protocol. This a great use case for transferring a large number of data from the servers running on-premise to AWS cloud
- Performance Mode of EFS
- General Purpose Mode:
- This is default mode for EFS
- This is optimized for latency-sensitive applications & general purpose file-based workloads
- This mode is best option for majority of the use cases
- Max I/O mode
- This is optimized for large-scale and data-heavy applications, where tens, hundreds or thousands of EC2 instances are accessing the file system
- This scales to higher level aggregate throughput and operations per second with a trade-off of slightly higher latencies for file operations

TechData-Infinity-Devops with MultiCloud



- EFS mounts don't work on Windows Servers so AWS has a service called FSx

Understand General Purpose Volumes-

- Here lets focus on GP2 which are SSD-backed, low cost and default in all regions.
- Lets look at the documented performance characteristics of a GP2 volume
- Max IOPS/Volume => 16000
- Max Throughput/Volume => 250 MB/s
- Max Burstable IOPS => 3000
- The actual max IOPS for any specific volume is based on size 3 IOPS/GB i.e only volume or size 5.3 TB or larger can hit a max IOPS
- Baseline IOPS and Burst Limits
- Lets consider a smaller volume 100 GB GP2 Volume, at this size the volume has baseline performance of 300 IOPS, but it can burst upto 3000 IOPS using a bucket credit system

Throughput:

- IOPS is a tricky metric and you typically care more about throughput-speed at which data comes on and off the volume
- GP2 Volumes have a throughput limit based on size
  |Volume Size| Throughput Limit | |< 170 GB| 128 MB/s | |170-334 GB| 250 MB/s (burst) | |>334 GB| 250 MB/s |
- So take, use GP3 volumes with higher throughputs to save costs rather than provisioned IOPS

Advantages-

- EFS can be used when our application needs a shared file system, which can be accessed by multiple instances at the same time.
- Reliable
- Pay as you go (cost effective)
- Performance that scales to support any workload.
- Dynamic elasticity
- Fully managed by AWS.
- Security and compliance.

Disadvantages-

- No Windows instances. Amazon EFSs are not supported on AWS Windows EC2 instances. EFS volumes can only be used with non-Windows instances, such as Linux, that support NFS volumes.
- No system boot volumes. Amazon EFS volumes also cannot be used for system boot volumes.

Use Cases-

- Web serving and content management.
- Application development and testing.
- Media and entertainment.
- Database backups.

NFS:

- Network File System is a distributed file system protocol originally developed by Sun Microsystems in 1984, allowing a user on a client computer to access files over a computer network much like local storage is accessed.