S3 Cloud Disk Best Practices

Without proper configuration, a SoftNAS instance leveraging S3-compatible cloud disk extenders can perform poorly. To get the best performance possible for a SoftNAS deployment with S3-compatible cloud disks, keep in mind the following:

Sizing

Sizing a solution involving use of Cloud Disk Extenders is very much the same as for a solution making use of a block-based implementation (VMDK or EBS). There is no change to storage space requirements. However, additional system resources may be required in order to handle the virtualization of the S3 storage required in order to present the S3 Cloud Disk as block storage. Stated another way, the number of buckets that are configured via cloud disk extender influences the amount of additional resources that are required to access the same overall capacity of storage.

CPU

If using cloud disk extenders in your instance/s, it is important to configure your instance with additional processing power (CPU), above and beyond what is required for traditional block-based storage access. Presenting S3 storage as block-based storage requires a number of additional functions to be executed, including, for example, SSL/TLS key exchange and encryption, MD5 block computations, network stack processing, as well as optional encryption options. To avoid performance issues:

- Do not use cloud disk extender on single vCPU instances.
- 2 vCPU instances may be suitable for test scenarios. Two vCPU instances may still prove insufficient if your S3-compatible test/POC environment requires decent performance metrics.
- For a production environment, a minimum of 4 vCPU instances is highly recommended. Many workloads will perform better with additional vCPU.
- For each 75 MB/s of throughput required to perform the same task with block-based storage, an additional two vCPU is highly recommended.
- CPU utilization should be monitored during proof-of-concept and initial production stages to verify that sufficient CPU has been provisioned for the provided workload.
- Monit email alerts should be monitored and indications of high CPU utilization should be reviewed with respect to the Cloud Disk Extender configuration.
- If operating in a trusted environment, and available as an option for the S3-compatible object storage being used, CPU usage can be reduced by using http rather than https.
- CPU usage can be further reduced by disabling optional encryption options.

Example:

A customer wants to use S3 object storage to save money over EBS. The current workload operates between 100-150MB/s of throughput and is running on an m4.xlarge instance. Evaluating the current workload, we know that it averages a healthy 50% CPU usage. To provide the same 150MB/s of S3 throughput, the general guideline requests 4 additional vCPU over and above the current instance's existing 4 vCPU base. As a result, the CPU recommendation points to an m4.2xlarge instance, in order to provide four additional vCPU.

RAM

As mentioned previously in this document, each instance of the cloud disk extender represents a process that is running inside of the SoftNAS instance for virtualizing the object storage as block storage.

- Cloud Disk Extender should not be used in production on systems with less than 8GB of RAM.
- Memory footprints less than 8GB of RAM may be suitable for test or PoC environments only.
- A general guideline of 512MB of RAM should be provisioned above the normal required memory for a given workload.
- Remember that half of the RAM is utilized for file-system caching. Additional resources are needed for the network file services and the base operating environment (~2GB of RAM).

Network

Cloud Disk Extender utilizes the network interface of an instance in order to access the object storage. Sufficient network bandwidth must be provisioned in order to reach maximum performance profiles using Cloud Disk Extender. When considering the desired available throughput to the object store also consider the amount of network throughput for network file services (NFS, CIFS, iSCSI, AFP) and SnapReplicate/SNAP HA^{TM} which, in most configurations and platforms, all come from the same pool of available network bandwidth.

- A somewhat safe calculation can be to determine the available network throughput being used for the instance, and to divide it divided by 3, in order tocalculate 1/3 for file services, 1/3 for replication, and 1/3 for object storage I/O.
- When calculating, consider that SnapReplicate only replicates the write bandwidth, not the read bandwidth.
- Be sure to convert properly between bits and bytes when comparing network throughput (usually expressed in bits) to disk throughput (usually expressed in bytes)
- There is inherent overhead in the protocols used on the network (request/response, headers, checksums, control data, etc) such that full network saturation does not yield the full bandwidth as useful throughput. Consider only anticipating 90% of the link-speed as usable throughput.
- Most clouds (and most data centers) do not provide full link-speed bandwidth on a sustained basis as systems are utilizing shared resources. Systems designed to run at full provisioned capacity (of any metric) should be assigned to dedicated hosts rather than shared tenancy.

Example:

A customer uses NFS, SnapReplicate and SNAP HA^{TM}, and would like to use object storage. Expected throughput is about 40MB/s with 90% reads. According to calculation, the network throughput for the source node reads as follows:

- 4MB/s writes to NFS (incoming)
- 36MB/s reads to NFS (outgoing)

- 4MB/s writes to SnapReplicate (outgoing)
- 4MB/s writes to Object Storage (outgoing)
- 36MB/s reads to Object Storage (incoming)

Total: 40MB/s incoming 44MB/s outgoing

Calculating the total throughput in bytes, this is 320mbps incoming and 352mbps outgoing.

According to calculation, the network throughput for the target node reads as follows:

- 4MB/s writes from SnapReplicate (incoming)
- 4MB/s writes to Object Storage (outgoing)

Total: 4MB/S incoming and 4MB/S outgoing

In bytes, this works out to 32mbps incoming 32mbps outgoing.

A 100 mbps network connection is certainly not sufficient for this configuration, however, a 1gbps connection should be enough, even considering protocol overhead and avoiding 100% saturation of the network.

Amazon AWS S3 Recommendation: VPC Endpoints

Customers on AWS within a VPC should be using VPC Endpoints for accessing S3 object stores. By using a VPC endpoint, a higher quality service level is provided to S3 object stores within a region, thereby improving the overall reliability and performance when accessing S3 object storage. Additionally, a VPC Endpoint can be used in order to to communicate with resources in other services via private IPs, without exposing instances to the internet.

For guidance on setting up VPC Endpoints via the Amazon AWS console, see **Amazon's help on the topic**.