**Lab 04: Replication and Deduplication**

In this lab, we setup and configured data replication and deduplication as a means of data protection and data recovery. These are useful in disaster-recovery situations.

**Replication**

Replication, as the name implies, is an exact copy of data. It can be point-in-time, meaning an administrator can restore or recover a set of data from a time in the past, or continuous, where a copy is saved as soon as any changes are made to the data.

Logging onto NAS01, a snapshot of the storage system was taken. The ZFS replication public key information was copied. Logging onto NAS02, the key information was copied into the SSH public key textbox for the root user. On NAS01, a replication task was added. It was confirmed that the SSH key scan automatically populated the NAS02 SSH key.

You can view the progress of the replication by navigating to View Volumes, and under the Storage tab you will be able to observe the available space and size of the volume. Note that the replication will initialise NAS02, and wipe all existing data. There are also graphs available in the Reporting tab to view the progress. Once the replication has completed, it will be listed under ZFS Snapshots on NAS02.

Some drawbacks to replication include increased storage and network bandwidth. When properly used, replication is a very useful tool in disaster recovery efforts as it could house hundreds of virtual machines.

**Deduplication**

Deduplication is a technique that is used to remove duplicate copies of the same data. This can be beneficial as it increases storage space, as well as being more cost effective.

Logging onto NAS01, a new volume was created for the demonstration. The deduplication setting was turned on. The volume was then shared using NFS, mounted, and the CentOS server data was copied across. In the FreeNAS interface, you can view the progress of the volume space as the data is being copied to the dedup volume. Note that it does not update automatically, the user must press the refresh button. There are also graphs available in the Reporting tab to view the progress. Note how the physical memory use increases when deduplication increases.

One drawback to using deduplication is that FreeNAS uses a lot of RAM to maintain a table of block hashes during the process, which can hinder performance.