Nutanix Files

Nutanix Tech Note

Version 3.0 • October 2018 • TN-2041

Copyright

Nutanix, Inc.
1740 Technology Drive, Suite 150
San Jose, CA 95110

Contents

1. Executive Summary

Nutanix Files is a software-defined, scale-out file storage solution that provides a repository for unstructured data, such as home directories, user profiles, departmental shares, application logs, backups, and archives. Flexible and responsive to workload requirements, Files is a fully integrated, core component of the Nutanix Enterprise Cloud.

You can deploy Nutanix Files on an existing cluster or a standalone cluster. Unlike standalone NAS appliances, Files consolidates VM and file storage, eliminating the need to create an infrastructure silo. Administrators can manage Files with Nutanix Prism, just like VM services, thus unifying and simplifying management. Integration with Active Directory enables support for quotas and access-based enumeration, as well as self-service restores with the Windows previous version feature. Nutanix Files also supports file server cloning, which lets you back up Files off-site, as well as run antivirus scans and machine learning without affecting production.

Nutanix Files can run on a dedicated cluster or be collocated on a cluster running user VMs. Beginning with AOS 5.0, Nutanix supports Files with both ESXi and AHV. Files includes native high availability and uses the Acropolis Distributed Storage Fabric (DSF) for intracluster data resiliency and intercluster asynchronous disaster recovery. The DSF also provides data efficiency techniques such as erasure coding (EC-X), compression, and deduplication.

2. Audience and Purpose

This tech note is part of the Nutanix Solutions Library and is intended for architects and systems engineers who want to use Nutanix Files as a simple way to deliver user and group file management. This document describes how to implement and operate Files in your datacenter.

We cover the following subject areas:

- Overview of the Nutanix architecture with Files.
- Load balancing of general and distributed shares (SMB) and exports (NFS).
- High availability.
- Backup and recovery.
- Quotas and permission management.
- Antivirus.

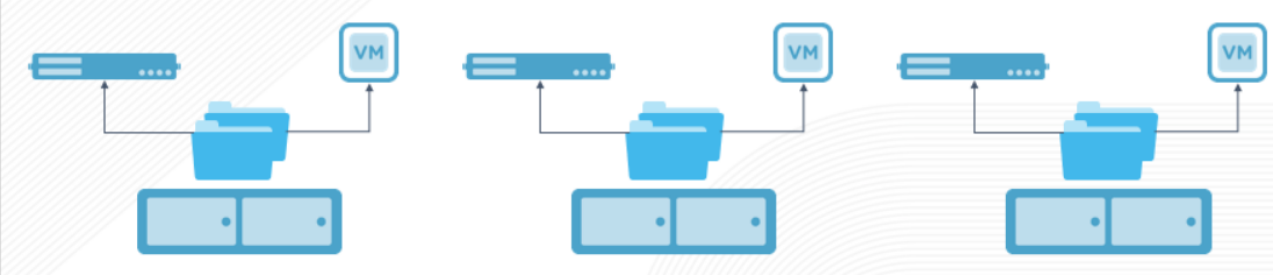

Figure 1: Nutanix Files Scales Out or Up on Existing Nutanix Clusters

Table 1: Document Version History

Version Number	Published	Notes
1.0	December 2016	Original publication.
1.1	May 2017	Updated Backup and Disaster Recovery section.
1.2	September 2017	Updated for version 2.2 features.
2.0	February 2018	Updated for version 3.0.
2.1	April 2018	Solution overview update.

Version Number	Published	Notes
2.2	August 2018	SMB share and NFS export updates.
3.0	October 2018	Updated for version 3.1 and updated product naming.

3. Nutanix Enterprise Cloud Overview

Nutanix delivers a web-scale, hyperconverged infrastructure solution purpose-built for virtualization and cloud environments. This solution brings the scale, resilience, and economic benefits of web-scale architecture to the enterprise through the Nutanix Enterprise Cloud Platform, which combines three product families—Nutanix Acropolis, Nutanix Prism, and Nutanix Calm.

Attributes of this Enterprise Cloud OS include:

- Optimized for storage and compute resources.
- Machine learning to plan for and adapt to changing conditions automatically.
- Self-healing to tolerate and adjust to component failures.
- API-based automation and rich analytics.
- Simplified one-click upgrade.
- Native file services for user and application data.
- Native backup and disaster recovery solutions.
- Powerful and feature-rich virtualization.
- Flexible software-defined networking for visualization, automation, and security.
- Cloud automation and life cycle management.

Nutanix Acropolis provides data services and can be broken down into three foundational components: the Distributed Storage Fabric (DSF), the App Mobility Fabric (AMF), and AHV. Prism furnishes one-click infrastructure management for virtual environments running on Acropolis. Acropolis is hypervisor agnostic, supporting three third-party hypervisors—ESXi, Hyper-V, and XenServer—in addition to the native Nutanix hypervisor, AHV.

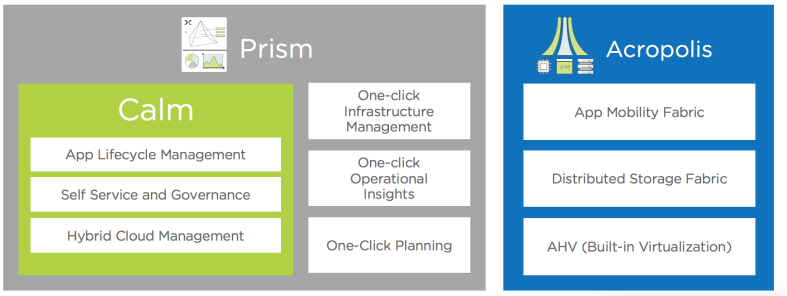

Figure 2: Nutanix Enterprise Cloud

3.1. Nutanix Acropolis Architecture

Acropolis does not rely on traditional SAN or NAS storage or expensive storage network interconnects. It combines highly dense storage and server compute (CPU and RAM) into a single platform building block. Each building block delivers a unified, scale-out, shared-nothing architecture with no single points of failure.

The Nutanix solution requires no SAN constructs, such as LUNs, RAID groups, or expensive storage switches. All storage management is VM-centric, and I/O is optimized at the VM virtual disk level. The software solution runs on nodes from a variety of manufacturers that are either all-flash for optimal performance, or a hybrid combination of SSD and HDD that provides a combination of performance and additional capacity. The DSF automatically tiers data across the cluster to different classes of storage devices using intelligent data placement algorithms. For best performance, algorithms make sure the most frequently used data is available in memory or in flash on the node local to the VM.

To learn more about the Nutanix Enterprise Cloud, please visit the Nutanix Bible and Nutanix.com.

4. Nutanix Files Architecture

Nutanix Files is a scale-out approach that provides Server Message Block (SMB) and Network File System (NFS) file services to clients. Nutanix Files server instances are composed of a set of VMs (called FSVMs). Files requires at least three FSVMs running on three nodes to satisfy a quorum for high availability.

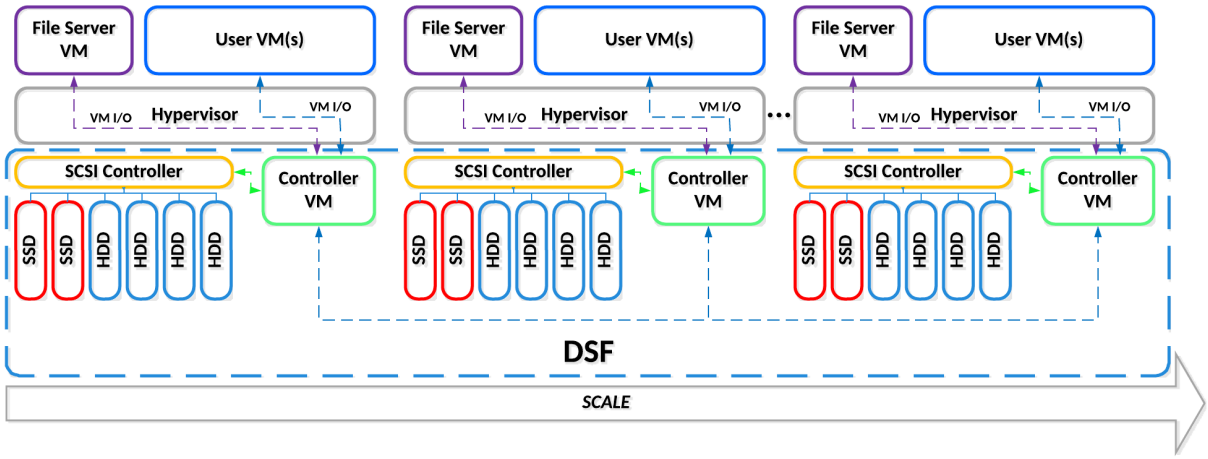

Figure 3: Nutanix Files Server Instances Run as VMs for Isolation from the DSF

4.1. Nutanix Files: File Server Virtual Machine

The File Server VM (FSVM) is based on CentOS and incorporates all the security and hardening that goes into the Nutanix Controller VM (CVM). All the FSVMs have the same configuration, starting with four vCPUs and 12 GiB of RAM. You can add more vCPUs, RAM, and FSVMs to the cluster. For each file server the number of FSVMs must be less than or equal to the number of nodes in the Nutanix cluster; however, you can create multiple file server deployments if needed. Nutanix Files 3.1 introduced single FSVM deployments intended for one-node and two-node Nutanix clusters.

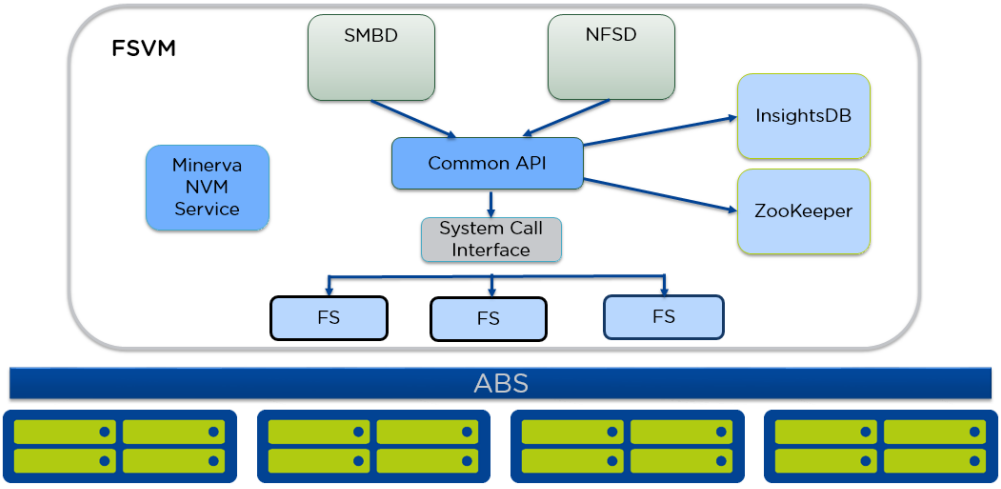

Figure 4: Data Path Architecture of Nutanix Files

Nutanix Files can support SMB and NFS from the same FSVM, but each individual share or export must be one or the other. Both SMB and NFS share a common library, allowing a modular approach. InsightsDB is a NoSQL database that maintains the stats and alerts for Files. Zookeeper is a centralized service that maintains configuration information, such as domain, share or export, and IP information. The Minerva NVM Service talks to the local CVM and sends heartbeats to share health information and to help with failover.

Each FSVM stores file server data on multiple file systems that store share-specific data. The individual file system provides snapshot capability that is used to provide Windows Previous Version (WPV) support to clients. By using separate file systems for each share or export, Nutanix Files can scale to support billions of files in one cluster.

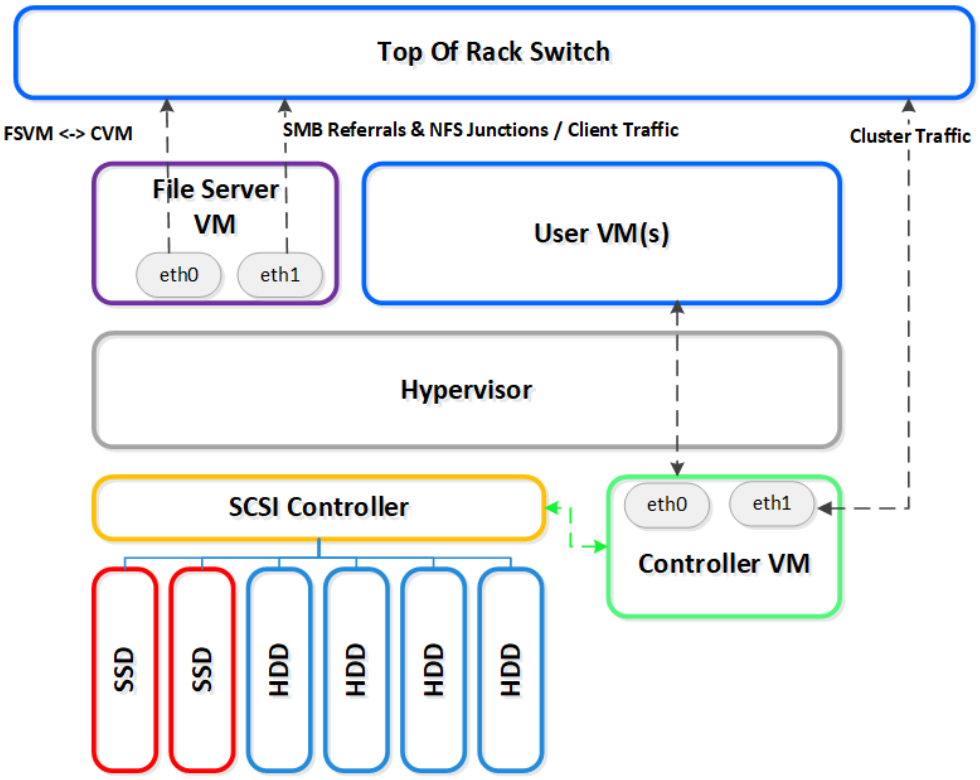

Figure 5: File Server VM Internal Communication on One Node

The above diagram shows one FSVM running on a node, but you can put multiple FSVMs on a node for multitenancy.

Networking

The FSVM has two network interfaces: the storage interface and the client interface. The storage interface is used by the FSVM service that talks to the CVMs, and it also provides access to Nutanix Volumes iSCSI vDisks in volume groups. The storage interface helps manage deployment, failover, and maintenance, and enables control over one-click upgrades. Integration with the CVM lets the FSVM determine if a storage fault has occurred and, if so, whether you must take action. The FSVM service sends a heartbeat to its local CVM service each second indicating its state.

The client interface allows clients to connect to SMB shares and NFS exports hosted on a FSVM. A client can connect to any FSVM client network interface to access their file data. If a different FSVM provides the data, the client connection automatically redirects to the correct FSVM interface. If an FSVM fails, the client network address for the failed FSVM moves to another to preserve data access.

Storage

As shown in the following figure, each FSVM uses three separate vDisks: a 12 GiB boot disk that contains the boot image, a 45 GiB disk (/home/nutanix) that contains logs and software state, and a 45 GiB disk for Cassandra data.

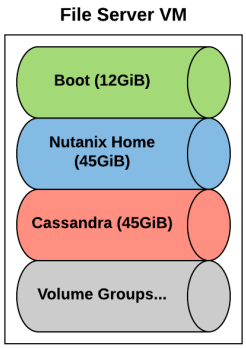

Figure 6: File Server VM vDisks and Volume Groups

Volume groups are used to make the individual file systems highly available. A volume group is a collection of logically related vDisks (or volumes) attached to the guest via iSCSI. When an FSVM is down for maintenance or if a fault occurs, one of the surviving FSVMs takes over volume group ownership and continues servicing requests.

Up to 10 volume groups back each FSVM. All storage for the FSVM and the volume groups is thin provisioned. Each volume group has four 10 TB vDisks for data and two 512 GB vDisks for metadata.

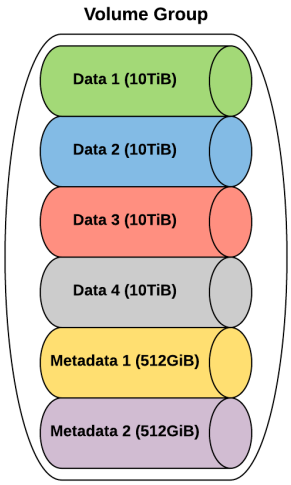

Figure 7: File Server VM Volume Group vDisks

Each time you add a share or export to the file server, you're also adding between 1 and 15 volume groups.

4.2. Exports and Shares

There are two types of SMB shares and two types of NFS exports.

SMB Shares

- Home directory and user profile
- General purpose share

NFS Exports

- Nonsharded directory
- Sharded directory

A nonsharded directory is an NFS export hosted by a single FSVM. A sharded directory export spreads the workload by distributing the hosting of top-level directories across FSVMs, which also simplifies administration.

Home directory and user profile SMB shares are sharded directories with additional features (see the Accessing Home Share Directories section below for more detail). In a sharded share or export, there are no files allowed in the root.

An SMB general purpose share is similar to an NFS nonsharded export. Both are more traditional shares in which a single FSVM hosts the data for the entire share. You scale out by adding shares or exports.

Sharded Directories

Sharded shares and exports distribute data by dividing the top-level directories across all the FSVMs that make up the file server. Nutanix Files maintains the directory mapping for each responsible FSVM using an internal scale-out database called InsightsDB. FSVMs use DFS referrals for SMB and junctions for NFS to make sure the clients can connect to the right top-level directories.

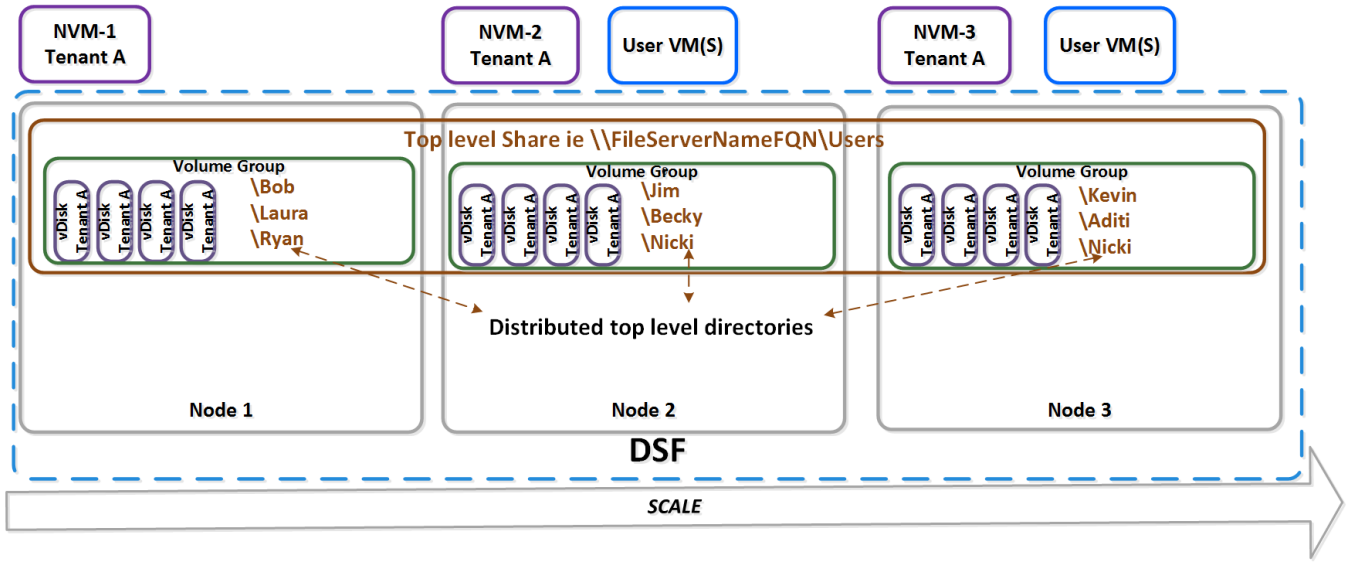

Figure 8: Distribution of Home Directory Shares

Sharded directories work well for home shares and exports because Nutanix Files automatically spreads the workload over multiple FSVMs per user (see the preceding figure, Distribution of Home Directory Shares). If a user creates a share called “\\FileServer1\Users” that contains top-level directories \Bob, \Becky, and \Kevin, \Bob may be on FSVM-1, \Becky on FSVM-2, \Kevin on FSVM-3, and so on. The FSVMs use a string hashing algorithm based on the directory names to distribute the top-level directories.

This distribution can accommodate a very large number of users or directories in a single share or export. The scaling limits of more traditional designs can force administrators to create multiple shares or exports in which, for example, one set of users whose last names begin with A through M run off one controller, and users whose names begin with N through Z run off another controller. This design limitation leads to management overhead headaches and unnecessary Active Directory complexity. For these reasons, Nutanix Files expects to have one SMB home

directory share for the entire cluster. If you need to have more than one home directory share, you can create it using nCLI.

The top-level directories act as a reparse point—essentially a shortcut. Consequently, administrators must create directories at the root of the share for optimal load balancing. Because it appears as a shortcut, we don't allow user files in the root level of the share; we recommend setting permissions at the share or export root before deploying user folders. This step allows newly created top-level directories to inherit permissions, rather than having to adjust them after the fact using the Nutanix Files Microsoft Management Console (MMC) plugin.

General-purpose shares and exports (not user directories) by default do not distribute top-level directories. The files and subfolders for general-purpose shares and exports are always owned by a single file server. The diagram below illustrates two general-purpose shares (for example, accounting and IT) on the same file server.

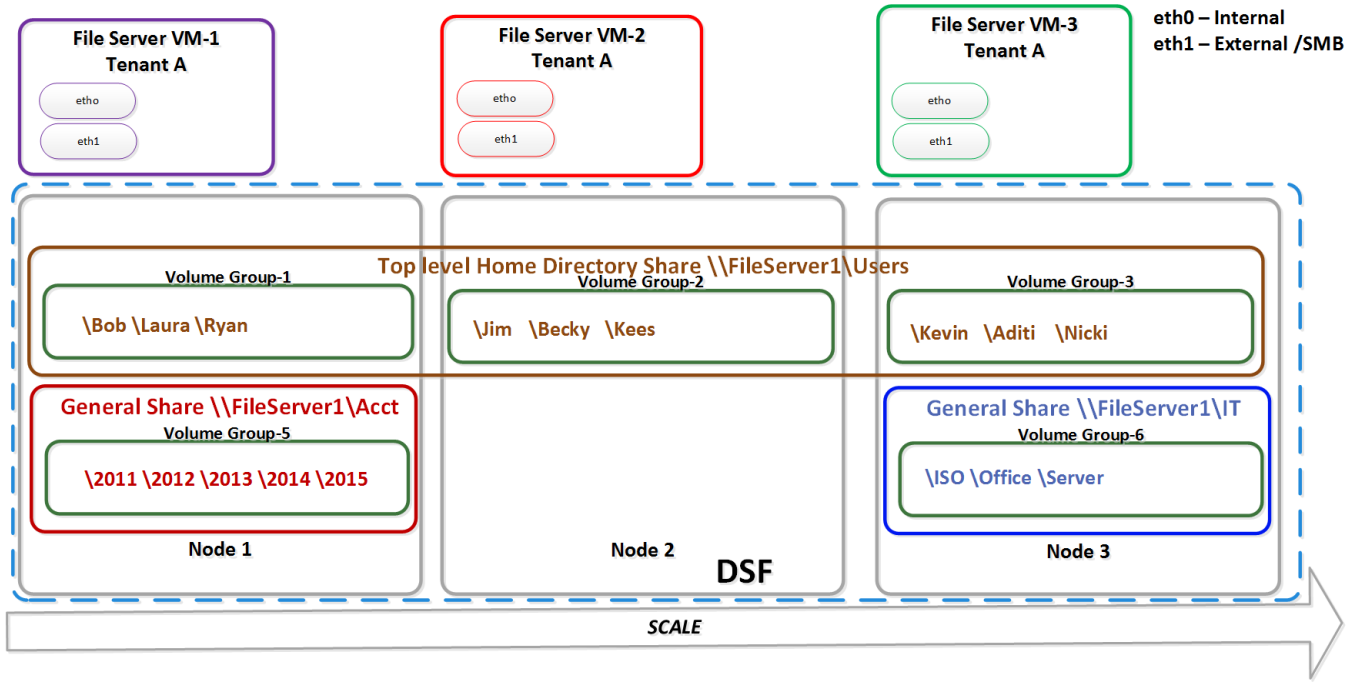

Figure 9: Two General-Purpose Shares on the Same File Server

Unlike home directory shares and exports, general-purpose shares and exports can store files in the root of the directory.

Managing NFS Sharded Directories

NFS sharded directories introduces some unique behaviors. To balance performance across FSVMs, each top-level directory you create becomes an automatically generated export. Nutanix Files mounts the export on demand when the directory is accessed. Since these are exports,

rather than standard directories, it takes a few steps to remove a top-level directory. Below we provide an example to demonstrate how sharded top-level directories behave and then walk through the process of deleting a top-level directory.

If you create a sharded share and mount it as /projects, you can see that mount on Linux using the **df** command.

```
# df /projects
Filesystem         1K-blocks      Used Available Use% Mounted on
1.1.1.10:/projects 1073741824 839455744 234286080  79% /projects
```

Next, create some project directories and access the directories.

```
mkdir /projects/project1
mkdir /projects/project2
ls /projects/project1
ls /projects/project2
```

Accessing the directory, using **ls** in this case, causes the automatically created top-level directory export to be mounted. If you run **df** again you will see two additional mount points.

```
# df | grep project
1.1.1.10:/projects          1073741824 839455744 234286080  79% /projects
1.1.1.11:/projects/project1 1073741824 839455744 234286080  79% /projects/project1
1.1.1.12:/projects/project2 1073741824 839455744 234286080  79% /projects/project2
```

These additional mount points allow a different FSVM to serve each export.

This behavior introduces additional steps when deleting a top-level directory. You can delete the project2 directory from any NFS client with the export mounted if you are logged in as a user with the appropriate permissions. There are three steps to deleting a sharded export.

- Delete the contents of the share:

```
rm -rf /projects/project2/*
```

- Unmount the project2 share:

```
umount /projects/project2
```

- Delete the project2 directory:

```
rmdir /projects/project2
```

At this point the top-level directory is deleted and becomes inaccessible to other clients mounting the export. Processes that access the export after it is deleted receive a Stale File Handle error.

vDisk and Volume Group Allocation

Home and sharded shares and exports begin with 15 volume groups and each FSVM starts with 5 volume groups. Nutanix Files distributes the 15 volume groups to different FSVMs within the file server cluster. When you have a large number of users, multiple volume groups improve load balancing across the file server cluster. Instead of physically copying data, a different storage controller can easily host a volume group based on detected hot spots. General-purpose shares store application and group directories. A single FSVM and volume group serves each nonsharded general-purpose share or export.

Once you reach the limit of 10 volume groups per FSVM, new shares and exports use existing volume groups of the same type. For example, if you have deployed a home directory share (15 volume groups) and 15 general-purpose shares (creating 15 additional volume groups) on a three-node physical cluster, each FSVM hosts 10 volume groups: 5 volume groups for the home directory share and 5 volume groups for the general-purpose shares (one per share). In this situation, because each FSVM is serving the maximum of 10 volume groups, the next share created uses an existing volume group.

Every file server maps one-to-one to a container. This mapping allows you to manage dedupe, compression, and erasure coding individually for each file server deployed. Inline compression is turned on by default to save capacity.

4.3. Load Balancing and Scaling

Load balancing occurs on two levels. First, a client can connect to any one of the FSVMs and users can add FSVMs as needed. Second, on the storage side, Nutanix Files can redistribute volume groups to different FSVMs for better load balancing across nodes. Following are situations that necessitate load balancing:

1. When removing an FSVM from the cluster, Files automatically load balances all its volume groups across the remaining FSVMs.
2. During normal operation, the distribution of top-level directories becomes poorly balanced due to changing client usage patterns or suboptimal initial placement.
3. When increased user demand necessitates adding a new FSVM, its volume groups are initially empty and may require rebalancing.

Nutanix Files addresses the second and third situations by maintaining usage statistics and patterns to detect per-FSVM load (in terms of CPU and memory utilization) and per-volume group load (in terms of user connections, number, and latency of operations). Nutanix Files uses these statistics to make a load balancing recommendation, but the administrator must accept the recommendation before Files carries out the action—Nutanix calls this feature one-click optimization.

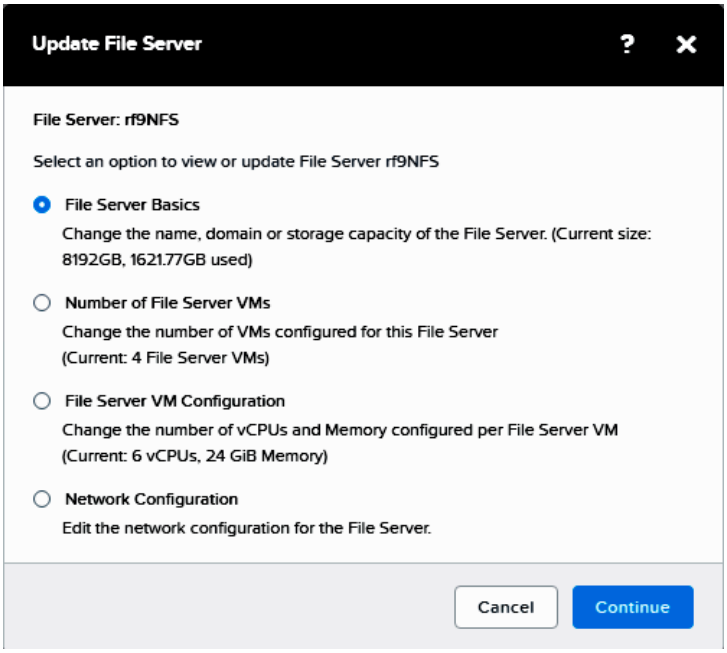

Figure 10: Update the File Server to Scale Up or Out

Load balancing through volume group redistribution may not always improve performance. For example, if clients target a low-level share directory that cannot be further distributed among FSVMs, performance does not improve. In such cases, Nutanix Files supports scaling up the FSVMs by adding vCPUs and memory. Scaling up is seamless to end users.

In the current implementation, there is a brief outage during volume group migration and FSVM scaling out. We don't currently support a durable file handle. The file share or export reconnects after scaling up or scaling out. When a volume group is moved, the time to live (TTL) slowly decreases to limit the disconnection from the client. Today most clients try to reconnect for 50 to 60 seconds, limiting the overall impact.

Load balancing volume groups with Nutanix Volumes requires the administrator to configure an iSCSI data services IP. The data services IP is a highly available virtual IP address used to help balance the load. Once this IP is configured, the Nutanix administrator doesn't need to worry about configuring anything else.

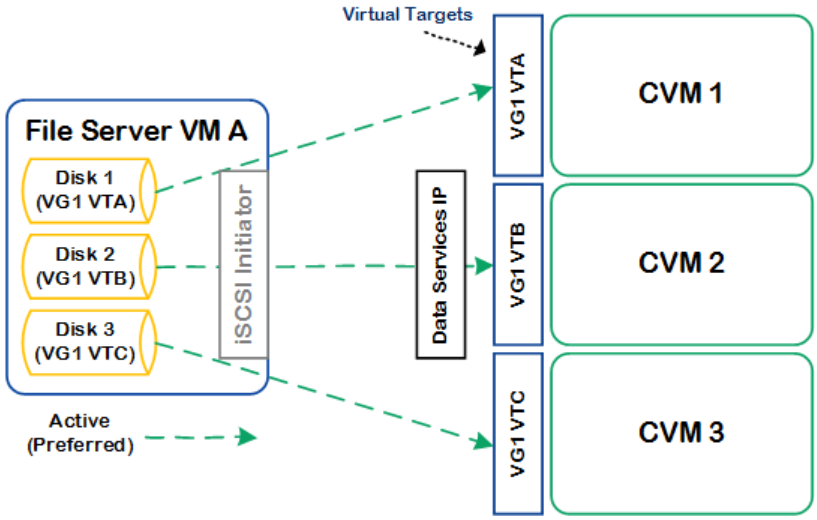

Figure 11: Load Balancing Volume Groups with Nutanix Volumes

For more detailed information, please refer to the Nutanix Volumes best practices guide.

4.4. High Availability

Nutanix designed Files to recover from a range of service disruptions, including when a local CVM or FSVM restarts or fails.

CVM Failure or Upgrade

If a CVM goes offline because of failure or planned maintenance, any active sessions against that CVM are disconnected, triggering the iSCSI client to log on again. The new logon occurs through the external data services IP, which redirects the session to a healthy CVM. The figure below shows this general process.

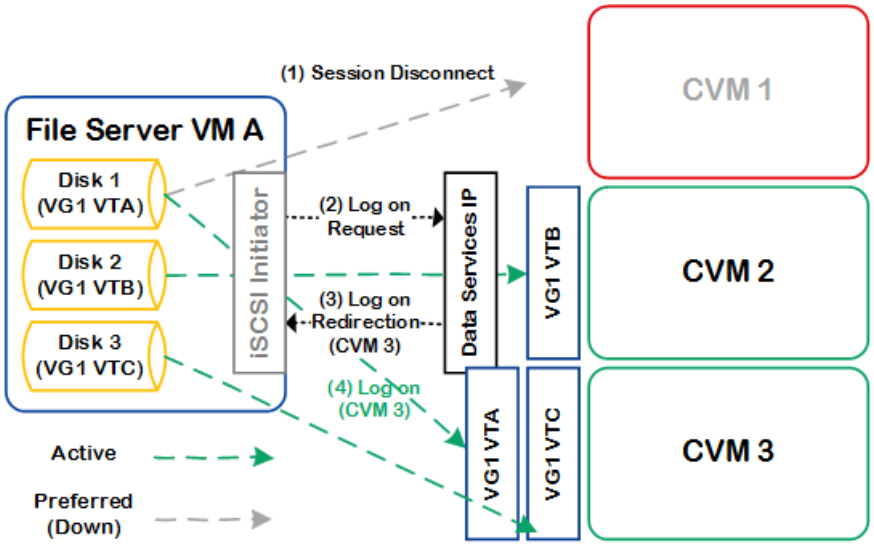

Figure 12: Nutanix Volumes Load Balancing for File Server Volume Groups

When the failed CVM returns to operation, the iSCSI session fails back. In the case of a failback, the FSVM is logged off and redirected to the appropriate CVM.

Node Failure

When a physical node fails completely, Nutanix Files uses leadership elections and the local Minerva CVM service to recover. The FSVM sends heartbeats to its local Minerva CVM service once per second, indicating its state and that it's alive. The Minerva CVM service keeps track of this information and can take action during a failover.

When an FSVM goes down, the Minerva CVM service unlocks the files from the downed FSVM and releases the external address from eth1. The downed FSVM's resources then appear on a running FSVM. The internal Zookeeper instances store this information so that they can send it to other FSVMs if necessary.

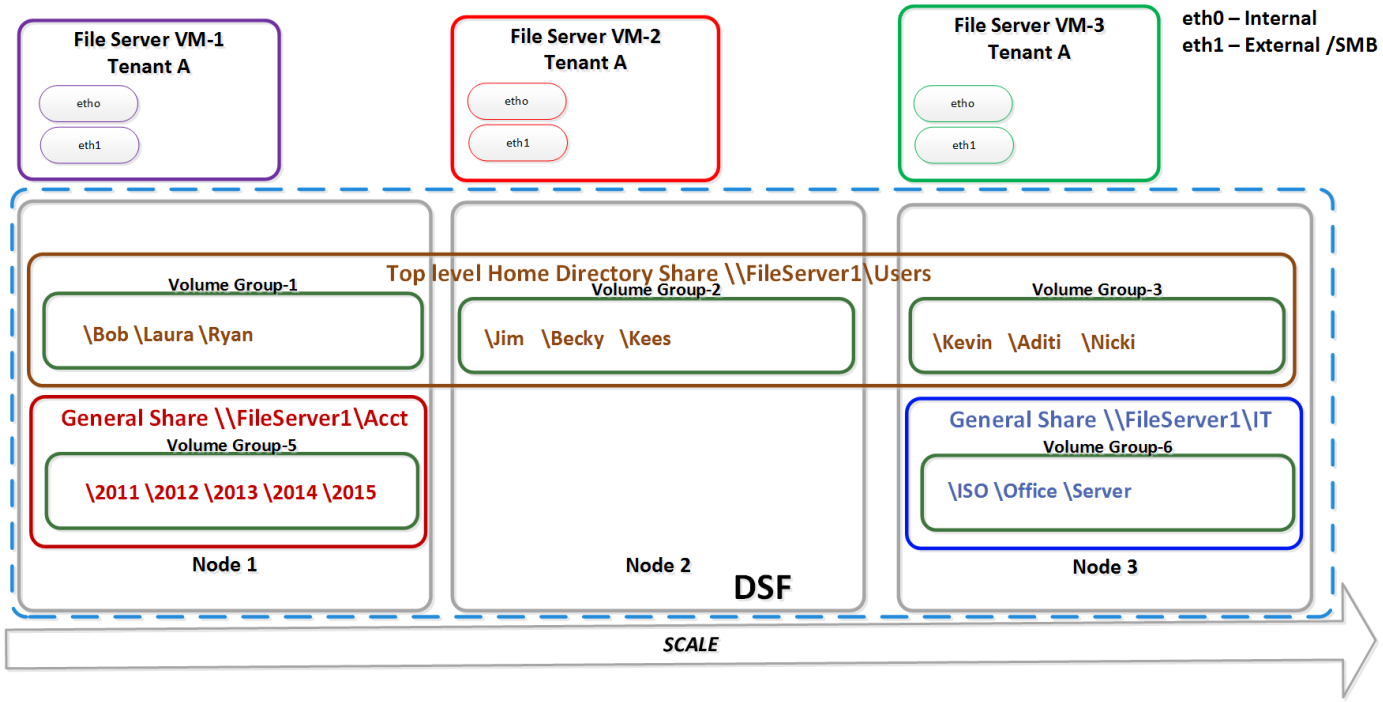

Figure 13: Each File Server VM Controls Its Own Volume Groups in a Healthy State

When an FSVM is unavailable, the remaining FSVMs volunteer for ownership of the shares and exports that were associated with the failed FSVM. The FSVM that takes ownership of the volume group informs the CVM that the volume group reservation has changed. If the FSVM that attempts to take control of the volume group is already the leader for a different volume group that it has volunteered for, it relinquishes leadership for the new volume group immediately. This arrangement ensures distribution of volume groups, even if multiple FSVMs fail.

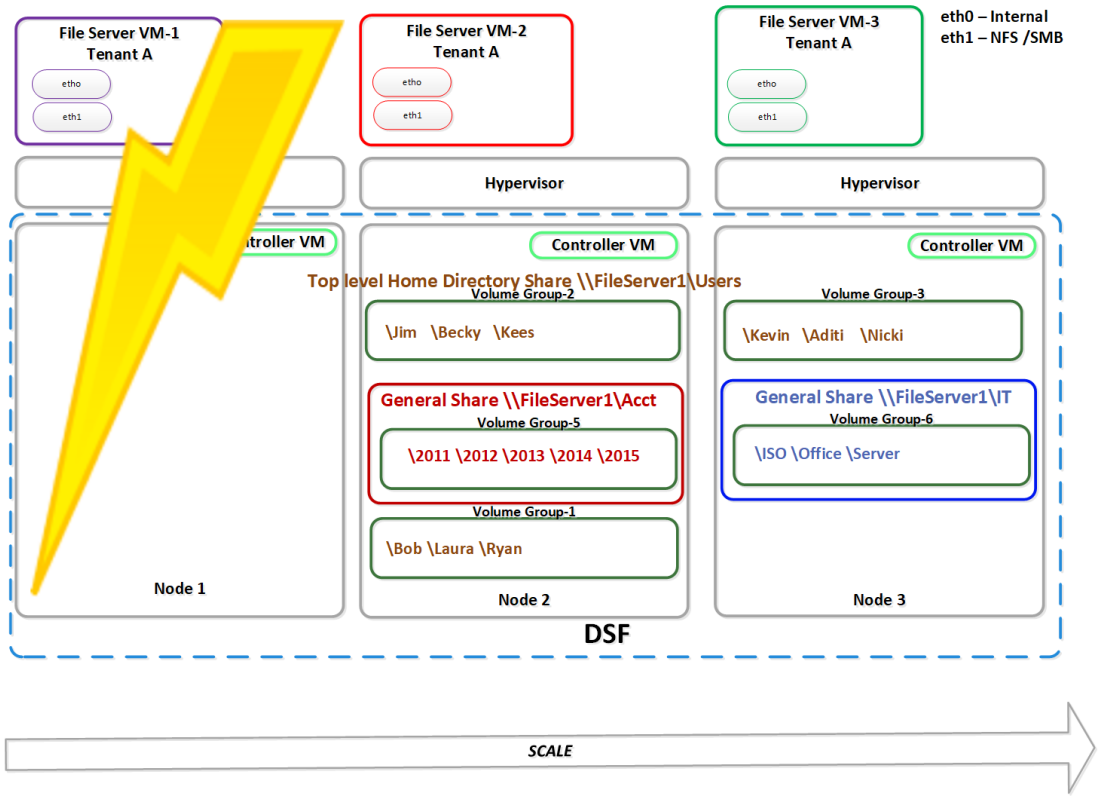

Figure 14: File Server VM-1 Failure

The Nutanix Files Zookeeper instance tracks the original FSVM's ownership using the storage IP address (eth0), which does not float from node to node. Because FSVM-1's client IP address from eth1 is now on FSVM-2, client connections persist. The volume group and its shares and exports are reregistered and locked to FSVM-2 until FSVM-1 can recover and a grace period has elapsed.

When FSVM-1 comes back up and finds that its shares and exports are locked, it assumes that an HA event has occurred. After the grace period expires, FSVM-1 regains control of the volume group through the Minerva CVM service.

To summarize, the process Nutanix Files goes through to reinstate control is:

1. Stop SMB and NFS services.
2. Disconnect the volume group.
3. Release the IP address and share and export locks.
4. Register the volume group with FSVM-1.
5. Present new shares and exports to FSVM-1 with eth1.

A node failure drops the client connection, but the client can recover based on common timeout values.

4.5. Active Directory and SMB Operations

Nutanix Files SMB works with Active Directory. To deploy a cluster, you must have domain administrator privileges; these privileges are necessary to create the machine account and the DNS entries for DNS referrals. The file server does not store these credentials. Domain administrator privileges are also necessary to remove a file server from Active Directory.

The maximum number of client connections depends primarily on the amount of RAM in the FSVM. See the table below for Nutanix configuration recommendations.

Table 2: Supported Active Client Connections

RAM per FSVM	Supported Client Connections per FSVM	Supported Client Connections for Four-FSVM File Server Cluster
Less than or equal to 12 GB	250	1,000
Between 12 GB and 16 GB	500	2,000
Between 16 GB and 24 GB	1,000	4,000
Between 24 GB and 32 GB	1,500	6,000
Between 32 GB and 40 GB	2,000	8,000
Between 40 GB and 60 GB	2,500	10,000
Greater than 60 GB	4,000	16,000

You can continue deploying additional FSVMs if you have free nodes; you can also deploy multiple file servers.

The following diagram shows what happens behind the scenes when a client sends a file access request.

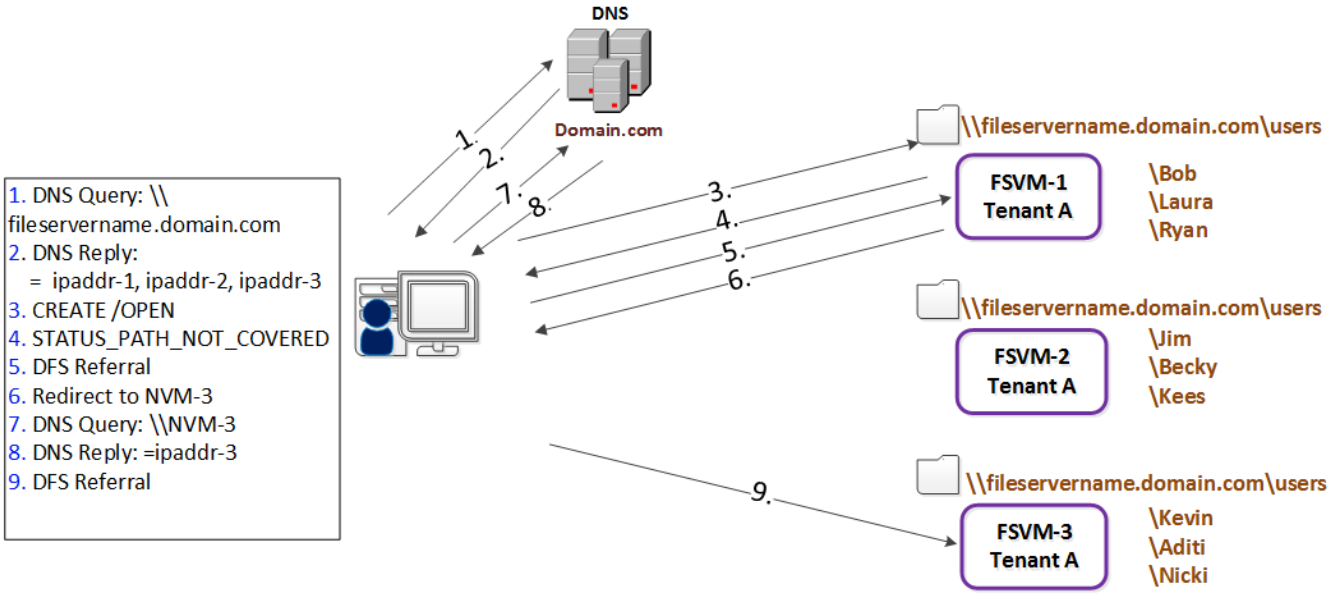

Figure 15: DNS Request for SMB

When the user “Nicki” accesses her files, she clicks on a shortcut that triggers a DNS request. The DNS request is first sent for the file server name.

1. Using DNS round robin, DNS replies with an FSVM IP address. In this example, the IP address for FSVM-1 returned first.
2. The client sends a create or open request to FSVM-1.
3. The \Nicki folder doesn’t exist on this file server, so a STATUS_PATH_NOT_COVERED is returned.
4. The client then requests a DFS referral for the folder.
5. FSVM-1 looks up the correct mapping in the file server’s Zookeeper and refers the client to FSVM-3.
6. A DNS request goes out to resolve FSVM-3.
7. The DNS request returns the IP address of FSVM-3.
8. The client gets access to the correct folder.

Managing Shares

You can manage general-purpose shares the same way that you manage traditional file servers. The home share has special requirements because of the way we use DFS referrals. DFS referrals have a single namespace even though the data contained within the share may be spread out over many FSVMs. Typically, to delete a user’s home directory, you select the top-level directory and delete the entire directory subtree. However, when using a sharded home export or share you cannot simply delete the top-level directory, because it is actually a separate

share created internally by Nutanix Files. Since you cannot directly delete a top-level directory, removing it requires additional steps.

There are three options for renaming or deleting home share folders:

- Find out which FSVM is hosting the folder and delete the folder directly out of the VM.
- For SMB shares, you can use Nutanix PowerShell scripts (found on the support portal) to handle mass operations.
- Use the Microsoft Management Console (MMC) snap-in to manage top-level directories (TLD). Any file server administrator can perform the MMC operations; there is no need to assign privileges manually. Establish a connection through shared folders, snap-in to the Files namespace, and perform the following SMB share management tasks:
  - # Create, delete, rename, and change permissions for top-level directories.
  - # Change permissions for shares.

Accessing Home Share Directories

You can access an SMB home directory without including the share name in the universal naming convention (UNC) path. You can access a user home directory directly via this path:

- Use \\Files_server\AccountName instead of \\Files_server\home_share\AccountName.
  - # For example, \\FileServer1\Laura instead of \\FileServer1\users\Laura.

In Nutanix Files, user home shares are visible with other shares when enumerating shares on the file server. This support is enabled by default, but it requires the following prerequisites to work:

- User home directories are the top-level directories in the home share.
- User home directories have the required permissions.

If a user's home directory exists in multiple home shares, Files matches to the first share created chronologically.

4.6. Active Directory, LDAP, and NFS Operations

Nutanix Files supports NFS version 4, which is more stateful than NFSv3, and includes advanced features such as strongly mandated security, firewall friendliness, ACLs (Access Control Lists), and DFS-like referrals. Moreover, most recent distributions of Linux, Solaris, and AIX use NFSv4 as the default client protocol.

To make the transition from NFSv3 easier, Files does not require administrators to configure Active Directory or LDAP. You can use AUTH_SYS or AUTH_NONE authentication. AUTH_SYS authenticates at the client, just like NFSv3.

There are three different levels of Kerberos authentication when you enable Active Directory or LDAP support. Each of the following options uses Kerberos version 5:

- krb5, DES symmetric key encryption, and an MD5 one-way hash for Nutanix Files credentials.
- krb5i, in addition to krb5, uses MD5-based MAC on every request and response.
- krb5p, on top of krb5 and krb5i, makes the connection between client and server private by applying DES encryption.

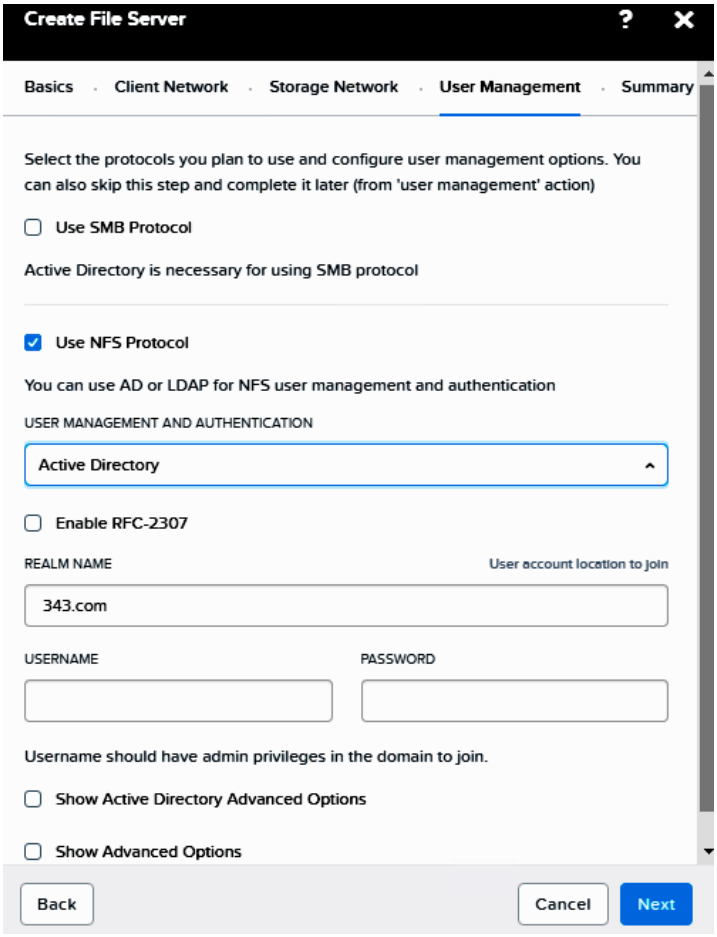

Figure 16: Nutanix Files User Management with NFS

When you deploy Nutanix Files for NFS, you select AD, LDAP, or leave it unmanaged. Files supports using Active Directory for SMB and LDAP for NFS. However, Files does not support both AD and LDAP for the same file server.

The following diagram shows what happens behind the scenes when a client sends a file access request using NFS.

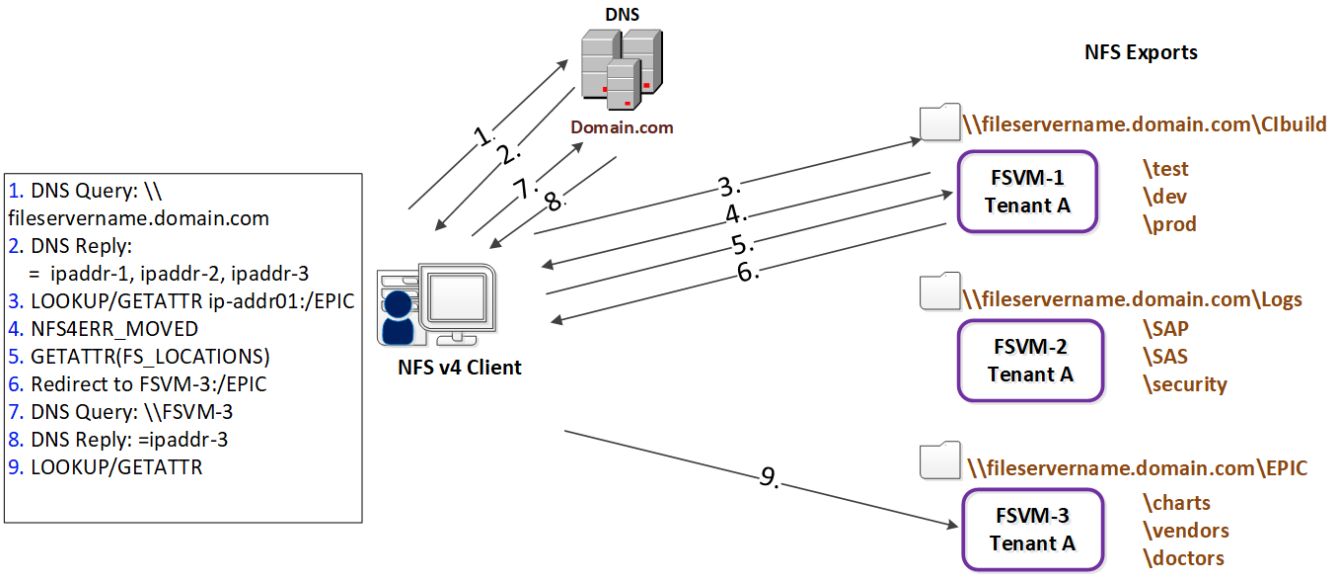

Figure 17: DNS Request for NFS

When the server wants to access files, the client first sends a DNS request for the file server name.

1. Using DNS round robin, a DNS reply returns with an FSVM address. In this example, the IP address for FSVM-1 returned first.
2. The client sends a create/open request to FSVM-1.
3. The \EPIC mount doesn't exist on this file server, so it returns NFS4ERR_MOVED.
4. The client then requests a GETATTR(FS_LOCATIONS).
5. FSVM-1 looks up the correct mapping in the file server's Zookeeper and refers the client to FSVM-3.
6. A DNS request goes out to resolve FSVM-3.
7. The DNS request returns the IP address of FSVM-3.
8. The client gets access to the correct mount point.

5. Backup and Disaster Recovery

Following modern data protection methodologies, Nutanix provides administrators and users quick restore access using Self-Service Restore (SSR) and site recovery with Nutanix-based snapshots.

5.1. Self-Service Restore

Administrators can enable SSR at any time for SMB shares, such as when they create a share.

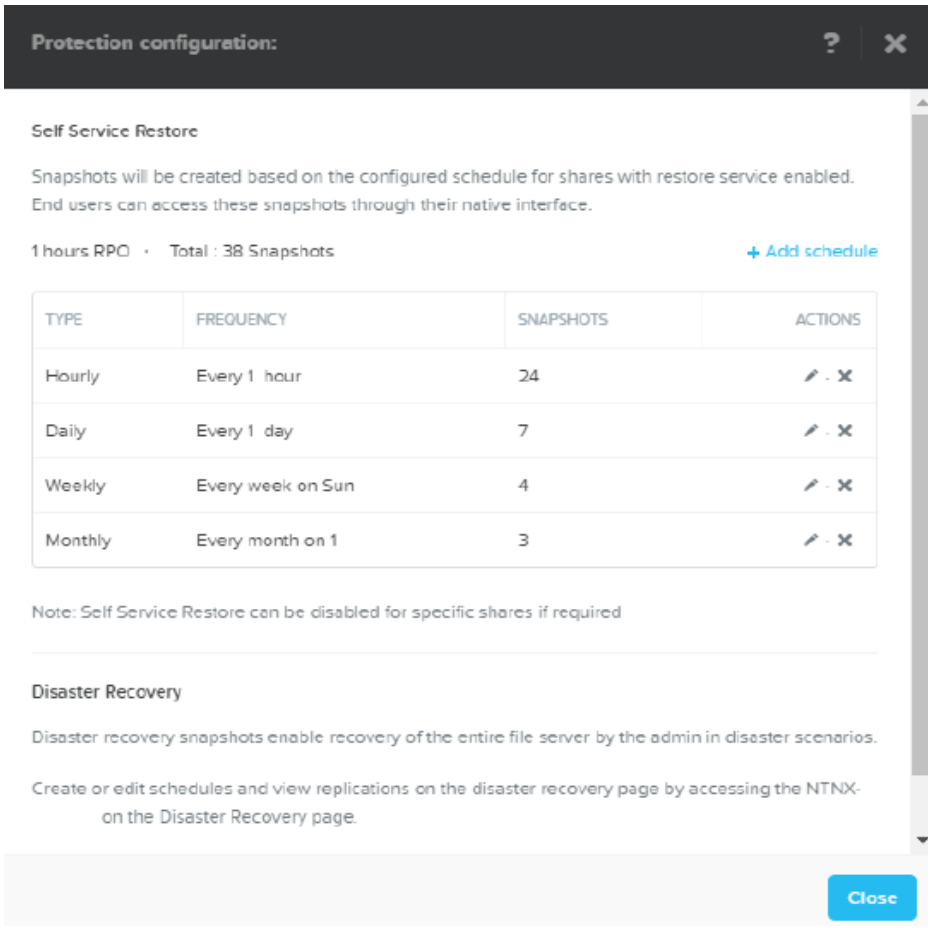

Figure 18: Enable Protection Using SSR

SSR allows you to create and view manual or automatic snapshots of SMB shares while the share is in use. The share snapshots are read-only, point-in-time copies (that is, copies created at a specified time).

You can view and restore removed or overwritten files, which allows you to choose a share snapshot from the same file at different times during the file's history. Administrators can configure a snapshot schedule at the file server level that applies to all shares in the file server. Nutanix Files supports 24-hour, daily, weekly, and monthly snapshots on a fixed schedule. The default snapshot policy includes:

- Hourly, 24 per day.
- Once daily.
- Three per month.

SSR schedule configuration is available for all shares and allows users to configure different schedules for individual shares. Snapshots can also be scheduled for each object in a share. Schedule frequency can vary to suit your requirements, including shorter intervals for same-day protection against accidental deletions. You can enable SSR during or after share creation.

SSR supports share updates for both general-purpose and home SMB shares. After share creation, the administrator can change the current settings using the share update workflow feature, which is supported on both home and general-purpose shares.

5.2. Protection Domains and Consistency Groups

Nutanix provides integrated, automated disaster recovery from a secondary running Nutanix cluster. A Nutanix Files cluster can be protected via Prism and uses the same asynchronous replication with protection domains and consistency groups as any other Nutanix cluster. A protection domain is a defined group of entities (VMs and volume groups) that you back up locally on a cluster and that may replicate to one or more remote sites. A consistency group is a subset of the entities in a protection domain. Consistency groups are configured to snapshot a group of VMs or volume groups in a crash-consistent manner.

When you create a file server, Prism automatically sets up a corresponding protection domain, which it annotates with the Nutanix Files cluster UUID and file server name. Prism also creates multiple consistency groups within a protection domain, including a group that includes all FSVMs.

Once Nutanix Files is protected, all future operations on it (such as adding or removing FSVMs or adding or deleting volume groups) automatically update the corresponding consistency group in the protection domain.

5.3. Cluster Failure and Restoration

In the event of a Nutanix Files cluster failure, restore Files by initiating the activate workflow. Activate restores from the last good snapshot. If you are moving your file services because the cluster needs to be shut down, as with a planned outage, the migrate workflow shuts down all the FSVMs and takes a final snapshot for replication.

Restoring the file server may require you to configure network VLANs on the replica site before Nutanix Files becomes operational again. The process is similar to the process for creating the file server. The system administrator must enter Active Directory information during the restoration process.

5.4. Cloning

Because Nutanix Files cloning does not affect the original Files cluster, it offers improved support for a variety of use cases, including:

- Backups at the primary and secondary sites.
- DR at the secondary site.
- File server recovery from a specific point in time.
- File server creation at the primary or remote site for testing or development.
- File server clone copies.

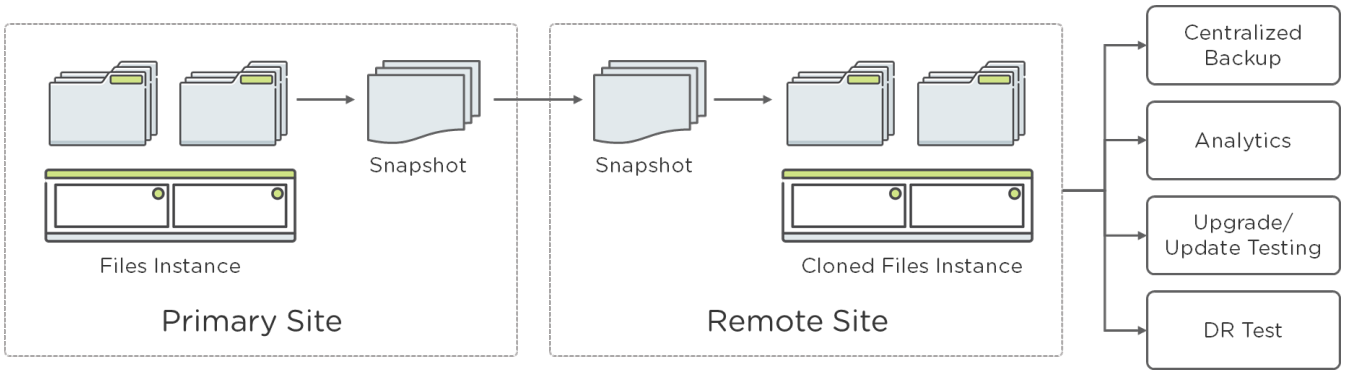

Figure 19: Nutanix Files Cloning Use Cases

Files uses Nutanix native snapshots to clone entire file servers. The clone is a thin copy that consumes minimal storage space. File server clones reside on the same container as the original and maintain the original security permissions. During the clone process you can specify new IP addresses and give the cloned file server a new name.

5.5. SMB Quotas

The administrator can configure the default, user, and group quota space for any share. The default level is the quota limit for every user unless the administrator specifies otherwise. A user-level quota policy sets a specific amount of storage for a single user. For example, if an administrator allocates 1 GB, then the user cannot take more than 1 GB. A group-level quota policy extends a user policy to include all users for an entire Active Directory group, where each user can use the assigned quota value. For example, if the administrator sets a group's quota to 10 GB, then each member of that group can use 10 GB.

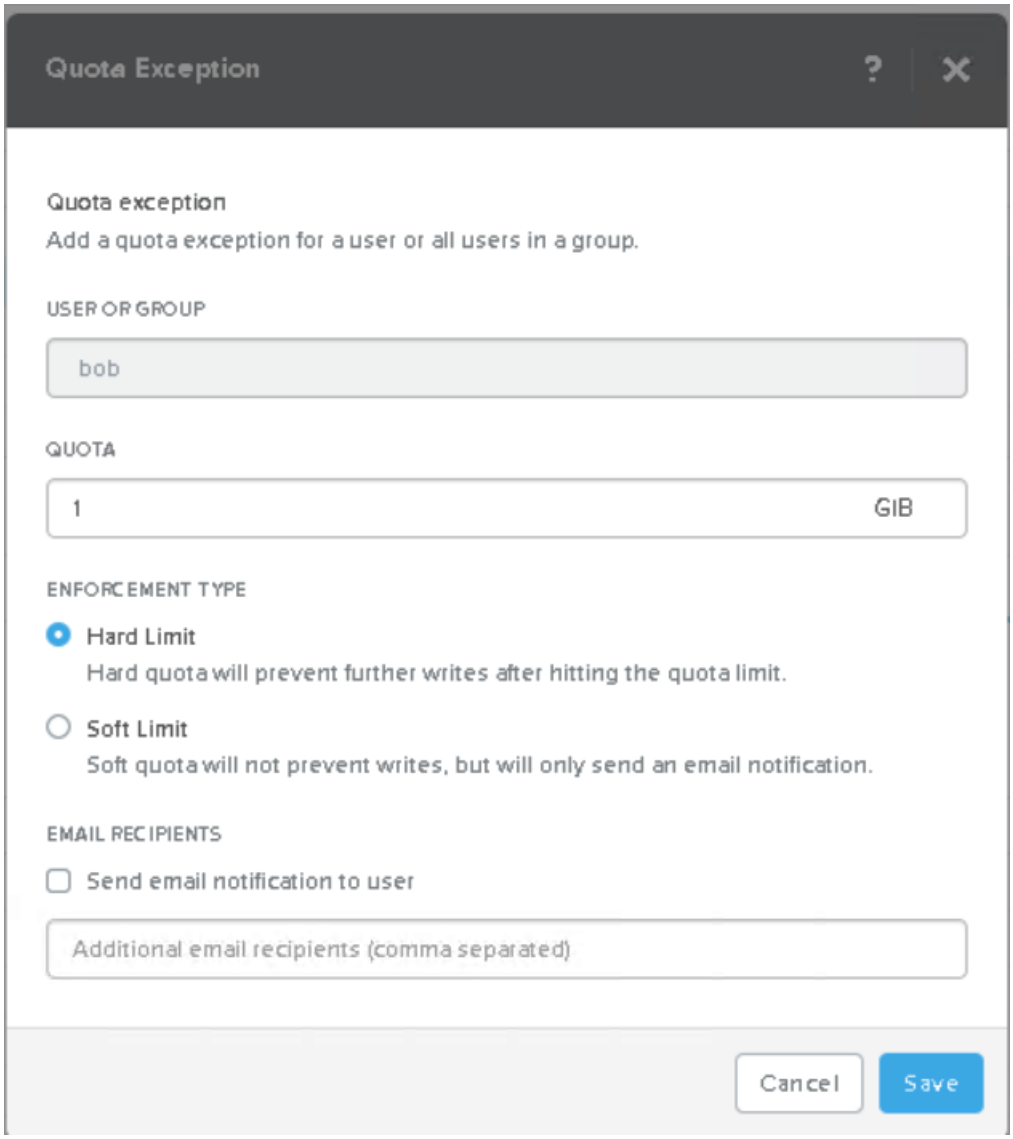

Figure 20: Setting a Share Quota

Notifications

Administrators can configure Prism to send email alerts to the user and to other recipients using the same engine that sends cluster alerts. Designated users receive email notifications when the quota is near maximum consumption—a warning email at 90 percent and an alert email at 100 percent. You can also add departmental share owners to the email notification list so they're aware that they may need to take action.

The table below shows the order of precedence when dealing with quotas for a share.

Table 3: Order of Precedence for Quotas

Order of Precedence	Policy
1	User Policy
2	Group Policy: The group policy with the highest quota wins
3	Default User Policy

Enforcement

The administrator can also configure enforcement types for each quota-level category. Enforcement types determine if a user or group can continue to use the share when they've consumed their quota. A hard enforcement type prevents the user from writing on the share when they reach their quota limit. If a user or application is writing to the share after it has reached the quota, the operation fails. A soft enforcement type allows a user to write even if they exceed the quota limit. Under either enforcement type, users over their quota receive an email notification every 24 hours until the issue is resolved.

5.6. Access-Based Enumeration

Access-based enumeration (ABE) is a Windows feature (SMB protocol) that filters the list of available files and folders on the file server to include only those that the requesting user can access. This filter helps both users and administrators, saving time for the user and worry for the administrator concerned about users accessing files not meant for them.

It's important to understand that ABE does not control security permissions, and that running ABE has associated overhead. Every time a user requests a browse operation, ABE must filter out objects for which the user doesn't have permission. Even if the user has permission to access all contents of the share, ABE still runs, causing additional CPU cycles and increased latency.

The home share is a great example of where not to enable ABE. Because most users get their home mapping on logon and always have access to their own contents, it doesn't make sense to enable ABE here.

5.7. Hypervisor-Specific Support

Nutanix supports ESXi and AHV for Files. For ESXi support, you need vCenter credentials to deploy Nutanix Files and to create DRS rules to make sure the FSVMs are on different nodes. The deployment process generates the DRS rules for AHV automatically.

5.8. Third-Party Integration

Antivirus

To protect users from malware and viruses, you need to address both the client and the file server. Nutanix currently supports third-party vendors that use Internet Content Adaptation Protocol (ICAP) servers. ICAP, which is supported by a wide range of security vendors and products, is a standard protocol that allows file and web servers to be integrated with security products. Nutanix chose this method to give customers wide latitude in selecting the antivirus solution that works best for their specific environment.

Following is the workflow for an ICAP-supported antivirus solution:

1. An SMB client submits a request to open or close a file.
2. The file server determines if the file needs to be scanned, based on the metadata and virus scan policies. If a scan is needed, the file server sends the file to the ICAP server and issues a scan request.
3. The ICAP server scans the file and reports the scan results back to the file server.
4. The file server takes an action based on the scan results:
   a. If the file is infected, the file server quarantines it and returns an “access denied” message to the SMB client.
   b. If the file is clean, it returns the file handle to the SMB client.

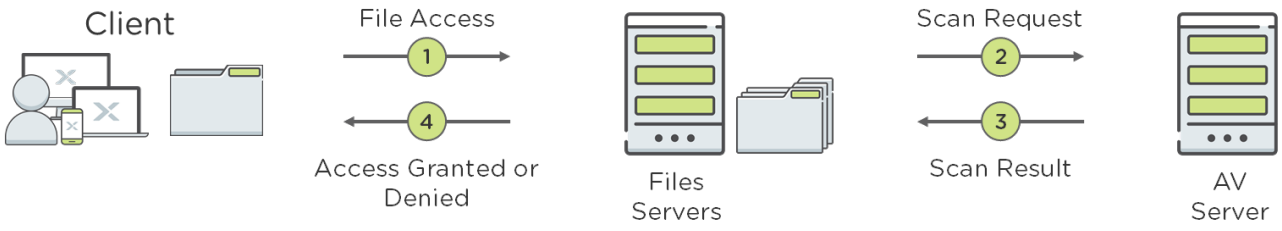

Figure 21: ICAP Workflow

The ICAP service runs on each Nutanix Files file server and can interact with more than one ICAP server in parallel to support horizontal scale-out of the antivirus server. We recommend configuring two or more ICAP servers for production. The scale-out nature of Files and one-click optimization greatly mitigates any antivirus scanning performance overhead. If the scanning affects Nutanix Files file server VM performance, one-click optimization recommends increasing the virtual CPU resources or scaling out the file server VMs. This feature also allows both the ICAP server and Files to scale out, ensuring fast responses from the customer’s antivirus vendor.

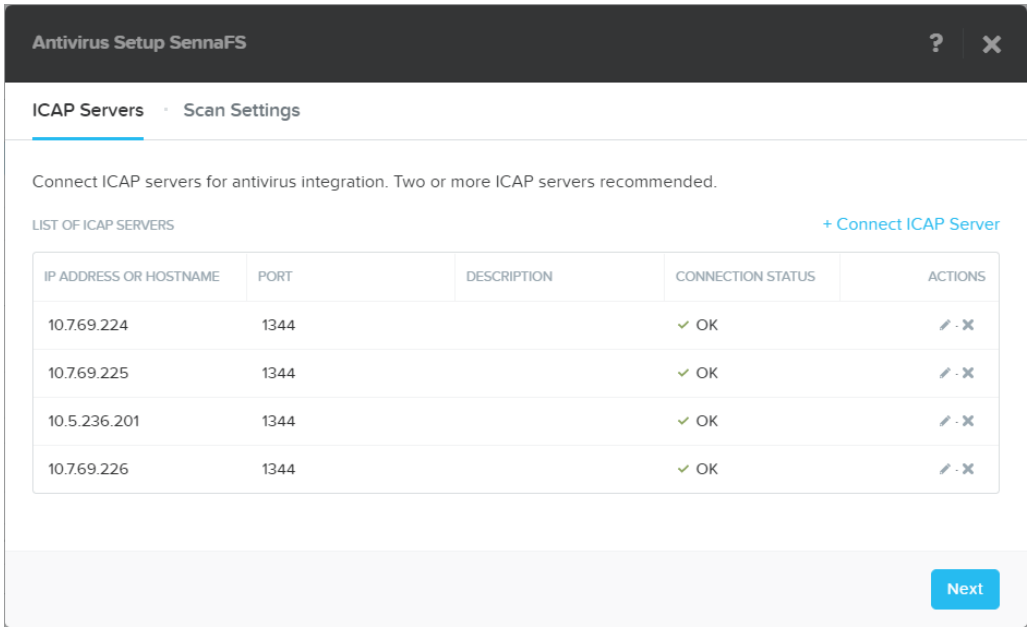

Figure 22: Configure Multiple ICAP Servers

Nutanix Files sets scanning defaults across the entire file server, but they are disabled by default per share when you enable file scanning. You can enable scan on write and scan on read. Scan on write begins when the file is closed, and scan on read occurs when the file is opened. You can also exclude certain file types and files over a certain size. Share scan polices can override any defaults set for the file server.

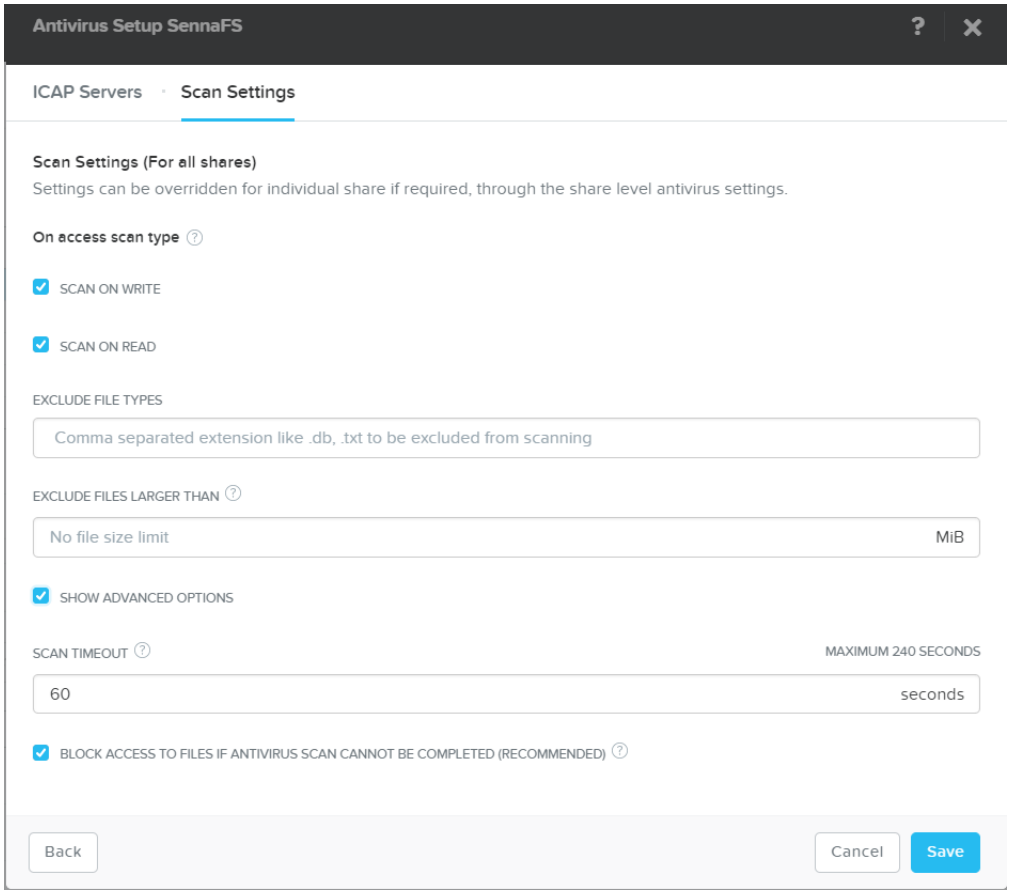

Figure 23: Default Scan Settings

For each ICAP server, we spin up no more than 10 parallel connections per FSVM and randomly dispatch the file scanning among all the ICAP servers. With heavier workloads, which may encounter many scan requests and use all connections, the scan servers with more processing power scan more files. As soon as the current scan finishes, the next file is picked up from the queue, which keeps the number of active connections at 10.

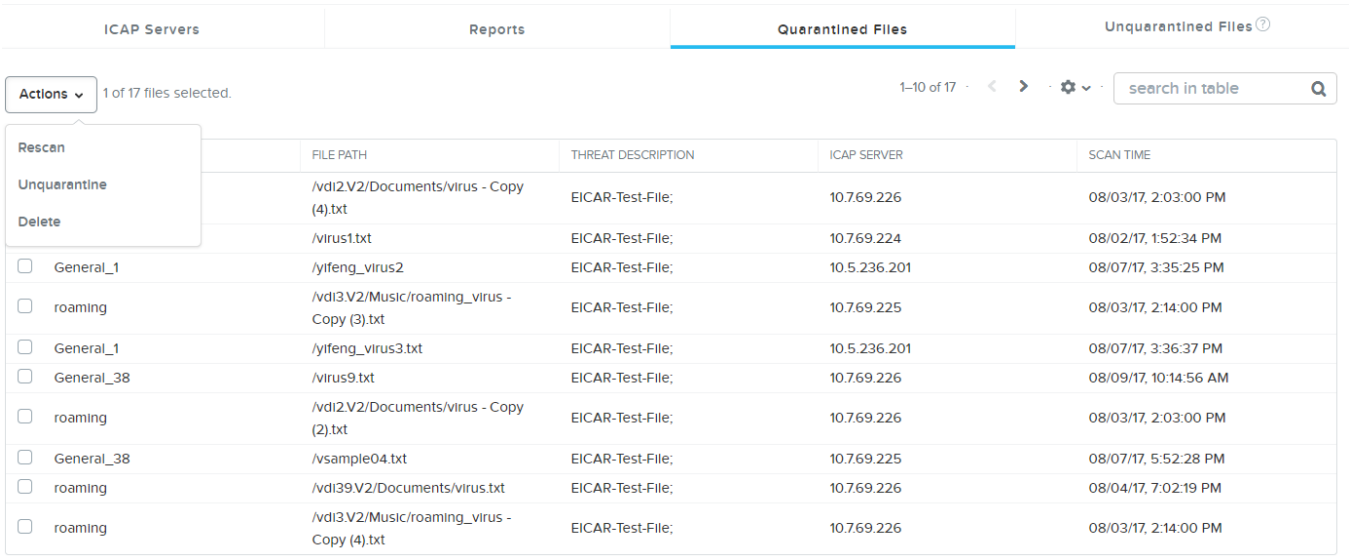

Figure 24: Quarantined Files

Once Nutanix Files quarantines a file, the admin can rescan, unquarantine, or delete the file. Quarantined files can be searched if it is necessary to restore a file quickly.

If your antivirus vendor doesn't support ICAP, you can scan the shares by installing an antivirus agent onto a Windows machine and then mounting all the shares from the file server. This approach allows you to schedule scans during periods of low usage. At the desktop or client level, you can set your antivirus solution to scan on write or scan only when files are modified. You can configure high-security environments to scan inline for both reads and writes.

5.9. File Operations Monitoring

File operations monitoring tracks Nutanix Files file notifications, which enables Files to send alerts to users about protocol file events that are happening for SMB shares.

File operations monitoring can be broken down into two major areas for third-party vendors:

- File activity
- Audit

Active-Active Deployments

Partner software can register with the Nutanix Files instance in their software. Once completed, the partner software makes the discover REST API call to Files, which returns the list of targets (file shares). Partner software can then make a REST call to Files to create the policy defining which file notifications they want to receive. The partner software uses a web client

and communicates with the Nutanix Files file server using HTTPS requests. The HTTPS communication relies on SSL authentication. The HTTP server runs with its unique self-signed SSL certificate or the option for a Transport Layer Security (TLS) connection with the partner, which then exchanges the keys between the two parties and sends messages to the partner server over this secure channel.

Intelligent Backup

Nutanix has developed a change file tracking (CFT) API that third-party backup vendors use. The new backup API speeds backup times by not doing a metadata scan across your file server, which could contain millions of files and directories. The API also reassures customers that they are not locked into any one backup vendor.

CFT uses internal snapshots to track differences between snapshots. The backup software tracks files that have changed between snapshots and lets Nutanix Files know which snapshot was the last one they backed up. Files returns a list of the changed files and proceeds to mount the share onto the backup server to be read.

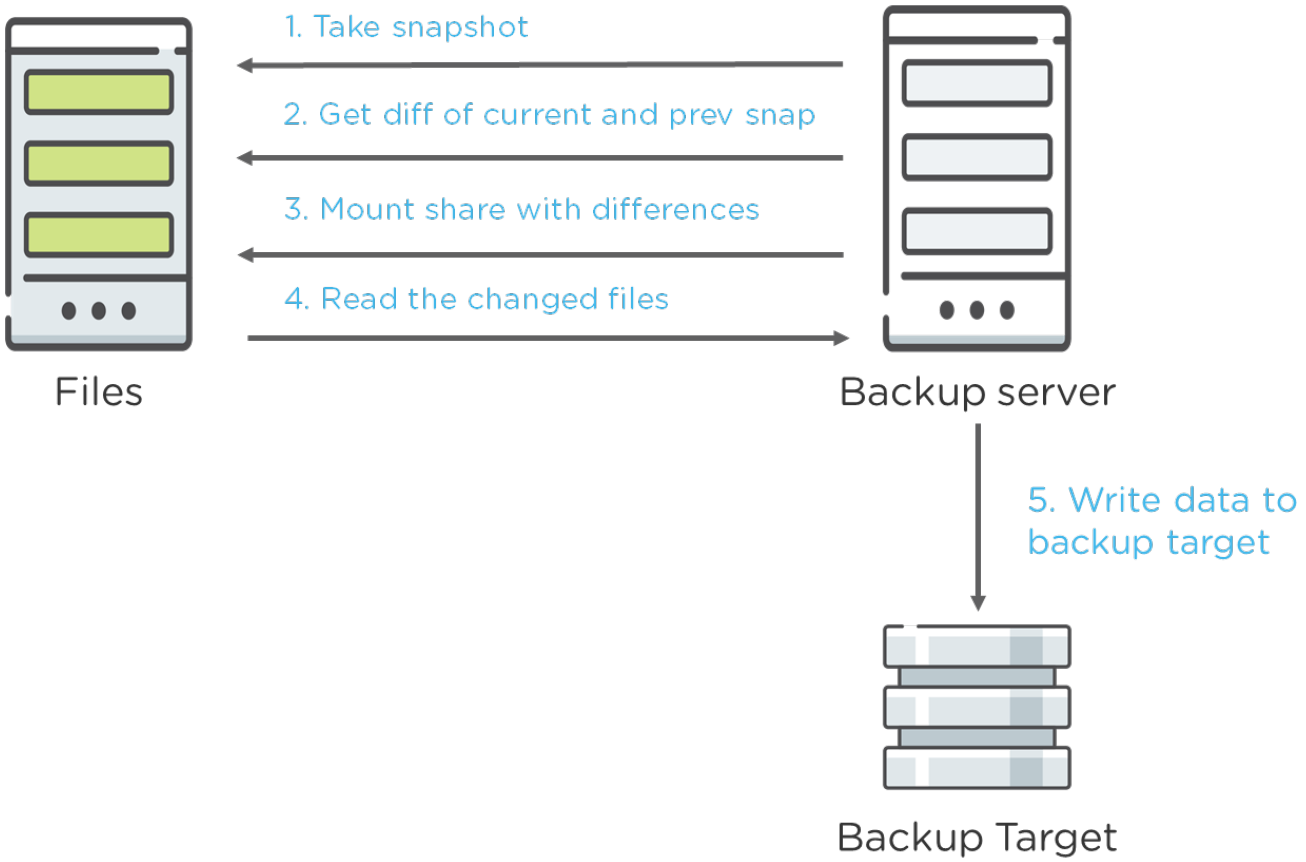

Figure 25: CFT Backup Process

Backup software can specify multiple shares and their respective snapshot information. Nutanix Files returns the list of URLs that map to the number of client streams that can start in parallel. Because shares are distributed evenly across the FSVMs based on load, the backup software can take advantage of all the FSVMs to drive throughput.

There are two ways to back up Files shares with software that doesn't support CFT. One option is to run the backup application on a Windows machine and map the UNC path of the share as a drive that needs to be backed up.

There are also vendors (such as Commvault, Rubrik, and Veritas) that provide support for backing up file shares without mounting to a guest VM. These applications can read directly from the UNC path. Because the system spreads different general shares across the cluster, try to back up multiple shares at the same time with multiple subclients. The home share allows you to configure multiple data readers to drive throughput.

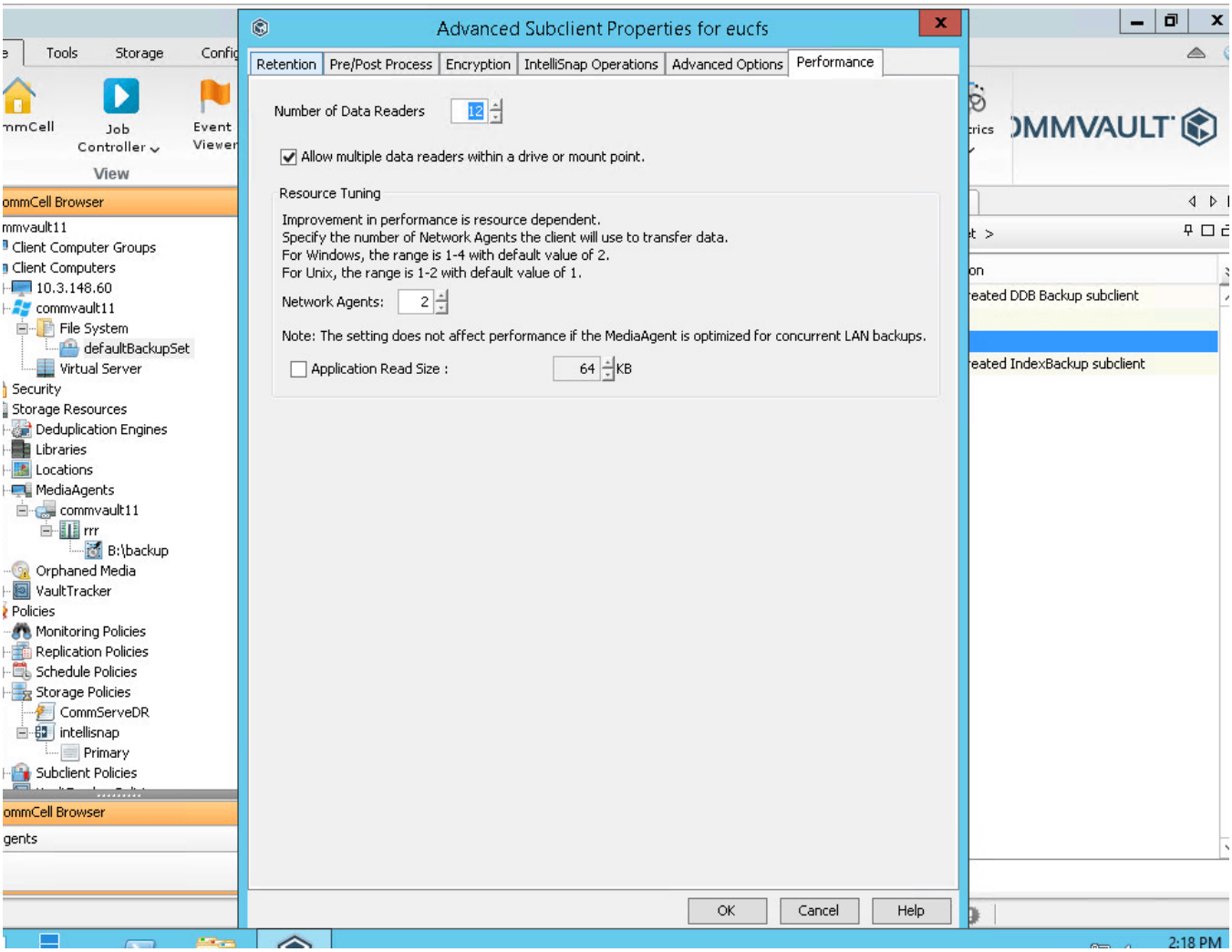

Figure 26: Adding More Readers to Drive Throughput on a Home Share

As an example of backup using Commvault, we tested 400 users spread out on three FSVMs, placing the data on a home share. We found that adding more readers for the backup job could increase performance. The bottleneck was the media agent, which was the backup destination for the files. The media agent was virtualized and configured with only eight vCPUs. Adding more CPUs to the media agent achieved a faster backup time.

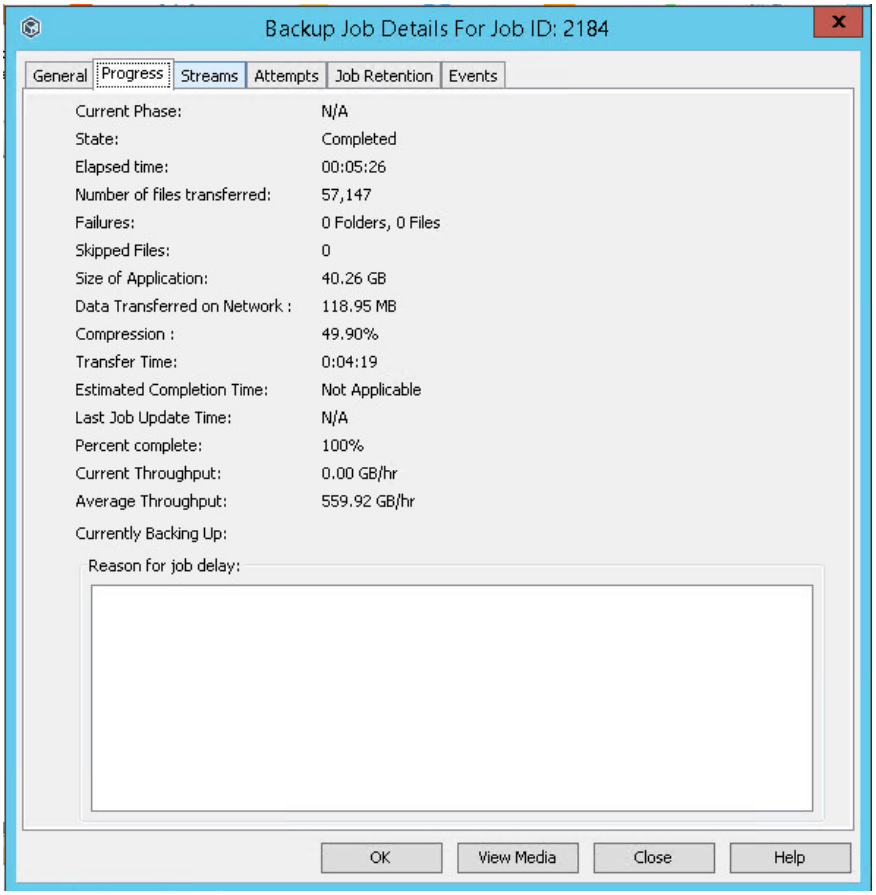

Figure 27: Backup Job Details

As the preceding figure shows, we achieved almost 600 GB per hour with a fairly small setup.

6. Conclusion

Nutanix Files delivers on-demand performance and automated provisioning to provide highly scalable file management. It reduces the administration and configuration time needed to deploy and maintain your environment, providing a public cloud experience within your private cloud.

The Acropolis Distributed Storage Fabric (DSF) and Prism make hosted file services highly resilient and easy to use—for example, you can configure the native data efficiency features of the DSF, such as compression, deduplication, and Erasure Coding-X (EC-X), for each individual file server. Prism lets you administer network and resource management, Active Directory, fault tolerance, and Nutanix Files share and export management all in one place, vastly improving operational efficiency.

You can deploy Files on a new or existing Nutanix environment, which means you have the flexibility to use your existing investment to take advantage of Files services.

Nutanix Files streamlines design, implementation, and maintenance, which provides an unrivaled user experience, an essential element of success for end-user computing projects that store data in shared environments.

Appendix

About Nutanix

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power their business. The Nutanix Enterprise Cloud OS leverages web-scale engineering and consumer-grade design to natively converge compute, virtualization, and storage into a resilient, software-defined solution with rich machine intelligence. The result is predictable performance, cloud-like infrastructure consumption, robust security, and seamless application mobility for a broad range of enterprise applications. Learn more at www.nutanix.com or follow us on Twitter @nutanix.

List of Figures

List of Tables