v1.0.1 | June 2022 | NVD-2151

NUTANIX VALIDATED DESIGN

Hybrid Cloud: Unified Storage Design

NUTANIX
YOUR ENTERPRISE CLOUD

Copyright

Contents

1. Executive Summary

Nutanix continues to innovate and engineer unified storage solutions that are simple to deploy and operate. To further improve customer experience and add value for customers, Nutanix uses robust validation to simplify the process of architecting and deploying storage solutions. This document details the design decisions that support the deployment of a scalable, resilient, and secure unified storage solution with two datacenters for high availability and disaster recovery.

> Note: In this design each datacenter resides in its own availability zone (AZ). This document denotes location exclusively by AZ.

Nutanix can deliver this Nutanix Validated Design (NVD) as a bundled unified storage solution that includes hardware, software, and services to accelerate and simplify the deployment and implementation process.

You can add this flexible and scalable design as a module to the core Hybrid Cloud NVD, or you can deploy it as a standalone unified storage solution. If you're deploying the Unified Storage NVD by itself, you might still need to reference parts of the core Hybrid Cloud NVD, as this document omits some of the information detailed there to avoid repetition.

You can have this validated unified storage solution up and running in weeks with minimal burden on your internal teams, allowing you to realize the full value of your infrastructure quickly. After you place your order, Nutanix takes care of the rest.

Audience

This guide is part of the Nutanix Solutions Library. We developed it for architects and engineers responsible for scoping, designing, installing, and testing file and object storage solutions.

Purpose

This document describes the components, integration, and configuration for the NVD packaged unified storage solution (an optional module of the core Hybrid Cloud NVD) and covers the following topics:

- Core Nutanix storage infrastructure and related technology.
- Backup and disaster recovery for Nutanix storage services.
- Test plan.
- Bill of materials.

Table 1: Document Version History

Version Number	Published	Notes
1.0	May 2022	Original publication.
1.0.1	June 2022	Updated the Bill of Materials section.

2. Storage Infrastructure Design

The following tables provide core storage infrastructure design requirements, risks, and restraints.

Table 2: Storage Infrastructure Design Requirements

Component	Description
Management	Deploy a unified management plane at the right scale to manage all clusters and workloads in the environment.
Management	Configure management to integrate with Active Directory for authentication.
Management	Use Active Directory–based groups for access control.
Monitoring	Enable platform fault monitoring and use email to send alerts.
Monitoring	Enable Data Lens anomaly detection, ransomware protection, and monitoring and use email to send alerts.
Monitoring	Monitor performance metrics in the Files and Objects control planes.
Monitoring	Monitor resources critical to Nutanix Files and Objects operations (for example, CPU, memory, storage, and network resources); resource usage that exceeds configured limits generates an alert.
Monitoring	Use email as the primary channel for event monitoring alerts.
Connectivity	Support the SMB, NFS, and S3 protocols.
Capacity and Performance	Usable storage capacity exceeds 200 TiB.
Capacity and Performance	Support a working set size of 1 TiB.

Component	Description
Capacity and Performance	Colocate applications and data to reduce unnecessary WAN traffic, minimize application latency, and reduce network congestion.
Business Continuity and Disaster Recovery	Achieve RPOs of 1 minute, 1 hour, and 24 hours.
Business Continuity and Disaster Recovery	Achieve RTO of 1 hour.
Business Continuity and Disaster Recovery	Provide nightly incremental backup, file-level and share-level granular restore, and self-service restore.
Infrastructure	Minimize cost of storing cold data.
Security and Analytics	Provide ransomware protection.
Security and Analytics	Provide network-level security to isolate protocol services for virus and worm threats.
Security and Analytics	Enforce segregation of duties: storage administrators shouldn't have management access for workloads that aren't storage.
Security and Analytics	Provide file system analytics for performance and capacity management, audit reporting, anomaly detection, and ransomware policies.

Table 3: Storage Infrastructure Design Risks

Component	Description
Monitoring	If Prism Central becomes unavailable for any reason, the platform can no longer send alerts. To mitigate this risk, configure each Prism Element instance to send alerts as well. As this approach results in duplicate alerts during normal operations, send Prism Element alerts to a different mailbox that you can monitor when Prism Central is unavailable.

Component	Description
Management	If Prism Central becomes unavailable in an AZ, the Objects service in that AZ continues to function, but object store creation, upgrade, and scale-out operations are not available. If the affected Prism Central instance is managing the Files clusters, Smart Disaster Recovery (Smart DR) replication continues to work, but you can't make changes to file server data protection policies.
Management	If the Prism Central instance managing the Files clusters becomes unavailable and the file server must fail over to the other AZ, you must issue a command from the file server virtual machine (FSVM) command-line interface (CLI).

Table 4: Storage Infrastructure Design Constraints

Component	Description
Files Clusters	The number of FSVMs per cluster in a fully scaled Files deployment doesn't exceed 16. The number of FSVMs doesn't exceed the number of physical nodes in the cluster.
Objects Clusters	The number of load balancer VMs doesn't exceed four. The number of worker VMs doesn't exceed the number of physical nodes in the cluster.

Component	Description
Business Continuity and Disaster Recovery	— A file server can only have one tiering profile. Therefore, in a scenario where both AZs contain active file servers, each must replicate to a target file server that doesn't already have a tiering profile (a passive file server). Because you're replicating to a file server that isn't already involved in a tiering relationship, the tiering profile can successfully fail over to the passive file server without conflict. — Inline reads of tiered data files are possible after failover, but data tiering and recall operations are blocked.
Business Continuity and Disaster Recovery	Smart DR requires that the same Prism Central instance manage both the source and target Files clusters. Smart DR supports a maximum of 80 standard shares and 20 distributed shares per file server pairing. The maximum number of shares that can be replicated with a frequency of less than 10 minutes is 5 with the distributed share type or 20 with the standard share type.
Monitoring	SMTP is an available channel in the environment that can receive event monitoring alerts. Because Objects doesn't natively integrate into the platform's SMTP capability, you must use a workaround. Syslog, which is supported by both Files and Objects, captures logs but doesn't generate alerts on events.

Storage Infrastructure Conceptual Design

The conceptual pod design has the following features:

- Two active-active datacenters in separate AZs.
- Two physical storage clusters in each AZ: one hosts Objects, and the other hosts Files.

- An instance of Prism Central hosted on the Objects cluster in each AZ.
  - The Prism Central instance in AZ1 manages the Objects cluster in AZ1.
  - The Prism Central instance in AZ2 manages the Objects cluster in AZ2, the Files cluster in AZ2, and the Files cluster in AZ1. Note that Smart DR requires that the same Prism Central instance manage both the source and target Files clusters.
- Smart tiering of cold file data between the Files and Objects clusters in each AZ.
- Bidirectional replication between the Files clusters in each AZ.
- Bidirectional replication between the Objects clusters in each AZ.

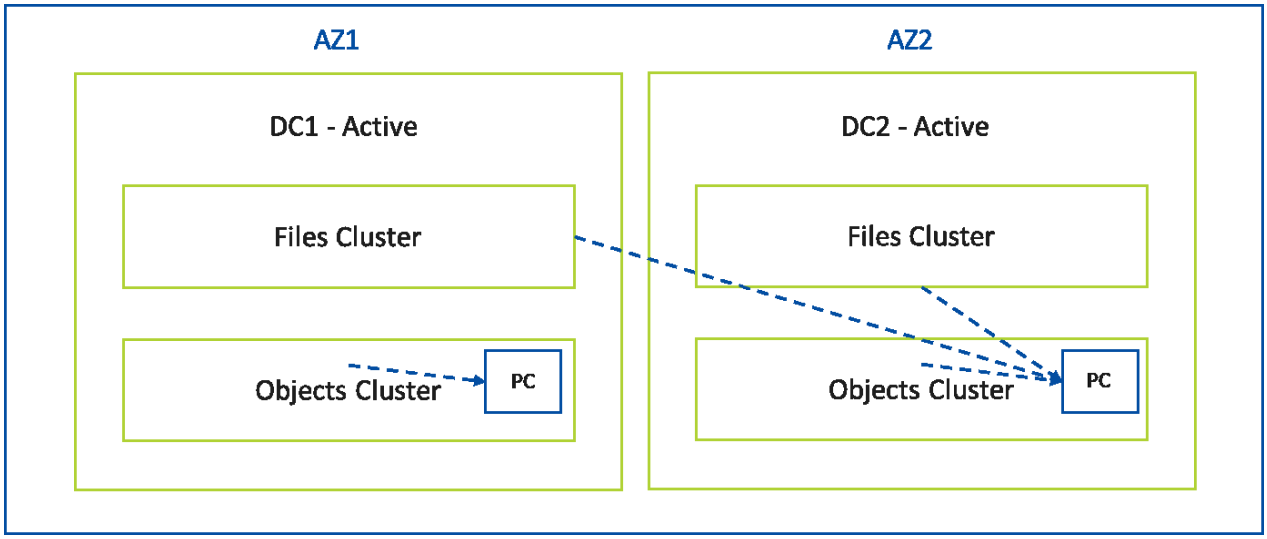

Figure 1: Conceptual Storage Design

Scalability

Scalability is one of the core concepts of the Nutanix platform and, in the context of unified storage, refers to the ability to increase storage space and storage processing capacity to meet both current and future SMB, NFS, and S3 workload demands. A well-designed cluster meets current requirements while providing a path to support future growth.

Scalability Conceptual Design

This unified storage NVD allows vertical and horizontal scaling within the boundaries set by running storage workloads in a single rack per AZ across two AZs. If the workload grows, you can add nodes and storage capacity to the cluster. This design has a maximum of 8 nodes per Files cluster and per Objects cluster, for a total of 16 nodes per AZ; if you need to scale beyond that number, you can deploy additional nodes in an adjacent rack.

Such expansion is beyond the scope of the current design. However, if you need to expand the Objects namespace beyond 8 nodes, you can deploy a new Nutanix cluster in another rack and use the Objects multicluster functionality to add it to the existing Objects namespace. With the Smart Tiering integration, expanding the Objects deployment in this way also makes more storage capacity available to the Files namespace.

> Note: If the infrastructure changes in one AZ, you must upgrade the other AZ accordingly to ensure that a failover can complete successfully. Identical configurations are required between AZs.

This NVD provides separate, dedicated clusters in each AZ for Files and Objects workloads, so you can expand allocated vCPU and memory for both service types concurrently without the risk of CPU or memory sharing between the Files and Objects services. Because the most common storage use cases tend to require storage density, this design uses a hybrid disk configuration with a mixture of SSDs and high-capacity HDDs.

The Nutanix model used in this design puts 8 workload nodes in 16 rack units (4 nodes in 8RU for Files and 4 nodes in 8RU for Objects). The design uses a single rack per AZ, with redundant top-of-rack network switches. This approach reduces operational complexity, but the large form factor of these storage-dense nodes limits the number that can fit in a single rack. Power and cooling limitations might introduce further constraints. For more information on hardware and network-related limitations, review the core Hybrid Cloud NVD.

Table 5: Scalability Design Decisions

Decision Name	Decision
Node memory population	Use 128 GB for Objects nodes, 256 GB for Files nodes
Node drive type	Use a mix of SSDs and HDDs
Node drive population	Fully populate the node drive
Single rack	Use one rack per AZ
Establish scalability boundaries	Use X-Ray to confirm load per node

At the software level, configuration maximums also constrain solution scalability. For the latest Files limits, refer to the configuration maximums in the Files Release Notes (portal account required). For the latest Objects limits, refer to the Limitations section of the Objects User Guide.

Table 6: Configuration Maximums or Maximum System Values

Entity	Architectural Maximum	Number Included in Unified Storage NVD
FSVMs per file server	16	4
Concurrent client connections	64,000	8,000
Single share size	5 PB (usable)**	205.2 TiB (usable)* **
Worker VMs per object store	Same as node count (up to 48)	3
Load balancers per object store	4	2
Object store metadata scaling	80 PB (raw)**	218.47 TiB (usable)* **

* Accounts for n + 1.

** Doesn't account for potential savings from compression and erasure coding.

Scaling Beyond Local Cluster Limits

The storage design uses the Smart Tiering feature to facilitate scaling file data capacity beyond the limits of the local physical Files cluster. The Data Lens service provides access to Smart Tiering. Once you've connected the file servers to the Data Lens service, you can define the tiering endpoint for each active file server and create a tiering policy.

Smart Tiering operates on file data based on last access time, moving cold file data out to an object storage endpoint and leaving a stub behind in the share's active file system. The tiering endpoint can be either a public cloud or an on-premises object store. In this NVD, the endpoint is the Nutanix Objects cluster residing in the same AZ as the file server.

Table 7: Smart Tiering Policy Decisions

Policy	Timeframe	Reads within timeframe	Capacity threshold
Tier	6 months	0	70%
Recall	1 day	3	N/A

Resilience

Nutanix provides many resilience features in the infrastructure where both Files and Objects run, including storage replication, snapshots, degraded node detection, and self-healing. These capabilities increase the resilience of unified storage workloads. Nutanix layers these software features on hardware designed with resilience in mind (for example, with redundant physical components and power supplies, many of which are hot-swappable or otherwise easily serviceable). Running workloads in a virtualized environment adds another kind of resilience, as you can perform many maintenance operations without application downtime. A resilient network fabric that can sustain individual link or node failures without significant impact completes the architecture. For more information on hardware and network-related limitations, review the core Hybrid Cloud NVD.

High Availability at the Files Level

File servers with three or more FSVMs achieve high availability during node failure and Files upgrades at the file server level. Resources under the control of a failed FSVM or an FSVM that's being upgraded temporarily move to another FSVM. For single-FSVM instances, hypervisor high availability provides protection against node failure events.

High Availability at the Objects Level

Objects provides built-in resilience in the event of worker VM failure. If a worker VM fails, the load balancers automatically redirect clients to surviving worker VMs. Note that any in-flight requests to the failed worker VM might fail, causing the client to resend the request.

To ensure that the system doesn't direct clients to a failed load balancer, each AZ has a global server load balancer (GSLB) deployed in front of the object store. The GSLB used in this NVD is F5 BIG-IP. All client lookups of the object store namespace are handled by the GSLB, which forwards the client requests to the Objects load balancers in the local AZ. The GSLB detects when a load balancer VM fails and ceases to forward client requests to that particular load balancer until it comes back into service. The GSLB also makes disaster recovery failovers between AZs much more seamless than would otherwise be possible by detecting a full AZ-level failure and forwarding all client requests to the load balancer VMs in the surviving AZ.

Resilience Conceptual Design

The workload cluster sizing allows for n + 1 failure redundancy. Monitoring and alerting ensure that any issues result in an alert; consistently monitoring workload growth ensures that sufficient headroom is available at any time.

Table 8: Resilience Design Decisions

Decision Name	Decision
Full redundancy of all components	Ensure the full redundancy of all components in the AZ

Decision Name	Decision
Established resilience boundaries	Use X-Ray to find resilience constraints

File Server Design and Object Store Design

To achieve high performance for both the Files and Objects workloads, this design minimizes overallocation of physical resources where possible. If Files and Objects are located on the same physical cluster (a supported configuration), sharing of certain resources (for example, CPU cores and resources for the Nutanix Controller Virtual Machine (CVM) between the environments is unavoidable. However, this storage NVD provides separate physical clusters for Files and Objects, removing any risk of contention between the two storage service types.

Files Deployment Sizes

The amount of compute resources that you can allocate to the FSVMs in Nutanix Files is configurable. This design replicates smart-tiered file servers between AZs. Because each file server can have only one tiering profile (see the earlier table Storage Infrastructure Design Constraints), you must deploy two file server instances on each physical Files cluster (one active and one passive) so that there are two FSVMs on every physical node. The base deployment allocates 6 vCPU and 32 GB of RAM to each FSVM because the compute specification of the underlying physical nodes is dual 12-core CPUs and FSVMs can potentially achieve better performance when they remain within the boundaries of a NUMA node.

If each of our coexisting FSVMs has 6 vCPU, they can coexist on the same 12-core NUMA node without having to share cores. For more information about the number of client connections a given FSVM configuration can support, refer to the Nutanix Files Performance tech note (Nutanix customer account required). The other NUMA node in these dual-socket nodes is dedicated to the Nutanix CVM.

You can add compute resources nondisruptively to the FSVMs to support more client sessions or deliver greater throughput. Scaling up the FSVMs in

this way introduces a degree of core sharing; however, even with both FSVMs at maximum size, the sharing doesn't exceed 2:1. As one of the two FSVMs only receives incoming replication traffic from the other AZ, sharing is unlikely to have significant effects. Note that you can also scale FSVMs down for full flexibility.

The following table details the minimum and maximum FSVM sizes. To scale, you can either expand the existing FSVM compute resources (scaling up) or add more FSVMs (scaling out).

Table 9: Supported FSVM Configurations

FSVM Size	Minimum	Maximum
Virtual CPU	4	24
Virtual memory	12 GB	512 GB
Maximum concurrent connections per FSVM	500	4,000
FSVMs per cluster (regardless of FSVM configuration)	3	16

Objects Deployment Sizes

Objects has one fixed size for the worker VM and one fixed size for the load balancer VM, both detailed in the following table. To scale Objects, add more worker VMs (scaling out).

This design deploys a single Prism Central VM (small size) allocated with 6 vCPU on each Objects cluster. This configuration results in a modest amount of core sharing with one of the Objects worker VMs. However, only one node is sharing; even with a load balancer VM on the same node, oversubscription doesn't exceed 2:1.

Table 10: Supported Objects VM Configurations

Item	Worker VM	Load Balancer VM
Virtual CPU	10	2
Virtual memory	32 GB	4 GB
Maximum per cluster	The total number of nodes in the AOS cluster	4

Cluster Design

This section defines the overall high-level cluster design, platform selection, capacity management, scaling, and resilience.

This design includes, in each AZ, one dedicated physical cluster for Objects and another for Files, with the Prism Central management layer located on the Nutanix Objects cluster. There are no dedicated physical clusters for management because the additional costs they incur aren't warranted. You could alternatively use the Prism Central instances from the core Hybrid Cloud NVD (on dedicated management infrastructure) to manage the Files and Objects clusters, but this approach lacks role-based access control (RBAC) for Files management.

Backup: Files

This design uses Nutanix Mine infrastructure, described in the core Hybrid Cloud NVD, to back up file data. Several backup vendors can very efficiently back up Files using the Changed File Tracking (CFT) API. This NVD uses HYCU as the backup application because HYCU supports the CFT API and provides a simplified, efficient, and well-integrated backup capability for Nutanix Files.

Backup: Objects

Instead of backing up the object stores, this NVD protects object data by replicating buckets between the Objects clusters in each AZ. Bucket replication is the conventional approach to protecting object stores, which often grow too large to be backed up using traditional methods.

Cluster Conceptual Design

In keeping with the core Hybrid Cloud NVD, this NVD uses one region with two separate AZs. Both AZs host active workloads, and each AZ provides a replication target for the other's data. Each Objects deployment is managed by a Prism Central instance in the respective local AZ.

The Files deployments are both managed by the same Prism Central instance, located in AZ2, because Smart DR requires both the source and target file servers to be under the management of the same Prism Central.

Table 11: Cluster Design Decisions

Decision Name	Decision
Number of regions	Use 1 region
Number of AZs	Use 2 AZs
Number of datacenters	Use 2 datacenters: 1 per AZ
Mixed workloads or dedicated workload per cluster	Use a dedicated workload per cluster
Minimum workload cluster size	Use at least 4 nodes per workload cluster (Files and Objects are on separate clusters)
Maximum workload cluster size for this design	Files: Use 8 nodes (single rack constraint)Objects: Use 8 nodes (single rack constraint)
Cluster node redundancy	Use n + 1 for redundancy (safely usable storage capacity)
Maximum usable nodes for storage capacity	Consume at most 3 usable nodes for storage to ensure that you can always rebuild data with replication factor 2
Maximum usable nodes for storage processing compute	Configure all 4 nodes to optimize the investment; the loss of a node might cause performance degradation
Workload clusters in one rack or split across multiple racks	Use one rack per AZ for both Files and Objects clusters
Cluster replication factor	Use replication factor 2

Decision Name	Decision
Cluster high availability configuration	Guarantee high availability
Percentage of client connections supported during disaster recovery failover	Support 100 percent of client connections; there might be some performance degradation when hardware resources experience increased demand
Maximum usable resource capacity per cluster to allow for disaster recovery failover	With workloads split evenly between AZs, 50 percent of the storage resources at each AZ can serve data locally while the other 50 percent accommodates incoming replication from the other AZ

Platform Selection

Table 12: Platform Selection

Cluster	Files	Objects
Node type	NX-8155-G8	NX-8155-G8
Node count	4	4
Rackspace (per node)	2RU	2RU
Processor	2 Intel Xeon Silver 4310 12-core 120 W, 2.1 GHz (Ice Lake)	2 Intel Xeon Silver 4310 12-core 120 W, 2.1 GHz (Ice Lake)
RAM	8 x 32 GB 3,200 MHz DDR4 RDIMM (256 GB total)	4 x 32 GB 3,200 MHz DDR4 RDIMM (128 GB total)
SSD	4 x 7.68 TB	2 x 3.84 TB
HDD	8 x 18 TB	10 x 18 TB
NIC	25 GbE Dual SFP+	25 GbE Dual SFP+
Support	3Yr Production	3Yr Production

Capacity Management

In this design, each Files cluster can grow from a minimum of four nodes (8RU) with three nodes of usable storage capacity to a maximum of eight nodes

(16RU) with seven nodes of usable storage capacity. Likewise, each Objects cluster can grow from a minimum of four nodes (8RU) with three nodes of usable storage capacity to a maximum of eight nodes (16RU) with seven nodes of usable storage capacity. You can expand both the Files and Objects clusters in single-node increments up to the maximum. In both cases each additional node provides increased throughput, support for more client connections, and more storage capacity available to the environment. If you reach the maximum number of nodes prescribed by this NVD and need more, you can either deploy a new cluster in a new rack or expand the existing cluster by deploying new nodes in a neighboring rack. This storage NVD, however, stops at eight nodes for each storage service, all in a single rack.

Scaling Beyond the Constraints of the Storage NVD

The proportion of usable storage decreases as the cluster size decreases because the system reserves one node per cluster for maintenance and failure. Moreover, the reduction in usable capacity after the n + 1 allowance can be significant when using storage-dense nodes. Coupled with the fact that expanding an existing file server namespace is significantly more transparent than creating an entirely new namespace (each new Files cluster results in a new namespace), these factors mean that if you need to expand the Files cluster beyond 8 nodes, you should scale the Files cluster into a new rack.

Objects, unlike Files, has multicluster support, so an object store's namespace can span multiple physical clusters. However, the same AOS n + 1 overheads apply. Furthermore, you can only add new worker VMs in the initial AOS cluster. Therefore, to scale performance, if you need to expand the Objects deployment beyond eight nodes, you should add more nodes (in a new rack) to the existing underlying AOS cluster rather than create a new cluster.

> Note: Although scaling beyond eight nodes per workload is beyond the scope of the current design, eight nodes is neither an architectural nor a practical limit for either Files and Objects.

In situations where you need to scale capacity to accommodate growth in file data, you might need to scale the Mine cluster backing up the Files environment proportionately. Refer to the core Hybrid Cloud NVD for the details of the backup cluster design.

Cluster Resilience

Replication factor 2 protects against the loss of a single component in case of failure or maintenance. During a failure or maintenance scenario, Nutanix rebuilds any data that falls out of compliance much faster than traditional RAID data protection methods. Rebuild performance increases linearly as the cluster grows.

In the Nutanix architecture, rapid recovery in the event of failure is the standard, and there are no single points of failure. You can configure the cluster to maintain either two or three copies of data; to maintain three copies, you need at least five nodes.

Storage Design

Nutanix uses a distributed, shared-nothing architecture for storage. For a discussion of Nutanix storage constructs, refer to the Storage Design section in the Nutanix Hybrid Cloud Reference Architecture. For information on node types, counts, and physical configurations, see the Cluster Design section of the core Hybrid Cloud NVD.

Creating a cluster automatically creates a number of storage containers; for more information, refer to the core Hybrid Cloud NVD.

When you create a new file server, the system automatically creates a storage container dedicated to that file server. The name of the container conforms to the following format:

- Nutanix_<fileserver_name>_ctr

When you deploy an Object store, the system creates two storage containers: one for data and the other for metadata. These containers use the following naming convention:

- Data container: objectsd<uniqueidentifier>
- Metadata container: objectsm<uniqueidentifier>

Data Reduction Options

To increase the effective capacity of both the Files and the Objects clusters, this design enables erasure coding with the default strip size on the Nutanix Files and Objects containers.

> Note: Erasure coding requires a minimum of four nodes.

Deduplication isn't enabled for either the Files or Objects containers. Unless the application writing to Objects is already compressing the data, enable inline compression on the Objects containers. Files manages inline compression at the file server level on a share-by-share basis (enabled by default); because of the compression at the file server level, compression at the container level isn't enabled when you deploy a file server. The data reduction settings in the following table apply across clusters in both AZs.

Table 13: Data Reduction Settings

Container	Compression	Deduplication	Erasure Coding
Nutanix_ <file_server_name>_ctr	Off (enabled at file server level instead)	Off	On
Objects containers	On	Off	On

Table 14: Storage Design Decisions

Decision Name	Decision
Sizing a cluster	Files: Deploy 4 large SSDs per node to provide enough usable hot tier capacity to support the file server's working set. Objects: Deploy 2 medium SSDs per node to support Objects metadata (object data payload is directly written to and read from the HDD tier)
Node type vendors	Use all Nutanix NX nodes

Decision Name	Decision
Node and disk types	Use identical node types equipped with similar disks
Sizing for node redundancy for storage	Size all clusters for n + 1 failover capacity
Fault tolerance and replication factor settings	Configure the cluster for fault tolerance 1 and configure the container for replication factor 2
Inline compression	Enable inline compression at the container level for Objects and at the file server level (not the container level) for Files Exception: In both cases, if data is already compressed at the client level, disable compression on the Nutanix side
Deduplication	Don't enable deduplication
Erasure coding	Enable erasure coding

Network Design

A Nutanix cluster can tolerate multiple simultaneous failures because it maintains a set redundancy factor. However, this level of resilience requires a highly available network connecting a cluster's nodes. Refer to the core Hybrid Cloud NVD for more information on deploying a network to achieve the necessary resilience levels. Nutanix Files and Objects have the following additional network requirements for a four-node deployment.

> Note: In both cases the storage network addresses should be on the same network as CVM eth0. Files 4.1 supports segmented iSCSI traffic, but this NVD didn't test it. Objects doesn't support segmented iSCSI traffic.

Files

The following table details IP address requirements covering eight FSVMs per AZ (four FSVMs per file server, with two file servers).

Table 15: Nutanix Files IP Address Requirements

Network Type	Minimum IP Addresses per File Server	Minimum IP Addresses per AZ
Client	4	8
Storage	5	10

When scaling out the environment, allow for one more client network address and one more storage network address for every FSVM you add.

Objects

The following table details IP address requirements covering four worker VMs and three load balancer VMs per AZ.

Objects runs as a containerized service on a Kubernetes microservices platform, which provides benefits such as increased velocity of new features. Several of the required storage IP addresses are for functions relating to the underlying microservices platform. Note that you must also manage these networks.

Table 16: Nutanix Objects IP Address Requirements

Network Type	Minimum IP Addresses per AZ
Public	2
Storage	10

When scaling out the environment, allow for one more storage network address for every worker VM added.

You can't scale out load balancers after deployment, so you don't need additional IP addresses.

General

This design includes disaster recovery replication. With both Files and Objects, replication traffic travels across the client or public network interfaces to the corresponding target system on the destination AZ's client or public

network. As neither service currently supports synchronous replication with high availability across AZs, you don't need stretched layer 2 networking.

The Prism Central instance in AZ1, which manages the local Objects cluster, must have connectivity to the storage network and the public network connected to the object store in AZ1.

The Prism Central instance in AZ2, which manages the local and remote (AZ1) Files clusters as well as the local Objects cluster, must connect to:

- The storage networks to which the file servers in AZ1 and AZ2 are connected.
- The storage network and the public network to which the object store in AZ2 is connected.

For more information on the specific functions that require the IP addresses in the preceding tables, refer to the appropriate user guide.

Management Components

Management components such as Prism Central, Active Directory, DNS, and NTP are critical services that must be highly available. Prism Central is the global control plane for Nutanix; its unified storage management responsibilities include:

- Object store deployment and management.
- Objects IAM user and administrator management.
- Files Smart DR management.
- Centralized monitoring and management for both Files and Objects.

You can deploy Prism Central in either a single-VM or scale-out (three-VM) configuration. This design uses a single-VM Prism Central at each AZ to minimize the resource overhead that the management layer incurs.

When you design your management components, decide how many Prism Central instances you need. This NVD deploys one Prism Central instance in the Objects cluster in each AZ for a total of two Prism Central instances.

Management Conceptual Design

Although you can use the dedicated management clusters in the core Hybrid Cloud NVD to manage Files and Objects, this configuration has implications for RBAC in Files. Therefore, this storage design deploys a Prism Central instance in the Objects cluster in each AZ. The AZ1 Prism Central manages the local Objects deployment, while the AZ2 Prism Central manages the local and remote Files deployments in addition to the local Objects deployment. By deploying Prism Central instances dedicated to Nutanix storage services, you don't need to rely on RBAC to restrict the resources that storage administrators can access and manage.

Management Detailed Design

In this iteration of the unified storage NVD, the Files and Objects workload clusters run AOS 5.20.3, which is a long-term service (LTS) version.

Table 17: Nutanix Management Component Software Versions

Component	Software Version
Prism Central	pc.2022.1
AOS	5.20.3 (LTS)
Objects	3.4.0.2
Objects Manager	3.4.0.2
MSP	2.4.0
Files	4.1
Files Manager	2.0.2
File Server Module	2.2

The following table lists the design decisions for the Nutanix management components.

Table 18: Management Component Design Decisions

Decision Name	Decision
Management cluster architecture	One Prism Central VM in the Objects cluster at each AZ
Prism Central deployment structure	Single-VM Prism Central. Because storage services don't rely heavily on Prism Central, they don't warrant the resources required for scale-out
Prism Central deployment size	Small: 1 VM with 6 vCPU, 26 GB of RAM, and 500 GiB of storage
Prism Central deployment locations	One in each AZ
Prism Central container name	Default container
Active Directory authentication	Use Active Directory authentication
Connection to Active Directory	Use SSL or TLS for Active Directory

Monitoring

Hardware monitoring works the same way for all Nutanix clusters, regardless of workload. For more information on hardware monitoring, refer to the core Hybrid Cloud NVD. This section covers monitoring for metrics specific to Files and Objects.

Storage monitoring falls into two categories: event monitoring and performance or usage monitoring. Each category addresses different needs and different issues.

In a highly available environment, you must monitor events to maintain high service levels. When faults occur, the system must raise alerts in a timely manner so that administrators can take remediation actions as soon as possible. This NVD configures the Nutanix platform's built-in capability to generate Files-related alerts; alerts for Objects are generated in the Object Manager UI. You

can use a workaround method to bring Objects alerts into Prism Central so you can review them alongside alerts for Files and AOS.

In addition to keeping the platform healthy, maintaining a healthy level of resource usage is also essential to delivering a high-performing environment. Performance monitoring continuously captures and stores metrics that are essential when you need to troubleshoot performance issues in the Files and Objects environments. A comprehensive monitoring approach should track metrics specific to the Objects and Files services as well as metrics that describe the underlying hyperconverged platform, the network, and the physical environment.

By tracking a variety of metrics in these areas, the Nutanix platform can also provide capacity monitoring across the stack. Most enterprise environments inevitably grow, so you need to understand resource utilization and the rate of expansion to anticipate changing capacity demands and avoid any business impact caused by lack of resources.

If you need to forward Files or Objects log data to third-party systems, syslog support in Files allows you to do so, and Objects supports both NATS messaging and syslog.

This design uses Data Lens, which provides file analytics as a service, to gain insights into file usage trends, auditing events including anomalies and ransomware, and the overall composition of the file estate (broken down by file type, file size, file age, file operation type, and so on). The wealth of information that the Data Lens service returns can prove very useful in future decision making related to the Files environment. Data Lens can also raise alerts when it observes suspicious activity. Because Data Lens is a software-as-a-service running in the public cloud, it requires internet access.

Monitoring Conceptual Design

Prism Central performs most of the event monitoring at the underlying infrastructure level. Objects-specific alerts appear in the Objects management console in Prism Central (the Alerts tab presents all notifications and events), while Files-specific alerts and events appear in the Files Manager component in Prism Central.

This design uses SMTP-based email alerts as the channel for infrastructure-level and Files-related notifications. You can use a workaround to configure similar email alerting for Objects.

> Note: This NVD uses syslog for log collection; for more information, refer to the Security and Compliance section. Alerts from Prism Central go to a primary email alert recipient that's always monitored.

To cover situations where Prism Central might be unavailable, each Nutanix cluster in this NVD (the two Files clusters and two Objects clusters) sends out notifications using SMTP as well. The individual Nutanix clusters send alerts to a different receiving mailbox that's only monitored when Prism Central isn't available. Data Lens, a resilient Nutanix service, also provides SMTP alerting that you should configure for both the primary and secondary email recipients.

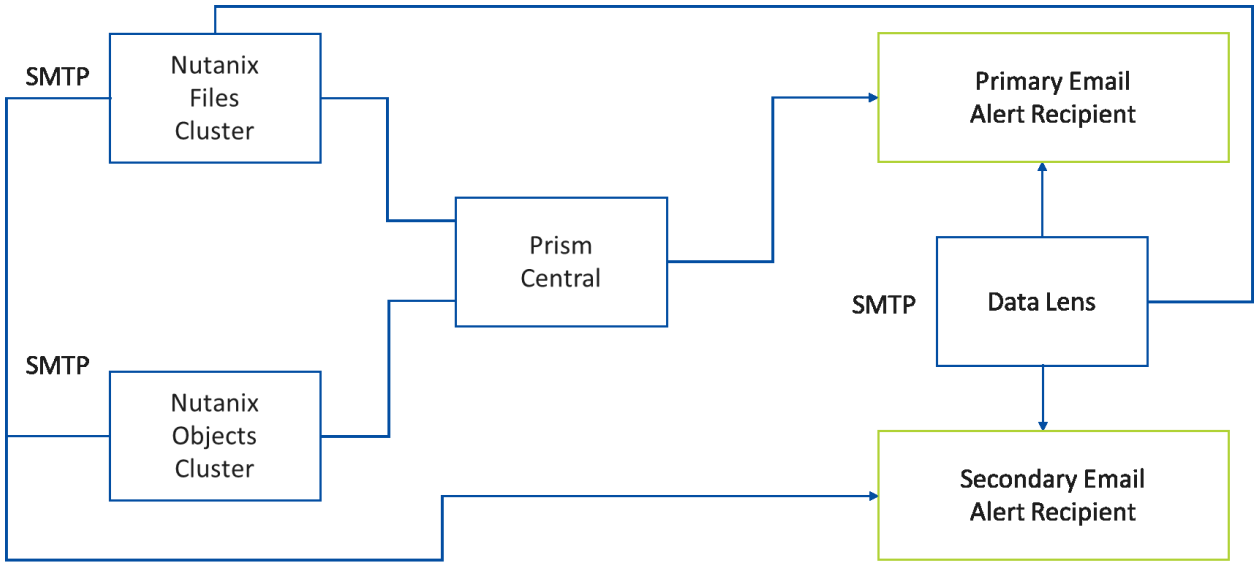

Figure 2: SMTP for Email Alerts from Prism Element and Prism Central

Prism Central monitors cluster performance in key areas such as CPU, memory, network, and storage utilization. The Objects management module monitors Objects-specific performance and usage rates, exposing metrics such as requests per second, time to first byte, throughput, and capacity consumption. Many of the reported metrics are available at both the object store level and the bucket level.

The Files Management Console monitors performance and usage information for Files, reporting on metrics such as file latency, throughput, IOPS, capacity consumption, open connections, and file count. Many of the metrics are available at both the file server level and the share level.

In all cases Nutanix captures these metrics, so you don't need to do much configuration.

The network switches that connect the cluster also play an important role in cluster performance. For more information on monitoring network switches, review the core Hybrid Cloud NVD.

Data Lens provides analytics-driven insights into file activity that can help you understand the composition of the file environment. Each Files cluster must have Pulse dial-home support enabled for the file servers to be visible to the Data Lens service. The file servers at each of the AZs can then have the service explicitly enabled in the Data Lens portal. Once Pulse and Data Lens are configured, you can set up email alerts in both the anomaly detection and ransomware modules and flag suspicious activity in real time.

The following table provides descriptions of the monitoring design decisions.

Table 19: Monitoring Design Decisions Specific to Unified Storage

Decision Name	Description
Objects performance monitoring	Objects Manager (a module in Prism Central)
Files performance monitoring	Files Management Console (a module in Prism Element)
Objects storage utilization monitoring	Objects Manager (a module in Prism Central)
Files storage utilization monitoring	Files Management Console (a module in Prism Element)
Cluster storage runway	Prism feature that predicts future cluster and storage container space utilization
Objects alerts	Objects Manager (a module in Prism Central) surfaces alert and event information

Decision Name	Description
Files alerts	Prism Central and Prism Element both surface Files-related alert and event information and email it according to Prism Central and Prism Element SMTP configuration settings
Data Lens alerts	Data Lens surfaces alerts relating to anomaly detection and ransomware attempts and sends them according to the configuration settings in the respective modules

Security and Compliance

Nutanix recommends a defense-in-depth strategy for layering security throughout any enterprise datacenter solution, which also applies to unified storage services. This design section focuses on validating the layers specific to Files and Objects that Nutanix can directly oversee at the control and data plane levels. Refer to the Network Design section in the core Hybrid Cloud NVD for more information on how the design delivers network-based security through microsegmentation.

Authentication and Authorization

Refer to the core Hybrid Cloud NVD for guidance on configuring the AOS and network components and integrating them with Active Directory authentication.

Because Active Directory is required for SMB authentication, this design uses it to apply user and group permissions.

For SMB file client authentication and permissions management, Nutanix recommends leaving the share permission settings at Full Control and managing access with NTFS permissions. Apply NTFS permissions using Active Directory–based security groups.

With NFS, Files supports Active Directory and LDAP for directory-based authentication. Files also supports leaving the environment unmanaged, using either system authentication or no authentication. This NVD tested system authentication.

With S3, the administrator generates client access keys for Objects using the IAMv1 service. You can generate these keys either for individuals or for all members of an Active Directory security group at once. Users then receive read, write, or no permission on a per-bucket basis.

You can integrate Files and Objects event logging into the syslog infrastructure as described in the core Hybrid Cloud NVD.

Prism and Object Store Certificates

With Objects, self-signed Secure Socket Layer (SSL) certificates are generated by default. For control plane security in Prism, the core Hybrid Cloud NVD describes replacing the default self-signed certificates with certificates signed by an internal certificate authority from a Microsoft public key infrastructure (PKI). You can choose alternative tools such as openssl for certificate generation and signing. These certificates secure communications between Objects components and Prism.

Generate additional sets of certificates in the same way, using the same certificate authority, and apply them to the object store in each AZ to provide strong security for S3 client connections using the HTTPS protocol. S3 clients that interact with Objects should have the trusted CA chain preloaded.

To facilitate disaster recovery, configure each object store with the FQDN of the other object store in addition to its own. This configuration allows each object store to respond to client requests intended for the other store. Therefore, to ensure strong security for S3 client connections in a disaster recovery scenario, apply certificates for both FQDNs to each of the object stores.

> Note: Certificate management is an ongoing activity, and certificates need to be rotated periodically. The NVD signs all certificates for one year of validity.

Data-at-Rest Encryption

Nutanix AOS can perform data-at-rest encryption (DaRE) at the cluster level; however, as the NVD doesn't have a stated requirement that warrants enabling it, this design doesn't use it on either the Files or Objects clusters. Note, however, that both Files and Objects support DaRE if it's required. You can also enable DaRE nondisruptively after cluster creation and data population. Once

you enable DaRE, existing data is encrypted in place and all new data is written in an encrypted format.

> Note: To enable DaRE, you must also deploy an encryption key management solution such as the key management functionality built into Prism Central or a third-party key management system.

Client-Server Traffic Encryption

Nutanix Objects supports HTTPS for secure, encrypted client communications according to the certificate considerations described earlier. This design uses HTTPS encryption because it doesn't incur a significant performance penalty.

Nutanix Files supports SMB3 encryption for SMB3 client-server traffic and Kerberos krb5p encryption for NFSv4 client-server traffic. This design doesn't include these encryption types because they incur a performance penalty and this NVD doesn't have a requirement that warrants enabling them.

Security Analytics

This design uses Data Lens analytics (delivered as a service) to provide additional levels of security for the Files environment. Data Lens adds security protection through the following capabilities:

- Auditing

  Searchable audit trail of all file operations performed by each user. Data Lens also tracks Files client workstation IP addresses.

- Anomalous event detection

  Tracks administrator-defined behavioral patterns that could signify potentially malicious activity. Email alerts are configurable.

- Ransomware protection

  Signature-based detection uses the Nutanix Files ransomware file blocking mechanism to identify potential ransomware attacks and block files with a ransomware extension from carrying out malicious operations. The

ransomware file blocking mechanism uses a curated list of more than 4,000 signatures that frequently appear in ransomware attacks.

Event-pattern-based ransomware protection looks for audit events in near real time to identify potential ransomware attacks. Configuring automatic remediation allows you to block users and clients suspected to be the source of a ransomware attack from accessing shares, thereby preventing the ransomware attack from further infecting the files.

You can configure email alerts on the ransomware page of the Data Lens UI to notify staff of ransomware attacks as they occur.

Reports let you know whether shares have self-service restore snapshots enabled. Using self-service restore snapshots is a best practice, and this design enables them for all shares.

Role-Based Access Control

Nutanix Objects uses the RBAC feature in Prism. Objects has two out-of-the-box management roles (Objects viewer and Objects admin), but you can also create custom roles. Use the RBAC capability if different Objects administrators have different responsibilities. You must enable IAMv2 authentication in Prism Central for Objects RBAC to be available.

> Note: Objects RBAC is specifically for defining administrator privileges and has nothing to do with S3 user permissions management.

Although the Objects RBAC feature is available, it's not part of this storage design. Because the Prism Central instances used to manage Files and Objects aren't used for managing any other Nutanix services, storage administrators are inherently limited to managing only storage.

Files currently lacks granular RBAC in Prism, so another approach is necessary to restrict storage administrators' sphere of influence in the wider environment. This design's dedicated storage clusters, with a Prism Central instance dedicated to those clusters, limits storage administrators to managing only storage resources.

RBAC for Data Lens is simple and straightforward. Assign user roles in the Admin Center at my.nutanix.com for users and admininstrators.

Data Immutability and Versioning

To protect against deletion, encryption, or other malicious modification of object data, Nutanix Objects supports write once, read many (WORM). Cohasset has validated the Nutanix WORM implementation against industry-recognized security standards. The core Hybrid Cloud NVD includes WORM as part of the backup repository to ensure immutable backups. You can also apply WORM to other object workloads if required; WORM is an optional part of this storage NVD.

Nutanix Files supports share-level WORM starting with the 4.1 release. This NVD didn't test Files WORM capabilities.

Objects also supports versioning, which provides access to previous states (versions) of an object before it was changed. Unlike with WORM, previous versions aren't immutable and can be deleted. However, because the previous versions provide restore points that can prove useful in the event of data tampering, this design uses versioning. Data Lens also requires versioning enablement on the Smart Tiering endpoint bucket.

Network Microsegmentation

The core Hybrid Cloud NVD describes how microsegmentation protects AHV VMs by controlling which network traffic can travel between VMs or groups of VMs as defined by categories. This storage design uses the same microsegmentation technology to protect the Files and Objects services. This design's policies allow only the necessary inbound communication over the required ports, as detailed in the following table. Inbound ports are restricted to the FSVMs and to the Objects load balancer VMs. Refer to the core Hybrid Cloud NVD for more details about microsegmentation. Refer to the Port Reference guide for additional port requirement details.

Table 20: Inbound Ports to Files FSVMs

Port Number	Description	Source	Destination	Transfer Protocol	Service
22	SSHD	CVM	FSVM	TCP or UDP	SSH

Port Number	Description	Source	Destination	Transfer Protocol	Service
2027	CVM to FSVM management	CVM	FSVM	TCP	Insights
2090	CVM to file server management and task status	CVM	FSVM	TCP	Ergon
2100	Cluster configuration	CVM	FSVM	TCP	Genesis
7502	Access services running on Files	CVM	FSVM	TCP	minerva_nvm
9440	REST API calls and Prism access	CVM Prism Central	FSVM	TCP or UDP	SSH Mercury
7515	Smart DR Replication	FSVM	Remote FSVM	TCP	Replicator
111	NFS port mapper	NFS client	FSVM	TCP or UDP	NFS
2049	NFSv4 support	NFS client	FSVM	TCP	NFS
7508	Statd output port for boot, reboot, and recovery functions	NFS client	FSVM	TCP	NFS (statd)
20048	Mountd access and service request monitoring port	NFS client	FSVM	TCP	NFS (mountd)
20049	Karbon	NFS client	FSVM	TCP	Karbon
20050	Lockd port for locking requests	NFS client	FSVM	TCP	NFS (lockd)
445	FSVM to SMB client communication	SMB client	FSVM	TCP	Active Directory or SMB

Table 21: Inbound Ports to Objects Load Balancers

Port Number	Description	Source	Destination	Transfer Protocol	Service
22	SSHD	Prism Central	Storage network	TCP	SSH
80	HTTP endpoint to access Objects	External S3 clients	Public network (load balancer)	TCP	HTTP
443	HTTPS endpoint to access Objects	External S3 clients	Public network (load balancer)	TCP	HTTPS
9901	Prism Central checks health of envoy	Prism Central	Storage network	TCP	Envoy
5553	Prism Central communicates with IAM service	Prism Central	Public network (load balancer)	TCP	IAM
7100	Objects replication	Storage network	Public network of the target object store	TCP	Objects replication

Table 22: Security Design Decisions Specific to Unified Storage

Decision Name	Description
Objects client connection security	Sign with an internal certificate authority
Certificates	Provision certificates with a yearly expiration date and rotate accordingly
DaRE	Disable DaRE, don't deploy a key management server

Decision Name	Description
Client-server traffic encryption	Enable HTTPS for Objects, don't enable SMB3 encryption or krb5p encryption
File share permissions	Leave share permission setting at Full Control and manage access with NTFS permissions
Security analytics	Enable Pulse to establish link to Data Lens and permit analytics for the file servers in AZ1 and AZ2. Define anomalies and enable ransomware protection with email alerting
Data immutability and versioning	Enable versioning policy on deployed buckets. Optionally enable WORM on buckets not associated with Files Smart Tiering
Microsegmentation	Use microsegmentation to restrict traffic to the Files FSVMs and Objects load balancers to the required ports

Datacenter Infrastructure

This design assumes that datacenters in the hosting region can sustain two AZs without intraregional fate-sharing—in other words, that failures in one datacenter's physical plant or supporting utilities don't affect the other datacenter.

Rack Design

A single rack might contain a Files cluster and an Objects cluster (four nodes for each cluster). You can expand both clusters in the rack to a total of eight nodes each. If you need additional storage expansion, you can go beyond the scope of this design and add more racks, depending on top-of-rack network switch density as well as the datacenter's power, weight, and cooling density capabilities per square foot. Subject to meeting these environmental requirements, existing Files and Objects clusters can expand into neighboring racks.

Refer to the Platform Selection section for the specific node models selected for this NVD. The following figure shows the initial density for this design, with the designated requirements, assumptions, and constraints.

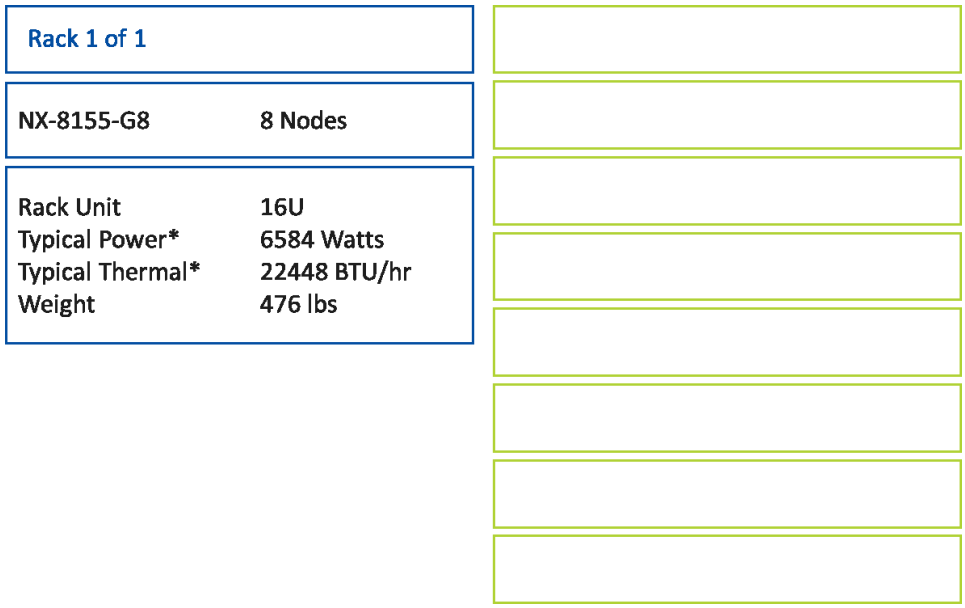

Figure 3: Rack Layout

When you scale the environment, consider physical rack space, network port availability, and the datacenter's power and cooling capacity. Consider scaling the backup clusters (described in the core Hybrid Cloud NVD) in proportion to the data growth in the Files environment.

In this design's physical rack space, one generic 42RU rack contains 16RU of systems with 3RU reserved for two data switches and one out-of-band switch, leaving 23RU of space available. Adding eight more nodes to the rack (four for Files, four for Objects) consumes an additional 16RU, leaving 7RU available.

For network ports, the eight nodes in this storage NVD consume 8 ports on each of the two data switches. With 48 port switches, two Inter-Switch Links (ISLs), and two uplinks to the upstream network, this configuration leaves 36 ports available per data switch. Adding eight more nodes to the rack (four for Files, four for Objects) consumes 8 more ports per switch, leaving 28 unused ports available per switch.

For power, cooling, and weight, you need the minimums specified in the previous figure and should assume at least double these values for a fully loaded rack including network switches. Datacenter selection is beyond the scope of this design; however, you should have a conversation about fully

loaded racks with datacenter management prior to initial deployment, as planning to properly support the environment's long-term growth may change where in the facility you want to set up the equipment.

3. Backup and Disaster Recovery

This storage NVD uses Files Smart DR and Objects streaming replication to provide a business continuity and disaster recovery (BCDR) solution to protect against different types of events. This section defines the overall high-level disaster recovery and backup designs. The core Hybrid Cloud NVD provides the details of the backup storage design.

This design provides three levels of recovery point objective (RPO) for file data protection:

- Gold Tier RPO: 1 minute
- Silver Tier RPO: 1 hour
- Bronze Tier RPO: 24 hours

The solution provides a recovery time objective (RTO) of 1 hour, which includes client-side cache operations that are outside the influence of Files Smart DR.

This design doesn't provide a specific RPO for Objects because Objects replication is streaming—data replicates as soon as the system writes the object on the source bucket. Depending on factors such as network speed and bandwidth, streaming replication can be nearly synchronous.

To protect workloads against security threats like ransomware attacks, this NVD copies data to an external backup system that provides immutability. The backup target is a Nutanix Mine cluster in each AZ, described in depth in the core Hybrid Cloud NVD.

Table 23: Unified Storage NVD BCDR Requirements

Requirement Description
Place file shares with different levels of criticality in separate replication policies.
Configure Files self-service restore snapshot schedules for share-level protection.
Provide an RPO of 1 min for critical file data.

Requirement Description
Provide an RPO of 1 hour for file data.
Provide an RPO of 24 hours for noncritical file data.
Support full failover (including networking).
Files: Support automated DNS update and Service Principal Name transfer.Objects: Manage client redirection with a third-party GSLB (F5 BIG-IP).
Provide maximum automation and orchestration for failover and failback.
Read-only access of replicated data for testing purposes
Simplify disaster recovery exercise, reducing human interaction to a minimum during disaster recovery.
Support the following disaster recovery events: — Datacenter outage. — Single cluster outage. — Ransomware attack. — Top-of-rack switch outage. — Single VLAN outage. — Human error. — Software bug. — Performance degradation caused by infrastructure (Nutanix cluster or network) or hardware components.
Ensure that tiered files remain accessible after AZ failover.
Ensure that file tiering activity resumes after AZ failback.
Choose a backup solution that supports Nutanix Files backup and restore using API, ideally CFT.
Choose a backup solution that supports Nutanix Files file-level backup and restore.
Choose a backup solution that supports S3-compatible storage as a backup target.
Choose a target backup storage system that supports ransomware protection.
Choose a target backup storage system that supports WORM.

Requirement Description
Choose a backup solution that supports replication to a secondary location.
Choose a backup solution that supports archiving to S3-compatible storage, including public cloud providers.

Note: The customer must confirm every assumption listed in the following table.

Table 24: Unified Storage NVD BCDR Assumptions

Assumption Description
Disaster recovery avoidance causes minimal storage service downtime.
Customer provides redundant WAN connectivity between AZs.
Supporting infrastructure elements like DNS, Active Directory, and IP Address Management (IPAM) are available in both AZs.
Solution doesn't provide partial failover capabilities.

Table 25: Unified Storage NVD BCDR Risks

Risk Description	Impact	Likelihood	Mitigation
Full outage of active AZ	Large	Unlikely	Fail over to remote AZ.
WAN link outage	Large	Unlikely	Provide redundant WAN connection.
Ransomware attack	Large	Likely	Implement backup solution with immutability. Replicate data to remote AZ.
Top-of-rack switch outage or misconfiguration	Large	Unlikely	Use two top-of-rack switches for redundancy.
Single Nutanix cluster outage	Medium	Unlikely	Replicate data and fail over to remote AZ.
Single VLAN outage or misconfiguration	Medium	Unlikely	Replicate data and fail over to remote AZ.

Risk Description	Impact	Likelihood	Mitigation
Human error	Large	Likely	Introduce automation. Replicate data and fail over to remote AZ.
Performance degradation caused by infrastructure or hardware components (Nutanix clusters, network)	Large	Unlikely	Replicate data and fail over to remote AZ.

Table 26: Unified Storage NVD BCDR Constraints

Constraint Description	Comment
Use Nutanix Mine for file backup	Although this NVD uses HYCU as the Mine backup partner, Nutanix Mine solutions are also available in partnership with Commvault, Arcserve, and Veeam.
Use Nutanix Smart DR for Files disaster recovery automation	Smart DR is the solution of choice to automate file share failover. Source and destination clusters must be under the same Prism Central instance.
Use GSLB for automated Objects client redirection during disaster recovery events	The GSLB integrates with DNS to direct clients to the live object store. Local and remote FQDNs applied to object stores in each AZ ensure that the object store in the local AZ can respond to S3 requests targeted at the remote AZ's object store. Apply the appropriate certificates on each Objects cluster.

Constraint Description	Comment
Use identical names for source and target buckets	Bucket names must be identical to preserve the fileserver-to-bucket tiering relationship in a disaster recovery failover event. This naming consistency, combined with the configuration of the source object store's FQDN as a secondary FQDN on the disaster recovery object store, ensures that the file server's tiering profile can locate the disaster recovery bucket.

Table 27: Unified Storage NVD BCDR Design Decisions

Decision Name	Decision
Automate disaster recovery failover and testing	Use Smart DR workflows for FilesUse a GSLB for Objects and apply a local and a remote FQDN to each object store
Provide solution to support RPOs of 1 min, 1 hour, and 24 hours	Control Files Smart DR with replication policies and Objects streaming replication with replication rules
Determine the maximum number of shares for Smart DR (Files) and bucketsfor streaming replication (Objects)	Maximum of 100 shares (20 distributed, 80 standard) Up to 10 min RPO (with 5 distributed and 20 standard shares) No limits for Objects replication
Determine the Nutanix local and remote snapshot retention policies	Keep 12 hourly snapshots and 7 daily snapshots for file shares at both local and remote AZs
Determine the Nutanix local and remote versioning policies	Enable versioning for Objects buckets at both local and remote AZs

Decision Name	Decision
Keep the relationship between Files Smart Tiering and Objects intact in a disaster recovery failover event	Use two file server instances per AZ: one passive (receiving data replicated from the alternate AZ) and one active. Only active file servers can have an associated tiering profile. After both the file share and object bucket layers fail over, the file share's reference to the bucket remains valid because: — The Objects instances have secondary FQDNs — The GSLB seamlessly directs files to the surviving object store
Back up workloads within AZs or across AZs	To optimize the backup window and save WAN bandwidth, Mine clusters back up file shares that are in the local AZ
Determine the RPO to set on backup policies	Set 24-hour RPO on backup policies
Determine how many backup policies to configure per HYCU BC	Single HYCU policy
Determine which storage solution to use as a backup repository	Use Nutanix Mine version 3 as the backup repository
Determine how many S3 buckets to use as the backup repository	Use one object store with one bucket as the backup repository
Determine which advanced features to enable on S3 storage	Enable WORM and set it for 365 days

Backup and Disaster Recovery Conceptual Design

Nutanix Prism Central is the management and control plane for Files and Objects disaster recovery capabilities. Files takes advantage of the Smart DR feature that uses replication jobs to define replication frequencies on a share-by-share basis. Nutanix Objects lets you configure streaming replication

between a source and a target Objects instance on a bucket-by-bucket basis. HYCU backup policies configured for each share use the CFT API for incremental backups of the Files environment.

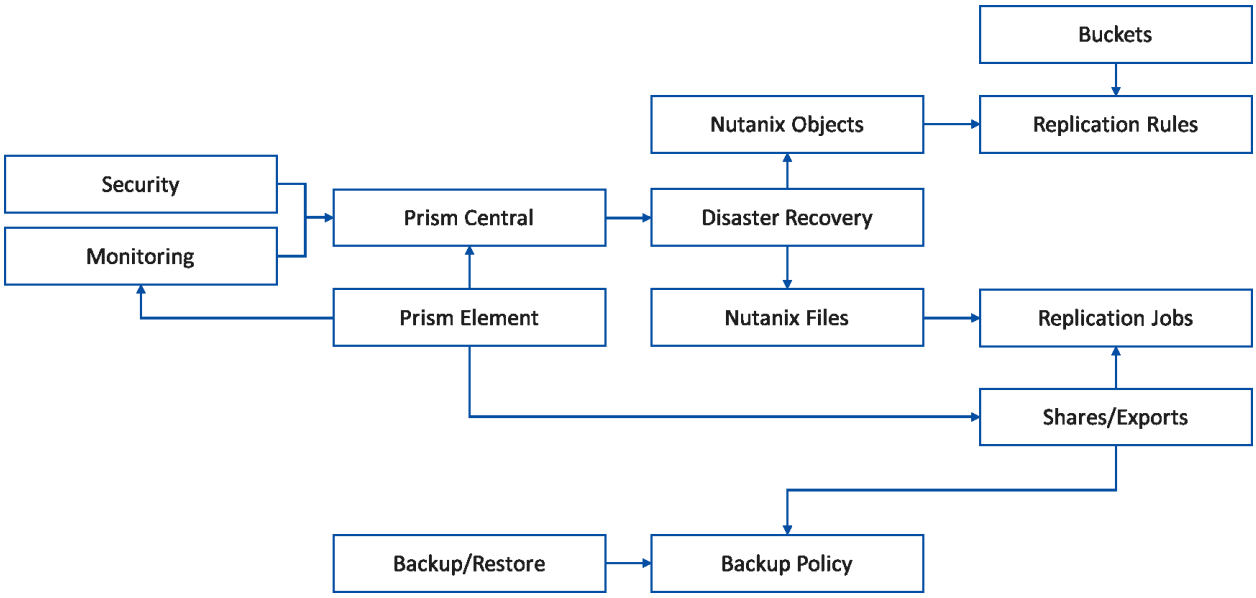

Figure 4: Unified Storage NVD BCDR Conceptual Design

Disaster Recovery

Disaster Recovery Logical Design

This storage NVD provides comprehensive disaster recovery protection for file shares and buckets across both AZs in a single region. Applications can take advantage of underlying infrastructure to provide disaster recovery resilience based on three different protection levels with bidirectional replication between AZs. The design provides disaster recovery to the file servers and objects instances.

Disaster recovery testing, failover, and failback are fully orchestrated and require only minimal human involvement. Files clusters rely on a single Prism Central instance to register both the source and target file servers for replication. Each Objects cluster is registered to the Prism Central instance in its respective AZ.

There are four file servers in the design, two in each AZ. In each AZ one file server is passive (no connected clients, only receiving data replicated from the alternate site), while the other actively serves clients in the AZ. This configuration preserves the Files-Objects share-to-bucket Smart Tiering relationship in the event of a failover between AZs. Because each file server can have only one tiering profile, only the active file servers have Smart Tiering configured. If both the Files and Objects services fail over, the file share's reference to its bucket remains valid because of two key aspects of the design:

- The Objects instances have remote FQDNs that allow them to respond to S3 requests from the file server intended for the object store in the other AZ.
- The GSLB seamlessly directs reads of tiered data initiated by the file server to the disaster recovery object store.

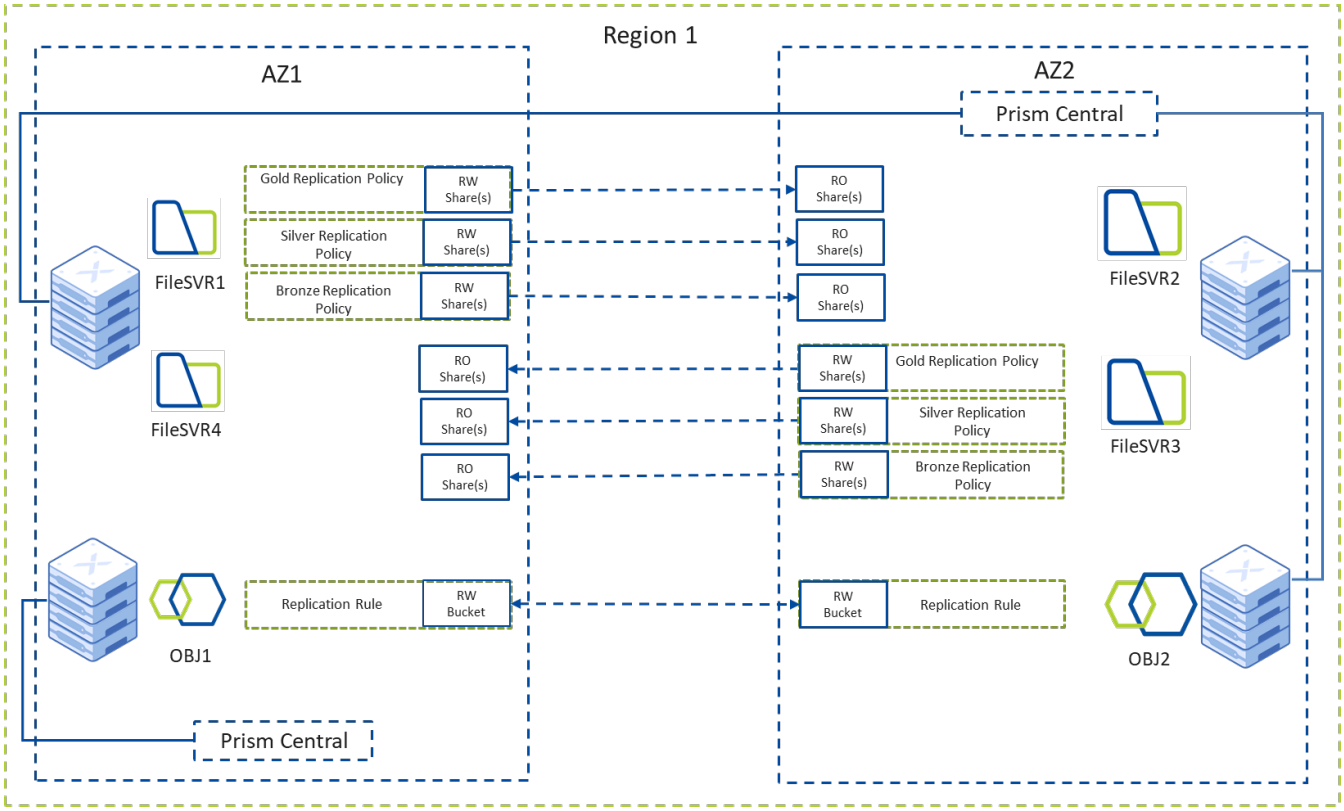

Figure 5: Unifed Storage NVD BCDR Logical Diagram

Disaster Recovery Detailed Design

The NVD provides three protection tiers for Files. Objects uses continuous replication, where RPO depends on bandwidth and latency between AZs.

Table 28: Files Disaster Recovery Protection Tiers

Tier	RPO	RTO
Gold	1 minute	1 hour
Silver	1 hour	1 hour
Bronze	1 day	1 hour

Table 29: Objects Disaster Recovery Protection Tiers

Tier	RPO	RTO
Continuous	Real Time (variable)	1 hour

The BCDR section of this storage NVD uses the following software versions.

Table 30: Software Versions for Disaster Recovery

Component	Software Version
Prism Central	pc.2022.1
AOS	5.20.3 (LTS)
Objects	3.4.0.2
Files	4.1
F5 BIG-IP	16.1.2-0.0.18

Table 31: Files Replication Policy Configuration

Policy Name	Source Cluster	Target Cluster	RPO
AZ01-AZ02-Bronze-01	AZ01-FS-01	AZ02-FS-02	1 day
AZ01-AZ02-Silver-01	AZ01-FS-01	AZ02-FS-02	1 hour
AZ01-AZ02-Gold-01	AZ01-FS-01	AZ02-FS-02	1 minute
AZ02-AZ01-Bronze-01	AZ02-FS-03	AZ01-FS-04	1 day
AZ02-AZ01-Silver-01	AZ02-FS-03	AZ01-FS-04	1 hour
AZ02-AZ01-Gold-01	AZ02-FS-03	AZ01-FS-04	1 minute

Table 32: Objects Replication Policy Configuration

Policy Name	Source Cluster	Target Cluster	RPO
Applied on source bucket	AZ01-OBJ-01	AZ02-OBJ-02	Continuous replication
Applied on source bucket	AZ02-OBJ-02	AZ01-OBJ-01	Continuous replication

You can run Files failover operations at the file server level from Prism Central. If Prism Central isn't available, you can run the same workflow using the CLI on the target file server. All shares in replication policies fail over together when you recover a file server on a target cluster.

A GSLB controls Objects failover. This storage NVD tested F5 BIG-IP as the GSLB. The F5 GSLB acts as a DNS server and forwards requests to the appropriate Objects instance based on availability. With both Files and Objects, clients automatically redirect to the target site because the namespaces move as a part of the disaster recovery orchestration. This approach also ensures that Smart Tiering relationships between the Files and Objects layers remain intact.

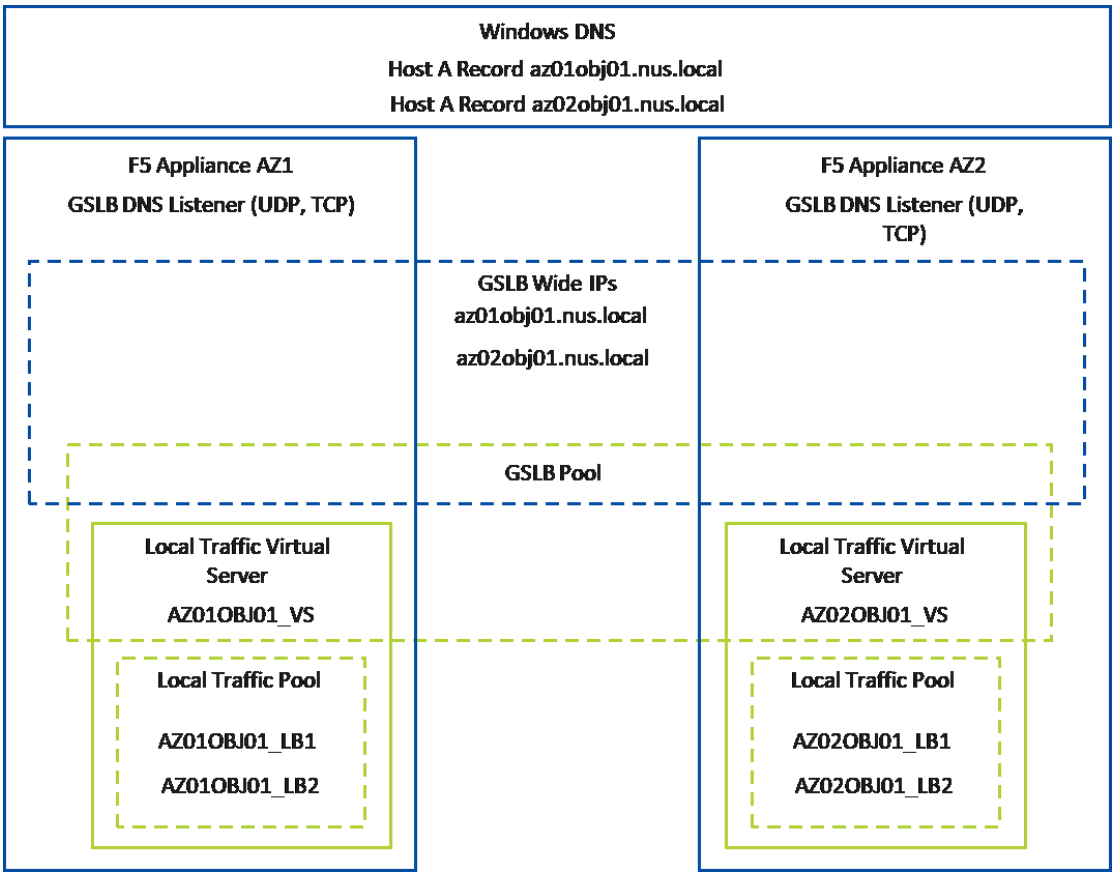

Figure 6: F5 BIG-IP Configuration for Objects

Backup

Backup Logical Design

The core Hybrid Cloud NVD provides a backup option for workloads running in both AZs. This design, which extends to file data, optimizes the backup solution to back up workloads that run locally to the backup cluster.

In Files you apply a backup policy to source shares and exports. The FSVMs that make up the Files instance are stateless and not backed up. HYCU integrates with the Files CFT API to ensure efficient incremental backup of shares as data changes. HYCU also supports backing up shares that are either the source or the target for Smart DR replication. This storage NVD uses

Smart DR to replicate Files data between AZ1 and AZ2 instead of HYCU-based replication. HYCU policies then back up the local copy of each file share, whether it's a source or target for Smart DR replication.

Objects backup is out of scope for this NVD, which protects Objects data using remote replication and disaster recovery capabilities.

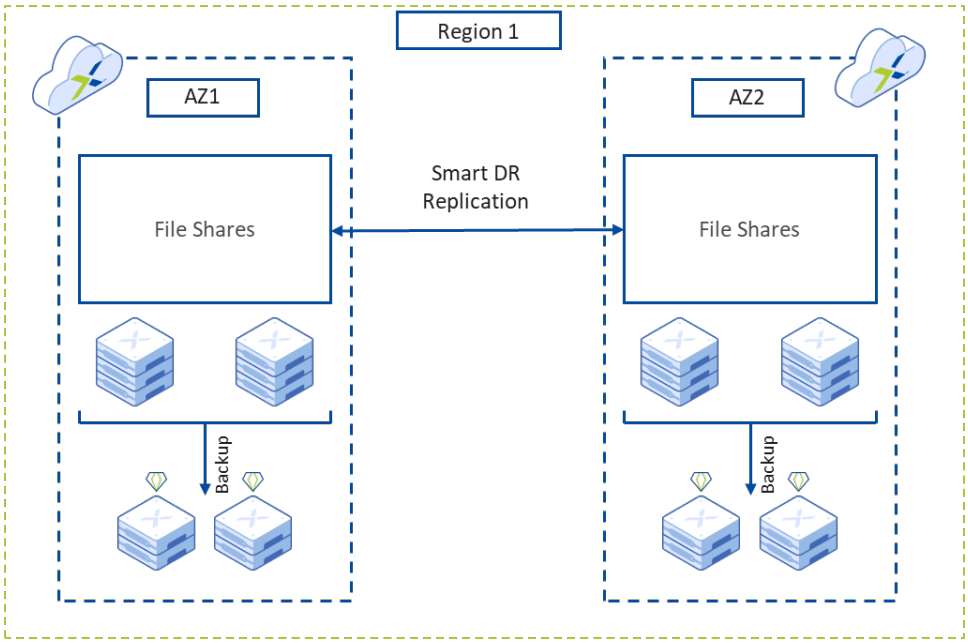

Figure 7: Nutanix Unified Storage NVD Backup Architecture Logical Design

4. Test Plan

An NVD provides a framework of components and certifies the operational functionality of the integrated modular design. This detailed test plan describes what Nutanix tested, along with the expected and actual results, to certify that Nutanix has tested the solution with the configuration specified and confirmed that it functions as designed.

Unified Storage Detailed Test Plan

Table 33: Core Infrastructure Tests

Summary	Validation Tasks	Expected Result
Verify app integration.	Launch Objects from Prism Central. Launch Files Manager from Prism Central.	Interfaces open without error.

Summary	Validation Tasks	Expected Result
Test platform availability, resilience, and integrity with X-Ray.	Deploy X-Ray on a different system than the infrastructure you're testing. Verify that the cluster you're testing isn't running any other workloads or VMs. Select the platform availability, resilience, and integrity scenario. Set test parameters: — 10 percent usable space consumption — 30 second I/O timeout — Full platform resilience test — Workload type: Workload.fio — Select the target cluster(s)	Standard deviations for all tests are 0.
Verify storage container creation.	Verify that two containers have been created for the object store. Verify that one container has been created for the file server.	Storage container created with inline compression enabled by default.
Verify file share access.	Create, read, and delete files in the SMB and NFS shares.	The client operations are successful.
Verify bucket access.	Upload and download files in the object store.	The client operations are successful.

Summary	Validation Tasks	Expected Result
Run X-Ray NFS Fileserver Peak Performance Microbenchmark.	Verify that the cluster you're testing isn't running any other workloads or VMs. Select the NFS Fileserver Peak Performance Microbenchmark.	Client operations (random and sequential read and write workloads) are successful.
Run X-Ray S3 Peak Performance Microbenchmark.	Verify that the cluster you're testing isn't running any other workloads or VMs. Select the S3 Object Storage Microbenchmark.	Client operations (random and sequential read and write workloads) are successful.
Tier file share data.	Create a tiering policy in Data Lens with a capacity limit and data age that allows tiering to the configured Objects instances.	File stubs and objects are created. Tiered files can be read.
Verify email receipt in primary mailbox.	Wait up to 24 hours for the first emails from Prism Central to arrive in the primary mailbox.	The primary mailbox receives alerts from Prism Central.
Verify email receipt in secondary mailbox.	Wait up to 24 hours for the first emails from Prism Element to arrive in the secondary mailbox.	The secondary mailbox receives alerts from Nutanix clusters (Prism Element).

Table 34: Security Tests

Summary	Validation Tasks	Expected Result
Verify Data Lens anomaly alerts.	Configure Data Lens anomaly alerts for 100 read and write events over 1 hour. Create more than 100 read or write events in less than 1 hour.	Anomaly events are logged. Data Lens sends email to configured primary and secondary mailboxes.
Verify Data Lens audit trails.	After running core infrastructure tests, query audit trails to confirm activity logging.	File, Folder, User, and Client audit trails are populated.
Create ransomware events.	Ensure that ransomware protection is enabled from Data Lens. Attempt to create files with known ransomware file extensions.	File creation is blocked, ransomware events are logged in Data Lens, and primary and secondary mailboxes receive email alerts. Optionally, confirm that user and client access is blocked (if you configured user and client blocking).
Verify Flow security policies.	Attempt to access a file share from a subnet not allowed in the configured security policies.	File share access is blocked.

Backup and Disaster Recovery Detailed Test Plan

Backup Integration

Summary	Validation Tasks	Expected Result
Back up and restore file shares.	Ensure that CFT integrated incremental backup functions.	Incremental share backup.
	Verify the interaction between tiering and backup.	Tiered files properly are backed up and not read inline once tiered.

Summary	Validation Tasks	Expected Result
	Restore individual files from backup.	Files are properly restored to alternate locations (or in-place with tiering).
	Restore file shares from backup.	The full content of a share is recovered.
	Restore individual files from self-service restore snapshots.	Recover files using self-service restore.

Planned Failover from AZ01 to AZ02

Perform planned failover for shares and buckets from all protection tiers.

Table 35: BCDR Test

Summary	Validation Tasks	Expected Result
Perform planned failover of Files and Objects instances.	Verify that files and objects are recovered at the remote AZ.	Files and objects are recovered at remote AZ.
	Validate DNS update.	Files and Objects clients received correct IP addresses.
	Verify that you can read and write files and objects over different protocols.	Successful read and write access over SMB, NFS, and S3 protocols.
	Verify that Data Lens and tiered data are accessible.	Tiered data is accessible.
	Verify that reverse replication occurs for Files and Objects.	New data is replicated from the remote AZ.
	Measure RTO.	RTO: 1 hour maximum.

Planned Failback from AZ02 to AZ01

Perform planned failback for shares and buckets from AZ02 to AZ01.

Table 36: BCDR Test: Planned Failback

Summary	Validation Tasks	Expected Result
Perform planned failback of Files and Objects instances.	Verify that files and objects are recovered at the local AZ.	Files and objects are recovered at local AZ.
	Validate DNS update.	Files and Objects clients received correct IP addresses.
	Verify that you can read and write files and objects over different protocols.	Successful read and write access over SMB, NFS, and S3 protocols.
	Verify that Data Lens and tiered data are accessible.	Tiered data is accessible.
	Verify that reverse replication occurs for Files and Objects.	New data is replicated from the local AZ.
	Measure RTO.	RTO: 1 hour maximum.

Unplanned Failover from AZ01 to AZ02

Table 37: BCDR Test: Unplanned Failover

Summary	Validation Tasks	Expected Result
Perform unplanned failover of Files and Objects instances.	Verify that files and objects are recovered at the remote AZ.	Files and objects are recovered at remote AZ.
	Validate DNS update.	Files and Objects clients received correct IP addresses.
	Verify that you can read and write files and objects over different protocols.	Successful read and write access over SMB, NFS, and S3 protocols.
	Verify that Data Lens and tiered data are accessible.	Tiered data is accessible.
	Verify that reverse replication occurs for Files and Objects on AZ01 recovery.	New data is replicated from the remote AZ.
	Measure RTO.	RTO: 1 hour maximum.

5. Ordering

This bill of materials (BoM) reflects the validated and tested hardware, software, and services that Nutanix recommends to achieve the outcomes described here. Consider the following points when you build your orders:

- All software licensing is based on storage capacity.
- Nutanix Xpert Services or an affiliated partner selected by Nutanix provides all services.
- Nutanix based the functional testing described in this document on NX series models with similar configurations to validate the interoperability of software and services.

Substitutions

- Nutanix recommends that you purchase the exact hardware configuration reflected in the BoM whenever possible. If a specific hardware configuration is unavailable, choose a similar option that meets or exceeds the recommended specification.
- You can make hardware substitutions to suit your preferences; however, such changes may result in a solution that doesn't follow the recommended Nutanix configuration.
- Avoid software product code substitutions except when:
  - You need different quantities to maintain software licensing compliance.
  - You prefer a higher license tier or support level for the same software product code.
- Adding any software or workloads that aren't specified in this design to the environment (including additional Nutanix products) may affect the validated density calculations and result in a solution that doesn't follow the recommended Nutanix configuration.

- Professional Services substitutions to accommodate customer preferences aren't possible.

Sizing Considerations

This NVD is based on one 4-node Files cluster and one 4-node Objects cluster in each AZ for BCDR. You can use backup clusters in the core NVD file backup.

A 4-node cluster is the minimum size, but you can increase the Files and Objects clusters incrementally up to 8 nodes each while remaining in a single rack. If you need even more storage capacity, you can expand the clusters into adjacent racks. Files can use up to 32 nodes of storage capacity and up to 16 nodes of compute resources for client I/O processing. Objects can use up to 48 nodes in a single cluster (both compute and storage) with the option to incorporate storage from up to four other physical clusters into the namespace.

Bill of Materials

The following tables show the BoMs for the primary and secondary Files clusters and the primary and secondary Objects clusters.

Table 38: Primary and Secondary Files Clusters: Hardware, Software, and Services

Product Code	Description	Quantity
	Hardware	
NX-NX8155-G8	NX-8155-G8, 1-node configuration Type: Hybrid Hardware support: — Support level: Production — NRDK support: No — NR node support: No	8
	Per-Node Hardware Configuration	

Product Code	Description	Quantity
	Processor: Intel Xeon-Silver 4310 processor (2.1 GHz / 12-core / 120 W, Ice Lake)	2
	Memory: 32 GB memory module (3,200 MHz DDR4 RDIMM)	8
	HDD: 18 TB	8
	SSD: 7.68 TB	4
	Network adapter: 25 GbE, 2-port (Mellanox ConnectX-5)	1
	Software	
SW-NUS-PRO	Subscription, contains Files Support level: Production	200 TiB
SWA-NUS-SEC	Security add-on	200 TiB
SW-NCM-STR	NCM Starter software license subscription for 1 CPU core Type: Core-based licensing License type: NCM Starter	192
HS-NDL-PR	Data Lens cloud-based file analytics service	200 TiB

Table 39: Primary and Secondary Objects Clusters: Hardware, Software, and Services

Product Code	Description	Quantity
	Hardware	

Product Code	Description	Quantity
NX-NX8155-G8	NX-8155-G8, 1-node configuration Type: Hybrid Hardware support: — Support level: Production — NRDK support: No — NR node support: No	8
	Per-Node Hardware Configuration	
	Processor: Intel Xeon-Silver 4310 processor (2.1 GHz / 12-core / 120 W, Ice Lake)	2
	Memory: 32 GB Memory Module (3,200 MHz DDR4 RDIMM)	4
	HDD: 18 TB	10
	SSD: 3.84 TB	2
	Network adapter: 25 GbE, 2-port (Mellanox ConnectX-5)	1
	Software	
SW-NUS-PRO	Subscription, contains Objects Support level: Production	200 TiB
SWA-NUS-SEC	Security add-on	200 TiB
SWA-NUS-ADR	Streaming replication	200 TiB
SW-NCM-STR	NCM Starter software license subscription for 1 CPU core Type: Core-based licensing License type: NCM Starter	192
HS-NDL-PR	Data Lens cloud-based file analytics service	200 TiB

Professional Services

The following professional services allow Nutanix to implement this storage NVD as designed, built, and tested. These services are outcome-based, with fixed prices for the scope described by the SKUs included in the BoM. See the Xpert Services information available on Nutanix.com for more details on each of the SKUs included.

Table 40: Professional Services for Platform

Product Code	Description	Quantity
CNS-INF-A-SVC-DEP-ULT	Deploy physical clusters in an AZ and configure appropriate storage service AZ01	8
CNS-INF-A-SVC-DEP-ULT	Deploy physical clusters in an AZ and configure appropriate storage service AZ02	8
CNS-INF-A-SVC-MCR-STD	Deploy Flow network security policies (covers up to 10 policies).	1
CNS-INF-A-WRK-STG	Files consolidation workshop: plan file migration to Nutanix Files.	1
CNS-INF-A-SVC-MIG-FIL	Files migration: packs of 100 TB of file data.	2
CNS-INF-A-SVC-DRD2	Set up and configure disaster recovery for Files and Objects. Note: F5 GSLB deployment isn't included. A customer must deploy and configure the F5 GSLB.	2
FLEX-CST-CR	Extend and integrate the HYCU backup solution to the Files environment (including deployment, testing, and validation).	5 (days)

For more information on these Nutanix Xpert Services, review the following datasheets:

1. HCI Cluster Deployment

2. Flow Network Security Deployment
3. Storage Consolidation Workshop
4. Files Migration Service
5. HCI Disaster Recovery Deployment

Appendix

References

1. Nutanix Hybrid Cloud Reference Architecture
2. Nutanix Hybrid Cloud Validated Design
3. Nutanix Files
4. Nutanix Objects
5. Physical Networking

About Nutanix

Nutanix is a global leader in cloud software and a pioneer in hyperconverged infrastructure solutions, making clouds invisible and freeing customers to focus on their business outcomes. Organizations around the world use Nutanix software to leverage a single platform to manage any app at any location for their hybrid multicloud environments. Learn more at www.nutanix.com or follow us on social media @nutanix.

List of Figures

List of Tables