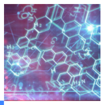
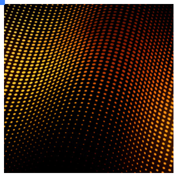
NetApp

eBook

Cloud Data Storage: The Promise and The Challenges





Introduction

With the explosion of applications and devices that are generating huge volumes of data, it is not surprising that data storage costs play a significant role in corporate IT budgets. The Osterman Research white paper Data Center Transformation Requires Software-Based Cloud Storage, sponsored by HubStor and Microsoft from December 2018, found that two out of three organizations spend in excess of 10% of their yearly IT budget on storage, while one in six spend more than 25%.

Currently 54% of the organizations surveyed for the Osterman Research white paper have more than 500 terabytes of data in storage, with the top 13% managing more than 10 petabytes. The anticipated growth in data will only continue to drive up data storage requirements and costs. By 2025, IDC predicts that there will be 175 zettabytes (175 ZB) of data in use worldwide, up from 32 ZB at the end of 2018—at a truly remarkable Compound Annual Growth Rate (CAGR) of 61%.

In order to meet their rising data storage needs, many enterprises saw the benefits of the cloud's usage-based price model as well as the other cloud storage benefits such as agility, scalability, global geographic distribution, and enhanced security. The IDC 2025 forecast noted above expects that half of the 175 ZB of data in use will be stored on the public cloud.

This guide explores the challenges that might be encountered when shifting to cloud storage, and how to address those challenges.

Table of Contents

Cloud Migration

The first challenge is shifting large data stores from on-premises storage appliances to cloud storage resources. Your data and your applications are closely intertwined, and you must consider how to move your large quantities of data to the cloud without disrupting the business. These more strategic onboarding challenges are discussed in the "Strategic Organizational Challenges To Be Considered" section below.

The tactical challenges of transferring large volumes of data include:

Transfer Speeds

Slow transfer speeds, determined mainly by the local uplink bit rate.

Migration Costs

Data transfer costs between services and regions during data distribution within your cloud environment.

Multiple Deployment Stages/Architectures

The full cloud migration may be so lengthy that it will require adopting a hybrid environment—either in the interim, or on an ongoing basis—as part of your architecture strategy.

Onboarding

Lengthy onboarding might interrupt your workloads and disrupt business continuity.

Application Refactoring

Time- and cost-consuming application refactoring or rearchitecting to make sure existing applications can function properly in the new cloud infrastructure.

Availability

Now that your data is residing on the cloud, your next challenge is to ensure that you can continue to meet the on-premises availability and business continuity SLAs. The public cloud providers commit to close to 100% availability SLAs, such as 99.91%. That does not mean, however, that their actual downtime is only 7.74 hours per year. What it means is that you will be compensated for downtime in excess of that figure.

> In fact, the three major cloud providers self-reported between 338 and close to 2,000 hours of downtime throughout 2018 and the first half of 2019.

Sometimes these outages also resulted in data loss, such as the AWS total hardware failure on May 31, 2018 in the North Virginia region that not only had customers offline for 30 minutes, but not all data could be restored afterwards.

> Critical enterprise workloads require strict SLAs of minimal downtime (RTO<60 seconds) and no data loss (RPO=0).

Any loss of data or downtime may be costly—in terms of lost revenue, reputation, customer churn, legal exposure, and more.

Thus, your critical workloads must meet the following challenges:

High availability (HA)

Support redundant storage where every “write” operation is confirmed on both instances, in parallel.

HA performance

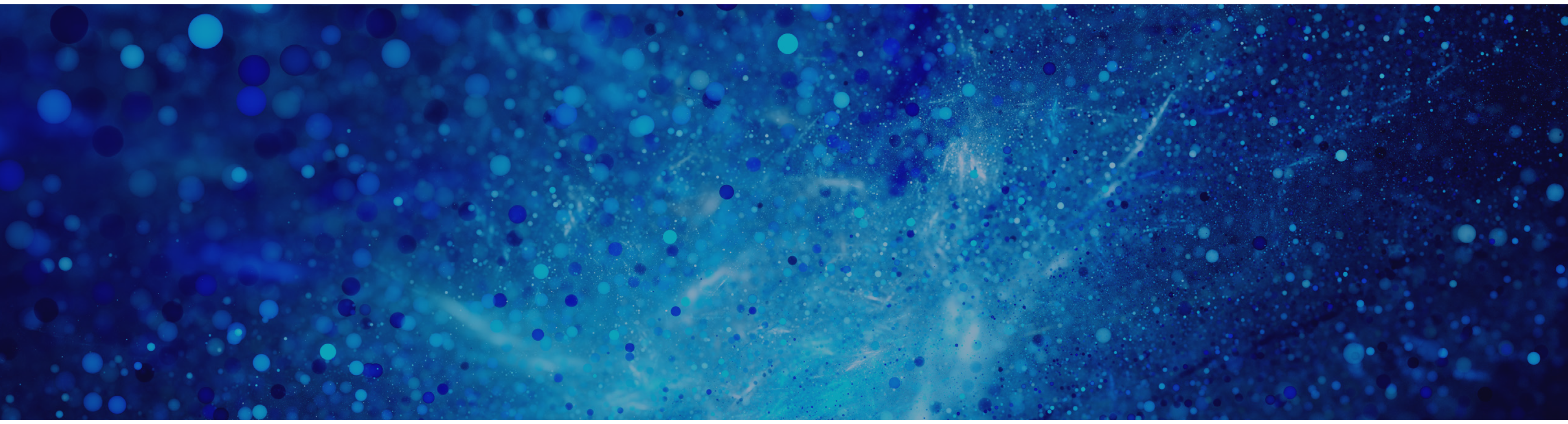
The redundancy of operations inherent to HA configurations can undermine or even disrupt performance through its added loads.

Multiple zone redundancy

Logical redundancy can be achieved using AWS placement groups, or across Azure resource groups, availability sets, fault domains and update domains. But the most resilient redundancy is achieved across multiple zones.

Seamless failover and failback

Cloud availability requires a failover process that automatically identifies a failure and instantaneously activates the relevant redundant resources or instances. Just as important is a failback process that seamlessly restores operations without loss of data or transactions.



Performance

Enterprise workloads often require high, consistent performance and low levels of latency. Data, no matter where it is stored, should be usable upon request, throughout the organization. It is important to be able to scale out or up on request, and move data between tiers without causing performance issues or disruptions.

There is a common misperception that cloud storage performance is inferior to that of on-premises.
Yet with the right services and optimal configuration, cloud storage can provide performance levels on par with—and in many cases exceeding—on-premises storage.
Storage architecture can also have a significant impact on performance. A hybrid cloud storage solution, for example, is often the most flexible for achieving optimal performance.

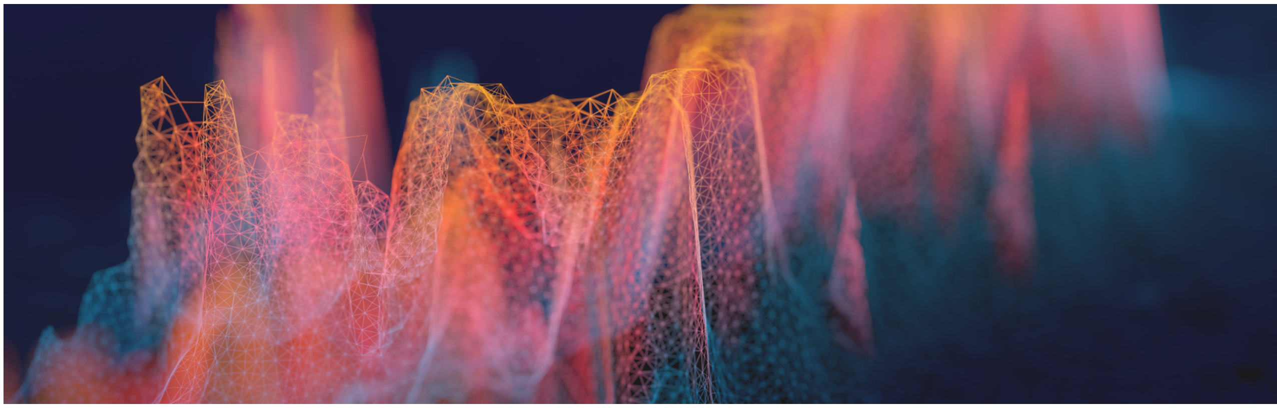
> Cloud storage can provide performance levels on par with—and in many cases exceeding—on-premises storage.

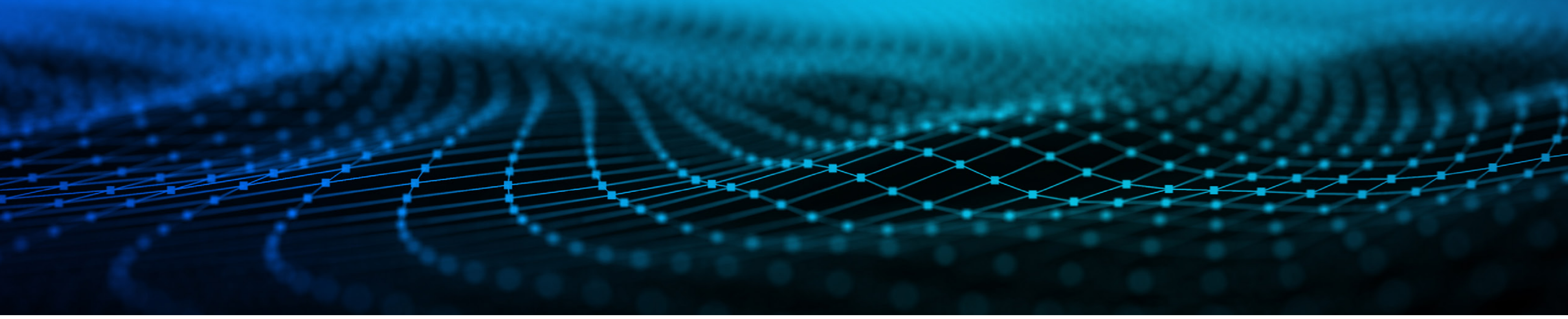
The following cloud performance issues need to be considered on a workload by workload basis:

Disparate Locations
Match server/host configurations and locations with your workloads' throughput, IOPS, and latency requirements.

Storage Type
Match the storage format (object, block, file) to the workload requirements, bearing in mind the different performance/cost trade-offs of each format.

Sizing
With varying capacity, throughput and IOPS attributes, cloud storage resources are optimized for different performance use cases. You need to find the optimal balance between cost and performance.





Data Protection

In the shared responsibility model, the cloud provider secures its own infrastructure, platform and software resources and services, but the customer is responsible for securing what runs **on** the cloud. Thus, the cloud provider is responsible for data durability, i.e., preventing the loss or corruption of data in the case of hardware failure, but the customer is responsible for all other data protection measures, such as snapshots, i.e., point-in-time copies for recovery in case of accidental deletion or data corruption, and disaster recovery (DR) systems.

Unfortunately, DR scenarios are becoming more commonplace. The Uptime Institute reports that climate change is accelerating the rate of extreme natural disasters such as storms, floods, droughts, lightning and wildfire—all of which pose threats to data centers in their way. For example, it took an entire weekend to recover Microsoft and AWS' European cloud services after their data centers were hit by lighting in Dublin in 2011. Similarly, at least eight New York data centers were disrupted when Hurricane Sandy struck in October 2012.

Disaster recovery plans must ensure that services are restored as quickly as possible, with minimal data loss. The cloud providers offer snapshots and DR tools as well as fully managed DRaaS platforms. However you cannot hand over data protection responsibilities to a third party.

Some of the key data protection challenges that you need to consider in the cloud are:

Scheduled synchronization

Define snapshot policies, automate scheduling, and deletion. Similarly, data replicas in standby DR sites must be automatically synchronized frequently enough to uphold the data loss objectives.

Failover/Failback processes

Seamless failover and failback procedures between primary and secondary DR sites that meet recovery time objectives.

Testing requirements

Regular testing to ensure that automated snapshots and DR systems will work as expected when needed.

Controlling costs

Snapshots can easily rack up significant storage costs if not managed carefully. Similarly, DR compute and storage resources effectively double the storage costs and should therefore be allocated in a cost-effective way, considering that the system is idle most of the time.

Performance disruptions

Data snapshots and DR copy synchronizations processes need to take place seamlessly, without affecting the production environment's performance, no matter how large the data sync size is.

Consistent snapshots

Snapshots technology should be application-aware and created when the application is in a quiescent state, in order to ensure data consistency.

Backups and Archiving

In many industries strict data backup and recovery requirements are mandated to protect against data loss or corruption. Similarly, many organizations are required to retain data for long periods of time in accessible archives. It is not surprising, therefore, to learn that more than one-third of enterprises spend 10+% of their IT budgets on backups (Osterman Research).

The cloud, with its storage tiers that offer trade-offs between performance and cost, is a good environment for hosting backup/recovery and archiving systems. Unitrends' 2019 survey on cloud storage adoption found that backing up from disk to cloud rose at a 15% Compound Annual Growth Rate (CAGR) between 2016 and 2019. During the same period the use of the cloud for long-term data retention, i.e., archiving, rose at an even steeper CAGR of 30%.

However, there are some challenges to keep in mind:

Aggregating Multiple Data Sources
Archival, cold data can be retained on more cost-effective, low-performance cloud storage tiers. The challenge is to seamlessly and continuously identify and aggregate data to be archived from all the relevant storage systems.

Backup and Restore Speed
Backup and restoration over the conventional NDMP (Network Data Management Protocol) can be slow. Find a backup and restore system that can quickly and easily restore a whole volume or a single file, as needed.

Visibility
In order to avoid excessive backup administration time, a cloud backup system should make it as easy to view and manage backup versions as the on-premises options.

Compliance and Governance

More than half of the enterprises surveyed in the Osterman Research white paper on data center transformation cited being compliant with privacy regulations as their top strategic concern for cloud storage. Laws and regulations being put in place by governments and industry to protect the ever-growing volumes of personal data gathered and stored have put the spotlight on effective data governance. Three prime examples are: HIPAA, for protecting personal health data in the US; GDPR, for protecting the personal data of European citizens anywhere in the world; and PCI-DSS, for protecting sensitive cardholder data. Most recently, the California Consumer Privacy Act (CCPA) came into effect on January 1, 2020, and more local regulations are expected to follow.

Governance of an organization's IT resources and activities to ensure that they align with corporate strategies and regulatory requirements is not new. However accountability frameworks and programs built for on-premises IT cannot be applied "as is" to cloud environments.

Some of the challenges inherent to compliance and governance in the cloud are:

Blurred lines of responsibility

The cloud providers are compliant with all of the major government and industry regulations, but it is important to understand where the lines of compliance responsibility are drawn between providers and customers in all of the cloud service models— IaaS, PaaS, SaaS.

Highly distributed environment

By definition cloud resources are highly distributed across geographies and zones, making it hard to understand exactly where an organization's data is residing.

Oversight

Frictionless cloud provisioning can create shadow IT that flies under the organization's governance and compliance radar.

Orchestration

Across complex environments it can be difficult to get full visibility into and control over data, such as in large hybrid and multi-cloud environments.

Security

One of the biggest benefits of cloud storage is that data becomes more easily accessible across teams, business units, and remote offices. But this benefit brings with it a major headache for security teams: how to establish effective and consistent security controls throughout often complex cloud environments.

According to Statista, three out of every four organizations around the globe consider security to be a challenge when using cloud computing services. Furthermore, Gartner predicts that by 2022 at least 95% of cloud security failures will be the customer's fault.

There are many attack vectors that threaten cloud assets. Ransomware attacks in 2019 as an example in the US alone were mounted on close to 1,000 government, educational and healthcare targets. The most damaging ransomware attack in 2019, however, was against Demant, a Danish hearing aid manufacturer that had to shut down for several weeks, resulting in a financial impact estimated at $80 to 95 million.

Establishing a robust cloud security posture is critical, and the following challenges need to be addressed:

Role-based access restrictions

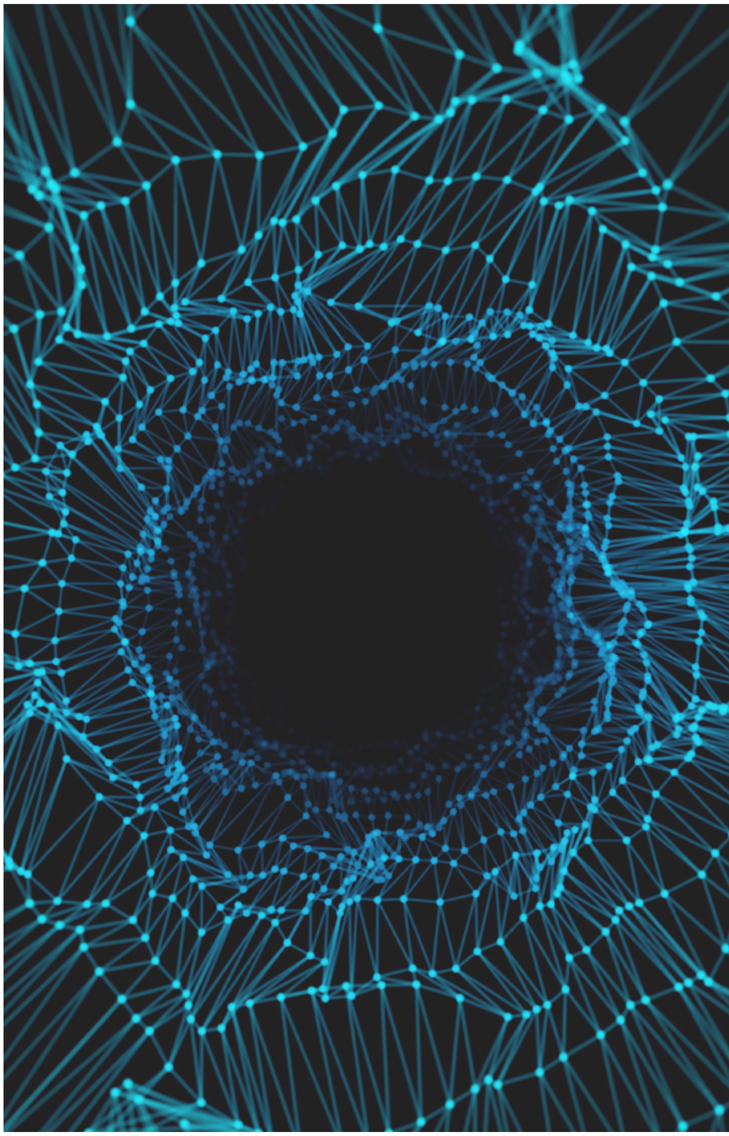
According to (ISC)², a leading international association for information security leaders, in 2019 unauthorized access through misuse of employee credentials and improper access controls were the biggest threats to cloud security. The organization must ensure that its IAM controls are robust and consistent across its entire data storage infrastructure.

Lack of visibility

One out of three organizations considers the lack of visibility into cloud infrastructure a significant security challenge at the operational level—including the difficulty of verifying the efficacy of security controls in the cloud.

Encryption

Both data-at-rest and data-in-transit need to be encrypted, with careful management of encryption keys.



Storage Footprint and Costs

Two out of three enterprises spend 10+% of their annual IT budget on storage. Of these, one in six spend 25+% on storage (Osterman Research). Containing storage costs is often a key motivator for enterprises to migrate some or all of their data storage to the cloud. However, without proper management, storage in the cloud can also suffer from sprawl and unexpectedly high monthly charges.

> Containing stoxrage costs is often a key motivator for enterprises to migrate

In order to reduce cloud storage costs, enterprises need to find solid solutions for the following challenges:

Storage consumption

Keep data footprint and data transfer bandwidth minimal through storage efficiency features such as data compression and deduplication. For example, in some sectors, such as Media & Entertainment or Advertisement, it is not unusual for the same image or frame to be saved many times in a different context or with slight variations. In these situations, deduplication can reduce storage footprint and costs by as much as 90%.

Over-Provisioning

Avoid the costs of over-provisioning storage resources. Agile and automated thin provisioning of storage resources is one way to make sure provisioned storage is always used.

Efficient snapshots

Point-in-time full snapshot copies are both slow and costly. Adopt incremental snapshots technology, which makes a full baseline copy once, and then only register the changes.

Storage tiering

Identify various storage performance requirements per use cases and usage patterns, and provision the most cost-efficient storage tier accordingly.

Inactive Data

Not all data is created equal. Some data sets have to be immediately available to active applications and services, while other data is not critical to production workflows. In 60% of organizations, 25+% of the data fits this latter category, i.e., inactive. In 24% of organizations, 50+% of the data is inactive (Osterman Research).

A good example is in the healthcare sector, where data is active while the patient is in treatment, but can be moved to a lower-cost tier shortly after the patient is released. However if the patient returns, even years later, the data needs to be retrieved seamlessly.

The cloud providers offer different cloud storage tiers for different use cases, with trade-offs between performance and cost. For example, Amazon Elastic Block Storage (Amazon EBS) costs $0.10 per GB-month, while Amazon Simple Storage Service (Amazon S3) offers object storage at only $0.023 per GB-month.

However, the cloud service providers don’t provide built-in mechanisms that support automated data tiering that is able to automatically identify and seamlessly move data between low-cost object storage and more performant disk storage as required—and doing so manually is not possible. The challenge is to ensure that active and inactive data can be detected and stored in the most cost and performance efficient storage tier automatically.

> The annual savings for storing inactive data in cost-effective storage can be in the hundreds of thousands of dollars.

However, the cloud service providers don’t provide built-in mechanisms that support automated data tiering that is able to automatically identify and seamlessly move data between low-cost object storage and more performant disk storage as required—and doing so manually is not possible. The challenge is to ensure that active and inactive data can be detected and stored in the most cost and performance efficient storage tier automatically.

Challenges to housing inactive data include:

Costs

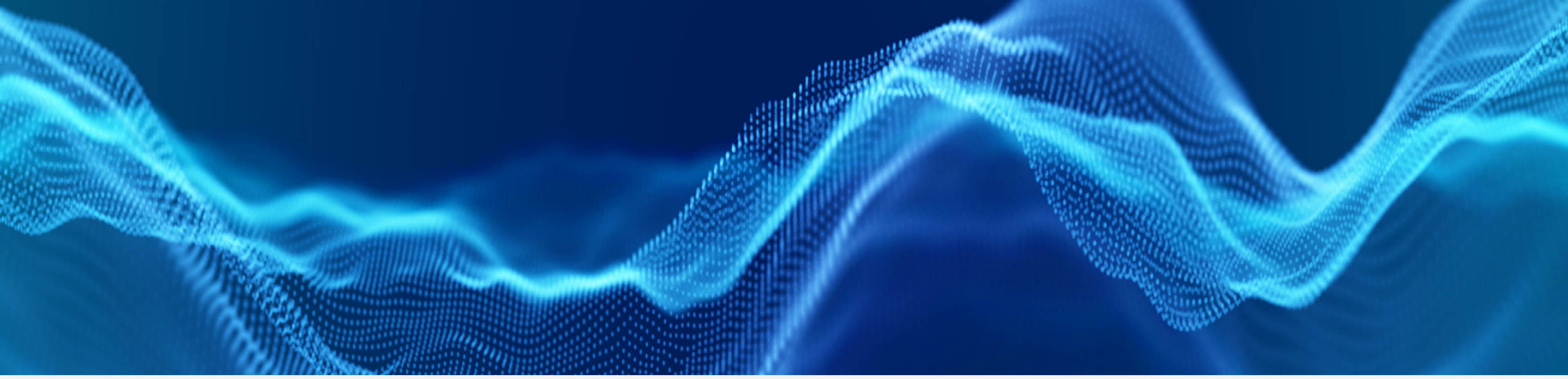
Inactive data can take up a considerable amount of storage footprint though by nature it is not in use. This is a major cost concern.

Storage Optimization

The cloud offers specific storage types for performant use and for inactive use, but moving data between those tiers is a process that the user needs to manage and carry out.

Oversight

Managing the data lifecycle can be a difficult, time-consuming task for storage admins, who may not be aware of what data is active, inactive, or archivable.



Automation

Automation is of great importance to organizations operating in the cloud at scale, in order to carry out complex tasks and workflows such as provisioning and managing volumes, snapshots, and clones, setting up replications, and so on. In addition, resource provisioning and infrastructure updates should not become a bottleneck or introduce the risk of errors in the face of dynamic business demands on the cloud environment. In general, automation makes tasks and processes standard and repeatable. The cloud providers offer some automation services and tools, but other functions will need to be deployed externally.

Some typical cloud automation use cases are:

DevOps
Continuously roll out multiple test, integration, and production environments by combining the deployment of storage services and other cloud infrastructure in a single operation.

Cloud Migration
Quickly and reliably create volumes, data replication relationships, file shares, and much more through automated templates and scripts.

Disaster Recovery
Continuous automated syncing to the secondary site, as well as fully orchestrated failover and failback procedures in the event of disaster.

Keys to achieving optimal data storage automation include finding solutions for:

Management
A centralized, single-pane management and control interface through which even the most complex automated workflows can be orchestrated.

Integration
Strong API integrations that turn diverse applications and services—including services from the cloud provider—into an end-to-end stack across which automated workflows can run.

Repeatability
Leverage Infrastructure as Code (IAC) wherever possible so that the whole enterprise is using the same infrastructure source code to deploy environments with the same configuration every time.

File Share Accessibility

Organizations rely on enterprise file sharing tools to provide access for critical files from multiple devices and platforms. An enterprise organization's IT landscape is usually a mix of Windows and Linux machines—in certain scenarios a single file share may need access by both Windows and Linux machines at the same time. Managing this could get complex, as Windows and Linux follow different security semantics and protocols.

Enterprises running file shares in the cloud have to address challenges such as:

Interoperability
To meet the demands of both Linux/Unix and Windows workloads, a file share solution should enable access with both NFS and SMB / CIFS protocols and any of those protocols various versions or flavors.

Multicloud environments
Accessibility and interoperability issues become even more complex if file shares are being accessed in different clouds.

Customized solutions
There is no standalone cloud-native solution that provides multi-protocol access. Configuring an in-house solution that uses multiple services and tools, however, can be prohibitively expensive and time consuming.

Container Platforms and Kubernetes Integration

The ability to manage applications independently of infrastructure holds great value for cloud deployments. In recent years, software developers and DevOps engineers have benefited from encapsulating applications into containers, each of which is a lightweight, independent unit. Kubernetes takes container deployment to a whole new level by providing a robust solution for managing and scaling containers and containerized applications and workloads across a cluster of machines.

According to a StackRox survey, as of the end of 2019, close to 70% of organizations have containerized more than 10% of their applications, and 13% have containerized more than half. Of the organizations running containers, 86% have adopted Kubernetes. A little more than half are running containers both on-premises and in the cloud.

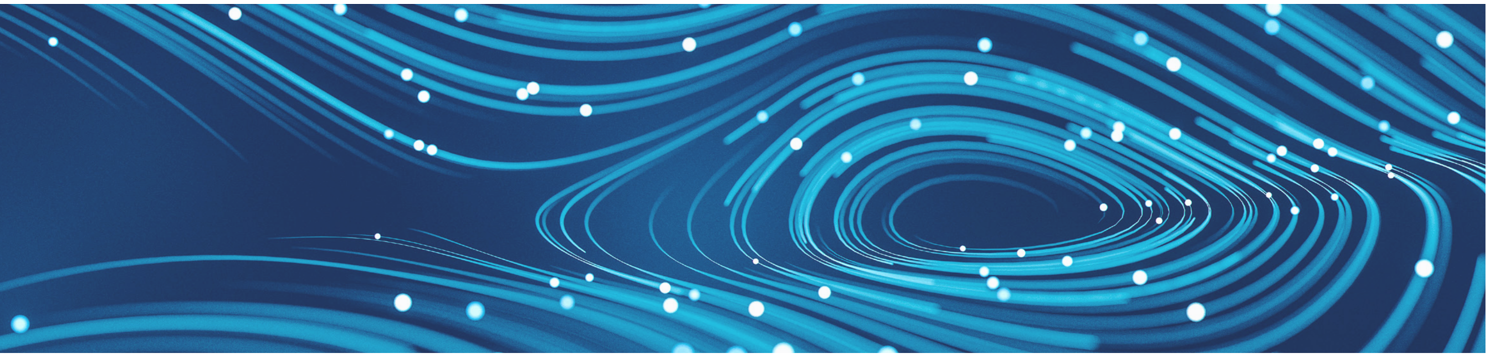
One of the biggest challenges for enterprises that wish to deploy containerized applications and services is managing persistent data storage. A storage volume is provisioned directly to each Kubernetes pod, and is destroyed when the parent pod is deleted. This type of ephemeral storage is fine for a stateless application; however, stateful applications, such as database services, require persistent volumes that exist independently of any specific pod and with its own lifetime.

Dynamic allocation of persistent volumes (PVs) requires the use of an external provisioner. Look for a provisioner that will know how to handle your PVs in the cloud, including: allocation, file access, data protection, backups, and more.

Container and Kubernetes Challenges:

Persistent Volume Provisioning
Whether using static or dynamic provisioning, Kubernetes does not include a built-in function to provide persistent volumes.

Over/under provisioning
Static provisioning will require you to know the exact amount of storage needed in advance, or else you risk under or over provisioning.



Multicloud and Hybrid

The #1 concern (by far) of enterprises with regard to cloud storage is how to connect and integrate with their existing infrastructure (Osterman Research). This concern is a reflection of the growing adoption of hybrid architectures. According to the RightScale (now Flexera) 2019 State of the Cloud Report, 84% of enterprises have embraced a multicloud strategy in general, of which 58% are managing both public and private clouds in a hybrid architecture.

This trend towards multicloud and hybrid environments is expected to continue because of their clear benefits such as agility, flexibility, scalability, and avoiding vendor lock-in.

However, there are a number of data storage challenges to be considered when building out a hybrid/multicloud strategy:

Visibility
How are you going to be able to identify, track and monitor data storage resources wherever they reside? Or as they move from one platform to another?

Unified management
How are you going to effectively manage data resources across fragmented data management stacks—corporate, different cloud providers?

Interoperability
Data synchronization and collaboration between different geographies, clouds, and within hybrid deployments needs to be seamless.

Privacy and security
The more dispersed your data storage resources, the more attack surfaces you are exposing. Corporate security and access controls must be applied robustly and consistently across different providers and environments.

Compliance
Some industries are subject to strict regulations as to where data can reside. Cloud customers never really know where their data storage resources are located geographically, unless special terms are established with the provider.

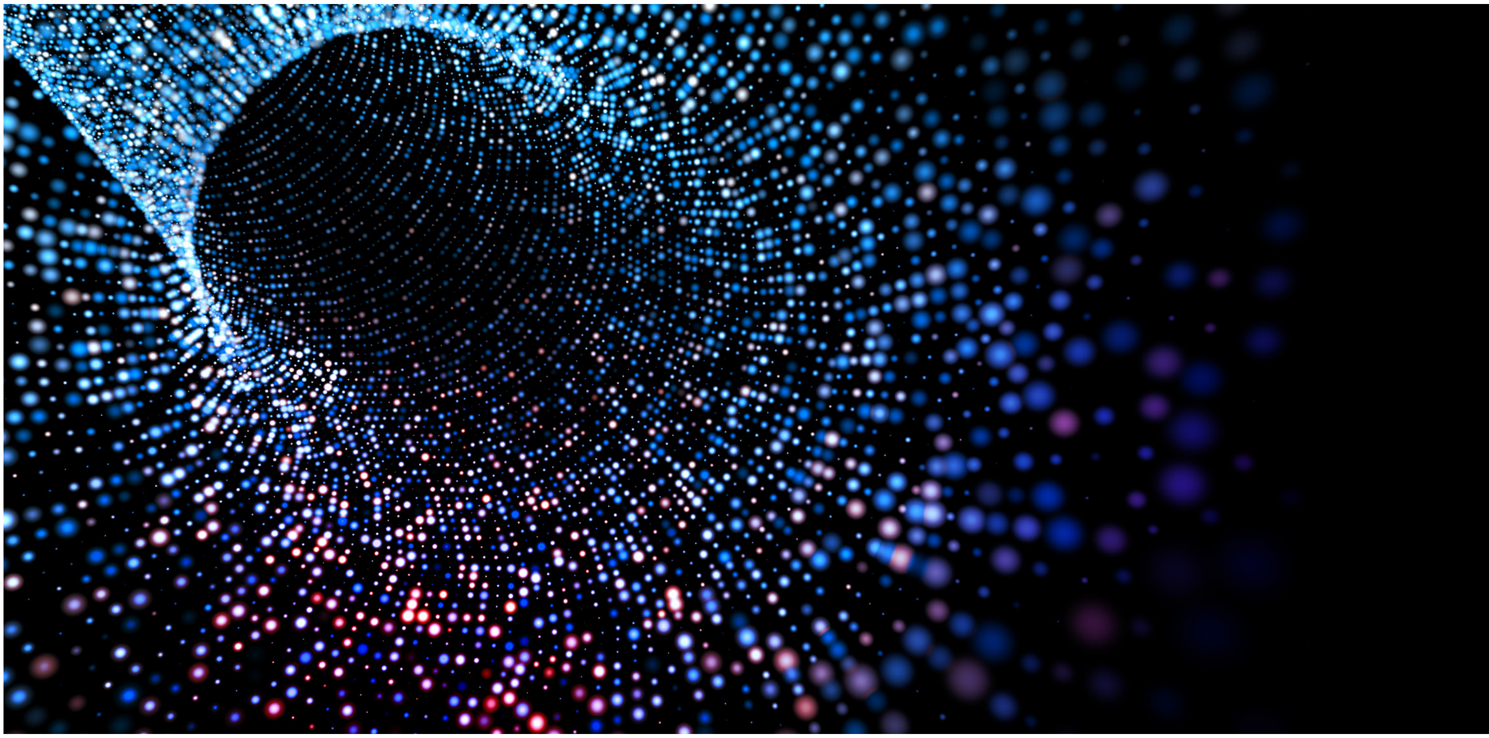
Strategic Organizational Challenges To Be Considered

In this section we look at three challenges that should not be overlooked as organizations plan and implement their cloud data strategies: which workloads will be migrated and how, how new skills will be brought into the organization, and potential impact on the enterprise's business processes.

Consider the migration Netflix carried out in 2008. They realized their on-premises backend could no longer cope with the company's growth and made a strategic decision to move to the cloud. As the migration picked up momentum, they shifted from a monolithic to a microservice application architecture—basically redesigning their organization from the ground up.

> Netflix' story is a prime example of the profound influence that a move to the cloud has on an entire enterprise.

NETFLIX



Migration Plan

The enterprise's cloud strategy must rest on a clear mapping of its current workloads and an understanding of what each requires to be moved to the cloud. Some legacy workloads will never make it to the cloud because the operational risk of moving them is too great. Until those workloads can be replaced with cloud-native alternatives, the enterprise will have to manage a hybrid environment.

For those workloads that are suitable for cloud migration, the three main options are IaaS, PaaS, and SaaS:

1

Infrastructure-as-a-Service (IaaS), otherwise known as a lift and shift migration, i.e., migrating applications and their associated data to the cloud with minimal or no changes to the application architecture, data flow, or authentication mechanisms. If the application's current compute, storage and network requirements can easily be mapped to matching cloud resources, then a lift and shift migration approach should be the first choice.

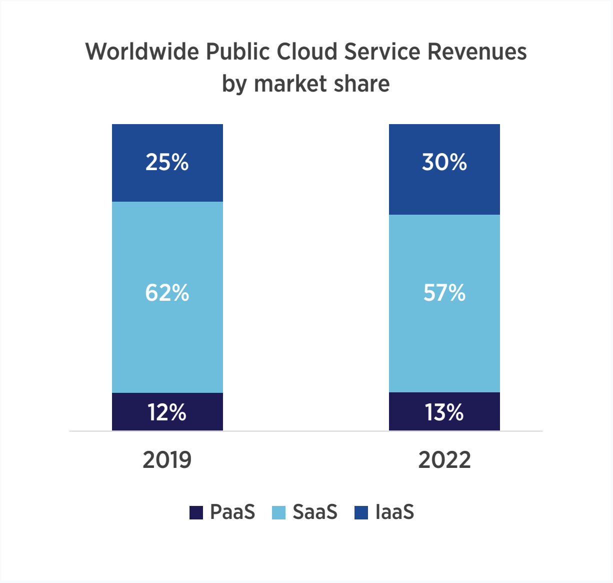
2

Platform-as-a-Service (PaaS), uses a platform configured by the cloud provider to run the application's code. This approach requires applications to be significantly refactored or rebuilt to fit the target cloud platform, including code rewrites, library updates, deployment pipeline modifications, and more.

3

Software-as-a-Service (SaaS), totally replaces functionalities or components of an existing workload or application with a third-party SaaS service. Although faster than a PaaS migration, it involves all the challenges of adopting a new technology such as restructuring parts of the architecture, creating new interfaces, educating teams on its use, and more. Other concerns include the complications of data migration, access management, vendor lock-in, and so on.

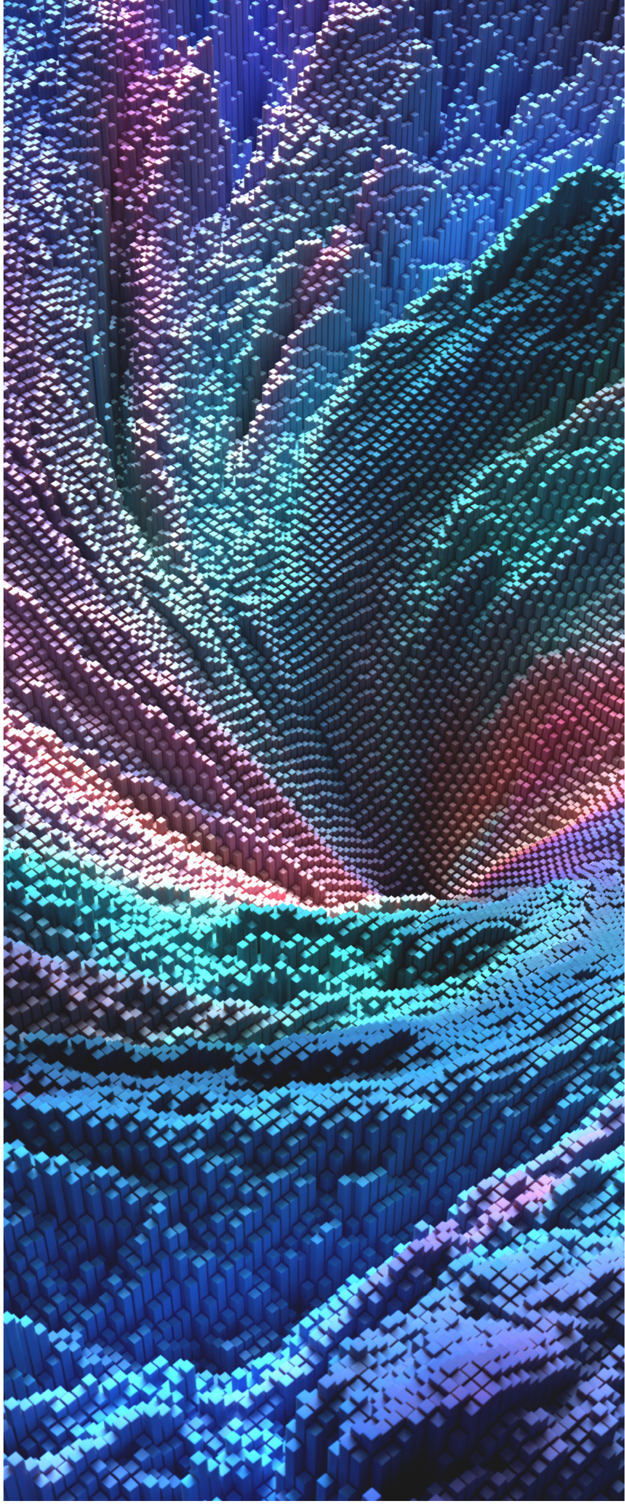
Most companies will find themselves using all three migration strategies during their move to the cloud, with the mix most likely reflecting the global market shares of these three service models. According to Gartner (April 2019), SaaS has the largest market share in 2019 and will continue to hold that position in 2022, but IaaS will gain 5% in market share over the next three years, at the expense of SaaS.



Training

Managing, securing and using data in the cloud requires experience, skills and certifications that the current development, operations and security teams do not have. Some of those gaps may be closed by new hires and others through training. Both recruiting and retraining require significant budget and time allocations that must be taken into account as the enterprise shifts workloads and data to the cloud.

Another cohort that requires training are the enterprise's cloud end-users. The shift to the cloud affects almost every department, from sales, marketing, and finance to human resources. These in-house end-users will need to be brought up to speed on how to provision the cloud resources and services that they need, how to track usage, how to secure data in the cloud, and more. A detailed cross-enterprise training program must be designed and implemented, with the required budget and time allotments for both trainers and trainees.



Impact on Enterprise Processes

Never underestimate the impact that moving workloads and data to the cloud at scale has on business processes and procedures throughout the enterprise. Whether it's security, compliance, backups, monitoring, workflow management, or infrastructure resources management—all standard procedures will have to be re-examined and aligned with how things are done in the cloud.

Two other factors to consider in this context are:

> Ensuring that SLAs already in place with customers and partners will be upheld in the cloud and, if not, agreements must be adjusted accordingly.

> The cloud onboarding process will at times disrupt and even temporarily interrupt business processes. Planned downtime must be scheduled carefully and in coordination with all of the departments affected.

A Single Solution for Your Cloud Challenges: Cloud Volumes ONTAP

Cloud Volumes ONTAP is an enterprise-grade data management solution that runs in AWS, Azure, Google Cloud and hybrid environments. Cloud Volumes ONTAP brings to the cloud the same rich set of data storage and protection features as NetApp's industry-leading on-premises platform. Current NetApp customers who adopt Cloud Volumes ONTAP benefit from a seamless Data Fabric managed from the Cloud Manager single-pane UI or API.

1 Cloud Migration

- NetApp's SnapMirror® data replication technology replicates data to the cloud from NetApp on-premises systems with no need to change applications (the "lift and shift" approach).
- NetApp Cloud Sync service can migrate any kind of file data, or object.

2 Availability

- Cloud Volumes ONTAP high availability (HA) configuration uses a redundant environment composed of two nodes across regions, supporting automatic failover and failback.
- No data loss (RPO=0) and minimal recovery time (RTO < 60 secs) guaranteed.
- Keeps testing, updates, and planned maintenance nondisruptive.

3 Performance

- Based on NetApp's Write Anywhere File Layout (WAFL) technology, Cloud Volumes ONTAP is able to grow rapidly along with support for enormous RAID arrays.
- Intelligent data caching technology brings users closer to data no matter where it's stored.
- Meet demand spikes non-disruptively and with no performance impacts when changing instance types, disk types, dynamically resizing or expanding volumes, and data tiering.
- Support for HA configurations, Snapshot schedules of 5 seconds, instant clone volumes, and data replication and sync between systems—all non-disruptively and with no performance impact.

4 Data Protection

- NetApp Snapshot™ copies are created instantly no matter how large the source volume.
- SnapMirror® transfers and synchronizes data from the primary data set to the DR copy across regions and clouds.

5 Backups and Archiving

- SnapMirror and SnapVault® create secondary systems across hybrid environments, regions and clouds.
- Archive data from multiple source storage systems to a single destination.
- Instantly restore complete volumes or single files.

6 Compliance and Governance

- NetApp Cloud Manager provides a single pane for integrated, unified, standard, and repeatable automation and orchestration policies and processes.
- Cloud WORM (Write Once Read Many) storage via NetApp SnapLock® technology.
- NetApp Cloud Compliance data mapping finds sensitive data to easily facilitate compliance with regulations such as GDPR, CCPA, HIPAA, etc.

7 Security

- Data Encryption: All data at rest is fully encrypted. Data in transit is encrypted when using SMB3+/ NFS4.1+protocols.
- By deploying within your VPC/VNet you can make sure that your data is always within your control.
- Support for ransomware protection, suspicious file auditing, Vscan integration, and multi-tenancy protection.

8 Storage Footprint and Costs

Cloud Volumes ONTAP reduces storage footprint and costs through an array of storage efficiency features, reducing footprint and network transfer costs:

- **Data Compression:** Automatically compresses data blocks, reducing space consumed, saving up to 87%, depending on the application.
- **Data Deduplication:** Eliminates duplicate data blocks by saving a pointer to the original block, reducing storage as much as 90%, depending on the application.
- **Data Compaction:** Combines small data into 4 KB physical blocks.
- **Thin Provisioning:** Allocates storage only when written to a volume rather than ahead of time.
- **NetApp Snapshot Copies:** Extremely space efficient, saving only data changes.
- **Data Clones:** Writable data clones, consuming storage only for changes that are made to the clone.

9 Inactive Data

- Automated tiering of data from performant disk storage to low-cost object capacity tiers.

10 Automation

- Automate provisioning volumes and disk aggregates, setting up new storage systems, creating snapshot schedules, provisioning clone environments, and creating new file shares via Cloud Manager UI or API.
- Integrated with automation tools such as Jenkins, Ansible, and OpenStack.

11 File Share Accessibility

- Supports all versions of NFS and SMB / CIFS file shares in all clouds or on-premises.
- Full Active Directory integration.
- Keep your existing domain credentials and group memberships.
- Multi-Protocol-Access grants users NFS and SMB / CIFS protocol access to the same files and sharing between Linux and Windows clients.

12 Container Platforms and Kubernetes Integration

- NetApp Trident is a dynamic provisioner that uses Cloud Volumes ONTAP as the storage back end for persistent volumes.
- Adds all of Cloud Volumes ONTAP's data storage management capabilities to Kubernetes clusters, OpenShift, Docker, and other container platforms.

13 Multicloud and Hybrid

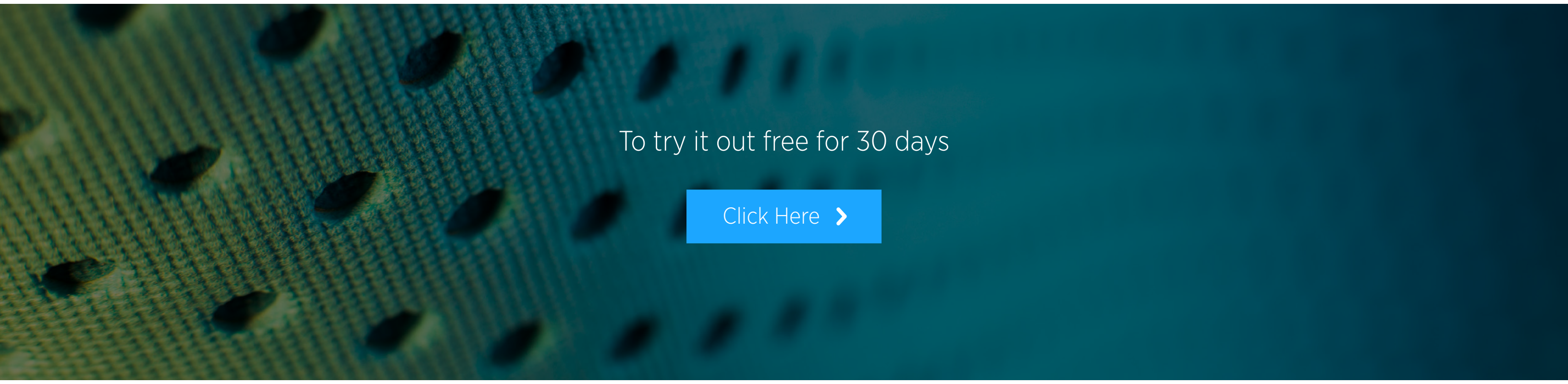
- Supports AWS, Azure and GCP, as well as private cloud and on-premises deployments.
- Single-pane view and control of all storage components.
- Creates a Data Fabric for seamless automation and orchestration of data workflows across all environments.

Summary

The cloud has brought huge benefits to give to every organization that migrates, but getting there isn't a one-step process. In this guidebook we've given you some insight into the challenges that await any business that is undertaking a cloud migration and beginning to research those first few steps involved in taking the enterprise beyond the data center. Those first few steps can make or break a cloud journey.

NetApp's solution for this is Cloud Volumes ONTAP.

Whether you're moving to AWS, Azure, Google Cloud, or any combination of clouds and on-prem deployment, Cloud Volumes ONTAP is the data management platform that thousands of enterprise businesses trust to facilitate, protect, develop, and enhance their IT needs from the data center and into the cloud.



To try it out free for 30 days

Click Here ›

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

Copyright Information

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

Trademark Information

NA-287-0218