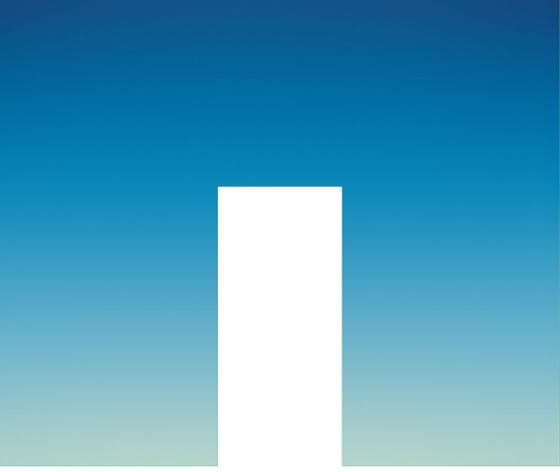


WHITE PAPER

File Services with Azure NetApp Files and Cloud Volumes Service



Executive Summary	3
Challenges of Using File Services in the Cloud	3
Major File Services in the Cloud	3
Azure Files	4
Google Cloud Platform	4
“Roll-Your-Own” NAS Solution	5
Azure NetApp Files and Cloud Volumes Service	5
Conclusion	6

Executive Summary

File shares provide data access and management capabilities for files to any client, in any geography and data format, at any scale. [Migrating file share workloads to the cloud](#) offers the benefit of limitless scalability, while transparently addressing concerns about high availability and resilience to system failure.

With SMB- and NFS-based file shares, file systems can give concurrent access to hundreds or even thousands of user machines. These cloud-based file services also support media processing workloads, [big data analytics](#), backups to offsite repositories, and more. With such a wide range of use cases, a large number of factors must be considered when deciding on a cloud-based file service. Will it provide the performance levels that enterprise workloads require? Would opting for a fully managed service be better than relying on a file share that needs resources and labor to maintain?

Answering these questions is an important first step in choosing file services. This white paper looks at shared storage for file services in the cloud. The first section explains the challenges of running shared storage for file services in the cloud; introduces some of the major offerings available from the major cloud providers AWS, Azure, and Google Cloud Platform; and concludes by showing how [NetApp® Cloud Volumes Service for AWS](#) and [GCP](#) and [Azure NetApp Files](#) provide the highest performance in file services in the cloud today.

NetApp Cloud Volumes Service and Azure NetApp Files offer something brand new: a fully-managed service built on NetApp technology running in the cloud. It supports SMB, NFSv3, and (coming soon) NFSv4 protocols and allows full data mobility between any end points for hybrid and multicloud flexibility. The service combines NetApp's data management capabilities and the cloud's unlimited availability and scalability, without any of the usual administrative tasks and overhead.

Challenges of Using File Services in the Cloud

When it comes to shared storage in the cloud, enterprises face all of the challenges of running a major workload in the cloud, in addition to the need to scale up unlimited amounts of storage with guaranteed availability while also ensuring compatibility with host clients' data formats and operating systems. This section looks at these challenges and requirements individually.

1. Extreme File Service Performance

File services are extremely dependent on the high level of performance of the storage environment in use. Some workloads may require a high-level SLA to maintain peak operability and IOPS levels.

2. Host Client Data Compatibility

An enterprise's file services need to be compatible with all their host clients' data formats and operating systems.

3. Guaranteed Business Continuity

When it comes to running an enterprise file service, any disruption in normal operation can negatively impact the business. Whether an outage is caused by a disaster or through an update process, it is essential to ensure the availability of the file share, with zero downtime and no data loss.

4. Data Protection

File shares require companies to comply with industry-specific data security, data protection, and disaster recovery regulations. NetApp Snapshot™ technology is key to meeting these requirements, because Snapshot copies can be used to restore systems in the event of malware or ransomware attacks. And because file shares can be the most important part of a business' operation, these copies must be able to be automated to meet stringent RPO, RTO, and backup requirements.

5. Migration

When using a cloud-based file share, it's important to find a solution for migrating data to the cloud without having to rewrite the applications that will depend on the data. If applications need to be recoded, the enterprise has to commit significant amounts of time and money, and to test the applications to ensure that the end user isn't negatively affected. Migration needs to be as seamless as possible.

6. Keeping Data in Sync

Following up on the previous point, any data that moves to the cloud in the file share must be kept synchronized among the disparate environments in use, whether it's data housed on premises or in one or more cloud environments.

7. Data Security

With shared storage, data security becomes a major concern. It's important to make sure that access to the file system is in the user's control at all times. Ensuring data security at all levels is crucial to preventing losses before they happen. Role-based access and proper key management within organizations can keep data more secure.

Major File Services in the Cloud

This section profiles the cloud-based file services available through the major public cloud service providers: Amazon EFS, Azure Files, the options of Google Cloud Platform, and IBM Cloud File Storage.

Amazon EFS

Amazon EFS (Elastic File System) greatly simplifies the setup of a horizontally scalable cloud-based file service, allowing access to the same files concurrently from hundreds or even thousands of different clients. Because Amazon EFS is a fully managed solution, it's very easy to get started. The wizard interface allows users to create a new, highly available file system in just seconds, with all files redundantly stored across availability zones. Amazon EC2 Linux hosts can use standard commands to connect to the file system over NFS.

[Amazon EFS](#) can be used to grant access to the stored data for other cloud services, such as big data analytics, and it can be used to support the shared data required by web applications and content management systems. It could also be used, for example, as a backup location for on-premises file servers.

The primary way to use Amazon EFS is with [Amazon EC2](#), by managing network access using security groups, which behave somewhat like a firewall for the file system. Mounting the file system from an on-premises file server requires the use of [AWS Direct Connect](#). This must be done over a connection outside of the internet, because Amazon EFS support for AWS virtual private network connections is not currently available.

Billing for each Amazon EFS file system in use is based on the amount of storage consumed per month, calculated for varying usage rates measured in GB hours.

The challenge of protecting the data in the file system must be considered when using Amazon EFS. Unlike Amazon EBS, Amazon's popular block storage offering, so far Amazon EFS has no built-in feature to facilitate the creation of Snapshot copies or backup. Workarounds, such as using [AWS Data Pipeline](#) to back up to another Amazon EFS file system, do exist, but such solutions require manual operations. However, enterprise-grade RTO and RPO demand higher levels of protection than it's possible to create manually. The secondary Amazon EFS file environment also doubles storage costs. One strategy for avoiding the high cost of a secondary copy is to store it on Amazon S3, where costs are lower. However, users still need to consume Amazon EFS burst credits to make the copy, and the copy's history and location must be maintained manually.

It's also important to be aware of capacity and throughput benchmarks. Amazon EFS uses a [burst credits](#) system to set the performance of the file system, which means that the amount of data stored is the deciding factor for performance. Larger file systems gain more burst credits and therefore are able to operate at peak rates, although smaller, active systems may find that their credits fall below their operating requirements.

Amazon EFS Quick Hits

Benefits:

- Managed by AWS
- Easy setup
- Built-in scalability and high availability

Considerations:

- Currently no native backup or snapshot features
- Support for NFSv4
- Latency, IOPS, and throughput levels
- For use in the Oregon, Ohio, Virginia, Ireland, and Sydney AWS regions

Azure Files

The complementary service to Amazon EFS in the Microsoft public cloud is [Azure Files](#), which provides file share services for SMB v3.0. Given that SMB is primarily for use with Microsoft systems, it's notable that Azure Files is also mountable for read/write operations with MacOS and Linux operating systems.

Support for the newer version of SMB enables features such as encryption in transit, meeting the file service demands for high levels of data security. However, the same level of protection is possible using the RESTful calls via HTTPS. Although they are easily creatable through the user interface, Azure Files file shares can also be created by using the Azure CLI or PowerShell.

In addition to Azure Files, Microsoft also offers [Azure File Sync](#). This service, now in preview, offers more opportunities for Azure Files to work with on-premises storage architectures. Through an agent installed in the on-premises system, Azure Files can cache data locally, increasing access speed and allowing writes to sync transparently with Azure Files. Configuring additional servers provides a way to unite storage environments located in multiple geographic areas.

Azure Files contains a backup feature, currently in preview. [Azure Files Share Snapshots](#) gives users the ability to create read-only snapshots of the Azure Files file share that can be used for backup.

Charges for Azure Files are based on two factors: storage consumed and file access. Additional costs can accrue for the use of the additional features, including Azure File Sync.

Azure Quick Hits

Benefits:

- Managed by Azure
- Easy setup
- New services adding backup and faster speed

Considerations:

- Currently supports SMB
- Snapshot capability
- Max volume size 5TB
- Native backup system remains manual
- Active Directory support through the use of Azure File Sync

Google Cloud Platform

[Cloud FileStore](#) is a new Google Cloud Platform file service, currently in beta. Because this service is still in beta, many users on GCP may already be using in-house file solutions, and NetApp Cloud Volumes Service can also be used for GCP. Both options are discussed later in the document.

Cloud FileStore, a new service for fully managed, high-performance NFS file shares, enables users to easily deploy NFSv3 file shares and to choose between Standard and Premium performance tiers. Each file share must be between 1TiB and just under 64TiB in size. The service is bound to a single zone without the option of failover and may experience short periods of downtime in response to error recovery. There is currently no built-in backup or snapshot capability, which means that users need to manage these data protection operations manually.

GCP Cloud FileStore Quick Hits

Benefits:

- Fully managed
- NFSv3 support
- Two performance tiers available

Considerations:

- Manual backup and data protection procedures required when deploying to Cloud FileStore
- Cloud FileStore is currently in beta and is not covered by any SLAs
- 1TiB to 64TiB file share size

“Roll-Your-Own” NAS Solution

An alternative to using a managed, cloud-based NAS service—and by far the most common file solution today—is to use cloud compute and storage resources to create a custom file share solution. At first glance, this kind of setup is appealing due to the high degree of control that it offers over the type of storage used, access protocols, file system features, and more. However, these solutions very often run into complications.

The first major issue with a custom NAS solution is ensuring reliable access to the files and guaranteeing cross-site durability. Most block-level cloud storage solutions provide redundancy within a single location; however, extra steps need to be taken to ensure durability beyond a single site. Asynchronous replication or copying the data on a schedule always means that the files being stored are susceptible to data loss.

The second issue is the level of manual administration that these solutions require. For example, allocating a new file system requires allocating new storage, mounting it to the compute nodes that will serve out the data, and potentially initializing the new share with existing data. If the file system needs to grow, this growth must be handled manually. If the performance of the underlying disks needs to be upgraded, the allocation of the new storage and migration of existing files need to be taken care of while still trying to minimize downtime.

The third issue is the complexity of managing the storage over time as the deployment grows. Storage administrators working with production file shares need to maintain uninterrupted access to the files, provide backup or snapshot facilities, allow test copies of the data to be created, and much more. Providing

robust support for this kind of functionality requires a high level of technical expertise.

Let’s look at one such solution. [The GlusterFS file share](#) can distribute files across both on-premises and virtual devices, offering scalability and reliability. GlusterFS can be used for GCP, AWS, and Azure. It is capable of making connections via NFS, SMB, and iSCSI, with the correct drivers and add-ons in place.

[Gluster storage configurations](#) offer a range of options for file sharing, including striped, replicated, dispersed, distributed, and a number of combinations of those. It also has the ability to take snapshots, as well as to clone snapshots.

As an open-source technology, all of the setup and configuration for GlusterFS are up to the user. This can be a considerable task that may require a large investment in time and money. Problems that arise through use of the platform require maintenance and active resolution. Although GlusterFS is an open-source technology, support for it is commercially available through [Red Hat](#).

“Roll-Your-Own” Quick Hits

Benefits:

- Full control over the deployment, including storage disks and access protocols
- Suitable for small deployments

Considerations:

- No built-in cross-site data redundancy
- Manual administration required
- Complex to support over time

Azure NetApp Files and Cloud Volumes Service

Cloud Volumes Service for AWS and GCP and Azure NetApp Files are NetApp’s fully managed, high-performance, cloud-native file services. Cloud Volumes Service is deployed out of the AWS or GCP cloud marketplace while Azure NetApp Files is a native Azure service, delivered and support by Microsoft built on NetApp technology. The services deliver multi-protocol support (SMB, NFSv3, and NFSv4 coming soon), advanced data services, and integrate directly with both platform and compute services.

Users can select one of three service levels: Standard, Premium, or Extreme with the following corresponding IOPS and throughput for each service level:

- **Standard service level:** Up to 1000 IOPS per TB (16k I/O) and 16MB of throughput per TB.
- **Premium service level:** Up to 4000 IOPS per TB (16k I/O) and 64MB of throughput per TB.
- **Extreme service level:** Up to 8000 IOPS per TB (16k I/O) and 128MB of throughput per TB.

High-performance service levels come with the ability to spin up 100TB of storage in under 10 seconds with a high level of data protection. The performance level enables companies

to move workloads to the cloud that would never have been possible before. Even more importantly, with Azure NetApp Files and Cloud Volumes Service, users no longer have to worry about storage management. NetApp takes care of all the setup, configuration, updates, performance, and service levels. The service is built on the NetApp infrastructure that has served enterprise customers for more than 25 years – faster and better than ever.

For data protection, no solution supports a file service better than NetApp Snapshot technology. These Snapshot copies give Azure NetApp Files and Cloud Volumes Service the ability to provide companies with point-in-time backups that can be created instantly and automatically, ensuring that important data protection goals are maintained. You also get powerful cloning technology. These quickly created, non-disruptive data clones can be used for a number of purposes, such as building test environments.

Azure NetApp Files and Cloud Volumes Service have low latency, which enables predictable performance that other file share services can't guarantee. Also, built-in always-on protection comes with encryption of data at rest, encryption in transit for SMB connections and over VPNs, high availability, and NetApp reliability.

- Fully managed service. Data management is handled completely by NetApp, not the customer.
- Popular clouds. Fully capable file service for AWS, Azure, and GCP.
- Scalability and performance. Spin up to 100TB at a time of extremely high-performance storage in just seconds.
- Multiple protocols. Support for SMB, NFSv3, and (coming soon) NFSv4 file shares
- Compatibility. Supports shared file access across Windows and UNIX/Linux operating systems for greater host client data operability.
- Integratable. Complete integration with file directory metadata, keeping domain credentials, access and authentication, and group memberships, including full compatibility with Microsoft Active Directory.
- Provides both high performance and low cost.
- Data protection. Data corruption or loss can be prevented with efficient, automatic data Snapshot copies.
- Migratable. Import data from on-premises and other storage repositories via Cloud Sync.
- Business continuity. High availability ensures business continuity with no data loss (RPO = 0) and short recovery times (RTO < 60 secs).
- Automation. Schedule tasks to meet file share demands with automation and orchestration capabilities. Even more automation and orchestration capabilities will be included in the upcoming release of the RESTful API.

Conclusion

File services in the cloud come with a range of options, but users should be cautious about deciding to move ahead with a file share. It's important to make sure that the file service has the capacity and features to meet their organization's needs. Most important in this case are scalability, data protection, and compatibility, all of which Azure NetApp Files and Cloud Volumes Service can provide better than any other option.

To get the most out of the cloud, sign up today for [Cloud Volumes Service for AWS or GCP](#), or [Azure NetApp Files](#).

NetApp is the data authority for hybrid cloud. We provide a full range of hybrid cloud data services that simplify management of applications and data across cloud and on-premises environments to accelerate digital transformation. Together with our partners, we empower global organizations to unleash the full potential of their data to expand customer touchpoints, foster greater innovation, and optimize their operations. For more information, visit www.netapp.com. #DataDriven

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

Copyright © 2018 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

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.

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.

WP-7285-1018