AWS Prescriptive Guidance

Deploying Amazon FSx for NetApp ONTAP in an enterprise environment

AWS Prescriptive Guidance: Deploying Amazon FSx for NetApp ONTAP in an enterprise environment

Copyright © 2023 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Deploying Amazon FSx for NetApp ONTAP in an enterprise environment

Luigi Seregni, Antonio Aga Rossi, and Giulio Dipace Amazon Web Services (AWS)

August 2023 (document history (p. 17))

Migrating workloads to the cloud supports organizational growth and helps you adapt to the changing market landscape. Cloud capabilities and features can provide scalability, agility, and resilience, which can increase the level of service for applications.

Every year, new rules and regulations are introduced, such as new International Financial Reporting Standards (IFRS) standards. Evolving standards often require increased computational power, and this can be difficult to achieve on premises. High performance computing (HPC) applications, which have extremely high storage throughput requirements, are the perfect candidates for migration to the cloud or a hybrid-cloud environment.

Amazon FSx for NetApp ONTAP is a cloud service that supports the high throughput requirements of these HPC applications and is backward compatible with on-premises ONTAP workloads. Using this guide, you can deploy a fully functional FSx for ONTAP solution in an enterprise environment. *Enterprise environment* means that the guide doesn't focus exclusively on the on FSx for ONTAP service. Instead, it adopts a holistic view and provides considerations for deploying this kind of file system in a complex environment.

This guide also reviews common deployment challenges, such as Active Directory integration, service account deployment, ONTAP command line troubleshooting, and storage virtual machine (SVM) configuration.

The guide provides four different sections:

- **Architecture** – This chapter provides an overview of a possible enterprise architecture using FSx for ONTAP, describing the interaction between the different components and the suitable usage patterns.
- **Creating a file system** – This section lists all the actions to create a fully working FSx for ONTAP environment. We provide step by step guidelines both to deploy the solution manually.
- **Best practices** – This chapter provides recommendations and best practices that are based on lessons learned during the deployment activities of FSx for ONTAP in an enterprise environment.
- **FAQ** – This section contains a set of questions that address common concerns about the technology.

Intended audience

This guide is intended to help cloud administrators and architects who need to deploy an FSx for ONTAP solution that supports HPC workloads and Active Directory integration.

Objectives

This guide can help you and your organization do the following:

- Understand the architecture and deploy a fully functional FSx for ONTAP solution in an enterprise environment

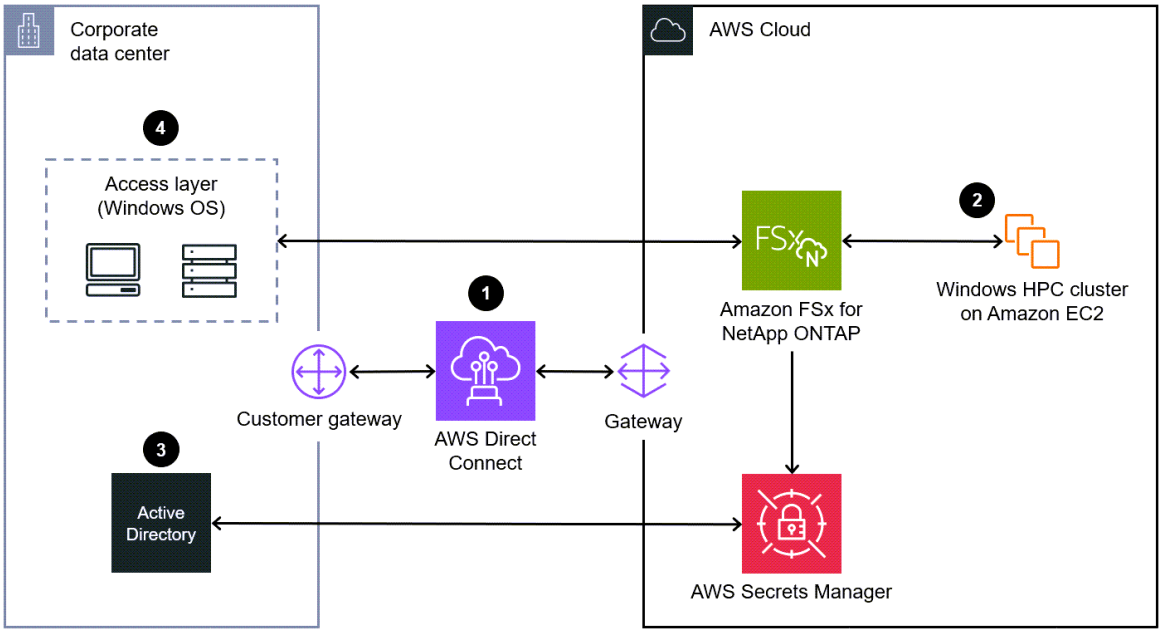
- Integrate FSx for ONTAP with Active Directory
- Create a service account to connect FSx for ONTAP to Active Directory in a production environment
- Manage storage tiers for Amazon FSx
- Troubleshoot through the ONTAP command line
- Troubleshoot configuration issues with storage virtual machines (SVMs) (NetApp documentation)
- Use Service Message Block (SMB) protocol to access data in your FSx for ONTAP file systems

Architecture for deploying FSx for ONTAP in an enterprise environment

Amazon FSx for NetApp ONTAP is a managed storage service that helps you launch and run fully managed NetApp ONTAP file systems in the AWS Cloud. FSx for ONTAP supports Windows or Linux operating systems (OSs), and it is accessible through industry-standard protocols, such as Network File System (NFS), Server Message Block (SMB), and Internet Small Computer System Interface (iSCSI). In addition, this file system supports compression and deduplication, which can reduce storage costs.

This guide focuses on deployment for a Windows workload. For example, you can use FSx for ONTAP as shared storage for an HPC third-party solution that is composed of hundreds of Windows nodes. These nodes have extremely high write and read throughput requirements and are connected to a grid scheduler.

The following diagram depicts a typical example of an enterprise HPC workload and FSx for ONTAP deployment in a hybrid-cloud environment. This architecture is referenced throughout the guide.



The following are the features of this architecture:

1. The on-premises data center and cloud environments are connected by using AWS Direct Connect.
2. The HPC workload, running Windows, is deployed in the AWS Cloud.
3. Active Directory is deployed in the on-premises environment.
4. The access layer systems, which are running on Windows, are deployed in the on-premises environment.

Customer access layer

Through the customer access layer, the end user accesses the workload in the AWS Cloud. Amazon WorkSpaces or Citrix are commonly used to access applications and access the data in Amazon FSx by using an SMB mount.

Active Directory

Typically, Microsoft Active Directory is installed and managed on premises. Many organizations want to join their FSx for ONTAP SVMs to their Active Directory domain in order to provide user authentication and access control at the file and folder level. SMB clients can then use their existing user identities in Active Directory to authenticate themselves and access SVM volumes. For more information, see Working with Microsoft Active Directory in FSx for ONTAP. You must establish proper networking rules to make sure that the SVMs can reach the Active Directory domain.

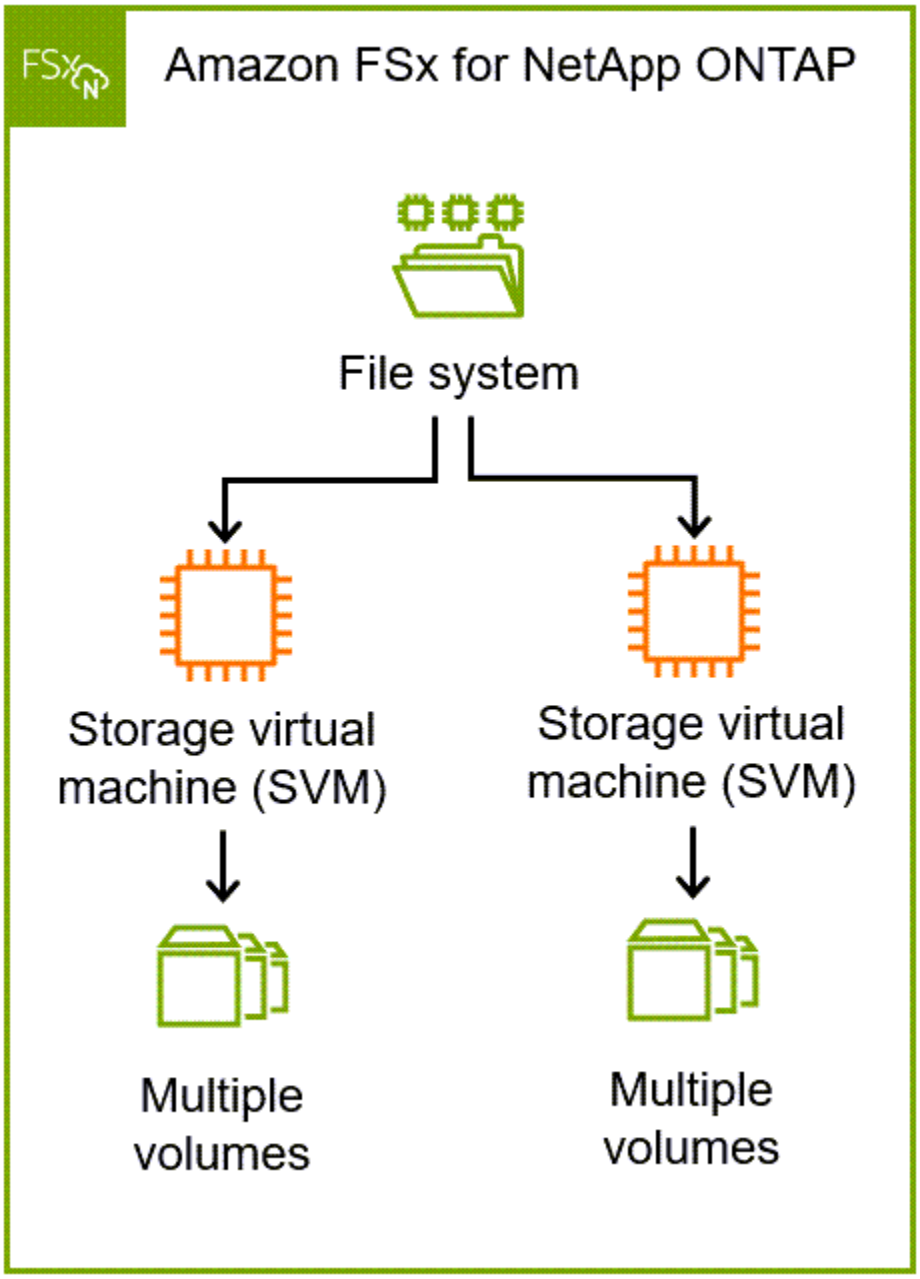
To allow the Amazon FSx file system to create, edit and delete files on the managed volumes, you need to create a service account for the Active Directory domain. For more information, see Delegating permissions to your Amazon FSx service account. Active Directory is a core component in many enterprise organizations, and the deployment of a new account—even with limited privileges—might require considerable time.

Amazon FSx resources

The following are the primary types of resources in FSx for ONTAP:

- A file system is the primary FSx for ONTAP resource, analogous to an on-premises NetApp ONTAP cluster. For troubleshooting, you can use NetApp CLI commands to establish an SSH connection with a file share endpoint. More information about troubleshooting commands is provided later in this guide.
- A storage virtual machine (SVM) is an isolated virtual file server with its own administrative and data access endpoints. The integration between FSx for ONTAP and an Active Directory domain is managed at the SVM level. Therefore, if you get an error regarding Active Directory, the SVM is a good starting point for troubleshooting.
- Volumes are virtual resources that you use to organize and group your data. These are logical containers, and data stored in them consumes physical capacity on your file system. Volumes are hosted on SVMs. You can configure each volume with different tiering policies. Tiering policies are powerful tools that help you manage performance and cost by defining whether data is stored in the performance-optimized SSD layer or in the cost-optimized capacity layer.

The following diagram explains the resource structure of an FSx for ONTAP file system. Amazon FSx fully manages all of the components.



You can join multiple volumes into a single logical namespace by using junction paths (NetApp documentation). To the client, a junction appears to be an ordinary directory. Junction paths provide the benefits of using multiple volumes (such as fine-grained control over snapshot and migration options) with the convenience of accessing data in multiple volumes through a single access point.

Windows HPC cluster on Amazon EC2

For the purposes of this guide, Amazon FSx acts as storage layer for a critical, high-throughput Windows HPC cluster composed of Amazon Elastic Compute Cloud (Amazon EC2) instances. There are multiple approaches for setting up an HPC cluster on Amazon EC2. For an example approach, see Tutorial: Set up

a Windows HPC cluster on Amazon EC2 in the Amazon EC2 documentation. The HPC cluster compute nodes, also known as *worker nodes*, interact with the Amazon FSx file system through SMB shares. You can automatically or manually create the SMB shares on the compute nodes.

AWS Secrets Manager

Enterprise architectures are usually deployed by using infrastructure as code (IaC) tools, such as HashiCorp Terraform. It is a security best practice not to include any sensitive information in IaC scripts. AWS Secrets Manager is commonly used to store sensitive information, such as passwords for Active Directory service accounts.

Creating an FSx for ONTAP file system that is joined to Active Directory

This section contains deployment instructions for creating an Amazon FSx for NetApp ONTAP file system in an enterprise environment and joining your storage virtual machines (SVMs) to an Active Directory domain in your on-premises data center. This chapter reviews the following high-level processes:

- Provisioning an Active Directory service account (p. 7) – To allow the file system to create, edit and delete files on the managed volumes, you create a service account for the Active Directory domain.
- Setting up the FSx for ONTAP file system (p. 8) – You set up the FSx for ONTAP file system, SVM, and volumes.
- Monitoring FSx for ONTAP (p. 10) – You configure logging and monitoring for FSx for ONTAP usage and activity.

Provisioning an Active Directory service account

If you want to join Amazon FSx for NetApp ONTAP SVMs to your on-premises Active Directory domain, you must maintain a valid Active Directory service account throughout the lifetime of the Amazon FSx file system. Amazon FSx must be able to fully manage the file system and perform tasks that require unjoining and rejoining your Active Directory domain, such as replacing a failed file SVM or patching NetApp ONTAP software. Keep your Active Directory configuration, including the service account credentials, updated in Amazon FSx.

This service account must have the following permissions in Active Directory:

- Permissions to join computers to the domain
- In the organizational unit (OU) that you are joining the file system, permissions to:
  - Reset passwords
  - Restrict accounts from reading and writing data
  - Write to the DNS hostname
  - Write to the service principal name
  - Create and delete computer objects
  - Read and write account restrictions

An Active Directory domain administrator can create the service account manually by using the **Active Directory User and Computers** MMC snap-in. For instructions, see Delegating permissions to your Amazon FSx service account in the FSx for ONTAP documentation. You could also configure this account programmatically. For example, you could use PowerShell, as shown in the following example.

```
param(
    [string] $DomainName,
    [string] $Username, #Service Account username
    [string] $Firstname, #Service Account Firstname
    [string] $Lastname, #Service Account Lastname
    [string] $saOU, #OU where Service Account is created
    [string] $delegateOrganizationalUnit #OU where Service Account has delegation
)
```

```
#Retrieve Active Directory domain credentials of a Domain Admin
$DomainCredential = ...

#Import Active Directory PowerShell module
...

#Create Service Account in specified OU
New-Active DirectoryUser -Credential $DomainCredential -SamAccountName $Username -
UserPrincipalName "$Username@$DomainName" -Name "$Firstname $Lastname" -GivenName
 $Firstname -Surname $Lastname -Enabled $True -ChangePasswordAtLogon $False -DisplayName
 "$Lastname, $Firstname" -Path $saOU -CannotChangePassword $True -PasswordNotRequired $True
$user = Get-Active Directoryuser -Identity $Username
$userSID = [System.Security.Principal.SecurityIdentifier] $user.SID

#Connect to Active Directory drive
Set-Location Active Directory:

$ACL = Get-Acl -Path $delegateOrganizationalUnit
$Identity = [System.Security.Principal.IdentityReference] $userSID

#GUID of Active Directory Class
$Computers = [GUID]"bf967a86-0de6-11d0-a285-00aa003049e2"
$ResetPassword = [GUID]"00299570-246d-11d0-a768-00aa006e0529"
$ValidatedDNSHostName = [GUID]"72e39547-7b18-11d1-adef-00c04fd8d5cd"
$ValidatedSPN = [GUID]"f3a64788-5306-11d1-a9c5-0000f80367c1"
$AccountRestrictions = [GUID]"4c164200-20c0-11d0-a768-00aa006e0529"

#Delegation list
$rules = @()
$rules += $(New-Object System.DirectoryServices.ActiveDirectoryAccessRule($Identity,
 "CreateChild, DeleteChild", "Allow", $Computers, "All"))
$rules += $(New-Object System.DirectoryServices.ActiveDirectoryAccessRule($Identity,
 "ExtendedRight", "Allow", $ResetPassword, "Descendents", $Computers))
$rules += $(New-Object System.DirectoryServices.ActiveDirectoryAccessRule($Identity,
 "ReadProperty, WriteProperty", "Allow", $AccountRestrictions, "Descendents", $Computers))
$rules += $(New-Object System.DirectoryServices.ActiveDirectoryAccessRule($userSID, "Self",
 "Allow", $ValidatedDNSHostName, "Descendents", $Computers))
$rules += $(New-Object System.DirectoryServices.ActiveDirectoryAccessRule($userSID, "Self",
 "Allow", $ValidatedSPN, "Descendents", $Computers))

#Set delegation
foreach($rule in $rules) {
    $ACL.AddAccessRule($rule)
}
Set-Acl -Path $delegateOrganizationalUnit -AclObject $ACL
```

Setting up the FSx for ONTAP file system

The Amazon FSx for NetApp ONTAP documentation contains instructions for setting up a file share by using the Quick create or Standard create options in the AWS Management Console. This guide includes additional recommendations and guidance for enterprises that need to support HPC workloads and Active Directory integration.

For instructions on creating an FSx for ONTAP file system in the AWS Management Console, see Creating FSx for ONTAP file systems. For each section, note the following setup recommendations and considerations for enterprise environments.

File system details section

1. For **Deployment type**, choose one of the following:

- Choose **Multi-AZ** for production workloads. This option helps maintain data availability in the event that an Availability Zone is inaccessible.
- Choose **Single-AZ** for disaster recovery purposes, for non-production workloads, or for workloads that have replication already built into the application layer and do not require additional storage-level redundancy. This option is cost-effective for these types of workloads.

For more information and recommendations about configuring the deployment type, see Best practices for choosing an Availability Zone deployment option (p. 12) in the *Best practices* section of this guide.

2. For **Provisioned SSD IOPS**, choose one of the following:
   - Choose **Automatic** if you want Amazon FSx to automatically provision 3 IOPS per GiB of SSD storage, up to up to a maximum of 160,000 SSD IOPS per file system.
   - Choose **User-provisioned mode** if you want to specify the number of IOPS. If you provision a higher level of IOPS, you pay for the average IOPS provisioned above your included rate for the month, measured in IOPS-months.

   For more information, see Storage capacity and IOPS, Considerations when updating storage and IOPS, and Impact of storage capacity on performance in the FSx for ONTAP documentation.

Default storage virtual machine configuration section

1. For **Root volume security style**, choose one of the following:
   - Choose **Unix** if you plan to access the file system mainly from Linux-based clients (NFS protocol).
   - Choose **NTFS** if you plan to access the file system mainly from Windows-based clients (SMB protocol)
   - Choose **Mixed** if you plan to access the file system equally from Linux-based and Windows-based clients. This is an advanced setting.

   For more information about these settings, see What the security styles and their effects are (NetApp documentation).
2. For **Active Directory**, choose **Join an Active Directory**.
3. For **NetBIOS name**, enter the NetBIOS name of the Active Directory computer object that will be created for your SVM. This is often the same as the SVM name but can be different. The NetBIOS name cannot exceed 15 characters.
4. For **Active Directory domain name**, enter the fully qualified domain name of your Active Directory. The domain name cannot exceed 255 characters.
5. For **DNS server IP addresses**, enter the IPv4 addresses of the DNS servers for your domain. You can enter up to three IP addresses.
6. For **Service account username** and **Service account password**, enter the username and password of the service account in your existing Active Directory. This is the service account that you previously set up in Provisioning an Active Directory service account (p. 7) in this guide.
7. For **Organization Unit (OU)**, enter the distinguished path name of the organizational unit to which you want to join your file system.
8. For **Delegated file system administrators group**, enter name of the group in your Active Directory that can administer your file system. The default group is `Domain Admins`.

Default volume configuration section

1. For **Volume name**, enter a name for the volume. The name cannot exceed 203 characters, and alphanumeric and underscore (_) characters are accepted.

2. For **Junction path**, enter the location within the file system where you want to mount the volume. The name must have a leading forward slash, such as `/vol1`. For more information, see the Amazon FSx resources (p. 4) section of this guide.
3. For **Volume size**, enter the storage capacity of your volume, in mebibyte (MiB). Enter any whole number in the range of 20–104857600 (100 TiB).
4. For **Storage efficiency**, choose whether to enable storage efficiency features, such as deduplication, compression, and compaction. For more information, see Storage efficiency in the FSx for ONTAP documentation. If these efficiency features are compatible with your HPC workload, we recommend them in enterprise environments.
5. For **Capacity pool tiering policy**, choose a tiering policy for the volume. For more information and recommendations about tiering policies, see Best practices for storage tiers and tiering policies (p. 11) in the *Best practices* section of this guide.

Monitoring FSx for ONTAP

Logging and monitoring Amazon FSx for NetApp ONTAP usage and activity is a best practice because it helps you understand the state of the file system. You can use the log data to troubleshoot any issues that affect the system's reliability, operations, and efficiency. This can be especially critical in enterprise environments because issues with the file system can jeopardize outcomes of processing tasks, produce inaccurate results, or break service level agreements. These can result in fines for the HPC workload owner or result in decisions that are made with incorrect data. For monitoring best practices for enterprise environments, see Best practices for monitoring FSx for ONTAP file systems (p. 12) in this guide.

There are a variety of AWS and third-party tools and services that you can use to log and monitor FSx for ONTAP usage and activity. For example, you can use Amazon CloudWatch, the Events Management System (EMS) in ONTAP, NetApp Cloud Insights Service, NetApp Harvest, NetApp Grafana, and AWS CloudTrail. For more information, see Monitoring Amazon FSx for NetApp ONTAP.

Best practices for FSx for ONTAP deployments in enterprise environments

This section provides some best practices and considerations for deploying and operating Amazon FSx for NetApp ONTAP in enterprise environments. These recommendations are based on the experiences of AWS Professional Services.

In addition to the recommendations in this guide, adhere to the following best practices:

- Best practices for working with Active Directory (FSx for ONTAP documentation)
- Data protection (FSx for ONTAP documentation)
- Security best practices in IAM (AWS Identity and Access Management (IAM) documentation)
- Best practices and implementation guide for NetApp ONTAP FlexGroup volumes (NetApp documentation)

Best practices for storage tiers and tiering policies

Storage tiers are the physical storage media for an Amazon FSx for NetApp ONTAP file system. The following storage tiers are available:

- The *SSD tier* is high-performance solid-state drive (SSD) storage designed for active data, and you choose the storage size for this tier.
- The *capacity pool tier* is fully elastic storage that is cost-optimized for infrequently accessed data. The SSD tier is significantly faster than the capacity pool tier. FSx for ONTAP SSD storage provides sub-millisecond file operational latencies, and the capacity pool tier provides tens of milliseconds of latency.

For more information about these tiers, see FSx for ONTAP storage tiers.

A *tiering policy*, which you configure at the volume level, determines if and when data that's stored in the SSD tier transitions to the capacity pool tier. FSx for ONTAP offers four different tiering policies: **Snapshot Only**, **Auto**, **All**, and **None**. For more information about each policy, see Tiering policies in the FSx for ONTAP documentation.

Consider the following recommendations when setting tiering policies for the volumes in your file share:

- HPC workloads should access data in the SSD tier to prevent performance bottlenecks. For volumes accessed by HPC workloads, we recommend setting the tiering policy to **None** or **Snapshot Only**.
- When migrating data to the file share, we recommend setting the target volume tiering policy to **All**. This reduces costs because all data migrates to the SSD tier and is then immediately moved to the capacity pool tier. In addition, if 98% or more of the SSD tier capacity is utilized, then writing to the tier is stopped. Setting the tiering policy to **All** prevents reaching this tiering threshold during the migration. After the migration is complete, you can change the tiering policy in order to balance

performance and costs. For more information, see Migrating file shares to Amazon FSx for NetApp ONTAP using AWS DataSync (AWS blog post).

Best practices for using the NetApp ONTAP maximum directory size

maxdirsize (NetApp documentation) is a NetApp ONTAP setting that determines the maximum number of files that can be stored in each directory. This setting applies to the volume, so all of the directories in a volume have the same `maxdirsize` setting. The default value is 320 MB, which allows you to store up to 4.3 million files in each directory.

You can increase the `maxdirsize` value to support larger directories. After the value has been increased, it cannot be decreased without recreating the directory. Because directories are loaded in memory, there is a tradeoff between the size of the directories and the performance of your file system. You can validate custom settings only through a test. NetApp recommends that you keep this value at its default. For more information, see Best practices and implementation guide for NetApp ONTAP FlexGroup volumes (NetApp documentation).

If you customize the `maxdirsize` setting, you can use the following formula to determine how many files can fit into a single folder.

```
max number of files in each directory = maxdirsize in MB × 53 × 0,25
```

Best practices for monitoring FSx for ONTAP file systems

Similar to other AWS services, FSx for ONTAP is integrated with Amazon CloudWatch. CloudWatch helps you monitor the metrics of your AWS resources in near real time. Metrics are available at the file system and volume levels, and *detailed monitoring* metrics for these resources help you analyze them with more granular reporting detail. For more information, see Monitoring with Amazon CloudWatch in the FSx for ONTAP documentation. Consider the following recommendations when monitoring FSx for ONTAP by using CloudWatch:

- We recommend that you use the `StorageUsed` file system metric so that you can filter monitoring results by storage tier.
- Use the `StorageCapacity` file system metric to configure a CloudWatch alarm that notifies you if more than 80% of the SSD tier capacity is utilized. This ensures that tiering functions properly for the volume, and it helps you maintain capacity for new data. For more information, see Tiering thresholds.

Best practices for choosing an Availability Zone deployment option

You can deploy Amazon FSx for NetApp ONTAP in a Single-AZ or Multi-AZ configuration. Each option provides different levels of availability and durability. For more information about these deployment options, see Availability and durability in the FSx for ONTAP documentation.

Multi-AZ deploys the FSx for ONTAP file system in an active-passive configuration. Therefore, all servers that connect to the file share use only the endpoint in the primary Availability Zone. The endpoint in

the secondary Availability Zone is for failover only, and it is not used to read or write unless the primary Availability Zone fails.

You cannot change the Availability Zone deployment option after you create the FSx for ONTAP file system. To change the Availability Zone configuration, you have to create a new file system and then migrate the data to the new file system.

However, even if you deployed a file share using the Single-AZ option, you can still access it from other Availability Zones. Your networking configuration, such as security groups and network access control list (network ACL) must allow the clients to connect to the file system endpoint. Using this approach, there is a charge for cross-AZ traffic in each direction (read and write). For more information, see Amazon FSx for NetApp ONTAP Pricing.

When choosing a deployment option, you must choose between the resiliency of the Multi-AZ configuration and the performance of the Single-AZ configuration. If practical for your use case, we recommend selecting Multi-AZ option because it provides high availability. However, the Single-AZ option can be more cost-effective and reduce latency. Consider the HPC workload and whether it can tolerate the additional latency.

FAQ

What does thin provisioned mean in regards to FSx for ONTAP volumes?

Generally speaking, *thin provisioned* means that the system uses virtualization technologies to show more resources as available than are actually provisioned. It is like the cash reserve in a bank—the amount of money physically kept in a bank is less than the total amount of the bank accounts. When you create a volume in Amazon FSx for NetApp ONTAP, the storage is not reserved in advance. Its size gradually increases as you add data to it. For more information, see FSx for ONTAP storage tiers.

What protocols are supported by FSx for ONTAP?

FSx for ONTAP file systems can be accessed through Network File System (NFS), Server Message Block (SMB), and Internet Small Computer System Interface (iSCSI) protocols. For more information, see Accessing data in the FSx for ONTAP documentation.

I'm using FSx for ONTAP in a Windows environment. Are there any prerequisites to enable integration with Active Directory?

Yes, you need to create a service account on the Active Directory domain. For more information, see Provisioning an Active Directory service account (p. 7) in this guide. You also need to ensure proper network connectivity. When setting up the file system, don't forget to specify the organizational unit (OU) that you want to join. For more information, see Prerequisites for joining an SVM to a self-managed Microsoft AD in the FSx for ONTAP documentation.

Can I change the volume tiering policy?

Yes, you can change the tiering policy at any time. For more information, see Setting a volume's tiering policy in the Amazon FSx documentation.

Tiering policy and write operations on the file system are not working, and metrics show >98% SSD storage tier utilization. What should I do?

All tiering functionality and write operations stop when the SSD storage tier is at or above 98% utilization. For more information, see Tiering thresholds in the Amazon FSx documentation. To resume

operations, increase the amount of SSD storage capacity. Consider changing the tiering policy to retain less data in the SSD tier. For more information, see Managing volume storage capacity and Setting a volume's tiering policy in the Amazon FSx documentation.

Does Multi-AZ deployment support active-active configuration?

No, Multi-AZ deployment is an active-passive configuration. For more information, see Best practices for choosing an Availability Zone deployment option (p. 12) in this guide.

Is the pricing the same for Single-AZ and Multi-AZ deployments of FSx for ONTAP?

No, Multi-AZ configuration costs about the double of the Single-AZ configuration. For Single-AZ deployments, there are changes for cross-AZ traffic. For more information, see Amazon FSx for NetApp ONTAP Pricing.

Resources

Amazon FSx for NetApp ONTAP documentation

- Storage tiers
- Supported clients
- Managing volumes
- Managing SMB shares
- Best practices for joining FSx for ONTAP SVMs to an Active Directory domain
- Volume data tiering and thresholds
- File system metrics for Amazon CloudWatch monitoring
- TieringPolicy (API reference)

Other AWS resources

- Choosing an AWS storage service
- Amazon FSx for NetApp ONTAP pricing
- Migrating file shares to Amazon FSx for NetApp ONTAP using AWS DataSync (AWS blog post)

NetApp resources

- NetApp ONTAP FlexGroup volumes: Best practices and implementation guide (NetApp PDF)
- What are junction paths (NetApp Knowledge Base)
- What is maxdirsize (NetApp Knowledge Base)

Document history

The following table describes significant changes to this guide. If you want to be notified about future updates, you can subscribe to an RSS feed.

Change	Description	Date
Initial publication (p. 17)	—	August 29, 2023

AWS Prescriptive Guidance glossary

The following are commonly used terms in strategies, guides, and patterns provided by AWS Prescriptive Guidance. To suggest entries, please use the **Provide feedback** link at the end of the glossary.

Migration terms

7 Rs

Seven common migration strategies for moving applications to the cloud. These strategies build upon the 5 Rs that Gartner identified in 2011 and consist of the following:

- Refactor/re-architect – Move an application and modify its architecture by taking full advantage of cloud-native features to improve agility, performance, and scalability. This typically involves porting the operating system and database. Example: Migrate your on-premises Oracle database to the Amazon Aurora PostgreSQL-Compatible Edition.
- Replatform (lift and reshape) – Move an application to the cloud, and introduce some level of optimization to take advantage of cloud capabilities. Example: Migrate your on-premises Oracle database to Amazon Relational Database Service (Amazon RDS) for Oracle in the AWS Cloud.
- Repurchase (drop and shop) – Switch to a different product, typically by moving from a traditional license to a SaaS model. Example: Migrate your customer relationship management (CRM) system to Salesforce.com.
- Rehost (lift and shift) – Move an application to the cloud without making any changes to take advantage of cloud capabilities. Example: Migrate your on-premises Oracle database to Oracle on an EC2 instance in the AWS Cloud.
- Relocate (hypervisor-level lift and shift) – Move infrastructure to the cloud without purchasing new hardware, rewriting applications, or modifying your existing operations. This migration scenario is specific to VMware Cloud on AWS, which supports virtual machine (VM) compatibility and workload portability between your on-premises environment and AWS. You can use the VMware Cloud Foundation technologies from your on-premises data centers when you migrate your infrastructure to VMware Cloud on AWS. Example: Relocate the hypervisor hosting your Oracle database to VMware Cloud on AWS.
- Retain (revisit) – Keep applications in your source environment. These might include applications that require major refactoring, and you want to postpone that work until a later time, and legacy applications that you want to retain, because there's no business justification for migrating them.
- Retire – Decommission or remove applications that are no longer needed in your source environment.

active-active migration

A database migration method in which the source and target databases are kept in sync (by using a bidirectional replication tool or dual write operations), and both databases handle transactions from connecting applications during migration. This method supports migration in small, controlled batches instead of requiring a one-time cutover. It's more flexible but requires more work than active-passive migration (p. 18).

active-passive migration

A database migration method in which in which the source and target databases are kept in sync, but only the source database handles transactions from connecting applications while data is replicated to the target database. The target database doesn't accept any transactions during migration.

application portfolio

A collection of detailed information about each application used by an organization, including the cost to build and maintain the application, and its business value. This information is key to the portfolio discovery and analysis process and helps identify and prioritize the applications to be migrated, modernized, and optimized.

artificial intelligence operations (AIOps)

The process of using machine learning techniques to solve operational problems, reduce operational incidents and human intervention, and increase service quality. For more information about how AIOps is used in the AWS migration strategy, see the operations integration guide.

AWS Cloud Adoption Framework (AWS CAF)

A framework of guidelines and best practices from AWS to help organizations develop an efficient and effective plan to move successfully to the cloud. AWS CAF organizes guidance into six focus areas called perspectives: business, people, governance, platform, security, and operations. The business, people, and governance perspectives focus on business skills and processes; the platform, security, and operations perspectives focus on technical skills and processes. For example, the people perspective targets stakeholders who handle human resources (HR), staffing functions, and people management. For this perspective, AWS CAF provides guidance for people development, training, and communications to help ready the organization for successful cloud adoption. For more information, see the AWS CAF website and the AWS CAF whitepaper.

AWS Workload Qualification Framework (AWS WQF)

A tool that evaluates database migration workloads, recommends migration strategies, and provides work estimates. AWS WQF is included with AWS Schema Conversion Tool (AWS SCT). It analyzes database schemas and code objects, application code, dependencies, and performance characteristics, and provides assessment reports.

business continuity planning (BCP)

A plan that addresses the potential impact of a disruptive event, such as a large-scale migration, on operations and enables a business to resume operations quickly.

change data capture (CDC)

The process of tracking changes to a data source, such as a database table, and recording metadata about the change. You can use CDC for various purposes, such as auditing or replicating changes in a target system to maintain synchronization.

Cloud Center of Excellence (CCoE)

A multi-disciplinary team that drives cloud adoption efforts across an organization, including developing cloud best practices, mobilizing resources, establishing migration timelines, and leading the organization through large-scale transformations. For more information, see the CCoE posts on the AWS Cloud Enterprise Strategy Blog.

cloud stages of adoption

The four phases that organizations typically go through when they migrate to the AWS Cloud:

- Project – Running a few cloud-related projects for proof of concept and learning purposes
- Foundation – Making foundational investments to scale your cloud adoption (e.g., creating a landing zone, defining a CCoE, establishing an operations model)
- Migration – Migrating individual applications
- Re-invention – Optimizing products and services, and innovating in the cloud

These stages were defined by Stephen Orban in the blog post The Journey Toward Cloud-First & the Stages of Adoption on the AWS Cloud Enterprise Strategy blog. For information about how they relate to the AWS migration strategy, see the migration readiness guide.

configuration management database (CMDB)

A repository that stores and manages information about a database and its IT environment, including both hardware and software components and their configurations. You typically use data from a CMDB in the portfolio discovery and analysis stage of migration.

epic

In agile methodologies, functional categories that help organize and prioritize your work. Epics provide a high-level description of requirements and implementation tasks. For example, AWS CAF security epics include identity and access management, detective controls, infrastructure security, data protection, and incident response. For more information about epics in the AWS migration strategy, see the program implementation guide.

flash-cut migration

A database migration method that uses continuous data replication through change data capture (CDC) (p. 19) to migrate data in the shortest time possible, instead of using a phased approach. The objective is to keep downtime to a minimum.

heterogeneous database migration

Migrating your source database to a target database that uses a different database engine (for example, Oracle to Amazon Aurora). Heterogeneous migration is typically part of a re-architecting effort, and converting the schema can be a complex task. AWS provides AWS SCT that helps with schema conversions.

homogeneous database migration

Migrating your source database to a target database that shares the same database engine (for example, Microsoft SQL Server to Amazon RDS for SQL Server). Homogeneous migration is typically part of a rehosting or replatforming effort. You can use native database utilities to migrate the schema.

hypercare period

Immediately following cutover, the period of time when a migration team manages and monitors the migrated applications in the cloud in order to address any issues. Typically, this period is 1–4 days in length. At the end of the hypercare period, the migration team typically transfers responsibility for the applications to the cloud operations team.

idle application

An application that has an average CPU and memory usage between 5 and 20 percent over a period of 90 days. In a migration project, it is common to retire these applications or retain them on premises.

incremental migration

A cutover strategy in which you migrate your application in small parts instead of performing a single, full cutover. For example, you might move only a few microservices or users to the new system initially. After you verify that everything is working properly, you can incrementally move additional microservices or users until you can decommission your legacy system. This strategy reduces the risks associated with large migrations.

IT information library (ITIL)

A set of best practices for delivering IT services and aligning these services with business requirements. ITIL provides the foundation for ITSM.

IT service management (ITSM)

Activities associated with designing, implementing, managing, and supporting IT services for an organization. For information about integrating cloud operations with ITSM tools, see the operations integration guide.

landing zone

A landing zone is a well-architected, multi-account AWS environment that is scalable and secure. This is a starting point from which your organizations can quickly launch and deploy workloads and applications with confidence in their security and infrastructure environment. For more information about landing zones, see Setting up a secure and scalable multi-account AWS environment.

large migration

A migration of 300 or more servers.

Migration Acceleration Program (MAP)

An AWS program that provides consulting support, training, and services to help organizations build a strong operational foundation for moving to the cloud, and to help offset the initial cost of migrations. MAP includes a migration methodology for executing legacy migrations in a methodical way and a set of tools to automate and accelerate common migration scenarios.

Migration Portfolio Assessment (MPA)

An online tool that provides information for validating the business case for migrating to the AWS Cloud. MPA provides detailed portfolio assessment (server right-sizing, pricing, TCO comparisons, migration cost analysis) as well as migration planning (application data analysis and data collection, application grouping, migration prioritization, and wave planning). The MPA tool (requires login) is available free of charge to all AWS consultants and APN Partner consultants.

Migration Readiness Assessment (MRA)

The process of gaining insights about an organization's cloud readiness status, identifying strengths and weaknesses, and building an action plan to close identified gaps, using the AWS CAF. For more information, see the migration readiness guide. MRA is the first phase of the AWS migration strategy.

migration at scale

The process of moving the majority of the application portfolio to the cloud in waves, with more applications moved at a faster rate in each wave. This phase uses the best practices and lessons learned from the earlier phases to implement a *migration factory* of teams, tools, and processes to streamline the migration of workloads through automation and agile delivery. This is the third phase of the AWS migration strategy.

migration factory

Cross-functional teams that streamline the migration of workloads through automated, agile approaches. Migration factory teams typically include operations, business analysts and owners, migration engineers, developers, and DevOps professionals working in sprints. Between 20 and 50 percent of an enterprise application portfolio consists of repeated patterns that can be optimized by a factory approach. For more information, see the discussion of migration factories and the Cloud Migration Factory guide in this content set.

migration metadata

The information about the application and server that is needed to complete the migration. Each migration pattern requires a different set of migration metadata. Examples of migration metadata include the target subnet, security group, and AWS account.

migration pattern

A repeatable migration task that details the migration strategy, the migration destination, and the migration application or service used. Example: Rehost migration to Amazon EC2 with AWS Application Migration Service.

migration strategy

The approach used to migrate a workload to the AWS Cloud. For more information, see the 7 Rs (p. 18) entry in this glossary and see Mobilize your organization to accelerate large-scale migrations.

offline migration

A migration method in which the source workload is taken down during the migration process. This method involves extended downtime and is typically used for small, non-critical workloads.

online migration

A migration method in which the source workload is copied to the target system without being taken offline. Applications that are connected to the workload can continue to function during the migration. This method involves zero to minimal downtime and is typically used for critical production workloads.

operational-level agreement (OLA)

An agreement that clarifies what functional IT groups promise to deliver to each other, to support a service-level agreement (SLA).

operations integration (OI)

The process of modernizing operations in the cloud, which involves readiness planning, automation, and integration. For more information, see the operations integration guide.

organizational change management (OCM)

A framework for managing major, disruptive business transformations from a people, culture, and leadership perspective. OCM helps organizations prepare for, and transition to, new systems and strategies by accelerating change adoption, addressing transitional issues, and driving cultural and organizational changes. In the AWS migration strategy, this framework is called *people acceleration*, because of the speed of change required in cloud adoption projects. For more information, see the OCM guide.

playbook

A set of predefined steps that capture the work associated with migrations, such as delivering core operations functions in the cloud. A playbook can take the form of scripts, automated runbooks, or a summary of processes or steps required to operate your modernized environment.

portfolio assessment

A process of discovering, analyzing, and prioritizing the application portfolio in order to plan the migration. For more information, see Evaluating migration readiness.

responsible, accountable, consulted, informed (RACI) matrix

A matrix that defines and assigns roles and responsibilities in a project. For example, you can create a RACI to define security control ownership or to identify roles and responsibilities for specific tasks in a migration project.

runbook

A set of manual or automated procedures required to perform a specific task. These are typically built to streamline repetitive operations or procedures with high error rates.

service-level agreement (SLA)

An agreement that clarifies what an IT team promises to deliver to their customers, such as service uptime and performance.

task list

A tool that is used to track progress through a runbook. A task list contains an overview of the runbook and a list of general tasks to be completed. For each general task, it includes the estimated amount of time required, the owner, and the progress.

workstream

Functional groups in a migration project that are responsible for a specific set of tasks. Each workstream is independent but supports the other workstreams in the project. For example, the

portfolio workstream is responsible for prioritizing applications, wave planning, and collecting migration metadata. The portfolio workstream delivers these assets to the migration workstream, which then migrates the servers and applications.

zombie application

An application that has an average CPU and memory usage below 5 percent. In a migration project, it is common to retire these applications.

Storage and backup terms

authoritative data source

A location where you store the primary version of data, which is considered to be the most reliable source of information. You can copy data from the authoritative data source to other locations for the purposes of processing or modifying the data, such as anonymizing, redacting, or pseudonymizing it.

cold data

Data that is rarely accessed and is typically historical. When querying this kind of data, slow queries are typically acceptable. Moving this data to lower-performing and less expensive storage tiers or classes can reduce costs.

disaster

An event that prevents a workload or system from fulfilling its business objectives in its primary deployed location. These events can be natural disasters, technical failures, or the result of human actions, such as unintentional misconfiguration or a malware attack.

disaster recovery (DR)

The strategy and process you use to minimize downtime and data loss caused by a disaster (p. 23). For more information, see Disaster Recovery of Workloads on AWS: Recovery in the Cloud in the AWS Well-Architected Framework.

high availability (HA)

The ability of a workload to operate continuously, without intervention, in the event of challenges or disasters. HA systems are designed to automatically fail over, consistently deliver high-quality performance, and handle different loads and failures with minimal performance impact.

hot data

Data that is frequently accessed, such as real-time data or recent translational data. This data typically requires a high-performance storage tier or class to provide fast query responses.

recovery point objective (RPO)

The maximum acceptable amount of time since the last data recovery point. This determines what is considered an acceptable loss of data between the last recovery point and the interruption of service.

recovery time objective (RTO)

The maximum acceptable delay between the interruption of service and restoration of service.

warm data

Data that is infrequently accessed. When querying this kind of data, moderately slow queries are typically acceptable.