



FlexCache 卷管理

ONTAP 9

NetApp
February 12, 2026

目录

FlexCache 卷管理	1
了解ONTAP FlexCache卷	1
视频	1
ONTAP FlexCache卷支持和不支持的功能	2
FlexCache卷和初始卷之间的ONTAP版本支持	3
支持的协议	3
支持的功能	3
ONTAP FlexCache卷规模估算准则	7
创建ONTAP FlexCache卷	8
FlexCache回写	13
了解ONTAP FlexCache回写	13
ONTAP FlexCache回写准则	14
ONTAP FlexCache回写架构	15
ONTAP FlexCache回写用例	19
ONTAP FlexCache回写前提条件	21
ONTAP FlexCache回写互操作性	22
启用并管理ONTAP FlexCache回写	23
有关ONTAP FlexCache回写的常见问题	27
FlexCache 二元性	28
关于 FlexCache 二元性的常见问题	28
启用 S3 对 NAS FlexCache 卷的访问	29
管理FlexCache卷	35
了解如何审核ONTAP FlexCache卷	35
从初始卷同步ONTAP FlexCache卷的属性	36
更新ONTAP FlexCache关系的配置	37
在ONTAP FlexCache卷上启用文件访问时间更新	37
在ONTAP FlexCache卷上启用全局文件锁定	38
预填充ONTAP FlexCache卷	39
删除ONTAP FlexCache关系	41
用于修复热点的FlexCache	41
在使用ONTAP FlexCache卷的高性能计算工作负载中修补热抽查	41
构建ONTAP FlexCache热点修复解决方案	42
确定ONTAP FlexCache密度	45
确定ONTAP的SVM间或SVM内HDFA选项	47
配置HDFA和ONTAP数据LUN	49
配置客户端以分布ONTAP NAS连接	51

FlexCache 卷管理

了解ONTAP FlexCache卷

NetApp FlexCache技术可加快读取密集型工作负载的数据访问速度、降低WAN延迟并降低WAN带宽成本、尤其是在客户端需要重复访问相同数据的情况下。创建FlexCache卷时、您需要为已有(初始)卷创建远程缓存、该卷仅包含初始卷的活动访问数据(热数据)。

当FlexCache卷收到对其包含的热数据的读取请求时、其响应速度会快于初始卷、因为数据不需要传输到客户端。如果FlexCache卷收到读取不常读取的数据(冷数据)的请求、则它会从初始卷检索所需的数据、然后在处理客户端请求之前存储这些数据。然后、该数据的后续读取请求将直接从FlexCache卷提供服务。在第一次请求后、数据不再需要在网络中传输、也不再需要从负载过重的系统提供。例如、假设您的集群在一个用于访问频繁请求的数据的单一访问点遇到瓶颈。您可以在集群中使用FlexCache卷为热数据提供多个挂载点、从而减少瓶颈并提高性能。另一个示例是、假设您需要减少从多个集群访问的卷的网络流量。您可以使用FlexCache卷在网络中的各个集群之间分发来自初始卷的热数据。这样可以为用户提供更近的访问点、从而减少WAN流量。

您还可以使用FlexCache技术提高云和混合云环境中的性能。FlexCache卷可以将数据从内部数据中心缓存到云、从而帮助您将工作负载过渡到混合云。您还可以使用FlexCache卷将数据从一个云提供商缓存到另一个云提供商或同一个云提供商的两个区域之间、从而消除云孤岛。

从ONTAP 9.10.1开始，您可以["启用全局文件锁定"](#)所有FlexCache卷。全局文件锁定可防止用户访问已由其他用户打开的文件。然后、对初始卷的更新会同时分发到所有FlexCache卷。

从ONTAP 9.9.1开始、FlexCache卷会维护一个未找到文件的列表。这样，当客户端搜索不存在的文件时，就无需向源发送多个调用，从而有助于减少网络流量。

其他的列表["FlexCache卷及其原始卷支持的功能"](#)还提供了按ONTAP版本列出的受支持协议。

您可以在[中](#)了解有关ONTAP FlexCache技术架构的更多信息["TR-4743 : FlexCache in ONTAP"](#)。

视频

[FlexCache 如何降低全局数据的 WAN 延迟和读取时间](#)

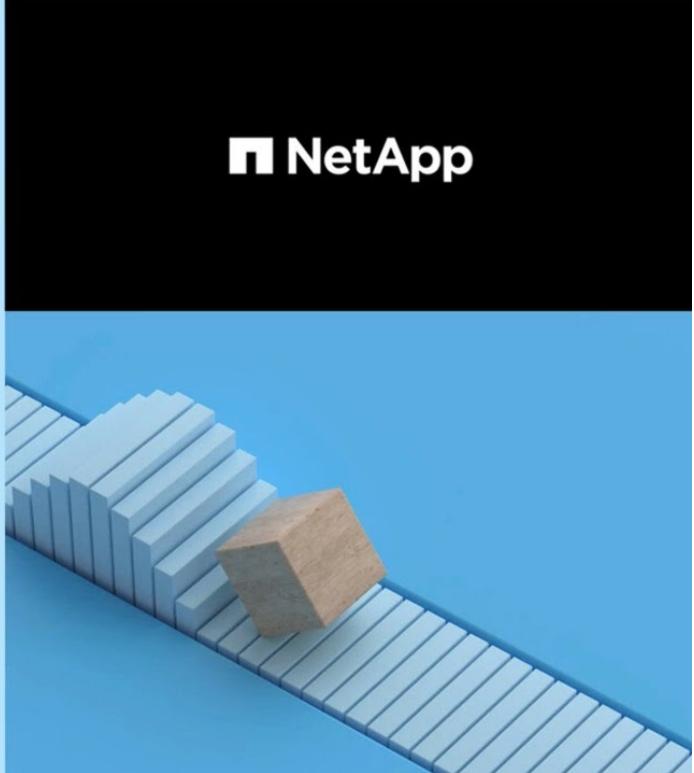
ONTAP FlexCache

Data Access Where You Need It

Use Case

© 2020 NetApp, Inc. All rights reserved.

NetApp



了解 ONTAP FlexCache 的性能优势！

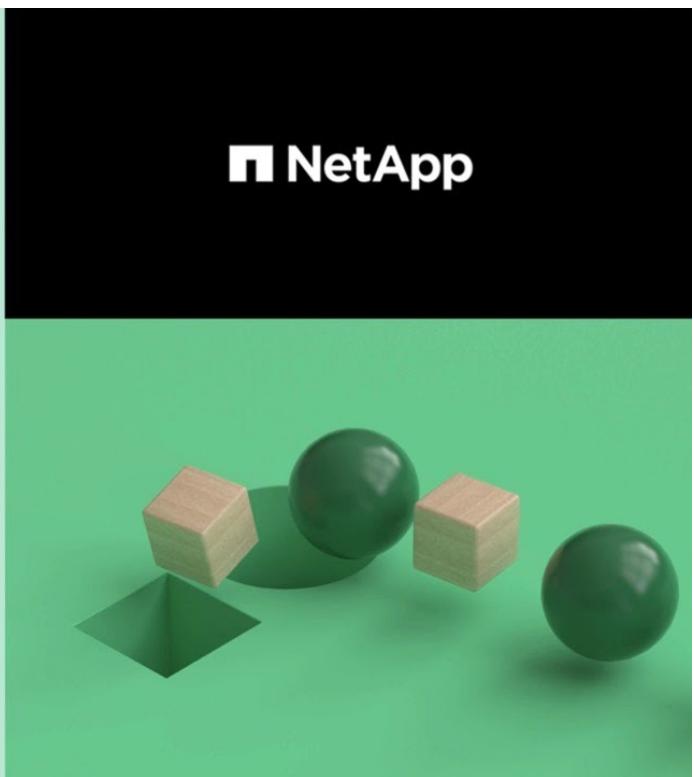
ONTAP FlexCache

Data Access Where You Need It

Tech Clip

© 2020 NetApp, Inc. All rights reserved.

NetApp



ONTAP FlexCache卷支持和不支持的功能

从ONTAP 9.5开始、您可以配置FlexCache卷。支持将FlexVol卷用作初始卷、并将FlexGroup卷用作FlexCache卷。从ONTAP 9.7开始、支持将FlexVol卷和FlexGroup卷作

为初始卷。初始卷和FlexCache卷支持的功能和协议会有所不同。



只要缓存卷和初始卷都运行在受支持的ONTAP版本上、它们就可以互操作。请注意、只有当缓存和源服务器运行的版本至少为引入了支持的ONTAP版本或更高的ONTAP版本时、才支持这些功能。

FlexCache卷和初始卷之间的ONTAP版本支持

初始卷和缓存卷之间支持的建议ONTAP版本不超过四个或四个以上。例如、如果缓存使用的是ONTAP 9.14.1、则原始卷可以使用的最早版本是ONTAP 9.10.1。

支持的协议

协议	是否在原始卷上受支持?	是否在 FlexCache 卷上受支持?
NFSv3	是的。	是的。
NFSv4	是的。 要使用NFSv4.x协议访问缓存卷、初始集群和缓存集群都必须使用ONTAP 9.10.1或更高版本。初始集群和FlexCache集群可以具有不同的ONTAP版本、但两者都应为ONTAP 9.10.1及更高版本、例如、初始集群可以具有ONTAP 9.10.1、而缓存可以具有ONTAP 9.11.1。	是的。 从ONTAP 9.10.1开始支持。 要使用NFSv4.x协议访问缓存卷、初始集群和缓存集群都必须使用ONTAP 9.10.1或更高版本。初始集群和FlexCache集群可以具有不同的ONTAP版本、但两者都应为ONTAP 9.10.1及更高版本、例如、初始集群可以具有ONTAP 9.10.1、而缓存可以具有ONTAP 9.11.1。
NFSv4.2	是的。	否
SMB	是的。	是的。 从 ONTAP 9.8 开始支持。

支持的功能

功能	是否在原始卷上受支持?	是否在 FlexCache 卷上受支持?
自主勒索软件保护	是的。 从ONTAP 9.10.1开始支持FlexVol初始卷、从ONTAP 9.13.1开始支持FlexGroup初始卷。请参见 " 自主勒索软件保护使用情形和注意事项 "。	否

防病毒	是的。 从ONTAP 9.7开始支持。	不适用 如果您在初始位置配置防病毒扫描、则不需要在缓存上执行此操作。无论写入源是什么、源站防病毒扫描都会在提交写入前检测到受病毒感染的文件。有关对FlexCache使用防病毒扫描的详细信息、请参见 " 采用ONTAP的FlexCache技术报告 "。
审核	是的。 从ONTAP 9.7开始支持。 您可以使用本机ONTAP审核功能审核FlexCache关系中的NFS文件访问事件。 有关详细信息，请参见 审核FlexCache 卷的注意事项	是的。 从ONTAP 9.7开始支持。 您可以使用本机ONTAP审核功能审核FlexCache关系中的NFS文件访问事件。 有关详细信息，请参见 审核FlexCache 卷的注意事项
Cloud Volumes ONTAP	是的。 从ONTAP 9.6开始支持	是的。 从ONTAP 9.6开始支持
数据缩减	是的。 从ONTAP 9.6开始支持	是的。 从ONTAP 9.7 开始支持
压缩	是的。 从ONTAP 9.6开始支持	是的。 从ONTAP 9.6开始支持
重复数据删除	是的。	是的。 从ONTAP 9.6 开始， FlexCache 卷支持实时重复数据删除。从ONTAP 9.7 开始， FlexCache 卷支持跨卷重复数据删除。
FabricPool	是的。	是的。 从ONTAP 9.7 开始支持  您可以创建一个FlexCache卷作为启用了FabricPool分层功能的源卷的缓存，但FlexCache卷本身不能分层。

FlexCache 灾难恢复	是的。	是的。 从ONTAP 9.1.1开始支持，仅支持NFSv3协议。FlexCache卷必须位于不同的SVM中或不同的集群中。
FlexGroup 卷	是的。 从ONTAP 9.7开始支持	是的。
FlexVol 卷	是的。	否
fpolicy	是的。 从ONTAP 9.7开始支持	是的。 从ONTAP 9.7开始支持NFS。 从ONTAP 9.14.1开始支持SMB。
MetroCluster 配置	是的。 从ONTAP 9.7开始支持	是的。 从ONTAP 9.7开始支持
Microsoft 卸载数据传输（Offloaded Data Transfer，ODX）	是的。	否
NetApp 聚合加密（NAE）	是的。 从ONTAP 9.6开始支持	是的。 从ONTAP 9.6开始支持
NetApp 卷加密（NVE）	是的。 从ONTAP 9.6开始支持	是的。 从ONTAP 9.6开始支持
ONTAP S3 NAS存储分段	是的。 从ONTAP 9.12.1开始支持	是的。 从ONTAP 9.18.1开始支持
QoS	是的。	是的。  FlexCache卷不支持文件级QoS。
qtree	是的。 从ONTAP 9.6开始，您可以创建和修改qtrees。在源上创建的qtrees可在缓存上访问。	否

配额	是的。 从ONTAP 9.6开始、支持对用户、组和qtrees在FlexCache原始卷上强制实施配额。	否 使用FlexCache writearound模式(默认模式)时、缓存上的写入将转发到初始卷。配额在源站强制实施。  从ONTAP 9.6开始，FlexCache卷支持远程配额(rquota)。
SMB更改通知	是的。	是的。 从ONTAP 9.14.1开始、缓存支持SMB更改通知。
SnapLock 卷	否	否
SnapMirror异步关系*	是的。	否
	<ul style="list-style-type: none"> • FlexCache来源： • 您可以从原始FlexVol创建FlexCache卷 • 您可以从原始FlexGroup创建FlexCache卷 • 在SnapMirror关系中、您可以从原始主卷创建FlexCache 卷。 • 从ONTAP 9.8开始，SnapMirror二级卷可以是FlexCache原始卷。SnapMirror二级卷必须处于空闲状态、并且没有活动的SnapMirror更新；否则、FlexCache创建将失败。 	SnapMirror同步关系
否	否	SnapRestore
是的。	否	快照
是的。	否	SVM IP 配置

是的。	否	存储级别访问防护（SLAG）
从ONTAP 9.5 开始支持。SVM DR 关系的主 SVM 可以具有源卷；但是，如果您运行的是ONTAP 9.5 之前的版本，当 SVM DR 关系断开时，必须使用新的源卷重新创建FlexCache关系。	FlexCache 卷可以位于主 SVM 中，但不能位于辅 SVM 中。主 SVM 中的任何 FlexCache 卷不会作为 SVM DR 关系的一部分进行复制。	
从ONTAP 9.18.1 开始，当源 SVM 发生故障转移时，DR 站点的缓存会自动切换到源。无需手动恢复步骤。		
了解如何创建FlexCache卷。		
否	否	精简配置
是的。	是的。 从 ONTAP 9.7 开始支持	卷克隆
是的。 从 ONTAP 9.6 开始，支持克隆初始卷和初始卷中的文件。	否	卷移动
是的。	是（仅适用于卷成分卷） ONTAP 9.6及更高版本支持移动FlexCache卷的卷成分卷。	卷重新托管
否	否	用于阵列集成的vStorage API (VAAI)



在 ONTAP 9.5 之前的版本中，原始 FlexVol 卷只能为在运行 7- 模式 Data ONTAP 8.2.x 的系统上创建的 FlexCache 卷提供数据。从 ONTAP 9.5 开始，原始 FlexVol 卷还可以为 ONTAP 9 系统上的 FlexCache 卷提供数据。有关从 7- 模式 FlexCache 迁移到 ONTAP 9 FlexCache 的信息，请参见["NetApp 技术报告 4743：《ONTAP 中的 FlexCache》"](#)。

ONTAP FlexCache 卷规模估算准则

在开始配置 FlexCache 卷之前，您必须了解卷的限制。

FlexVol 卷的大小限制适用于原始卷。FlexCache 卷的大小可以小于或等于原始卷。对于 FlexCache 卷的大小，最佳做法是至少为原始卷大小的 10%。

此外，您还必须了解 FlexCache 卷的以下附加限制：

limit	ONTAP 9.8及 更高版本	ONTAP 9.7	ONTAP 9.6 - 9.5
可以从原始卷创建的最大 FlexCache 卷数	100	10.	10.
每个节点的建议最大原始卷数	100	100	10.
每个节点的建议最大 FlexCache 卷数	100	100	10.
每个节点的 FlexCache 卷中的建议最大 FlexGroup 成分卷数	800	800	40.
每个节点每个 FlexCache 卷的最大成分卷数	32.	32.	32.

相关信息

- ["NetApp 互操作性"](#)

创建ONTAP FlexCache卷

您可以在同一个 ONTAP 群集中创建 FlexCache 卷，以提高访问热对象时的性能。如果您在不同位置拥有数据中心，则可以在远程 ONTAP 群集上创建 FlexCache 卷，以加速数据访问。

关于此任务

- 从 ONTAP 9.18.1 开始，您可以通过在创建 FlexCache 卷时将 `'-is-s3-enabled` 选项设置为 `true` 来在卷上启用 NAS S3 存储桶访问。此选项默认处于禁用状态。
- 从 ONTAP 9.18.1 开始，FlexCache 支持为具有 SVM-DR 关系的 SVM 的源卷创建缓存卷。

如果您运行的是 ONTAP 9.18.1 或更高版本，则存储管理员必须先将缓存 SVM 与 SVM-DR 关系中的主源 SVM 和辅助源 SVM 建立对等关系，然后再创建属于 SVM-DR 关系的源卷的缓存卷。

- 从 ONTAP 9.14.0 开始、您可以从加密源创建未加密的 FlexCache 卷。
- 从 ONTAP 9.7 开始，FlexVol 卷和 FlexGroup 卷都支持作为源卷。
- 从 ONTAP 9.5 开始、FlexCache 支持将 FlexVol 卷用作初始卷、并将 FlexGroup 卷用作 FlexCache 卷。

开始之前

- 您必须运行 ONTAP 9.5 或更高版本。
- 如果您运行的是 ONTAP 9.6 或更早版本，则必须 ["添加 FlexCache 许可证"](#)。

ONTAP 9.7 或更高版本不需要 FlexCache 许可证。从 ONTAP 9.7 开始、ONTAP 随附了 FlexCache 功能、不再需要许可证或激活。



HA 对使用时 ["加密 SAS 或 NVMe 驱动器 \(SED, NSE, FIPS\)"](#)，您必须按照主题中的说明进行操作 ["将 FIPS 驱动器或 SED 恢复到未受保护的模式"](#) 初始化系统之前 HA 对中的所有驱动器（启动选项 4 或 9）。如果不这样做，则在重新利用驱动器时，可能会导致未来数据丢失。

示例 1. 步骤

System Manager

1. 如果 FlexCache 卷位于与原始卷不同的ONTAP 群集上, 请创建群集对等关系:
 - a. 在本地集群中, 单击 * 保护 > 概述 *。
 - b. 展开*Intercluster Settings*, 单击*Add Network Interfaces*并为集群添加集群间网络接口。
在远程集群上重复此步骤。
 - c. 在远程集群中, 单击 * 保护 > 概述 *。单击 :: "Cluster Peers"(集群对等方)部分, 然后单击*Generate Pas密切 相关*。
 - d. 复制生成的密码短语并将其粘贴到本地集群中。
 - e. 在本地集群中的集群对等方下, 单击 * 对等集群 * 并对本地和远程集群建立对等关系。
2. 创建SVM对等关系:

在Storage VM对等方下、依次单击 :: 和*对等Storage VM*以对等Storage VM。

3. 选择 * 存储 > 卷 *。
4. 选择 * 添加 *。
5. 选择*更多选项*, 然后选择*为远程卷添加缓存*。



如果您运行的是ONTAP 9.8或更高版本, 并且要禁用QoS或选择自定义QoS策略, 请单击*更多选项*, 然后在*存储和优化*下选择*性能服务级别*。

命令行界面

1. 如果要创建的 FlexCache 卷位于其他集群中, 请创建集群对等关系:

- a. 在目标集群上、创建与数据保护源集群的对等关系:

```
cluster peer create -generate-passphrase -offer-expiration
MM/DD/YYYY HH:MM:SS|1...7days|1...168hours -peer-addrs
<peer_LIF_IPs> -initial-allowed-vserver-peers <svm_name>,..|*
-ipspace <ipspace_name>
```

从ONTAP 9.6开始, 默认情况下, 在创建集群对等关系时会启用TLS 加密。支持使用TLS 加密在初始卷和FlexCache 卷之间进行集群间通信。如果需要, 您还可以为集群对等关系禁用TLS 加密。

```
cluster02::> cluster peer create -generate-passphrase -offer  
-expiration 2days -initial-allowed-vserver-peers *  
  
        Passphrase: UCa+6lRVICXeL/gq1WrK7ShR  
        Expiration Time: 6/7/2017 08:16:10 EST  
        Initial Allowed Vserver Peers: *  
        Intercluster LIF IP: 192.140.112.101  
        Peer Cluster Name: Clus_7ShR (temporary generated)
```

Warning: make a note of the passphrase - it cannot be displayed again.

- a. 在源集群上，将源集群身份验证到目标集群：

```
cluster peer create -peer-addrs <peer_LIF_IPs> -ipspace <ipspace>
```

```
cluster01::> cluster peer create -peer-addrs  
192.140.112.101,192.140.112.102
```

Notice: Use a generated passphrase or choose a passphrase of 8 or more characters.

To ensure the authenticity of the peering relationship, use a phrase or sequence of characters that would be hard to guess.

Enter the passphrase:

Confirm the passphrase:

Clusters cluster02 and cluster01 are peered.

2. 如果FlexCache卷与原始卷位于不同的SVM中、请使用创建SVM对等关系 `flexcache` 作为应用程序：

- a. 如果SVM位于不同集群中、请为对等SVM创建SVM权限：

```
vserver peer permission create -peer-cluster <cluster_name>  
-vserver <svm-name> -applications flexcache
```

以下示例说明了如何创建适用于所有本地 SVM 的 SVM 对等权限：

```
cluster1::> vserver peer permission create -peer-cluster cluster2  
-vserver "*" -applications flexcache  
  
Warning: This Vserver peer permission applies to all local Vservers.  
After that no explicit  
"vserver peer accept" command required for Vserver peer relationship  
creation request  
from peer cluster "cluster2" with any of the local Vservers. Do you  
want to continue? {y|n}: y
```

a. 创建SVM对等关系：

```
vserver peer create -vserver <local_SVM> -peer-vserver  
<remote_SVM> -peer-cluster <cluster_name> -applications flexcache
```

3. 创建 FlexCache 卷：

```
volume flexcache create -vserver <cache_svm> -volume  
<cache_vol_name> -auto-provision-as flexgroup -size <vol_size>  
-origin-vserver <origin_svm> -origin-volume <origin_vol_name> -is-s3  
-enabled true|false
```

以下示例将创建一个 FlexCache 卷并自动选择要配置的现有聚合：

```
cluster1::> volume flexcache create -vserver vs_1 -volume fc1 -auto  
-provision-as flexgroup -origin-volume vol_1 -size 160MB -origin  
-vserver vs_1  
[Job 443] Job succeeded: Successful
```

以下示例将创建一个 FlexCache 卷并设置接合路径：

```
cluster1::> volume flexcache create -vserver vs34 -volume fc4 -aggr  
-list aggr34,aggr43 -origin-volume origin1 -size 400m -junction-path  
/fc4  
[Job 903] Job succeeded: Successful
```

以下示例在 FlexCache 卷上启用 S3 访问：

```
cluster1::> volume flexcache create -vserver vs3 -volume  
cache_vs3_vol33 -origin-volume vol33 -origin-vserver vs3 -junction  
-path /cache_vs3_vol33 -is-s3-enabled true
```

4. 验证 FlexCache 卷和原始卷的 FlexCache 关系。

a. 查看集群中的FlexCache关系：

```
volume flexcache show
```

```
cluster1::> volume flexcache show  
Vserver Volume      Size      Origin-Vserver Origin-Volume  
Origin-Cluster  
-----  
-----  
vs_1   fc1          160MB    vs_1        vol_1  
cluster1
```

b. 查看初始集群中的所有FlexCache关系：

```
volume flexcache origin show-caches
```

```
cluster::> volume flexcache origin show-caches  
Origin-Vserver Origin-Volume Cache-Vserver Cache-Volume  
Cache-Cluster  
-----  
-----  
vs0       ovoll       vs1        cfg1  
clusA  
vs0       ovoll       vs2        cfg2  
clusB  
vs_1     vol_1       vs_1        fc1  
cluster1
```

结果

已成功创建 FlexCache 卷。客户端可以使用 FlexCache 卷的接合路径挂载卷。

相关信息

["集群和 SVM 对等"](#)

FlexCache回写

了解ONTAP FlexCache回写

FlexCache回写是在ONTAP 9.151中推出的一种备用操作模式、用于在缓存中写入数据。回写允许将写入提交到缓存中的稳定存储、并向客户端确认、而无需等待数据传输到源站。数据会异步转储回源站。因此、可以构建一个全球分布式文件系统、使特定工作负载和环境能够以接近本地的速度执行写入、从而显著提高性能。



ONTAP 9.12.1引入了一项回写功能、可作为公共预览。这称为回写版本1 (wbv1)、不应视为与ONTAP 9.15.1中的回写相同、后者称为回写版本2 (wbv2)。

回写与绕写

自FlexCache在ONTAP 9.5中推出以来、它一直是一种可读写缓存；但是、它在绕写模式下运行。缓存中的写入操作会发送到初始卷、以提交到稳定存储。初始服务器成功将写入提交到稳定存储后、它会确认写入缓存。然后、缓存会向客户端确认写入。这样、每次写入都会导致在缓存和源站之间遍历网络。FlexCache回写将更改此设置。



升级到ONTAP 9.15.1后、您可以将传统的绕写缓存转换为回写缓存、并在必要时返回到绕写缓存。但是、如果出现问题、这可能会使读取诊断日志变得更加困难。

	绕写	回写
ONTAP 版本	9.6多个	9.15.1+
用例	读取繁重的工作负载	写入量大的工作负载
数据提交时间	来源	缓存
客户体验	类似于WAN	类似于LAN
限制	每个原产地100个	每个原产地10个
"Cap Theorem"	对分区可用并容错	可用且一致

FlexCache回写术语

了解使用FlexCache回写的关键概念和术语。

期限	定义
dirty 数据	已提交到缓存中的稳定存储、但尚未转储到源站的数据。
专用锁定委派(XLD)	按文件授予缓存的协议级锁定权限。此授权允许缓存向客户端分发专用写入锁定、而无需联系源站。
共享锁定委派(SLD)	按文件授予缓存的协议级锁定权限。此授权允许缓存向客户端分发共享读取锁定、而无需联系源站。
回写	一种FlexCache操作模式、其中、对缓存的写入会提交到该缓存中的稳定存储、并立即向客户端确认。数据会异步写入源站。

期限	定义
绕写式	一种FlexCache操作模式、其中、对缓存的写入会转发到初始卷以提交到稳定存储。提交后、初始服务器将确认对缓存的写入、而缓存将确认对客户端的写入。
脏数据记录系统(DDRS)	一种专有机制、可在启用了回写功能的缓存中按文件跟踪脏数据。
原产地	包含所有FlexCache缓存卷的源数据的FlexGroup或FlexVol。它是唯一的数据来源、可协调锁定、并确保100%的数据一致性、货币性和一致性。
缓存	FlexGroup是FlexCache初始卷的稀疏缓存卷。

一致、当前和一致

FlexCache是NetApp的解决方案、可随时随地提供正确的数据。FlexCache始终保持100%一致、当前和一致：

- *一致：*无论访问数据的位置、数据都是相同的。
- *current：*数据始终是最新的。
- *一致：*数据正确/未损坏。

ONTAP FlexCache回写准则

FlexCache写回涉及源服务器和缓存服务器之间的许多复杂交互。为了获得最佳性能，您应该确保您的环境遵循以下准则。这些指南基于内容创建时可用的最新主要ONTAP版本（ONTAP 9.17.1）。

作为最佳实践、请在非生产环境中测试生产工作负载。如果您在这些准则之外实施FlexCache回写、则这一点更为重要。

以下准则已在NetApp内部经过充分测试。*仅适用于您待在他们里面。否则、可能会发生意外行为。

- ONTAP 9.17.1P1 引入了对FlexCache回写功能的重大增强。强烈建议您在源集群和缓存集群上，在 9.17.1P1 版本之后运行当前推荐版本。如果您无法运行 9.17.1 代码线，建议使用最新的 P 版本 9.16.1。
◦ ONTAP 9.15.1 尚未包含FlexCache回写所需的全部修复和改进，因此不建议用于生产工作负载。
- 在当前迭代过程中、应为FlexCache回写缓存配置一个成分卷来容纳整个FlexCache卷。多成分卷FlexCaches可能会导致从缓存中意外删除数据。
- 测试针对小于 100GB 的文件进行，缓存和源之间的 WAN 往返时间不超过 200ms。超出这些限制的任何工作负载都可能导致意想不到的性能特征。
- 写入SMB备用数据流会导致主文件从缓存中逐出。需要先将主文件的所有脏数据转储到原始文件、然后才能对该文件执行任何其他操作。备用数据流也会转发到源。
- 重命名文件会导致文件从缓存中逐出。需要先将文件的所有脏数据转储到源、然后才能对该文件执行任何其他操作。
- 此时、只能对启用了回写的FlexCache卷上的文件更改或设置以下属性：
 - 时间戳
 - 模式位
 - NT ACL

- 所有者
- 组
- Size

更改或设置的任何其他属性都会转发到原始属性、这可能会导致从缓存中逐出文件。如果您需要在缓存中更改或设置其他属性、请要求您的客户团队创建PVR。

- 在初始卷创建的快照会导致从与该初始卷关联的每个已启用回写的缓存中调用所有未处理的脏数据。如果正在进行大量回写活动、则可能需要多次重试此操作、因为清除这些脏文件可能需要一段时间。
- 启用回写功能的FlexCache卷不支持 SMB 机会锁（Oplocks）写入。
- 原产地必须保持在80%满量以下。如果初始卷中没有至少20%的剩余空间、则不会为缓存卷授予独占锁定委派。在这种情况下、对启用了回写的缓存的调用将转发到源站。这有助于防止源站空间不足、从而导致在启用了回写的缓存中孤立脏数据。
- 低带宽和/或有损的集群间网络会对FlexCache写回性能产生显著的负面影响。虽然没有具体的带宽要求，因为它很大程度上取决于您的工作负载，但强烈建议您确保缓存和源之间的集群间链路的健康状况。

ONTAP FlexCache回写架构

FlexCache在设计时充分考虑了一致性、包括两种写入操作模式：回写和绕写。传统的绕写操作模式和ONTAP 9.15.1中引入的新回写操作模式均可确保访问的数据始终保持100%一致、最新且一致。

以下概念详细介绍了FlexCache回写的操作方式。

委派

锁定委派和数据委派有助于FlexCache保持回写和绕写缓存数据的一致性、一致性和最新性。源服务器会编排这两个委派。

锁定委派

锁定委派是指初始服务器按文件授予缓存的协议级锁定授权、用于根据需要向客户端发出协议锁定。其中包括[专用锁定委派\(XD\)](#) 和 [共享锁定委派\(SLD\)](#)。

XLD和回写

为了确保ONTAP不必协调冲突写入、将向客户端请求写入文件的缓存授予XLD。重要的是、任何文件在任何时候只能存在一个XLD、这意味着一个文件一次不会有多个写入程序。

当写入文件的请求进入启用了回写的缓存时、将执行以下步骤：

1. 缓存会检查是否已为请求的文件创建了XLD。如果是这样、只要另一个客户端未向缓存中的文件写入数据、它就会为客户端授予写入锁定。如果缓存没有用于请求文件的XLD、则它将从原始位置请求一个。这是一个经过集群间网络的专有调用。
2. 从缓存收到XLD-请求后、初始服务器将检查另一个缓存中的文件是否存在未完成的XLD-。如果是这样、它将调用该文件的XLD、从而触发从该缓存转储到原始缓存的任何 [脏数据](#)。
3. 一旦将该缓存中的脏数据转储回来并提交到源站的稳定存储、源站将为请求缓存授予文件的XLD。
4. 收到文件的XLD后、缓存会将锁定授予客户端、写入将开始。

序列图中介绍了涵盖其中某些步骤的高级序列图 [\[write-back-sequence-diagram\]](#)。

从客户端的角度来看、所有锁定都将像写入标准FlexVol或FlexGroup一样工作、在请求写入锁定时可能会有一个小延迟。

在当前迭代中、如果启用了回写的缓存保存了文件的XLD、则ONTAP将阻止其他缓存对该文件的*任何*访问、包括 READ 操作。



每个原始成分卷的XLD限制为170。

数据委派

数据委派是指初始卷为缓存提供的每个文件的保证、保证为该文件缓存的数据是最新的。只要缓存具有文件的数据委派、它就可以向客户端提供该文件的缓存数据、而无需联系源站。如果缓存没有文件的数据委派、它必须联系源站以接收客户端请求的数据。

在回写模式下、如果在另一个缓存或原始位置为文件创建了XLD、则文件的数据委派将被撤消。这样可以有效地将文件与所有其他缓存中的客户端和源服务器隔开、即使是读取也是如此。这是一种权衡、必须确保旧数据永远不会被访问。

启用了回写的缓存上的读取操作通常类似于绕写缓存上的读取操作。在启用了绕写和回写的缓存中、如果请求的文件在启用了回写的缓存(而不是发出读取的缓存)具有排他写锁定、则初始性能可能会 READ 受到影响。必须撤消XLD、并且必须先将脏数据提交到源站、然后才能处理其他缓存上的读取。

跟踪脏数据

从缓存回写源站的操作是异步进行的。这意味着脏数据不会立即写入源站。ONTAP使用脏数据记录系统来跟踪每个文件的脏数据。每个脏数据记录(脏数据记录、DDR)代表特定文件的大约20 MB脏数据。当正在写入文件时、ONTAP将在填充两个DDRR并写入第三个DDR后开始将脏数据转储回。这样、写入期间缓存中会保留大约40 MB的脏数据。对于有状态协议(NFSv4.x、SMB)、在关闭文件后、剩余的40 MB数据将转储回原始卷。对于无状态协议(NFSv3)、当在其他缓存上请求访问该文件时、或者在该文件闲置两分钟或更长时间(最长为五分钟)后、将转储40 MB的数据。有关计时器触发或空间触发的异常数据转储的详细信息，请参见 [\[缓存擦除器\]](#)。

除了CDR和擦除器之外、某些前端NAS操作还会触发文件的所有脏数据的转储：

- SETATTR
 - `sETATTR`s只能在缓存中处理修改mtime、atime和/或ctime的操作，避免了WAN的影响。
- CLOSE
- OPEN 在另一个缓存中
- READ 在另一个缓存中
- REaddir 在另一个缓存中
- REaddirplus 在另一个缓存中
- WRITE 在另一个缓存中

已断开连接模式

如果文件的XLD保存在绕写缓存中、并且该缓存与原始缓存断开连接、则仍允许在其他缓存和原始缓存中读取该文件。如果XLD由启用了回写的缓存保留、则此行为会有所不同。在这种情况下、如果缓存断开连接、则对文件

的读取将在任何地方挂起。这有助于确保保持100%的一致性、一致性和一致性。允许在绕写模式下进行读取，因为保证源站拥有已向客户端写入确认的所有可用数据。在断开连接期间的回写模式下、初始卷无法保证已启用回写的缓存中写入并确认的所有数据都在断开连接之前写入到初始卷。

如果文件的具有XLD的缓存长时间断开连接、系统管理员可以手动在原始位置撤消XLD。这样、该文件的IO便可 在运行正常的缓存和原始卷上恢复。



手动恢复XLD将导致断开连接的缓存中文件的任何脏数据丢失。只有在缓存和源站之间发生灾难性中断时、才应手动删除XLD。

缓存擦除器

ONTAP中存在一些擦除了程序、它们会根据特定事件运行、例如计时器即将过期或违反空间阈值。擦除程序会 对正在擦除的文件进行独占锁定、从而有效冻结该文件的IO、直到擦除完成为止。

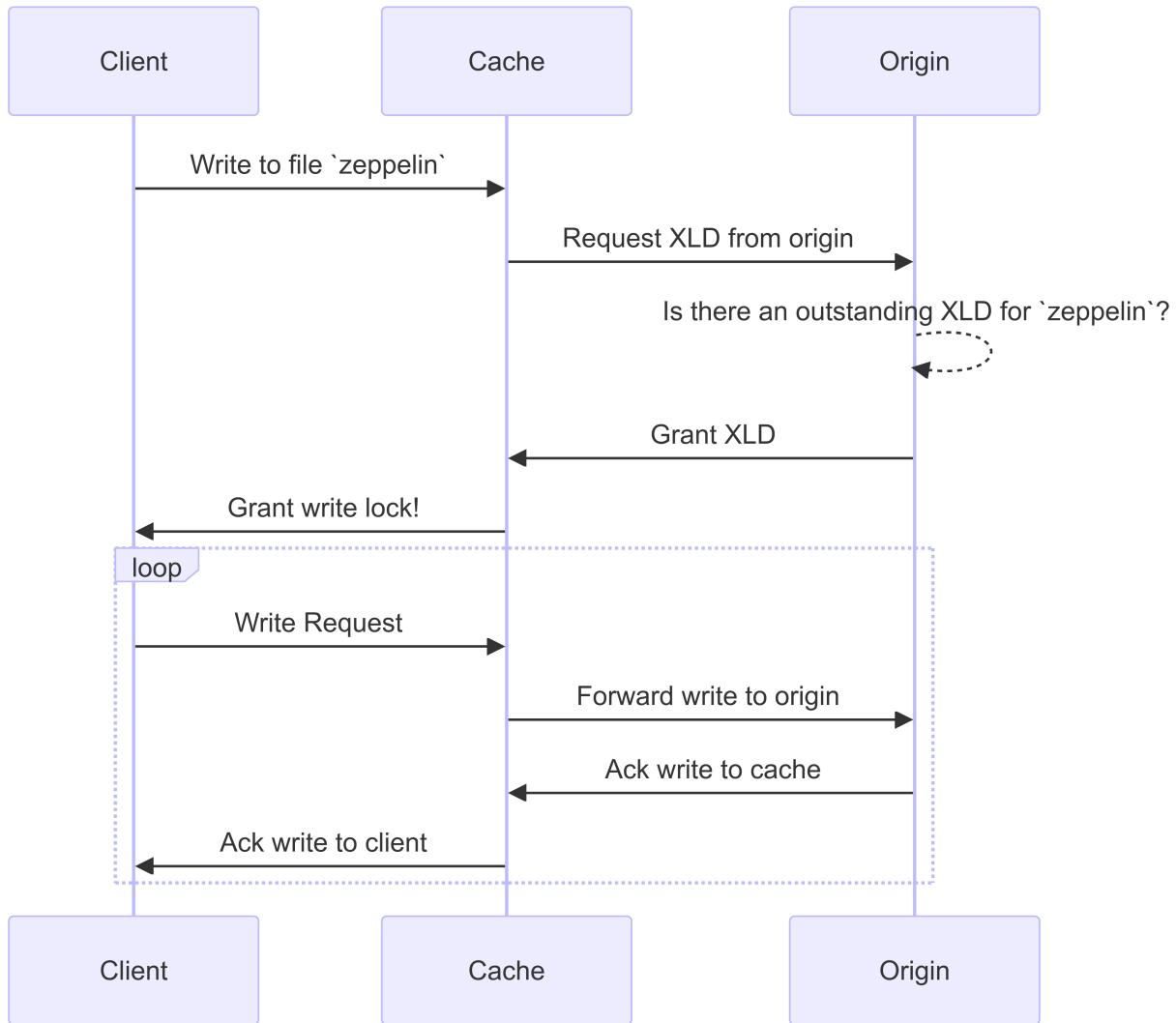
擦尘器包括：

- 缓存上的*基于时间的清理程序：*此清理程序每五分钟启动一次，并擦除任何未经修改的文件两分钟。如果 此文件的任何脏数据仍在缓存中、则此文件的IO将被静置并触发回写。回写完成后、IO将恢复。
- 与缓存中基于mtime的清理程序非常相似，该程序也每五分钟运行一次。但是、它会将未修改的任何文件进 行15分钟的反转、从而重新调用索引节点的委派。此擦除程序不会启动任何回写。
- **RW**基于限制的清理程序在源站：ONTAP可监控每个源站成分卷分发的RW锁定委派的数量。如果此数字超 过170、则ONTAP将开始擦除最近最少使用的写入锁定委派(Least Recently Used、LRU)。
- *缓存上基于空间的擦除程序：*如果FlexCache卷已满90%、则会擦除缓存、并按LRU逐出。
- *基于空间的擦除程序：*如果FlexCache初始卷已满90%、则会擦除缓存、并按LRU逐出。

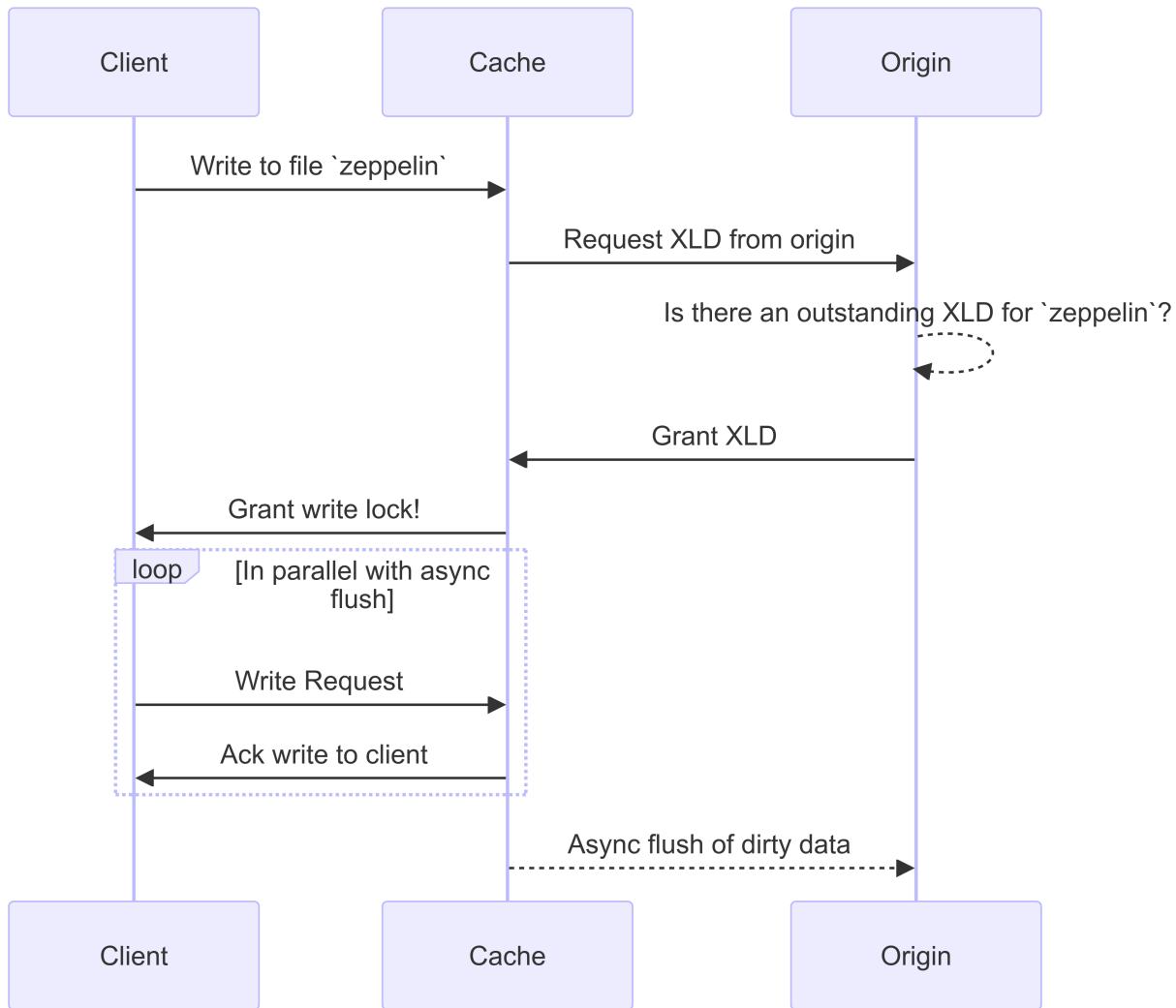
序列图

这些顺序图显示了绕写模式与回写模式之间的写入确定差异。

绕写



回写



ONTAP FlexCache回写用例

这些写入配置文件最适合启用了回写的FlexCache。您应测试工作负载、以确定回写或绕写是否可提供最佳性能。



回写不能替代绕写。虽然回写是针对写入量大的工作负载设计的、但对于许多工作负载来说、绕写仍然是更好的选择。

目标工作负载

文件大小

文件大小的重要性低于在和调用文件之间发出的写入次数 OPEN CLOSE。小型文件本身调用次数较少 WRITE、因此不太适合回写。在和调用之间、大型文件可能会写入更多数据 OPEN CLOSE、但这并不能保证。

有关最大文件大小的最新建议、请参见 "[FlexCache回写准则](#)" 页面。

写入大小

从客户端写入时、除了写入调用之外、还会涉及其他修改NAS调用。其中包括但不限于：

- CREATE
- OPEN
- CLOSE
- SETATTR
- SET_INFO

SETATTR 和 `SET_INFO` 设置、`atime`、`ctime`、`owner`、`group` 或的 `size` 呼叫 `mtime` 在缓存中处理。其余这些调用必须在源站进行处理，并对在启用了回写功能的缓存中为正在运行的文件累积的任何脏数据触发回写操作。文件的IO将暂停，直到回写完成。

了解这些呼叫必须遍历WAN有助于确定适合回写的工作负载。通常，`OPEN`、`CLOSE` 在与调用之间可以执行的写入越多、而不发出上述其他调用之一、性能增益回写效果就越好。

写后读取

写入后读取工作负载以往在FlexCache中的性能较差。这是由于9.15.1之前的绕写操作模式造成的。`WRITE` 对文件的调用必须在初始位置提交、后续 `READ` 调用必须将数据提取回缓存。这会导致这两种操作都会对WAN造成影响。因此、对于采用绕写模式的FlexCache、不建议在写入后读取工作负载。随着9.15.1中引入回写功能、数据现在会提交到缓存中、并可立即从缓存中读取、从而消除了WAN方面的负面影响。如果您的工作负载在FlexCache卷上包括后读写、则应将缓存配置为在回写模式下运行。



如果写入后读取是工作负载的关键部分、则应将缓存配置为在回写模式下运行。

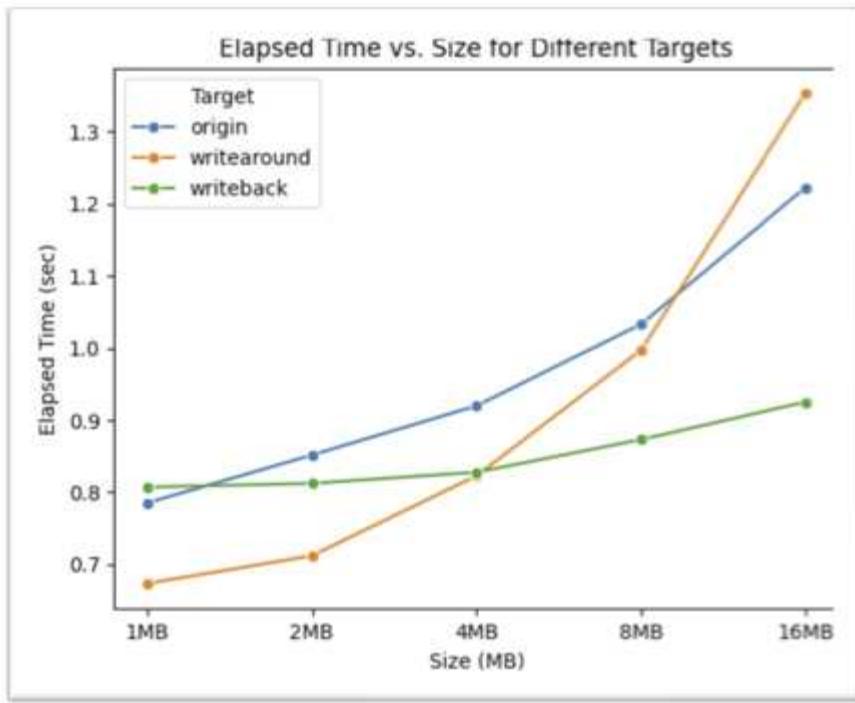
写后写入

当文件在缓存中累积脏数据时、缓存会将数据异步写入源站。这自然会导致客户端关闭文件时、脏数据仍在等待转储回源。如果刚刚关闭但仍包含脏数据的文件再次打开或写入、则写入操作将暂停、直到所有脏数据都转储到源站为止。

延迟注意事项

当FlexCache在回写模式下运行时、随着延迟的增加、它对NAS客户端更有利。但是、在一定程度上、回写开销会超过在低延迟环境中获得的优势。在某些NetApp测试中、回写优势从缓存与初始服务器之间的最小延迟8毫秒开始。此延迟因工作负载而异、因此请务必进行测试、以了解工作负载的返回点。

下图显示了NetApp实验室测试中回写的返回点。x轴是文件大小、y轴是经过的时间。此测试使用NFSv3、挂载的 `rsize` 和 `wsize` 为256 KB、WAN延迟为64毫秒。执行此测试时、会对缓存和初始卷使用一个小型ONTAP Select实例、并执行单线程写入操作。结果可能有所不同。



回写不应用于集群内缓存。如果初始缓存和缓存位于同一集群中，则会发生集群内缓存。

ONTAP FlexCache回写前提条件

在以回写模式部署FlexCache之前、请确保满足这些性能、软件、许可和系统配置要求。

CPU和内存

强烈建议每个原始集群节点至少具有 128GB 的 RAM 和 20 个 CPU 来吸收启用回写功能的缓存发起的回写消息。这相当于A400或更高版本。如果初始集群用作多个启用了回写的FlexCaches的初始集群，则需要更多CPU和RAM。



为工作负载使用大小不足的源站可能会对启用了回写的缓存或源站的性能产生深远影响。

ONTAP 版本

- 原点*Must*运行的是ONTAP 9.15.1.或更高版本。
- 任何需要在回写模式*Must*下运行的缓存集群都运行ONTAP 9.15.1.或更高版本。
- 任何不需要在回写模式下运行的缓存集群都可以运行任何普遍支持的ONTAP版本。

许可

您购买的ONTAP附带了FlexCache (包括回写模式操作)。不需要额外的许可证。

对等

- 初始集群和缓存集群必须为 "[已建立集群对等](#)"
- 原始集群和缓存集群上的服务器虚拟机(SVM)必须 "[已为Vserver对等](#)" 具有FlexCache选项。



您无需将缓存集群与另一个缓存集群建立对等关系。也无需将缓存SVM与另一个缓存SVM建立对等关系。

ONTAP FlexCache回写互操作性

了解在回写模式下部署FlexCache时的这些互操作性注意事项。

ONTAP 版本

要使用回写操作模式，缓存和原始服务器*必须*都运行ONTAP 9.15.1.或更高版本。



不需要启用回写缓存的集群可以运行早期版本的ONTAP、但该集群只能在绕写模式下运行。

您的环境可以混合使用多种ONTAP版本。

集群	ONTAP 版本	是否支持回写？
原产地	ONTAP 9.15.1.	不适用†
集群1	ONTAP 9.15.1.	是的。
集群2	ONTAP 9.14.1	否

集群	ONTAP 版本	是否支持回写？
原产地	ONTAP 9.14.1	不适用†
集群1	ONTAP 9.15.1.	否
集群2	ONTAP 9.15.1.	否

†原点 不是缓存、因此回写或绕写支持均不适用。



在中 [example2-table]，两个集群都不能启用回写模式，因为初始集群未运行ONTAP 9.151或更高版本，这是一项严格要求。

客户端互操作性

ONTAP通常支持的任何客户端均可访问FlexCache卷、而不管该卷是在绕写模式还是回写模式下运行。有关受支持客户端的最新列表，请参阅NetApp的 "[互操作性表](#)"。

尽管客户端版本无关紧要、但客户端必须足够新才能支持NFS3、NFSv4.0、NFSv4.1、SMB2.x或SMB3.xSMB1和NFSv2是已弃用的协议、不受支持。

回写和绕写

如中所示 [example1-table]，在回写模式下运行的FlexCache可以与在绕写模式下运行的缓存共存。建议将绕写与回写与特定工作负载进行比较。



如果工作负载在回写和绕写之间的性能相同、请使用绕写。

ONTAP功能互操作性

有关FlexCache功能互操作性的最新列表，请参阅 "[FlexCache卷支持和不支持的功能](#)"。

启用并管理ONTAP FlexCache回写

从ONTAP 9.15.1开始，您可以在FlexCache卷上启用FlexCache回写模式，以便为具有大量写入工作负载的边缘计算环境和缓存提供更好的性能。您还可以根据需要确定是在FlexCache卷上启用回写还是在卷上禁用回写。

在缓存卷上启用回写后，写入请求将发送到本地缓存，而不是原始卷。

开始之前

您必须处于高级权限模式。

创建启用了回写的新FlexCache卷

步骤

您可以使用ONTAP系统管理器或ONTAP命令行界面创建启用了回写的新FlexCache卷。

System Manager

1. 如果FlexCache卷与初始卷位于不同的集群上、请创建集群对等关系：
 - a. 在本地集群上，单击“保护>概述”。
 - b. 展开“Intercluster Settings”，单击“Add Network Interfaces”，然后将集群间接口添加到集群中。
在远程集群上重复此操作。
 - c. 在远程集群上、单击“保护>概述”。单击：“Cluster Peers”(集群对等方)部分，然后单击“Generate Pas密钥相关”。
 - d. 复制生成的密码短语并将其粘贴到本地集群中。
 - e. 在本地集群上的“Cluster Peers”下、单击“对等集群”并为本地集群和远程集群建立对等关系。
2. 如果FlexCache卷与原始卷位于不同的集群上、请创建SVM对等关系：
在“Storage VM对等方”下、依次单击 和“对等Storage VM”以建立Storage VM对等关系。
如果FlexCache卷位于同一集群上、则不能使用System Manager创建SVM对等关系。
 3. 选择“存储 > 卷”。
 4. 选择“添加”。
 5. 选择“更多选项”，然后选择“为远程卷添加缓存”。
 6. 选择“启用FlexCache回写”。

命令行界面

1. 如果要创建的 FlexCache 卷位于其他集群中，请创建集群对等关系：

- a. 在目标集群上、创建与数据保护源集群的对等关系：

```
cluster peer create -generate-passphrase -offer-expiration
MM/DD/YYYY HH:MM:SS|1...7days|1...168hours -peer-addrs
<peer_LIF_IPs> -initial-allowed-vserver-peers <svm_name>,..|*
-ipspace <ipspace_name>
```

从 ONTAP 9.6 开始，默认情况下，在创建集群对等关系时会启用 TLS 加密。支持使用 TLS 加密在初始卷和 FlexCache 卷之间进行集群间通信。如果需要，您还可以为集群对等关系禁用 TLS 加密。

```
cluster02::> cluster peer create -generate-passphrase -offer  
-expiration 2days -initial-allowed-vserver-peers *  
  
        Passphrase: UCa+6lRVICXeL/gq1WrK7ShR  
        Expiration Time: 6/7/2017 08:16:10 EST  
        Initial Allowed Vserver Peers: *  
        Intercluster LIF IP: 192.140.112.101  
        Peer Cluster Name: Clus_7ShR (temporary generated)  
  
Warning: make a note of the passphrase - it cannot be displayed  
again.
```

- a. 在源集群上，将源集群身份验证到目标集群：

```
cluster peer create -peer-addrs <peer_LIF_IPs> -ipspace <ipspace>
```

```
cluster01::> cluster peer create -peer-addrs  
192.140.112.101,192.140.112.102
```

Notice: Use a generated passphrase or choose a passphrase of 8 or more characters.

To ensure the authenticity of the peering relationship, use a phrase or sequence of characters that would be hard to guess.

Enter the passphrase:

Confirm the passphrase:

Clusters cluster02 and cluster01 are peered.

2. 如果FlexCache卷与原始卷位于不同的SVM中、请使用创建SVM对等关系 `flexcache` 作为应用程序：

- a. 如果SVM位于不同集群中、请为对等SVM创建SVM权限：

```
vserver peer permission create -peer-cluster <cluster_name>  
-vserver <svm-name> -applications flexcache
```

以下示例说明了如何创建适用于所有本地 SVM 的 SVM 对等权限：

```
cluster1::> vserver peer permission create -peer-cluster cluster2  
-vserver "*" -applications flexcache  
  
Warning: This Vserver peer permission applies to all local Vservers.  
After that no explicit  
"vserver peer accept" command required for Vserver peer relationship  
creation request  
from peer cluster "cluster2" with any of the local Vservers. Do you  
want to continue? {y|n}: y
```

a. 创建SVM对等关系：

```
vserver peer create -vserver <local_SVM> -peer-vserver  
<remote_SVM> -peer-cluster <cluster_name> -applications flexcache
```

3. 创建启用了回写的FlexCache卷：

```
volume flexcache create -vserver <cache_vserver_name> -volume  
<cache_flexgroup_name> -aggr-list <list_of_aggregates> -origin  
-volume <origin flexgroup> -origin-vserver <origin_vserver name>  
-junction-path <junction_path> -is-writeback-enabled true
```

在现有FlexCache卷上启用FlexCache回写

您可以使用ONTAP系统管理器或ONTAP命令行界面在现有FlexCache卷上启用FlexCache回写。

System Manager

1. 选择*存储>卷*并选择现有FlexCache卷。
2. 在卷的"概述"页面上，单击右上角的*Edit*。
3. 在*Edit Volumes*窗口中，选择*Enable FlexCache write-back*。

命令行界面

1. 在现有FlexCache卷上启用回写：

```
volume flexcache config modify -volume <cache_flexgroup_name> -is  
-writeback-enabled true
```

检查是否已启用**FlexCache**回写

步骤

您可以使用System Manager或ONTAP命令行界面来确定是否已启用FlexCache回写。

System Manager

1. 选择“存储>卷”并选择一个卷。
2. 在卷“Overview”中，找到“FlexCache详细信息”，并检查FlexCache卷上的FlexCache回写是否设置为“Enabled”。

命令行界面

1. 检查是否已启用FlexCache回写：

```
volume flexcache config show -volume <cache_flexgroup_name> -fields  
is-writeback-enabled
```

在**FlexCache**卷上禁用回写

在删除FlexCache卷之前、您需要禁用FlexCache回写。

步骤

您可以使用System Manager或ONTAP命令行界面禁用FlexCache回写。

System Manager

1. 选择“存储>卷”并选择已启用FlexCache回写的现有FlexCache卷。
2. 在卷的“概述”页面上，单击右上角的“Edit”。
3. 在“Edit Volumes”窗口中，取消选择“Enable FlexCache write-back”。

命令行界面

1. 禁用回写：

```
volume flexcache config modify -volume <cache_vol_name> -is  
-writeback-enabled false
```

有关**ONTAP FlexCache**回写的常见问题

此常见问题解答有助于您快速找到问题的答案。

我想使用回写。我需要运行哪个版本的**ONTAP**？

缓存和初始服务器都必须运行9.15.1 9.151或更高版本。强烈建议您运行最新的P版本。工程部门正在不断提高

启用了回写的缓存的性能和功能。

访问源站的客户端是否会影响访问启用了回写的缓存的客户端？

是。源服务器与任何缓存具有相同的数据访问权限。如果对某个文件执行的操作要求将该文件从缓存中逐出、或者撤消锁定/数据委派、则缓存中的客户端可能会在访问该文件时出现延迟。

是否可以将**QoS**应用于启用了回写的**FlexCaches**？

是。每个缓存和源站都可以应用独立的QoS策略。这不会对任何回写启动的集群间流量产生直接影响。通过QoS限制启用了回写的缓存中的前端流量、您可以间接降低集群间回写流量的速度。

启用了回写的**FlexCaches**是否支持多协议**NAS**？

是。启用了回写的FlexCaches完全支持多协议。目前、在绕写或回写模式下运行的FlexCache不支持NFSv4.2和S3。

启用了回写的**FlexCaches**是否支持**SMB**备用数据流？

支持SMB备用数据流(ADS)、但不能通过回写来加速。对ADS的写入会转发到源站、从而会降低WAN延迟。写入操作还会从缓存中清除ADS所属的主文件。

创建缓存后、是否可以在绕写模式和回写模式之间切换缓存？

是。您只需切换 `is-writeback-enabled`` 链接中的标志：`./vs-writback/vs. FlexCache FlexCache-writeback-enable-task.html [`flexcache modify command]`。

对于缓存和源服务器之间的集群间链路，我应该注意哪些带宽方面的因素？

是FlexCache 的写回功能高度依赖于缓存和源之间的集群间链路。低带宽和/或丢包网络会对性能产生显著的负面影响。没有具体的带宽要求，因为它很大程度上取决于您的工作负载。

FlexCache 二元性

关于 FlexCache 二元性的常见问题

此常见问题解答回答了有关 ONTAP 9.18.1 中引入的 FlexCache 二元性的常见问题。

常见问题解答

什么是"二元性"？

Duality 支持使用文件 (NAS) 和对象 (S3) 协议统一访问相同的数据。在 ONTAP 9.12.1 中引入时不支持 FlexCache，duality 在 ONTAP 9.18.1 中得到扩展以包括 FlexCache 卷，允许 S3 协议访问 FlexCache 卷中缓存的 NAS 文件。

在 FlexCache S3 存储桶上支持哪些 S3 操作？

标准 S3 NAS 存储区上支持的 S3 操作在 FlexCache S3 NAS 存储区上受支持，但 COPY 操作除外。有关标准 S3 NAS 存储区不支持的操作的最新列表，请访问 "[互操作性文档](#)"。

我可以在具有 FlexCache 二元性的回写模式下使用 FlexCache 吗？

不可以。如果在 FlexCache 卷上创建了 FlexCache S3 NAS 存储桶，则该 FlexCache 卷*必须*处于写入绕过模式。如果尝试在回写模式的 FlexCache 卷上创建 FlexCache S3 NAS 存储桶，则操作将失败。

由于硬件限制，我无法将其中一个集群升级到 ONTAP 9.18.1。如果只有缓存集群运行 ONTAP 9.18.1，我的集群中的双重性是否仍然有效？

否。缓存集群和原始集群都必须具有最低有效集群版本 9.18.1。如果尝试在缓存集群上创建一个 FlexCache S3 NAS 存储桶，该集群与运行早于 9.18.1 的 ONTAP 版本的原始集群对等，则操作将失败。

我有 **MetroCluster** 配置。我可以使用 **FlexCache** 二元性吗？

不，MetroCluster 配置中不支持 FlexCache 的双重性。

我可以审核 **S3** 对 **FlexCache S3 NAS** 存储桶中文件的访问权限吗？

S3 审核由 NAS 审核功能 FlexCache 卷使用提供。有关 FlexCache 卷的 NAS 审核的详细信息，请参见 "[详细了解 FlexCache 审核](#)"。

如果缓存集群与原始集群断开连接，我应该期待什么？

如果缓存集群与原始集群断开连接，则对 FlexCache S3 NAS 存储桶的 S3 请求将失败并出现 `503 Service Unavailable` 错误。

我可以将多部分 **S3** 操作与 **FlexCache** 双重性一起使用吗？

要使多部分 S3 操作正常工作，底层 FlexCache 卷必须将 granular-data 字段设置为 'advanced'。此字段设置为源卷设置的任何值。

FlexCache 双重性是否支持 **HTTP** 和 **HTTPS** 访问？

是。默认情况下，HTTPS 是必需的。如果需要，您可以将 S3 服务配置为允许 HTTP 访问。

启用 **S3** 对 **NAS FlexCache** 卷的访问

从 ONTAP 9.18.1 开始，您可以启用 S3 对 NAS FlexCache 卷的访问，也称为"双重性"。这允许客户端使用 S3 协议访问存储在 FlexCache 卷中的数据，以及 NFS 和 SMB 等传统 NAS 协议。您可以使用以下信息设置 FlexCache 双重性。

前提条件

在开始之前，您必须确保完成以下先决条件：

- 确保在 SVM 上许可并配置了 S3 协议和所需的 NAS 协议（NFS、SMB 或两者）。
- 验证是否已配置 DNS 和任何其他必需服务。
- 集群和 SVM 已建立对等关系
- FlexCache 卷创建
- 已创建 Data-lif



有关 FlexCache 二元性的更全面的文档，请参见 "[ONTAP S3 多协议支持](#)"。

步骤 1：创建并签署证书

要启用对 FlexCache 卷的 S3 访问，需要为托管 FlexCache 卷的 SVM 安装证书。此示例使用自签名证书，但在生产环境中，应使用由受信任的证书颁发机构 (CA) 签名的证书。

1. 创建 SVM 根 CA：

```
security certificate create -vserver <svm> -type root-ca -common-name  
<arbitrary_name>
```

2. 生成证书签名请求:

```
security certificate generate-csr -common-name <dns_name_of_data_lif>  
-dns-name <dns_name_of_data_lif> -ipaddr <data_lif_ip>
```

输出示例:

```
-----BEGIN CERTIFICATE REQUEST-----  
MIICzjCCAbYCAQAwHzEdMBsGA1UEAxMUY2FjaGUxZy1kYXRhLm5hcy5sYWIwggEi  
MA0GCSqGSIb3DQEBAQUAA4IBDwAwggEKAoIBAQCuJk07508Uh329cHI6x+BaRS2  
w5wrqvzoYlidXtYmdCH3m1DDprBiAyfIwBC0/iU3Xd5NpB7nc1wK1CI2VEkrXGUg  
...  
vMIGN351+FgzLQ4X5lKfoMXCV70NqIakxzEmkTIUDKv7n9EVZ4b5DTT1rL03X/nK  
+Bim2y2y180PaFB3NauZHTnIIzIc8zCp2IEqmFWyMDcdBjP9KS0+jNm4QhuXiM8F  
D7gm3g/O70qa5OxbAEa15o4Nb0l95U0T0rwqTaSzFG0XQnK2PmA1OIwS5ET35p3Z  
dLU=  
-----END CERTIFICATE REQUEST-----
```

私钥示例:

```
-----BEGIN PRIVATE KEY-----  
MIIEvAIBADANBgkqhkiG9w0BAQEFAASCBKYwggSiAgEAAoIBAQCuJk07508Uh32  
9cHI6x+BaRS2w5wrqvzoYlidXtYmdCH3m1DDprBiAyfIwBC0/iU3Xd5NpB7nc1wK  
1CI2VEkrXGUgwBtx1K4IlrCTB829Q1aLGAQXVyWnzhQc4tS5PW/DsQ8t7o1Z9zEI  
...  
rXGEDDaqp7jQGNXUGlbxO3zcBil1/A9Hc6oalNECgYBKwe3PeZamiwhIHLY9ph7w  
dJffCshsPalMuAp2OuKIANNa916fT9y5kf9tIbskT+t5Dth8bmV9pwe8UzaK5eC4  
Svxm19jHT5Qql0DaZVUmMXFKyKoqPDdfvcDk2Eb5gMfIIb0a3TPC/jqqpDn9BzuH  
TO02fuRvRR/G/HUz2yRd+A==  
-----END PRIVATE KEY-----
```



保留证书请求和私钥的副本，以备将来参考。

3. 签署证书:

```
`root-ca` 是您在 <><anchor1-step, 创建 SVM 根 CA>> 中创建的。
```

```
certificate sign -ca <svm_root_ca> -ca-serial <svm_root_ca_sn> -expire  
-days 364 -format PEM -vserver <svm>
```

4. 粘贴在 [生成证书签名请求](#) 中生成的证书签名请求 (CSR)。

示例

```
-----BEGIN CERTIFICATE REQUEST-----  
MIICzjCCAbYCAQAwHzEdMBsGA1UEAxMUY2FjaGUxZy1kYXRhLm5hcy5sYWIwggEi  
MA0GCSqGSIb3DQEBAQUAA4IBDwAwggEKAoIBAQCuJk075O8Uh329cHI6x+BaRS2  
w5wrqvzoYlidXtYmdCH3m1DDprBiAyfIwBC0/iU3Xd5NpB7nc1wK1CI2VEkrXGUg  
...  
vMIGN351+FgzLQ4X51KfoMXCV70NqIakxzEmkTIUDKv7n9EVZ4b5DTTlrl03X/nK  
+Bim2y2y180PaFB3NauZHTnIIzIc8zCp2IEqmFWyMDcdBjP9KS0+jNm4QhuXiM8F  
D7gm3g/O70qa5OxbAEa15o4Nb0195U0T0rwqTaSzFG0XQnK2PmA1OIwS5ET35p3Z  
dLU=  
-----END CERTIFICATE REQUEST-----
```

此操作会将签名证书打印到控制台，类似于以下示例。

签名证书示例：

```
-----BEGIN CERTIFICATE-----  
MIIDdzCCA1+gAwIBAgIIGHolbgv5DPowDQYJKoZIhvcNAQELBQAwLjEfMB0GA1UE  
AxMWY2FjaGUtMTY0Zy1zdm0tcm9vdC1jYTELMakGA1UEBhMCVVMwHhcNMjUxMTIx  
MjIxNTU4WhcNMjYxMTIwMjIxNTU4WjAfMR0wGwYDVQQDExRjYWNoZTFnLWRhdGEu  
...  
qS7zhj3ikWE3Gp9s+QijKWxx/0HDd1UuGqy0QZNqNm/M0mqVnokJNk5F4fBFxMiR  
1o63BxL8xGIRdtTCjjb2Gq2Wj7EC1Uw6CykEkxAcVk+XrRtArGkNtcYdtHfUsKVE  
wswwv0rNydrNnWhJLhS18TW5Tex+OMyTXgk9/3K8kB0mAMrtxxYjt8tm+gztkivf  
J0eo1uDJhaNxqwEZRzFyGaa4k1+56oFzRfTc  
-----END CERTIFICATE-----
```

5. 复制证书以供下一步使用。
6. 在 SVM 上安装服务器证书：

```
certificate install -type server -vserver <svm> -cert-name flexcache-  
duality
```

7. 从 [签署证书](#) 粘贴签名的证书。

示例

```
Please enter Certificate: Press <Enter> [twice] when done
-----BEGIN CERTIFICATE-----
MIIDdzCCA1+gAwIBAgIIGHolbgv5DPowDQYJKoZIhvcNAQELBQAwLjEfMB0GA1UE
AxMWY2FjaGUtMTY0Zy1zdm0tcm9vdC1jYTELMAkGA1UEBhMCVVMwHhcNMjUxMTIx
MjIxNTU4WhcNMjYxMTIxMjIxNTU4WjAfMR0wGwYDVQQDExRjYWNoZTFnLWRhdGEu
bmFzLmxhYjCCASIwDQYJKoZIhvcNAQEBBQADggEPADCCAQoCggEBAK6wmTTvk7xS
...
qS7zhj3ikWE3Gp9s+QijKWxx/0HDd1UuGqy0QZNqNm/M0mqVnokJNk5F4fBFxMiR
1o63BxL8xGIRdtTCjbjb2Gq2Wj7EC1Uw6CykEkxAcVk+XrRtArGkNtcYdtHfUsKVE
wswwv0rNydrNnWhJLhS18TW5Tex+OMyTXgk9/3K8kB0mAMrtxxYjt8tm+gztkivf
J0eo1uDJhaNxqwEZRzFyGaa4k1+56oFzRfTc
-----END CERTIFICATE-----
```

8. 粘贴在 [生成证书签名请求](#) 中生成的专用密钥。

示例

```
Please enter Private Key: Press <Enter> [twice] when done
-----BEGIN PRIVATE KEY-----
MIIEvAIBADANBgkqhkiG9w0BAQEFAASCBKYwggSiAgEAAoIBAQCuSJk075O8Uh32
9cHI6x+BaRS2w5wrqvzoYlidXtYmdCH3m1DDprBiAyfIwBC0/iU3Xd5NpB7nc1wK
1CI2VEkrXGUGwBtx1K4I1rCTB829Q1aLGAQXVyWnzhQc4ts5PW/DsQ8t7o1Z9zEI
W/gaEIajgpXIwGNWZ+weKQK+yoolxC+gy4IUE7WvnEUiezaIdoqzyPhYq5GC4XWF
0johpQuqOPe0/w2nVFRWJoFQp3ZP3NZAXC8H0qkRB6SjaM243XV2jnuEzX2joXvT
WHHH+IBAQ2JDs7s1TY0I20e49J2Fx2+HvUxDx4BHao7CCHA1+MnmEl+9E38wTaEk
NLsU724ZAgMBAECggEABHUy06wxcIk5ho3S9Ik1FDZV3JWzsu5gGdLSQOHrd5W+
...
rXGEDDaqp7jQGNXUGlbxO3zcB1l/A9Hc6oalNECgYBKwe3PeZamiwhIHLY9ph7w
dJffCshsPalMuAp2OuKIAAnNa916ft9y5kf9tIbskT+t5Dth8bmV9pwe8UZaK5eC4
Svxm19jHT5Qql0DaZVUmMXFKyKoqPDdfvcDk2Eb5gMfIIb0a3TPC/jqqpDn9BzuH
TO02fuRvRR/G/HUz2yRd+A==
-----END PRIVATE KEY-----
```

9. 输入构成服务器证书的证书链的证书颁发机构 (CA) 的证书。

这从颁发服务器证书的 CA 证书开始，范围可达根 CA 证书。

```
Do you want to continue entering root and/or intermediate certificates  
{y|n}: n
```

You should keep a copy of the private key and the CA-signed digital certificate for future reference.

The installed certificate's CA and serial number for reference:

```
CA: cache-164g-svm-root-ca  
serial: 187A256E0BF90CFA
```

10. 获取 SVM 根 CA 的公钥:

```
security certificate show -vserver <svm> -common-name <root_ca_cn> -ca  
<root_ca_cn> -type root-ca -instance  
  
-----BEGIN CERTIFICATE-----  
MIIDgTCCAmgAwIBAgIIGHokTnbshKEwDQYJKoZIhvcNAQELBQAwLjEfMB0GA1UE  
AxMWY2FjaGUtMTY0Zy1zdm0tcm9vdC1jYTELMAkGA1UEBhMCVVMwHhcNMjUxMTIx  
MjE1NTIxWhcNMjYxMTIxMjE1NTIxWjAuMR8wHQYDVQQDEXZjYWNoZS0xNjRnLXN2  
bS1yb290LWNhMQswCQYDVQQGEwJVUzCCASIwDQYJKoZIhvcNAQEBBQADggEPADCC  
...  
DoOL7vZFFt44xd+rp0DwafhSnLH5HNhdIAfa2JvZW+eJ7rgevH9wmOzyclvaih13  
Ewtb6cz1a/mtESSYRNBMGkIGM/SFCy5v1ROZXcxF96XPbYQN4cW0AYI3AHYZP0A  
H1NzDR8iml4k9IuKf6BHLFA+VwLTJJZKrdf5Jvjgh0trGAbQGI/Hp2Bjuiopkui+  
n4aa5Rz0JFQopqQddAYnMuvcq10CyNn7S0vF/XLd3fJaprH8kQ==  
-----END CERTIFICATE-----
```



这是配置客户端信任由 SVM root-ca 签名的证书所必需的。公钥会输出到控制台。复制并保存公钥。此命令中的值与您在 [创建 SVM 根 CA](#) 中输入的值相同。

步骤 2: 配置 S3 服务器

1. 启用 S3 协议访问:

```
vserver show -vserver <svm> -fields allowed-protocols
```



默认情况下，S3 在 SVM 级别是允许的。

2. 克隆现有策略:

```
network interface service-policy clone -vserver <svm> -policy default-data-files -target-vserver <svm> -target-policy <any_name>
```

3. 将 S3 添加到已克隆的策略:

```
network interface service-policy add-service -vserver <svm> -policy <any_name> -service data-s3-server
```

4. 将新策略添加到 data lif:

```
network interface modify -vserver <svm> -lif <data_lif> -service-policy duality
```



修改现有 LIF 的服务策略可能会造成中断。它要求将 LIF 取下，并使用新服务的侦听器重新启动。TCP *应该*能够快速从中恢复，但要注意潜在的影响。

5. 在 SVM 上创建 S3 对象存储服务器:

```
vserver object-store-server create -vserver <svm> -object-store-server <dns_name_of_data_lif> -certificate-name flexcache-duality
```

6. 在 FlexCache 卷上启用 S3 功能:

该 `flexcache config` 选项 `-is-s3-enabled` 必须设置为 `true`，然后才能创建存储桶。您还必须将选项 `-is-writeback-enabled` 设置为 `false`。

以下命令将修改现有 FlexCache:

```
flexcache config modify -vserver <svm> -volume <fcache_vol> -is-writeback-enabled false -is-s3-enabled true
```

7. 创建 S3 存储桶:

```
vserver object-store-server bucket create -vserver <svm> -bucket <bucket_name> -type nas -nas-path <flexcache_junction_path>
```

8. 创建存储桶策略:

```
vserver object-store-server bucket policy add-statement -vserver <svm>  
-bucket <bucket_name> -effect allow
```

9. 创建 S3 用户：

```
vserver object-store-server user create -user <user> -comment ""
```

输出示例：

```
Vserver: <svm>>  
User: <user>>  
Access Key: WCOT7...Y7D6U  
Secret Key: 6143s...pd__P  
Warning: The secret key won't be displayed again. Save this key for  
future use.
```

10. 为 root 用户重新生成密钥：

```
vserver object-store-server user regenerate-keys -vserver <svm> -user  
root
```

输出示例：

```
Vserver: <svm>>  
User: root  
Access Key: US791...2F1RB  
Secret Key: tgYmn...8_3o2  
Warning: The secret key won't be displayed again. Save this key for  
future use.
```

步骤 3：设置客户端

有许多 S3 客户端可用。AWS CLI 是一个很好的起点。有关详细信息，请参见 ["安装 AWS CLI"](#)。

管理FlexCache卷

了解如何审核ONTAP FlexCache卷

从 ONTAP 9.7 开始，您可以使用原生 ONTAP 审核以及通过 FPolicy 管理文件策略来审核 FlexCache 关系中的 NFS 文件访问事件。

从ONTAP 9.14.1开始、使用NFS或SMB的FlexCache卷支持FPolicy。以前、使用SMB的FlexCache卷不支持FPolicy。

原生审核和 FPolicy 使用与 FlexVol 卷相同的命令行界面命令进行配置和管理。但是， FlexCache 卷存在一些不同的行为。

- * 原生审核 *

- 您不能使用 FlexCache 卷作为审核日志的目标。
- 如果要审核对 FlexCache 卷的读取和写入，则必须同时在缓存 SVM 和初始 SVM 上配置审核。

这是因为文件系统操作会在处理位置进行审核。也就是说，读取会在缓存 SVM 上进行审核，而写入则会在原始 SVM 上进行审核。

- 为了跟踪写入操作的起源， SVM UUID 和 MSID 会附加在审核日志中，以确定写入的源 FlexCache 卷。

- * FPolicy*

- 尽管对 FlexCache 卷的写入是在原始卷上提交的，但 FPolicy 配置会监控对缓存卷的写入。这一点与原生审核不同，在审核中，会对初始卷上的写入进行审核。
- 虽然 ONTAP 不需要对缓存和初始 SVM 进行相同的 FPolicy 配置，但建议您部署两个类似的配置。为此，您可以为缓存创建一个新的 FPolicy 策略，该策略与原始 SVM 的策略配置相似，但新策略的范围仅限于缓存 SVM 。
- FPolicy配置中的扩展大小限制为20 KB (20480字节)。如果FlexCache卷上FPolicy配置中使用的扩展大小超过20 KB、则会触发EMS消息。 nblade.fpolicy.extn.failed

从初始卷同步ONTAP FlexCache卷的属性

FlexCache 卷的某些卷属性必须始终与原始卷的属性保持同步。如果在原始卷上修改属性后 FlexCache 卷的卷属性无法自动同步，则可以手动同步这些属性。

关于此任务

FlexCache 卷的以下卷属性必须始终与原始卷的属性同步：

- 安全风格 (-security-style)
- Volume name (-volume-name)
- 最大目录大小 (-maxdir-size)
- 最小预读 (-min-readahead)

步骤

1. 在 FlexCache 卷中，同步卷属性：

```
volume flexcache sync-properties -vserver svm_name -volume flexcache_volume
```

```
cluster1::> volume flexcache sync-properties -vserver vs1 -volume fc1
```

更新ONTAP FlexCache关系的配置

发生卷移动、聚合重新定位或存储故障转移等事件后，原始卷和 FlexCache 卷上的卷配置信息将自动更新。如果自动更新失败，则会生成一条 EMS 消息，然后您必须手动更新 FlexCache 关系的配置。

如果初始卷和 FlexCache 卷处于断开连接模式，则可能需要执行一些额外的操作来手动更新 FlexCache 关系。

关于此任务

如果要更新 FlexCache 卷的配置，必须从原始卷运行命令。如果要更新初始卷的配置，必须从 FlexCache 卷运行命令。

步骤

1. 更新 FlexCache 关系的配置：

```
volume flexcache config-refresh -peer-vserver peer_svm -peer-volume  
peer_volume_to_update -peer-endpoint-type [origin | cache]
```

在ONTAP FlexCache卷上启用文件访问时间更新

从ONTAP 9.11.1开始、您可以启用 `-atime-update` 字段FlexCache、以允许更新文件访问时间。您还可以使用设置访问时间更新期限 `-atime-update-period` 属性。。

`-atime-update-period` 属性控制访问时间更新的频率以及更新传播到初始卷的时间。

概述

ONTAP提供了一个名为的卷级字段 `-atime-update`，管理使用Read、readlink和REaddir读取的文件和目录的访问时间更新。atime用于为不常访问的文件和目录制定数据生命周期决策。不常访问的文件最终会迁移到归档存储、并且以后通常会移动到磁带。

默认情况下、现有和新创建的FlexCache 卷上的atime-update字段处于禁用状态。如果您将FlexCache卷与9.11.1之前的ONTAP版本结合使用、则应禁用"ATI-UPDAtE"字段、以便在对初始卷执行读取操作时不会必要地逐出缓存。但是、对于大型FlexCache 缓存、管理员会使用专用工具来管理数据、并帮助确保热数据保留在缓存中、冷数据也会被清除。禁用atime-update时无法执行此操作。但是、从ONTAP 9.11.1开始、您可以启用`-atime-update`和`-atime-update-period`、并使用管理缓存数据所需的工具。`

开始之前

- 所有FlexCache 卷都必须运行ONTAP 9.11.1或更高版本。
- 您必须使用 `advanced` 权限模式。

关于此任务

正在设置 ... `-atime-update-period` 到86400秒后、无论对文件执行的类似读取操作数如何、每个24小时时间段的访问时间更新都不超过一次。

设置 `-atime-update-period` 设置为0时、会针对每个读取访问向源站发送消息。然后、原始卷会通知每个FlexCache 卷、atime已过时、这会影响性能。

步骤

1. 将权限模式设置为 advanced:

```
set -privilege advanced
```

2. 启用文件访问时间更新并设置更新频率:

```
volume modify -volume vol_name -vserver <SVM name> -atime-update true -atime-update-period <seconds>
```

以下示例将启用 -atime-update 和设置 -atime-update-period 至86400秒或24小时:

```
c1: volume modify -volume origin1 vs1_c1 -atime-update true -atime-update-period 86400
```

3. 请验证 -atime-update 已启用:

```
volume show -volume vol_name -fields atime-update,atime-update-period
```

```
c1::>*> volume show -volume cachel_origin1 -fields atime-update,atime-update-period
vserver volume      atime-update atime-update-period
-----
vs2_c1  cachel_origin1  true          86400
```

4. 启用后 -atime-update、您可以指定是否可以自动擦除FlexCache卷上的文件以及擦洗间隔:

```
volume flexcache config modify -vserver <SVM name> -volume <volume_name> -is-atime-scrub-enabled <true|false> -atime-scrub-period <integer>
```

有关参数的详细信息 -is-atime-scrub-enabled, 请参见["ONTAP 命令参考"](#)。

在ONTAP FlexCache卷上启用全局文件锁定

从 ONTAP 9.10.1 开始, 可以应用全局文件锁定, 以防止读取所有相关缓存文件。

启用全局文件锁定后、对初始卷的修改将暂停、直到所有FlexCache卷联机为止。只有当您能够控制缓存与初始卷之间连接的可靠性时, 才应启用全局文件锁定, 因为在 FlexCache 卷脱机时, 由于暂停以及修改可能超时而导致此连接的可靠性。

开始之前

- 全局文件锁定要求包含原始缓存以及所有关联缓存的集群运行 ONTAP 9.9.1 或更高版本。可以在新的或现有的 FlexCache 卷上启用全局文件锁定。此命令可在单个卷上运行、并适用于所有关联的FlexCache卷。
- 要启用全局文件锁定, 您必须处于高级权限级别。
- 如果还原到ONTAP 9.9.1之前的版本、则必须先在初始缓存和关联缓存上禁用全局文件锁定。要禁用此功能、请从初始卷运行: `volume flexcache prepare-to-downgrade -disable-feature-set`

9.10.0

- 启用全局文件锁定的过程取决于源站是否具有现有缓存：

- [enable-gfl-new]
- [enable-gfl-existing]

在新 FlexCache 卷上启用全局文件锁定

步骤

- 使用创建FlexCache卷 -is-global-file-locking 设置为true：

```
volume flexcache create volume volume_name -is-global-file-locking-enabled true
```



的默认值 -is-global-file-locking 为"false"。任意后续操作 volume flexcache create 命令在卷上运行、必须使用传递 -is-global-file-locking enabled 设置为"true"。

在现有 FlexCache 卷上启用全局文件锁定

步骤

- 必须从原始卷设置全局文件锁定。
- 源服务器不能具有任何其他现有关系（例如 SnapMirror）。任何现有关系都必须解除关联。必须在运行命令时连接所有缓存和卷。要检查连接状态，请运行：

```
volume flexcache connection-status show
```

列出的所有卷的状态均应显示为 connected。有关详细信息，请参见 "[查看 FlexCache 关系的状态](#)" 或 "[从原始卷同步 FlexCache 卷的属性](#)"。

- 对缓存启用全局文件锁定：

```
volume flexcache origin config show/modify -volume volume_name -is-global-file-locking-enabled true
```

相关信息

- ["ONTAP 命令参考"](#)

预填充ONTAP FlexCache卷

您可以预先填充 FlexCache 卷，以缩短访问缓存数据所需的时间。

开始之前

- 您必须是高级权限级别的集群管理员
- 您为预填充传递的路径必须存在、否则预填充操作将失败。

关于此任务

- 仅预填充读取文件并遍历目录
- -isRecursion 标记适用场景 传递给预填充的整个目录列表

步骤

1. 预先填充 FlexCache 卷：

```
volume flexcache prepopulate -cache-vserver vserver_name -cache-volume -path-list path_list -isRecursion true|false
```

◦ -path-list 参数表示从原始根目录开始要预先填充的相对目录路径。例如、如果原始根目录名为/origin、并且包含目录/origin/dir1和/origin/dir2，则可以按如下所示指定路径列表：-path-list dir1, dir2 或 -path-list /dir1, /dir2。

◦ 的默认值 -isRecursion 参数为True。

此示例预先填充单个目录路径：

```
cluster1::*:> flexcache prepopulate start -cache-vserver vs2 -cache-volume fg_cachevol_1 -path-list /dir1  
(volume flexcache prepopulate start)  
[JobId 207]: FlexCache prepopulate job queued.
```

以下示例预先填充了多个目录中的文件：

```
cluster1::*:> flexcache prepopulate start -cache-vserver vs2 -cache-volume fg_cachevol_1 -path-list /dir1,/dir2,/dir3,/dir4  
(volume flexcache prepopulate start)  
[JobId 208]: FlexCache prepopulate job queued.
```

此示例预先填充单个文件：

```
cluster1::*:> flexcache prepopulate start -cache-vserver vs2 -cache-volume fg_cachevol_1 -path-list /dir1/file1.txt  
(volume flexcache prepopulate start)  
[JobId 209]: FlexCache prepopulate job queued.
```

此示例将预先填充原始文件中的所有文件：

```
cluster1::*:> flexcache prepopulate start -cache-vserver vs2 -cache-volume fg_cachevol_1 -path-list / -isRecursion true  
(volume flexcache prepopulate start)  
[JobId 210]: FlexCache prepopulate job queued.
```

此示例包含一个无效的预填充路径：

```

cluster1::*> flexcache prepopulate start -cache-volume
vol_cache2_vs3_c2_vol_origin1_vs1_c1 -cache-vserver vs3_c2 -path-list
/dirl, dir5, dir6
(volume flexcache prepopulate start)

Error: command failed: Path(s) "dir5, dir6" does not exist in origin
volume
"vol_origin1_vs1_c1" in Vserver "vs1_c1".

```

2. 显示读取的文件数：

```
job show -id job_ID -ins
```

相关信息

- ["作业显示"](#)

删除ONTAP FlexCache关系

如果不再需要 FlexCache 卷，则可以删除 FlexCache 关系和 FlexCache 卷。

开始之前

如果您启用了 FlexCache 回写，则必须先将其禁用，然后才能删除 FlexCache 卷。请参阅["在FlexCache卷上禁用回写"](#)。

步骤

1. 从包含 FlexCache 卷的集群中，使 FlexCache 卷脱机：

```
volume offline -vserver svm_name -volume volume_name
```

2. 删除 FlexCache 卷：

```
volume flexcache delete -vserver svm_name -volume volume_name
```

FlexCache 关系详细信息将从原始卷和 FlexCache 卷中删除。

用于修复热点的FlexCache

在使用ONTAP FlexCache卷的高性能计算工作负载中修补热抽查

许多高性能计算工作负载(例如动画渲染或EDA)的一个常见问题是HotSpoting。如果集群或网络的特定部分遇到的负载比其他区域高得多、则会出现热抽样、这会由于该位置集中的数据流量过多而导致性能瓶颈和整体效率降低。例如、正在运行的作业对一个或多个文件的需求较高、这会导致用于(通过卷关联性)处理该文件的请求的CPU出现瓶颈。FlexCache可以帮助缓解此瓶颈、但必须正确设置。

本文档介绍如何设置FlexCache以修复热点。



从2024年7月开始、以前以PDF格式发布的技术报告中的内容已与ONTAP产品文档集成。此ONTAP热点修复技术报告内容自发布之日起为全新内容、之前从未生成过任何格式。

关键概念

在规划热点修复时、了解这些基本概念非常重要。

- **高密度FlexCache (HDF)**: 一种FlexCache，它会根据缓存容量要求进行精简，以尽可能少地覆盖节点
- **HDF阵列(HDFA)**: 一组HDFS，它们是位于集群中的相同来源的缓存
- **SVM间HDFA**: 每个服务器虚拟机(SVM)从HDFA中获取一个HDF
- **SVM内HDFA**: HDFA中的所有HDFS都位于一个SVM中
- **东西向流量**: 间接数据访问生成的集群后端流量

下一步行动

- "[了解如何使用高密度FlexCache进行架构设计以帮助修复热点](#)"
- "[确定FlexCache阵列密度](#)"
- "[确定HDFS的密度、并确定是否要使用NFS以及SVM间HDFA和SVM内HDFA访问HDFS](#)"
- "[配置HDFA和数据lifs、以实现在ONTAP配置中使用集群内缓存的优势](#)"
- "[了解如何配置客户端以使用客户端配置分布ONTAP NAS连接](#)"

构建ONTAP FlexCache热点修复解决方案

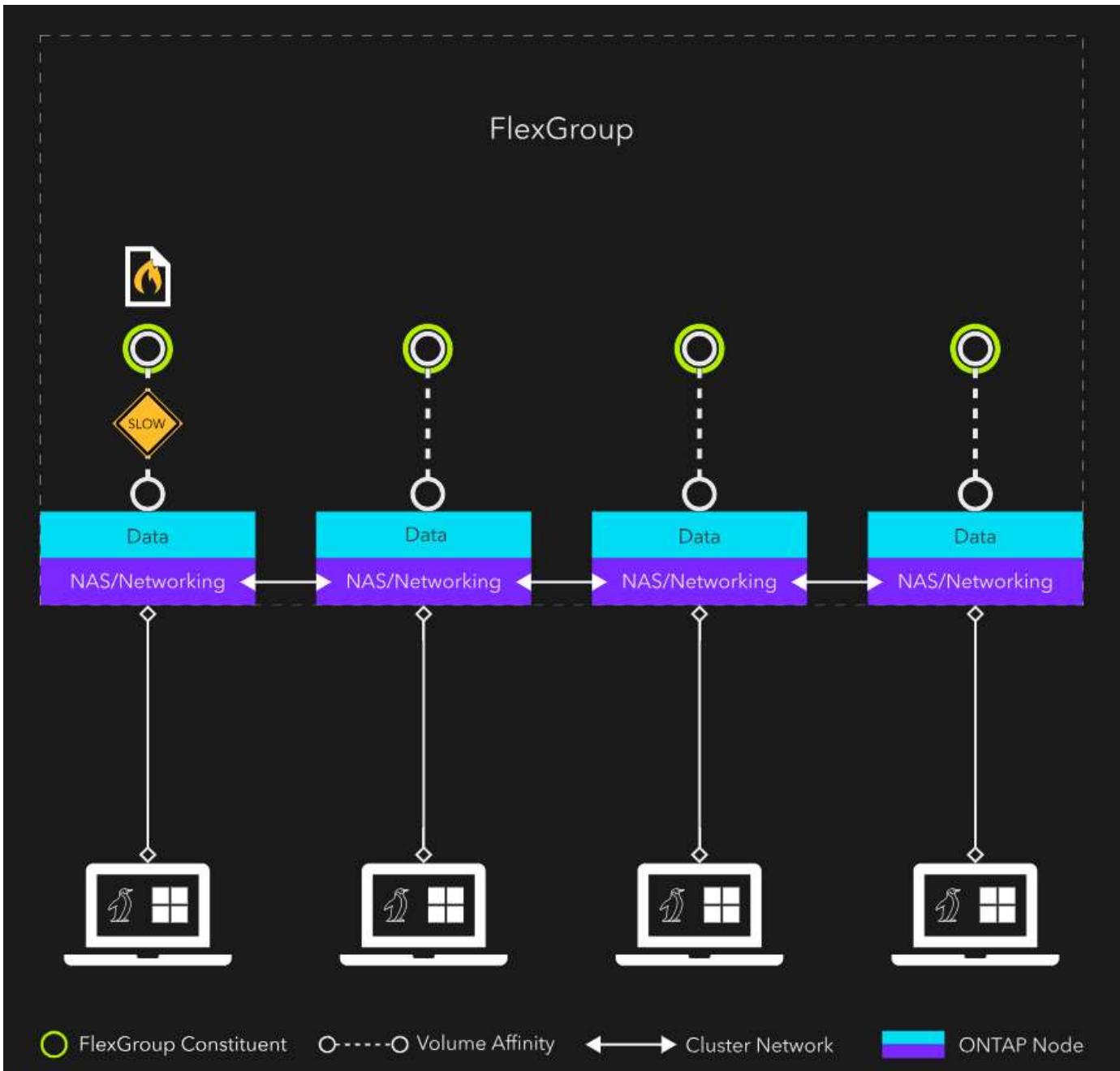
要修复热点问题、请了解瓶颈的根本原因、自动配置的FlexCache不足的原因以及有效构建FlexCache解决方案所需的技术详细信息。通过了解和实施高密度FlexCache阵列(HDFA)、您可以优化高需求工作负载的性能并消除瓶颈。

了解瓶颈

下面[图像](#)显示了典型的单文件热抽查情形。此卷是一个FlexGroup、每个节点具有一个成分卷、文件驻留在节点1上。

如果将所有NAS客户端的网络连接分布在集群中的不同节点上、则为热文件所在的卷关联性提供服务的CPU仍会出现瓶颈。此外、对于来自连接到文件所在节点以外的节点的客户端的呼叫、还会引入集群网络流量(东西向流量)。东西向流量开销通常很小、但对于高性能计算工作负载来说、每一个小位都很重要。

图1：FlexGroup单文件热点场景

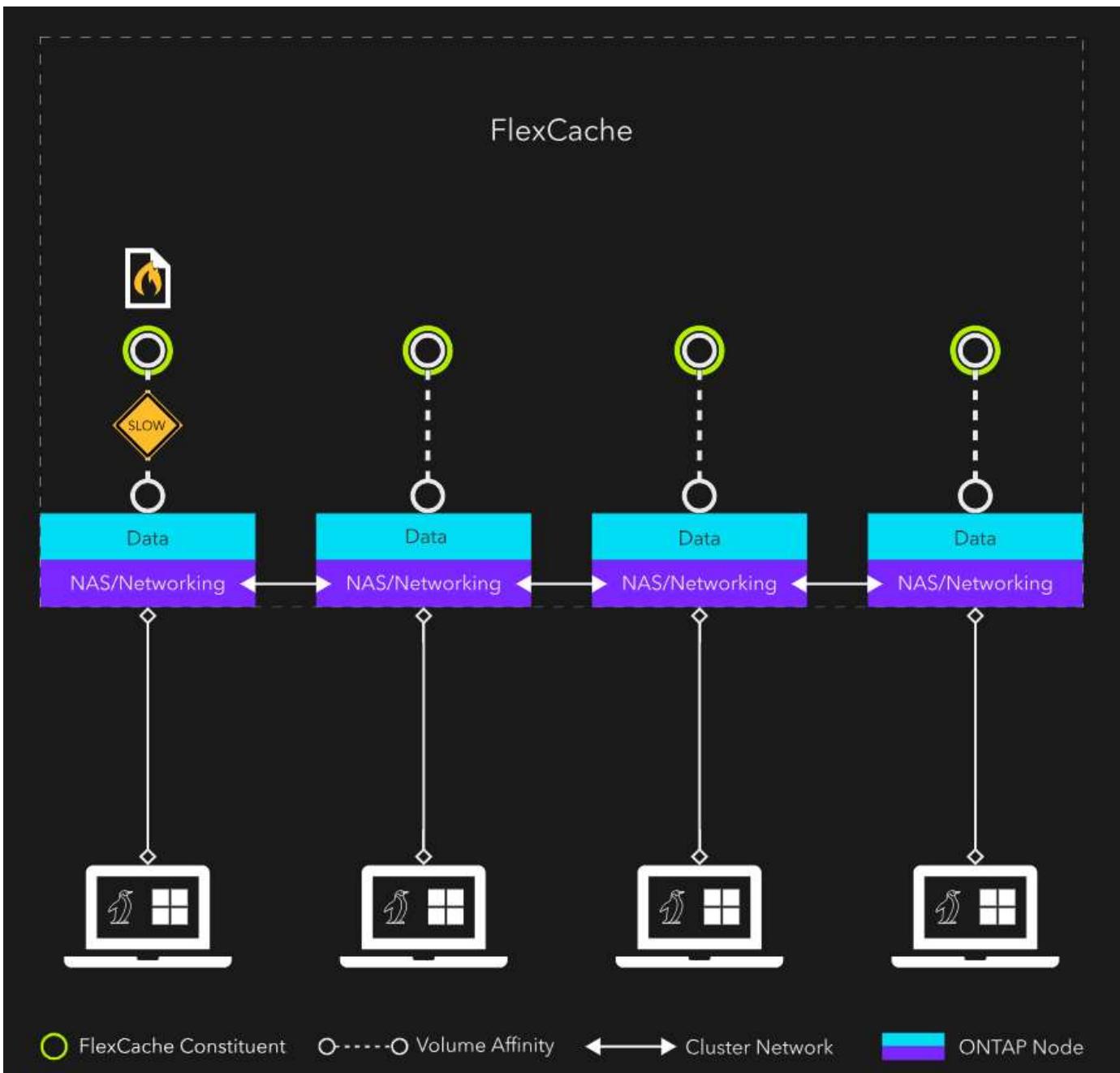


为什么自动配置的**FlexCache**不是答案

要补救HotSpoting、请消除CPU瓶颈、最好也消除东西向流量。如果设置正确、FlexCache可以提供帮助。

在以下示例中，FlexCache使用系统管理器、NetApp控制台或默认 CLI 参数自动配置。[图1.](#)和[图2.](#)乍一看它们是一样的：都是四节点、单成分 NAS 容器。唯一的区别是，图 1 的 NAS 容器是FlexGroup，而图 2 的 NAS 容器是FlexCache。每个图都描述了相同的瓶颈：节点 1 的 CPU 用于卷亲和性服务访问热文件，以及导致延迟的东西向流量。自动配置的FlexCache并未消除瓶颈。

图2：自动配置的**FlexCache**方案



FlexCache解剖结构

要有效地构建FlexCache以进行热点修复、您需要了解有关FlexCache的一些技术详细信息。

FlexCache始终是稀疏FlexGroup。一个FlexGroup由多个FlexVol组成。这些FlexVol称为FlexGroup成分卷。在默认FlexGroup布局中、集群中的每个节点都有一个或多个成分卷。这些成分卷在抽象层下“缝合在一起”、并作为一个大型NAS容器提供给客户端。将文件写入FlexGroup后、通过数据启发法确定要将文件存储在哪个成分卷上。它可能是包含客户端NAS连接的成分卷、也可能是其他节点。位置无关紧要、因为所有内容都在抽象层下运行、并且对客户端不可见。

让我们将对FlexGroup的了解应用于FlexCache。由于FlexCache是基于FlexGroup构建的，因此默认情况下、您有一个，该FlexCache在集群中的所有节点上都有成分卷，如图1所示。在大多数情况下、这是一件好事。您正在利用集群中的所有资源。

但是、对于热文件的解决、这并不理想、因为存在两个瓶颈：单个文件的CPU和东西向流量。如果为热文件创

建的FlexCache在每个节点上都有成分卷、则该文件仍将仅驻留在其中一个成分卷上。这意味着、只有一个CPU可以处理对热文件的所有访问。此外、您还希望限制访问热文件所需的东西向流量。

该解决方案是一组高密度FlexCaches。

高密度FlexCache剖析

高密度FlexCache (HDF)在缓存数据的容量要求允许的数量下、在尽可能少的节点上包含成分卷。其目标是使缓存驻留在单个节点上。如果容量要求导致无法实现这一点、则只能在少数节点上创建成分卷。

例如、一个24节点集群可以具有三个高密度FlexCaches：

- 一个跨节点1到8
- 一秒、跨越节点9到16
- 第三个节点、跨越节点17到24

这三个HDFS构成一个高密度FlexCache阵列(HDFA)。如果文件在每个HDF中均匀分布、则客户端请求的文件可能会驻留在前端NAS连接的本地、这种可能性为八分之一。如果有12个HDFS、每个HDFS仅跨越两个节点、则文件位于本地的几率为50%。如果您可以将HDF折叠到一个节点并创建其中24个节点、则可以保证该文件为本地文件。

此配置将消除所有东西向流量、最重要的是、将为访问热文件提供24个CPU/volume关联性。

下一步是什么？

["确定FlexCache阵列密度"](#)

相关信息

["有关FlexGroup和TRs的文档"](#)

确定ONTAP FlexCache密度

您的第一个热点修复设计决策是确定FlexCache密度。以下示例是四节点集群。假设文件数量均匀分布在每个HDF中的所有成分卷之间。此外、还假定在所有节点之间均匀分布前端NAS连接。

尽管这些示例并不是您可以使用的唯一配置、但您应了解指导设计原则、以便根据空间要求和可用资源的允许创建尽可能多的HDFS。

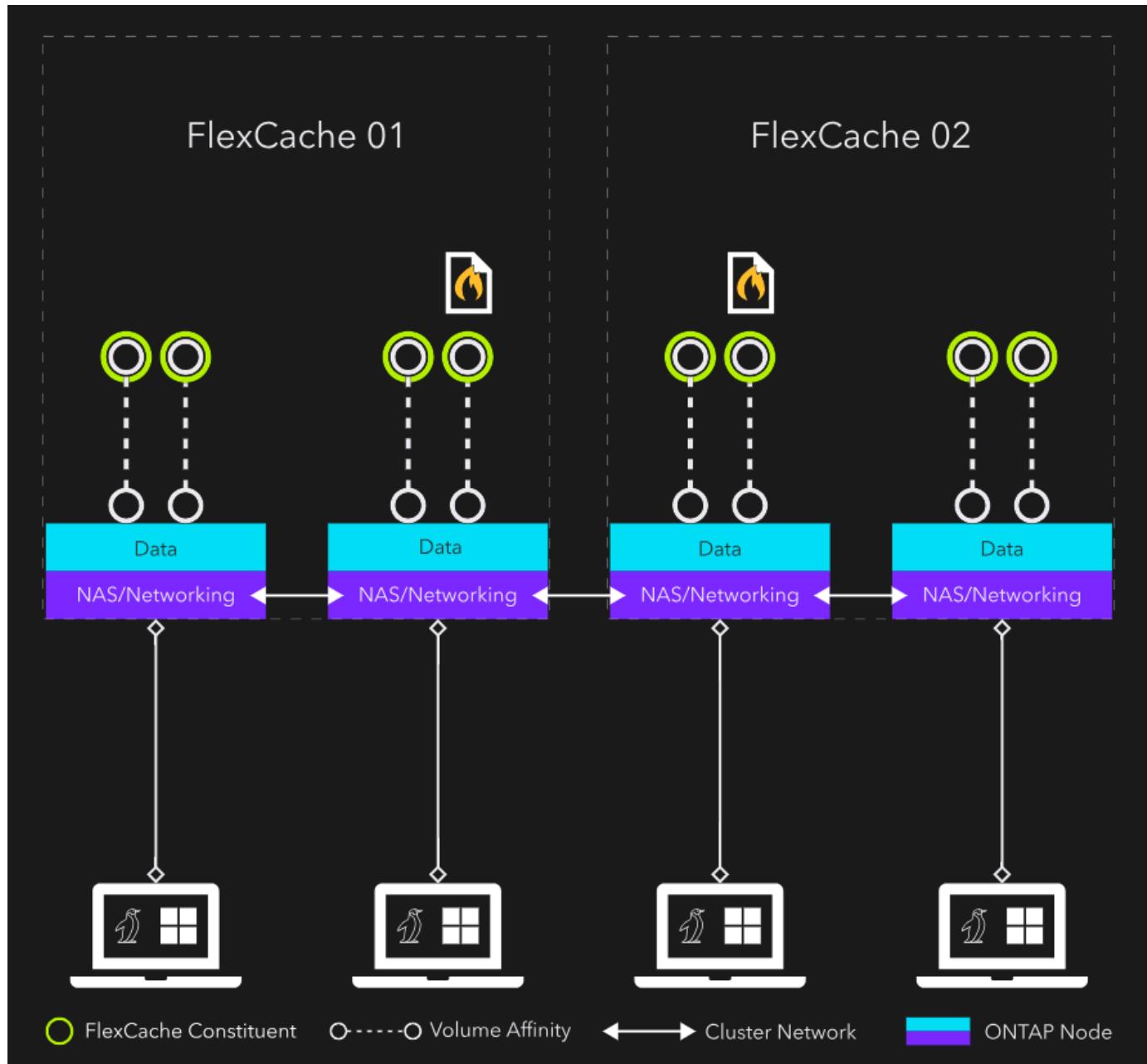


HDFA使用以下语法表示：HDFs per HDFA x nodes per HDF x constituents per node per HDF

2x2x2 HDFA配置

图1.是一个2x2X2 HDFA配置示例：两个HDFS、每个HDFS跨越两个节点、每个节点包含两个成分卷。在此示例中、每个客户端有50%的机会直接访问热文件。四个客户端中有两个具有东西向流量。重要的是、现在有两个HDFS、这意味着热文件有两个不同的缓存。现在、每个卷有两个CPU/Affinities、用于访问热文件。

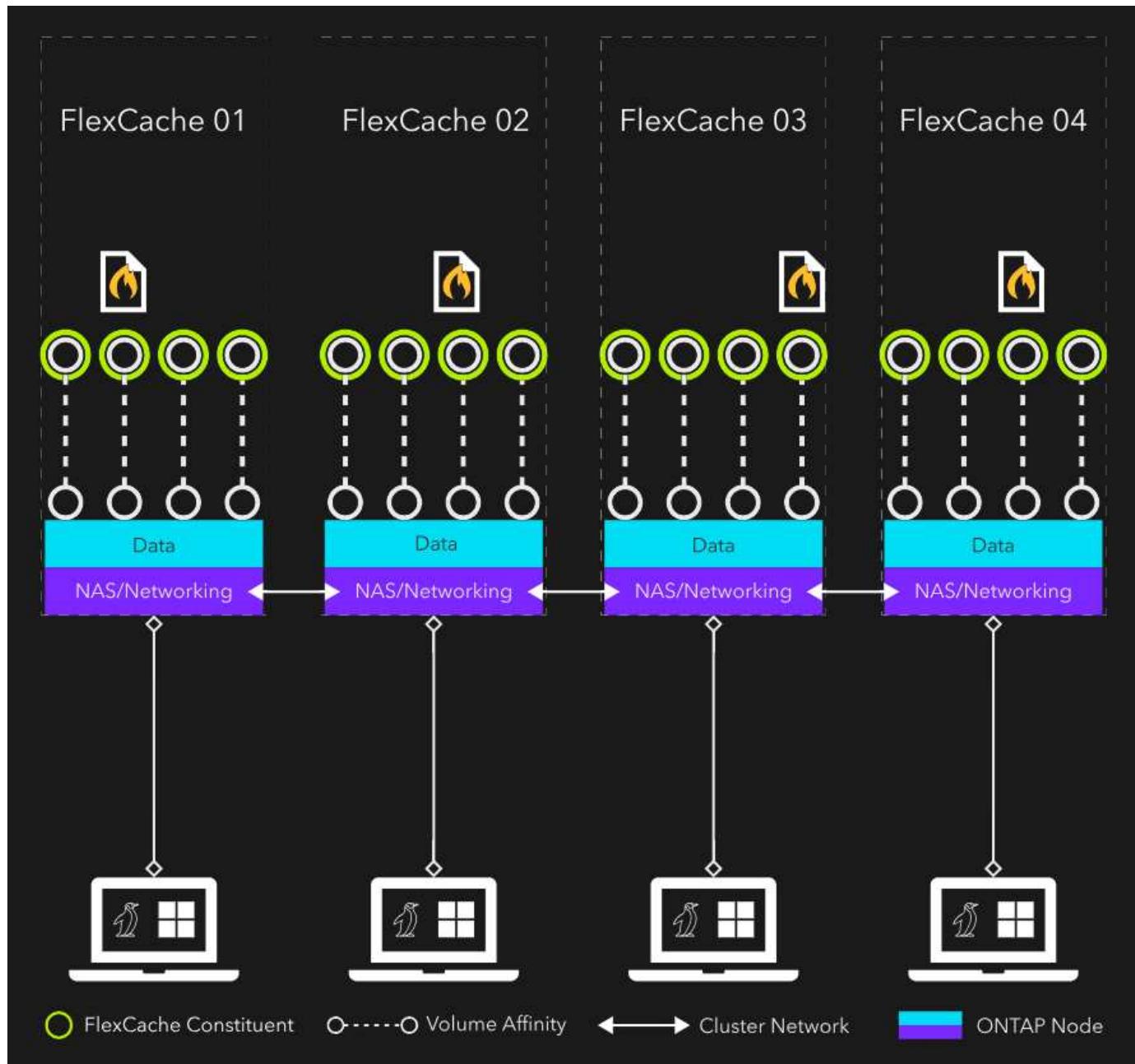
图1：2x2X2 HDFA配置



4x1x4 HDFA配置

图2.表示最佳配置。下面是4x1x4 HDFA配置的一个示例：四个HDFS、每个HDFS包含在一个节点中、每个节点包含四个成分卷。在此示例中、每个客户端都保证可以直接访问热文件的缓存。由于四个不同节点上有四个缓存文件、因此四个不同的CPU/卷关联有助于服务对热文件的访问。此外、不会生成任何东西流量。

图2：4x1x4 HDFA配置



下一步行动

在确定HDFS的密度后，如果要使用通过NFS访问HDFS，则必须做出另一个设计决定“[SVM间HDFA和SVM内HDFA](#)”。

确定ONTAP的SVM间或SVM内HDFA选项

确定HDFS的密度后、请确定是否要使用NFS访问HDFS、并了解SVM间HDFA和SVM内HDFA选项。



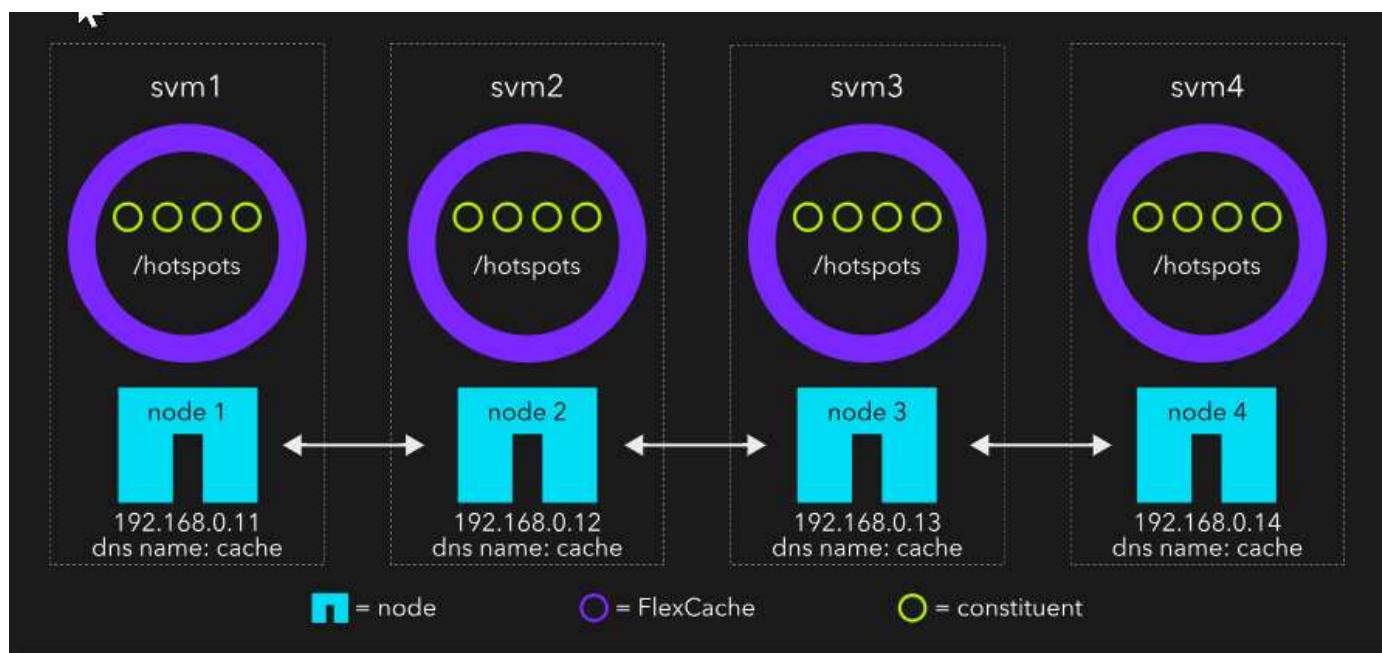
如果只有SMB客户端要访问HDFS，则应在一个SVM中创建所有HDFS。请参见Windows客户端配置、了解如何使用DFS目标进行负载平衡。

SVM间HDFA部署

SVM间HDFA要求为HDF中的每个HDF创建一个SVM。这样、HDFA中的所有HDFS都可以具有相同的接合路径、从而更轻松地在客户端进行配置。

在此示例中图1.、每个HDF都位于各自的SVM中。这是SVM间HDFA部署。每个HDF都有一个接合路径 /hotspots。此外、每个IP都有一条DNS A主机名缓存记录。此配置可利用DNS循环对不同HDFS的挂载进行负载平衡。

图1：4x1x4 SVM间HDFA配置

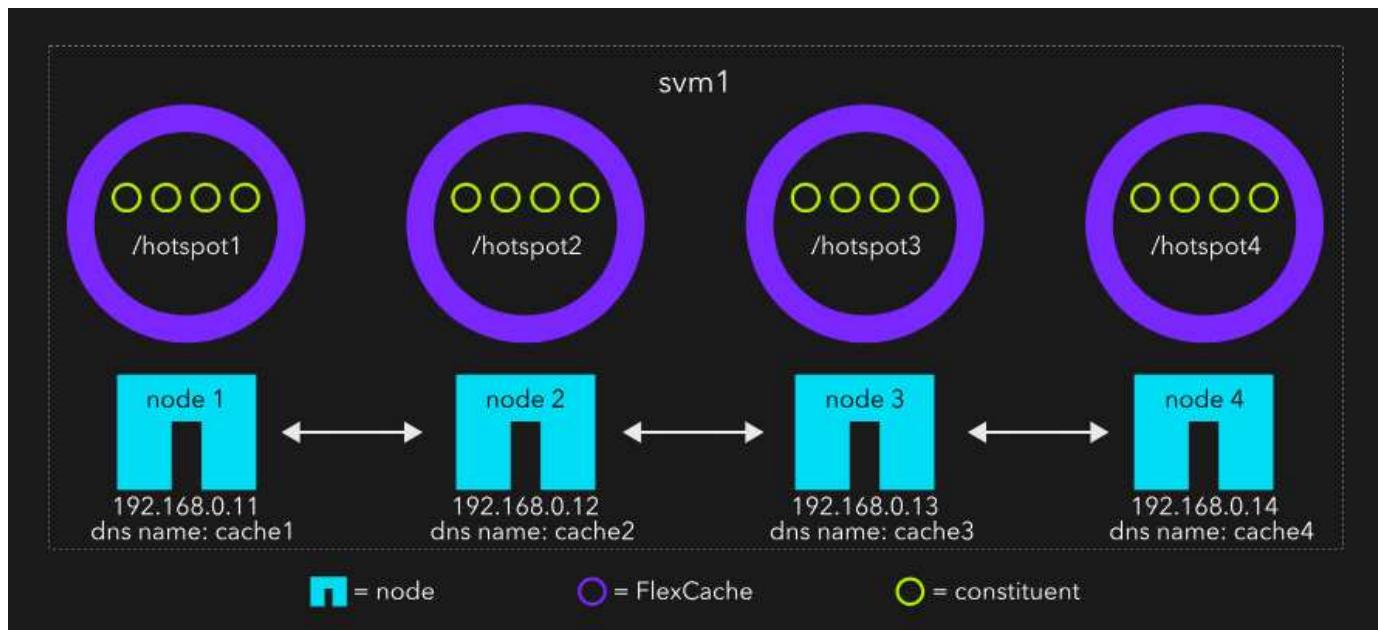


SVM内HDFA部署

SVM内需要每个HDF都具有唯一的接合路径、但所有HDFS都位于一个SVM中。在ONTAP中进行此设置更容易、因为它只需要一个SVM、但需要在Linux端进行更高级的配置 autofs、并在ONTAP中放置数据LIF。

在此示例中图2.、每个HDF都位于同一个SVM中。这是SVM内HDFA部署、需要接合路径是唯一的。要使负载平衡正常工作、您需要为每个IP创建一个唯一的DNS名称、并仅将主机名解析为的数据SIF放置在HDF所在的节点上。您还需要配置 `autofs` 多个条目，如中所述["Linux客户端配置"](#)。

图2：4x1x4 SVM内HDFA配置



下一步行动

现在，您已经了解了如何部署HDFA，["部署HDFA并将客户端配置为以分布式方式访问它们"](#)。

配置HDFA和ONTAP数据LUN

您需要相应地配置HDFA和数据Mifs、才能实现此热点修复解决方案的优势。此解决方案使用集群内缓存、并在同一集群中使用原始缓存和HDFA。

以下是两个HDFA示例配置：

- 2x2个SVM间HDFA
- 4x1x4 SVM内HDFA

关于此任务

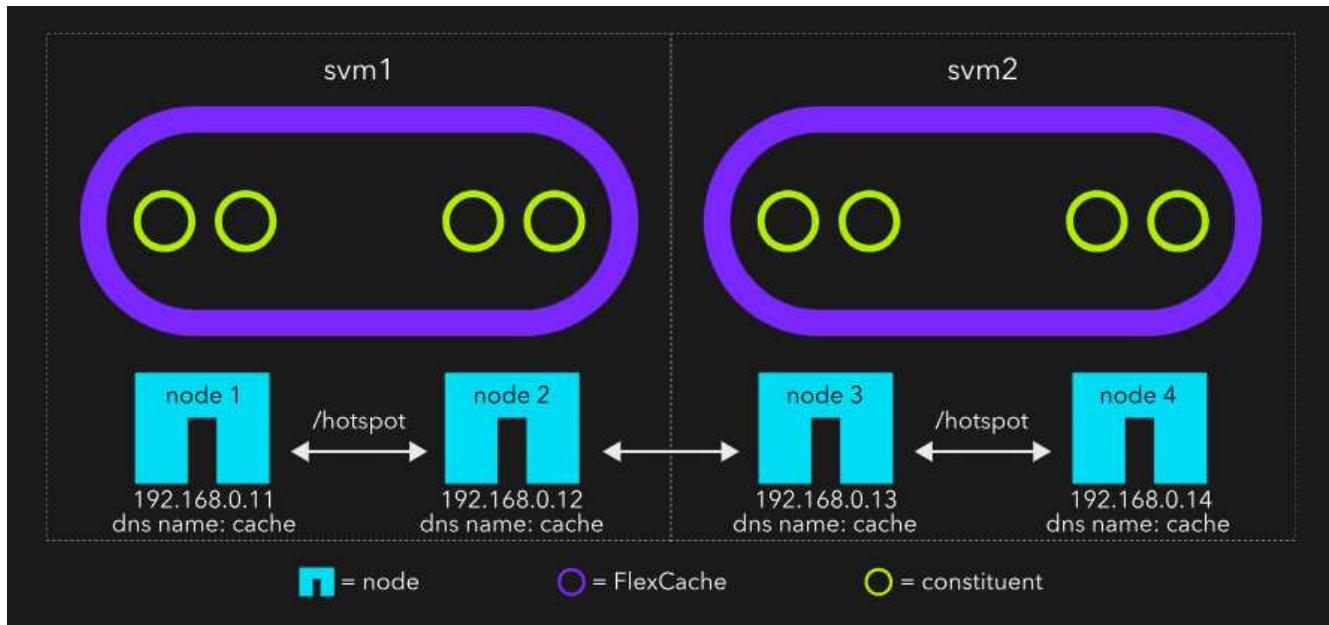
使用ONTAP命令行界面执行此高级配置。命令中必须使用两种配置、其中一种配置必须`flexcache create`确保未配置：

- `-aggr-list`: 提供位于要将HDF限制到的节点或节点子集上的聚合或聚合列表。
- `-aggr-list-multiplier`: 确定要为选项中列出的每个聚合创建多少个成分卷 `aggr-list`。如果列出了两个聚合，并将此值设置为 2，则最终将包含四个成分卷。NetApp建议每个聚合最多包含8个成分卷、但16个也足以满足要求。
- `-auto-provision-as`: 如果您选择退出，CLI将尝试自动加载并将值设置为 `flexgroup`。确保未配置此项。如果出现、请将其删除。

创建2x2X2 SVM间HDFA配置

1. 要帮助配置2x2X2 SVM间HDFA (如图1所示)、请完成准备工作表。

图1：2x2个SVM间HDFA布局



SVM	每个HDF的节点数	聚合	每个节点的成分卷数	Junction path	数据LIF IP
svm1	node1、node2	aggr1、aggr2	2.	/hotspot	192.168.0.11、192.168.0.12
svm2	node3、node4	aggr3、aggr4	2.	/hotspot	192.168.0.13、192.168.0.14

2. 创建HDFS。运行以下命令两次、对准备工作表中的每行运行一次。确保为第二次迭代调整 `vserver` 和 `aggr-list` 值。

```
cache::> flexcache create -vserver svm1 -volume hotspot -aggr-list aggr1,aggr2 -aggr-list-multiplier 2 -origin-volume <origin_vol> -origin -vserver <origin_svm> -size <size> -junction-path /hotspot
```

3. 创建数据生命周期。运行命令四次、在准备工作表中列出的节点上为每个SVM创建两个数据文件。确保为每次迭代适当调整值。

```
cache::> net int create -vserver svm1 -home-port e0a -home-node node1 -address 192.168.0.11 -netmask-length 24
```

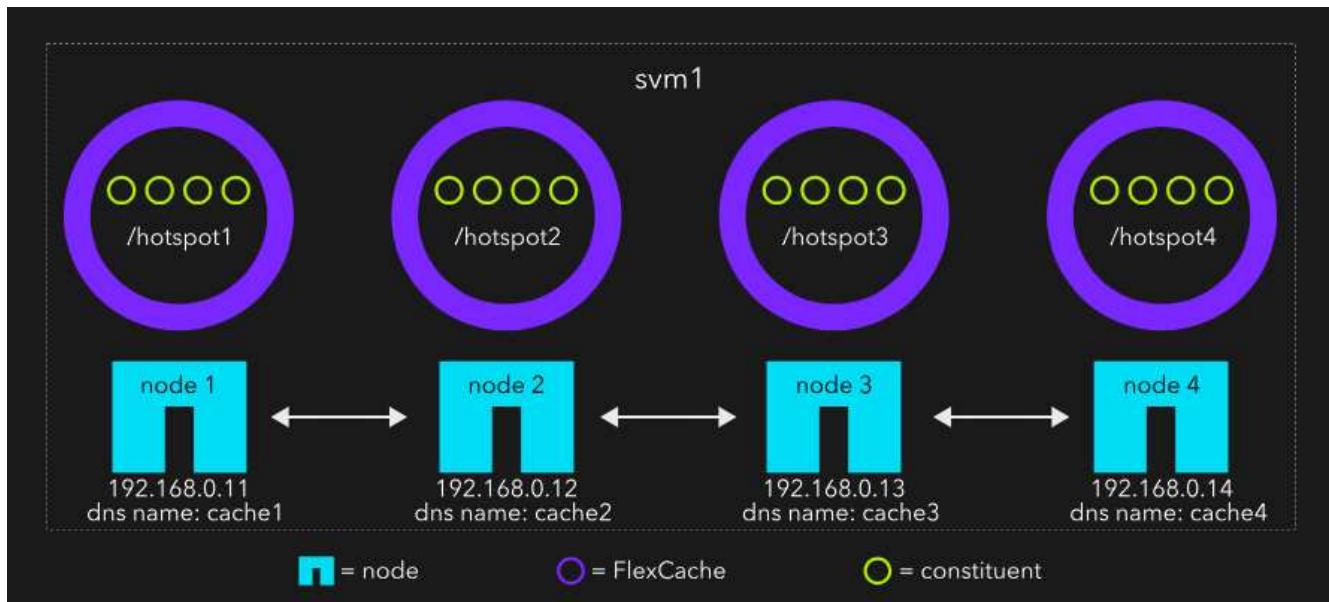
下一步行动

现在、您需要对客户端进行配置、以便适当地利用HDFA。请参阅。 "["客户端配置"](#)"

创建4x1x4 SVM内HDFA

1. 要帮助配置4x1x4 SVM间HDFA (如图2所示)、请填写准备表。

图2：4x1x4 SVM内HDFA布局



SVM	每个HDF的节点数	聚合	每个节点的成分卷数	Junction path	数据LIF IP
svm1	节点 1	aggr1	4.	/hotspot1	192.168.0.11
svm1	节点 2.	aggr2	4.	/hotspot2	192.168.0.12
svm1	node3	aggr3	4.	/hotspot3	192.168.0.13
svm1	node4	aggr4	4.	/hotspot4	192.168.0.14

2. 创建HDFS。运行以下命令四次、对准备工作表中的每行运行一次。请确保调整 `aggr-list` 每次迭代的和 `junction-path` 值。

```
cache::> flexcache create -vserver svm1 -volume hotspot1 -aggr-list aggr1 -aggr-list-multiplier 4 -origin-volume <origin_vol> -origin-vserver <origin_svm> -size <size> -junction-path /hotspot1
```

3. 创建数据生命周期。运行命令四次、在SVM中总共创建四个数据生命周期。每个节点应有一个数据LIF。确保为每次迭代适当调整值。

```
cache::> net int create -vserver svm1 -home-port e0a -home-node node1 -address 192.168.0.11 -netmask-length 24
```

下一步行动

现在、您需要对客户端进行配置、以便适当地利用HDFA。请参阅。 "["客户端配置"](#)"

配置客户端以分布ONTAP NAS连接

要补救"热点"、请正确配置客户端、使其在防止CPU瓶颈方面发挥作用。

Linux客户端配置

无论您选择SVM内还是SVM间HDFA部署、都应在Linux中使用`autofs`来确保客户端在不同HDFS之间实现负载平衡。`autofs` SVM间和SVM内的配置有所不同。

开始之前

您需要`autofs`安装和相应的依赖项。有关此操作的帮助、请参见Linux文档。

关于此任务

所述步骤将使用一个示例`/etc/auto_master`文件、其中包含以下条目：

```
/flexcache auto _hotspot
```

此选项会配置`autofs`为在进程尝试访问目录`/flexcache`中`/etc`调用的文件`auto_hotspot`。文件内容`auto_hotspot`将指定要挂载到目录中的NFS服务器和接合路径`/flexcache`。所述示例是文件的不同配置`auto_hotspot`。

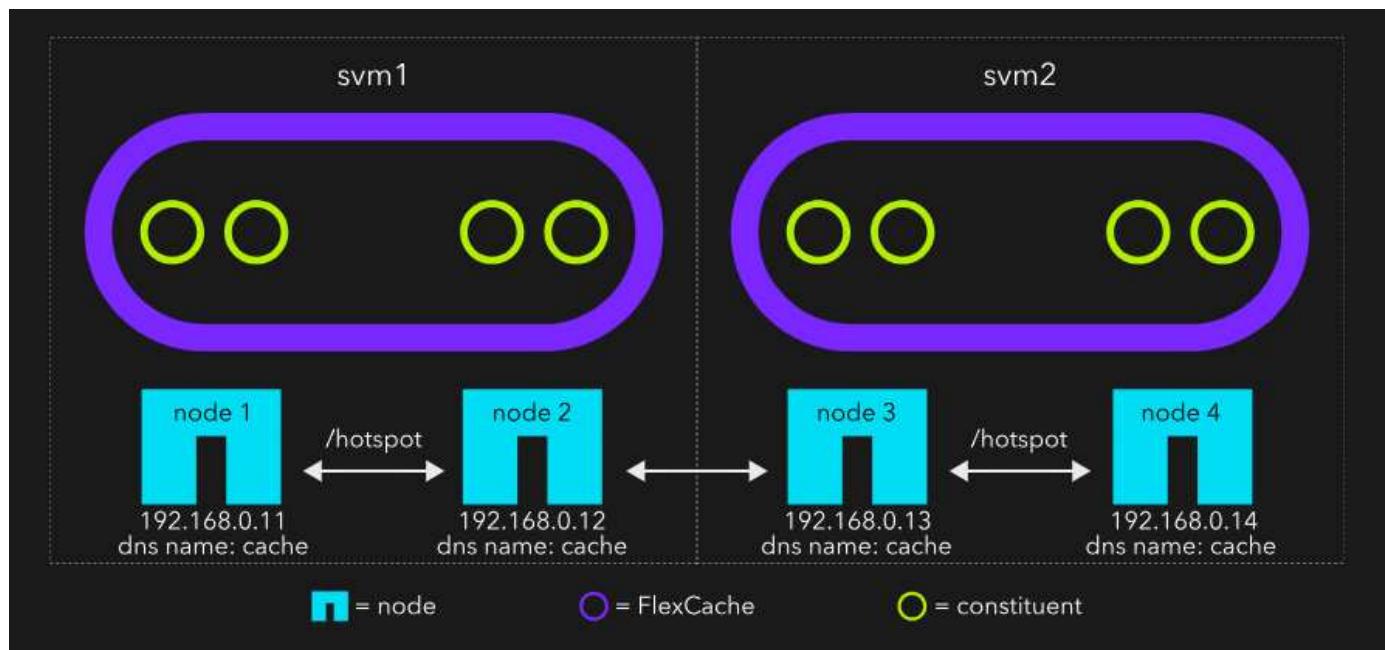
SVM内HDFA自动配置

在以下示例中，我们将在中为图创建一个`autofs`地图[图1](#)。由于每个缓存都有相同的接合路径、并且主机名`cache`有四条DNS A记录、因此我们只需要一行：

```
hotspot cache:/hotspot
```

这一行很简单，会使NFS客户端对主机名进行DNS查找`cache`。DNS已设置为以循环方式返回IP。这样可以均匀分布前端NAS连接。客户端收到IP后，将在上`/flexcache/hotspot`挂载接合路径`/hotspot`。它可以连接到SVM1、SVM2、SVM3或SVM4、但特定SVM无关紧要。

图1：2x2个SVM间HDFA



SVM内HDFA自动配置

在以下示例中，我们将在中为图创建一个 `autofs` 地图^{图2}。我们需要确保NFS客户端装载HDF接合路径部署中的IP。换言之、我们不希望使用IP 192.168.0.11以外的任何其他内容进行挂载 `/hotspot1`。为此、我们可以在映射中列出一个本地挂载位置的所有四个IP/接合路径对 `auto_hotspot`。



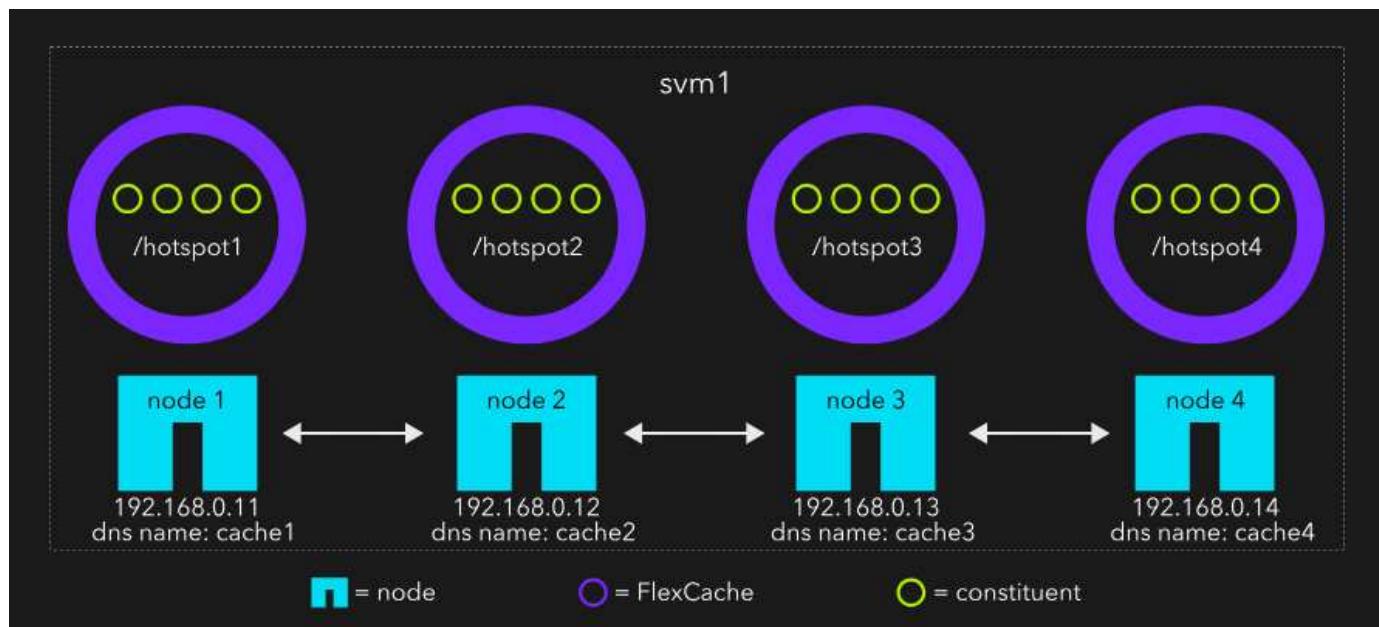
(\`在以下示例中，反斜杠将继续条目到下一行，使其更易于阅读。

```
hotspot      cache1:/hostspot1 \
              cache2:/hostspot2 \
              cache3:/hostspot3 \
              cache4:/hostspot4
```

当客户端尝试访问时 `/flexcache/hotspot`，`autofs` 将对所有四个主机名进行正向查找。假设所有四个IP都与客户端位于同一子网中或位于不同子网中、`autofs` 则会对每个IP发出NFS NULL ping。

此空ping操作需要ONTAP的NFS服务处理数据包、但不需要任何磁盘访问。要返回的第一个ping将是IP、接合路径 `autofs` 选择挂载。

图2：4x1x4 SVM内HDFA



Windows客户端配置

对于Windows客户端、应使用SVM内HDFA。要在SVM中的不同HDFS之间进行负载平衡、必须为每个HDF添加一个唯一的共享名称。之后、按照中的步骤[Microsoft 文档](#)为同一文件夹实施多个DFS目标。

版权信息

版权所有 © 2026 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。