



使用 **VMware** 和 **Microsoft** 副本卸载实现存储虚拟化 ONTAP 9

NetApp
September 12, 2024

目录

- 使用 VMware 和 Microsoft 副本卸载实现存储虚拟化 1
 - 使用 VMware 和 Microsoft 副本卸载的存储虚拟化概述 1
 - LUN 访问在虚拟化环境中的工作原理 1
 - 提高 ESX 主机的 VMware VAAI 性能 3
 - Microsoft 卸载数据传输（ Offloaded Data Transfer ， ODX ） 4

使用 VMware 和 Microsoft 副本卸载实现存储虚拟化

使用 VMware 和 Microsoft 副本卸载的存储虚拟化概述

VMware 和 Microsoft 支持副本卸载操作，以提高性能和网络吞吐量。您必须根据 VMware 和 Windows 操作系统环境的要求对系统进行配置，以使用各自的副本卸载功能。

在虚拟化环境中使用 VMware 和 Microsoft 副本卸载时，必须对齐 LUN。未对齐的 LUN 可能会降低性能。

使用虚拟化 SAN 环境的优势

通过使用 Storage Virtual Machine（SVM）和 LIF 创建虚拟化环境，您可以将 SAN 环境扩展到集群中的所有节点。

- 分布式管理

您可以登录到 SVM 中的任何节点来管理集群中的所有节点。

- 增加数据访问

通过 MPIO 和 ALUA，您可以通过 SVM 的任何活动 iSCSI 或 FC LIF 访问数据。

- 受控制的 LUN 访问

如果使用 SLM 和端口集，则可以限制启动程序可用于访问 LUN 的 LIF。

LUN 访问在虚拟化环境中的工作原理

在虚拟化环境中，LIF 使主机（客户端）能够通过优化和非优化路径访问 LUN。

LIF 是一个逻辑接口，用于将 SVM 连接到物理端口。尽管多个 SVM 可以在同一端口上具有多个 LIF，但一个 LIF 属于一个 SVM。您可以通过 SVM LIF 访问 LUN。

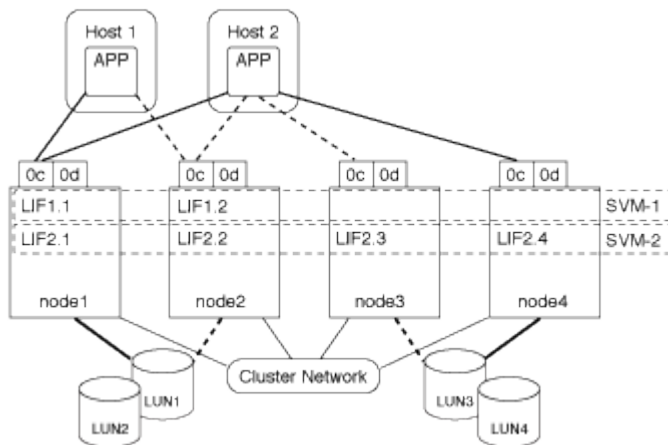
集群中单个 SVM 的 LUN 访问示例

在以下示例中，主机 1 连接到 SVM-1 中的 LIF1.1 和 LIF1.2 以访问 LUN1。LIF1.1 使用物理端口 node1：0c，LIF1.2 使用 node2：0c。LIF1.1 和 LIF1.2 仅属于 SVM-1。如果在节点 1 或节点 2 上为 SVM-1 创建了新的 LUN，则它可以使用这些相同的 LIF。如果创建了新的 SVM，则可以使用两个节点上的物理端口 0c 或 0d 创建新的 LIF。



一个集群中多个SVM的LUN访问示例

一个物理端口可以支持多个为不同 SVM 提供服务的 LIF。由于 LIF 与特定 SVM 关联，因此集群节点可以将传入数据流量发送到正确的 SVM。在以下示例中，1 到 4 之间的每个节点都具有一个 SVM-2 LIF，并使用每个节点上的物理端口 0c。主机 1 连接到 SVM-1 中的 LIF1.1 和 LIF1.2 以访问 LUN1。主机 2 连接到 SVM-2 中的 LIF2.1 和 LIF2.2 以访问 LUN2。两个 SVM 共享节点 1 和 2 上的物理端口 0c。SVM-2 具有其他 LIF，主机 2 使用这些 LIF 访问 LUN 3 和 4。这些 LIF 正在节点 3 和 4 上使用物理端口 0c。多个 SVM 可以共享节点上的物理端口。



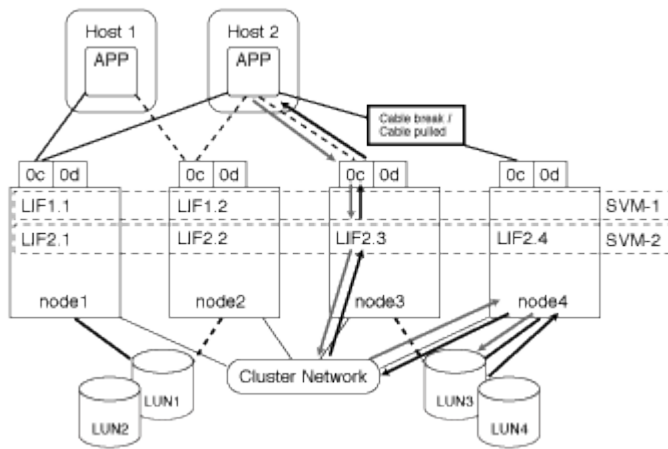
从主机系统到LUN的活动或优化路径示例

在活动路径或优化路径中，数据流量不会通过集群网络传输，而是通过最直接的路由传输到 LUN。到 LUN1 的活动或优化路径通过 node1 中的 LIF1.1，并使用物理端口 0c。主机 2 有两个活动路径或优化路径，一个路径指向节点 1 LIF2.1，该节点共享物理端口 0c，另一个路径指向节点 4 LIF2.4，该节点使用物理端口 0c。



从主机系统到LUN的活动或非优化路径(间接)示例

在主动或非优化路径（间接）路径中，数据流量通过集群网络传输。只有当主机中的所有活动或优化路径均不可用于处理流量时，才会发生此问题描述。如果从主机 2 到 SVM-2 LIF2.4 的路径丢失，则对 LUN3 和 LUN4 的访问将遍历集群网络。从主机 2 访问时，会在节点 3 上使用 LIF2.3。然后，流量将进入集群网络交换机并备份到 node4 以访问 LUN3 和 LUN4。然后，它会通过集群网络交换机返回，然后通过 LIF2.3 返回到主机 2。在节点 4 上的另一个物理端口上为 SVM-2 还原 LIF2.4 路径或建立新的 LIF 之前，系统将使用此活动路径或非优化路径。



= :allow-uri-read:

提高 ESX 主机的 VMware VAAI 性能

如果 ESX 主机运行的是 ESX 4.1 或更高版本，则 ONTAP 支持某些 VMware vStorage APIs for Array Integration（VAAI）功能。这些功能有助于将操作从 ESX 主机卸载到存储系统，并提高网络吞吐量。ESX 主机在正确的环境中自动启用这些功能。

VAAI 功能支持以下 SCSI 命令：

- EXTENDED_COPY

通过此功能，主机可以在 LUN 之间或 LUN 内部启动数据传输，而无需主机参与数据传输。这样可以节省 ESX CPU 周期并提高网络吞吐量。扩展副本功能也称为“副本卸载”，用于克隆虚拟机等情形。ESX 主机调用副本卸载功能时，该功能会复制存储系统中的数据，而不是通过主机网络进行复制。副本卸载可通过以下方式传输数据：

- 在 LUN 中
- 卷中的 LUN 之间
- 在 Storage Virtual Machine （ SVM ）中不同卷上的 LUN 之间
- 在集群中不同SVM上的LUN之间 如果无法调用此功能、ESX主机将自动使用标准读写命令执行复制操作。

- WRITE_SAME

此功能可减轻向存储阵列写入重复模式（例如全零）的工作负担。ESX 主机在零填充文件等操作中使用此功能。

- COMPARE_AND_WRITE

此功能可绕过某些文件访问并发限制，从而加快启动虚拟机等操作的速度。

使用 VAAI 环境的要求

VAAI 功能是 ESX 操作系统的一部分，设置正确的环境后，ESX 主机会自动调用这些功能。

环境要求如下：

- ESX 主机必须运行 ESX 4.1 或更高版本。
- 托管 VMware 数据存储库的 NetApp 存储系统必须运行 ONTAP 。
- （仅限副本卸载） VMware 副本操作的源和目标必须托管在同一集群中的同一存储系统上。



当前，副本卸载功能不支持在不同存储系统上托管的 VMware 数据存储库之间复制数据。

确定 ESX 是否支持 VAAI 功能

要确认 ESX 操作系统是否支持 VAAI 功能，您可以检查 vSphere Client 或使用任何其他方式访问主机。默认情况下，ONTAP 支持 SCSI 命令。

您可以检查 ESX 主机的高级设置以确定是否已启用 VAAI 功能。下表说明了哪些 SCSI 命令与 ESX 控制名称相对应。

SCSI命令	ESX 控制名称（ VAAI 功能）
extended_copy	HardwareAcceleratedMove
write_same	HardwareAcceleratedInit
compare 和 _write	HardwareAcceleratedLocking

Microsoft 卸载数据传输（ Offloaded Data Transfer ， ODX ）

Microsoft 卸载数据传输（ Offloaded Data Transfer ， ODX ）也称为 *copy offload* ，可在

存储设备内部或兼容存储设备之间直接传输数据，而无需通过主机计算机传输数据。

ONTAP支持对SMB和SAN协议使用ODX。

在非 ODX 文件传输中，数据将从源读取，并通过网络传输到主机。主机通过网络将数据传输回目标。在 ODX 文件传输中，数据会直接从源复制到目标，而无需通过主机。

由于ODX卸载副本是直接在源和目标之间执行的、因此、如果在同一卷中执行副本、则可以获得显著的性能优势、包括加快相同卷副本的复制速度、降低客户端上的CPU和内存利用率以及降低网络I/O带宽利用率。如果副本跨卷、则与基于主机的副本相比、性能可能不会显著提高。

对于 SAN 环境，只有当主机和存储系统都支持 ODX 时，ODX 才可用。支持 ODX 且启用了 ODX 的客户端计算机在移动或复制文件时会自动透明地使用卸载文件传输。无论您是通过 Windows 资源管理器拖放文件还是使用命令行文件复制命令，还是客户端应用程序启动文件复制请求，系统都会使用 ODX。

使用 ODX 的要求

如果您计划使用 ODX 进行副本卸载，则需要熟悉卷支持注意事项，系统要求和软件功能要求。

要使用 ODX，您的系统必须具备以下条件：

- ONTAP

ODX 会在支持的 ONTAP 版本中自动启用。

- 源卷最小为 2 GB

为了获得最佳性能，源卷应大于 260 GB。

- Windows 客户端上的 ODX 支持

Windows Server 2012 或更高版本以及 Windows 8 或更高版本均支持 ODX。互操作性表包含有关受支持的 Windows 客户端的最新信息。

["NetApp 互操作性表工具"](#)

- ODX 的复制应用程序支持

执行数据传输的应用程序必须支持 ODX。支持 ODX 的应用程序操作包括：

- Hyper-V 管理操作，例如创建和转换虚拟硬盘（VHD），管理 Snapshot 副本以及在虚拟机之间复制文件
- Windows 资源管理器操作
- Windows PowerShell copy 命令
- Windows 命令提示符复制命令 Microsoft TechNet库包含有关Windows服务器和客户端上支持的ODX应用程序的详细信息。

- 如果使用的是压缩卷，则压缩组大小必须为 8K。

不支持 32 K 压缩组大小。

ODX 不适用于以下卷类型：

- 容量小于 2 GB 的源卷
- 只读卷
- "FlexCache 卷"



FlexCache初始卷支持ODX。

- "半厚配置卷"

特殊的系统文件要求

您可以删除在 qtree 中找到的 ODX 文件。除非技术支持要求您删除或修改任何其他 ODX 系统文件，否则不得删除或修改此类文件。

使用 ODX 功能时，系统的每个卷中都存在 ODX 系统文件。这些文件可以对 ODX 传输期间使用的数据进行时间点表示。以下系统文件位于包含 LUN 或将数据卸载到的文件的每个卷的根级别：

- .copy-offload (隐藏目录)
- .tokens (文件位于隐藏的下 .copy-offload 目录)

您可以使用 `copy-offload delete-tokens -path dir_path -node node_name` 用于删除包含 ODX 文件的 qtree 的命令。

ODX 的用例

您应了解在 SVM 上使用 ODX 的使用情形，以便确定 ODX 在何种情况下可为您带来性能优势。

支持 ODX 的 Windows 服务器和客户端会使用副本卸载作为在远程服务器之间复制数据的默认方式。如果 Windows 服务器或客户端不支持 ODX，或者 ODX 副本卸载在任何时刻失败，则复制或移动操作将回退为复制或移动操作的传统读写操作。

以下使用情形支持使用 ODX 副本和移动：

- 卷内

源文件或 LUN 与目标文件或 LUN 位于同一个卷中。

- 卷间，同一节点，同一 SVM

源文件或 LUN 和目标文件或 LUN 位于同一节点上的不同卷上。数据属于同一个 SVM。

- 卷间，不同节点，相同 SVM

源文件或 LUN 和目标文件或 LUN 位于不同节点上的不同卷上。数据属于同一个 SVM。

- SVM 间，同一节点

源和目标文件或 LUN 位于同一节点上的不同卷上。数据属于不同的 SVM。

- SVM 间，不同节点

源和目标文件或 LUN 位于不同节点上的不同卷上。数据属于不同的 SVM。

- 集群间

源 LUN 和目标 LUN 位于集群中不同节点上的不同卷上。只有 SAN 支持此功能、SMB 不支持此功能。

还有一些其他特殊使用情形：

- 在 ONTAP ODX 实施中，您可以使用 ODX 在 SMB 共享与 FC 或 iSCSI 连接的虚拟驱动器之间复制文件。

您可以使用 Windows 资源管理器，Windows 命令行界面或 PowerShell，Hyper-V 或其他支持 ODX 的应用程序，在 SMB 共享和连接的 LUN 之间使用 ODX 副本卸载功能无缝复制或移动文件，但前提是 SMB 共享和 LUN 位于同一集群上。

- Hyper-V 还提供了一些 ODX 副本卸载的其他使用情形：

- 您可以使用 ODX 副本卸载直通与 Hyper-V 在虚拟硬盘（VHD）文件内部或之间复制数据，或者在同一集群中映射的 SMB 共享和连接的 iSCSI LUN 之间复制数据。

这样，子操作系统中的副本就可以传递到底层存储。

- 创建固定大小的 VHD 时，ODX 用于使用众所周知的置零令牌以零初始化磁盘。
- 如果源存储和目标存储位于同一集群上，则使用 ODX 副本卸载进行虚拟机存储迁移。



要利用 Hyper-V ODX 副本卸载直通的使用情形，子操作系统必须支持 ODX，而子操作系统的磁盘必须是 SCSI 磁盘，并由支持 ODX 的存储（SMB 或 SAN）提供支持。子操作系统上的 IDE 磁盘不支持 ODX 直通。

版权信息

版权所有 © 2024 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。