



Proxmox VE 工作负载的业务连续性 NetApp virtualization solutions

NetApp
October 23, 2026

This PDF was generated from <https://docs.netapp.com/zh-cn/netapp-solutions-virtualization/proxmox/proxmox-business-continuity.html> on October 23, 2026. Always check docs.netapp.com for the latest.

目录

Proxmox VE 工作负载的业务连续性	1
了解使用 SnapMirror 主动同步的 Proxmox VE 业务连续性	1
客户挑战	1
解决方案概述	2
解决方案如何应对故障	5
在 SnapMirror 活动同步和 MetroCluster 之间进行选择	6
Proxmox VE 延伸集群的 Corosync 注意事项	9
与 Storage Manager for Proxmox 集成	10
具有适用于 NAS 的 SnapMirror 活动同步功能的 NFS 共享存储	11
相关信息	15
使用 NAS 存储配置 Proxmox VE 业务连续性	15
步骤 1: 配置调解器	16
步骤 2: 使用 SnapMirror 活动同步保护 NFS SVM	17
步骤 3: 为第 3 层网络配置带有 BGP 的 VIP LIF	18
步骤 4: 使用 QDevice 创建 Proxmox VE stretch 集群	20
第 5 步: 配置受保护的 NFS 共享存储	21
第 6 步: 配置 HA 和 VM 备份	22
第 7 步: 测试计划的故障转移	23
相关信息	24
配置和使用 Storage Manager for Proxmox	24
了解 Storage Manager for Proxmox	24
安装和设置 Storage Manager for Proxmox	26
为 Proxmox 配置 Storage Manager	27
配置存储	32
保护虚拟机	43
还原虚拟机或数据存储库	48
克隆虚拟机	51
监控操作	54
摘要	55

Proxmox VE 工作负载的业务连续性

了解使用 SnapMirror 主动同步的 Proxmox VE 业务连续性

在 Proxmox 虚拟环境（VE）上运行关键业务虚拟机（VM）的企业，以及在 Proxmox VE 上托管客户 VM 的服务提供商，都需要这些 VM 在存储系统或整个数据中心离线时继续运行。ONTAP SnapMirror 活动同步会同步复制跨两个站点的 Proxmox VE 集群的共享存储，并自动将存储故障转移到正常运行的站点。您可以获得零恢复点目标（RPO）、无需脚本或 runbook 的自动恢复，以及两个站点上相同的共享存储，因此 Proxmox VE 高可用性（HA）可以在不进行任何存储重新配置的情况下重新启动 VM。对于 NFS 共享存储，该解决方案使用用于 NAS 的 SnapMirror 活动同步。

客户挑战

仅使用原生工具为 Proxmox VE 构建存储和站点弹性非常困难：

- *站点故障后数据丢失。*Proxmox VE 存储复制和 Proxmox Backup Server 是异步的。主站点发生故障时，自上次复制或备份以来所做的更改将丢失。
- *手动，恢复速度慢。*在辅助站点恢复通常意味着提升副本，在不同地址下挂载存储，更新 `/etc/pve/storage.cfg`，以及重新启动虚拟机。每个手动步骤都会增加停机时间以及在中断期间出错的风险。
- *与虚拟机竞争主机资源的延伸存储。*超融合延伸存储（例如处于延伸模式的 Ceph）在 Proxmox VE 节点上运行，并在每个站点保留每个对象的两个副本。复制、恢复和重新平衡流量会消耗运行虚拟机的相同节点上的 CPU、内存和网络资源。
- *必须保持一致的两个仲裁决策。*任何双站点设计都需要在第三个故障域中为 Proxmox VE 集群和存储提供见证节点。如果集群和存储在站点间链路故障后选择了不同的幸存站点，则 VM 将在无法访问活动存储的节点上重新启动。
- *必须在两个站点上都可用的存储地址。*在存储故障切换后，Proxmox VE 节点必须通过相同地址访问共享存储。该地址必须在两个站点上都访问，这通常意味着在站点之间延伸的第 2 层 VLAN。

服务提供商面临的挑战

在 Proxmox VE 上构建云服务的服务提供商也面临同样的挑战，此外还要满足租户服务级别协议（SLA）：

- *存储系统是共享的故障点。*Proxmox VE HA 保护租户虚拟机免受主机故障的影响，但不能使其免受托管其磁盘的存储系统故障的影响。一次存储中断会影响该系统上的每个租户。
- *无法销售更高的可用性。*如果没有同步复制和自动存储故障转移，提供商就无法提供零数据丢失和更强正常运行时间承诺的高级可用性层。
- *恢复过程中的租户隔离。*恢复、维护和灾难恢复测试必须仅影响相关租户，而不是整个平台。

企业挑战

将关键任务工作负载从其他虚拟机管理程序转移到 Proxmox VE 的企业希望拥有与以前相同的站点弹性，例如数据库和业务线应用程序的零数据丢失以及无需手动运行手册的恢复。仅使用 Proxmox VE 原生工具满足这一期望就需要大量的设计、脚本和运维工作。

解决方案概述

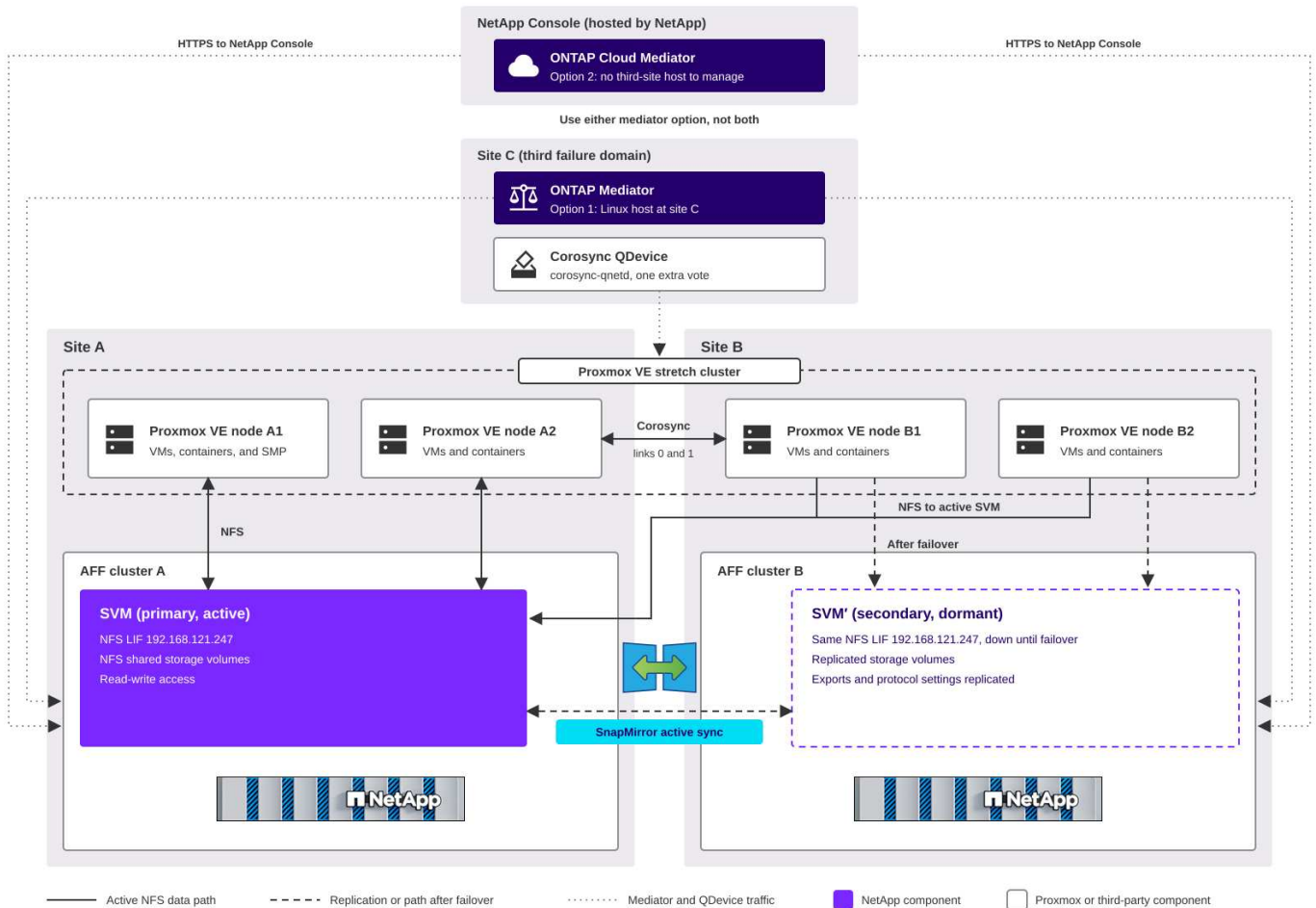
该解决方案将 Proxmox VE 延伸集群与 SnapMirror 主动同步（位于两个站点的 NetApp AFF 系统上）相结合：

- *零 RPO。*ONTAP 将对共享存储的每次写入从站点 A 的主集群同步复制到站点 B 的辅助集群。
- *两个站点使用相同的共享存储。*故障转移后，辅助集群向 Proxmox VE 提供相同的存储，因此 `/etc/pve/storage.cfg` 中的存储定义保持有效。
- *自动故障转移。*调解器在两个站点之外充当见证者。您可以在第三个站点安装 ONTAP Mediator，也可以使用托管在 NetApp Console 中的 ONTAP Cloud Mediator 服务。当主存储或站点发生故障时，ONTAP 会故障转移到站点 B，而无需人工干预。
- *精细控制。*您可以选择要保护的工作负载，并且可以单独故障转移工作负载，例如移动一个租户或应用程序进行站点维护。
- *二级副本的使用。*您可以克隆二级副本以进行开发、测试或灾难恢复测试，而不会影响复制。

解决方案如何应对挑战

挑战	解决方案如何提供帮助	您仍在计划的内容
数据丢失和手动恢复	同步复制可实现零 RPO。ONTAP 会自动进行故障转移，并且辅助 SVM 会提供相同的 NFS 地址，因此 <code>/etc/pve/storage.cfg</code> 保持有效。	Proxmox VE HA 仍会从故障站点重新启动虚拟机。
主机上的延伸存储	复制在 AFF 控制器上运行，而不是在 Proxmox VE 节点上运行，因此主机 CPU、内存和网络仍然可用于虚拟机。ONTAP 在每个站点保留一个副本，存储效率适用于两个副本。	与任何同步复制一样，解决方案需要在站点 B 提供完整的第二个副本，并且复制会向两个 AFF 群集添加负载。为复制工作负载调整两个群集的大小。
两个仲裁决定	ONTAP Mediator 或 ONTAP Cloud Mediator 作为 NetApp Console 中的服务，是数据路径之外的轻量级见证。此解决方案演示如何将 corosync QDevice 与主集群对齐，以便两者都选择相同的站点。	QDevice 仍然需要第三个位置。ONTAP Mediator 也需要它，但 ONTAP Cloud Mediator 不需要。请参阅 [调解器选项] 。
两个站点的存储地址	辅助 SVM 接管相同的 NFS LIF 地址，因此无需重新挂载或重新配置存储。	NFS 子网必须在两个站点都可访问，通常使用延伸的第 2 层 VLAN。有关 VIP LIF 的资格状态，请参阅 具有 VIP LIF 和 BGP 的第 3 层网络 。

下图显示了使用 NFS 共享存储的解决方案架构。



站点 A 和站点 B

每个站点都有一个 AFF 双节点集群和一半的 Proxmox VE 节点。站点 A 的主集群为所有 Proxmox VE 节点提供共享存储。站点 B 的节点通过站点间网络访问主集群。使用 NFS 存储，站点 B 的辅助 SVM 保持休眠状态，直到发生故障转移。

站点 C

第三个故障域托管 corosync QDevice 守护进程(corosync-qnetd)，如果选择选项 1，还托管 ONTAP Mediator。QDevice 提供额外的投票，以便 Proxmox VE 集群在一个站点出现故障时保持法定人数。

NetApp Console

如果选择选项 2，ONTAP Cloud Mediator 将在 NetApp Console 中作为服务运行，该控制台由 NetApp 托管和管理。两个 AFF 集群均通过 HTTPS 连接到它，因此您无需安装或维护中介主机。

调解器

ONTAP 调解器或 ONTAP 云调解器可帮助 ONTAP 确定由哪个集群提供数据。请使用其中之一，而不是同时使用两者。有关比较，请参见 [\[调解器选项\]](#)。

Proxmox VE 延伸集群

所有 Proxmox VE 节点都属于一个集群，并共享相同的存储定义。Proxmox VE HA 在幸存节点上重新启动来自故障站点的虚拟机。Storage Manager for Proxmox (SMP) 在集群上作为 HA 虚拟机运行，以配置共享存储并备份、还原和克隆虚拟机。

调解器选项

SnapMirror 主动同步使用调解器作为被动见证。在网络分区或一个集群发生故障期间，ONTAP 使用调解器来决定哪个副本继续提供 I/O 服务。调解器不在数据路径中，也不会自行启动故障转移。您可以使用以下选项之一，但不能同时使用这两个选项。

注意事项	选项 1: ONTAP Mediator	选项 2: ONTAP Cloud Mediator
运行位置	您在第三个故障域（例如站点 C）中安装和管理的具有独立电源和网络基础架构的 Linux 主机。	NetApp Console 中的云服务。NetApp 托管和管理该服务，因此您不需要在第三个站点配置中介主机。
ONTAP 版本	适用于 NAS 的 SnapMirror 活动同步的 ONTAP 9.19.1 或更高版本。	ONTAP Cloud Mediator 从 ONTAP 9.17.1 开始提供。用于 NAS 的 SnapMirror 活动同步需要 ONTAP 9.19.1 或更高版本。
集群连接	这两个集群都连接到调解器主机。默认服务端口为 31784。	这两个集群都直接或通过 HTTP 代理，通过 HTTPS 出站连接到 NetApp Console。到 NetApp Console 的建议 ping 延迟低于 200 毫秒。
设置	在 Linux 主机上安装该软件包，然后使用主机地址和调解器用户将其添加到集群中。	创建具有 ONTAP Mediator Setup Role 的 NetApp Console 服务帐户，然后使用组织 ID、客户端 ID 和客户端密钥将 Mediator 添加到集群中。
操作	您对主机进行修补和升级，并保持其可用性。	没有要修补或升级的主机。您依赖于从两个站点访问互联网以及 NetApp Console 服务的可用性。
最佳匹配	具有第三个站点或不允许存储集群访问 Internet 的环境。	没有第三个站点用于部署调解器，或者希望避免运行调解器的环境。

选择时请记住以下几点：

- *QDevice 仍然需要第三个位置。*ONTAP 云中介移除了中介主机，而不是 corosync QDevice 主机。QDevice 需要一个所有 Proxmox VE 节点都可以在 TCP 端口 5403 上访问的小型 Linux 主机，并且它没有 corosync 的低延迟要求。
- *两个群集都需要访问 NetApp Console。*如果群集使用代理，请在两个群集上安装代理的根 CA 证书。NetApp 建议为每个群集使用不同的代理服务器。
- *使用任一选项计划相同的故障行为。*当中介程序不可用时，复制将继续而不会丢失数据，但在中介程序再次可用之前，无法进行自动和计划的故障转移。使用 ONTAP Cloud Mediator，这包括从集群到 NetApp Console 的连接断开。



本节中的调解器选项适用于 SnapMirror 活动同步。对于 MetroCluster，此解决方案使用 ONTAP Mediator 或 MetroCluster Tiebreaker。在部署之前，请确认您的 ONTAP 版本、调解器选项和 Proxmox VE 版本一起列在 "[NetApp 互操作性矩阵工具 \(IMT\)](#)" 中。

服务提供商的优势

- *虚拟机可用性超出了主机 HA。*当托管其磁盘的存储系统出现故障时，租户虚拟机会继续运行。ONTAP 故障转移到第二个 AFF 集群，并且虚拟机在不重启的情况下恢复相同共享存储上的 I/O。

- *高级可用性层。*您可以提供零 RPO、具有存储弹性的虚拟机服务作为附加服务，同时在相同集群中未受保护的存储上提供标准层。
- *每租户保护和故障转移。*通过为每个租户或服务层提供单独的受保护存储，您可以独立地为每个租户进行保护、故障转移和灾难恢复测试，而不会影响其他租户。
- *更好地利用这两个站点。*您可以将站点 A 的某些工作负载设置为主工作负载，将站点 B 的其他工作负载设置为主工作负载，以便两个数据中心都能处理活动工作负载。

对企业的益处

- 在存储或站点故障期间，业务关键型虚拟机零数据丢失。
- 自动恢复，无需脚本、存储重新挂载或更改 `/etc/pve/storage.cfg`。
- 内置于 **ONTAP**，Proxmox VE 节点上没有额外的软件层。
- *从一个界面进行日常操作。*Storage Manager for Proxmox 提供受保护的共享存储，并使用 ONTAP 快照和 FlexClone 技术备份、恢复和克隆虚拟机。

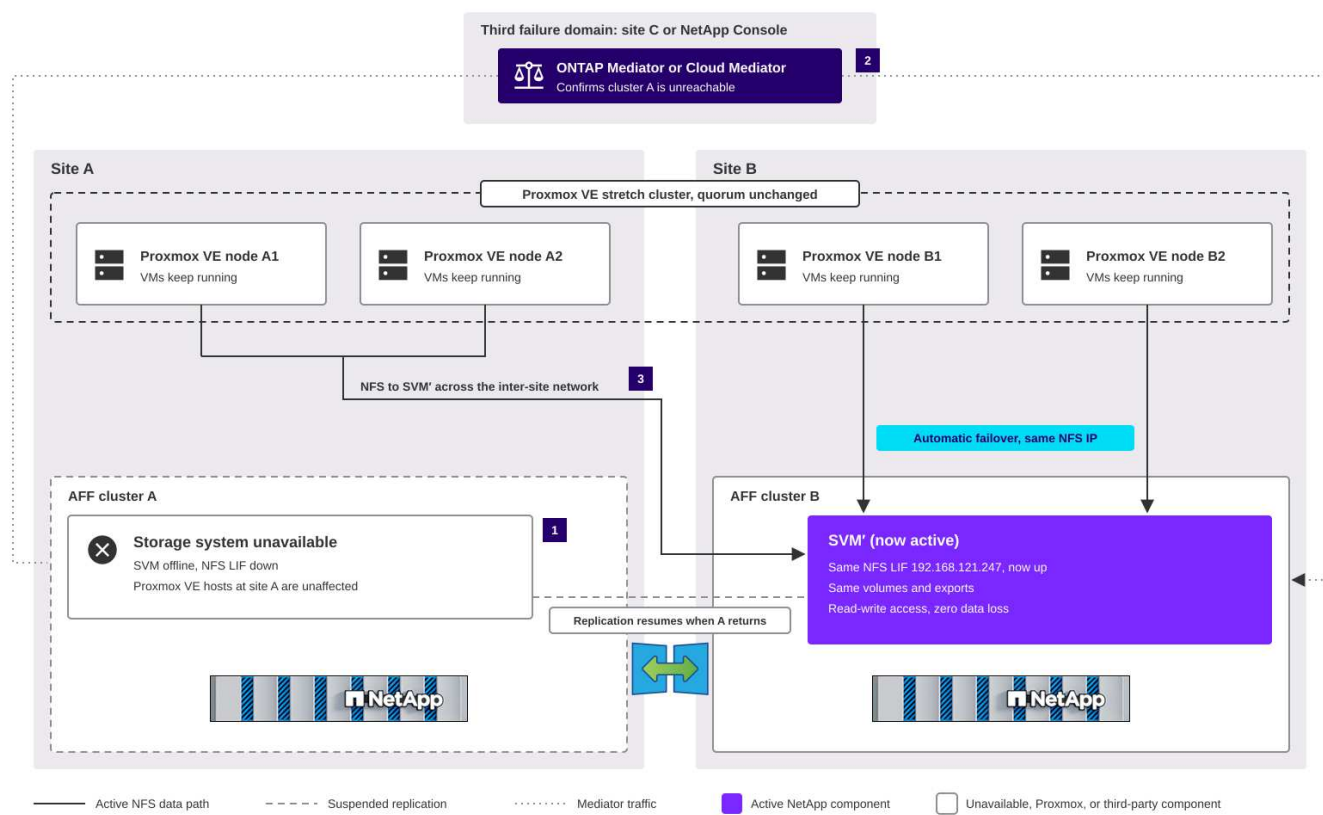
解决方案如何应对故障

ONTAP 和 Proxmox VE 做出独立的法定人数决定。该设计使它们保持一致，以便两者选择相同的幸存站点。

事件	ONTAP 行为	Proxmox VE 行为
站点 A 故障	调解器确认故障，ONTAP 自动将共享存储故障转移到集群 B。	站点 B 节点和 QDevice 保持仲裁。HA 将站点 A 节点隔离，并在站点 B 节点上重新启动其虚拟机。已在站点 B 的虚拟机在存储故障转移后恢复 I/O。
AFF 集群 A 发生故障，站点 A 的 Proxmox VE 节点继续运行	调解器确认集群 A 不可达，ONTAP 自动将共享存储故障转移到集群 B。	没有仲裁更改，也不会重新启动虚拟机。所有节点上的虚拟机在存储故障转移后恢复同一共享存储上的 I/O，站点 A 节点通过站点间网络到达群集 B。
A 站点的一个 AFF 控制器出现故障	配对控制器在本地 HA 接管中接管。未发生 SnapMirror 活动同步故障转移。	无影响。虚拟机在接管期间会短暂暂停 I/O。
站点 B 发生故障	集群 A 继续提供数据。当站点 B 恢复时，复制恢复。	站点 A 节点和 QDevice 保持仲裁。HA 在站点 A 节点上重新启动站点 B 虚拟机。
站点间链路失败	集群 A 继续提供数据。不会发生故障转移。	当仲裁器与主集群对齐时，QDevice 会将其选票投给站点 A 分区。站点 B 节点失去仲裁并自我隔离，HA 会在站点 A 上重新启动其虚拟机。
站点 C 发生故障，或集群无法访问 NetApp Console	复制在零数据丢失的情况下继续进行，但在中介可用之前，无法进行自动和计划的故障转移。使用 ONTAP Cloud Mediator，站点 C 故障仅影响 QDevice。	集群通过四个节点的投票保持法定人数。没有 VM 受到影响。
一个 Proxmox VE 节点发生故障	无影响。	HA 会在其他节点上重新启动该节点上的虚拟机，首选同一站点的节点。

存储系统故障

存储系统故障是主机级 HA 无法单独处理的情况，也是 SnapMirror 活动同步为服务提供商带来最大价值的情况。下图显示了在使用 NFS 共享存储的情况下，AFF 集群 A 发生故障而每个 Proxmox VE 节点保持在线后的状态。



1. AFF 集群 A 变得不可用。站点 A 的 Proxmox VE 节点继续运行，因此集群不会丢失仲裁，HA 也不会隔离任何节点。
2. 调解器，即站点 C 的 ONTAP 调解器或 NetApp Console 中的 ONTAP 云调解器，确认集群 A 不可访问，并且 ONTAP 在集群 B 上激活具有相同 NFS LIF 地址、卷和导出策略的辅助 SVM。
3. 所有 Proxmox VE 节点都在同一共享存储上恢复 I/O。站点 B 节点在本地到达群集 B，站点 A 节点通过站点间网络到达群集 B。虚拟机在故障转移期间暂停 I/O，并在不重新启动的情况下继续运行。

当群集 A 返回时，复制恢复。在维护窗口期间计划对群集 A 的回切，以便站点 A 虚拟机再次访问本地存储。

站点故障

在站点故障期间，该站点的存储系统和 Proxmox VE 节点都会脱机。ONTAP 将共享存储故障转移到幸存的集群，并且 Proxmox VE HA 在幸存节点上重新启动来自故障站点的虚拟机，如表中所述。

在 SnapMirror 活动同步和 MetroCluster 之间进行选择

NetApp MetroCluster 配置还为 Proxmox VE stretch 集群提供零 RPO 和自动恢复。这两种解决方案在保护内容上有所不同：

- *MetroCluster 保护所有工作负载。*MetroCluster 将基于阵列的集群与同步复制相结合。它将每个镜像聚合中的数据、非易失性内存 (NVRAM) 以及每个 SVM 的配置镜像到伙伴站点，因此无需按工作负载设置即可

保护集群上的所有共享存储。

- *SnapMirror 活动同步可保护一部分工作负载。*您可以选择要保护的 SVM 或卷，并且可以独立地对每个 SVM 或卷进行故障切换。未受保护的工作负载可以共享相同的群集。

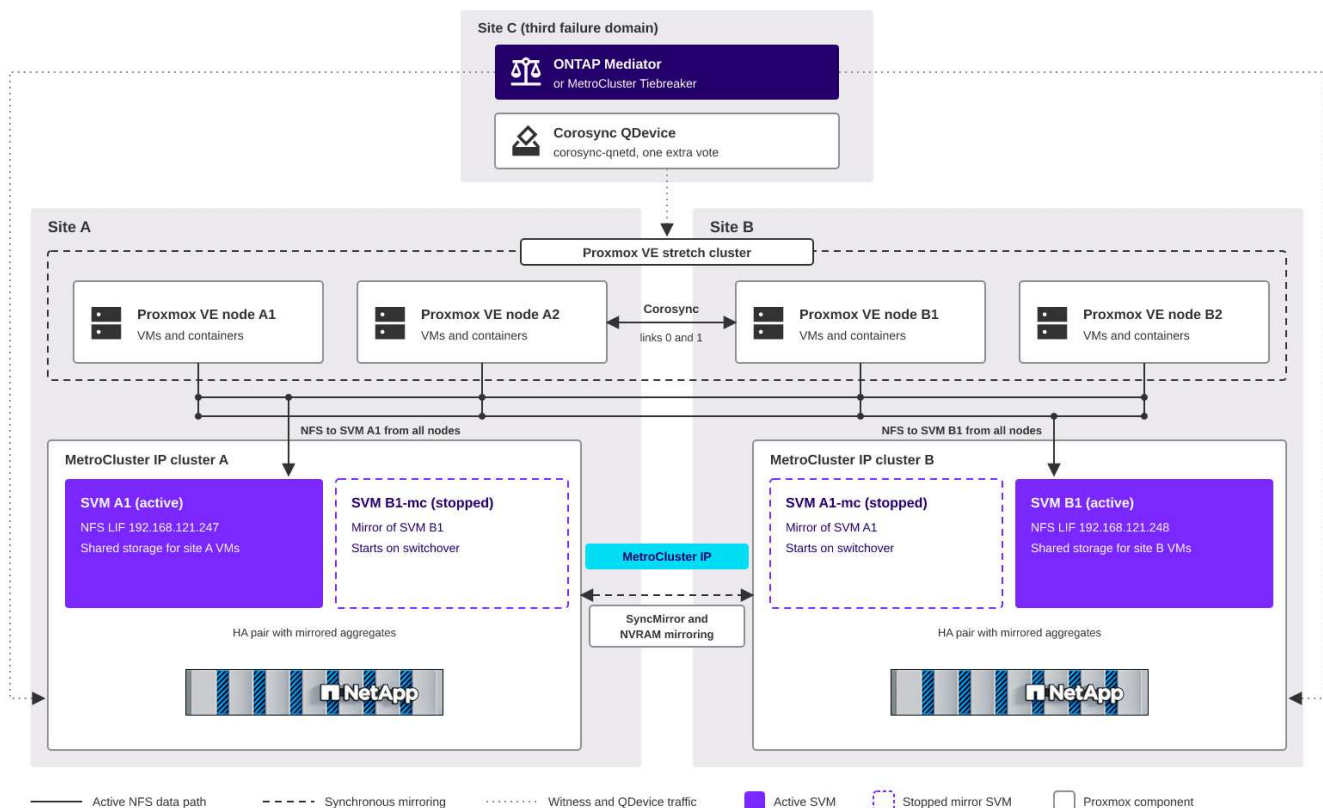
这两种解决方案都会在每个站点上保留受保护数据的完整副本，需要在第三个故障域中有见证，并将复制负载添加到存储控制器。它们在保护范围、故障转移粒度和群集设计方面有所不同。



在部署任一解决方案之前，请验证您的 ONTAP 版本、MetroCluster 或 SnapMirror 活动同步配置以及 Proxmox VE 版本是否一起列在 "[NetApp 互操作性矩阵工具 \(IMT\)](#)" 中。

采用 Proxmox VE 的 MetroCluster 架构

下图显示了具有四节点 MetroCluster IP 配置和 NFS 共享存储的 Proxmox VE stretch 集群。



MetroCluster IP 集群

每个站点都有一个具有镜像聚合的 HA 对。MetroCluster IP 后端网络将数据（SyncMirror）和 NVRAM 同步镜像到伙伴站点。

两个站点的 SVM

每个站点都可以为自己的 SVM 提供服务。对于每个 SVM，MetroCluster 会在配对站点保留一个带有 `~mc` 后缀的已停止镜像 SVM，并具有相同的配置和 LIF 地址。

从所有节点进行 NFS 访问

每个 Proxmox VE 节点都挂载两个活动 SVM（SVM A1 和 SVM B1）的共享存储，以便 HA 可以在任何节点上重新启动任何虚拟机。站点 B 的节点通过站点间网络访问 SVM A1，站点 A 的节点通过站点间网络访问 SVM B1。这与 SnapMirror 活动同步的访问模式相同。

切换

如果站点发生故障，站点 C 的 ONTAP Mediator 或 MetroCluster Tiebreaker 会启用自动切换。幸存站点上的镜像 SVM 启动并使用相同的 NFS 地址提供所有共享存储服务，并且 Proxmox VE HA 会重新启动故障站点上的虚拟机。单个控制器故障由本地 HA 接管处理，无需切换。修复故障站点后，您将修复聚合并切回。

MetroCluster 如何应对故障

下表显示了 MetroCluster IP 配置和 Proxmox VE stretch 集群如何响应与 [\[解决方案如何应对故障\]](#) 中相同的故障。在此示例中，SVM A1 在站点 A 处于活动状态，SVM B1 在站点 B 处于活动状态。

事件	ONTAP 行为	Proxmox VE 行为
站点 A 故障	ONTAP Mediator 或 MetroCluster Tiebreaker 确认故障，ONTAP 切换站点 A。SVM A1-mc 在站点 B 使用相同的 NFS LIF 地址启动。SVM B1 继续提供数据。	站点 B 节点和 QDevice 保持仲裁。HA 将站点 A 节点隔离，并在站点 B 节点上重新启动其虚拟机。已在站点 B 节点上运行的虚拟机继续运行，并且 SVM A1 上的虚拟机在切换后恢复 I/O。
MetroCluster IP 群集 A 发生故障，站点 A 的 Proxmox VE 节点继续运行	ONTAP Mediator 或 MetroCluster Tiebreaker 确认集群 A 不可达，ONTAP 切换站点 A。SVM A1-mc 在站点 B 启动，并使用相同的 NFS LIF 地址。	不会更改仲裁，也不会重新启动虚拟机。SVM A1 上的虚拟机在切换后恢复 I/O，站点 A 节点通过站点间网络到达站点 B。
站点 A 的一个控制器出现故障	站点 A 的 HA 合作伙伴在本地 HA 接管中进行接管。不会发生切换。	无影响。虚拟机在接管期间会短暂暂停 I/O。
站点 B 发生故障	ONTAP 切换站点 B。SVM B1-mc 在站点 A 以相同的 NFS LIF 地址启动。SVM A1 继续提供数据。	站点 A 节点和 QDevice 保持仲裁。HA 在站点 A 节点上重新启动站点 B 虚拟机，这些虚拟机在站点 A 访问 SVM B1。
站点间链路失败	不会发生切换，因为两个群集仍在运行。每个站点继续为其自己的 SVM 提供服务，并暂停向配对站点的镜像，直到链路恢复。	当仲裁器与站点 A 对齐时，QDevice 将其投票权授予站点 A 分区。站点 B 节点失去仲裁并自我隔离。HA 在站点 A 节点上的 SVM A1 上重新启动虚拟机。SVM B1 上的虚拟机在链路恢复之前无法重新启动，因为站点 A 节点无法访问 SVM B1。
站点 C 发生故障	镜像继续进行，零数据丢失，但在见证可用之前无法自动切换。您仍然可以手动切换。	集群通过四个节点的投票保持法定人数。没有 VM 受到影响。
一个 Proxmox VE 节点发生故障	无影响。	HA 会在其他节点上重新启动该节点上的虚拟机，首选同一站点的节点。

站点间链路故障是与单主站点设计的主要区别。每当两个站点都托管活动 SVM 时，无论是使用 MetroCluster 还是在相反方向上使用两个 SnapMirror active sync SVM，corosync tie breaker 都只能偏向一个站点。活动 SVM 位于另一个站点的虚拟机将保持停机状态，直到链路恢复。为了限制影响，请在单独的物理路径上使用冗余的站点间链路，并将最关键的虚拟机放置在 tie breaker 所偏向站点的 SVM 上。

功能对比

注意事项	SnapMirror 适用于 NAS 的活动同步	MetroCluster IP
保护范围	选定的 SVM 和卷	配置中的所有镜像聚合和所有 SVM

注意事项	SnapMirror 适用于 NAS 的活动同步	MetroCluster IP
故障转移粒度	每个 SVM 独立	整个站点，通过切换
未受保护的工作负载	在相同集群中未受保护的 SVM 上运行	可以使用未镜像的聚合，这些聚合在切换后离线，直到您将其上线
集群设计	两个独立的集群，通过 IP 网络建立对等关系	包含两个配对集群的一种配置，具有专用的 MetroCluster IP 后端网络
站点 C 的见证	ONTAP Mediator 或 ONTAP Cloud Mediator	ONTAP Mediator 或 MetroCluster Tiebreaker
受保护数据的副本	每个站点一个	每个站点一个，位于镜像聚合中
复制路径	基于集群间 LIF 的 SnapMirror 同步	SyncMirror 和 NVRAM 通过 MetroCluster IP 后端网络进行镜像
两个站点的 NFS 地址	延伸的第 2 层 VLAN。具有待认证 BGP 的 VIP LIF	延伸第 2 层 VLAN，或使用 BGP 的 VIP LIF
站点故障后恢复	故障回复每个 SVM	修复聚合，然后切换回站点

甄选指南

- 当存储上的所有或几乎所有工作负载都需要零 RPO 时，请选择 MetroCluster，例如，当企业仅在集群上运行任务关键型虚拟机时，或者服务提供商为每个租户提供相同的可用性层时。您无需按工作负载配置即可获得一个保护域。
- 仅当工作负载的子集需要零 RPO（例如，某些租户的高级层）时，或者需要一次为一个租户或应用程序进行故障转移或测试灾难恢复时，请选择 SnapMirror 主动同步。
- SnapMirror 主动同步不支持 MetroCluster 配置上的 NAS。每对集群使用一种解决方案。
- 此解决方案中的 corosync、QDevice 和 HA 指南适用于两者。使用 MetroCluster 时，将 QDevice 仲裁器与在站点间链路发生故障时继续提供数据的站点对齐，并使用 HA 节点亲和性规则将虚拟机与其活动 SVM 保持在同一站点。
- MetroCluster IP 还支持带有 BGP 的 VIP LIF，用于站点之间的第 3 层网络。

Proxmox VE 延伸集群的 Corosync 注意事项

Proxmox VE 将 corosync 用于集群成员身份和仲裁。Corosync 对站点间网络比 SnapMirror 主动同步更敏感，因此请先规划集群网络。

延迟和网络设计

- SnapMirror 活动同步要求群集间往返时间（RTT）小于 10 毫秒。Proxmox 建议所有节点之间的 corosync 延迟小于 5 毫秒，因此 corosync 要求为解决方案设置了距离限制。
- Corosync 需要低且一致的延迟，而不是高带宽。为 corosync 使用专用网络，将其与存储和实时迁移流量分开，以免存储突发导致 corosync 超时。
- 在站点之间的单独物理路径上配置至少两个 corosync 链路（例如，链路 0 和链路 1）。如果一条路径出现故障，Corosync 会使用 Kronosnet 自动切换链路。

在第三个站点使用 **QDevice** 的法定人数

在站点 A 和站点 B 之间均匀分配 Proxmox VE 节点，并在站点 C 添加 QDevice。凭借四个节点和一个 QDevice，集群拥有五票，需要三票才能达到法定人数，因此任何一个站点发生故障都不会丢失法定人数。QDevice 通过 TCP 端口 5403 连接到集群，并且没有 corosync 的低延迟要求，因此它可以与 ONTAP Mediator 共享站点 C。如果您使用 ONTAP Cloud Mediator，您仍然需要为 QDevice 提供第三个位置，但它可以是一台不托管 Mediator 的小型 Linux 主机。

仲裁器与主集群对齐

当站点间链路发生故障但两个站点仍然可以到达站点 C 时，ONTAP 会让站点 A 的主集群继续提供数据服务。QDevice 必须将其选票投给同一个站点。默认情况下，QDevice 会选择包含最低节点 ID 的分区。

最佳实践：在加入站点 B 节点之前，在站点 A 节点上创建 Proxmox VE 集群并加入其他站点 A 节点，以便站点 A 拥有最低的节点 ID。如果主存储长时间移动到站点 B，则将 QDevice tie_breaker 设置更新为站点 B 节点 ID。

如果 QDevice 转而倾向于站点 B，则站点 A 节点将自我隔离，并且 HA 会在无法访问站点 A 活动存储的站点 B 节点上重新启动虚拟机。

HA 行为和 VM 放置

- Proxmox VE HA 通常会检测到故障节点，并在大约两分钟内恢复其虚拟机，包括 60 秒的监视器隔离超时。在站点故障期间，SnapMirror 活动同步故障转移和 Proxmox VE HA 恢复并行运行。
- 使用具有节点优先级的 HA 节点关联规则，将虚拟机放置在其主存储位于同一站点的节点上，并允许它们故障切换到其他站点。
- 在站点故障切换后，将受影响虚拟机的 HA failback 属性设置为 0，直到存储故障恢复。否则，HA 会在站点 A 节点恢复后立即将虚拟机迁移回这些节点，而活动存储仍在站点 B。
- 确保来宾操作系统磁盘超时时间长于预期的故障转移时间，以便来宾在存储故障转移期间不会报告 I/O 错误。

与 **Storage Manager for Proxmox** 集成

Storage Manager for Proxmox（SMP）为 Proxmox VE 管理员提供了一个界面，可通过使用 ONTAP 快照和 FlexClone 技术来配置 ONTAP 共享存储以及备份、还原和克隆虚拟机。SMP 通过其 API 连接到 Proxmox VE 集群和 ONTAP，因此不需要在 Proxmox VE 节点上安装代理。

SMP 和 SnapMirror 主动同步可防止不同类型的故障。SnapMirror 主动同步可使虚拟机在存储或站点故障期间保持运行，但它会同步复制每次写入，包括意外删除、来宾更新失败或由勒索软件加密的数据。SMP 快照为您提供从这些逻辑故障中恢复的时间点副本。它们结合在一起，可从相同的 ONTAP 存储中提供连续的可用性和快速的操作恢复。

保护层	防止	恢复点	通过以下方式管理
SnapMirror主动同步	存储系统故障和站点故障	零数据丢失	ONTAP
SMP 备份（ONTAP 快照）	意外删除、数据损坏、更新失败和勒索软件	上次计划备份或按需备份	SMP
SnapMirror 异步扇出或 Proxmox Backup Server	两个站点的丢失和长期保留	上次复制或备份	ONTAP 或 Proxmox 备份服务器

备份和恢复

SMP 按需或按策略将崩溃一致性或文件一致性备份创建为存储卷的 ONTAP 快照，并恢复单个虚拟机、存储或资源组：

- 使用资源组和备份策略将相同的计划和保留应用于租户或应用程序的虚拟机。
- 对运行 QEMU 访客代理的虚拟机使用文件一致性备份，以便在 ONTAP 创建快照之前使访客文件系统静默。
- SMP 管理处于活动状态的存储。在 SnapMirror 活动同步故障转移后，在 SMP 中运行发现，以便在运行备份或还原操作之前将存储映射到集群 B。
- 对于两个站点之外的第三个副本，或为了更长的保留时间，请从受保护卷中添加 SnapMirror 异步扇出关系，或使用 Proxmox Backup Server。
- SnapMirror 活动同步不支持防篡改快照，因此请勿在 SMP 中为受保护卷启用快照锁定。

VM 克隆

SMP 使用 ONTAP FlexClone 技术创建节省空间的 VM 克隆，用于开发、测试和快速配置。您在受保护的存储中创建的克隆会消耗两个站点上的复制带宽和容量。将不需要业务连续性的克隆和测试存储从 SnapMirror 活动同步保护中排除。

SMP 放置

- 将 SMP VM 作为 HA 资源在受 SnapMirror 活动同步保护的共享存储上运行，以便在存储或站点发生故障后 SMP 保持可用。
- 在 SMP 中注册 Proxmox VE 集群和两个 ONTAP 集群，以便 SMP 可以访问任一站点的活动存储。

具有适用于 NAS 的 SnapMirror 活动同步功能的 NFS 共享存储

从 ONTAP 9.19.1 开始，SnapMirror 主动同步保护 NAS 工作负载。以下注意事项仅适用于 NFS 共享存储。要使用 NFS 存储部署解决方案，请按照["NAS 配置工作流"](#)进行操作。

SVM 级保护和网络身份

- 用于 NAS 的 SnapMirror 活动同步可保护整个存储虚拟机（SVM）。您可以保护 SVM 中的所有卷或排除选定的卷。仅支持 FlexVol 卷，并且 SVM 不能包含 LUN。
- 辅助 SVM 保持休眠状态，其 LIF 处于关闭状态，直到发生故障转移。ONTAP 复制 SVM 配置，包括 LIF IP 地址、导出策略和 NFS 设置，因此辅助 SVM 在故障转移后呈现相同的 NFS 端点。
- 对于 NFSv3 及更高版本，故障转移是无中断的。
- SnapMirror NAS 的主动同步是主动/被动的，因此站点 B 节点上的虚拟机通过站点间网络读取和写入主 SVM。对于两个站点的活动工作负载，请使用两个具有相反复制方向的 SVM，每个站点一个主 SVM，并将每个 VM 的节点关联规则与其主 SVM 的站点对齐。
- AFF A 系列和 C 系列双节点集群支持适用于 NAS 的 SnapMirror 主动同步。

NFS 客户端设置

对 NFS 存储使用 NFSv3 或 NFSv4.1。保留默认的硬挂载行为，不要添加 `soft` 挂载选项，以便 NFS I/O 在故障转移期间重试而不是失败。

使用多个 TCP 连接到存储，以获得更好的性能和容错。请使用 nConnect 或会话中继，而不是同时使用两者：

- *nConnect.* 打开到单个 NFS LIF 地址的多个 TCP 连接，例如 `nconnect=4`。nConnect 适用于 NFSv3 和 NFSv4.1，以及单个 VIP LIF 地址，因此是 NAS 的 SnapMirror 主动同步（包括第 3 层网络）的推荐选项。
- *会话中继.* 将多个数据 LIF 分组到一个 NFSv4.1 会话中，例如 `vers=4.1, trunkdiscovery, max_connect=16`。会话中继需要 NFSv4.1 或更高版本。仅当受保护的 SVM 具有客户端可以在两个站点（例如在延伸的第 2 层网络中）访问的多个数据 LIF 时，才使用它。有关详细信息，请参见["NFS 中继"](#)。

有关挂载选项命令和验证步骤，请参见 ["为 Proxmox VE 配置 NFS 存储"](#)。

使用 SMP 进行 NFS 存储资源调配

SMP 配置 ONTAP 卷，创建或选择导出策略，并将该卷作为 NFS 共享存储挂载到您选择的 Proxmox VE 节点上。对于受保护 SVM 中的存储：

- 在主 SVM 中配置卷，并选择 NFS LIF 或第 3 层网络中的 VIP LIF 作为首选 LIF。辅助 SVM 在故障转移后使用相同的地址，因此存储定义保持有效。
- 在两个站点的所有 Proxmox VE 节点上挂载存储，以便 HA 可以在任何节点上重新启动虚拟机。
- 确保导出策略允许两个站点上的 Proxmox VE 节点。ONTAP 将导出策略复制到辅助 SVM。
- 配置后，请验证卷是否包含在 SnapMirror 活动同步保护中。

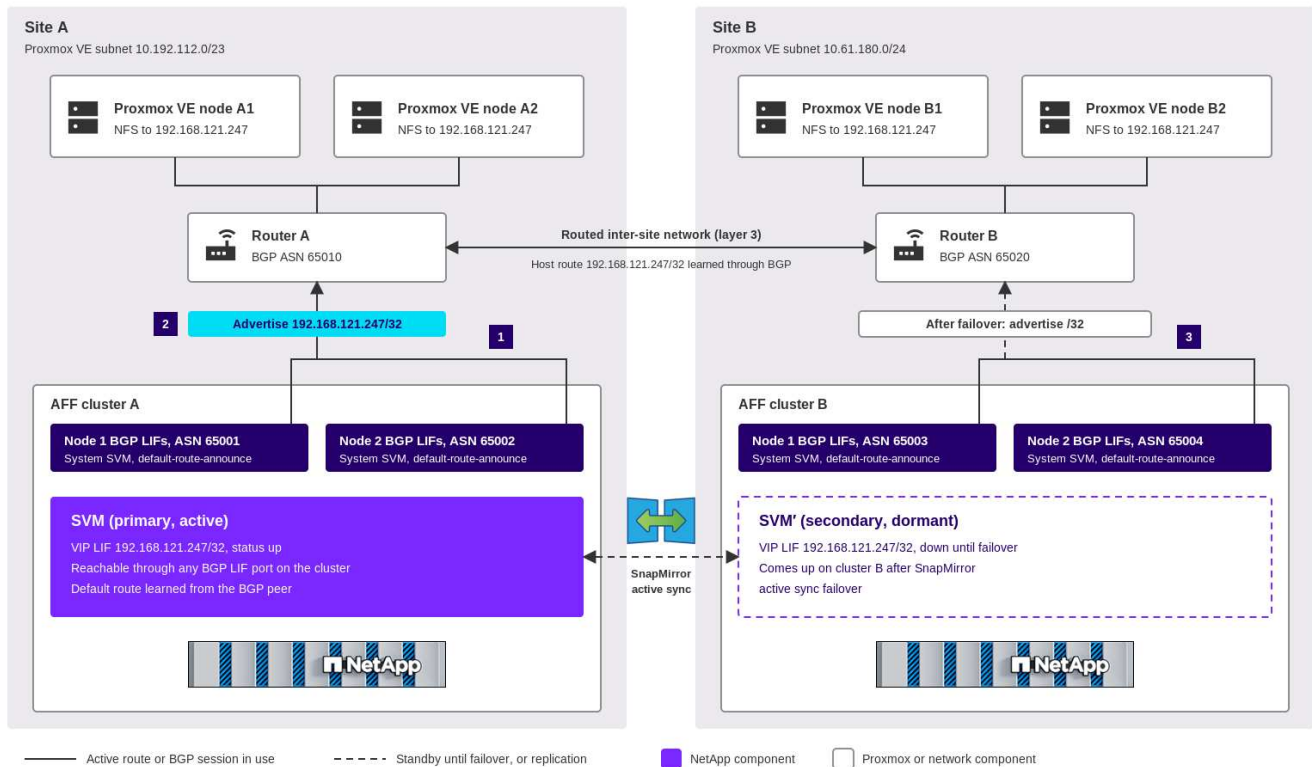
具有 VIP LIF 和 BGP 的第 3 层网络

NAS 的 SnapMirror 主动同步要求主 SVM 和辅助 SVM 上的 IP 地址相同。当 NFS VLAN 在站点之间延伸时，LIF IP 地址会随 SVM 一起移动。当站点通过路由的第 3 层网络连接时，站点 B 上不存在站点 A LIF 的子网。



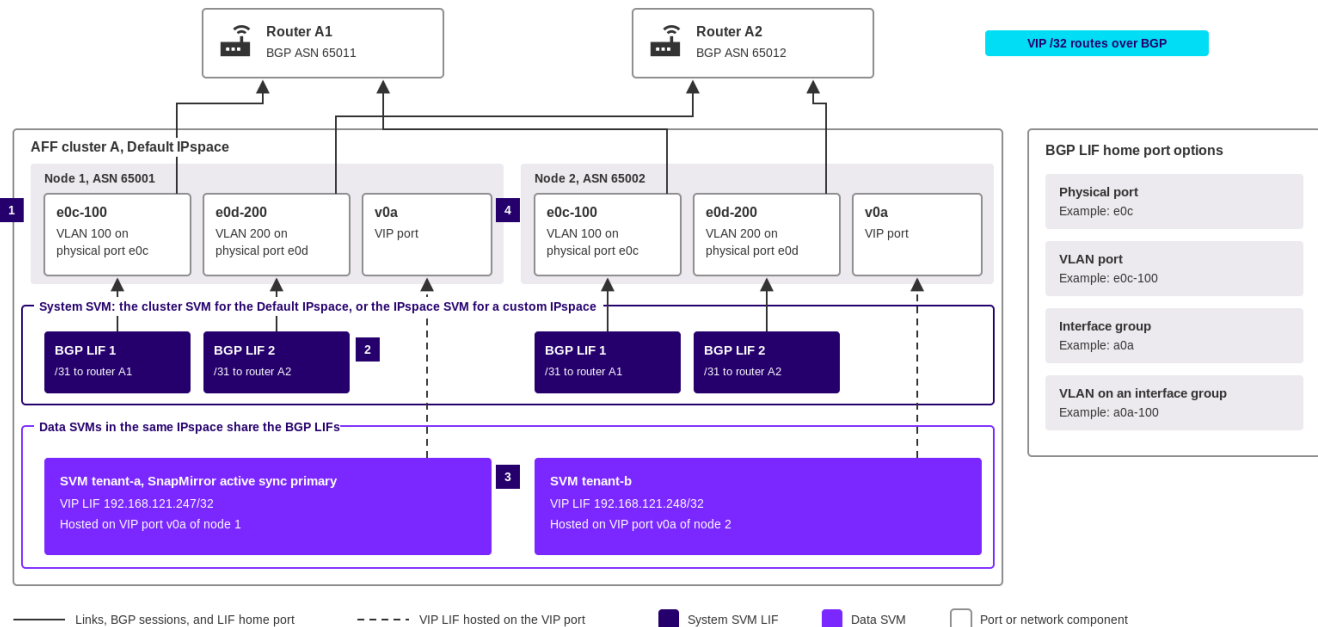
在发布时，NetApp 尚未将具有 BGP 的 VIP LIF 记录为 NAS 的 SnapMirror 活动同步的合格配置。在规划第 3 层设计之前，请向 NetApp 确认当前的支持情况。否则，请在站点之间延伸 NFS VLAN。MetroCluster IP 配置支持具有 BGP 的 VIP LIF。

ONTAP 虚拟 IP (VIP) LIF 解决了这个问题。VIP LIF 不属于任何子网。ONTAP 通过使用边界网关协议 (BGP)，将 VIP 地址作为 /32 主机路由通告给对等路由器，因此该地址可以在没有延伸 VLAN 的情况下在站点之间移动。



1. 两个站点的每个节点在系统 SVM 中都有 BGP LIF，与其所在站点的路由器建立对等关系。BGP LIF 是特定于节点的，不会进行复制，因此您需要在两个集群上配置它们。
2. 主 SVM 托管 NFS VIP LIF。群集 A 通告 /32 主机路由，两个站点的 Proxmox VE 节点通过路由网络到达存储。
3. 在 SnapMirror 主动同步故障转移后，VIP LIF 出现在辅助 SVM 上，群集 B 通告主机路由，路由器在站点 B 上收敛。

首次在 ONTAP 中规划 BGP 时，了解哪个 SVM 拥有每个 LIF 会有所帮助。BGP LIF 属于 IPspace 的系统 SVM，而不是数据 SVM。每个集群上的一组 BGP LIF 为该 IPspace 中每个数据 SVM 的 VIP LIF 提供服务，因此您无需为每个租户创建 BGP LIF。下图显示了一个集群上的 LIF 和端口布局，以及 VLAN 端口上的 BGP LIF。



1. 每个节点都有两个 BGP 端口，在本示例中为 VLAN 端口 e0c-100 和 e0d-200。每个端口直接连接到不同的路由器。
2. 每个节点在系统 SVM 中有两个 BGP LIF，每个端口一个，具有一个 /31 子网和一个连接到一个路由器的 BGP 对等组。对于默认 IPspace，系统 SVM 是集群 SVM。对于自定义 IPspace，它是与 IPspace 同名的 SVM。
3. 同一个 IPspace 中的数据 SVM 托管 VIP LIF，并且不需要自己的 BGP LIF。BGP 对等组为 IPspace 中的所有 SVM 通告 VIP 路由。
4. 当您创建第一个 BGP 对等组时，ONTAP 会在每个节点上创建一个 VIP 端口，例如 v0a。VIP LIF 托管在 VIP 端口上，可通过 IPspace 中托管 BGP LIF 的节点上的每个端口进行访问。

BGP LIF 的主端口可以是物理端口、VLAN 端口、接口组或接口组上的 VLAN。不需要接口组。NetApp 技术报告 TR-4949 建议为每个 IPspace 在两个独立的端口上配置两个 BGP LIF，每个端口直接连接到不同的第一跳路由器。

请记住以下准则：

- BGP LIF 不会故障转移。冗余来自每个节点的两个 BGP LIF，每个节点与不同的路由器对等。如果节点丢失了 SVM 的所有 BGP 会话，ONTAP 会将 VIP LIF 移动到具有该 SVM 的活动 BGP 会话的另一个节点。
- 为每个 ONTAP 节点和每个路由器使用具有唯一专用自治系统编号（ASN）的外部 BGP（eBGP）。
- 为每个 BGP LIF 及其路由器使用 /31 子网（网络掩码 255.255.255.254）。将奇数地址分配给 BGP LIF，将偶数地址分配给路由器。
- ONTAP 不会处理来自路由器的路由。将路由器配置为接受来自 ONTAP 的 /32 主机路由，并且不向集群发送路由更新。
- Proxmox VE 节点不得与 VIP 地址共享子网。冗余的第 3 层路由器（例如使用 VRRP 或 HSRP）必须位于节点和存储之间的路径中。
- 使用 `-use-peer-as-next-hop true` 配置 BGP 对等组，以便每个集群从其本地路由器学习其默认路由。故障转移后，指向站点 A 网关的路由在站点 B 无效。
- 启用多路径路由，以便传出 VIP 流量使用所有可用的 BGP LIF 端口。

- 将 BGP 保持时间从默认的 180 秒缩短。TR-4949 建议保持时间为 3 秒。只有当保持计时器到期时，路由器才会从出现故障的集群中撤回路由，而在存储系统出现故障期间，站点 A 的路由器将保持在线状态。

相关信息

- ["ONTAP SnapMirror 主动同步架构"](#)
- ["ONTAP SnapMirror 活动同步的先决条件"](#)
- ["准备配置 ONTAP Cloud Mediator"](#)
- ["配置 ONTAP Cloud Mediator 以进行 SnapMirror 活动同步"](#)
- ["虚拟 IP 和 BGP 的 ONTAP MetroCluster 注意事项"](#)
- ["MetroCluster 数据保护和灾难恢复"](#)
- ["配置和使用 Storage Manager for Proxmox"](#)
- ["TR-4949：数据中心中使用 ONTAP 的 BGP/VIP"](#)
- ["Proxmox VE 集群管理器"](#)
- ["Proxmox VE 高可用性"](#)

使用 NAS 存储配置 Proxmox VE 业务连续性

通过将 NAS 的 ONTAP SnapMirror 活动同步与 Proxmox VE 延伸群集相结合，保护 Proxmox 虚拟环境（VE）虚拟机免受站点故障的影响。您可以在 NetApp Console 中配置调解器（ONTAP 调解器或 ONTAP 云调解器），保护 NFS 存储虚拟机（SVM），可选择使用虚拟 IP（VIP）LIF 和边界网关协议（BGP）使其 NFS 地址在路由网络中可移植，使用 QDevice 构建延伸群集，然后使用 Storage Manager for Proxmox 配置共享存储并配置高可用性（HA）和虚拟机备份。结果是零数据丢失以及虚拟机在幸存站点的自动恢复。

关于此工作流

开始之前，请查看["解决方案架构、故障行为和 corosync 设计注意事项"](#)。将尖括号中的值（例如 <primary_svm>）替换为您环境的值。

开始之前

- 两个 AFF A 系列或 C 系列双节点集群必须运行 ONTAP 9.19.1 或更高版本，一个在站点 A，一个在站点 B。
- SnapMirror 和 SnapMirror 同步许可证（包含在 ONTAP One 中）必须安装在两个集群上。
- 集群必须在默认 IPspace 中建立对等关系，且集群间往返时间（RTT）小于 10 毫秒。
- 您必须选择一个调解器选项。有关比较，请参见 ["调解器选项"](#)。
 - 对于 ONTAP Mediator，Linux 主机必须安装在站点 C 或另一个第三故障域，并且可以从两个集群的集群和节点管理 LIF 访问。
 - 对于 ONTAP Cloud Mediator，两个群集都必须直接或通过代理对 NetApp Console 具有出站 HTTPS 访问权限，ping 延迟低于 200 毫秒。您还需要 NetApp Console 组织和权限才能创建服务帐户。
- 站点 A 的 SVM 必须向 Proxmox VE 节点提供 NFS。该 SVM 必须仅包含用于 NAS 的 FlexVol 卷，不包含 LUN，其导出策略必须允许所有 Proxmox VE 节点。

- 必须安装偶数个 Proxmox VE 9.0 或更高版本的节点，在站点 A 和站点 B 之间平均分配，所有节点之间的 corosync 延迟小于 5 毫秒。
- 站点 C 或其他第三个位置的 Linux 主机必须可供 corosync QDevice 守护程序使用，并且所有 Proxmox VE 节点均可通过 TCP 端口 5403 访问该主机。无论使用哪种调解器选项，您都需要此主机。
- NFS VLAN 必须在站点 A 和站点 B 之间延伸。对于路由的第 3 层网络，每个站点都必须有一个支持 BGP 的路由器，并且您必须向 NetApp 确认，NAS 的 SnapMirror 主动同步是否支持带有 BGP 的 VIP LIF。
- 您的 ONTAP 和 Proxmox VE 版本必须一起列在 "[NetApp 互操作性矩阵工具 \(IMT\)](#)" 中。
- 要调配共享存储和备份虚拟机，必须安装 Storage Manager for Proxmox，并确保其具有对 Proxmox VE 节点和两个 ONTAP 集群的网络访问权限。

步骤 1：配置调解器

在初始化 SnapMirror 活动同步关系之前，将中介添加到集群对等关系中。ONTAP 需要中介来执行计划内和自动故障转移。您可以从任一集群添加中介，ONTAP 会自动将其添加到另一个集群上。请使用其中一个中介选项，不要同时使用两个。

从下列选项中选择一项。

ONTAP 调解器

如果在站点 C 的 Linux 主机上安装调解器，请使用此选项。有关安装，请参见 "[安装或升级 ONTAP Mediator](#)"。

步骤

1. 在任一集群上，添加 mediator：

```
snapmirror mediator add -peer-cluster <peer_cluster> -type on-prem
-mediator-address <mediator_ip> -username <mediator_user>
```

出现提示时，输入调解器密码。

2. 验证调解器已连接且仲裁状态为 true：

```
snapmirror mediator show
```

ONTAP Cloud Mediator

使用此选项在 NetApp Console 中将调解器作为服务托管，这样您就不需要在第三个站点使用调解器主机。

开始之前

- 两个站点的防火墙必须允许从两个集群到 NetApp Console API 端点的 HTTPS 流量。有关当前端点主机名，请参见 "[准备配置 ONTAP Cloud Mediator](#)"。
- 如果群集使用代理服务器，请准备好代理 IP 地址、端口和凭据。请为每个群集使用不同的代理服务器。代理的根 CA 证书必须安装在 ONTAP 中。
- ONTAP 包含已知的根 CA 证书，因此通常不需要安装 NetApp Console 证书。要查看，请运行

```
security certificate show -common-name Sectigo*|USERTrust*|AAA*。
```

步骤

1. 在 NetApp Console 中，复制您的组织 ID。
2. 在 NetApp Console 中，创建成员服务帐户：
 - a. 将组织设置为使用 ONTAP Cloud Mediator 的组织。
 - b. 将类别设置为*应用程序*。
 - c. 将角色设置为 **ONTAP Mediator Setup Role**。
 - d. 创建帐户时保存客户端 ID 和客户端密钥。
3. 在任一集群上，添加 mediator：

```
snapmirror mediator add -peer-cluster <peer_cluster> -type cloud  
-org-id <console_org_id> -service-account-client-id <client_id> -use  
-http-proxy-local <true|false> -use-http-proxy-remote <true|false>
```

出现提示时输入客户端密码。

您还可以使用 System Manager。选择 保护 > 概述 > 调解器 > 添加，选择 云 作为调解器类型，然后输入组织 ID、客户端 ID 和客户端密钥。

4. 验证介质类型是否为 cloud，连接状态是否为 connected，仲裁状态是否为 true：

```
snapmirror mediator show
```

步骤 2：使用 SnapMirror 活动同步保护 NFS SVM

在站点 B 集群上创建休眠目标 SVM，并将站点 A SVM 复制到其中。ONTAP 同步复制数据并复制 SVM 配置，包括 NFS LIF 地址和导出策略。您也可以在 System Manager 中完成此步骤，方法是为 SVM 选择 保护 并选择 `AutomatedFailOver` 策略。

步骤

1. 在站点 B 群集上，创建目标 SVM：

```
vserver create -vserver <secondary_svm> -sub-type dp-destination
```

2. 在站点 B 集群上，创建与主 SVM 的 SVM 对等关系：

```
vserver peer create -vserver <secondary_svm> -peer-vserver <primary_svm>  
-peer-cluster <site_a_cluster> -applications snapmirror
```

3. 在站点 A 集群上，接受 SVM 对等请求：

```
vserver peer accept -vserver <primary_svm> -peer-vserver <secondary_svm>
```

4. 在站点 B 群集上，创建 SnapMirror 活动同步关系：

```
snapmirror create -source-path <primary_svm>: -destination-path  
<secondary_svm>: -policy AutomatedFailOver
```

5. 初始化关系：

```
snapmirror initialize -destination-path <secondary_svm>:
```

6. 验证关系已同步：

```
snapmirror show -destination-path <secondary_svm>:
```

`Mirror State` 是 `Snapmirrored`，并且 `Relationship Status` 是 `InSync`。

7. （可选）排除不需要保护的卷，例如暂存卷或测试卷：

```
volume modify -vserver <primary_svm> -volume <volume_name> -smas  
-protection unprotected
```

步骤 3：为第 3 层网络配置带有 BGP 的 VIP LIF



在发布时，NetApp 尚未将具有 BGP 的 VIP LIF 记录为 NAS SnapMirror 活动同步的合格配置。在完成此步骤之前，请向 NetApp 确认当前的支持情况。否则，请在站点之间延伸 NFS VLAN 并跳过此步骤。

如果 NFS VLAN 在站点 A 和站点 B 之间延伸，请跳过此步骤。如果站点通过路由网络连接，请在两个集群的每个节点上创建 BGP LIF，然后在主 SVM 的 VIP LIF 上托管 NFS 地址。VIP 地址不得属于 Proxmox VE 节点的子网。

开始之前

- 必须将每台路由器配置为接受来自其站点的 BGP LIF 的 BGP 会话，接受 /32 主机路由，并且不向 ONTAP 发送路由更新。
- 每个 ONTAP 节点和每个路由器必须具有唯一的私有 ASN，并且每个 BGP LIF 和路由器对必须具有 /31 子网。

步骤

1. 在每个集群上，更改为高级权限级别：

```
set -privilege advanced
```

2. 在两个集群的每个节点上，创建一个 BGP 配置，其中包含该节点唯一的 ONTAP ASN、3 秒的保持时间，并将节点管理 IP 地址作为路由器 ID：

```
network bgp config create -node <node_name> -asn <node_asn> -holdtime 3  
-routerid <node_mgmt_ip>
```

3. 可选地，在两个集群的每个节点上，在连接到路由器的两个端口中的每个端口上创建一个 VLAN 端口：

```
network port vlan create -node <node_name> -vlan-name <port_name>-  
<vlan_id>
```

您还可以使用物理端口、接口组或接口组上的 VLAN 作为 BGP LIF 主端口。端口必须属于 IP 空间中的广播域。

4. 在两个集群的每个节点上，在系统 SVM 中创建两个 BGP LIF，每个端口一个，并带有连接到该端口上路由器的 /31 子网：

```
network interface create -vserver <system_svm> -lif <bgp_lif_name>  
-service-policy default-route-announce -home-node <node_name> -home-port  
<vlan_or_port> -address <bgp_lif_ip> -netmask-length 31
```

对于默认 IPspace，`<system\_svm>` 是集群名称。对于自定义 IPspace，它是 IPspace 名称。您为每个节点创建一次 BGP LIF，它们为 IPspace 中所有数据 SVM 的 VIP LIF 提供服务。

5. 对于每个 BGP LIF，使用其路由器创建一个 BGP 对等组，并从该路由器学习默认路由：

```
network bgp peer-group create -peer-group <peer_group_name> -ipspace  
<ipspace> -bgp-lif <bgp_lif_name> -peer-address <router_ip> -peer-asn  
<router_asn> -use-peer-as-next-hop true
```

6. 验证每个对等组是否已启动：

```
network bgp peer-group show -ipspace <ipspace>
```

7. 在每个集群上，启用多路径路由：

```
network options multipath-routing modify -is-enabled true
```

8. 在站点 A 集群上，在主 SVM 中创建 NFS VIP LIF：

```
network interface create -vserver <primary_svm> -lif <vip_lif_name>  
-role data -data-protocol nfs -home-node <node_name> -address  
<vip_address> -is-vip true -failover-policy broadcast-domain-wide
```

9. 在站点 A 集群上，验证每个节点上的主 SVM 的 BGP 状态是否为 up：

```
network bgp vserver-status show
```

10. 在两台路由器上，验证 VIP 地址的 /32 路由是从群集 A 获知的。

步骤 4：使用 QDevice 创建 Proxmox VE stretch 集群

从站点 A 节点创建集群，使站点 A 节点接收最低的节点 ID，并在站点 C 添加 QDevice。当站点间链路发生故障时，QDevice 会将投票权授予站点 A，这与 ONTAP 保持活动状态的 SVM 相匹配。

步骤

1. 在第一个站点 A 节点上，在单独的网络上创建具有两条 corosync 链路的集群：

```
pvecm create <cluster_name> --link0 <node_link0_ip> --link1  
<node_link1_ip>
```

2. 在每个剩余的站点 A 节点上，然后在每个站点 B 节点上，加入集群：

```
pvecm add <first_node_ip> --link0 <node_link0_ip> --link1  
<node_link1_ip>
```

3. 验证站点 A 节点是否具有最低的节点 ID：

```
pvecm nodes
```

4. 在站点 C 或其他第三位置的 QDevice 主机上，安装 QDevice 守护进程：

```
apt install corosync-qnetd
```

5. 在每个 Proxmox VE 节点上，安装 QDevice 客户端：

```
apt install corosync-qdevice
```

6. 在一个 Proxmox VE 节点上，将 QDevice 添加到集群：

```
pvecm qdevice setup <qdevice_ip>
```

7. 验证集群是否通过 QDevice 投票达到法定人数：

```
pvecm status
```

对于四节点集群，Expected votes 为 5，Quorum 为 3，并且 Flags 包括 Quorate Qdevice。

第 5 步：配置受保护的 NFS 共享存储

在主 SVM 中配置 NFS 存储，并使用两个 SVM 共享的 NFS LIF 或 VIP 地址将其挂载到两个站点的所有 Proxmox VE 节点上。Storage Manager for Proxmox (SMP) 在一个工作流中创建 ONTAP 卷、导出策略和 Proxmox VE 存储定义。如果不使用 SMP，请在 ONTAP 中创建卷，并使用 Proxmox VE CLI 将其添加为 NFS 存储。

选择以下方法之一来配置存储。

Proxmox 的存储管理器

使用 SMP 调配存储。必须安装 SMP，并将 Proxmox VE 集群和两个 ONTAP 集群添加为计划。有关详细信息，请参见 ["配置和使用 Storage Manager for Proxmox"](#)。

1. 在 SMP 中，选择 **Provision > Add Volume**，然后选择 **NFS**。
2. 在 **ONTAP** 详细信息 中，选择主 SVM，输入卷名称和大小，然后选择 NFS LIF 或 VIP LIF 作为首选 LIF。
3. 选择允许两个站点上的 Proxmox VE 节点的导出策略，或创建一个。
4. 在 **Proxmox** 详细信息 中，选择两个站点的集群和所有节点，选择 NFS v3 或 NFS v4.1，然后选择内容类型。
5. 确保已选择*挂载数据存储*，然后选择*配置*。
6. 在站点 A 集群上，验证卷是否包含在 SnapMirror 活动同步保护中：

```
volume show -vserver <primary_svm> -volume <volume_name> -fields  
smas-protection
```

命令行界面

使用 Proxmox VE CLI 添加主 SVM 中已存在的卷作为 NFS 存储。

1. 在一个 Proxmox VE 节点上，添加 NFS 存储：

```
pvesm add nfs <storage_id> --server <nfs_lif_or_vip_address>
--export <export_path> --content images,rootdir --options
vers=<nfs_version>,nconnect=4
```

对 <nfs_version> 使用 `3` 或 `4.1`。请勿添加 `soft` 挂载选项。

2. 验证每个节点上的存储是否处于活动状态：

```
pvesm status --storage <storage_id>
```

设置 NFS 挂载选项

使用 nConnect 打开到 NFS LIF 或 VIP 地址的多个 TCP 连接，以获得更好的性能和容错。如果受保护的 SVM 具有客户端可以在两个站点上访问的多个数据 LIF，则可以改用 NFSv4.1 会话中继。请使用 nConnect 或会话中继，而不是同时使用两者。有关详细信息，请参见 ["为 Proxmox VE 配置 NFS 存储"](#)。

步骤

1. 如果您使用 SMP 或在不使用 nConnect 的情况下配置了存储，请在 Proxmox VE 节点上设置挂载选项：

```
pvesm set <storage_id> --options vers=<nfs_version>,nconnect=4
```

要改用会话中继，请设置：

```
pvesm set <storage_id> --options vers=4.1,trunkdiscovery,max_connect=16
```

Proxmox VE 在下次装载存储时应用新的装载选项。在将虚拟机迁移出该节点后，在维护窗口期间重新装载每个节点上的存储。

2. 验证 /etc/pve/storage.cfg 中的选项。
3. 在每个 Proxmox VE 节点上，验证是否存在到 NFS 服务器地址的多个连接：

```
ss -an | grep :2049
```

第 6 步：配置 HA 和 VM 备份

使用首选站点 A 的节点亲和性规则将虚拟机添加到 HA，其中主 SVM 提供数据。然后使用 SMP 备份策略保护虚拟机免受逻辑故障（例如意外删除或勒索软件）的影响。SnapMirror 活动同步会同步复制这些故障，因此不会替换时间点备份。

步骤

1. 将每个受保护的虚拟机添加到 HA:

```
ha-manager add vm:<vm_id> --state started
```

2. 创建首选站点 A 节点并允许故障转移到站点 B 节点的节点亲和性规则:

```
ha-manager rules add node-affinity <rule_name> --resources  
vm:<vm_id>,vm:<vm_id> --nodes  
"<site_a_node1>:2,<site_a_node2>:2,<site_b_node1>:1,<site_b_node2>:1"
```

3. (可选) 增加每个客户机操作系统中的磁盘超时时间, 使其长于预期的存储故障转移时间。
4. 在 SMP 中, 选择 保护 > 备份 > 资源组, 并为每个租户或应用程序的虚拟机或存储创建一个资源组。
5. 选择*保护* > 备份 > 策略 > 添加, 并为资源组创建包含备份频率和保留期限的策略。

对于运行 QEMU 来宾代理的虚拟机, 在*高级选项*中选择*虚拟机一致性*。

6. 运行按需备份, 然后从中还原测试 VM 以验证策略。

第 7 步: 测试计划的故障转移

在将生产虚拟机放入共享存储之前, 请验证设计。计划的故障转移会将活动 SVM 移动到站点 B 而不会丢失数据, 因此您可以在维护窗口期间对其进行测试, 然后进行故障恢复。

步骤

1. 将测试虚拟机实时迁移到站点 B 节点。
2. 在故障转移之前, 请防止 HA 将虚拟机移回站点 A 节点:

```
ha-manager set vm:<vm_id> --failback 0
```

3. 在站点 B 集群上, 启动计划的故障转移:

```
snapmirror failover start -destination-path <secondary_svm>:
```

4. 监控故障转移, 直到其完成:

```
snapmirror failover show
```

5. 验证 VM I/O 是否继续运行, 以及存储是否在所有 Proxmox VE 节点上保持活动状态。
6. 如果使用 VIP LIF, 请验证两台路由器都从群集 B 学习 /32 路由。
7. 在 SMP 中, 运行 ONTAP 计划的发现, 然后运行按需备份以验证 SMP 是否保护站点 B 的 SVM 中的 VM。

8. 测试完成后，故障恢复到站点 A，然后为测试 VM 将 `failback`` 设置回 ``1`。

结果

当活动 SVM 在站点之间移动时，Proxmox VE VM 继续在相同的 NFS 共享存储上运行，并且集群从计划外的站点故障中自动恢复。

相关信息

- ["使用适用于 NAS 的 ONTAP SnapMirror 活动同步进行保护"](#)
- ["配置 ONTAP VIP LIF"](#)
- ["准备安装 ONTAP Mediator"](#)
- ["准备配置 ONTAP Cloud Mediator"](#)
- ["配置 ONTAP Cloud Mediator 以进行 SnapMirror 活动同步"](#)
- ["snapmirror mediator add"](#)
- ["为 Proxmox VE 配置 NFS 存储"](#)
- ["Proxmox VE corosync 外部投票支持"](#)

配置和使用 Storage Manager for Proxmox

Storage Manager for Proxmox (SMP) 为使用 NetApp ONTAP 存储系统的 Proxmox 环境中的虚拟机提供全面的存储配置和生命周期管理。SMP 使管理员能够直接管理存储，通过自动配置、备份、还原和克隆功能简化操作。

了解 Storage Manager for Proxmox

Storage Manager for Proxmox (SMP) 为使用 NetApp 存储系统的 Proxmox 环境中的虚拟机提供端到端的存储管理。通过允许管理员直接管理存储，SMP 简化了 Proxmox 环境中的存储和数据管理任务。

主要特点

- 端到端存储管理：SMP 为虚拟机提供完整的生命周期管理，从调配存储到备份和恢复。
- 直接存储管理：Proxmox 管理员可以直接管理存储，简化操作并降低复杂性。
- 独立安装程序：SMP 作为独立安装程序交付，通过 API 无缝连接到 Proxmox 节点。

功能

SMP 的每个组件都旨在提高存储管理效率：

- 存储配置：轻松为 Proxmox 主机和相关虚拟机分配和管理存储资源。
 - **NFS** 存储配置：设置和配置要与 Proxmox 环境一起使用的 ONTAP NFS 存储。
 - **iSCSI** 存储配置：实施具有多路径的厚 LVM，以实现强大而高效的 iSCSI 存储管理。
- 备份和恢复：执行可靠的备份和快速恢复，以确保数据的完整性和可用性。
 - 崩溃一致性和文件一致性备份：通过设置崩溃一致性和文件一致性备份机制来确保数据完整性。

- **VM 还原操作**：执行虚拟机还原流程以恢复数据并维护系统可用性。
- **存储效率**：ONTAP 提供全面的存储效率技术套件，旨在最大限度地减少数据所需的存储。这些技术不仅可以降低存储消耗，还可以提高系统性能，从而提供：
 - **快速 VM 配置**：简化虚拟机配置，以加快部署。
 - **更快的数据集拷贝**：加快数据集的复制，实现高效的数据管理。

前提条件

在继续之前，请确保您已完成以下操作：

- 工作的 Proxmox Virtual Environment (PVE) 9.x 或更高版本集群。
- ONTAP All Flash FAS (AFF) 存储，在 SVM 级别启用 NFS 和 iSCSI 协议。有关说明，请参见 ["使用 ONTAP 配置 NFS 数据存储"](#) 和 ["使用 ONTAP 存储为 iSCSI 配置 Proxmox VE 9.x"](#)。

对于 iSCSI LUN 配置，应启用以下内容：

在所有 **Proxmox** 主机上安装多路径和 **iSCSI**：

```
apt update
apt install open-iscsi multipath-tools -y
systemctl enable --now iscsid multipathd
```

配置 **iSCSI** 启动程序并通过 **iSCSI** 连接（在 **Proxmox** 主机上）：

创建或更新 `/etc/multipath.conf`：



```
touch /etc/multipath.conf
```

或使用：

```
cat << 'EOF' > /etc/multipath.conf
defaults {
    user_friendly_names yes
    find_multipaths yes
}
blacklist {
    devnode "^sd[a]$"
}
EOF
```

- Proxmox 主机和 ONTAP 存储可从 SMP 主机进行网络访问。

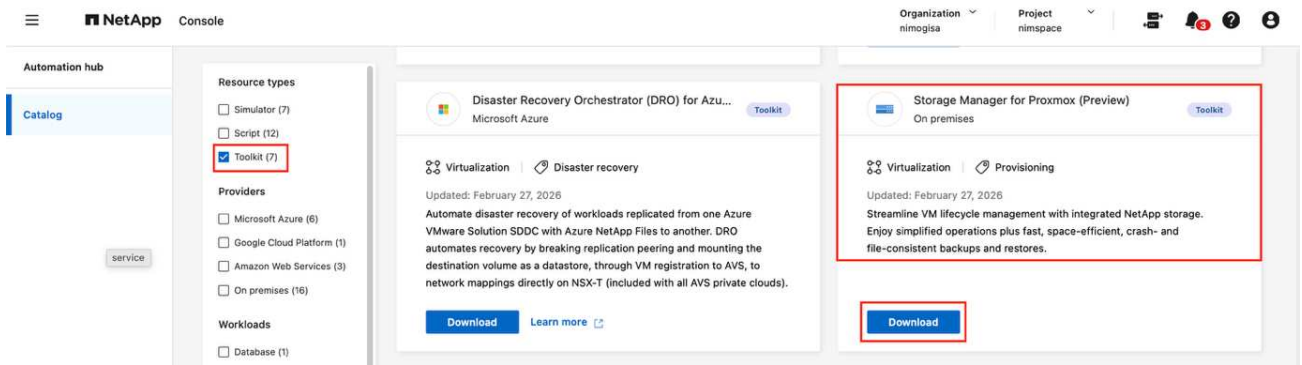
安装和设置 Storage Manager for Proxmox

要开始使用 SMP，请在指定的虚拟机上使用 Windows 操作系统。SMP 可以在 Windows Server 2019、2022 和 2025 上运行。

步骤

1. 从 "NetApp Console 自动化中心" 下载 SMP 包（约 17MB）。

▼ 查看 SMP 下载页面

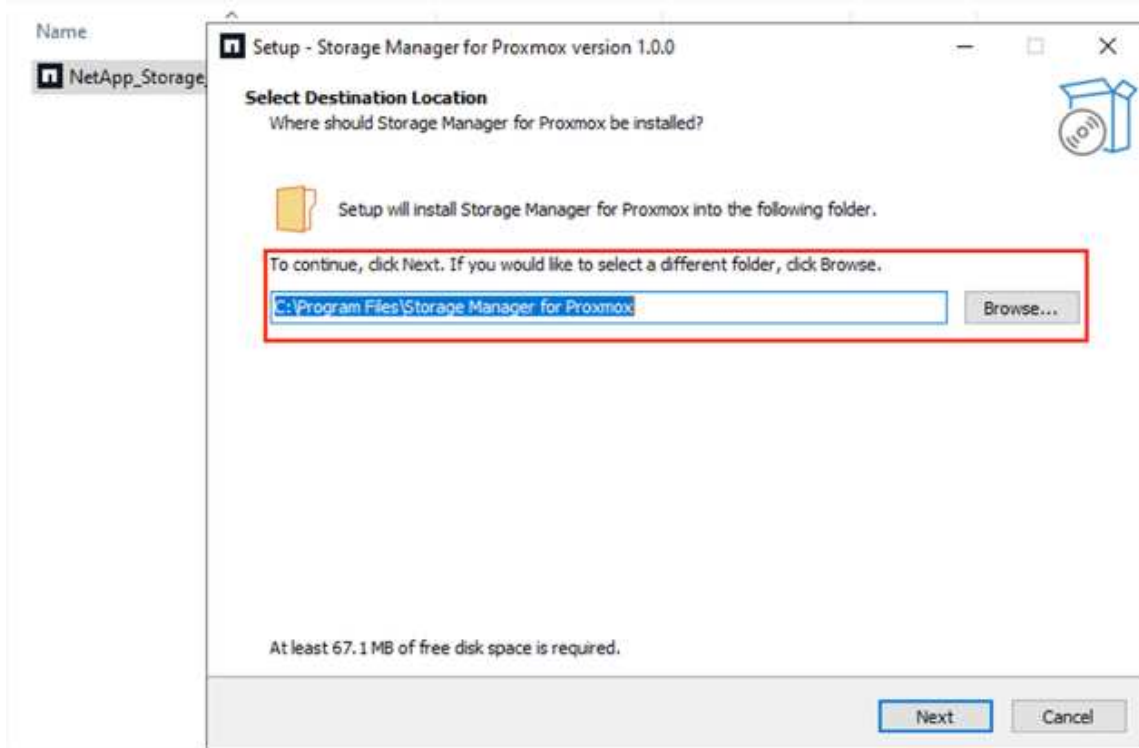


2. 解压缩软件包。
3. 运行可执行文件以安装并启动服务。



不需要额外的包依赖项，因为 SMP 利用可用的 API 来协调配置和数据保护 workflow。

▼ 查看 SMP 安装向导



使用 **Shift** 工具包的替代安装

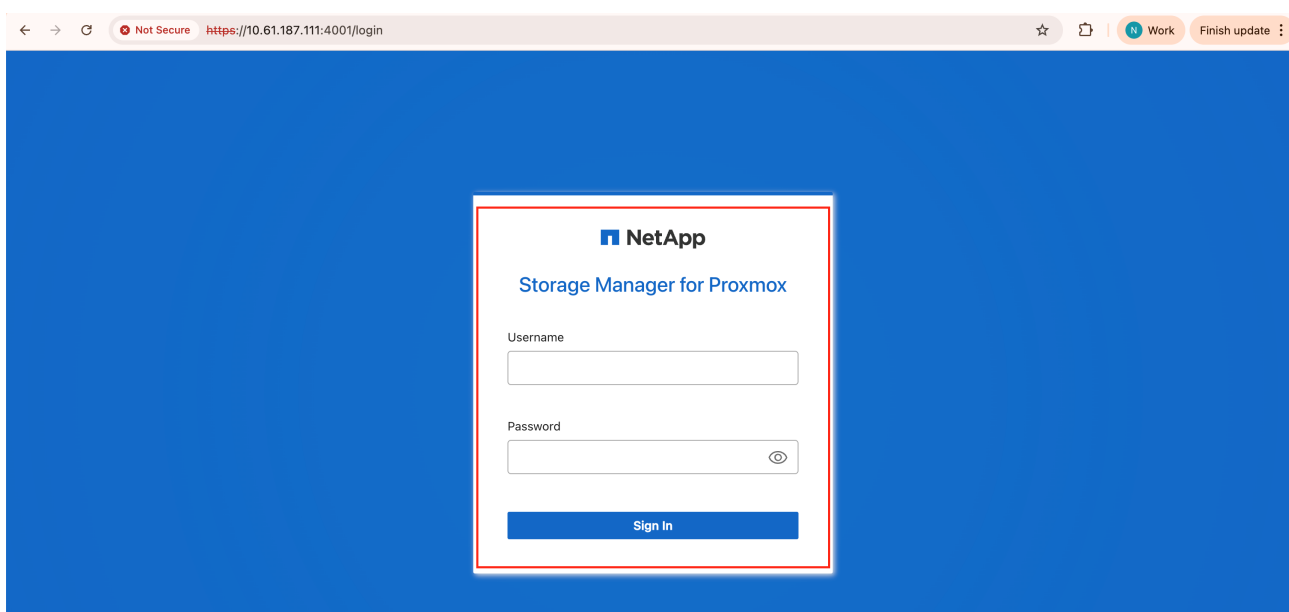
作为独立安装程序的替代方案，您可以使用 Shift 工具包安装 Proxmox VE Storage Manager：

- Open Shift Toolkit。
- 导航到 **Settings > Advanced > Storage Manager**。
- 使用切换开关启用 Proxmox VE。
- 等待系统自动安装必要的组件。
- 安装完成后，通过单击 查看 或打开 Web 浏览器访问 Storage Manager
https://<Shift_server_IP>:4001

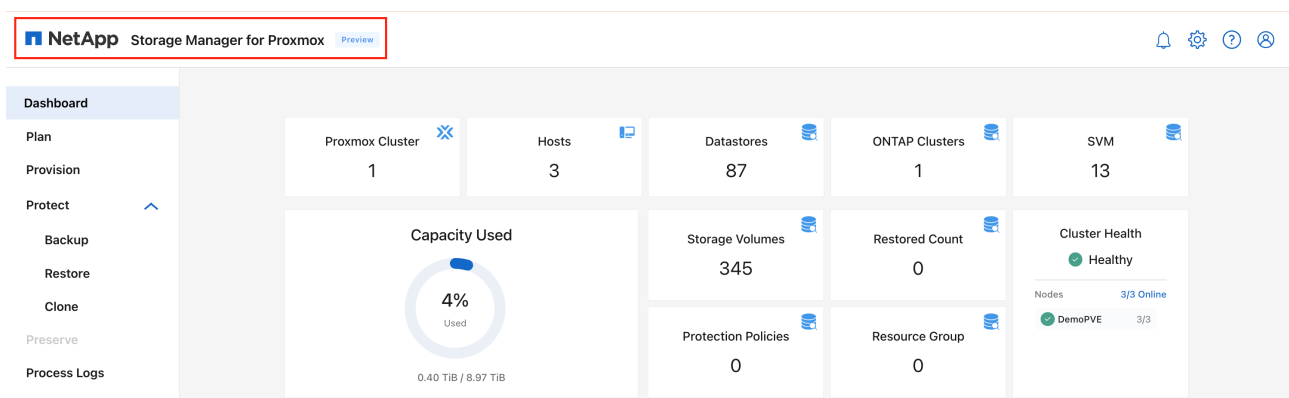
例如： <https://10.61.187.117:4001>

4. 打开 Web 浏览器，通过输入 `http://<IP address specified during installation>:4001` 访问 SMP UI

▼ 查看 SMP 登录屏幕



▼ 查看 SMP 主屏幕



为 Proxmox 配置 Storage Manager

正确配置到存储和 Proxmox 节点的存储和连接后，配置 SMP 以利用 ONTAP Snapshot 和 FlexClone 功能自动配置和保护虚拟机。

第一步是发现并添加 ONTAP 存储端点，然后将目标 Proxmox 节点添加到 SMP。

开始之前

确保满足以下先决条件：

- 已安装 SMP
- SMP 实例可连接到 Proxmox 节点和存储系统
- ONTAP 集群提供 NFS 或 iSCSI 存储池
- 使用 DNS 名称时，DNS 解析已到位。否则，请使用 IP 地址进行发现。
- 环境具有受支持版本的 Proxmox Virtual Environment 服务器

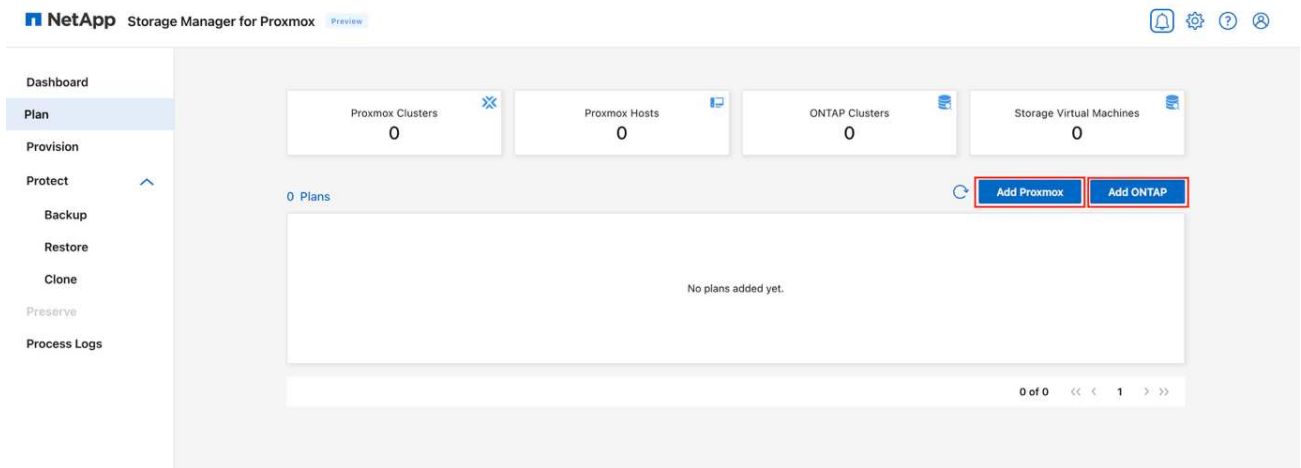
关于此任务

在虚拟机管理程序和存储之间建立连接后，只需单击几下即可进行配置，大约需要 2 分钟。

步骤

1. 登录到 SMP。
2. 在 SMP 左侧导航栏中，选择 **Plan**。
3. 单击 **Add Plan** 。

▼ 查看添加计划



4. 单击 **Add ONTAP** 添加 ONTAP 存储系统的凭据。

您可以使用集群端点或 SVM 管理 LIF 添加 ONTAP。

5. 要使用集群端点添加，请提供以下详细信息：
 - a. 集群 IP 地址或 FQDN
 - b. 用于访问 ONTAP 集群的用户名
 - c. 用于访问 ONTAP 集群的密码
 - d. 选择 **Accept self-signed certificate** ，然后单击 **Add Storage Cluster**

▼ 查看 **Add ONTAP**：集群端点

NetApp Storage Manager for Proxmox

ONTAP Cluster Details

Connection Type: **Cluster** (selected) | SVM

Cluster Name: SMPProxClus

Cluster Management IP: 10.192.102.41

Admin Username: admin

Password:

☒ Accept self-signed certificates

Buttons: Cancel | **Add Storage Cluster**

6. 要使用 SVM 管理 LIF 添加，请提供以下详细信息：

- SVM 管理 LIF 或 FQDN
- 用于访问 SVM 的用户名
- 用于访问 SVM 的密码
- 选择 **Accept self-signed certificate**，然后单击 **Add Storage Cluster**

▼ 查看 Add ONTAP：SVM 管理 LIF

NetApp Storage Manager for Proxmox

ONTAP SVM Details

Connection Type: Cluster | **SVM** (selected)

SVM Name: SMPProxSVM

SVM Management LIF IP: 10.61.187.155

Admin Username: vasdmin

Password:

☒ Accept self-signed certificates

Buttons: Cancel | **Add Storage SVM**

7. 单击 **Add Proxmox** 以添加要在 SMP 中注册的 Proxmox 节点的凭据。

要添加 Proxmox 节点，请指定其中一个现有 Proxmox 节点的 IP 地址。发现作业会自动查询群集以收集所有基本配置详细信息。

8. 请提供以下详细信息：

- a. Proxmox 独立 IP 地址或 FQDN
- b. 用于访问 Proxmox 的用户名
- c. 访问 Proxmox 主机以执行资源库存的密码
- d. 选择 **Accept self-signed certificate** 并单击 **Continue**

▼ 查看添加 **Proxmox**：节点凭据

NetApp Storage Manager for Proxmox

Proxmox Details

Plan Name: SMPProxmox-Username

Proxmox Endpoint: 10.193.172.225

Authentication Method: ☒ Username / Password ☐ API Token

Proxmox Username: root

Proxmox Password: NetApp123

☒ Accept self-signed certificates

Cancel Add Site

或者，使用 API 令牌通过提供以下内容来访问 API：

- API 令牌 ID
- API 令牌密钥



在此版本中，API 令牌仅允许用于 NFS 操作。

▼ 查看添加 **Proxmox**：API 令牌

NetApp Storage Manager for Proxmox

Dashboard

Plan

Provision

Protect

Backup

Restore

Clone

Preserve

Process Logs

Proxmox Details

Node

Plan Name: SMPProxmox-APIToken

Proxmox Endpoint: 10.193.172.221

Authentication Method: ☒ API Token

API Token ID: root@pam!My_Token_123

API Token Secret: [Masked]

☒ Accept self-signed certificates

Cancel Add Site

9. 验证发现作业是否成功，每个计划的状态是否标记为 **Success**。

添加后，SMP 会自动执行发现扫描并显示虚拟机及其关联的元数据。还会检测到任何现有的存储卷或映射的 LUN。

▼ 查看 Plan 发现结果

NetApp Storage Manager for Proxmox

Dashboard

Plan

Provision

Protect

Backup

Restore

Clone

Preserve

Process Logs

Proxmox Clusters: 2

Proxmox Hosts: 1

ONTAP Clusters: 1

Storage Virtual Machines: 21

3 Plans

Name	Endpoint	Type	Discover Status
SMPProxmox-Cluster	10.192.102.41	ONTAP	Success
SMPProxmox-APIToken	10.193.172.221:8006	Proxmox	In Progress
SMPProxmox-Username	10.193.172.225:8006	Proxmox	Success

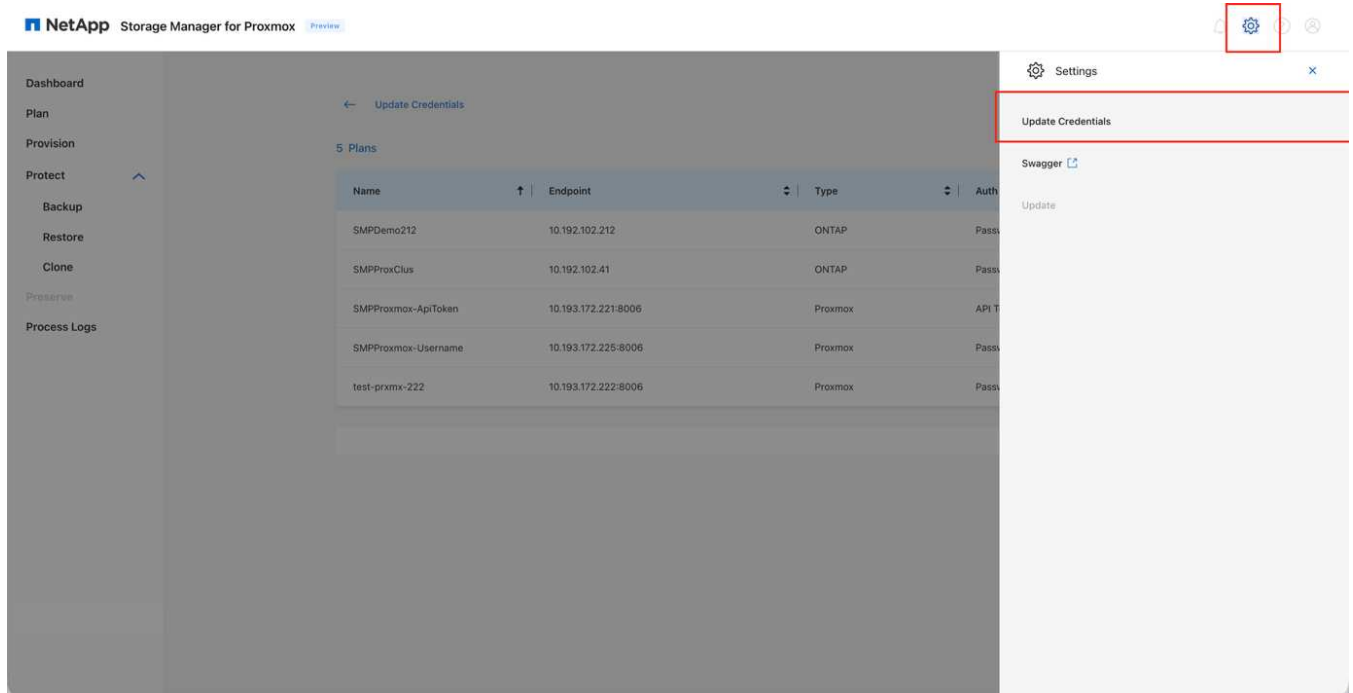
1 - 3 of 3

完成后

- 如果对 Proxmox 节点或存储进行了任何修改，请运行发现以获取最新信息。单击计划名称上的三个点，然后单击 发现。
- 确保首先添加 ONTAP 存储系统，以便发现作业可以相应地映射现有资源。如果首先添加了 Proxmox，则针对 ONTAP 运行发现作业。

- 如果凭据或 API 令牌稍后需要更新，请导航到*设置*并选择*更新凭据*选项。

▼ 查看更新凭据



配置存储

将卷设置为存储池（通常称为数据存储）可为虚拟机及其虚拟磁盘（QCOW2 或 RAW）创建逻辑容器。SMP 可以调配数据存储区，然后将其连接到单个主机或群集中的所有主机。

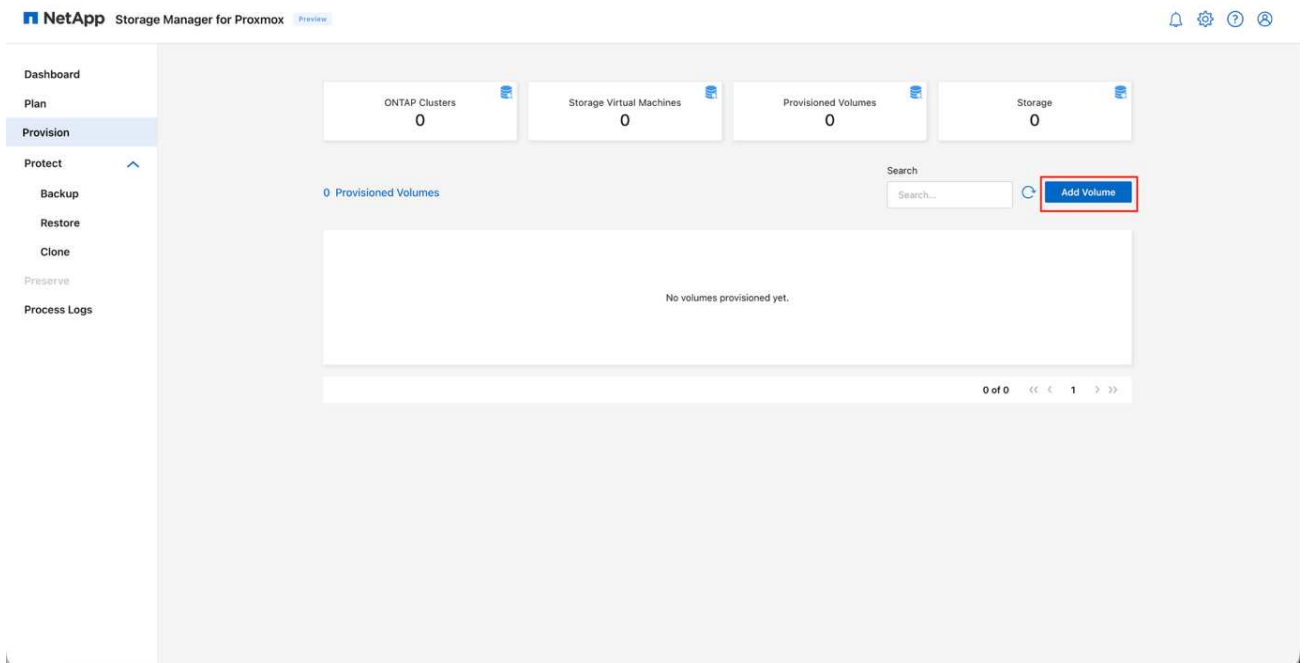
配置 NFS 存储

使用 SMP 将 ONTAP 卷配置和挂载为 NFS 数据存储库。

步骤

1. 在 SMP GUI 中，转到 **Provision** 选项卡。
2. 导航到 **Add Volume**。

▼ 查看配置选项卡：添加卷

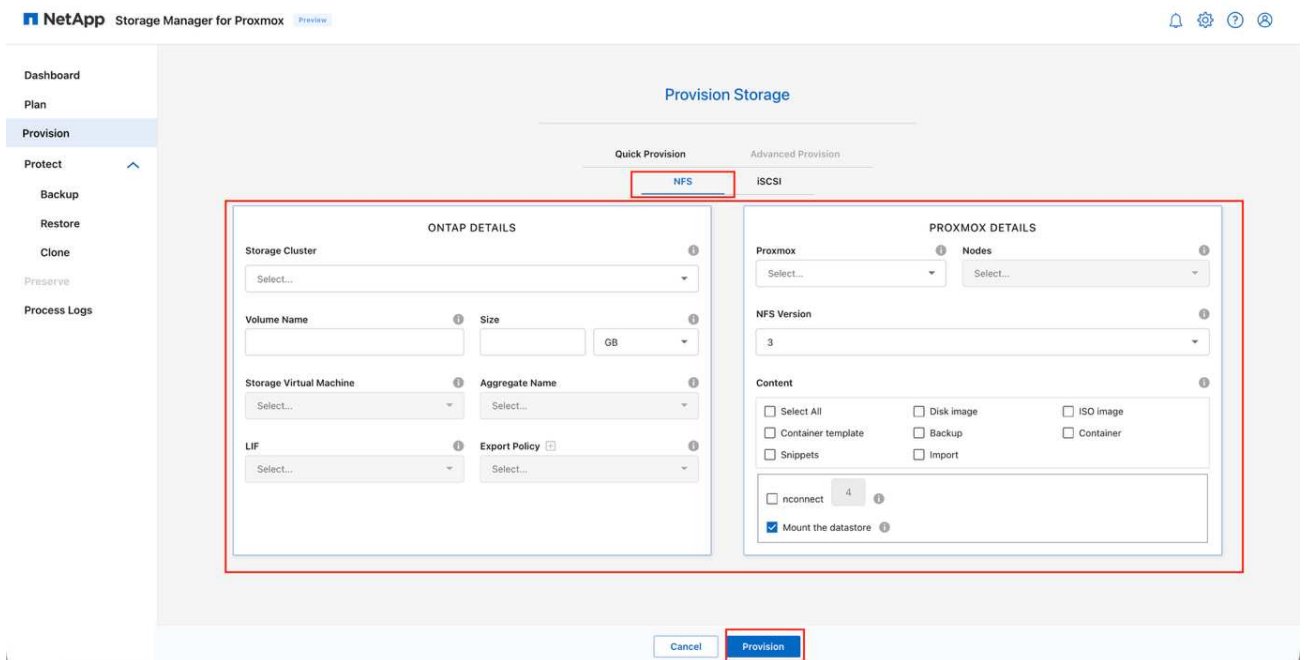


3. 在"新建卷"向导的"快速设置"步骤中，选择 **NFS**。



默认情况下选择 NFS。快速配置可使用标准部署的默认设置提供更快的设置。高级配置为 HA 和 BCDR 功能提供了扩展选项，将在未来版本中引入。

▼ 查看新卷向导：NFS

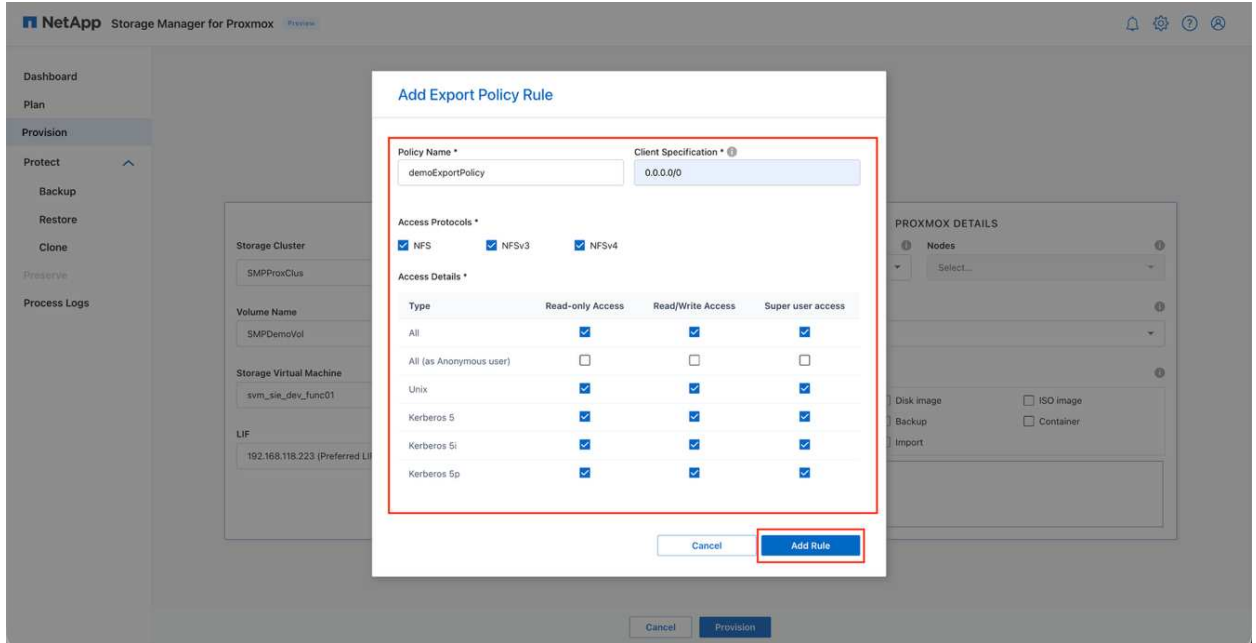


4. 在 **ONTAP** 详细信息 部分中，提供以下信息：

- 从 **Storage Cluster** 下拉列表中选择 ONTAP 集群或指定的 Storage Virtual Machine (SVM)
- 在 **Volume Name** 字段中提供一个唯一值
- 在 **Size** 字段中输入数据存储卷所需的容量

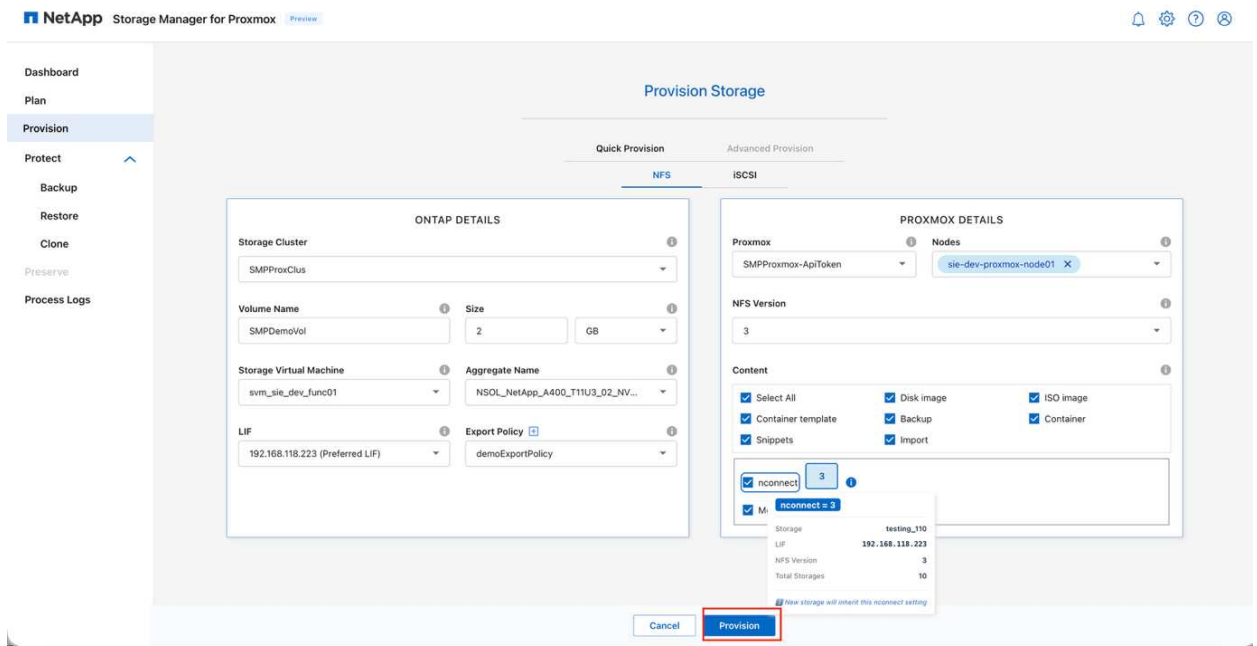
- d. 通过选择适当的选项来选择 ONTAP 存储资源：
- 聚合名称
 - Storage Virtual Machine (SVM)
 - 要用于数据存储访问的首选 LIF
- e. 选择现有的 **Export Policy** 或根据需要单击 (+) 创建自定义策略

▼ 查看 **ONTAP** 详细信息：NFS



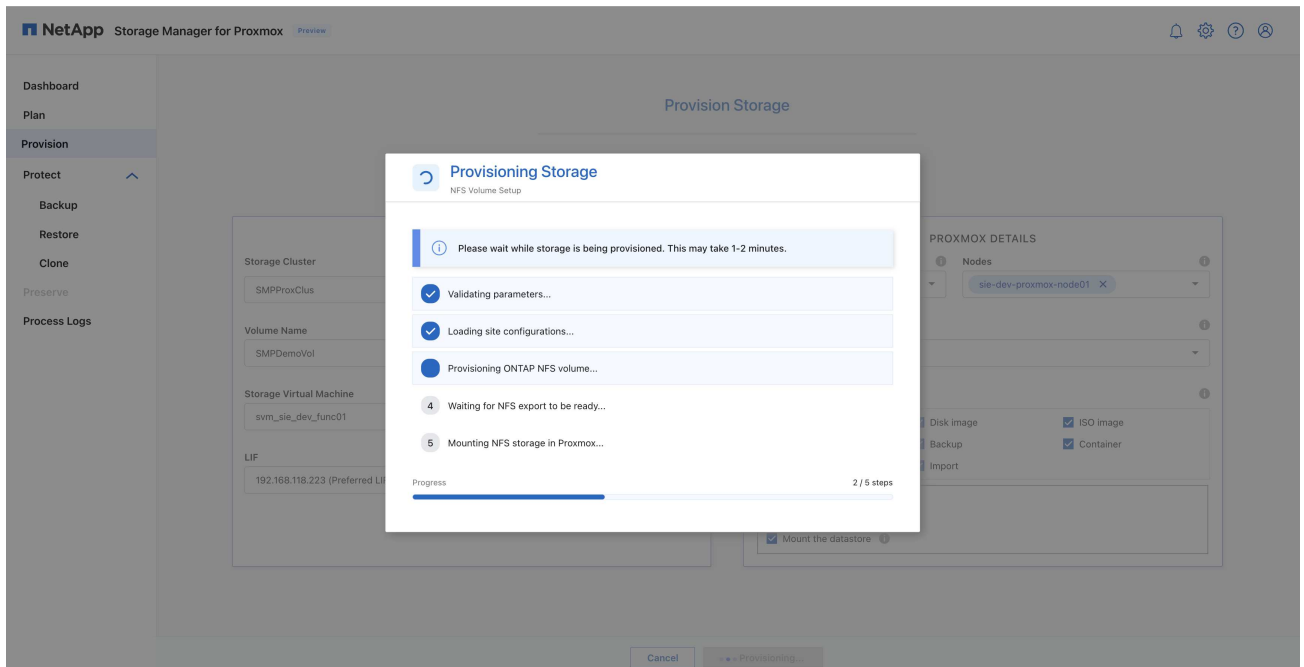
5. 在 **Proxmox** 详细信息 部分中，提供以下信息：
- 在 **Proxmox** 字段中，选择目标集群，然后选择所有需要访问数据存储的节点
 - 在 **NFS Version** 字段中，选择首选 NFS 协议（NFS v3 或 NFS v4.1）
 - 在 **Content** 部分中，选择要为此数据存储区启用的内容类型：
 - 磁盘映像
 - ISO 映像
 - 容器模板
 - 备份
 - 片段
 - 导入
 - 或选择 所有 内容类型
 - （可选）如果需要多个 TCP 连接，请激活 **nconnect** 选项
 - 确保选中 **Mount the datastore**，以便将数据存储库挂载到选定的 Proxmox 节点上

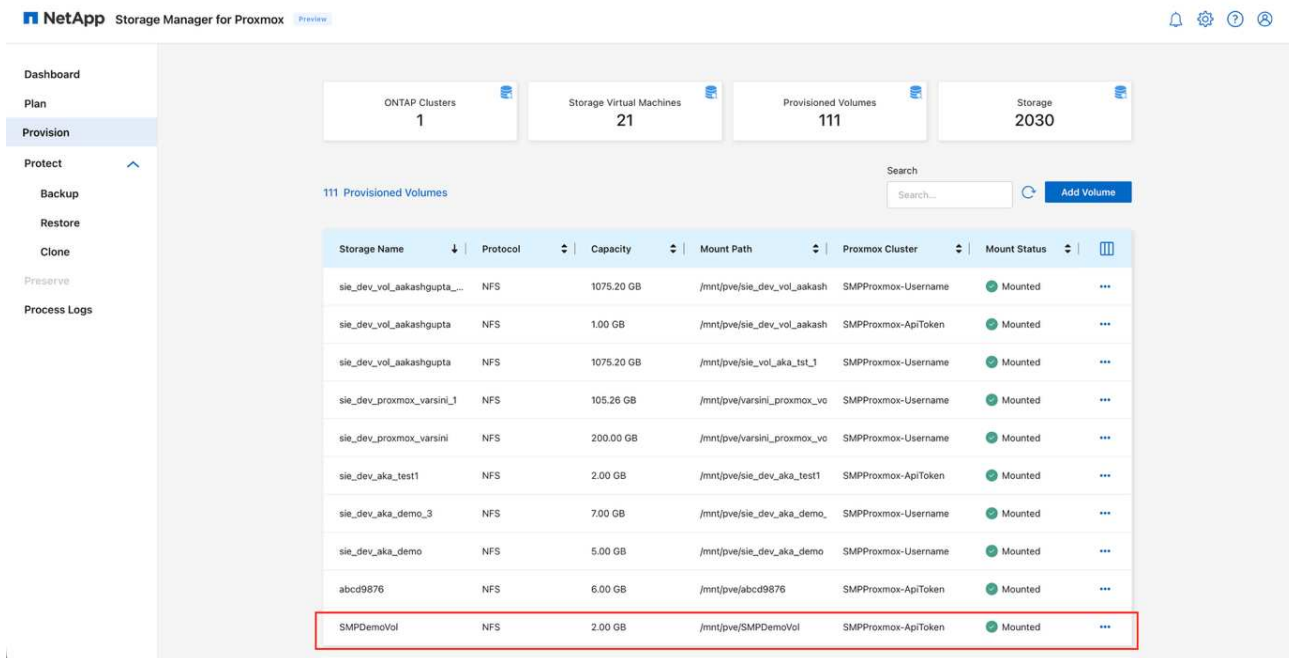
▼ 查看 **Proxmox** 详细信息：NFS



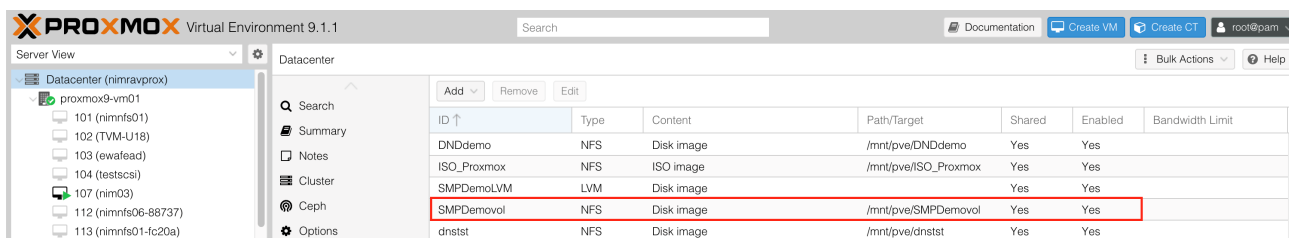
6. 单击 **Provision** 以创建 ONTAP 卷并在选定的 Proxmox 集群中注册数据存储区（存储池）。

▼ 查看 NFS 配置结果





▼ 在 Proxmox 中查看 NFS 数据存储



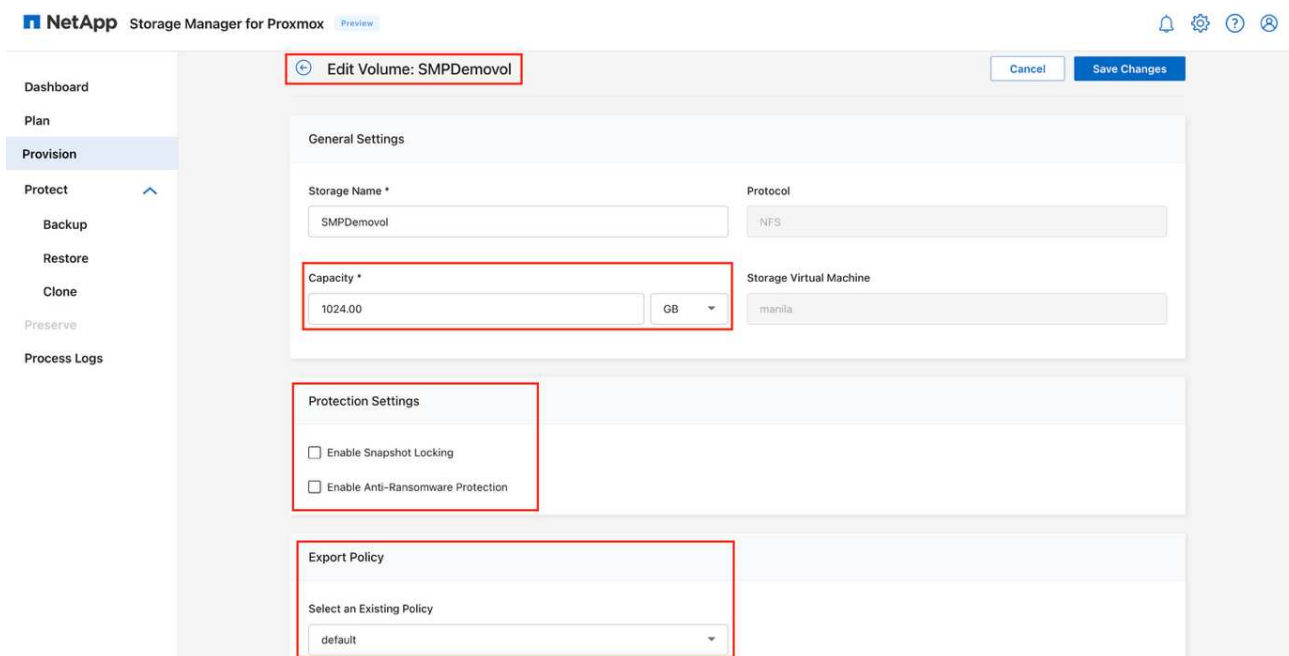
编辑已配置的卷

使用 SMP 在配置后修改卷设置。

步骤

1. 在 SMP GUI 中，在 **Provision** 选项卡中找到卷。
2. 单击卷旁边的三个垂直圆点 (⋮)。
3. 从上下文菜单中选择 **编辑** 以打开卷设置。

▼ 查看编辑卷对话框



4. 根据需要修改卷设置：

- 调整卷大小：调整卷的存储容量



调整大小仅适用于 NFS 卷。

- **Snapshot 锁定**：启用以防止删除或更改快照



这需要在集群上安装 SnapLock 许可证，并且必须初始化集群的合规时钟。

- **Anti-Ransomware Protection**：启用以主动检测并警告可能表明勒索软件攻击的异常活动
- 导出策略：更新或修改导出策略

5. 单击 **Save** 以应用更改。

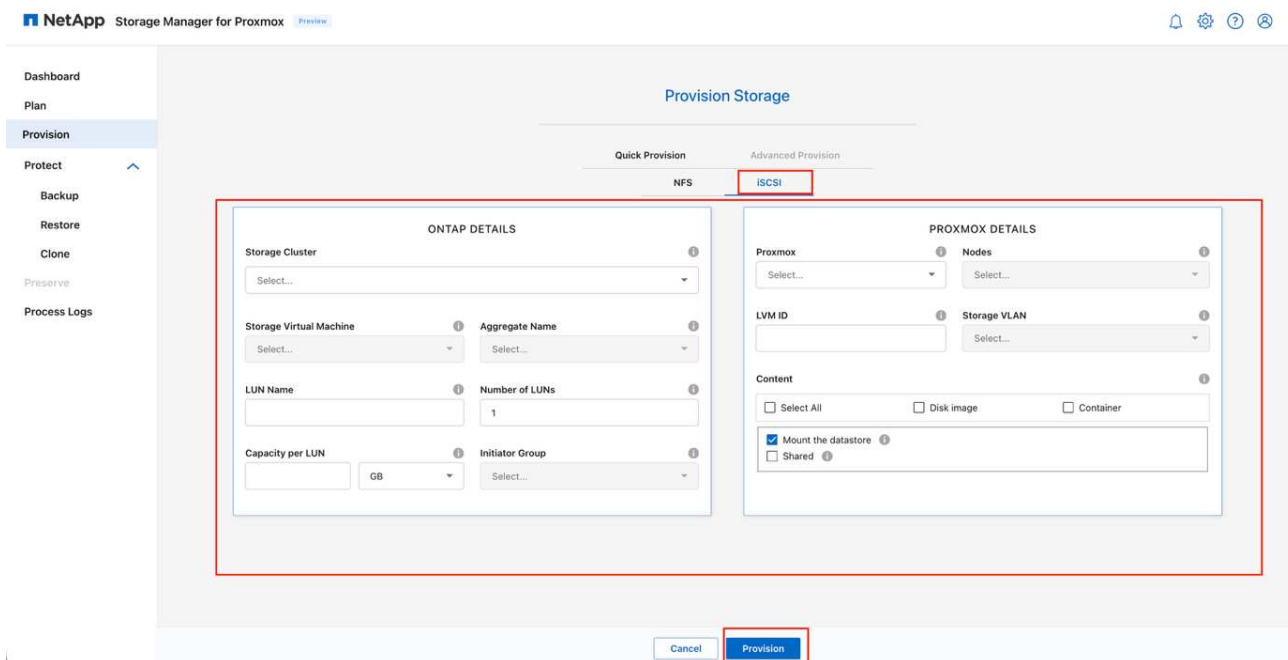
预配 iSCSI 存储

使用 SMP 为虚拟机设置 iSCSI 后端。

步骤

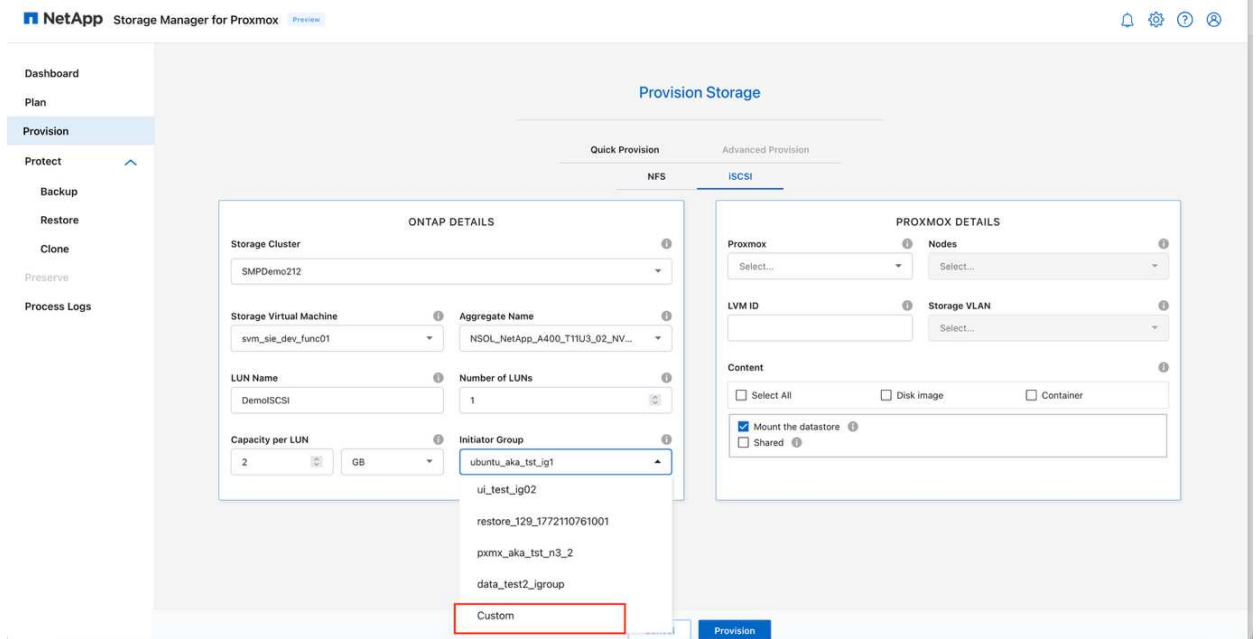
1. 在 SMP GUI 中，转到 **Provision** 选项卡。
2. 导航到 **Add Volume**。
3. 在“新建卷”向导的“快速设置”步骤中，选择 **iSCSI**。

▼ 查看新卷向导：iSCSI

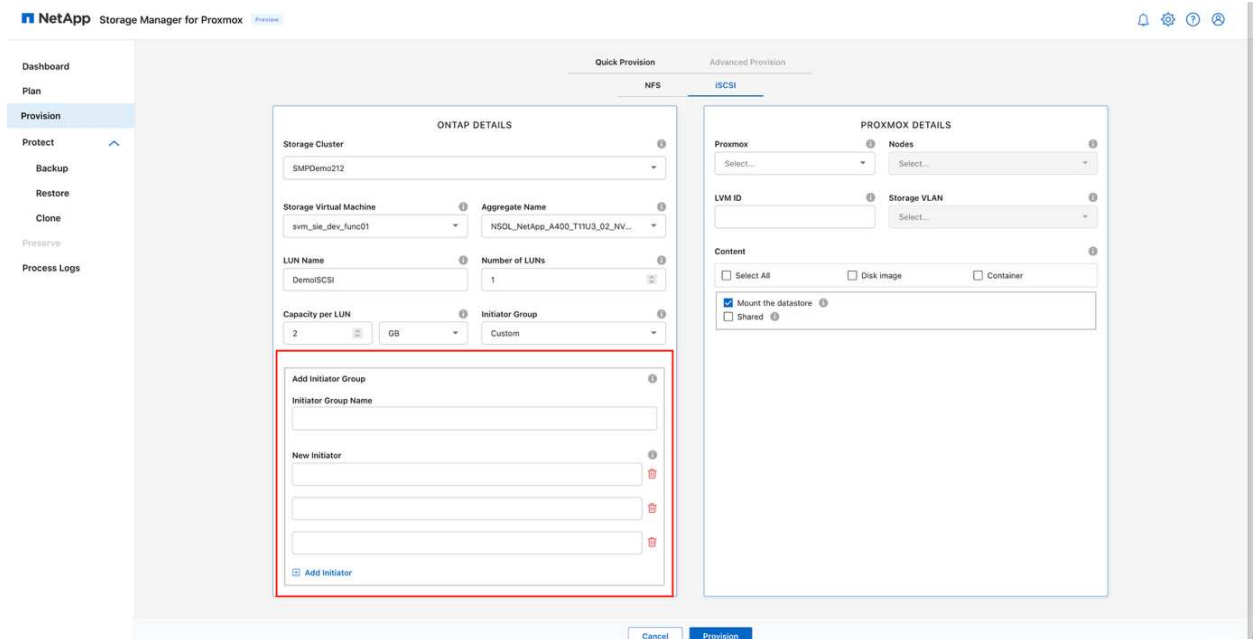


4. 在 iSCSI 协议下，在 **ONTAP** 详细信息 部分中，提供以下信息：
- 从 **Storage Cluster** 下拉列表中选择适当的 ONTAP 集群或指定的 Storage Virtual Machine (SVM)
 - 选择以下项的选项：
 - 聚合名称
 - Storage Virtual Machine (SVM)
 - 在 **LUN 名称** 字段中输入唯一值
 - 在 **Number of LUNs** 字段中提供所需的 LUN 数量（默认值为 1；可以并行创建相同大小的其他 LUN）
 - 在 **每个 LUN 的容量** 字段中输入每个 LUN 所需的大小
 - 为 iSCSI 访问控制选择一个*启动器组*（igroup）。可以选择现有的 igroup 或创建新的 igroup。

▼ 查看 **ONTAP** 详细信息：**iSCSI LUN 配置**



▼ 查看 iSCSI igroup 选择

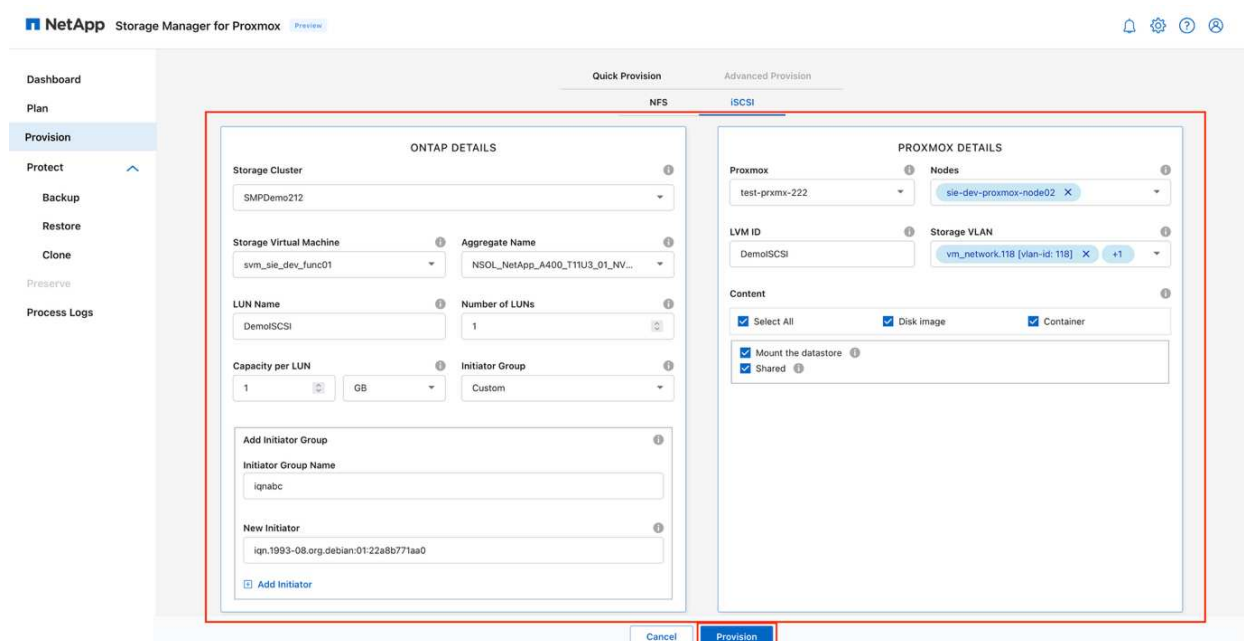


5. 在 **Proxmox** 详细信息 部分中，提供以下信息：

- 从 **Proxmox** 下拉列表中，选择目标集群，然后选择需要访问存储的所有节点
- 选择适当的 **LVM ID**，它将代表 Proxmox 中支持 iSCSI 的存储池
- 在 **节点** 字段中，选择所有需要访问显示的 iSCSI 目标的节点
- （可选）如果需要 VLAN 分段，请选择所需的 **Storage VLAN**
- 在 **Content** 部分中，选择此数据存储区支持的内容类型：
 - 磁盘映像
 - 容器
 - 或选择 所有 内容类型

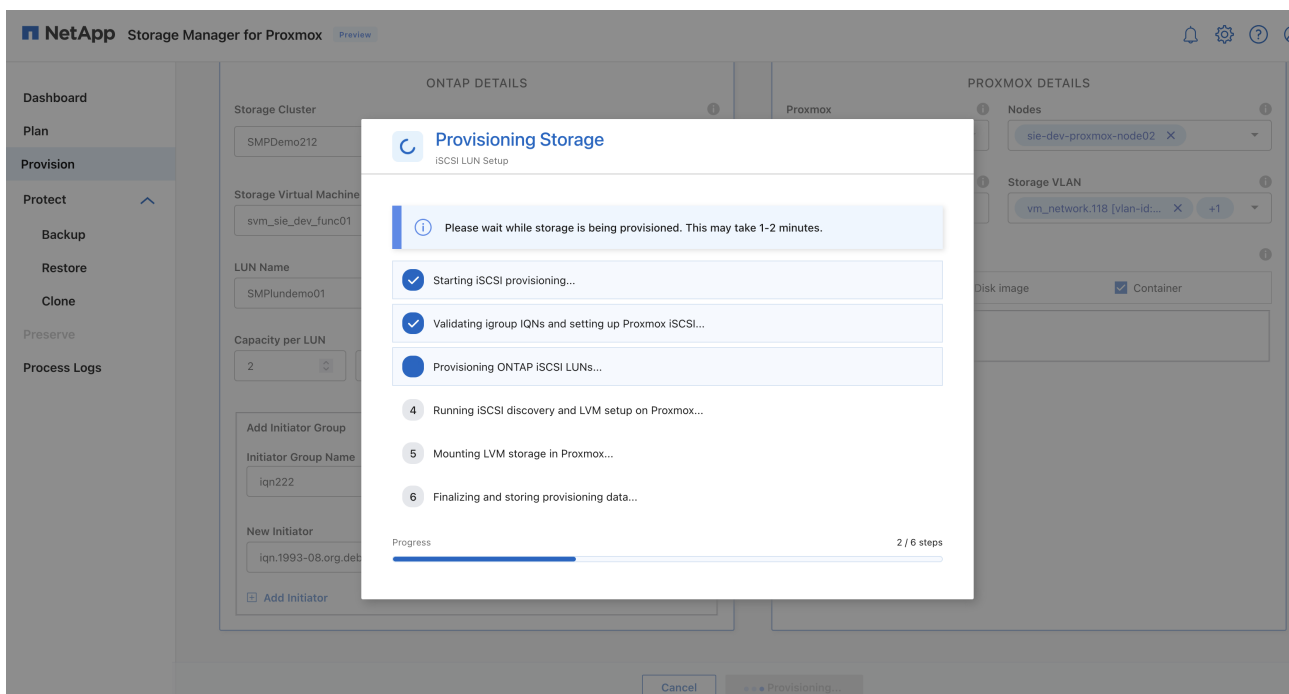
- f. 确保已选择 **Mount the datastore**，以便存储在选定的 Proxmox 节点上可用
- g. （可选）如果存储池将被多个节点同时访问，则启用 **Shared**

▼ 查看 Proxmox 详细信息：iSCSI

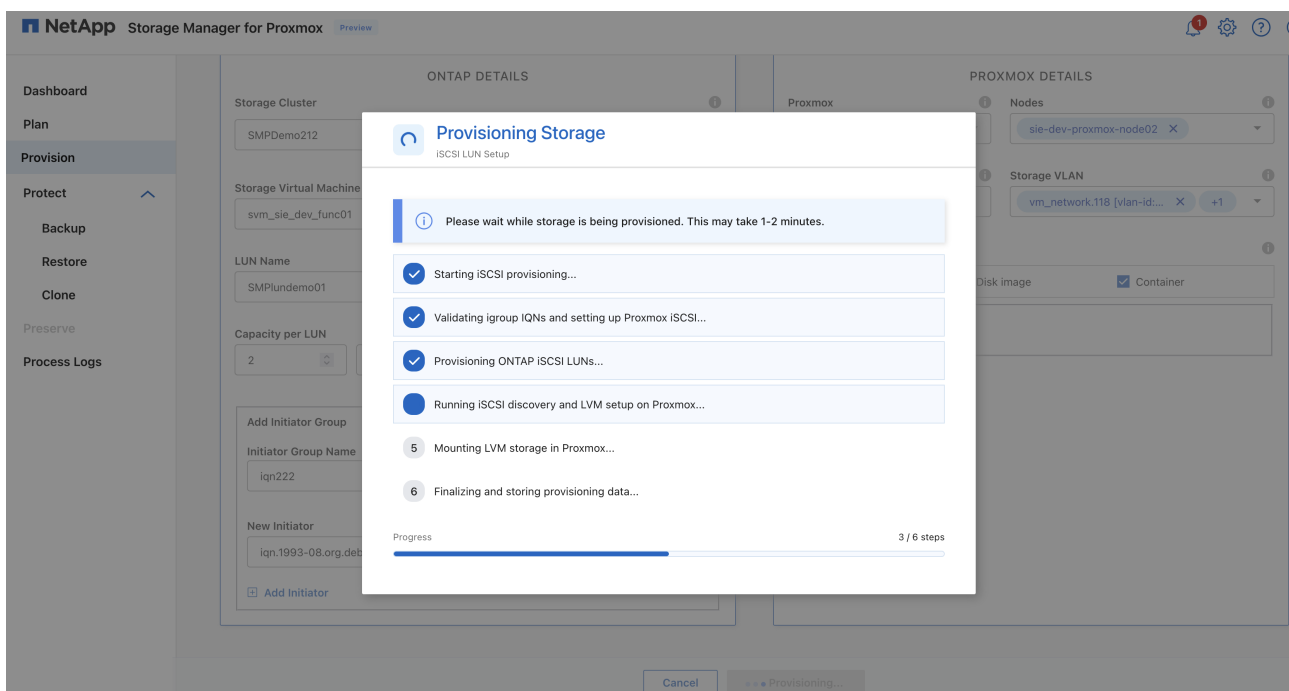


6. 单击 **Provision** 以创建 ONTAP LUN 并在选定的 Proxmox 集群中注册数据存储区（存储池）。

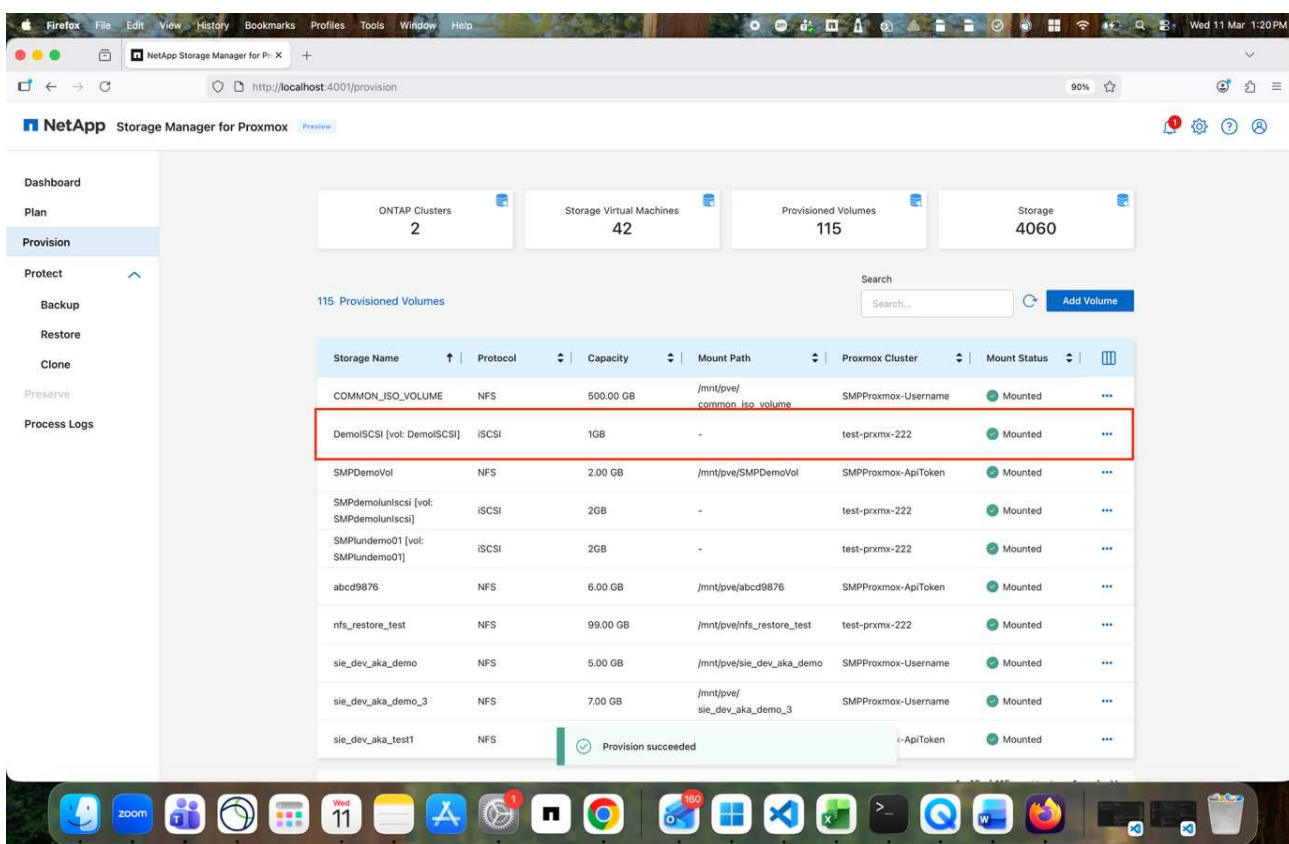
▼ 查看 iSCSI LUN 配置



▼ 查看 iSCSI 配置已完成



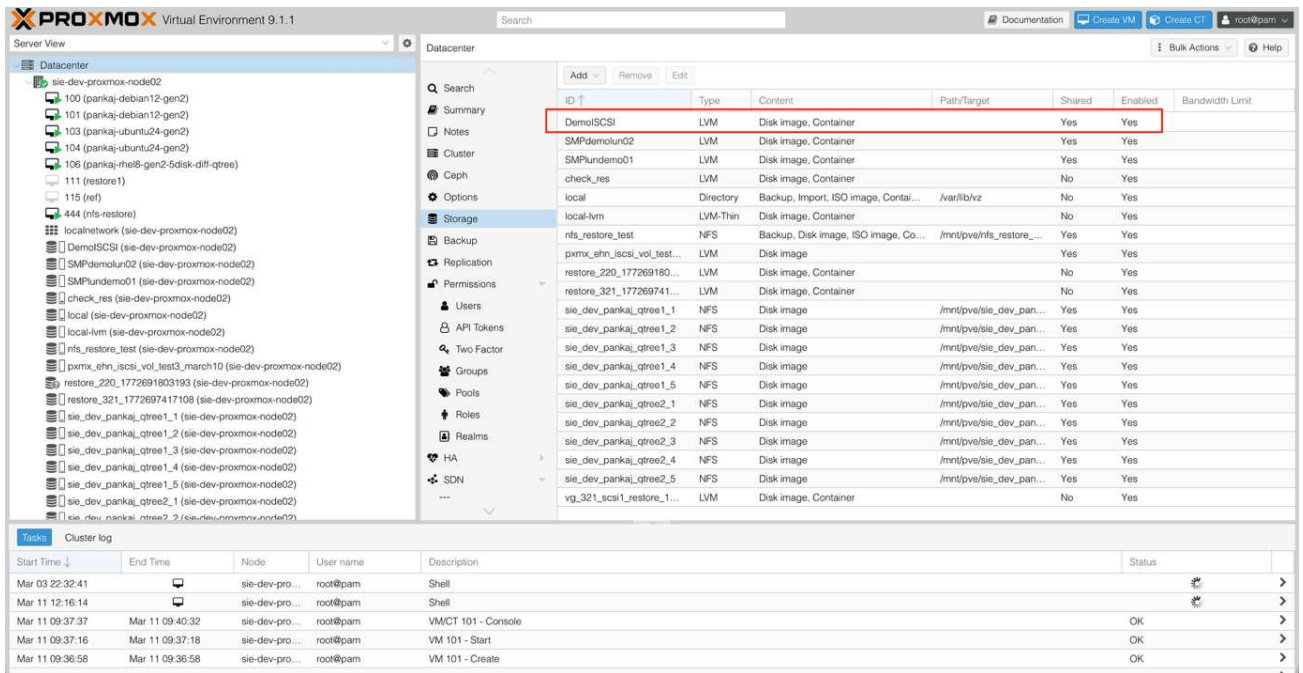
▼ 在 Proxmox 中查看 iSCSI 数据存储



7. 通过访问 Proxmox 端点验证已连接 NFS 和 iSCSI LUN。

现在，连接的存储可用于配置虚拟机。

▼ 查看 Proxmox 存储视图



卷配置选项

当 SMP 配置新的 ONTAP FlexVol 卷和 LUN 时，会自动应用几个默认配置值：

ONTAP 设置	默认值
Snapshot 预留 (-percent-snapshot-space)	0%
分数预留 (-fractional-reserve)	0%
访问时间更新 (-atime-update)	已禁用
最小预读数 (-min-readahead)	已禁用
计划快照	无
存储效率	已启用
卷保证	无（精简配置）
卷自动调整	grow_shrink
LUN 空间预留	已禁用
LUN 空间分配	已启用

NConnect 行为

将 NFS 存储与 Proxmox VE 一起使用时，请注意 Linux NFS 客户端（由 Proxmox VE 的基本操作系统 Debian 使用）中的特定行为。Linux 内核可能会重用首次在节点上挂载 NFS 导出时的挂载参数。如果使用不同的挂载选项再次挂载相同的导出，则可能不会应用这些新选项。

作为最佳做法，创建和装载单独的导出并使用不同的 NFS 服务器 IP 地址，例如不同的数据 LIF。SMP 自动检测现有连接及其关联的 nconnect 值，并将相应的 nconnect 值应用于任何新连接。

保护虚拟机

当虚拟机在 ONTAP 支持的 NFS 数据存储上运行时，最简化的保护方法是让 Proxmox 建立文件系统一致性，而 ONTAP 在存储层捕获实际备份点。ONTAP 快照提供数据存储卷的快速、节省空间的时间点副本，以后可用于恢复整个数据存储、单个文件或 VM 内容。

SMP 为崩溃一致和 VM 一致的数据保护和克隆管理提供了一个简单的协调器。此软件通过创建可应用于多个作业的备份策略来简化备份、还原和克隆生命周期。这些策略可以定义计划、保留、复制和其他功能。它们允许可选地选择文件一致性快照；这种方法使用 QEMU 在拍摄 ONTAP 快照之前停止 I/O 的能力。

SMP 允许保护整个数据存储库，其中一个 ONTAP NFS 卷包含多个虚拟机、可能跨越多个数据存储库的虚拟机组或单个虚拟机级别。

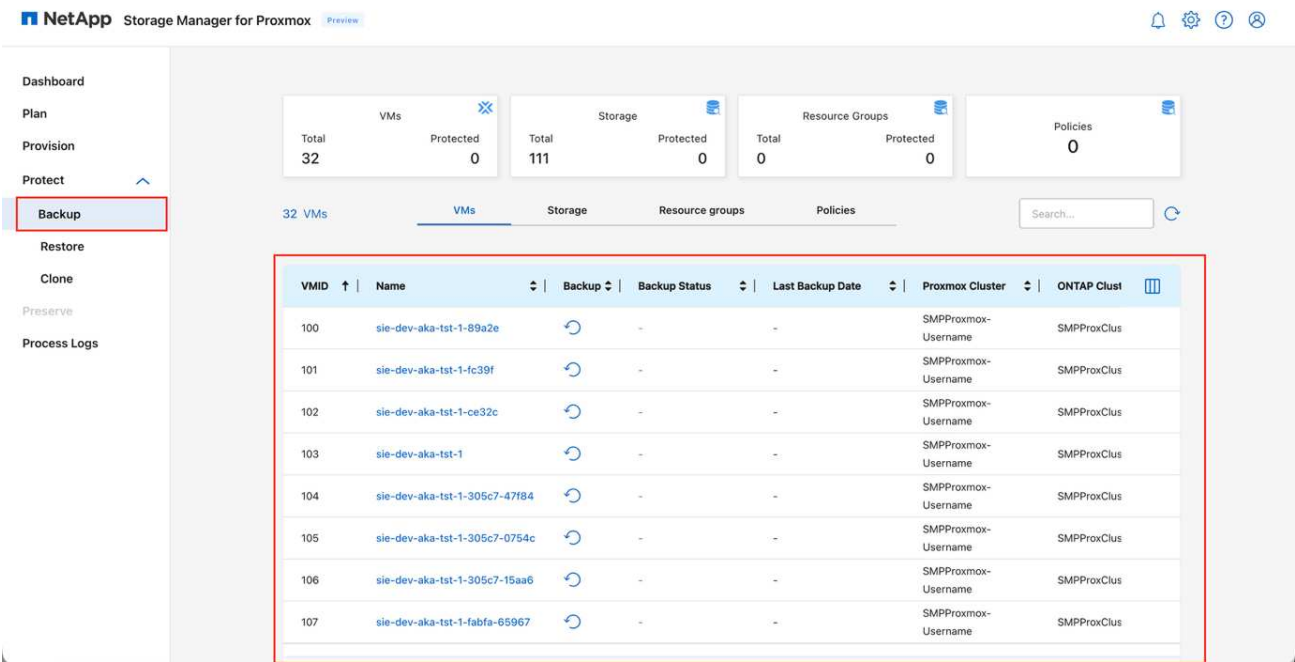
备份虚拟机

配置 ONTAP 卷后，使用备份选项列出在 Proxmox 集群上运行的所有虚拟机，并对虚拟机执行崩溃一致性或文件一致性备份。

步骤

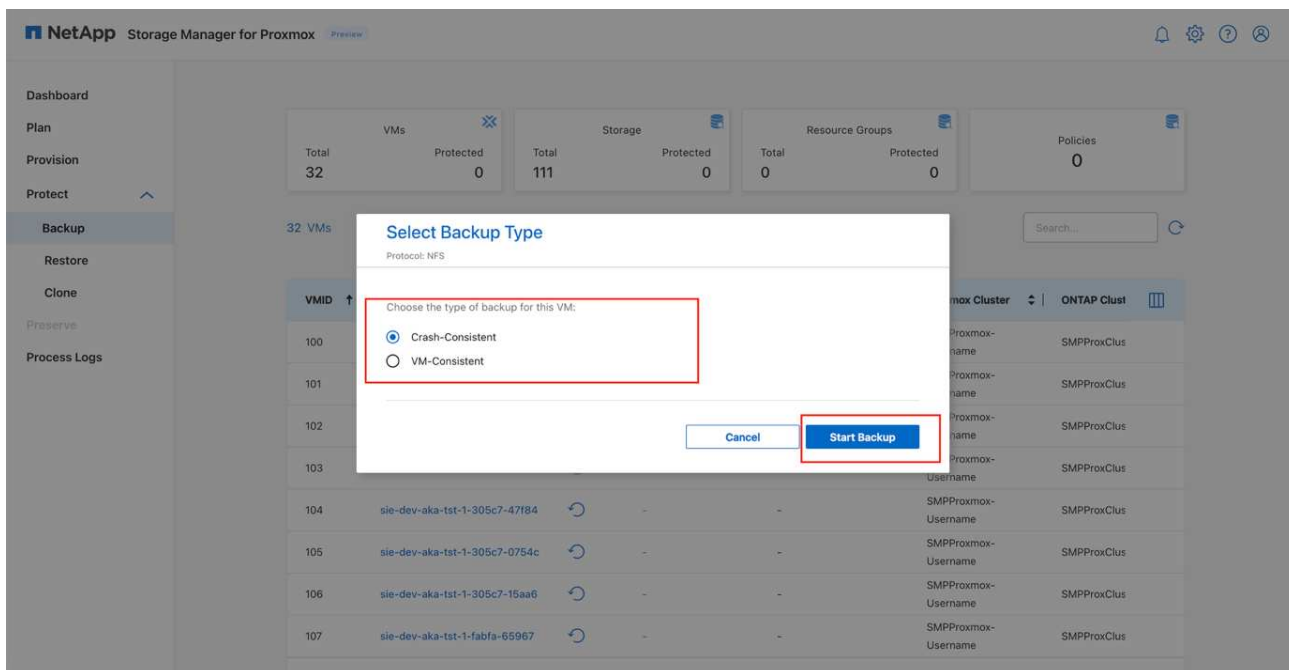
- 1. 从左侧导航窗格访问 **Protect > Backup**。

▼ 查看 Protect Backup 页面



- 2. 选择要备份的虚拟机，然后单击 **Backup**。

▼ 查看备份对话框：VM 选择



3. 备份工作流允许在崩溃一致性保护和文件一致性保护之间进行选择。

文件一致模式可选地使用 QEMU fsfreeze 来确保在拍摄快照之前来宾文件系统处于静止状态。如果 fsfreeze 失败，系统将回退到使用 Proxmox 快照。



此回退机制仅适用于存储在 NFS 卷上的虚拟机。

▼ 查看备份对话框：一致性选项

Start Time	End Time	Node	User name	Description	Status
Feb 25 02:55:28		proxmox9-...	root@pam	Shell	
Mar 10 08:51:01	Mar 10 08:51:01	proxmox9-...	root@pam	VM 116 - Delete Snapshot	OK
Mar 10 08:50:56	Mar 10 08:50:57	proxmox9-...	root@pam	VM 116 - Snapshot	OK

▼ 查看备份进度

VMID	Name	Backup	Backup Status	Last Backup Date	Proxmox Cluster	ONTAP Clust
100	sie-dev-aka-tst-1-89a2e		Success	11/03/2026, 11:44:00	SMPProxmox-Username	SMPProxClus
101	sie-dev-aka-tst-1-fc39f		-	-	SMPProxmox-Username	SMPProxClus
102	sie-dev-aka-tst-1-ce32c		-	-	SMPProxmox-Username	SMPProxClus
103	sie-dev-aka-tst-1		-	-	SMPProxmox-Username	SMPProxClus
104	sie-dev-aka-tst-1-305c7-47f84		-	-	SMPProxmox-Username	SMPProxClus
105	sie-dev-aka-tst-1-305c7-0754c		-	-	SMPProxmox-Username	SMPProxClus
106	sie-dev-aka-tst-1-305c7-15aa6		-	-	SMPProxmox-Username	SMPProxClus
107	sie-dev-aka-tst-1-fabfa-65967		VM backup completed: 1 successful, 0 failed		SMPProxmox-Username	SMPProxClus



ONTAP 快照是在卷级别拍摄的，在单个快照中捕获 NFS 数据存储区内的所有 VM 磁盘和配置文件。当选择 VM 一致性保护时，仅对所选 VM 发出 fsfreeze；卷中的所有其他 VM 都作为崩溃一致性包含在同一个 ONTAP 快照中。

其他备份选项

- 要备份单个数据存储库，请选择 **Storage** 选项卡并选择 **Backup** 选项。
- 要将多个数据存储区备份为一个单元或保护跨越不同数据存储区的 VM 组，请使用 **Resource Group** 选项。资源组允许将数据存储区或 VM 组织成逻辑保护集。

▼ 查看资源组选项卡

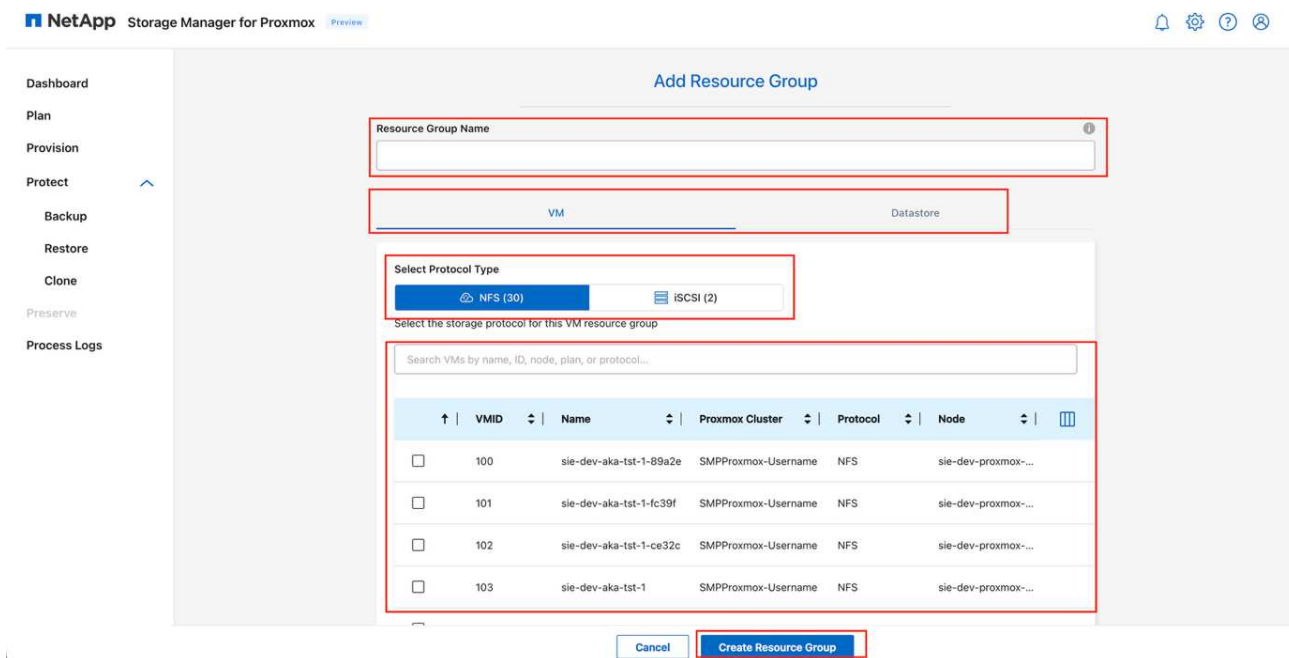
The screenshot shows the NetApp Storage Manager for Proxmox interface. The left sidebar contains navigation options: Dashboard, Plan, Provision, Protect (selected), Backup, Restore, Clone, Preserve, and Process Logs. The main content area displays four summary cards: VMs (Total: 32, Protected: 0), Storage (Total: 111, Protected: 0), Resource Groups (Total: 0, Protected: 0), and Policies (Total: 0, Protected: 0). Below these cards, there are tabs for VMs, Storage, Resource groups (highlighted with a red box), and Policies. A search bar is located to the right of the tabs. The main table lists VMs with columns: VMID, Name, Backup, Backup Status, Last Backup Date, Proxmox Cluster, and ONTAP Cluster. The table shows 8 VMs, with the first one having a successful backup status.

VMID	Name	Backup	Backup Status	Last Backup Date	Proxmox Cluster	ONTAP Cluster
100	sie-dev-aka-tst-1-89a2e		Success	11/03/2026, 11:44:00	SMPProxmox-Username	SMPProxClus
101	sie-dev-aka-tst-1-1c39f		-	-	SMPProxmox-Username	SMPProxClus
102	sie-dev-aka-tst-1-ce32c		-	-	SMPProxmox-Username	SMPProxClus
103	sie-dev-aka-tst-1		-	-	SMPProxmox-Username	SMPProxClus
104	sie-dev-aka-tst-1-305c7-47f84		-	-	SMPProxmox-Username	SMPProxClus
105	sie-dev-aka-tst-1-305c7-0754c		-	-	SMPProxmox-Username	SMPProxClus
106	sie-dev-aka-tst-1-305c7-15aa6		-	-	SMPProxmox-Username	SMPProxClus
107	sie-dev-aka-tst-1-fabfa-65967		-	-	SMPProxmox-Username	SMPProxClus

创建资源组

1. 打开 **Resource Group** 选项卡并选择 **Add**。
2. 输入新资源组的名称。
3. 选择资源类型：**VM** 或 **Datastore**。
4. 选择适当的 **Protocol**。
5. 根据所选资源类型选择要包含的特定虚拟机或数据存储区。
6. 单击 **Create Resource Group** 完成设置。

▼ 查看"创建资源组"对话框



在数据存储区（卷）级别进行的备份始终是崩溃一致的。一个资源组只能包含一种资源类型：虚拟机或数据存储区。对于基于 iSCSI 的存储，备份始终是崩溃一致的。

备份 workflow 详细信息

高级备份 workflow 包括：

1. 定义保护范围：选择保护目标是数据存储区（ONTAP 卷）、资源组（多个卷或 VM）还是单个 VM。
2. 确认一致性方法：选择在 ONTAP 快照之前创建一致性的方式：
 - QEMU 来宾代理冻结/解冻，用于来宾文件系统静止和 Windows VSS 集成。
3. 转储 **VM** 配置文件。
4. 获取 **ONTAP** 快照：在虚拟机处于一致状态时捕获数据存储卷的 ONTAP 快照。这将成为保留的备份副本。
5. 删除临时 **Proxmox** 快照：在 ONTAP 快照完成后删除 Proxmox 快照，以防止增量文件增长。
6. 应用保留和复制策略（使用策略）：如果启用了灾难恢复复制，则应用配置的 ONTAP 保留策略并复制快照集。

这种方法通过仅存储更改的信息来最大限度地减少存储空间和网络带宽。使用 Snapshot 副本可快速复制虚拟机或数据存储，而不会影响性能，然后使用 NetApp SnapMirror 技术将其发送到辅助系统，以实现长期的异地数据保护。

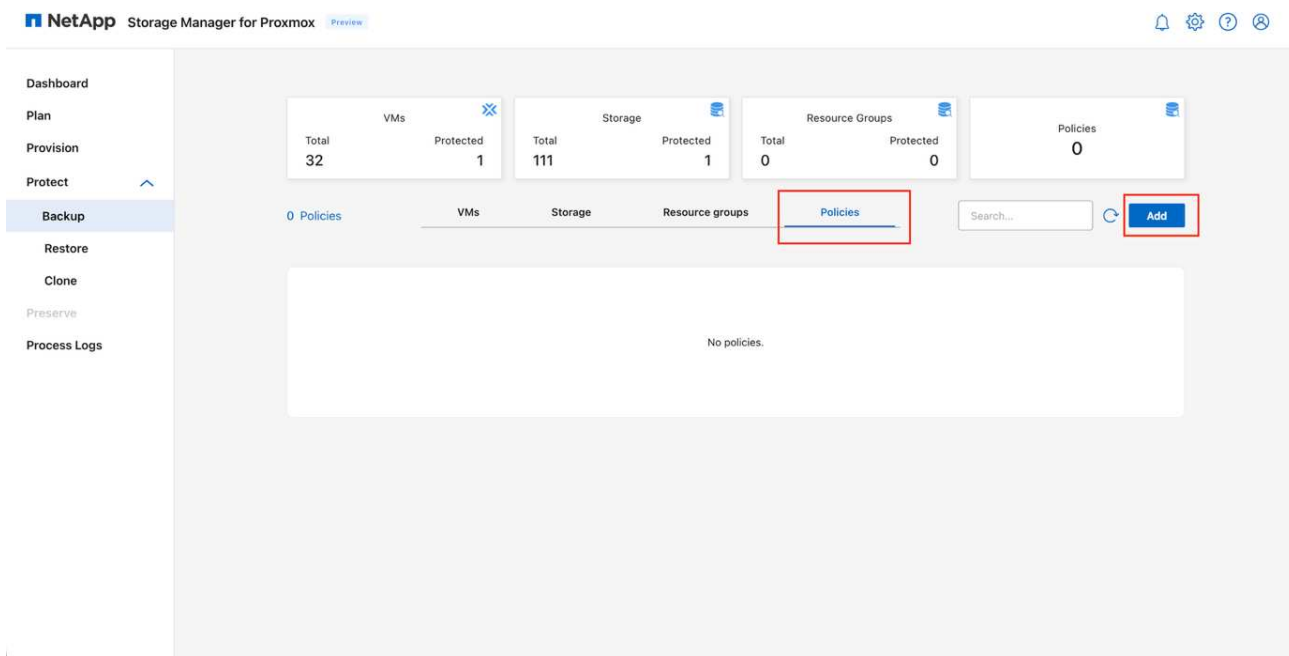
自动化 VM 备份任务

通过配置计划自动执行 Proxmox VM 备份。创建定期备份对于数据保护和灾难恢复至关重要。

步骤

1. 转到 **Protect > Backup > Policies**，然后单击 **Add** 以添加新的备份策略和计划。

▼ 查看备份策略页面



2. 配置常规设置：

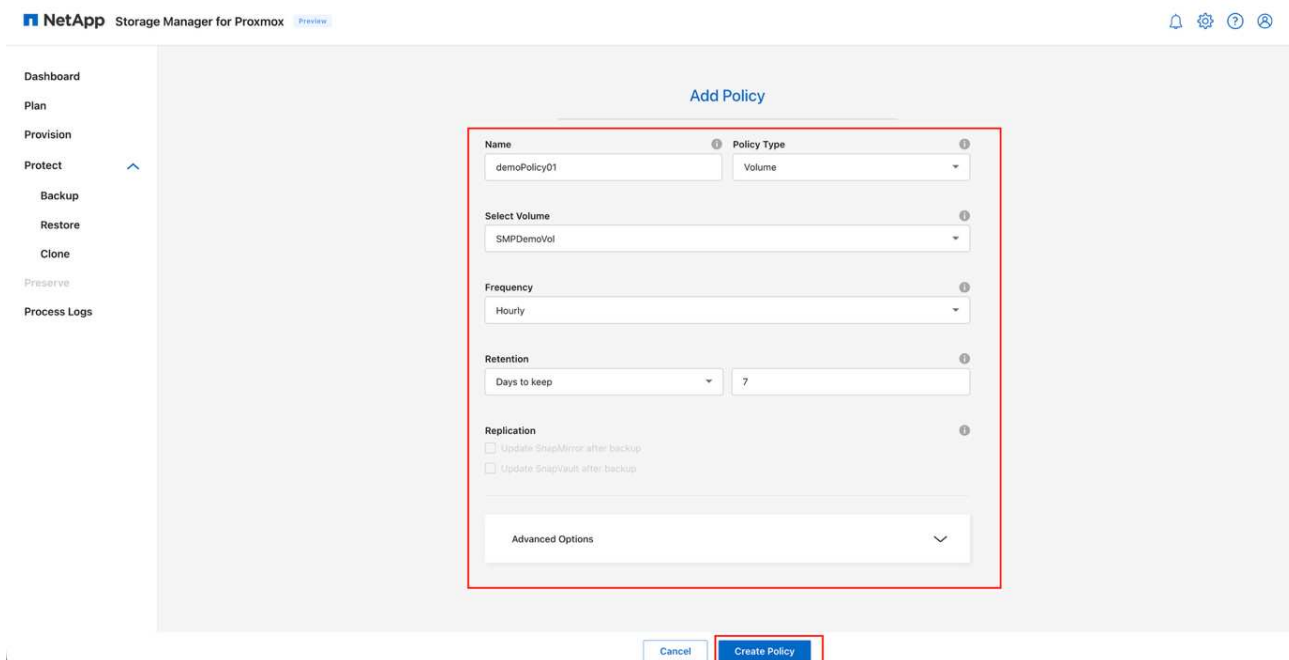
- 输入策略的名称
- 选择 策略类型（卷或资源组）
- 根据选定的类型选择 **Target**
- 设置 备份频率（每小时、每天、每周或每月）
- 使用"保留天数"或"保留数量"配置*保留*（默认保留值为 7）
- 选择 复制模式（在当前版本中禁用）：
 - 备份后更新 Snapmirror
 - 备份后更新 SnapVault



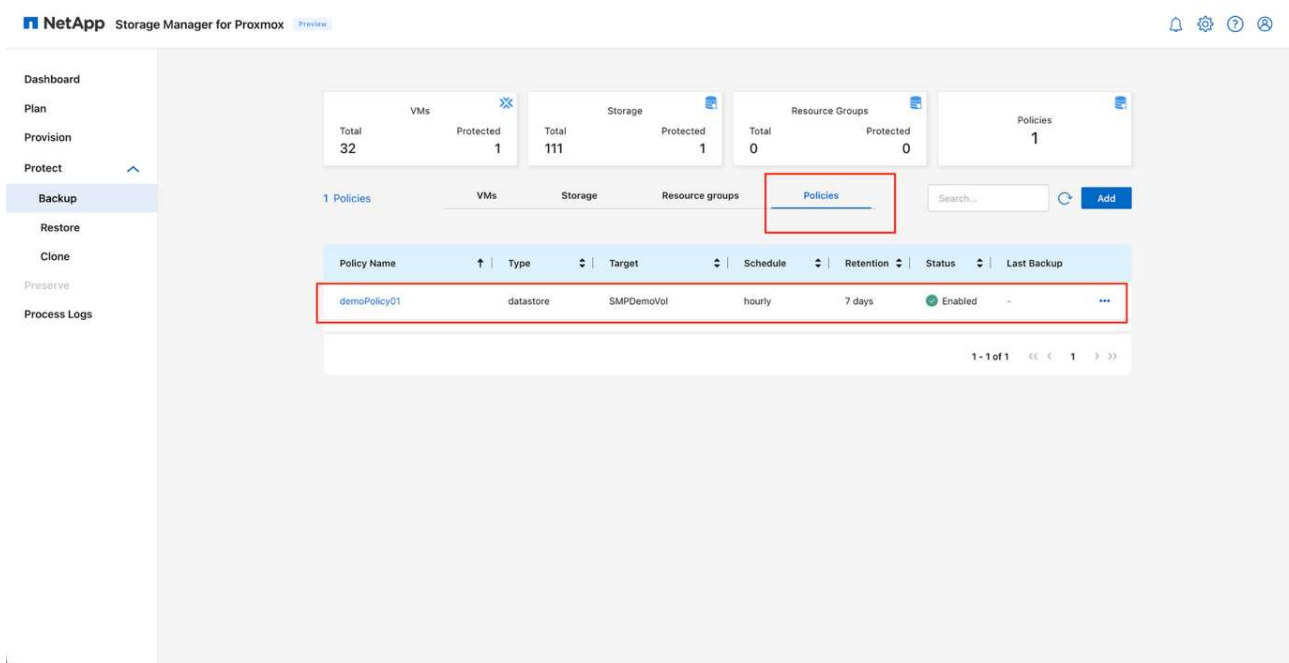
必须使用 CLI 或 System Manager 在 ONTAP 中配置 SnapMirror 复制。

- 在 **Advanced Options** 中，如果应在触发 ONTAP 快照之前使用 Proxmox 快照或 fsfreeze，请启用 **VM Consistency**。
- 单击 **Create Policy** 保存配置。

▼ 查看备份策略：计划和保留



▼ 查看备份策略：复制和高级选项



VM 备份作业也可以通过单击 **Backup** 立即触发，而不是等待资源组的计划执行时间。

还原虚拟机或数据存储库

使用 SMP 从备份还原 Proxmox VM。

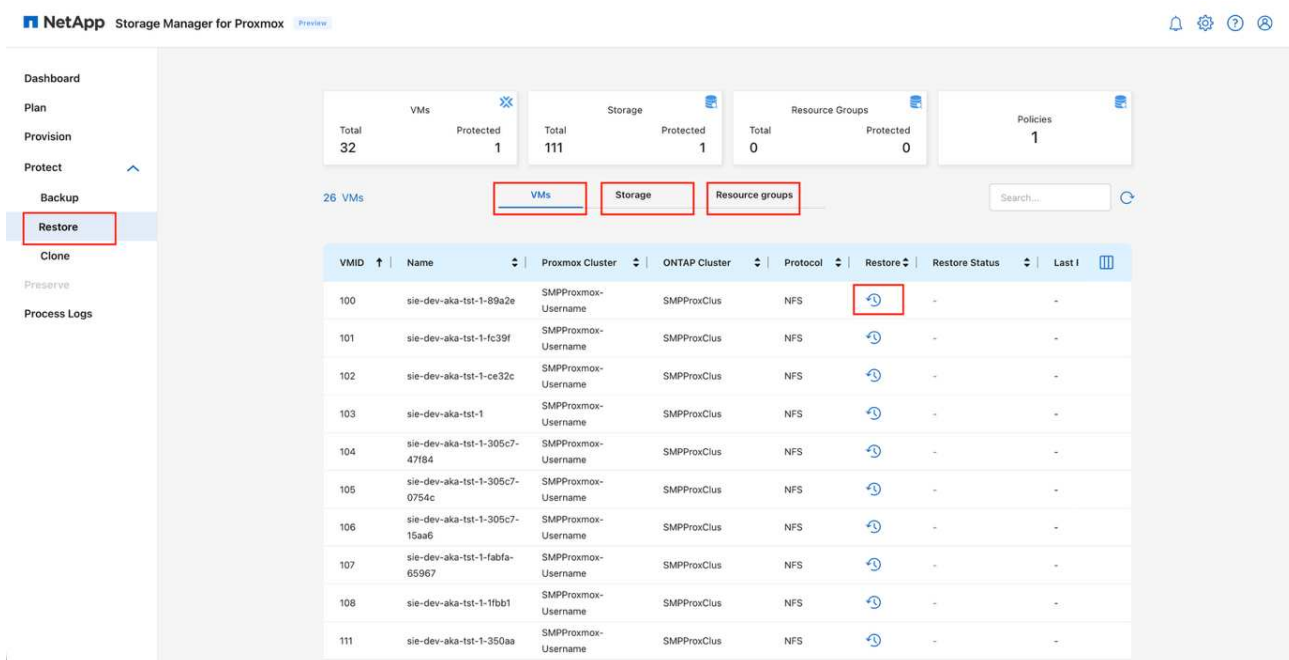
开始之前

在执行还原操作之前，必须关闭虚拟机的电源。

步骤

1. 在 Web 界面左侧的导航窗格中选择 **Restore** 选项。
2. 选择需要还原的虚拟机、存储或资源组。

▼ 查看还原页面



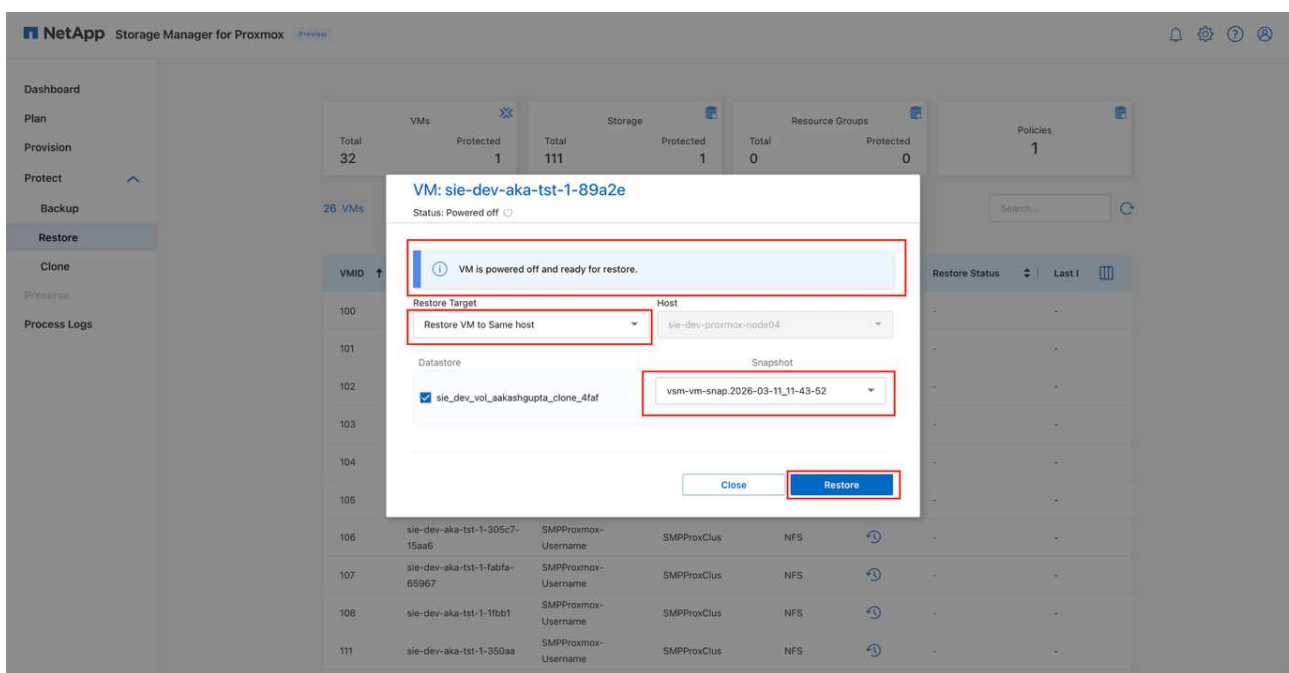
3. 单击 **Restore** 图标。

根据所选内容（无论是 VM、存储还是资源组），还原弹出窗口将打开。



根据所选资源，将显示为相应组件创建的快照。例如，在资源组级别创建的快照将仅在选择资源组时显示。这同样适用于存储。但是，对于虚拟机，将显示跨组件创建的所有快照以供选择。

▼ 查看还原对话框：快照选择

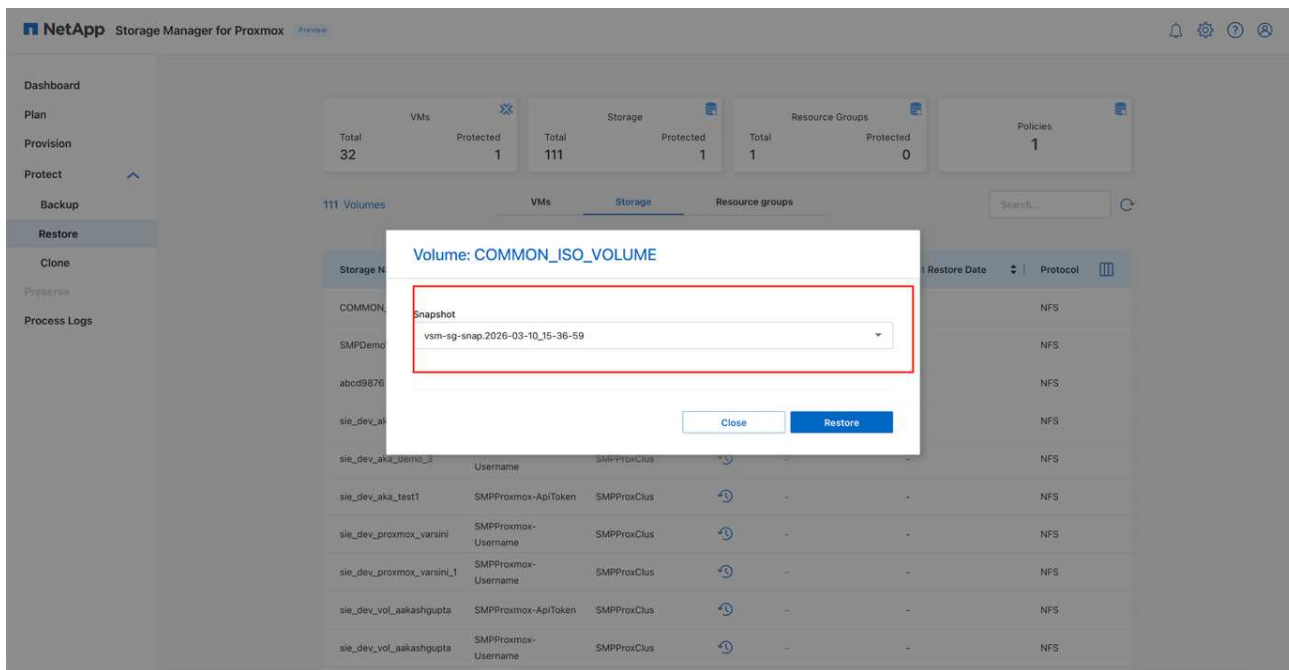


4. 选择所需选项：

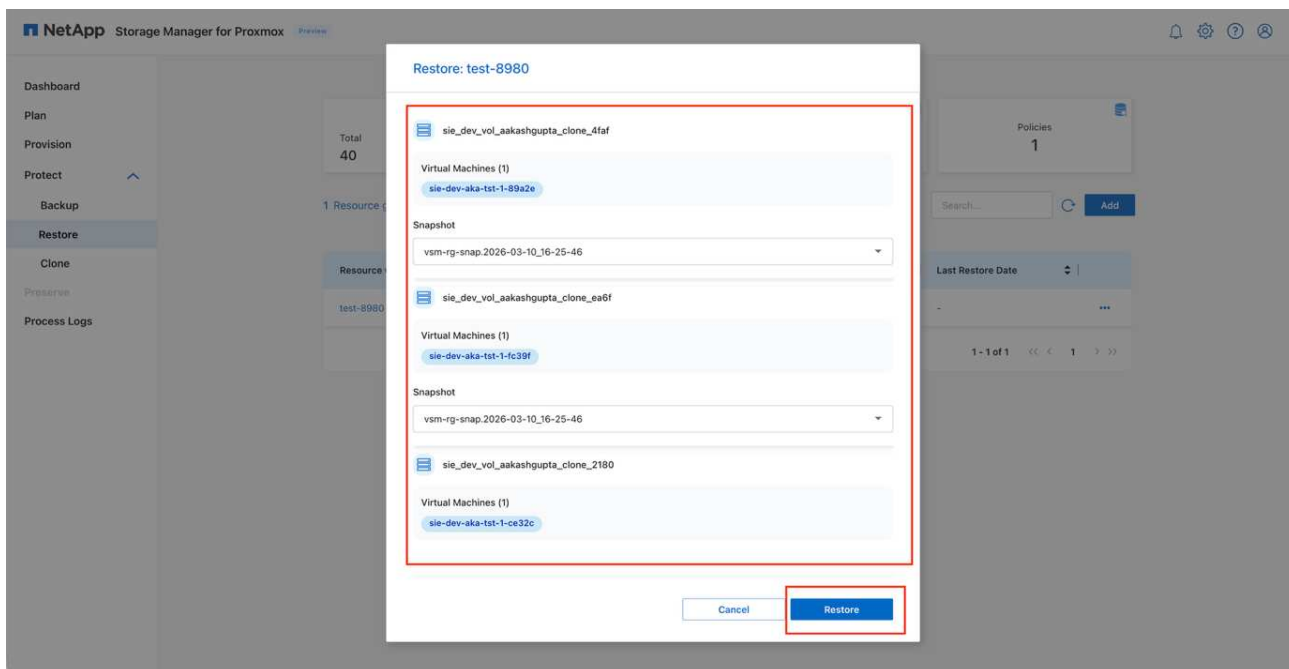
- a. 还原目标：还原到同一主机或备用主机的选项
- b. **Host**：如果选择了备用主机，则可以通过下拉菜单选择集群中的任何可用主机
- c. **Snapshot**：为时间点恢复选择适当的快照

5. 单击 还原。

▼ 查看正在进行的还原

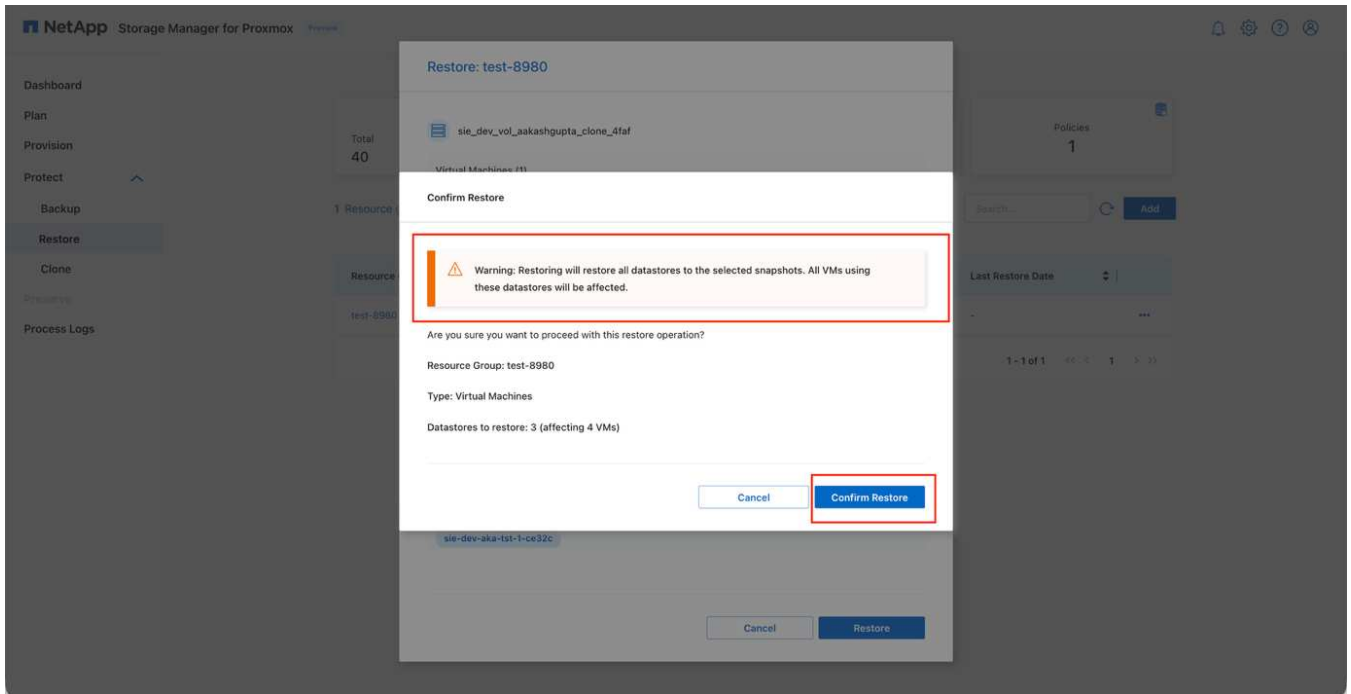


▼ 查看还原已完成



启动存储级还原时，该卷或 LUN 上的所有虚拟机都将断电，系统将作为一个整体还原整个存储实体。

▼ 查看存储级还原



还原 workflow 详细信息

还原 workflow 因基础存储类型而异，涉及以下关键步骤：

1. 定义还原范围：选择保护目标是数据存储区（ONTAP 卷）、资源组（多个卷或 VM）还是单个 VM。
2. 执行还原操作：VM 还原过程因存储协议而异：
 - 基于 **NFS** 的恢复：在 NFS 上恢复卷或数据存储时，SMP 会在目标卷上启动卷快照恢复操作。对于托管在 NFS 卷上的虚拟机，SMP 会验证还原参数，包括 VM ID、快照名称、站点 ID 和存储类型。如果 VM 正在运行，SMP 通过轮询其状态并发出停止命令来停止它。然后，系统使用文件级还原操作从快照中还原 VM 磁盘。
 - 基于 **iSCSI** 的恢复：对于 iSCSI 数据存储或 LUN，SMP 从快照恢复 LUN，验证恢复的设备，跨 Proxmox 节点执行 iSCSI 重新扫描，并确保设备可访问性。对于驻留在 iSCSI 存储上的 VM，SMP 克隆快照 LUN，将克隆的 LUN 映射到 Proxmox 主机，并执行 iSCSI 发现以导入它们。然后添加恢复的存储，并更新 VM 配置。VM 磁盘被移回其原始存储，每个 VM 都跟踪进度。该过程以清理活动结束，包括移除临时存储和删除克隆的 LUN。

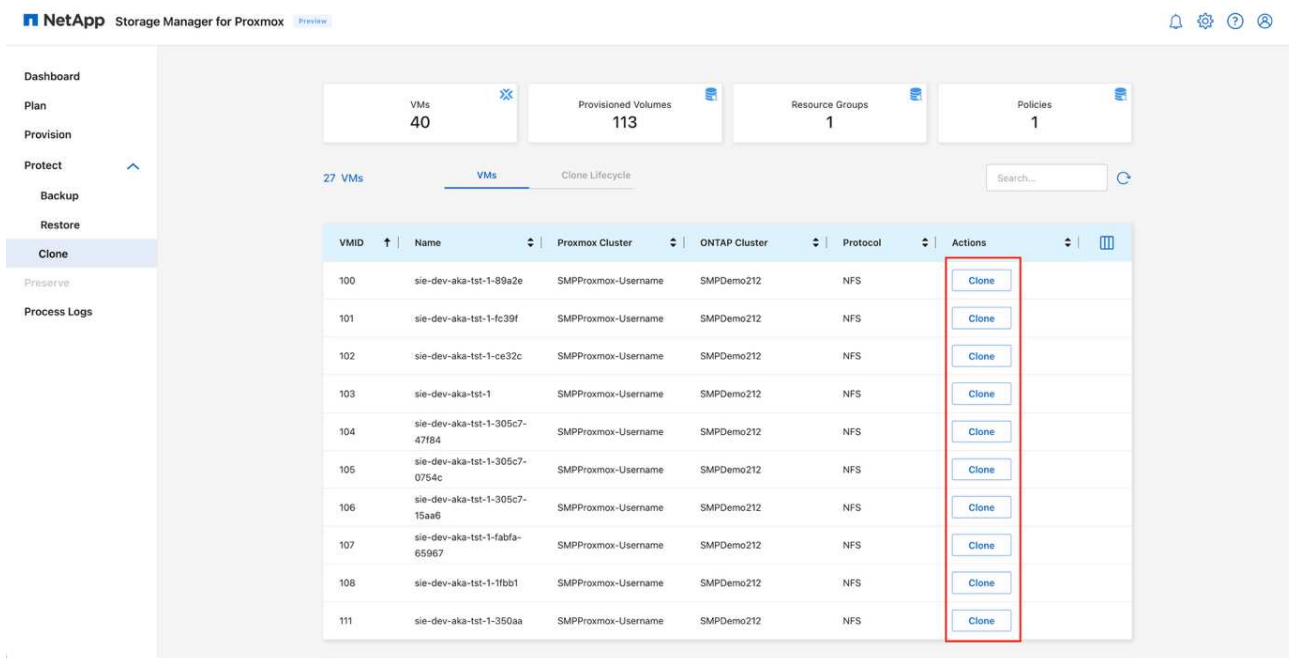
克隆虚拟机

使用 SMP 创建虚拟机克隆以进行开发、测试或快速配置。

步骤

1. 打开 **Protect** 部分，然后从左侧导航菜单中选择 **Clone**。
2. 在虚拟机列表中，找到所需的虚拟机，然后单击 Actions 列中的 **Clone**。
3. 在克隆虚拟机窗口中，选择创建克隆的目标节点。

▼ 查看克隆 VM 对话框



4. 选择要用作克隆源的 **Snapshot**。

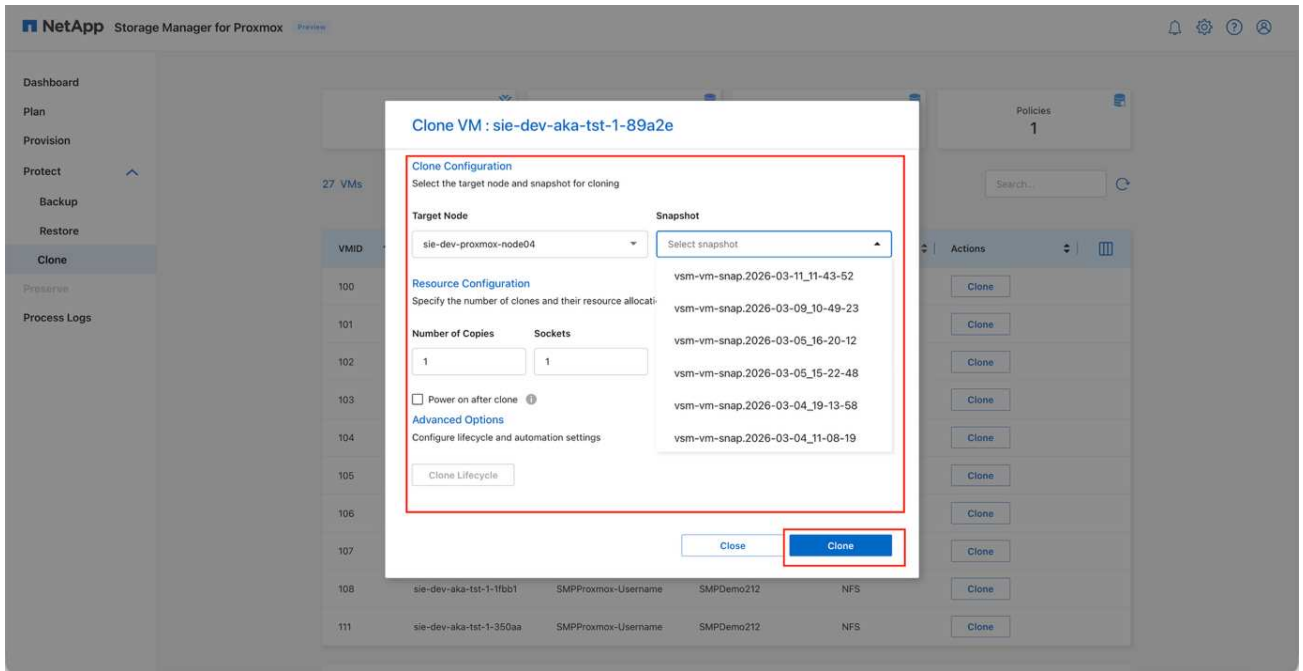


在当前版本中，对活动文件系统执行克隆。在即将发布的版本中，克隆 workflow 将允许选择快照，在这种情况下，使用从所选快照创建的新卷执行文件克隆。

- 指定克隆虚拟机的*副本数*、**CPU** 核心、套接字*和*内存。
- (可选) 如果克隆的虚拟机应在创建后自动启动，请启用 **Power on after clone**。
- (可选) 如果克隆应放置在新创建的数据存储上，请选择 **Create a new datastore**。
- 在 **高级选项** 下，如果网络设置应自动应用于克隆，请启用 **网络配置**。
- 单击 **Clone** 开始克隆过程。

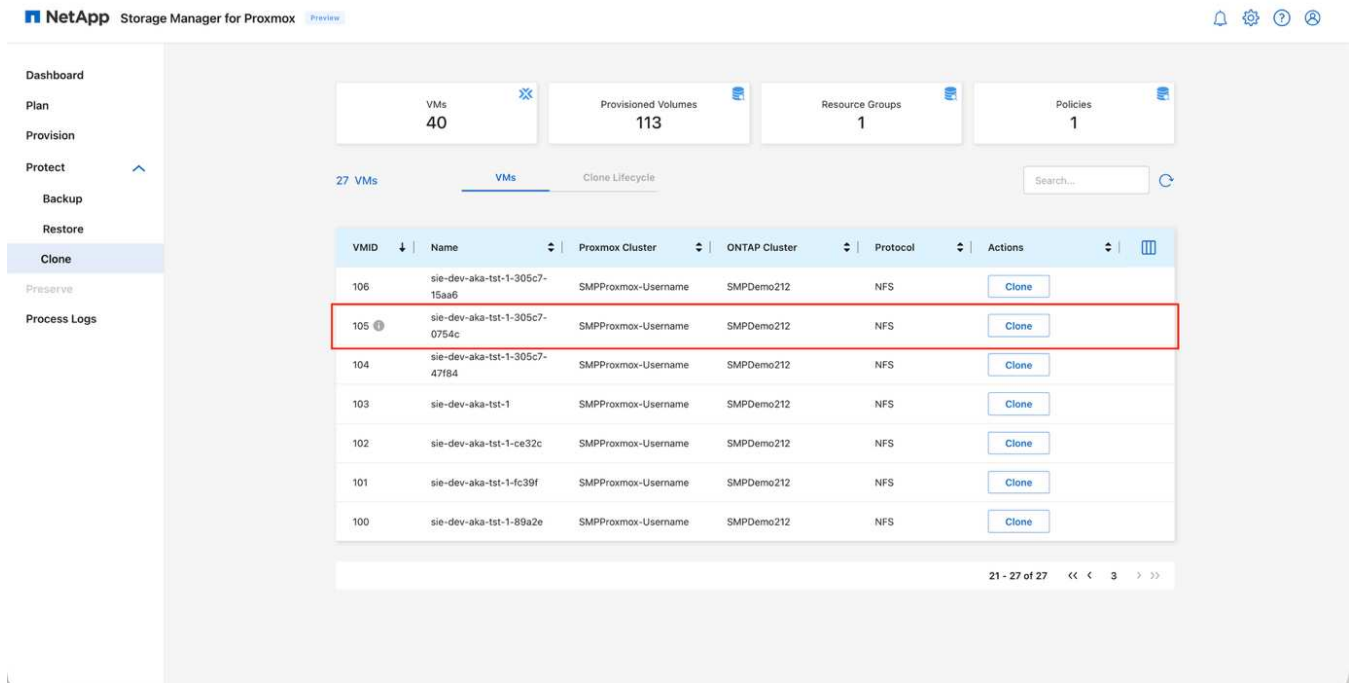
克隆 workflow 使用 ONTAP 的快速克隆 API 配置新虚拟机，并为克隆实例分配唯一的五位数 GUID。

▼ 查看克隆操作



克隆的虚拟机同时出现在备份和克隆列表中，可以使用现有备份策略进行保护。要使克隆的虚拟机在这些列表中可见，请对 Proxmox 和 ONTAP 计划运行发现。

▼ 在列表中查看克隆的 VM



克隆 workflow 详细信息

SMP 通过将繁重的工作卸载到 ONTAP 来执行存储端克隆而不是主机端文件复制，从而实现快速 VM 克隆的自动化。在底层，SMP 利用 ONTAP FlexClone 语义（基于快照、可写、节省空间的克隆），因此克隆可以快速创建，最初消耗的容量很少；它尽可能避免通过 Proxmox 主机进行数据读取/写入。

这种方法通过将克隆工作推入 ONTAP，提供适合开发/测试、实验室和快速刷新周期的快速、精简、存储效率高的克隆，从而加速 VM 配置。

默认情况下，对活动文件系统执行克隆。克隆 workflow 还允许选择快照，在这种情况下，使用从所选快照创建的新卷执行文件克隆。

监控操作

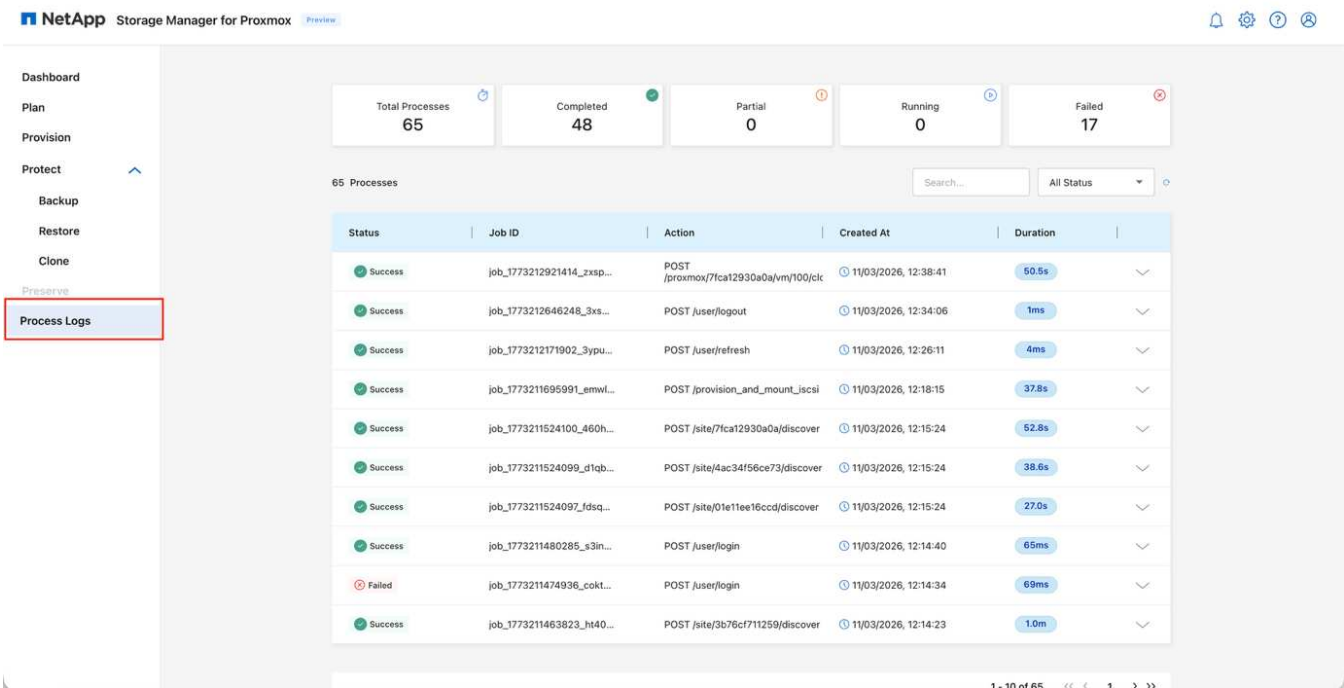
使用过程日志和仪表板视图来监控和跟踪 SMP 操作。

处理日志

Process Logs 视图提供了通过 Storage Manager for Proxmox 执行的所有操作的合并记录。每个条目显示作业状态、作业 ID、执行的操作、开始时间和持续时间。这样可以轻松跟踪已完成的任务，识别任何失败或部分操作，并查看在环境中执行的操作顺序。

筛选器和搜索选项可用于快速定位特定作业或按状态缩小列表。

▼ 查看 Process Logs 视图



信息板

仪表板一目了然地概述了整个 Storage Manager for Proxmox 环境。它总结了关键清单，例如 Proxmox 群集、主机、数据存储区、ONTAP 群集和受保护虚拟机的数量。还会显示容量使用情况、总存储卷、保护策略和资源组，以清晰显示系统利用率。

群集健康状况突出显示，指示所有连接节点的状态。仪表板的下半部分突出显示虚拟机分配（总计、运行和停止），并提供用于内存、存储和 CPU 的实时 Proxmox 资源利用率。此视图有助于快速评估系统运行状况、识别资源压力并监控整体运营状态。

▼ 查看信息板

Dashboard

Plan

Provision

Protect

Backup

Restore

Clone

Preserve

Process Logs

Proxmox Cluster

1

Hosts

3

Datastores

81

ONTAP Clusters

2

SVM

19

Capacity Used



Storage Volumes

512

Protection Policies

1

Restored Count

0

Cluster Health

Healthy

Nodes 3/3 Online

SMPProxClus 3/3

Resource Group

2

Virtual Machines



Proxmox Utilization



摘要

Storage Manager for Proxmox 提供了一种简化和统一的方式来管理 Proxmox 环境中的保护、克隆和恢复操作。通过与 ONTAP 直接集成，该解决方案将企业级存储的效率、可靠性和性能融入到日常 VM 工作流程中。ONTAP 的节省空间的快照、快速克隆功能和内置复制提供了一致和可预测的数据保护，而 SMP 将这些功能简化为直观的、以任务为中心的界面。

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。