最佳实践建议

NetApp Solutions

NetApp
September 26, 2024

目录

最佳实践建议

针对Red Hat OpenShift虚拟化中VM的最佳实践建议

作者：Banu Sunzhar、NetApp

本节介绍在部署新VM或将现有VM从VMware vSphere导入到OpenShift容器平台上的OpenShift虚拟化时应考虑的不同因素。

虚拟机性能

在OpenShift虚拟化中创建新VM时、您需要考虑要在VM上运行的工作负载的访问模式以及性能(IOPS和吞吐量)要求。这将影响在OpenShift容器平台中OpenShift虚拟化上运行所需的VM数量以及VM磁盘所需使用的存储类型。

要为VM磁盘选择的存储类型受以下因素影响：

- 访问工作负载的数据所需的协议访问
- 所需的访问模式(rwo与rwx)
- 工作负载所需的性能特征

有关详细信息、请参见下面的"存储配置"部分。

VM工作负载的高可用性

OpenShift虚拟化支持实时迁移虚拟机。通过实时迁移、可以在不中断工作负载的情况下将正在运行的虚拟机实例(Virtual Machine Instance、VM)移至另一个节点。迁移有助于在集群升级期间或在需要耗尽节点以进行维护或配置更改时实现平稳过渡。实时迁移需要使用一个共享存储解决方案、该解决方案可提供Read任意(rwx)访问模式。VM磁盘应使用提供rwx访问模式的存储选项作为后备存储。OpenShift虚拟化将检查一个VIF是否为实时迁移，如果是， **evicalStrategy**将设置为实时迁移。有关详细信息、请参见。 "有关实时迁移一节"

使用支持**rwx**访问方式的驱动程序非常重要。有关哪些ONTAP驱动程序支持rwx访问模式的详细信息、请参见下面的存储配置部分。

存储配置

Trident CSI配置程序为配置NetApp存储选项支持的存储提供了多种驱动程序(NAS、NAS经济型、NAS经济型、FlexGroup、SAN和SAN经济型)。

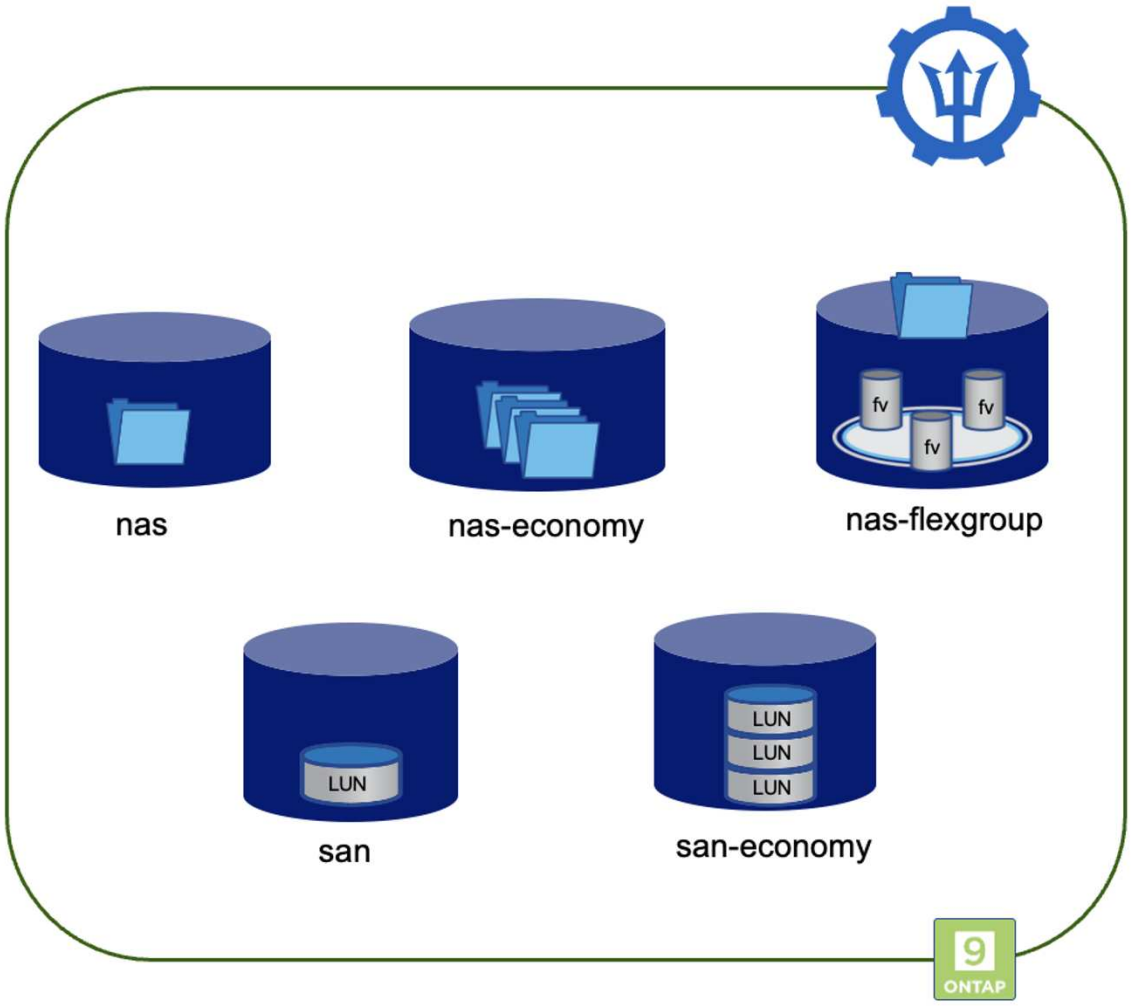
使用的协议：* NAS驱动程序使用NAS协议(NFS和SMB)* SAN驱动程序使用iSCSI或NVMe/TCP协议

以下内容可帮助您根据工作负载要求和存储利用率确定存储配置的方式。

- **NAS** 驱动程序在一个FlexVolume上创建一个永久性卷(PV)。
- **NAS**经济驱动程序在共享FlexVolume的qtree上创建一个PV。(每200个PIV对应一个FlexVolume、可在50到300之间配置)
- **NAS**－**PV**驱动程序在一个FlexGroup的一个FlexGroup上创建

- SAN驱动程序会在专用FlexVolume上的LUN上创建一个PV
- **SAN**经济驱动程序在共享FlexVolume上的LUN上创建一个PV (每100个PV创建一个FlexVolume，可在50到200之间配置)

下图对此进行了说明。



此外、驱动程序支持的访问模式也不同。

- ONTAP NAS驱动程序支持**
  - 文件系统访问和RwO、ROX、rwx、RWONP访问模式。
- ONTAP SAN驱动程序支持原始块和文件系统模式**
  - 在原始块模式下，它可以支持RWO、ROX、Rwx、RWONP访问模式。
  - 在文件系统模式下，只允许使用RwO、RwoP访问模式。

要实时迁移OpenShift虚拟化VM、磁盘必须具有rwx访问模式。因此、在原始块卷模式下选择NAS驱动程序或SAN驱动程序以创建由ONTAP提供支持的PVC和PV非常重要。

存储配置最佳实践

专用**Storage Virtual Machine (SVM)**

Storage Virtual Machine （ SVM ）可在 ONTAP 系统上的租户之间实现隔离和管理隔离。通过将SVM专用于OpenShift容器和OpenShift虚拟化VM、可以委派Privileges并应用最佳实践来限制资源消耗。

限制**SVM**上的最大卷数

要防止 Trident 占用存储系统上的所有可用卷，您应对 SVM 设置限制。您可以从命令行执行此操作：

```
vserver modify -vserver <svm_name> -max-volumes <num_of_volumes>
```

最大卷数值是在ONTAP集群中的所有节点上配置的总卷数、而不是在单个ONTAP节点上配置的总卷数。因此，在某些情况下， ONTAP 集群节点所配置的 Trident 卷可能远远多于或少于其他节点。要避免这种情况、请确保为Trident使用的SVM分配的集群中每个节点的聚合数量相等。

限制**Trident**创建的卷的最大大小

您可以在ONTAP中为每个SVM设置最大卷大小限制：

1. 使用vserver create命令创建SVM并设置存储限制：

```
vserver create -vserver vserver_name -aggregate aggregate_name -rootvolume
root_volume_name -rootvolume-security-style {unix|ntfs|mixed} -storage
-limit value
```

1. 要修改现有SVM上的存储限制、请执行以下操作：

```
vserver modify -vserver vserver_name -storage-limit value -storage-limit
-threshold-alert percentage
```

> 不能为包含数据保护卷、SnapMirror关系中的卷或MetroCluster 配置中的任何SVM配置存储限制。

除了控制存储阵列上的卷大小之外，您还应利用 Kubernetes 功能。

1. 要配置可由Trident创建的卷的最大大小，请使用backend.json定义中的 **LimitvolumeSize**参数。
2. 要配置用作ONTAP SAN经济型驱动程序和ONTAP NAS经济型驱动程序池的FlexVol的最大大小、请在backend.json定义中使用 **LimitVolumePoolSize**参数。

使用**SVM QoS**策略

将服务质量(QoS)策略应用于SVM、以限制Trident配置的卷可使用的IOPS数量。这有助于防止使用Trident配置的存储的工作负载影响Trident SVM外部的工作负载。

ONTAP QoS策略组可为卷提供QoS选项、并使用户能够为一个或多个工作负载定义吞吐量上限。有关QoS策略组的详细信息、请参阅"ONTAP 9.15 QoS 命令"

限制对**Kubbernetes**集群成员的存储资源访问

使用命名空间限制对Trident创建的NFS卷和iSCSI LUN的访问是Kubenetes部署安全防护的一个重要组成部分。这样可以防止不属于 Kubernetes 集群的主机访问卷并可能意外修改数据。

此外、容器中的进程可以访问挂载到主机但并非用于容器的存储。使用命名空间为资源提供逻辑边界可以避免此问题。但是、

请务必了解命名空间是 Kubernetes 中资源的逻辑边界。因此、请务必确保在适当时使用名称空间进行分隔。但是、运行具有特权的容器时所使用的主机级权限明显多于正常情况。因此，请使用禁用此功能"POD 安全策略"。

使用专用导出策略对于具有专用基础架构节点或无法计划用户应用程序的其他节点的OpenShift部署，应使用单独的导出策略进一步限制对存储资源的访问。其中包括为部署到这些基础架构节点的服务（例如 OpenShift 指标和日志记录服务）以及部署到非基础架构节点的标准应用程序创建导出策略。

{\f270可以自动创建和管理导出策略} {\f151。} 通过这种方式， Trident 会限制对其配置给 Kubernetes 集群中节点的卷的访问，并简化节点的添加 / 删除。

但是、如果您选择手动创建导出策略、请使用一个或多个导出规则来处理每个节点访问请求。

为应用程序**SVM**禁用**showmount**部署到Kubbernetes集群的Pod可以对数据LIF发出showmount -e命令，并接收可用挂载列表，包括它无权访问的挂载。要防止出现这种情况、请使用以下命令行界面禁用showmount功能：

```
vserver nfs modify -vserver <svm_name> -showmount disabled
```

有关存储配置和Trident使用最佳实践的其他详细信息、请查看"Trident 文档"

OpenShift虚拟化-调整和扩展指南

Red Hat已对此进行了说明"OpenShift集群扩展建议和限制"。

此外，它们还记录了"OpenShift虚拟化调整指南"和"支持的OpenShift虚拟化4.x限制"。

要访问上述内容、需要有效的Red Hat订阅。

调整指南包含许多调整参数的相关信息、包括：

- 调整参数以一次性或批量创建多个VM
- 实时迁移VM
- "为实时迁移配置专用网络"
- 通过包含工作负载类型自定义VM模板

支持的限制记录了在OpenShift上运行VM时测试的对象最大值

虚拟机最大值，包括

- 每个VM的最大虚拟CPU数
- 每个VM的最大和最小内存
- 每个VM的最大单磁盘大小
- 每个VM的最大热插拔磁盘数

最多主机数，包括*同时实时迁移(按节点和集群)

集群最大值，包括最大已定义VM数

从**VMware**环境迁移**VM**

适用于OpenShift虚拟化的迁移工具包是Red Hat提供的一个操作员、可从OpenShift容器平台的OperatorHub获得。此工具可用于从vSphere、Red Hat虚拟化、OpenStack和OpenShift虚拟化迁移虚拟机。

有关从vSphere迁移VM的详细信息、请参见"工作流 ONTAP的Red Hat OpenShift虚拟化"

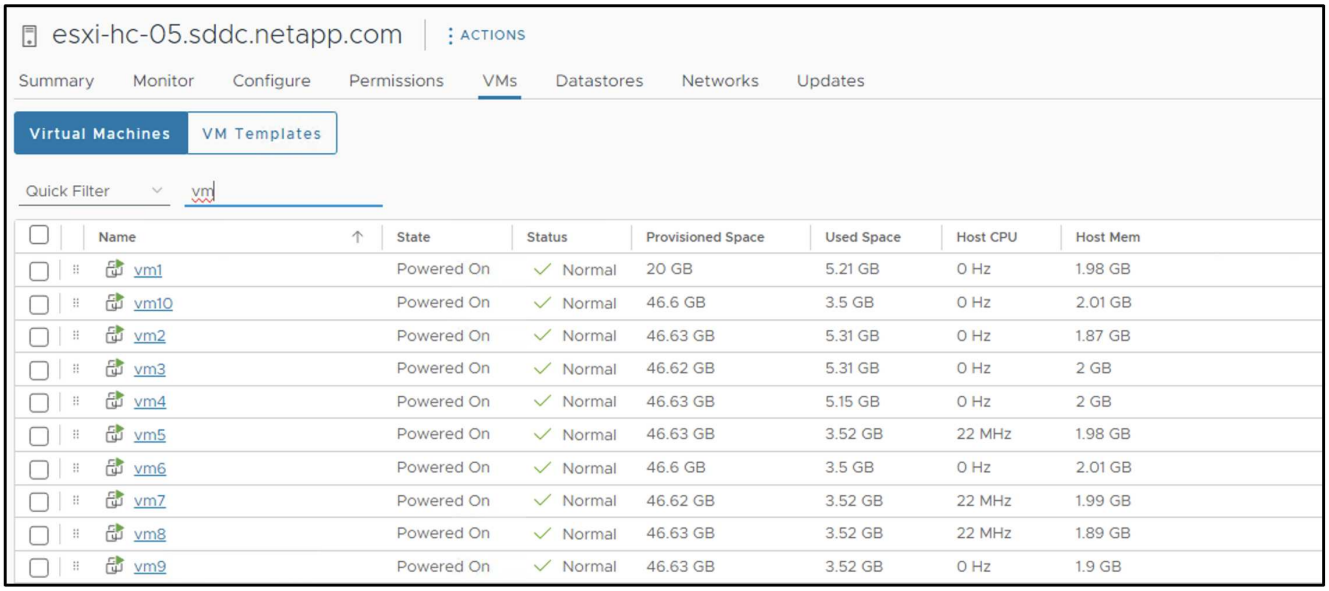
您可以从命令行界面或迁移Web控制台为各种参数配置限制。下面提供了一些示例

1. 最多并发虚拟机迁移数用于设置可同时迁移的最大虚拟机数。默认值为20个虚拟机。
2. 预复制间隔(分钟)用于控制在启动热迁移之前请求新快照的间隔。默认值为60分钟。
3. Snapshot轮询间隔(秒)用于确定系统在oVirt热迁移期间检查快照创建或删除状态的频率。默认值为10秒。

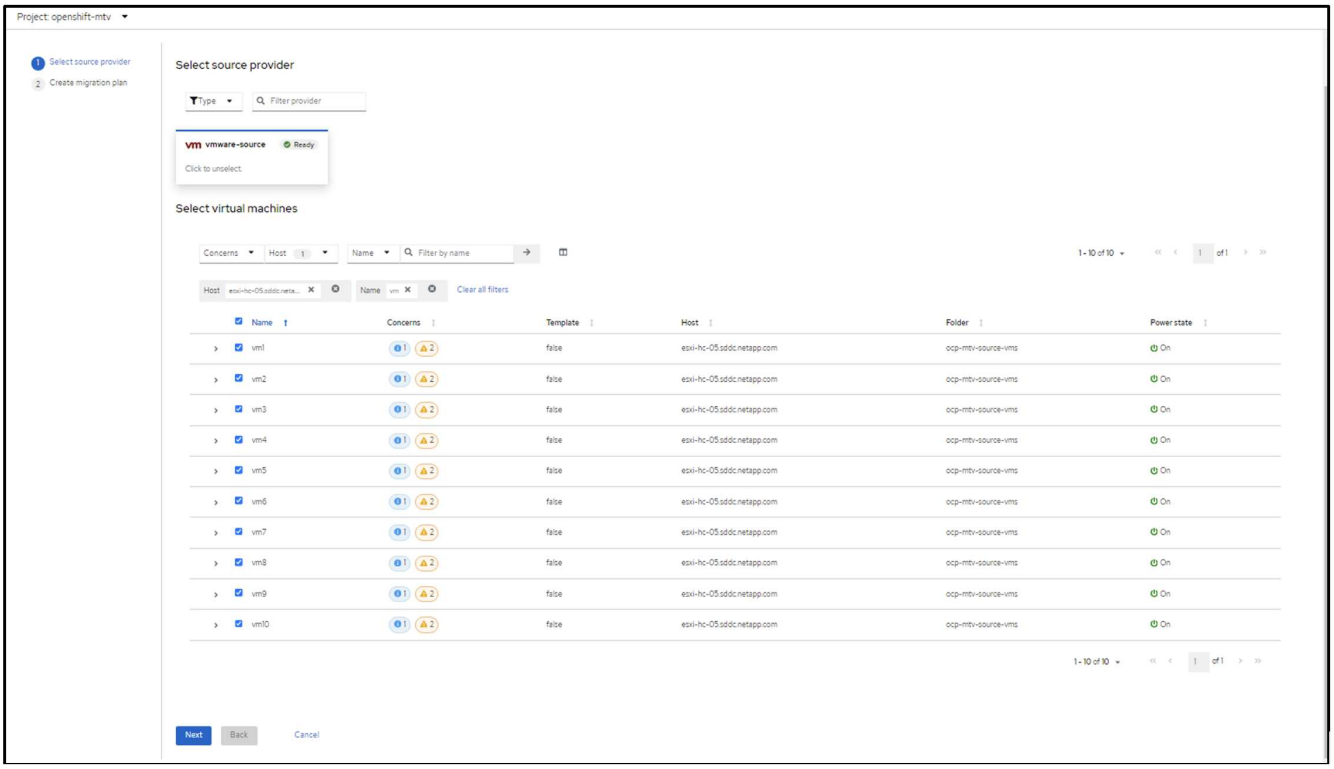
如果要在同一迁移计划中从ESXi主机迁移10个以上的VM、则必须增加此主机的NFC服务内存。否则、迁移将失败、因为NFC服务内存限制为10个并行连接。有关更多详细信息、请参见Red Hat文档："增加ESXi主机的NFC服务内存"

下面将使用适用于虚拟化的迁移工具包成功地将10个VM从vSphere中的同一主机并行迁移到OpenShift虚拟化。

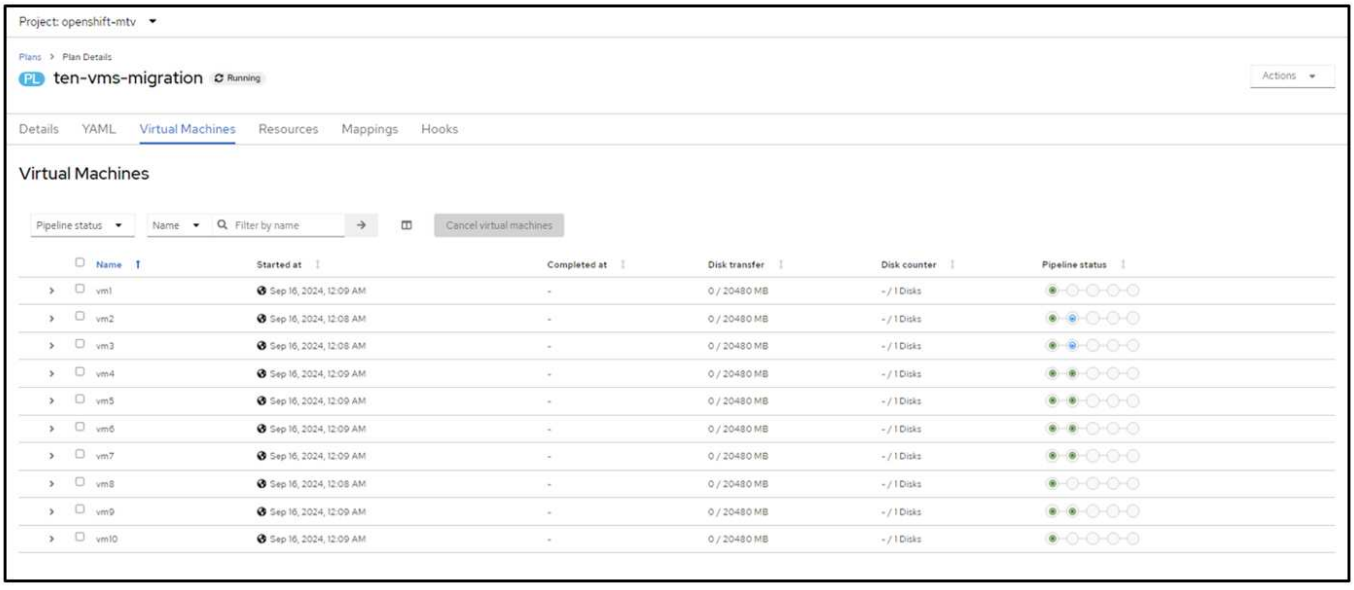
同一**ESXi**主机上的**VM**



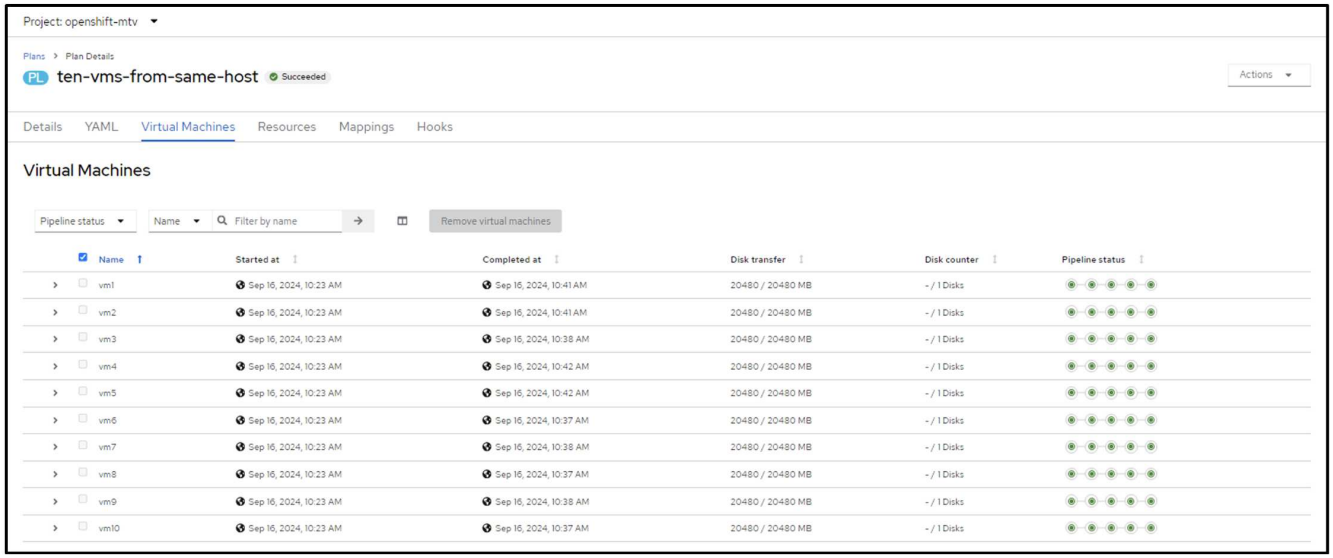
首先制定了从**VMware**迁移10个VM的计划



迁移计划已开始执行



所有**10**个**VM**均已成功迁移



在**OpenShift**虚拟化中，所有**10**个**VM**均处于运行状态