



ONTAP 配置

Enterprise applications

NetApp
May 09, 2024

目录

ONTAP 配置	1
RAID和Oracle数据库	1
Oracle数据库和存储容量管理	1
Oracle数据库和Storage Virtual Machine	2
使用ONTAP QoS进行Oracle数据库性能管理	2
Oracle数据库和ONTAP效率功能	3
使用Oracle数据库进行精简配置	6
Oracle数据库和ONTAP控制器故障转移/切换	8

ONTAP 配置

RAID和Oracle数据库

RAID是指使用冗余来保护数据不受驱动器丢失的影响。

在配置用于Oracle数据库和其他企业应用程序的NetApp存储时、有时会出现有关RAID级别的问题。有关存储阵列配置的许多传统Oracle最佳实践都包含有关使用RAID镜像和(或)避免使用某些类型的RAID的警告。尽管这些来源可以提供有效的支持、但它们并不适用于RAID 4以及ONTAP中使用的NetApp RAID DP和RAID-TEC技术。

RAID 4、RAID 5、RAID 6、RAID DP和RAID-TEC均使用奇偶校验来确保驱动器故障不会导致数据丢失。与镜像相比、这些RAID选项提供的存储利用率要高得多、但大多数RAID实施都有一个影响写入操作的缺点。在其他RAID实施中完成写入操作可能需要多次驱动器读取才能重新生成奇偶校验数据、此过程通常称为RAID惩罚。

但是、ONTAP不会受到这种RAID惩罚。这是因为NetApp WAFL (任意位置写入文件布局)与RAID层集成在一起。写入操作会在RAM中进行聚合、并准备为完整的RAID条带、包括奇偶校验生成。ONTAP无需执行读取即可完成写入、这意味着ONTAP和WAFL可以避免RAID惩罚。重做日志记录等延迟关键型操作的性能不受阻碍、随机数据文件写入不会因需要重新生成奇偶校验而产生任何RAID影响。

在统计可靠性方面、即使RAID DP也能提供比RAID镜像更好的保护。主要问题是RAID重建期间对驱动器的需求。使用镜像RAID集时、在重建到RAID集中的配对驱动器时、由于驱动器故障而导致数据丢失的风险远高于RAID DP集中三驱动器故障的风险。

Oracle数据库和存储容量管理

要使用可预测、可管理的高性能企业存储管理数据库或其他企业应用程序、需要在驱动器上留出一些可用空间来管理数据和元数据。所需的可用空间量取决于使用的驱动器类型和业务流程。

可用空间是指未用于实际数据的任何空间、其中包括聚合本身上的未分配空间以及成分卷中的未使用空间。此外、还必须考虑精简配置。例如、卷可能包含1 TB的LUN、其中实际数据仅利用了50%的空间。在精简配置环境中、这将正确地显示为占用500 GB的空间。但是、在完全配置的环境中、1 TB的全部容量似乎正在使用中。500 GB的未分配空间将被隐藏。实际数据未使用此空间、因此应将此空间计入总可用空间。

对于用于企业级应用程序的存储系统、NetApp建议如下：

SSD聚合、包括AFF系统



* NetApp建议*至少留出10%的可用空间。这包括所有未使用的空间、包括聚合或卷中的可用空间、以及由于使用完全配置而分配但实际数据未使用的任何可用空间。逻辑空间并不重要、问题是可用于数据存储的实际可用物理空间有多少。

建议10%的可用空间非常保守。SSD聚合可以支持利用率更高的工作负载、而不会对性能产生任何影响。但是、随着聚合利用率的提高、如果不仔细监控利用率、用尽空间的风险也会增加。此外、在以99%的容量运行系统时、可能不会影响性能、但可能会导致管理工作、试图在订购额外硬件时阻止系统完全填满、并且可能需要一些时间来采购和安装额外的驱动器。

HDD聚合、包括Flash Pool聚合



* NetApp建议*使用旋转驱动器时至少留出15%的可用空间。这包括所有未使用的空间、包括聚合或卷中的可用空间、以及由于使用完全配置而分配但实际数据未使用的任何可用空间。当言论自由接近10%时、性能将受到影响。

Oracle数据库和Storage Virtual Machine

Oracle数据库存储管理集中在Storage Virtual Machine (SVM)上

SVM (在ONTAP命令行界面中称为Vserver)是存储的基本功能单元、将SVM与VMware ESX服务器上的子系统进行比较非常有用。

首次安装时、ESX没有预配置的功能、例如托管操作系统或支持最终用户应用程序。在定义虚拟机(VM)之前、此容器为空容器。ONTAP与此类似。首次安装ONTAP时、在创建SVM之前、它无法提供数据。SVM特性用于定义数据服务。

与存储架构的其他方面一样、SVM和逻辑接口(LIF)设计的最佳选项在很大程度上取决于扩展要求和业务需求。

svms

目前尚无为ONTAP配置SVM的官方最佳实践。正确的方法取决于管理和安全要求。

大多数客户都会运行一个主SVM来满足大多数日常需求、但随后会创建少量SVM来满足特殊需求。例如、您可能希望创建：

- 由专业团队管理的关键业务数据库的SVM
- 开发组的SVM、该开发组拥有完全的管理控制权、因此可以独立管理自己的存储
- 用于敏感业务数据(例如人力资源或财务报告数据)的SVM、必须限制管理团队

在多租户环境中、可以为每个租户的数据提供一个专用SVM。每个集群、HA对和节点的SVM和LIP数量限制取决于所使用的协议、节点型号和ONTAP版本。请参见 "[NetApp Hardware Universe](#)" 的限制。

使用ONTAP QoS进行Oracle数据库性能管理

安全高效地管理多个Oracle数据库需要有效的QoS策略。原因在于现代存储系统的性能不断提高。

具体而言、随着全闪存存储的采用率不断提高、工作负载得以整合。由于旧旋转驱动器技术的IOPS功能有限、依赖旋转介质的存储阵列往往仅支持数量有限的I/O密集型工作负载。早在存储控制器达到其限制之前、一两个高度活跃的数据库就会使底层驱动器饱和。这种情况已经改变。即使是功能最强大的存储控制器、数量相对较少的SSD驱动器的性能也可能会饱和。这意味着、可以充分利用控制器的全部功能、而不必担心随着旋转介质延迟峰值而导致性能突然崩溃。

作为一个参考示例、一个简单的双节点HA AFF A800系统能够在延迟超过1秒之前提供高达100万次随机IOPS。只有极少数单个工作负载才能达到此级别。要充分利用此AFF A800系统阵列、需要托管多个工作负载、而安全地执行此操作、同时确保可预测性、则需要QoS控制。

ONTAP中有两种类型的服务质量(QoS)：IOPS和带宽。QoS控制可应用于SVM、卷、LUN和文件。

IOPS QoS

IOPS QoS控制显然基于给定资源的总IOPS、但IOPS QoS的许多方面可能并不直观。最初、一些客户对达到IOPS阈值时延迟明显增加感到很不明白。限制IOPS自然会导致延迟增加。从逻辑上讲、它的功能类似于令牌系统。例如、如果包含数据文件的给定卷具有10000 IOPS限制、则到达的每个I/O都必须先接收令牌才能继续处理。只要在给定的一秒内使用的令牌不超过10000个、就不会出现延迟。如果IO操作必须等待接收其令牌、则此等待将显示为额外的延迟。工作负载越难超过QoS限制、每个IO在队列中等待处理的时间就越长、这在用户看来是延迟越高。



对数据库事务/重做日志数据应用QoS控制时要小心。虽然重做日志记录的性能需求通常要比数据文件低很多、但重做日志活动会变得突发。IO会以短暂的脉冲发生、并且似乎适合平均重做IO级别的QoS限制对于实际要求可能过低。结果可能会造成严重的性能限制、因为QoS会与每个重做日志突发事件结合使用。通常、重做和归档日志记录不应受QoS的限制。

带宽QoS

并非所有I/O大小都相同。例如、数据库可能会执行大量小块读取、从而导致达到IOPS阈值、但是、数据库可能还会执行完整的表扫描操作、其中包含非常少量的大型块读取、占用的带宽非常大、但IOPS相对较少。

同样、VMware环境可能会在启动期间产生大量随机IOPS、但在外部备份期间执行的IO会更少、但会更大。

有时、有效管理性能需要IOPS或带宽QoS限制、甚至两者都需要。

最低/有保障的QoS

许多客户都希望解决方案能够提供有保障的QoS、而这种服务质量比看起来更难实现、而且可能会造成大量浪费。例如、如果要将10个数据库放置在10000 IOPS保证下、则需要对系统进行规模估算、以应对所有10个数据库同时以10000 IOPS运行的情形、总共需要100K IOPS。

最低QoS控制的最佳用途是保护关键工作负载。例如、假设一个ONTAP控制器的最大可能IOPS为50万次、并混合了生产和开发工作负载。您应将最大QoS策略应用于开发工作负载、以防止任何给定数据库独占控制器。然后、您可以对生产工作负载应用最低QoS策略、以确保它们在需要时始终具有所需的可用IOPS。

自适应 QoS

自适应服务质量(QoS)是指ONTAP功能、其中服务质量(QoS)限制基于存储对象的容量。它很少用于数据库、因为数据库大小与其性能要求之间通常没有任何关联。大型数据库可能几乎处于无活动状态、而小型数据库则可能是IOPS密集型最高的数据库。

自适应QoS对于虚拟化数据存储库非常有用、因为此类数据集的IOPS要求往往与数据库的总大小相关。包含1 TB VMDK文件的较新数据存储库所需的性能可能是2 TB数据存储库的一半左右。自适应QoS允许您在数据存储库中填充数据时自动增加QoS限制。

Oracle数据库和ONTAP效率功能

ONTAP空间效率功能针对Oracle数据库进行了优化。在几乎所有情况下、最佳方法都是保留默认值并启用所有效率功能。

数据压缩、数据缩减和重复数据删除等空间效率功能旨在增加给定物理存储量所需的逻辑数据量。这样可以降低成本和管理开销。

从较高层面来看、数据压缩是一个数学过程、通过该过程、可以检测数据模式并对其进行编码、从而减少空间需求。相反、重复数据删除会检测实际重复的数据块并删除无关的副本。数据缩减允许多个逻辑数据块共享介质上的同一物理块。



有关存储效率与预留百分比之间交互的说明、请参见以下有关精简配置的章节。

压缩

在全闪存存储系统推出之前、基于阵列的数据压缩的价值有限、因为大多数I/O密集型工作负载都需要大量磁盘轴才能提供可接受的性能。由于驱动器数量众多、存储系统所含容量总是远远超出所需容量。随着固态存储的兴起、这种情况发生了变化。不再需要纯粹为了获得良好的性能而大量过度配置驱动器。存储系统中的驱动器空间可以与实际容量需求相匹配。

与旋转驱动器相比、固态驱动器(SSD)的IOPS功能提高几乎始终可以节省成本、但数据压缩可以通过增加固态介质的有效容量来进一步节省成本。

数据压缩方法有多种。许多数据库都具有自己的数据压缩功能、但在客户环境中很少出现这种情况。原因通常是对压缩数据*进行更改*会对性能造成影响、此外、对于某些应用程序、数据库级数据压缩的许可成本较高。最后、还会对数据库操作产生整体性能影响。为执行数据压缩和解压缩的CPU支付较高的每CPU许可证成本毫无意义、而不是实际的数据库工作。更好的选择是将压缩工作负载分流到存储系统。

自适应数据压缩

自适应数据压缩已针对企业级工作负载进行了全面测试、未观察到对性能的影响、即使在延迟以微秒为单位的闪存环境中也是如此。一些客户甚至报告说、使用数据压缩后性能会提高、因为数据会在缓存中保持压缩状态、从而有效地增加了控制器中的可用缓存量。

ONTAP以4 KB为单位管理物理块。自适应数据压缩使用默认的压缩块大小8 KB、这意味着数据以8 KB单位进行压缩。这与关系数据库最常使用的8 KB块大小匹配。随着将更多数据作为一个单元进行压缩、数据压缩算法的效率也会提高。32 KB压缩块大小比8 KB压缩块单元更节省空间。这确实意味着、使用默认8 KB块大小的自适应数据压缩确实会使效率略低、但使用更小的数据压缩块大小也会有显著优势。数据库工作负载包含大量覆盖活动。要覆盖经过压缩的32 KB数据块中的8 KB、需要回读整个32 KB逻辑数据、对其进行解压缩、更新所需的8 KB区域、重新压缩、然后将整个32 KB写入驱动器。这对存储系统来说是一项非常昂贵的操作、因此、某些基于较大压缩块大小的竞争存储阵列也会对数据库工作负载的性能造成严重影响。



自适应数据压缩使用的块大小最多可以增加到32 KB。这可能会提高存储效率、对于事务日志和备份文件等不活动的文件、如果阵列上存储了大量此类数据、则应考虑使用此方法。在某些情况下、使用16 KB或32 KB块大小的活动数据库也可以通过增加要匹配的自适应数据压缩的块大小来受益。请咨询NetApp或合作伙伴代表、了解这是否适合您的工作负载。



在流式备份目标上、不应同时使用大于8 KB的数据压缩块大小和重复数据删除。原因是、对备份的数据所做的微小更改会影响32 KB数据压缩窗口。如果窗口发生变化、则生成的压缩数据会在整个文件中有所不同。重复数据删除在数据压缩后进行、这意味着重复数据删除引擎对每个压缩备份的看法不同。如果需要对流式备份进行重复数据删除、则只应使用8 KB块自适应数据压缩。最好使用自适应数据压缩、因为它的块大小较小、不会影响重复数据删除的效率。出于类似的原因、主机端压缩也会影响重复数据删除效率。

数据压缩对齐

数据库环境中的自适应数据压缩需要在一定程度上考虑数据压缩块对齐问题。对于随机覆盖非常特定的块的数据来说、这样做只是一个问题。这种方法在概念上类似于整体文件系统对齐、即文件系统的起点必须与4 k设备边

界对齐、文件系统的块大小必须是4 k的倍数。

例如、只有当8 KB写入文件与文件系统本身内的8 KB边界对齐时、才会对其进行压缩。这一点意味着它必须位于文件的前8 KB、文件的后8 KB、依此类推。要确保正确对齐、最简单的方法是使用正确的LUN类型、创建的任何分区都应与设备起始位置偏移8K的倍数、并使用数据库块大小的倍数作为文件系统块大小。

备份或事务日志等数据是跨多个块按顺序写入的操作、所有这些块都会进行压缩。因此、无需考虑对齐。唯一关注的I/O模式是随机覆盖文件。

数据缩减

数据缩减是一项可提高数据压缩效率的技术。如前文所述、自适应数据压缩本身最多可节省2: 1的空间、因为它仅限于在4 KB WAFL块中存储8 KB I/O。块大小越大、压缩方法的效率越高。但是、它们不适用于受到小块覆盖的数据。解压缩32 KB数据单元、更新8 KB部分、重新压缩以及回写驱动器会产生开销。

数据缩减的工作原理是、允许将多个逻辑块存储在物理块中。例如、具有高度可压缩数据(例如文本或部分全满块)的数据库可以从8 KB压缩到1 KB。如果不进行数据缩减、这1 KB的数据仍会占用整个4 KB块。实时数据缩减允许将1 KB的压缩数据与其他压缩数据一起存储在仅1 KB的物理空间中。它不是一种压缩技术;它只是一种在驱动器上分配空间的更高效的方式、因此不会产生任何可检测的性能影响。

节省的资金数额各不相同。已压缩或加密的数据通常无法进一步压缩、因此、数据集无法从数据缩减中受益。相比之下、新初始化的数据文件包含的块元数据和零数据略多、数据压缩率高达80: 1。

对温度敏感的存储效率

温度敏感型存储效率(TSSE)在ONTAP 9.8及更高版本中提供、它依靠块访问热图来识别不常访问的块并以更高的效率对其进行压缩。

重复数据删除

重复数据删除是指从数据集中删除重复的块大小。例如、如果10个不同文件中存在相同的4 KB块、则重复数据删除会将所有10个文件中的4 KB块重定向到相同的4 KB物理块。结果是、这些数据的效率将提高10: 1。

VMware子系统启动LUN等数据的重复数据删除效果通常非常好、因为它们包含同一操作系统文件的多个副本。我们观察到的效率为100: 1甚至更高。

某些数据不包含重复数据。例如、Oracle块包含数据库全局唯一的标头和几乎唯一的尾部。因此、对Oracle数据库进行重复数据删除很少能节省超过1%的空间。对MS SQL数据库执行重复数据删除略有改进、但块级别的唯一元数据仍是一个限制。

在少数情况下、使用16 KB和大型块的数据库可节省多达15%的空间。每个块的初始4 KB包含全局唯一标头、而最后4 KB块包含接近唯一的尾部。内部块是重复数据删除的候选数据、但实际上、这几乎完全是由于对置零数据进行重复数据删除。

许多争用资源的阵列都声称可以根据数据库被复制多次的假设对数据库进行重复数据删除。在这方面、也可以使用NetApp重复数据删除、但ONTAP提供了一个更好的选择: NetApp FlexClone技术。最终结果是相同的;系统会为一个数据库创建多个副本、这些副本共享大多数底层物理块。与花时间复制数据库文件并对其进行重复数据删除相比、使用FlexClone的效率要高得多。实际上、它是无重复数据删除、而不是重复数据删除、因为从一开始就不会创建重复数据。

效率和精简配置

效率功能是精简配置的一种形式。例如、占用100 GB卷的100 GB LUN可能会压缩到50 GB。由于卷仍为100 GB、因此尚未实现实际节省。必须先减小卷大小、以便节省的空间可用于系统上的其他位置。如果稍后更改100 GB LUN会导致数据的可压缩性降低、则LUN大小会增大、卷可能会填满。

强烈建议使用精简配置、因为它可以简化管理、同时显著提高可用容量并节省相关成本。原因很简单—数据库环境通常包含大量空空间、大量卷和LUN以及可压缩数据。厚配置会为卷和LUN预留存储空间、以防它们最终达到100%全满并包含100%不可压缩数据。这种情况不大可能发生。通过精简配置、可以回收这些空间并将其用于其他位置、并可以基于存储系统本身进行容量管理、而不是基于许多较小的卷和LUN。

有些客户更喜欢对特定工作负载使用厚配置、或者通常根据既定的运营和采购实践使用厚配置。

*注意：*如果卷配置厚配置、则必须小心操作、以便完全禁用该卷的所有效率功能、包括使用解压缩和删除重复数据删除 `sis undo` 命令：此卷不应显示在中 `volume efficiency show` 输出。如果配置了效率功能、则仍会为卷部分配置效率功能。因此、覆盖保证的工作方式有所不同、这会增加配置忽略发生原因卷以意外用尽空间的可能性、从而导致数据库I/O错误。

效率最佳实践

NetApp建议执行以下操作：

AFF默认值

在纯闪存AFF系统上运行的ONTAP上创建的卷经过精简配置、并启用了所有实时效率功能。尽管数据库通常不会从重复数据删除中受益、并且可能包含不可压缩的数据、但默认设置适用于几乎所有工作负载。ONTAP旨在高效处理所有类型的数据和I/O模式、无论它们是否可节省空间。只有在完全了解原因且有优势的情况下、才应更改默认值。

一般建议

- 如果卷和(或) LUN未进行精简配置、则必须禁用所有效率设置、因为使用这些功能不会节省空间、而将厚配置与已启用空间效率相结合会发生原因发生意外行为、包括空间不足错误。
- 如果数据不会被覆盖(例如使用备份或数据库事务日志)、则可以通过在较低的冷却期启用TSSE来提高效率。
- 某些文件可能包含大量不可压缩数据、例如、在应用程序级别已启用数据压缩时、文件已加密。如果出现上述任一情况、请考虑禁用数据压缩、以便在包含可压缩数据的其他卷上执行更高效的操作。
- 不要在数据库备份中同时使用32 KB数据压缩和重复数据删除。请参见一节 [\[自适应数据压缩\]](#) 了解详细信息。

使用Oracle数据库进行精简配置

对Oracle数据库进行精简配置需要仔细规划、因为这样会导致在存储系统上配置的空间超过实际可用的空间。这是非常值得的努力、因为如果操作正确、可以显著节省成本并提高易管理性。

精简配置有多种形式、是ONTAP为企业级应用程序环境提供的许多功能不可或缺的组成部分。精简配置也与效率技术密切相关、原因相同：效率功能允许存储的逻辑数据比存储系统上的技术数据多。

几乎任何快照使用都涉及精简配置。例如、NetApp存储上的典型10 TB数据库包含大约30天的快照。这种安排

会使活动文件系统中显示大约10 TB的数据、并将300 TB专用于快照。总存储量为312 TB、通常占用大约12 TB到15 TB的空间。活动数据库会占用10 TB的空间、其余300 TB的数据仅需要2 TB到5 TB的空间、因为系统仅会存储对原始数据所做的更改。

克隆也是精简配置的一个示例。一家主要NetApp客户为一个80 TB数据库创建了40个克隆、以供开发使用。如果使用这些克隆的所有40位开发人员都覆盖了每个数据文件中的每个块、则需要3.2 PB以上的存储。实际上、周转率较低、并且总空间需求接近40 TB、因为驱动器上仅存储更改。

空间管理

对应用程序环境进行精简配置时必须格外小心、因为数据变更率可能会意外增加。例如、如果为数据库表重新编制索引或对VMware子系统应用大规模修补、则快照占用的空间会快速增长。放错位置的备份可能会在很短的时间内写入大量数据。最后、如果文件系统意外用尽可用空间、则很难恢复某些应用程序。

幸运的是、这些风险可以通过仔细配置来解决 `volume-autogrow` 和 `snapshot-autodelete` 策略。正如其名称所暗示的那样、这些选项使用户能够创建策略、以自动清除快照占用的空间或增加卷以容纳更多数据。有多种选项可供选择、不同客户的需求也会有所不同。

请参见 ["逻辑存储管理文档"](#) 有关这些功能的完整讨论。

预留百分比

预留百分比是指卷中LUN在空间效率方面的行为。选项 `fractional-reserve` 设置为100%时、卷中的所有数据在使用任何数据模式时均可实现100%的周转率、而不会耗尽卷上的空间。

例如、假设数据库位于1 TB卷中的一个250 GB LUN上。创建快照会立即在卷中预留额外的250 GB空间、以保证卷不会因任何原因用尽空间。使用预留百分比通常会造成浪费、因为数据库卷中的每个字节极不可能需要覆盖。没有理由为从未发生的事件预留空间。但是、如果客户无法监控存储系统中的空间消耗、并且必须确保空间永远不会用尽、则需要100%预留百分比才能使用快照。

数据压缩和重复数据删除

数据压缩和重复数据删除都是精简配置的两形式。例如、50 TB的数据占用空间可能会压缩为30 TB、从而节省20 TB的空间。要使数据压缩产生任何优势、必须将这20 TB中的一部分用于其他数据、或者购买的存储系统必须小于50 TB。这样、存储的数据就会超过存储系统上的技术可用数据。从数据角度来看、数据容量为50 TB、尽管它在驱动器上仅占用30 TB。

数据集的可压缩性总是有可能发生变化、从而导致实际空间消耗增加。这种消耗量的增加意味着、在监控和使用方面、必须像其他形式的精简配置一样管理数据压缩 `volume-autogrow` 和 `snapshot-autodelete`。

有关数据压缩和重复数据删除的详细信息、请参见链接[efficiency.html](#)

数据压缩和预留百分比

数据压缩是一种精简配置形式。预留百分比会影响数据压缩的使用、但需要注意的一点是、空间是在创建快照之前预留的。通常、只有当存在快照时、预留百分比才重要。如果没有快照、则预留百分比并不重要。而数据压缩则不是这种情况。如果在已进行数据压缩的卷上创建了LUN、则ONTAP会保留空间以容纳快照。此行为在配置期间可能会令人困惑、但这是预期行为。

例如、假设一个10 GB的卷具有一个5 GB的LUN、该LUN已压缩为2.5 GB、并且没有快照。请考虑以下两种情形：

- 预留百分比= 100会导致利用率达到7.5 GB
- 预留百分比= 0会导致利用率达到2.5 GB

第一种情形包括：当前数据占用2.5 GB空间、而源在预计快照使用时的周转率为100%时占用5 GB空间。第二种情形不会预留任何额外空间。

虽然这种情况可能看起来令人困惑、但在实践中不太可能遇到。数据压缩意味着精简配置、而在LUN环境中进行精简配置需要预留百分比。压缩的数据始终可以被不可压缩的内容覆盖、这意味着必须对卷进行精简配置、才能进行压缩、从而节省空间。



- NetApp建议*采用以下预留配置：
- 设置 `fractional-reserve` 如果已实施基本容量监控、则为0 `volume-autogrow` 和 `snapshot-autodelete`。
- 设置 `fractional-reserve` 如果没有监控能力或在任何情况下都无法排空空间、则为100。

可用空间和LVM空间分配

随着数据被删除、文件系统环境中活动LUN的精简配置效率可能会逐渐降低。除非删除的数据被零覆盖(另请参见 "ASMRU" 或者通过TRIM/UNMAP空间回收释放空间、“已擦除”的数据会在文件系统中占用越来越多的未分配空格。此外、活动LUN的精简配置在许多数据库环境中的用途有限、因为数据文件在创建时会初始化为其完整大小。

仔细规划LVM配置可以提高效率、并最大限度地减少存储配置和LUN大小调整的需求。使用Veritas VLVM或Oracle ASM等LVM时、底层LUN会划分为仅在需要时才使用的块区。例如、如果数据集的大小从2 TB开始、但随着时间的推移可能会增长到10 TB、则可以将此数据集放置在LVM磁盘组中组织的10 TB精简配置LUN上。在创建时、它只会占用2 TB的空间、并且只会为在满足数据增长而分配块区时占用额外空间。只要空间受到监控、此过程就会很安全。

Oracle数据库和ONTAP控制器故障转移/切换

要确保Oracle数据库操作不会被这些操作中断、需要了解存储接管和切换功能。此外、如果使用不当、接管和切换操作使用的参数可能会影响数据完整性。

- 在正常情况下、传入给定控制器的写入会同步镜像到其配对控制器。在NetApp MetroCluster环境中、写入操作也会镜像到远程控制器。写入操作在所有位置的非易失性介质中存储之前、不会向主机应用程序确认。
- 存储写入数据的介质称为非易失性内存或NVMEM。它有时也称为非易失性随机存取存储器(NVRAM)、尽管它充当日志、但也可视为写入缓存。在正常操作下、不会读取NVMEM中的数据；该数据仅用于在发生软件或硬件故障时保护数据。将数据写入驱动器后、数据将从系统中的RAM传输、而不是从NVMEM传输。
- 在接管操作期间、高可用性(HA)对中的一个节点会从其配对节点接管操作。切换本质上是相同的、但在IT适用场景 MetroCluster配置中、远程节点接管本地节点的功能。

在日常维护操作期间、存储接管或切换操作应该是透明的、而不是在网络路径发生更改时可能会短暂暂停操作。但是、网络连接可能很复杂、容易出错、因此NetApp强烈建议在将存储系统投入生产之前对接管和切换操作进行全面测试。这样做是确保所有网络路径配置正确的唯一方法。在SAN环境中、请仔细检查命令的输出 `sanlun lun show -p` 确保所有预期的主路径和辅助路径均可用。

发出强制接管或切换命令时必须小心。使用这些选项强制更改存储配置意味着、拥有驱动器的控制器的状态将被

忽略、而备用节点将强制接管驱动器。不正确地强制执行接管可能会导致数据丢失或损坏。这是因为强制接管或切换可能会丢弃NVMEM的内容。接管或切换完成后、如果丢失这些数据、则从数据库的角度来看、存储在驱动器上的数据可能会还原到稍旧的状态。

很少需要使用普通HA对强制接管。在几乎所有故障情形下、节点都会关闭并通知配对节点、以便进行自动故障转移。在某些边缘情况下、例如、发生滚动故障时、节点之间的互连断开、然后一个控制器断开、此时需要强制接管。在这种情况下、节点之间的镜像会在控制器发生故障之前丢失、这意味着无故障控制器将不再具有正在进行的写入的副本。然后、需要强制执行接管、这意味着数据可能会丢失。

相同的逻辑适用场景会执行MetroCluster切换。在正常情况下、切换几乎是透明的。但是、灾难可能会导致运行正常的站点与灾难站点之间的连接断开。从运行正常的站点的角度来看、问题可能只是站点之间的连接中断、而原始站点可能仍在处理数据。如果节点无法验证主控制器的状态、则只能执行强制切换。



- NetApp建议*采取以下预防措施：
- 请格外小心、以免意外强制执行接管或切换。通常、不需要强制执行、强制执行更改可能会导致发生原因数据丢失。
- 如果需要强制接管或切换、请确保关闭应用程序、卸载所有文件系统并更改逻辑卷管理器(LVM)卷组。必须卸载ASM磁盘组。
- 如果发生强制MetroCluster切换、请将故障节点与所有运行正常的存储资源隔离。有关详细信息、请参见相关ONTAP版本的《MetroCluster管理和灾难恢复指南》。

MetroCluster和多个聚合

MetroCluster是一种同步复制技术、如果连接中断、则会切换到异步模式。这是客户最常见的请求、因为有保障的同步复制意味着站点连接中断会导致数据库I/O完全停止、从而使数据库无法使用。

借助MetroCluster、聚合可以在连接恢复后快速重新同步。与其他存储技术不同、MetroCluster不应要求在站点发生故障后进行完整的重新镜像。只需发送增量变更。

在跨聚合的数据集中、在滚动灾难场景中需要执行额外的数据恢复步骤的风险很小。具体而言、如果：(a)站点之间的连接中断、(b)连接恢复、(c)聚合达到一种状态、其中一些聚合已同步、而另一些聚合则未同步、然后(d)主站点丢失、结果是聚合未彼此同步的运行正常的站点。如果发生这种情况、数据集的部分内容会彼此同步、如果不进行恢复、则无法启动应用程序、数据库或数据存储器。如果数据集跨越多个聚合、NetApp强烈建议使用多种可用工具之一利用基于快照的备份来验证在这种异常情况下的快速可恢复性。

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。