



ASA R2系统上的存储配置

Enterprise applications

NetApp
February 11, 2026

目录

ASA R2系统上的存储配置	1
FC SAN	1
LUN对齐	1
LUN大小调整和LUN计数	2
LUN放置	2
LUN大小调整和LVM大小调整	4
LVM条带化	5
NVFAIL	6
ASM 回收实用程序 (ASRU)	7

ASA R2系统上的存储配置

FC SAN

LUN对齐

LUN对齐是指针对底层文件系统布局优化I/O。

ASA r2 系统使用与AFF/ FAS相同的ONTAP架构，但配置模型更简化。ASA r2 系统使用存储可用区 (SAZ) 而不是聚合，但对齐原则保持不变，因为ONTAP跨平台一致地管理块布局。但是，请注意以下ASA特有的几点：

- ASA r2 系统为所有 LUN 提供主动-主动对称路径，从而消除了对齐期间的路径不对称问题。
- 存储单元 (LUN) 默认采用精简配置；对齐方式不会改变此行为。
- 在创建 LUN 期间可以配置快照保留和自动快照删除（ONTAP 9.18.1 及更高版本）。

在ONTAP系统上、存储以4 KB为单位进行组织。一个数据库或文件系统的8 KB块应正好映射到两个4 KB块。如果LUN配置错误使对齐在任一方向上移动1 KB，则每个8 KB块将位于三个不同的4 KB存储块上、而不是两个。这种安排会增加发生原因延迟、并在存储系统中执行发生原因额外的I/O。

对齐也会影响LVM架构。如果在整个驱动器设备上定义了逻辑卷组中的物理卷(不创建分区)、则LUN上的第一个4 KB块与存储系统上的第一个4 KB块对齐。这是正确的对齐方式。分区会出现问题、因为它们会移动操作系统使用LUN的起始位置。只要偏移量以4 KB的整数单位移动、LUN就会对齐。

在Linux环境中、在整个驱动器设备上构建逻辑卷组。如果需要分区、请运行并验证每个分区的起始位置是否为八的倍数、以检查对齐 `fdisk -u` 情况。这意味着分区从八个512字节扇区的倍数开始、即4 KB。

另请参见一节中有关压缩块对齐的讨论"[效率](#)"。与8 KB压缩块边界对齐的任何布局也与4 KB边界对齐。

未对齐警告

数据库重做/事务日志记录通常会生成未对齐的I/O、此I/O可能会导致发生原因发出有关ONTAP上LUN错位的警告、从而使人产生误解。

日志记录会使用不同大小的写入顺序写入日志文件。不与4 KB边界对齐的日志写入操作通常不会出现发生原因性能问题、因为下一个日志写入操作会完成块。因此、ONTAP几乎能够将所有写入作为完整的4 KB块进行处理、即使某些4 KB块中的数据是在两个单独的操作中写入的。

使用诸如此类的工具来验证对齐情况。`sio` 或者 `dd` 能够以规定的块大小生成 I/O。可以使用以下命令查看存储系统的 I/O 对齐统计信息：`stats` 命令。看 "[WAFL对齐验证](#)" 了解更多信息。

Solaris环境中的对齐更为复杂。请参见 "[ONTAP SAN 主机配置](#)" 有关详细信息 ...



在Solaris x86环境中，请格外注意正确对齐，因为大多数配置都有多个分区层。Solaris x86分区片通常位于标准主引导记录分区表之上。

其他最佳实践：

- 使用NetApp互操作性矩阵工具 (IMT) 验证 HBA 固件和操作系统设置。

- 使用 `sanlun` 工具确认路径健康状况和对齐情况。
- 对于 Oracle ASM 和 LVM，请确保配置文件（`/etc/lvm/lvm.conf`、`/etc/sysconfig/oracleasm`）设置正确，以避免出现对齐问题。

LUN大小调整和LUN计数

要获得Oracle数据库的最佳性能和易管理性、选择最佳LUN大小和要使用的LUN数量至关重要。

LUN 是ONTAP上的虚拟化对象，它存在于ASA r2 系统上托管存储可用区 (SAZ) 的所有驱动器中。因此，LUN 的性能不受其大小的影响，因为无论选择什么大小，LUN 都能充分利用 SAZ 的全部性能潜力。

为了方便起见、客户可能希望使用特定大小的LUN。例如、如果数据库是基于LVM或Oracle ASM磁盘组构建的、其中每个磁盘组包含两个1 TB的LUN、则该磁盘组必须以1 TB为增量进行增长。最好使用八个500 GB的LUN来构建磁盘组、以便可以以较小的增量来增加磁盘组。

建议不要建立通用标准LUN大小、因为这样做会使易管理性复杂化。例如、如果数据库或数据存储库的大小介于1 TB到2 TB之间、则100 GB的标准LUN大小可能效果良好、但20 TB的数据库或数据存储库需要200个LUN。这意味着、服务器重新启动时间会更长、需要在各种用户界面中管理更多对象、SnapCenter等产品必须对许多对象执行发现。使用更少、更大的LUN可避免此类问题。

- ASA r2 注意事项：*
- ASA r2 的最大 LUN 大小为 128TB，这允许使用更少但更大的 LUN，而不会影响性能。
- ASA r2 使用存储可用区 (SAZ) 而不是聚合，但这并不会改变 Oracle 工作负载的 LUN 大小调整逻辑。
- 默认情况下启用精简配置；调整 LUN 大小不会造成中断，也不需要将其脱机。

LUN计数

与LUN大小不同、LUN计数会影响性能。应用程序性能通常取决于通过SCSI层执行并行I/O的能力。因此、两个LUN的性能优于一个LUN。使用Veritas VLVM、Linux LVM2或Oracle ASM等LVM是提高并行性的最简单方法。

对于ASA r2，LUN 计数的原则与AFF/ FAS相同，因为ONTAP在跨平台处理并行 I/O 的方式类似。然而， ASA r2 的 SAN 专用架构和主动-主动对称路径确保了所有 LUN 的一致性能。

虽然对随机I/O非常繁重的100% SSD环境进行的测试表明、LUN数量最多可增加到64个、但一般来说、NetApp 客户从LUN数量增加到16个以上所获得的优势微乎其微。

- NetApp建议*：



一般来说，4 到 16 个 LUN 足以满足任何给定 Oracle 数据库工作负载的 I/O 需求。由于主机 SCSI 实现方面的限制，少于四个 LUN 可能会造成性能限制。除了极端情况（例如非常高的随机 I/O SSD 工作负载）之外，将 LUN 数量增加到 16 个以上很少能提高性能。

LUN放置

ASA r2 系统中数据库 LUN 的最佳放置位置主要取决于ONTAP 的各种功能将如何使用。

在ASA r2 系统中，存储单元（LUN 或 NVMe 命名空间）由称为存储可用性区域 (SAZ) 的简化存储层创建

, SAZ 充当 HA 对的公共存储池。



通常每个 HA 对只有一个存储可用区 (SAZ)。

存储可用区 (SAZ)

在ASA r2 系统中, 卷仍然存在, 但它们会在创建存储单元时自动创建。存储单元 (LUN 或 NVMe 命名空间) 直接在存储可用区 (SAZ) 中自动创建的卷内进行配置。这种设计消除了手动卷管理的需要, 使 Oracle 数据库等块工作负载的配置更加直接和精简。

安全区域区和存储单元

相关存储单元 (LUN 或 NVMe 命名空间) 通常位于同一个存储可用区 (SAZ) 内。例如, 一个需要 10 个存储单元 (LUN) 的数据库, 通常会将所有 10 个单元放置在同一个 SAZ 中, 以简化操作并提高性能。



- 使用 1:1 的存储单元与卷的比例, 即每个卷对应一个存储单元 (LUN), 是ASA r2 的默认行为。
- 如果ASA r2 系统中存在多个 HA 对, 则给定数据库的存储单元 (LUN) 可以分布在多个 SAZ 中, 以优化控制器利用率和性能。



在 FC SAN 的上下文中, 存储单元指的是 LUN。

一致性组 (CG)、LUN 和快照

在ASA r2 中, 快照策略和计划是在一致性组级别应用的, 一致性组是一个逻辑结构, 它将多个 LUN 或 NVMe 命名空间分组, 以实现协调的数据保护。由 10 个 LUN 组成的数据集只需要一个快照策略, 前提是这些 LUN 属于同一个一致性组。

一致性组确保所有包含的 LUN 上的原子快照操作。例如, 如果将底层 LUN 分组到同一个一致性组中, 则可以将驻留在 10 个 LUN 上的数据库或由 10 个不同操作系统组成的基于 VMware 的应用程序环境作为单个一致的对象进行保护。如果快照被放置在不同的一个一致性组中, 即使在同一时间安排, 快照也可能无法完全同步。

在某些情况下, 由于恢复要求, 可能需要将一组相关的 LUN 分成两个不同的一个一致性组。例如, 一个数据库可能有四个 LUN 用于数据文件, 两个 LUN 用于日志。在这种情况下, 包含 4 个 LUN 的数据文件一致性组和包含 2 个 LUN 的日志一致性组可能是最佳选择。原因在于独立可恢复性: 数据文件一致性组可以有选择地恢复到较早的状态, 这意味着所有四个 LUN 都将恢复到快照的状态, 而包含关键数据的日志一致性组将不受影响。

CG、LUN 和SnapMirror

SnapMirror策略和操作与快照操作一样, 是在一致性组上执行的, 而不是在 LUN 上执行的。

将相关的 LUN 放在同一个一致性组中, 可以创建单个SnapMirror关系, 并通过一次更新更新所有包含的数据。与快照一样, 此次更新也将是一个原子操作。SnapMirror目标位置将保证拥有源 LUN 的单一时间点副本。如果 LUN 分布在多个一致性组中, 则副本之间可能一致, 也可能不一致。

在ASA r2 系统上使用SnapMirror进行复制存在以下限制：

- 不支持SnapMirror同步复制。
- SnapMirror主动同步仅支持在两个ASA r2 系统之间进行。
- SnapMirror异步复制仅支持在两个ASA r2 系统之间进行。
- SnapMirror异步复制不支持在ASA r2 系统与ASA、 AFF或FAS系统或云之间进行。

了解更多 "[ASA r2 系统支持SnapMirror复制策略](#)"。

CG、LUN 和 QoS

虽然 QoS 可以有选择地应用于单个 LUN，但通常在一致性组级别设置 QoS 更容易。例如，可以将给定 ESX 服务器中所有客户机使用的所有 LUN 放在一个一致性组中，然后应用ONTAP自适应 QoS 策略。最终结果是，每 TiB 的 IOPS 具有自扩展性，并且适用于所有 LUN。

同样地，如果一个数据库需要 100K IOPS 并占用 10 个 LUN，那么在单个一致性组上设置一个 100K IOPS 限制比在每个 LUN 上设置 10 个单独的 10K IOPS 限制要容易得多。

多种CG布局

在某些情况下，将 LUN 分布到多个一致性组中可能是有益的。主要原因是控制器条带化。例如，HA ASA r2 存储系统可能托管单个 Oracle 数据库，此时需要每个控制器的全部处理和缓存能力。在这种情况下，典型的设计是将一半的 LUN 放在控制器 1 上的一个一致性组中，将另一半 LUN 放在控制器 2 上的一个一致性组中。

同样地，对于托管多个数据库的环境，将 LUN 分布在多个一致性组中可以确保控制器利用率的均衡。例如，一个 HA 系统托管 100 个数据库，每个数据库有 10 个 LUN，则每个数据库可能将 5 个 LUN 分配给控制器 1 上的一个一致性组，将 5 个 LUN 分配给控制器 2 上的一个一致性组。这样可以保证在配置更多数据库时实现对称加载。

不过，这些例子都不涉及 1:1 LUN 与一致性组的比例。目标仍然是通过将相关的 LUN 在逻辑上分组到一致性组中来优化可管理性。

1:1 LUN 与一致性组比例的一个合理例子是容器化工作负载，其中每个 LUN 实际上可能代表一个单独的工作负载，需要单独的快照和复制策略，因此需要单独管理。在这种情况下，1:1 的比例可能是最佳选择。

LUN大小调整和LVM大小调整

当基于 SAN 的文件系统或 Oracle ASM 磁盘组在ASA r2 上达到容量限制时，有两种方法可以增加可用空间：

- 增加现有 LUN（存储单元）的容量
- 向现有 ASM 磁盘组或 LVM 卷组添加新的 LUN，并扩展其包含的逻辑卷。

虽然ASA r2 支持 LUN 调整大小，但通常最好使用逻辑卷管理器 (LVM)，例如 Oracle ASM。LVM 存在的主要原因之一是避免频繁调整 LUN 大小。使用 LVM，可以将多个 LUN 组合成一个虚拟存储池。从该存储池中划分出的逻辑卷可以轻松调整大小，而不会影响底层存储配置。

使用 LVM 或 ASM 的其他好处包括：

- 性能优化：将 I/O 分布到多个 LUN 上，减少热点。
- 灵活性：添加新的 LUN 而不会中断现有工作负载。
- 透明迁移：ASM 或 LVM 可以将扩展区迁移到新的 LUN 以进行均衡或分层，而无需主机停机。

ASA r2 的关键考虑因素：

-  • LUN 调整大小是在存储虚拟机 (SVM) 内的存储单元级别执行的，使用来自存储可用区 (SAZ) 的容量。
- 对于 Oracle 而言，最佳实践是将 LUN 添加到 ASM 磁盘组，而不是调整现有 LUN 的大小，以保持条带化和并行性。

LVM条带化

LVM条带化是指在多个LUN之间分布数据。结果是、许多数据库的性能显著提高。

在闪存驱动器时代之前、条带化用于帮助克服旋转驱动器的性能限制。例如、如果操作系统需要执行1 MB的读取操作、则从单个驱动器读取1 MB的数据将需要大量的驱动器磁头查找和读取、因为1 MB的传输速度较慢。如果在8个LUN上对1 MB的数据进行条带化、则操作系统可以并行执行8个128 K读取操作、从而减少完成1 MB传输所需的时间。

使用旋转硬盘进行条带化更加困难，因为必须预先知道 I/O 模式。如果条带化没有针对真实的 I/O 模式进行正确调整，则条带化配置可能会损害性能。对于 Oracle 数据库，尤其是全闪存存储配置，条带化配置起来要容易得多，并且已被证明可以显著提高性能。

默认情况下、逻辑卷管理器(例如Oracle ASM)会进行条带化、但本机操作系统LVM则不会进行条带化。其中一些会将多个LUN绑定在一起、形成一个串联设备、从而导致数据文件只存在于一个LUN设备上。这会导致热点。其他LVM实施默认使用分布式块区。这与条带化类似、但更粗。卷组中的LUN会被划分为多个大块、称为块区、通常以MB为单位进行测量、然后逻辑卷会分布在这些块区中。结果是、文件的随机I/O应在各个LUN之间分布良好、但顺序I/O操作的效率不如所能达到的高。

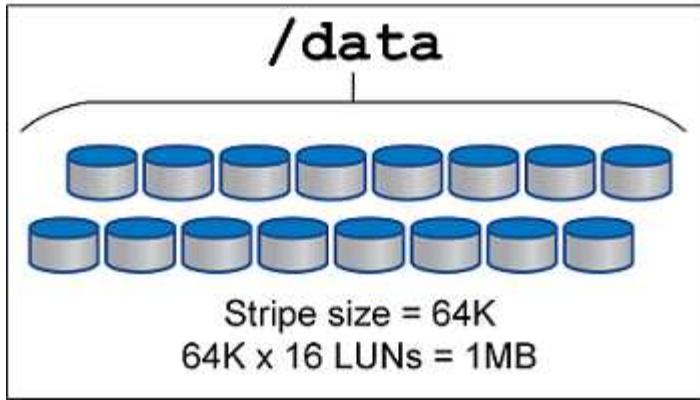
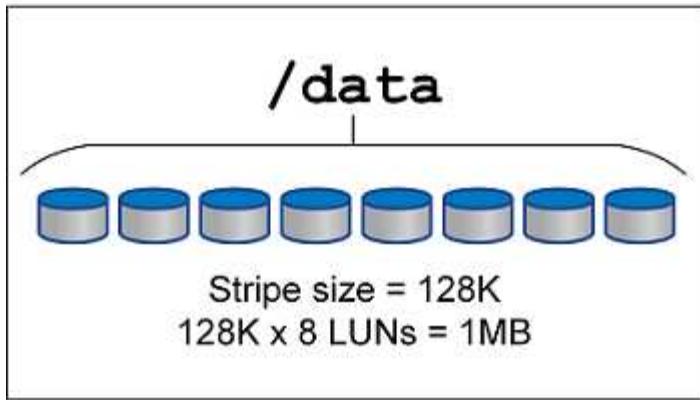
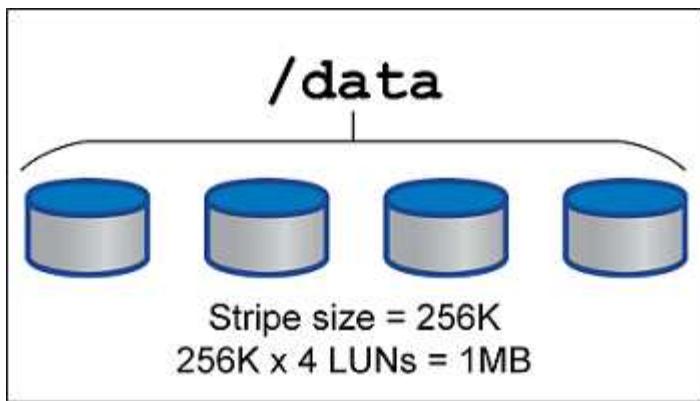
性能密集型应用程序I/O几乎始终为(a)基本块大小单位或(b) 1兆字节。

条带化配置的主要目标是确保单文件I/O可作为一个单元执行、多块I/O (大小应为1 MB)可在条带化卷中的所有LUN之间均匀并行。这意味着条带大小不能小于数据库块大小、条带大小乘以LUN数量应为1 MB。

Oracle 数据库 LVM 条带化的最佳实践：

-  • 条带大小 \geq 数据库块大小。
- 条带大小 * LUN 数量 \approx 1MB，以实现最佳并行性。
- 每个 ASM 磁盘组使用多个 LUN 以最大限度地提高吞吐量并避免热点。

下图显示了三个可能的条带大小和宽度调整选项。选择LUN数量是为了满足上述性能要求、但在所有情况下、单个条带内的总数据均为1 MB。



NVFAIL

NVFAIL 是ONTAP 的一项功能，可在灾难性故障转移情况下确保数据完整性。

即使ASA r2 系统采用简化的 SAN 架构 (SAZ 和存储单元而不是卷) ，此功能仍然适用于ASA r2 系统。

由于数据库维护着大量的内部缓存，因此在存储故障转移事件期间容易发生数据损坏。如果发生灾难性事件需要强制ONTAP故障转移，无论整体配置的健康状况如何，结果都是先前已确认的更改可能会被有效地丢弃。存储阵列的内容向前跳跃，数据库缓存的状态不再反映磁盘上的数据状态。这种不一致会导致数据损坏。

缓存可以在应用层或服务器层进行。例如，Oracle Real Application Cluster (RAC) 配置中，服务器在主站点和远程站点上都处于活动状态，并将数据缓存在 Oracle SGA 中。强制故障转移操作导致数据丢失，会使数据库面临损坏的风险，因为存储在 SGA 中的数据块可能与磁盘上的数据块不匹配。

缓存还有一个不太明显的用途，那就是在操作系统文件系统层。基于位于主站点的 LUN 的集群文件系统可以挂载到远程站点的服务器上，并且可以再次缓存数据。在这些情况下， NVRAM故障或强制接管可能会导致文件系

统损坏。

ONTAP 使用 NVFAIL 及其相关设置来保护数据库和操作系统免受此情况的影响，这些设置会向主机发出信号，使缓存数据失效，并在故障转移后重新挂载受影响的文件系统。该机制适用于 ASA r2 LUN 和命名空间，就像适用于 AFF/ FAS 一样。

ASA r2 的关键考虑因素：

- NVFAIL 在 LUN 级别（存储单元）运行，而不是在 SAZ 级别运行。
- 对于 Oracle 数据库，应在所有托管关键组件（数据文件、重做日志、控制文件）的 LUN 上启用 NVFAIL。
- ASA r2 不支持 MetroCluster，因此 NVFAIL 主要适用于本地 HA 故障转移场景。
- ASA r2 不支持 NFS，因此 NVFAIL 注意事项仅适用于基于 SAN 的工作负载（FC/iSCSI/NVMe）。

ASM 回收实用程序 (ASRU)

当启用在线压缩时，ONTAP on ASA r2 可以有效地删除写入 LUN（存储单元）的零块。Oracle ASM 回收实用程序 (ASRU) 等实用程序的工作原理是将零写入未使用的 ASM 区段。

这样，数据库管理员就可以在数据删除后回收存储阵列上的空间。ONTAP 拦截零值并从 LUN 中释放空间。由于存储系统中没有实际写入任何数据，因此回收过程非常快。

从数据库角度来看，ASM 磁盘组包含零、读取这些 LUN 区域会产生零流、但 ONTAP 不会将零存储在驱动器上。而是进行简单的元数据更改、以便在内部将 LUN 的置零区域标记为任何数据为空。

出于类似的原因、涉及置零数据的性能测试无效、因为零块实际上不会在存储阵列中作为写入进行处理。

使用 ASA r2 ONTAP 需要考虑的关键 ASRU 因素：

- 由于 ASA r2 仅支持块存储，因此其工作方式与 AFF/ FAS 相同，适用于 SAN 工作负载。
- 适用于在 SAZ 内配置的 LUN 和 NVMe 命名空间。
- FlexVol 卷不存在，但零块回收行为是相同的。



使用 ASRU 时、请确保已安装 Oracle 建议的所有修补程序。

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。