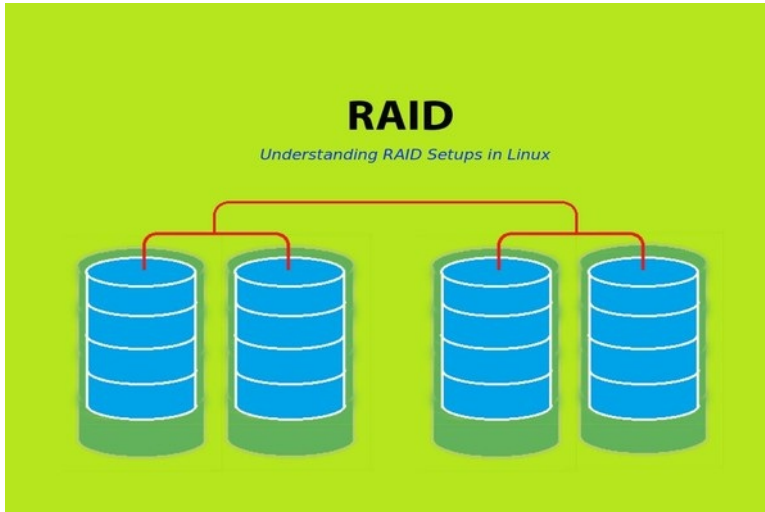


1 在 Linux 下使用 RAID (一) : 介绍 RAID 的级别和概念

RAID 的意思是廉价磁盘冗余阵列 (Redundant Array of Inexpensive Disks)，但现在它被称为独立磁盘冗余阵列 (Redundant Array of Independent Drives)。早先一个容量很小的磁盘都是非常昂贵的，但是现在我们可以很便宜的买到一个更大的磁盘。Raid 是一系列放在一起，成为一个逻辑卷的磁盘集合。



在 Linux 中理解 RAID 设置

RAID 包含一组或者一个集合甚至一个阵列。使用一组磁盘结合驱动器组成 RAID 阵列或 RAID 集。将至少两个磁盘连接到一个 RAID 控制器，而成为一个逻辑卷，也可以将多个驱动器放在一个组中。一组磁盘只能使用一个 RAID 级别。使用 RAID 可以提高服务器的性能。不同 RAID 的级别，性能会有所不同。它通过容错和高可用性来保存我们的数据。

这个系列被命名为“在 Linux 下使用 RAID”，分为9个部分，包括以下主题：

- 第1部分：介绍 RAID 的级别和概念
- 第2部分：在Linux中如何设置 RAID0 (条带化)
- 第3部分：在Linux中如何设置 RAID1 (镜像化)
- 第4部分：在Linux中如何设置 RAID5 (条带化与分布式奇偶校验)
- 第5部分：在Linux中如何设置 RAID6 (条带双分布式奇偶校验)
- 第6部分：在Linux中设置 RAID 10 或 1 + 0 (嵌套)
- 第7部分：增加现有的 RAID 阵列并删除损坏的磁盘
- 第8部分：在 RAID 中恢复 (重建) 损坏的驱动器
- 第9部分：在 Linux 中管理 RAID

这是9篇系列教程的第1部分，在这里我们将介绍 RAID 的概念和 RAID 级别，这是在 Linux 中构建 RAID 需要理解的。

软件 RAID 和硬件 RAID

软件 RAID 的性能较低，因为其使用主机的资源。需要加载 RAID 软件以从软件 RAID 卷中读取数据。在加载 RAID 软件前，操作系统需要引导起来才能加载 RAID 软件。在软件 RAID 中无需物理硬件。零成本投资。

硬件 RAID 的性能较高。他们采用 PCI Express 卡物理地提供有专用的 RAID 控制器。它不会使用主机资源。他们有 NVRAM 用于缓存的读取和写入。缓存用于 RAID 重建时，即使出现电源故障，它会使用后备的电池电源保持缓存。对于大规模使用是非常昂贵的投资。

硬件 RAID 卡如下所示：



硬件 RAID

重要的 RAID 概念

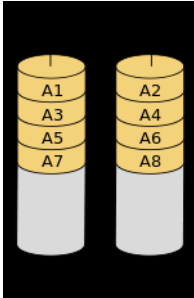
- **校验方式**用在 RAID 重建中从校验所保存的信息中重新生成丢失的内容。RAID 5，RAID 6 基于校验。
- **条带化**是将切片数据随机存储到多个磁盘。它不会在单个磁盘中保存完整的数据。如果我们使用2个磁盘，则每个磁盘存储我们的一半数据。
- **镜像**被用于 RAID 1 和 RAID 10。镜像会自动备份数据。在 RAID 1 中，它会保存相同的内容到其他盘上。
- **热备份**只是我们的服务器上的一个备用驱动器，它可以自动更换发生故障的驱动器。在我们的阵列中，如果任何一个驱动器损坏，热备份驱动器会自动用于重建 RAID。
- **块**是 RAID 控制器每次读写数据时的最小单位，最小 4KB。通过定义块大小，我们可以增加 I/O 性能。

RAID有不同的级别。在这里，我们仅列出在真实环境下的使用最多的 RAID 级别。

- RAID0 = 条带化
- RAID1 = 镜像
- RAID5 = 单磁盘分布式奇偶校验
- RAID6 = 双磁盘分布式奇偶校验
- RAID10 = 镜像 + 条带。（嵌套RAID）

RAID 在大多数 Linux 发行版上使用名为 mdadm 的软件包进行管理。让我们先对每个 RAID 级别认识一下。

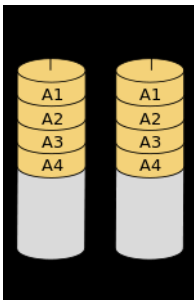
RAID 0 / 条带化



条带化有很好的性能。在 RAID 0（条带化）中数据将使用切片的方式被写入到磁盘。一半的内容放在一个磁盘上，另一半内容将被写入到另一个磁盘。假设我们有2个磁盘驱动器，例如，如果我们将数据“TECMINT”写到逻辑卷中，“T”将被保存在第一盘中，“E”将保存在第二盘，‘C’将被保存在第一盘，“M”将保存在第二盘，它会一直继续此循环过程。（LCTT 译注：实际上不可能按字节切片，是按数据块切片的。）在这种情况下，如果驱动器中的任何一个发生故障，我们就会丢失数据，因为一个盘中只有一半的数据，不能用于重建 RAID。不过，当比较写入速度和性能时，RAID 0 是非常好的。创建 RAID 0（条带化）至少需要2个磁盘。如果你的数据是非常宝贵的，那么不要使用此 RAID 级别。

- 高性能。
- RAID 0 中容量零损失。
- 零容错。
- 写和读有很高的性能。

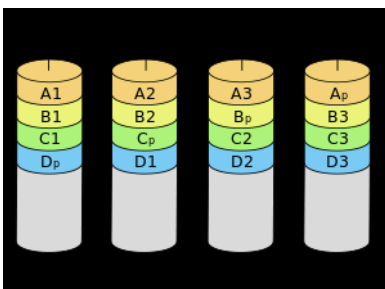
RAID 1 / 镜像化



镜像也有不错的性能。镜像可以对我们的数据做一份相同的副本。假设我们有两个2TB的硬盘驱动器，我们总共有4TB，但在镜像中，但是放在 RAID 控制器后面的驱动器形成了一个逻辑驱动器，我们只能看到这个逻辑驱动器有2TB。当我们保存数据时，它将同时写入这两个2TB驱动器中。创建 RAID 1（镜像化）最少需要两个驱动器。如果发生磁盘故障，我们可以通过更换一个新的磁盘恢复 RAID。如果在 RAID 1 中任何一个磁盘发生故障，我们可以从另一个磁盘中获取相同的数据，因为另外的磁盘中也有相同的数据。所以是零数据丢失。

- 良好的性能。
- 总容量丢失一半可用空间。
- 完全容错。
- 重建会更快。
- 写性能变慢。
- 读性能变好。
- 能用于操作系统和小规模的数据库。

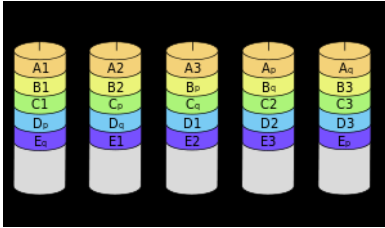
RAID 5 / 分布式奇偶校验



RAID 5 多用于企业级。RAID 5 的以分布式奇偶校验的方式工作。奇偶校验信息将被用于重建数据。它从剩下的正常驱动器上的信息来重建。在驱动器发生故障时，这可以保护我们的数据。假设我们有4个驱动器，如果一个驱动器发生故障而后我们更换发生故障的驱动器后，我们可以从奇偶校验中重建数据到更换的驱动器上。奇偶校验信息存储在所有4个驱动器上，如果我们有4个 1TB 的驱动器。奇偶校验信息将被存储在每个驱动器的256G中，而其它768GB是用户自己使用的。单个驱动器故障后，RAID 5 依旧正常工作，如果驱动器损坏个数超过1个会导致数据的丢失。

- 性能卓越
- 读速度将非常好。
- 写速度处于平均水准，如果我们不使用硬件 RAID 控制器，写速度缓慢。
- 从所有驱动器的奇偶校验信息中重建。
- 完全容错。
- 1个磁盘空间将用于奇偶校验。
- 可以被用在文件服务器，Web服务器，非常重要的备份中。

RAID 6 双分布式奇偶校验磁盘

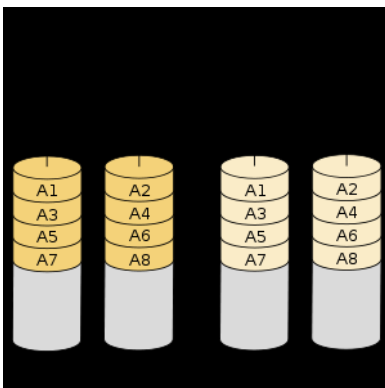
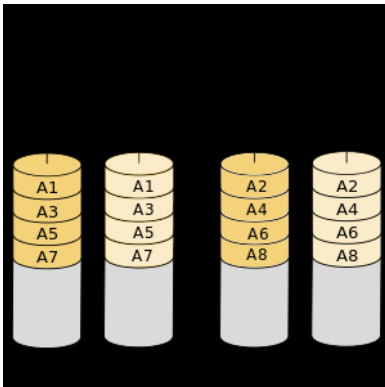


RAID 6 和 RAID 5 相似但它有两个分布式奇偶校验。大多用在大量的阵列中。我们最少需要4个驱动器，即使有2个驱动器发生故障，我们依然可以更换新的驱动器后重建数据。

它比 RAID 5 慢，因为它将数据同时写到4个驱动器上。当我们使用硬件 RAID 控制器时速度就处于平均水准。如果我们6个的1TB驱动器，4个驱动器将用于数据保存，2个驱动器将用于校验。

- 性能不佳。
- 读的性能良好。
- 如果我们不使用硬件 RAID 控制器写的性能会很差。
- 从两个奇偶校验驱动器上重建。
- 完全容错。
- 2个磁盘空间将用于奇偶校验。
- 可用于大型阵列。
- 用于备份和视频流中，用于大规模。

RAID 10 / 镜像+条带



RAID 10 可以被称为1 + 0或0 + 1。它将做镜像+条带两个工作。在 RAID 10 中首先做镜像然后做条带。在 RAID 01 上首先做条带，然后做镜像。RAID 10 比 01 好。假设，我们有4个驱动器。当我逻辑卷上写数据时，它会使用镜像和条带的方式将数据保存到4个驱动器上。

如果我在 RAID 10 上写入数据“TECMINT”，数据将使用如下方式保存。首先将“T”同时写入两个磁盘，“E”也将同时写入另外两个磁盘，所有数据都写入两块磁盘。这样可以每个数据复制到另外的磁盘。

同时它使用 RAID 0 方式写入数据，遵循将“T”写入第一组盘，“E”写入第二组盘。再次将“C”写入第一组盘，“M”到第二组盘。

- 良好的读写性能。
- 总容量丢失一半的可用空间。
- 容错。
- 从副本数据中快速重建。
- 由于其高性能和高可用性，常被用于数据库的存储中。

结论

在这篇文章中，我们已经了解了什么是 RAID 和在实际环境大多采用哪个级别的 RAID。希望你已经学会了上面所写的。对于 RAID 的构建必须了解有关 RAID 的基本知识。以上内容可以基本满足你对 RAID 的了解。