

# RAID（磁盘阵列）

## 【内容导航】

- 第 1 页：RAID 基本介绍
- 第 2 页：RAID 0
- 第 3 页：RAID 1
- 第 4 页：RAID 10 或 RAID 0+1
- 第 5 页：RAID 3
- 第 6 页：RAID 5
- 第 7 页：RAID 6
- 第 8 页：RAID 7
- 第 9 页：RAID 5E 和 RAID 5EE
- 第 10 页：Matrix RAID
- 第 11 页：NV RAID
- 第 12 页：RAID 1E、RAID DP、RAID ADG
- 第 13 页：产品详细信息

RAID 是英文 Redundant Array of Independent Disks 的缩写，翻译成中文意思是“独立磁盘冗余阵列”，实际上也是我们经常所说的“磁盘阵列”。

简单的说，RAID 是一种把多块独立的硬盘（物理硬盘）按不同的方式组合起来形成一个硬盘组（逻辑硬盘），从而提供比单个硬盘更高的存储性能和提供数据备份技术。组成磁盘阵列的不同方式成为 RAID 级别（RAID Levels）。数据备份的功能是在用户数据一旦发生损坏后，利用备份信息可以使损坏数据得以恢复，从而保障了用户数据的安全性。在用户看起来，组成的磁盘组就像是一个硬盘，用户可以对它进行分区，格式化等等。总之，对磁盘阵列的操作与单个硬盘一模一样。不同的是，磁盘阵列的存储速度要比单个硬盘高很多，而且可以提供自动数据备份。

RAID 技术的两大特点：一是速度、二是安全，由于这两项优点，RAID 技术早期被应用于高级服务器中的 SCSI 接口的硬盘系统中，随着近年计算机技术的发展，P C 机的 CPU 的速度已进入 GHz 时代。IDE 接口的硬盘也不甘落后，相继推出了 ATA66 和 ATA100 硬盘。这就使得 RAID 技术被应用于中低档甚至个人 P C 机上成为可能。RAID 通常是由在硬盘阵列塔中的 RAID 控制器或电脑中的 RAID 卡来实现的。

RAID 技术经过不断的发展，现在已拥有了从 RAID 0 到 6 七种基本的 RAID 级别。另外，还有一些基本 RAID 级别的组合形式，如 RAID 10（RAID 0 与 RAID 1 的组合），RAID 50（RAID 0 与 RAID 5 的组合）等。不同 RAID 级别代表着不同的存储性能、数据安全性和存储成本。但我们最为常用的是下面的几种 RAID 形式。

RAID 级别	RAID-0	RAID-1	RAID-3	RAID-5	RAID-10
别名	条带	镜像	专用奇偶位条带	分布奇偶位条带	镜像阵列条带
容错性	没有	有	有	有	有
冗余类型	没有	复制	奇偶校验	奇偶校验	复制
热备盘选项	没有	有	有	有	有
读性能	高	低	高	高	中间
随机写性能	高	低	最低	低	中间
连续写性能	高	低	低	低	中间
需要的磁盘数	一个或多个	只需 2 个或 2xN 个	三个或更多	三个或更多	只需 4 个或 4xN 个
可用容量	总的磁盘的容量	只能用磁盘容量的 50%	$(n-1)/n$ 的磁盘容量。其中 n 为磁盘数	$(n-1)/n$ 的总磁盘容量。其中 n 为磁盘数	磁盘容量的 50%
典型应用	无故障的迅速读写，要求安全性不高，如图形工作站等	随机数据写入，要求安全性高，如服务器、数据库存储领域	连续数据传输，要求安全性高，如视频编辑、大型数据库等	随机数据传输，要求安全性高，如金融、数据库、存储等	要求数据量大，安全性高，如银行、金融等领域

RAID 级别的选择有三个主要因素：可用性（数据冗余）、性能和成本。如果不要求可用性，选择 RAID0 以获得最佳性能。如果可用性和性能是重要的而成本不是一个主要因素，则根据硬盘数量选择 RAID 1。如果可用性、成本和性能都同样重要，则根据一般的数据传输和硬盘的数量选择 RAID 3、RAID 5。

## 内容导航

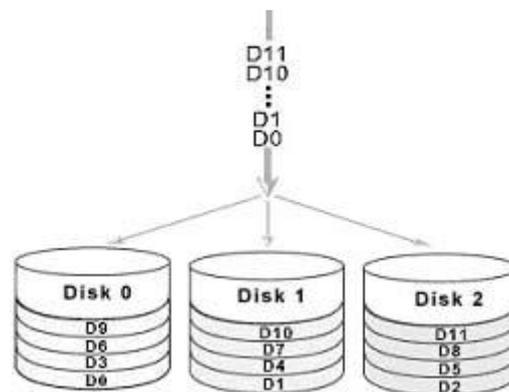
RAID 0 又称为 Stripe（条带化）或 Striping，它代表了所有 RAID 级别中最高存储性能。RAID 0 提高存储性能的原理是把连续的数据分散到多个磁盘上存取，这样，系统有数据请求就可以被多个磁盘并行的执行，每个磁盘执行属于它自己的那部分数据请求。这种数据上的并行操作可以充分利用总线的带宽，显著提高磁盘整体存取性能。

如图所示：系统向三个磁盘组成的逻辑硬盘（RAID 0 磁盘组）发出的 I/O 数据请求被转化为 3 项操作，其中的每一项操作都对应于一块物理硬盘。我们从图中可以清楚的看到通过建立 RAID 0，原先顺序的数据请求被分散到所有的三块

硬盘中同时执行。从理论上讲，三块硬盘的并行操作使同一时间内磁盘读写速度提升了3倍。但由于总线带宽等多种因素的影响，实际的提升速率肯定会低于理论值，但是，大量数据并行传输与串行传输比较，提速效果显著显然毋庸置疑。

RAID 0的缺点是不提供数据冗余，因此一旦用户数据损坏，损坏的数据将无法得到恢复。

RAID 0具有的特点，使其特别适用于对性能要求较高，而对数据安全不太在乎的领域，如图形工作站等。对于个人用户，RAID 0也是提高硬盘存储性能的绝佳选择。



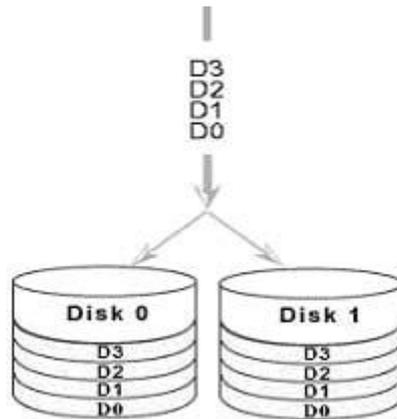
## 内容导航

RAID 1 又称为 Mirror 或 Mirroring（镜像），它的宗旨是最大限度的保证用户数据的可用性和可修复性。RAID 1 的操作方式是把用户写入硬盘的数据百分之百地自动复制到另外一个硬盘上。

如图所示：当读取数据时，系统先从 RAID 0 的源盘读取数据，如果读取数据成功，则系统不去管备份盘上的数据；如果读取源盘数据失败，则系统自动转而读取备份盘上的数据，不会造成用户工作任务的中断。当然，我们应当及时地更换损坏的硬盘并利用备份数据重新建立 Mirror，避免备份盘在发生损坏时，造成不可挽回的数据损失。

由于对存储的数据进行百分之百的备份，在所有 RAID 级别中，RAID 1 提供最高的数据安全保障。同样，由于数据的百分之百备份，备份数据占了总存储空间的一半，因而 Mirror(镜像)的磁盘空间利用率低，存储成本高。

Mirror 虽不能提高存储性能，但由于其具有的高数据安全性，使其尤其适用于存放重要数据，如服务器和数据库存储等领域。



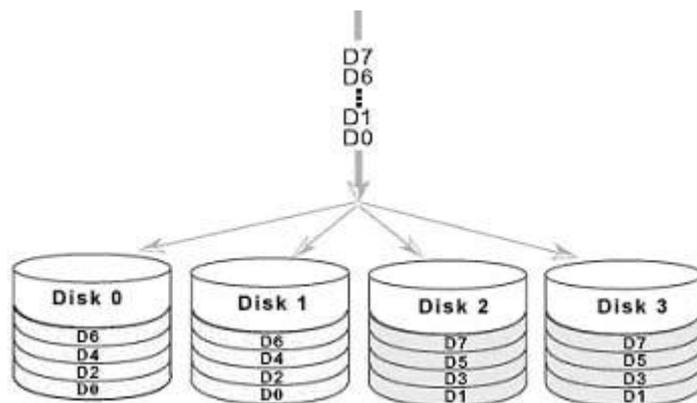
## 内容导航

正如其名字一样 RAID 0+1 是 RAID 0 和 RAID 1 的组合形式，也称为 RAID 10。

以四个磁盘组成的 RAID 0+1 为例，其数据存储方式如图所示：RAID 0+1 是存储性能和数据安全兼顾的方案。它在提供与 RAID 1 一样的数据安全保障的同时，也提供了与 RAID 0 近似的存储性能。

由于 RAID 0+1 也通过数据的 100% 备份功能提供数据安全保障，因此 RAID 0+1 的磁盘空间利用率与 RAID 1 相同，存储成本高。

RAID 0+1 的特点使其特别适用于既有大量数据需要存取，同时又对数据安全性要求严格的领域，如银行、金融、商业超市、仓储库房、各种档案管理等。



## 内容导航

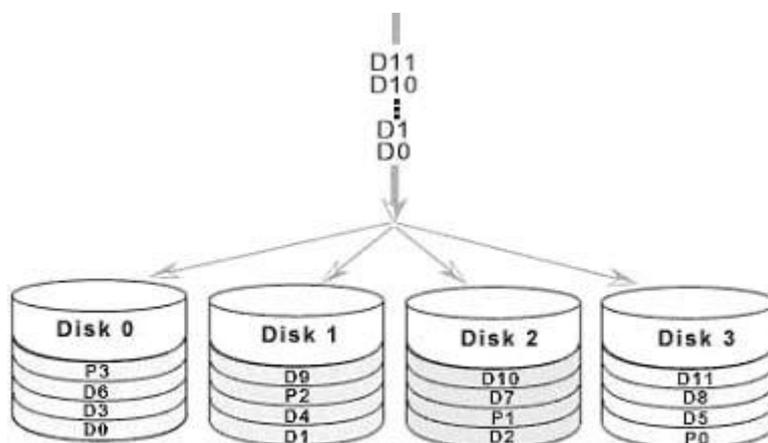
RAID 3 是把数据分成多个“块”，按照一定的容错算法，存放在 N+1 个硬盘上，实际数据占用的有效空间为 N 个硬盘的空间总和，而第 N+1 个硬盘上存储的数据是校验容错信息，当这 N+1 个硬盘中的其中一个硬盘出现故障时，从其它 N

个硬盘中的数据也可以恢复原始数据，这样，仅使用这 N 个硬盘也可以带伤继续工作（如采集和回放素材），当更换一个新硬盘后，系统可以重新恢复完整的校验容错信息。由于在一个硬盘阵列中，多于一个硬盘同时出现故障率的几率很小，所以一般情况下，使用 RAID3，安全性是可以得到保障的。与 RAID0 相比，RAID3 在读写速度方面相对较慢。使用的容错算法和分块大小决定 RAID 使用的应用场合，在通常情况下，RAID3 比较适合大文件类型且安全性要求较高的应用，如视频编辑、硬盘播出机、大型数据库等。

## 内容导航

RAID 5 是一种存储性能、数据安全和存储成本兼顾的存储解决方案。以四个硬盘组成的 RAID 5 为例，其数据存储方式如图 4 所示：图中，P0 为 D0，D1 和 D2 的奇偶校验信息，其它以此类推。由图中可以看出，RAID 5 不对存储的数据进行备份，而是把数据和相对应的奇偶校验信息存储到组成 RAID5 的各个磁盘上，并且奇偶校验信息和相对应的数据分别存储于不同的磁盘上。当 RAID5 的一个磁盘数据发生损坏后，利用剩下的数据和相应的奇偶校验信息去恢复被损坏的数据。

RAID 5 可以理解为是 RAID 0 和 RAID 1 的折衷方案。RAID 5 可以为系统提供数据安全保障，但保障程度要比 Mirror 低而磁盘空间利用率要比 Mirror 高。RAID 5 具有和 RAID 0 相近似的数据读取速度，只是多了一个奇偶校验信息，写入数据的速度比对单个磁盘进行写入操作稍慢。同时由于多个数据对应一个奇偶校验信息，RAID 5 的磁盘空间利用率要比 RAID 1 高，存储成本相对较低。

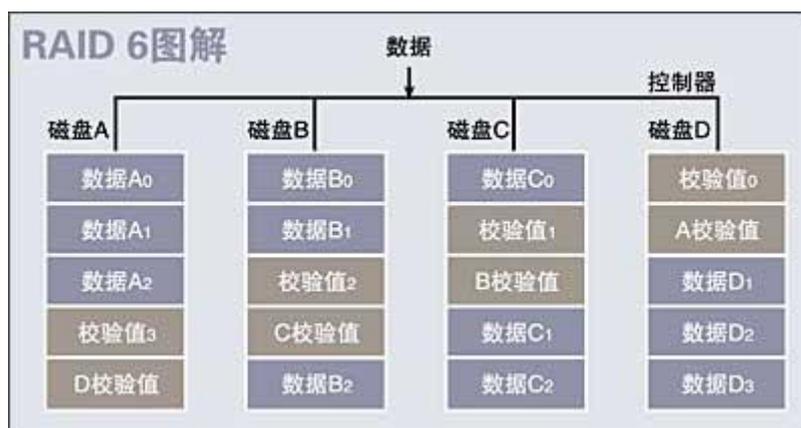


## 内容导航

下面要说的是一些并不常见的 RAID 级别，这些 RAID 级别有些是将来发展的方向，有些是单个公司提出的标准，它们提出的这些 RAID 级别比我们常见的这些 RAID 级别在性能和数据恢复能力上要强大，因此，也被一些公司所采用。

## RAID 6

RAID 6是由一些大型企业提出来的私有 RAID 级别标准，它的全名叫“Independent Data disks with two independent distributed parity schemes（带有两个独立分布式校验方案的独立数据磁盘）”。这种 RAID 级别是在 RAID 5 的基础上发展而成，因此它的工作模式与 RAID 5 有异曲同工之妙，不同的是 RAID 5 将校验码写入到一个驱动器里面，而 RAID 6 将校验码写入到两个驱动器里面，这样就增强了磁盘的容错能力，同时 RAID 6 阵列中允许出现故障的磁盘也就达到了两个，但相应的阵列磁盘数量最少也要 4 个。下图是 RAID 6 的图解。

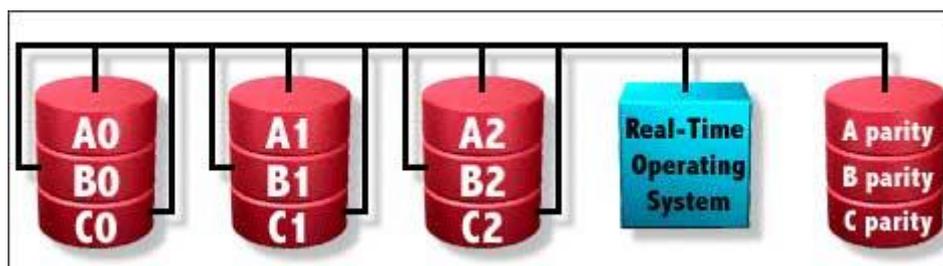


从图中我们可以看到每个磁盘中都具有两个校验值，而 RAID 5 里面只能为每一个磁盘提供一个校验值，由于校验值的使用可以达到恢复数据的目的，因此多增加一位校验位，数据恢复的能力就越强。不过在增加一位校验位后，就需要一个比较复杂的控制器来进行控制，同时也使磁盘的写能力降低，并且还需要占用一定的磁盘空间。因此，这种 RAID 级别应用还比较少，相信随着 RAID 6 技术的不断完善，RAID 6 将得到广泛应用。RAID 6 的磁盘数量为  $N+2$  个。

内容导航

## RAID 7

RAID 7 全名叫“Optimized Asynchrony for High I/O Rates as well as High Data Transfer Rates（最优化的异步高 I/O 速率和高数据传输率）”，它与以前我们见到 RAID 级别具有明显的区别。RAID 7 完全可以理解为一个独立存储计算机，它自身带有操作系统和管理工具，完全可以独立运行。RAID 7 的图解如下：



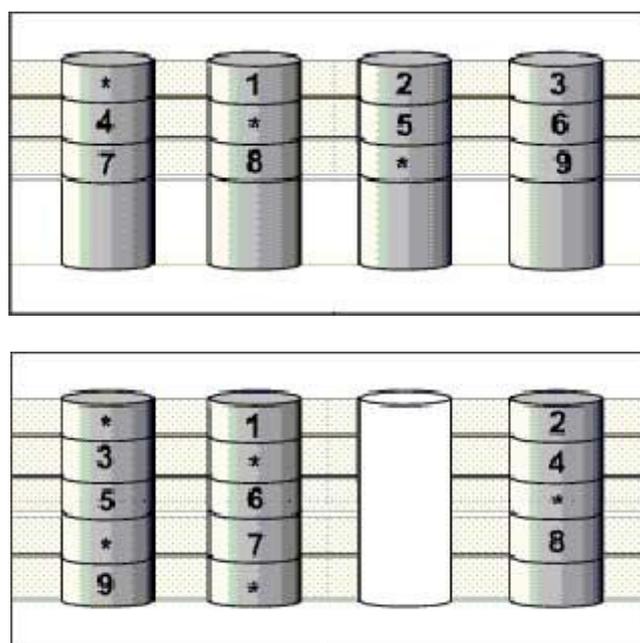
图中每个“柱体”是由多个磁盘构成，而不是我们以前看到的一个磁盘表示一个“柱体”。从上图我们可以看出，每个磁盘都有一个独立的 I/O 通道，它们与主通道相连，操作系统可以直接对每个磁盘的访问进行控制，可以让每个磁盘在不同的时段进行数据读写，这样就大大改善了 I/O 的应用，同时也提高了数据读写的能力，而这种磁盘访问方式也叫做非同步访问。在 RAID 7 中，提供了一个磁盘作为专门的校验盘，它适合于任何一个磁盘进行数据恢复。

总的来说，RAID 7 与我们传统的 RAID 级别有很大区别，它的优点很多，但缺点也非常明显，那就是价格非常高，对于普通企业用户并不实用。

内容导航

## RAID 5E

RAID 5E 是由 IBM 公司提出的一种私有 RAID 级别，没有成为国际标准。这种 RAID 级别也是从 RAID 5 的基础上发展而来的，它与 RAID 5 不同的地方是将数据校验信息平均分布在每一个磁盘中，并且每个磁盘都要预留一定的空间，这部分空间没有进行条带化（条带是指数据为了保存在 RAID 中，被划分成的最小单元。通过对条带进行调整，可以使支持 RAID 的磁盘阵列性能更加优异）。当一个磁盘出现故障时，这个磁盘上的数据将被压缩到其他磁盘预留没有条带化的空间内，达到数据保护的作用，而这时候的 RAID 级别则从 RAID 5E 转换成了 RAID 5，继续保护磁盘数据。RAID 5E 允许两个磁盘出错，最少也需要 4 个磁盘才能实现 RAID 5E。下图是 RAID 5E 的图解：



## RAID 5EE

RAID 5EE 也是由 IBM 公司提出的一种私有 RAID 级别，它也没有成为国际标准。RAID 5EE 的工作原理与 RAID 5E 基本相同，它也是在每个磁盘中预留一部分空间作为分布的热备盘，当一个硬盘出现故障时，这个磁盘上的数据将被压缩到分布的热备盘中，达到数据的保护作用。不过与 RAID 5E 不同的是 RAID 5EE 内增加了一些优化技术，使 RAID 5EE 的工作效率更高，压缩数据的速度也更快。RAID 5EE 允许两个磁盘出错，最少需要 4 个磁盘实现。

内容导航

## Matrix RAID

Matrix RAID 即所谓的“矩阵 RAID”，是 ICH6R 南桥所支持的一种廉价的磁盘冗余技术，是一种经济性高的新颖 RAID 解决方案。Matrix RAID 技术的原理相当简单，只需要两块硬盘就能实现了 RAID 0 和 RAID 1 磁盘阵列，并且不需要添加额外的 RAID 控制器，这正是我们普通用户所期望的。Matrix RAID 需要硬件层和软件层同时支持才能实现，硬件方面目前就是 ICH6R 南桥以及更高阶的 ICH6RW 南桥，而 Intel Application Accelerator 软件和 Windows 操作系统均对软件层提供了支持。

Matrix RAID 的原理就是将每个硬盘容量各分成两部分(即：将一个硬盘虚拟成两个子硬盘，这时子硬盘总数为 4 个)，其中用两个虚拟子硬盘来创建 RAID0 模式以提高效能，而其它两个虚拟子硬盘则透过镜像备份组成 RAID 1 用来备份数据。在 Matrix RAID 模式中数据存储模式如下：两个磁盘驱动器的第一部分被用来创建 RAID 0 阵列，主要用来存储操作系统、应用程序和交换文件，这是因为磁盘开始的区域拥有较高的存取速度，Matrix RAID 将 RAID 0 逻辑分割区置于硬盘前端(外圈)的主因，是可以让需要效能的模块得到最好的效能表现；而两个磁盘驱动器的第二部分用来创建 RAID1 模式，主要用来存储用户个人的文件和数据。

例如，使用两块 120GB 的硬盘，可以将两块硬盘的前 60GB 组成 120GB 的逻辑分割区，然后剩下两个 60GB 区块组成一个 60GB 的数据备份分割区。像需要高效能、却不需要安全性的应用，就可以安装在 RAID 0 分割区，而需要安全性备份的数据，则可安装在 RAID 1 分割区。换言之，使用者得到的总硬盘空间是 180GB，和传统的 RAID 0+1 相比，容量使用的效益非常的高，而且在容量配置上有着更高的弹性。如果发生硬盘损毁，RAID 0 分割区数据自然无法复原，但是 RAID 1 分割区的数据却会得到保全。

可以说，利用 Matrix RAID 技术，我们只需要 2 个硬盘就可以在获取高效数据存取的同时又能确保数据安全性。这意味着普通用户也可以低成本享受到 RAID 0+1 应用模式。



内容导航

## NV RAID

NV RAID 是 nVidia 自行开发的 RAID 技术，随着 nForce 各系列芯片组的发展也不断推陈出新。相对于其它 RAID 技术而言，目前最新的 nForce4 系列芯片组的 NV RAID 具有自己的鲜明特点，主要是以下几点：

(1)交错式 RAID(Cross-Controller RAID)：交错式 RAID 即俗称的混合式 RAID，也就是将 SATA 接口的硬盘与 IDE 接口的硬盘联合起来组成一个 RAID 模式。交错式 RAID 在 nForce3 250 系列芯片组中便已经出现，在 nForce 4 系列芯片组身上该功能得到延续和增强。

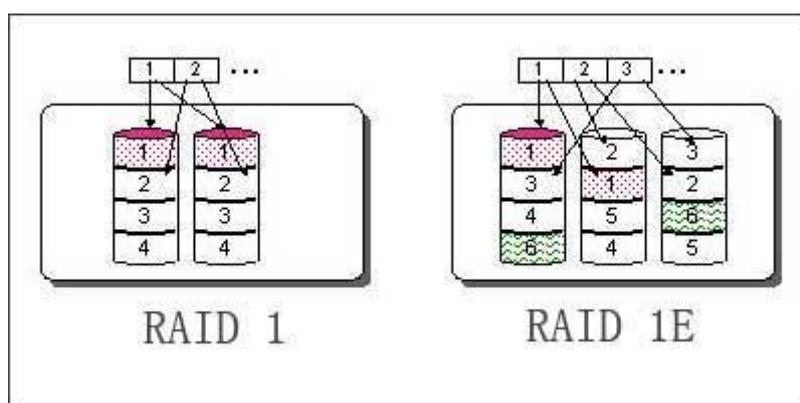
(2)热冗余备份功能：在 nForce 4 系列芯片组中，因支持 Serial ATA 2.0 的热插拔功能，用户可以在使用过程中更换损坏的硬盘，并在运行状态下重新建立一个新的镜像，确保重要数据的安全性。更为可喜的是，nForce 4 的 nVIDIA RAID 控制器还允许用户为运行中的 RAID 系统增加一个冗余备份特性，而不必理会系统采用哪一种 RAID 模式，用户可以在驱动程序提供的“管理工具”中指派任何一个多余的硬盘用作 RAID 系统的热备份。该热冗余硬盘可以让多个 RAID 系统(如一个 RAID 0 和一个 RAID1)共享，也可以为其中一个 RAID 系统所独自占有，功能类似于时下的高端 RAID 系统。

(3)简易的 RAID 模式迁移：nForce 4 系列芯片组的 NV RAID 模块新增了一个名为“Morphing”的新功能，用户只需要选择转换之后的 RAID 模式，而后执行“Morphing”操作，RAID 删除和模式重设的工作可以自动完成，无需人为干预，易用性明显提高。

内容导航

## RAID 1E

RAID 1E 是 RAID 1 的增强版本，它并不是我们通常所说的 RAID 0+1 的组合。RAID 1E 的工作原理与 RAID 1 基本上是一样的，只是 RAID 1E 的数据恢复能力更强，但由于 RAID 1E 写一份数据至少要两次，因此，RAID 处理器的负载得到加强，从而造成磁盘读写能力的下降。RAID 1E 至少需要 3 块硬盘才能实现。RAID 1E 和 RAID 1 的工作原理图如下：



## RAID DP

RAID DP 也属于一种私有的 RAID 标准，它实际上也就是双 RAID 3 技术，所谓双 RAID 3 技术主要是说在同一磁盘阵列中组建两个独立的校验磁盘，在单校验磁盘下工作原理与 RAID 3 一样，但增加了一个校验盘之后，则使整个磁盘阵列的安全性得到提高，并且它的性能比 RAID 3 和 RAID 5 都要好。

## RAID ADG

RAID ADG 相当于双 RAID 5 技术，是 HP 提出来的一种 RAID 技术。这种技术部署了 2 个奇偶校验集，并提供了 2 个硬盘的容量存储这些奇偶校验信息，能同时允许 2 块硬盘出现故障，有效提升了磁盘内数据的可靠性。不过这种技术会严重影响系统速度，所以并没有得到推广。

## 总结

以上提到的 RAID 技术都还不是规范的技术，因此，还有许多厂商也以定义了同样名字的 RAID 标准，但它们的原理有所区别，所以，大家在见到这些技术时，一定要根据厂商方面的解释为主。