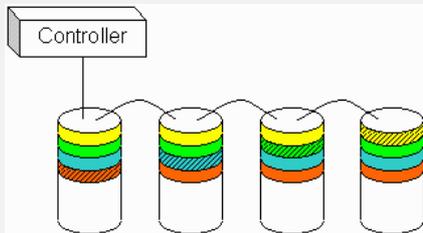


RAID 3采用的是一种较为简单的校验实现方式，使用一个专门的磁盘存放所有的校验数据，而在剩余的磁盘中创建带区集分散数据的读写操作。例如，在一个由4块硬盘构成的RAID 3系统中，3块硬盘将被用来保存数据，第四块硬盘则专门用于校验。这种配置方式可以用3+1的形式表示，具体如图：



在上图中，我们用相同的颜色表示使用同一个校验块的所有数据块，斜线标出的部分为校验块。校验块和所有对应的数据块一起构成一个带区。

第四块硬盘中的每一个校验块所包含的都是其它3块硬盘中对应数据块的校验信息。RAID 3的成功之处就在于不仅可以象RAID 1那样提供容错功能，而且整体开销从RAID 1的50%下降为25% (RAID 3+1)。随着所使用磁盘数量的增多，成本开销会越来越小。举例来说，如果我们使用7块硬盘，那么总开销就会将到12.5% (1/7)。

在不同情况下，RAID 3读写操作的复杂程度不同。最简单的情况就是从一个完好的RAID 3系统中读取数据。这时，只需要在数据存储盘中找到相应的数据块进行读取操作即可，不会增加任何额外的系统开销。

当向RAID 3写入数据时，情况会变得复杂一些。即使我们只是向一个磁盘写入一个数据块，也必须计算与该数据块同处一个带区的所有数据块的校验值，并将新值重新写入到校验块中。例如，当我们向上图中的绿色数据块写入数据时，必须重新计算所有3个绿色数据块的校验值，然后重写位于第四块硬盘的绿色校验块。由此我们可以看出，一个写入操作事实上包含了数据读取（读取带区中的关联数据块），校验值计算，数据块写入和校验块写入四个过程。系统开销大大增加。

我们可以通过适当设置带区的大小使RAID系统得到简化。如果某个写入操作的长度恰好等于一个完整带区的大小（全带区写入），那么我们就不要再读取带区中的关联数据块计算校验值。我们只需要计算整个带区的校验值，然后直接把数据和校验信息写入数据盘和校验盘即可。

到目前为止，我们所探讨的都是正常运行状况下的数据读写。下面，我们再来看一下当硬盘出现故障时，RAID系统在降级模式下的运行情况。

RAID 3虽然具有容错能力，但是系统会受到影响。当一块磁盘失效时，该磁盘上的所有数据块必须使用校验信息重新建立。如果我们是从好盘中读取数据块，不会有任何变化。但是如果我们所要读取的数据块正好位于已经损坏的磁盘，则必须同时读取同一带区中的所有其它数据块，并根据校验值重建丢失的数据。

当我们更换了损坏的磁盘之后，系统必须一个数据块一个数据块的重建坏盘中的数据。整个过程包括读取带区，计算丢失的数据块和向新盘写入新的数据块，都是在后台自动进行。重建活动最好是在RAID系统空闲的时候进行，否则整个系统的性能会受到严重的影响。

RAID 3的性能问题

除了我们在上文讨论过的有关数据写入和降级模式的问题之外，在使用RAID 3的过程中还有其他一些性能上的问题需要引起我们的注意。RAID 3所存在的最大一个不足同时也是导致RAID 3很少被人们采用的原因就是校验盘很容易成为整个系统的瓶颈。

我们已经知道RAID 3会把数据的写入操作分散到多个磁盘上进行，然而不管是向哪一个数据盘写入数据，都需要同时重写校验盘中的相关信息。因此，对于那些经常需要执行大量写入操作的应用来说，校验盘的负载将会很大，无法满足程序的运行速度，从而导致整个RAID系统性能的下降。鉴于这种原因，RAID 3更加适合应用于那些写入操作较少，读取操作较多的应用环境，例如数据库和WEB服务器等。