**[RAID](https://www.prepressure.com/library/technology/raid" \o "Permanent Link to RAID)**

RAID is a technology that is used to increase the performance and/or reliability of data storage. The abbreviation stands for *Redundant Array of Inexpensive Disks*. A RAID system consists of two or more drives working in parallel. These disks can be hard discs, but there is a trend to also use the technology for SSD (solid state drives). There are different RAID levels, each optimized for a specific situation. These are not standardized by an industry group or standardization committee. This explains why companies sometimes come up with their own unique numbers and implementations. This article covers the following RAID levels:

* [**RAID 0**](https://www.prepressure.com/library/technology/raid#raid-0) – striping
* [**RAID 1**](https://www.prepressure.com/library/technology/raid#raid-1) – mirroring
* [**RAID 5**](https://www.prepressure.com/library/technology/raid#raid-5) – striping with parity
* [**RAID 6**](https://www.prepressure.com/library/technology/raid#raid-6) – striping with double parity
* [**RAID 10**](https://www.prepressure.com/library/technology/raid#raid-10) – combining mirroring and striping

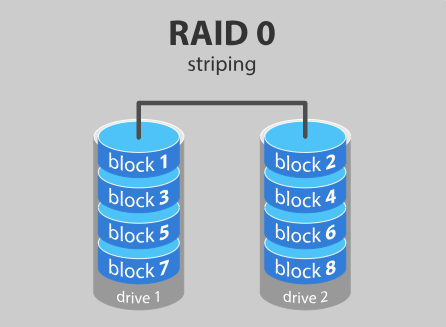
The software to perform the RAID-functionality and control the drives can either be located on a separate controller card (a hardware RAID controller) or it can simply be a driver. Some versions of Windows, such as Windows Server 2012 as well as Mac OS X, include software RAID functionality. Hardware RAID controllers cost more than pure software, but they also offer better performance, especially with RAID 5 and 6.

RAID-systems can be used with a number of interfaces, including SCSI, IDE, SATA or FC (fiber channel.) There are systems that use SATA disks internally, but that have a FireWire or SCSI-interface for the host system.

Sometimes disks in a storage system are defined as JBOD, which stands for *‘Just a Bunch Of Disks’*. This means that those disks do not use a specific RAID level and acts as stand-alone disks. This is often done for drives that contain swap files or spooling data.

Below is an overview of the most popular RAID levels:

RAID level 0 – Striping



In a RAID 0 system data are split up into blocks that get written across all the drives in the array. By using multiple disks (at least 2) at the same time, this offers superior I/O performance. This performance can be enhanced further by using multiple controllers, ideally one controller per disk.

Advantages

* RAID 0 offers great performance, both in read and write operations. There is no overhead caused by parity controls.
* All storage capacity is used, there is no overhead.
* The technology is easy to implement.

Disadvantages

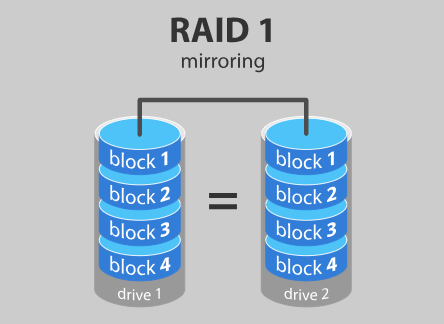
* RAID 0 is not fault-tolerant. If one drive fails, all data in the RAID 0 array are lost. It should not be used for mission-critical systems.

Ideal use

RAID 0 is ideal for non-critical storage of data that have to be read/written at a high speed, such as on an image retouching or video editing station.

If you want to use RAID 0 purely to combine the storage capacity of twee drives in a single volume, consider mounting one drive in the folder path of the other drive. This is supported in Linux, OS X as well as Windows and has the advantage that a single drive failure has no impact on the data of the second disk or SSD drive.

RAID level 1 – Mirroring



Data are stored twice by writing them to both the data drive (or set of data drives) and a mirror drive (or set of drives). If a drive fails, the controller uses either the data drive or the mirror drive for data recovery and continues operation. You need at least 2 drives for a RAID 1 array.

Advantages

* RAID 1 offers excellent read speed and a write-speed that is comparable to that of a single drive.
* In case a drive fails, data do not have to be rebuild, they just have to be copied to the replacement drive.
* RAID 1 is a very simple technology.

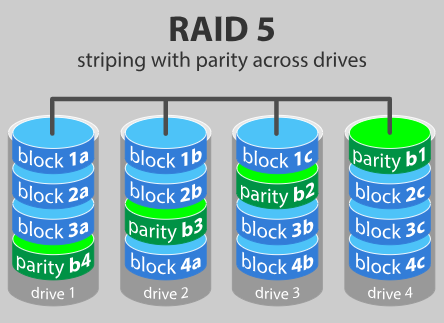
Disadvantages

* The main disadvantage is that the effective storage capacity is only half of the total drive capacity because all data get written twice.
* Software RAID 1 solutions do not always allow a hot swap of a failed drive. That means the failed drive can only be replaced after powering down the computer it is attached to. For servers that are used simultaneously by many people, this may not be acceptable. Such systems typically use hardware controllers that do support hot swapping.

Ideal use

RAID-1 is ideal for mission critical storage, for instance for accounting systems. It is also suitable for small servers in which only two data drives will be used.

RAID level 5



RAID 5 is the most common secure RAID level. It requires at least 3 drives but can work with up to 16. Data blocks are striped across the drives and on one drive a parity checksum of all the block data is written. The parity data are not written to a fixed drive, they are spread across all drives, as the drawing below shows. Using the parity data, the computer can recalculate the data of one of the other data blocks, should those data no longer be available. That means a RAID 5 array can withstand a single drive failure without losing data or access to data. Although RAID 5 can be achieved in software, a hardware controller is recommended. Often extra cache memory is used on these controllers to improve the write performance.

Advantages

* Read data transactions are very fast while write data transactions are somewhat slower (due to the parity that has to be calculated).
* If a drive fails, you still have access to all data, even while the failed drive is being replaced and the storage controller rebuilds the data on the new drive.

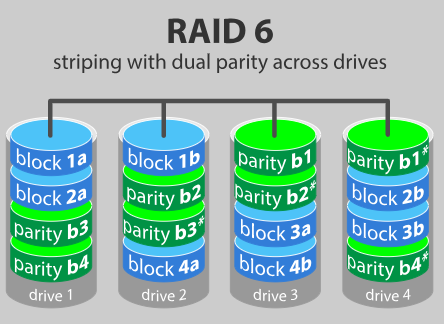
Disadvantages

* Drive failures have an effect on throughput, although this is still acceptable.
* This is complex technology. If one of the disks in an array using 4TB disks fails and is replaced, restoring the data (the rebuild time) may take a day or longer, depending on the load on the array and the speed of the controller. If another disk goes bad during that time, data are lost forever.

Ideal use

RAID 5 is a good all-round system that combines efficient storage with excellent security and decent performance. It is ideal for file and application servers that have a limited number of data drives.

RAID level 6 – Striping with double parity



RAID 6 is like RAID 5, but the parity data are written to two drives. That means it requires at least 4 drives and can withstand 2 drives dying simultaneously. The chances that two drives break down at exactly the same moment are of course very small. However, if a drive in a RAID 5 systems dies and is replaced by a new drive, it takes hours or even more than a day to rebuild the swapped drive. If another drive dies during that time, you still lose all of your data. With RAID 6, the RAID array will even survive that second failure.

Advantages

* Like with RAID 5, read data transactions are very fast.
* If two drives fail, you still have access to all data, even while the failed drives are being replaced. So RAID 6 is more secure than RAID 5.

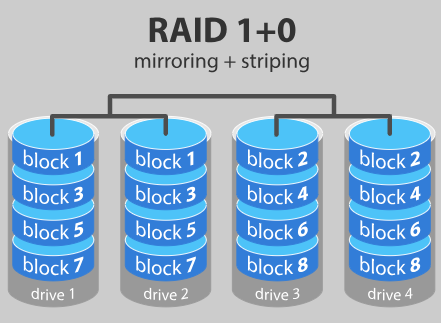
Disadvantages

* Write data transactions are slower than RAID 5 due to the additional parity data that have to be calculated. In one report I read the write performance was 20% lower.
* Drive failures have an effect on throughput, although this is still acceptable.
* This is complex technology. Rebuilding an array in which one drive failed can take a long time.

Ideal use

RAID 6 is a good all-round system that combines efficient storage with excellent security and decent performance. It is preferable over RAID 5 in file and application servers that use many large drives for data storage.

RAID level 10 – combining RAID 1 & RAID 0



It is possible to combine the advantages (and disadvantages) of RAID 0 and RAID 1 in one single system. This is a nested or hybrid RAID configuration. It provides security by mirroring all data on secondary drives while using striping across each set of drives to speed up data transfers.

Advantages

* If something goes wrong with one of the disks in a RAID 10 configuration, the rebuild time is very fast since all that is needed is copying all the data from the surviving mirror to a new drive. This can take as little as 30 minutes for drives of  1 TB.

Disadvantages

* Half of the storage capacity goes to mirroring, so compared to large RAID 5  or RAID 6 arrays, this is an expensive way to have redundancy.

What about RAID levels 2, 3, 4 and 7?

These levels do exist but are not that common (RAID 3 is essentially like RAID 5 but with the parity data always written to the same drive). This is just a simple introduction to RAID-systems. You can find more in-depth information on the pages of [**Wikipedia**](http://en.wikipedia.org/wiki/RAID) or **[ACNC](http://acnc.com/raid" \o "A visual tool to discover how RAID storage works)**.

RAID is no substitute for back-up!

All RAID levels except RAID 0 offer protection from a single drive failure. A RAID 6 system even survives 2 disks dying simultaneously. For complete security, you do still need to back-up the data from a RAID system.

* That back-up will come in handy if all drives fail simultaneously because of a power spike.
* It is a safeguard when the storage system gets stolen.
* Back-ups can be kept off-site at a different location. This can come in handy if a natural disaster or fire destroys your workplace.
* The most important reason to back-up multiple generations of data is user error. If someone accidentally deletes some important data and this goes unnoticed for several hours, days or weeks, a good set of back-ups ensure you can still retrieve those files.