

Introduction RAID : -

Storage is an important consideration when setting up a server. Almost all of the important information that you and your users care about will at one point be written to a storage device to save for later retrieval.

What is RAID ?

RAID stands for Redundant Arrays of Independent Disks. By combining drives in different patterns, administrators can achieve greater performance or redundancy than the collection of drives can offer when operated individually. RAID is implemented as a layer in between the raw drives or partitions and the file system layer.

When Is RAID a Good Idea?

The primary values that RAID provides are data redundancy and performance gains.

Redundancy is meant to help increase the availability of your data. This means that during certain failure conditions, like when a storage drive becomes faulty, your information is still accessible and the system as a whole can continue to function until the drive is replaced.

Hardware RAID, Software RAID, and Hardware-Assisted Software RAID

RAID arrays can be created and managed using a few different technologies.

Hardware RAID

Dedicated hardware called RAID controllers or RAID cards can be used to set up and manage RAID independent from the operating system. This is known as hardware RAID. True hardware RAID controllers will have a dedicated processor for managing RAID devices.

This has a number of advantages:

Performance :-

Genuine hardware RAID controllers do not need to take up CPU cycles to manage the underlying disks. This means no overhead for the management of the storage devices attached. High quality controllers also provide extensive caching, which can have a huge impact on performance.

Abstracting away complexity :-

Another benefit of using RAID controllers is that they abstract the underlying disk arrangement from the operating system. Hardware RAID can present the entire group of drives as a single logical unit of storage. The operating system does not have to understand the RAID arrangement; it can just interface with the array as if it were a single device.

Availability at boot:

Because the array is managed entirely outside of software, it will be available at boot time, allowing the root file system itself to easily be installed on a RAID array.

Hardware RAID also has a few significant disadvantages.

Vendor lock-in :-

Because the RAID arrangement is managed by the proprietary firmware on the hardware itself, an array is somewhat locked to the hardware used to create it. If a RAID controller dies, in almost all cases, it must be replaced with an identical or a compatible model. Some administrators recommend purchasing one or more backup controllers to use in the event that the first has a problem.

High cost :-

Quality hardware RAID controllers tend to be fairly expensive.

Software RAID

RAID can also be configured by the operating system itself. Since the relationship of the disks to one another is defined within the operating system instead of the firmware of a hardware device, this is called software RAID.

Some advantages of software RAID :-

Flexibility :-

Since RAID is managed within the operating system, it can easily be configured from available storage without reconfiguring hardware, from a running system. Linux software RAID is particularly flexible, allowing many different types of RAID configuration.

Open source :-

Software RAID implementations for open source operating systems like Linux and FreeBSD are also open source. The RAID implementation is not hidden, and can easily be read and implemented on other systems. For instance, RAID array created on an Ubuntu machine can easily be imported into a CentOS server at a later time. There is little chance of losing access to your data due to software differences.

No additional costs :-

Software RAID requires no specialty hardware, so it adds no additional cost to your server or workstation.

Some disadvantages of software RAID are:

Implementation-specific :-

Although software RAID is not tied to specific hardware, it tends to be tied to the specific software implementation of RAID. Linux uses mdadm, while FreeBSD uses GEOM-based RAID, and Windows has its own version of software RAID. While the open source implementations can be ported over or read in some cases, the format itself will likely not be compatible with other software RAID implementations.

Performance overhead :-

Historically, software RAID has been criticized for creating additional overhead. CPU cycles and memory are required to manage the array, which could be used for other purposes. Implementations like mdadm on modern hardware largely negates these concerns, however. CPU overhead is minimal and in most cases insignificant.

Hardware-Assisted Software RAID (Fake RAID)

A third type of RAID called hardware-assisted software RAID, firmware RAID, or fake RAID, is also available. Typically, this is found in RAID functionality within motherboards themselves or in inexpensive RAID cards. Hardware-assisted software RAID is an implementation that uses firmware on the controller or card to manage the RAID, but uses the regular CPU to handle the processing.

Advantages of hardware-assisted software RAID:

Multi-operating system support :-

Since the RAID is brought up during the early boot and then handed off to the operating system, multiple operating systems can use the same array, which might not be possible with software RAID.

Disadvantages of hardware-assisted software RAID:

Limited RAID support :-

Usually, only RAID 0 or RAID 1 are available.

Requires specific hardware :-

Like hardware RAID, hardware-assisted software RAID is tied to the hardware used to create and manage it. This issue is even more problematic when included in a motherboard, because a failure of the RAID controller can mean that you have to replace the entire motherboard to access the data again.

Performance overhead :-

Like software RAID, no CPU is dedicated to managing the RAID. Processing must be shared with the rest of the operating system.

Most administrators stay away from hardware-assisted software RAID as it suffers from a combination of the pitfalls of the other two implementations.