adaptec®

Hardware RAID vs. Software RAID: Which Implementation is Best for my Application?

Contents

Introduction

In the last couple of years, RAID (Redundant Array of Independent Disks) technology has grown from a server option to a data protection requirement. The first implementations of RAID in 1990 were very expensive controller boards with high-performance I/O processors that were as powerful as the host CPU. At that time, when hardware-based RAID solutions were the only option, the cost of a RAID controller limited the usage to high-priced servers

Today, RAID is found everywhere – from an operating system software feature to a stand-alone controller providing advanced data integrity in high-end storage area networks. It can be found in mobile environments such as laptops, as well as desktops, workstations, servers, and external enclosures with a larger number of hard disk drives. RAID is even included in TV set-top boxes or personal storage devices.

This white paper provides a high-level overview of the various RAID solutions. It begins by defining "Software RAID" vs. "Hardware RAID", explains how these different RAID implementations work and the benefits of each implementation, and helps you make a decision on which RAID solution is best for your application.

What Is RAID?

RAID, an acronym for Redundant Arrays of Inexpensive Disks, is a way to virtualize multiple, independent hard disk drives into one or more arrays to improve performance, capacity and reliability (availability). The total array capacity depends on the type of RAID array you build and the number and size of disk drives. This total array capacity is independent of whether you use software or hardware RAID. The following sections look at the different implementations, the strengths and weaknesses and their impact to system performance and effectiveness in enhancing data availability.

RAID Level Comparison

Features	RAID 0	RAID 1	RAID 1E	RAID 5	RAID 5EE	RAID 6	RAID 10	RAID 50	RAID 60
Minimum # Drives	2	2	3	3	4	4	4	6	8
Data Protection	No Protection	Single-drive failure	Single-drive failure	Single-drive failure	Single-drive failure	Two-drive failure	Up to one disk failure in each sub-array	Up to one disk failure in each sub-array	Up to two disk failures in each sub-array
Read Performance	High	High	High	High	High	High	High	High	High
Write Performance	High	Medium	Medium	Low	Low	Low	Medium	Medium	Medium
Read Performance (degraded)	N/A	Medium	High	Low	Low	Low	High	Medium	Medium
Write Performance (degraded)	N/A	High	High	Low	Low	Low	High	Medium	Low
Capacity Utilization	100%	50%	50%	67% - 94%	50% - 88%	50% - 88%	50%	67% - 94%	50% - 88%
Typical Applications	High end workstations, data logging, real-time rendering, very transitory data	Operating system, transaction databases	Operating system, transaction databases	Data warehousing, web serving, archiving	Data warehousing, web serving, archiving	Data archive, backup to disk, high availability solutions, servers with large capacity requirements	Fast databases, application servers	Large databases, file servers, application servers	Data archive, backup to disk, high availability solutions, servers with large capacity requirements

SERVER STORAGE WHITE PAPER

Hardware RAID vs. Software RAID: Which Implementation is Best for my Application?   2

Software RAID

A simple way to describe software RAID is that the RAID task runs on the CPU of your computer system.
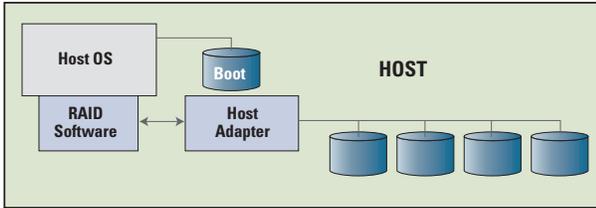
Some software RAID implementations include a piece of hardware, which might make the implementation seem like a hardware RAID implementation, at first glance. Therefore, it is important to understand that software RAID code utilizes the CPU's calculating power. The code that provides the RAID features runs on the system CPU, sharing the computing power with the operating system and all the associated applications.
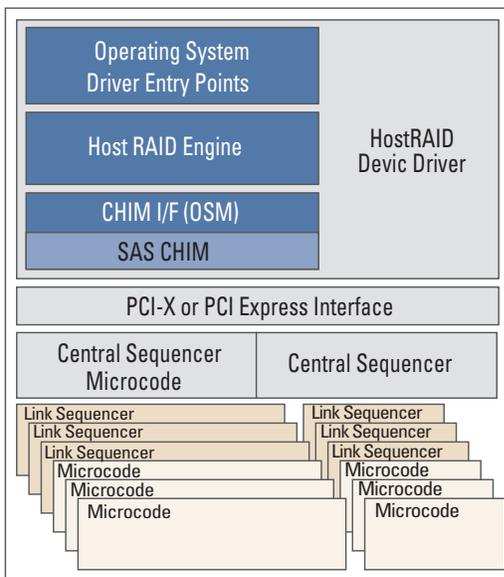
Software RAID Implementations

Software RAID can be implemented in a variety of ways: 1) as a pure software solution, or 2) as a hybrid solution that includes some hardware designed to increase performance and reduce system CPU overhead.

1) Pure Software Model – Operating System Software RAID

In this case, the RAID implementation is an application running on the host without any additional hardware. This type of software RAID uses hard disk drives which are attached to the computer system via a built-in I/O interface or a processor-less host bus adapter (HBA). The RAID becomes active as soon as the operating system has loaded the RAID driver software. Such pure software RAID solutions often come integrated into the server OS and usually are free of additional cost for the user. Low cost is the primary advantage of this solution.

Benefits and drawbacks of "pure," Operating System Software RAID:

- **Low cost:** No extra charge for the RAID functionality, as it is built into the OS. The only cost is the additional disk drives.

  - **Unprotected at boot (cannot manage or protect data at boot):** Drive failure or corrupted data during boot and before the RAID software become active leads to an inoperable system.

  - **Additional performance load on server:** System performance is impacted by the RAID application. The more drives involved and the more complex the RAID system is (e.g. involving parity on a RAID 5), the more impact on the overall performance. This solution is better suited for simple RAID 0, 1, 10 scenarios.

  - **Limited operating system migration:** RAID functionality might be limited to the current OS. There is no way to migrate the array to other OSes or other versions of the same OS, if not all versions of the OS support RAID functionality).

  - **Vulnerable to viruses:** Because RAID is running as an application on the computer system, viruses and other harmful software could impact RAID functionality.

  - **Data integrity issues due to system crashes:** Software or hardware problems on the server can impact data consistency and integrity.

  - **No write-back cache:** Software RAID runs only in write-through mode, but hardware RAID can run in write-back mode if it has a battery, adding another level of data protection. Write-back mode significantly enhances the write performance of a RAID array. There is no way to add a battery for software RAID.
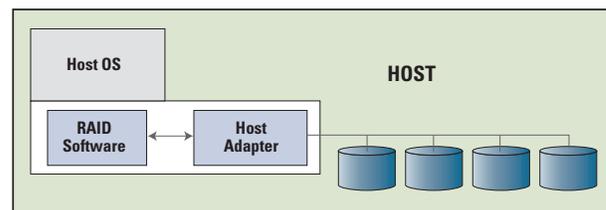
2) Hybrid Model – Hardware-Assisted Software RAID

While this is still software RAID, the hardware assist helps to overcome some of the weaknesses of pure software RAID. Such solutions usually come with additional hardware (e.g. an HBA with a RAID BIOS or just a RAID BIOS integrated onto the motherboard). The additional BIOS makes the RAID functionality available when the system is switched on, providing redundancy during boot that reduces the impact of medium errors on RAID which otherwise could lead to data corruption or an inoperable system. In addition, most of these solutions provide a BIOS setup software which is available at system boot. This allows for easy setup and maintenance of the RAID array without the need to install or boot an OS from hard disk or CD-

SERVER STORAGE WHITE PAPER

Hardware RAID vs. Software RAID: Which Implementation is Best for my Application?   3

ROM. Additionally, hardware-assisted software RAID usually comes with a variety of drivers for the most popular operating systems, and therefore, is more OS independent than pure software RAID.

Benefits and Drawbacks of hybrid, Hardware-Assisted Software RAID:

- **Moderate cost:** Only an HBA (plug-in card) or an additional flash memory for the BIOS down on the motherboard is needed; it may also include a hardware XOR accelerator if the controller supports RAID 5.

- **Protected at boot:** No negative impact on data availability when boot drive has medium errors or fails completely.

- **Dedicated GUI and software to build and maintain the RAID:** Easy setup and maintenance of the RAID array

  - **Additional performance load on server.** Systems performance is impacted by the RAID application. The more drives involved and the more complex the RAID system is (e.g. involving parity on a RAID 5), the more impact on the overall performance. This solution is better suited for simple RAID 0, 1, 10 scenarios.

  - **Limited operating system migration:** RAID functionality is still dependent on the OS as the driver runs on top of the OS. However, multiple drivers for a variety of OSes do allow migration of the array to other OSes. This might be limited by the availability of a RAID driver for very new OSes (e.g. newer versions of an OS might need a new RAID driver – RAID drivers are more complex than normal HBA driver programs, and therefore, it might take longer to develop them)

  - **Vulnerable to viruses:** Because RAID is running as an application on the computer system, viruses and other harmful software might impact RAID functionality

  - Data integrity issues due to system crashes: Software or hardware problems on the server can impact data consistency and integrity.

  - **No write-back cache:** Hardware-assisted software RAID runs only in write-through mode, but hardware RAID can run in write-back mode if it has a battery, adding another level of data protection. Write-back mode significantly enhances the write performance of a RAID array There is no way to add a battery for hardware assisted software RAID.

Hardware RAID

A hardware RAID solution has its own processor and memory to run the RAID application. In this implementation, the RAID system is an independent small computer system dedicated to the RAID application, offloading this task from the host system.
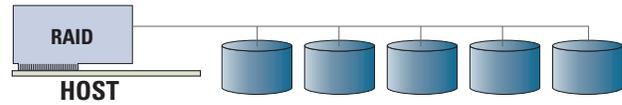
Hardware RAID can be found as an integral part of the solution (e.g. integrated in the motherboard) or as an add-in card. If the necessary hardware is already integrated in the system solution, then hardware RAID might become a software upgrade to your existing system. So like software RAID, hardware RAID might not be identified as such at first glance.

The simplest way to identify whether a solution is software or hardware RAID is to read the technical specification or data sheet of the RAID solution. If the solution includes a microprocessor (usually called I/O processor, processor or sometimes ROC – which means 'RAID on Chip'), then is the solution is a hardware RAID solution. If there is no processor, it is a software RAID solution.

This is important for your selection because of the system impacts of the software RAID vs. hardware RAID implementation. These impacts include:

- CPU utilization and performance when other applications are running

- Scalability of disk drives that can be added to a system

- Ease of recovery after a data loss

- Capability for advanced data management/monitoring

- Ability to manage disk drives consistently across different operating systems

- Ability to add a battery backup option that allow to enable write caching on the controller to enhance write performance of the system

Hardware RAID Implementations

Hardware RAID can be implemented in a variety of ways: 1) as a discrete RAID Controller Card, or 2) as integrated hardware based on RAID-on-Chip technology.

1) A Discrete RAID Controller Card

This is a plug-in expansion card that usually has a built-in RAID processor (I/O processor) and its own interfaces to the drives (I/O controllers). It usually plugs into PCI-X or PCIe slots of the computer system's motherboard. Such plug-in cards are usually the most costly, but also the most flexible and best performing RAID solutions. Here, the RAID functionality is completely independent of the host (computer system). The discrete character of this card allows the usage of the highest performing I/O processor and fastest memory. The RAID card fully offloads the task of creating a redundant storage subsystem from the computer system and does not impact the performance of the rest of the system, even in the case of a drive failure. More complex and space-efficient RAID levels (like RAID 5 or RAID 6) can be used without any impact on the system.

The additional I/O interfaces on the card usually allow for greater system expandability (adding more hard disk drives and more capacity). Even multiple RAID arrays do not impact the performance of the host system. These can easily be migrated to another operating system or even another host system or platform.

Benefits and Drawbacks of a Discrete RAID Controller Card:

- Protected at boot: No negative impact on data availability when boot drive has medium errors or fails completely.

- Performance independent of workload on server: Fast memory, fast processor and no impact on the performance of the application(s) running on the host system.

- RAID application independent of host: No data integrity issues when system crashes.

- Enhanced protection in case of power loss: Hardware RAID implementations typically keep track of in-progress writes in non-volatile hardware. Software RAID implementations lack this protection, which makes it difficult to recover from a power loss during a write.

- Not vulnerable to viruses: RAID arrays are completely independent of the host system and OS. No data integrity issues occur if the host system fails.

- Offloads the RAID task from the host: Best suited for complex RAID 5 or RAID 6 scenarios which usually offers best cost/performance ratio.

- Dedicated GUI and software to build and maintain the RAID: Easy setup and maintenance of the RAID array.

- Easy to migrate and replace: Card can be plugged in any system and easily replaced or upgraded with the latest, best performing variant. It is also easier to migrate from one operating system to another.

- Supports advanced RAID features: Typical examples include disk hot plug, array-level migration and online capacity expansion.

- On-controller caching: Accelerates access times by usage of cache memory including the ability to use write-back caching if memory is protected by a battery.

  - Highest cost: I/O processor and additional memory on a plug-in card increases cost

2) Integrated Hardware RAID solutions based on RAID-on-Chip (ROC) technology

In ROC solutions, the RAID processor, memory controller, host interface, I/O interfaces for hard disk drive connection, and sometimes even the memory, are all integrated into one single chip. This chip can be integrated into the motherboard and offers hardware RAID capabilities with reduced cost (as just one, highly integrated ASIC, is needed).

The ROC replaces the I/O interface chip you find on many server motherboards (e.g. a SCSI controller chip). That means that the ROC solution integrates more than just the connectivity for the hard disk drives.
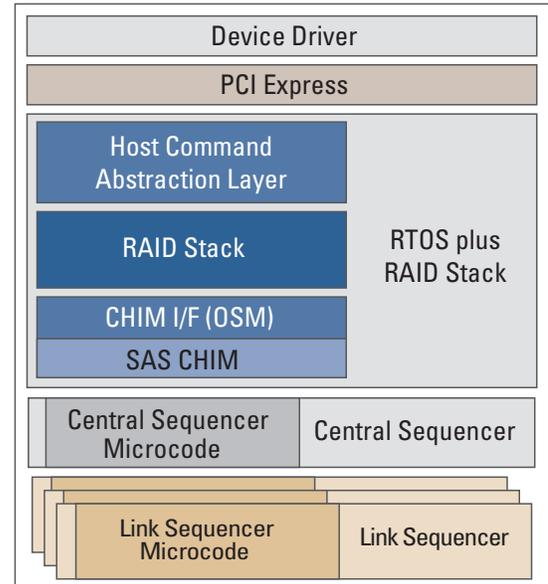
Benefits and Drawbacks of ROC-based Hardware RAID:

- Protected at boot: No negative impact on data availability when boot drive has medium errors or fails completely.

- RAID application independent of host: No data integrity issues when system crashes.

- Not vulnerable to viruses: RAID array completely independent of the host system and OS. No data integrity issues if host system fails.

- Enhanced protection in case of power loss: Hardware RAID implementations typically keep track of in-progress writes in non-volatile hardware. Software RAID implementations lack this protection, which makes it difficult to recover from a power loss during a write.

- Offloads the RAID task from the host: Best suited for complex RAID 5 or RAID 6 scenarios which usually offers best cost / performance ratio

- Dedicated GUI and software to build and maintain the RAID: Easy setup and maintenance of the RAID array

- Supports advanced RAID features: Typical examples include disk hot plug, array-level migration and online capacity expansion.

- On-controller caching: Accelerates access times by usage of cache memory including the ability to use write-back caching if memory is protected by a battery.

  - Medium cost: Lower chip count reduces cost and adds reliability compared to a discrete hardware RAID solution. Due to the high integration of an I/O processor and I/O interfaces into one single chip, the clock frequency of these complex chips might be limited for now. Next-generation ROC solutions may overcome this.

  - Limited flexibility & migration: RAID migration to other systems is only possible of they are equipped with a similar (compatible) ROC solution.

SERVER STORAGE **WHITE PAPER**

Hardware RAID vs. Software RAID: Which Implementation is Best for my Application?   5

Is Hardware or Software RAID Best for Your Application?

Now that we understand the benefits and drawbacks of the different RAID implementations, we can look at some typical server deployment scenarios and optimize the RAID subsystem to meet the overall server price performance goals.

Pure Software RAID solutions: Entry-level RAID 0 or 1 for performance or data availability is key. However, software RAID cannot be used on the boot drive as the operating system RAID functionality is not available until after the OS boots.

Targeted Applications:

1. Workstations without large data storage requirements

2. Entry-level servers without boot protection requirements

Hybrid solutions: Cost-sensitive solutions similar to pure software RAID, but with bootability requirements.

Targeted Applications:

1. Entry-level servers without large storage requirements

2. Compute engines connected to networked storage

Hardware RAID solutions: Most feature-rich and highest performance solution. This may be implemented as RAID on the Motherboard (ROMB) or with plug-in cards to scale performance and availability (e.g. controller redundancy).

Targeted Applications:

1. High Performance workstations with large data storage requirements

2. Entry level to enterprise servers requiring performance and scalability from the storage subsystem.

In general, hardware RAID offers more advantages over software-based solutions for more complex RAID algorithms. Measurements indicate that, for instance, software RAID-6 implementations impose a heavy load on system compute resources, especially in degraded mode. This makes hardware

RAID attractive for such configurations.

Features of the different RAID implementations

Features	Pure Software RAID	Hardware-assisted SW RAID	HW RAID controller – ROC or add-in card
Data protection during boot	No	Yes	Yes
Write-back caching possible	No	No	Yes
Enhanced protection in case of power loss	No	No	Yes
RAID independent of host operating system	No	No	Yes
RAID performance	Depends on server load	Depends on server load	High: Independent of server load
RAID functionality vulnerable to viruses	Yes	Yes	No
Setup during boot	No	Yes	Yes
Ability to migrate to other OS versions	No	Limited	Yes
Typical RAID applications	RAID 0, 1	RAID 0, 1	Advanced RAID: RAID 5 or RAID 6

Closing

This paper has shown the benefits that hardware RAID offers over software RAID implementations.

Advances in silicon technology are allowing for integration of the processors required for hardware RAID into single chip solutions and soon into commodity server chipsets, which will shrink the implementation cost. This reduced cost will enable a broader deployment of hardware RAID solutions into lower-priced servers, making them available to more users.

Currently, new levels of data protection and manageability are available – RAID 6 dual drive failure protection and data encryption are two examples. In many cases, these hardware RAID implementations can provide higher performance and lower cost solutions than external RAID configurations.

adaptec®