Impact of RAID Settings on Performance Configuring RAID for Optimal Performance

Write Policy

The Write Policy can have a very big impact on write performance. There are two modes available:

1. Write Back

2. Write Thru.

Write Back Mode

This mode provides better performance in most cases. In Write-Back mode, the RAID controller acknowledges write I/O requests immediately after the data loads into the controller cache. The application can continue working without waiting for the data to be physically written to the hard drives.

If a power loss occurs in write-back mode, there is a risk of losing data in the RAID cache. The data loss may be fatal and may require restoring data from a backup device. It is critical to have protection against power failures. Using a UPS with redundant system power supplies is highly recommended. RAID Backup Battery Unit can provide additional protection.

Write Thru Mode

This mode does not utilize the RAID cache for accelerating write I/O requests.

In most cases it will be slower than Write-Back mode. However, Write Thru mode allows achieving the highest sequential write bandwidth with RAID 0 or RAID 10.

Disk Cache Policy

Disk Cache Policy determines whether the hard-drive write cache is enabled or disabled. When Write Policy is set to Write Thru mode, Disk Cache Policy can have very big impact on write performance. When Write Policy is set to Write Back mode, impact of Disk Cache Policy is much smaller and in many cases negligible.

When Disk Cache Policy is enabled, there is a risk of losing data in the hard drive cache if a power failure occurs. The data loss may be fatal and may require restoring the data from a backup device. It is critical to have protection against power failures. Using a UPS with redundant system power supplies are highly recommended. RAID Backup Battery Unit can provide additional protection.

Note: A RAID Backup Battery Unit does not protect the hard drive cache. Impact of RAID Settings on Performance Configuring RAID for Optimal Performance Read Ahead Policy.

The Read Ahead Policy determines whether the RAID controller will read just a block of data that an application has requested, or whether it will read the whole stripe from the hard-drives. This setting can have big impact on read performance.

No Read Ahead (Normal)

The RAID controller will read only the block of data that the application has requested. This mode is preferred when read requests are primarily random. Also this mode is recommended when measuring sequential read throughput with I/O meter* under Windows.

Always Read Ahead

The RAID controller will read the whole stripe containing the requested data block and will keep it in cache. Each read operation will consume more hard drive resources, but if the read requests are primarily sequential it can substantially reduce the amount of read requests to the hard drives and can substantially increase performance.

Note: This setting will only make difference if the typical read request size is smaller than the stripe width.

Adaptive Read Ahead

The RAID controller automatically adjusts the read policy based on the current pattern of read requests. It combines the benefits of No Read Ahead and Always Read Ahead modes. This mode is recommended if the workload has mixed sequential and random patterns, or if the pattern is unknown.

I/O Policy

The I/O Policy determines whether the RAID controller will keep data in the cache, which can reduce the access time if subsequent read requests are made to the same data blocks.

Direct I/O

Direct IO mode is recommended in most cases. Most file systems and many applications have their own cache and do not require caching data at the RAID controller level.

Cached I/O

In Cached I/O mode the controller caches both read and write requests.

If there are subsequent read requests to the same data blocks, they are read from the RAID cache instead of the hard drives. This mode may be required if the application or file system does not cache read requests.

Strip Size

Strip size determines how data is distributed across hard drives. It also determines how many drives are accessed to service a single I/O request. Typically, sequential workloads benefit from using large strip sizes (512 KB or 1MB). With random types of access, the strip size depends on the typical access block size and on data alignment.

For example,
If a database is using 16 KB records with 16 KB alignment, the optimal strip size can be 16KB. For file- or web-server a large (512 KB or 1 MB) strip size can be optimal.

Software vendor documentation often provides recommendations on how to select RAID strip sizes.

Note: Matching the strip size to the file system cluster size does not usually provide any benefit. Data block or file sizes used by the application are usually more important. However, setting the strip size smaller than the cluster size is not recommended.

Other Performance Factors

When measuring performance of a RAID subsystem, it is important to remove factors that can limit the performance or cause variations in the performance.

Backup Battery Status

When doing write performance measurements in Write Back mode, it is important to check the status of the battery and the Current Write Policy. When the battery is not fully charged or is in the process of relearning, Write Policy will be automatically switched to Write Thru. This will have big impact on write performance.

You can disable the Write Thru for a failed or missing battery option to make sure that the Write Back mode is used regardless of the current battery status. In a production environment disabling this option may cause data loss if power failure occurs when the battery does not have sufficient level of charge.