Technical white paper

HP Smart Array Controllers and basic RAID performance factors

Technology brief

Table of contents

Click here to verify the latest version of this document.

Created December 2012.

Abstract

RAID storage technology continues to make advancements in drives, storage interfaces, RAID controller technology, and processing power. The features of the HP Smart Array Controller, the RAID level that you use, and drive technologies can all affect overall RAID performance. This technology brief provides an overview of the basic factors affecting RAID performance today.

Benefits of drive arrays

Drive arrays address several basic issues with drive-based storage:

- Drive arrays allow the creation of large storage volumes using multiple, smaller drives.
- Drive arrays increase the I/O capabilities and maximum throughput of the storage subsystem over that of individual drives.
- Drive arrays increase the reliability of data storage by using redundancy techniques (RAID levels) to ensure that the failure of one or more physical drives does not result in a permanent loss of data.

Factors that affect performance

Many variables influence the overall performance of drive arrays. The primary factors that influence performance are:

- **RAID level**

  The processing that the RAID level requires for high-level and low-level read and write operations affects performance.
- **Array controller**

  The processor and memory required to manage and execute the RAID operations as well as the read and write cache used to optimize read/write performance all influence performance.
- **Number of drives in the drive array**

  The number of drives in an array inclucences performance because the Smart Array controller can execute more read and write operations in parallel.
- **Drive performance**

  Drive throughput capability (MiB/s) influences RAID performance when performing random reads and writes (I/Os per second or IOPS).
- **Storage interface performance**

  Storage interface performance, including the protocols (SAS and SATA) and the speed of the physical links between the drives and the controller (3 GiB/s or 6 GiB/s) are factors that affect performance.

Each of these variables influences RAID performance. In addition, depending on the type of storage operation, these factors can determine the upper limit of the drive array's performance in a particular application environment.

HP Smart Array Controller performance

The Smart Array controller includes a processor, cache, and device driver that all contribute to providing optimum RAID performance.

Smart Array processor

The processor on the Smart Array controller manages the RAID system and transforms high-level read or write requests from an application into the complex series of individual instructions for the drive array.

Its capabilities are particularly critical to complex RAID operations, particularly write operations for redundant RAID modes, such as RAID 5 and RAID 6. Both RAID 5 and RAID 6 use mathematical XOR (Exclusive or) operations to calculate parity data. Redundant parity data provides data recovery capability if a physical drive failure occurs. Because the processor writes the parity data to the drive array, the performance of the XOR operations is a key contributor to overall write performance of parity-based arrays.

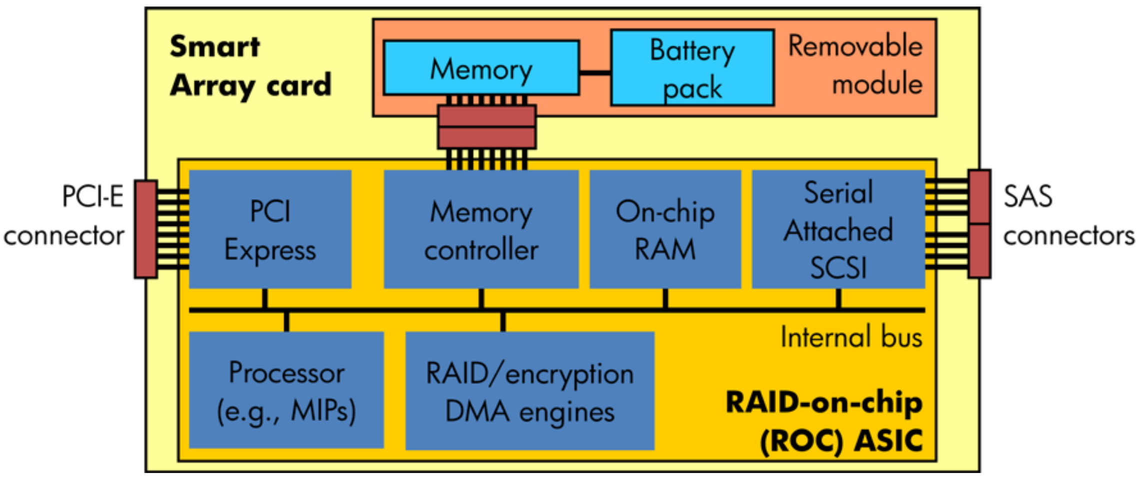
Performance improvements are most apparent in arrays with larger drive counts. The aggregate I/O of smaller drive counts, not the bandwidth of the Smart Array processor, constrains drive array performance.

The following Smart Array controllers use an embedded RAID-on-Chip (RoC) processor running at 1000 MHz:

- HP Smart Array P222 Controller
- HP Smart Array P420 Controller
- HP Smart Array P421 Controller
- HP Smart Array P822 Controller

While it is not a direct measure of overall RAID performance, the RoC processor can support up to 200,000 4 KB random Input/Output Operations Per Second (IOPS). Previous generations of the processors support up to 60,000 4 KB random IOPS.

Figure 1. Smart Array controller architecture



Smart Array cache

Smart Array controllers use cache to improve the overall performance of drive arrays for both read and write operations. You can use the HP Array Configuration Utility (ACU) to configure the percentage of the cache to use for read caching and write caching.

Cache width

Gen8 Smart Array controllers support the following cache modules:

- 512 MiB, 40-bit wide (32 bits data + 8 bits parity) cache module
- 1 GiB, 72-bit wide (64 bits data + 8 bits parity) cache module
- 2 GiB, 72-bit wide (64 bits data + 8 bits parity) cache module

The 1 GiB and 2 GiB cache modules improve array performance because they provide significantly more cache for read and write operations, and double the bandwidth for moving cache data to and from the storage system.

Read cache

Read cache does not improve array read performance significantly because simply reading from a drive array is already fast. The default configuration on Smart Array controllers assigns only 10% of the available cache space for read cache.

Read cache is most effective in increasing the performance for sequential small-block read workloads and, in particular, read workloads at low queue depth. The Smart Array controller differentiates between sequential and random workloads. It uses read cache in a predictive capacity to pre-fetch data when it detects sequential workloads. It identifies the pattern of the read commands, and then reads ahead on the drives. After reading the data, the Smart Array controller puts that data into the cache, so it is available if the upcoming read commands call for it.

Write Cache

Through a process known as "posted writes" or "write-back caching," Smart Array controllers use the write cache as an output buffer. Applications post write commands to the Smart array controller, and then continue without waiting for completion of the write operation to the drive. The application sees the write as completed in a matter of microseconds instead of milliseconds. In high workload environments, the write cache typically fills up and remains full most of the time.

The Smart array controller writes the data to the drive as it works through the list of write commands in its write cache. It analyzes the pending write commands, and then determines how to handle them most efficiently. Two techniques for improving efficiency are:

- Write coalescing: The controller combines small writes of adjacent logical blocks into a single, larger write.
- Command reordering: The controller rearranges the execution order of the writes in the cache to reduce overall drive latency.

When the Smart Array controller has a large cache memory size, it can coalesce and reorder commands efficiently, which improves overall array performance.

Zero Memory RAID

Smart Array controllers can ship without cache as part of their standard configuration (known as Zero Memory RAID). Zero Memory RAID provides entry-level RAID functionality.

Zero Memory RAID affects more than performance. It also limits the functionality that the controller can support. For example, the Smart Array controller with cache uses a significant amount of the cache to execute advanced RAID functions. These functions include performing XOR operations to calculate parity for RAID 5 and RAID 6 logical drives. Without cache, the controller cannot perform these operations and therefore cannot support those RAID levels. As a result, Zero Memory RAID supports RAID 0 (no fault tolerance), RAID 1, and a limited number of physical drives in an array.

Overall effect of cache on Smart Array performance

Using cache, particularly write cache, improves performance significantly. Read cache may provide modest performance gains for read operations; however, write cache is crucial to improving the write performance of drive arrays. This is because advanced RAID levels may require up to six individual read and write operations to physical drives in order to complete a single array-level "write" to a logical drive.

Figures 2 through 4 show how various cache levels affect the relative performance as follows:

- Figure 2 shows 4 KiB Random Write
- Figure 3 shows a mixture of 67% random read and 33% random write, OLTP
- Figure 4 shows the cache impact on RAID 5 sequential write

The configuration for this testing included an HP ProLiant GL380p Gen8 Server with an HP Smart Array P421 Controller with various RAID 5 logical drive configurations using a strip size of 256 KB.

Each chart value is the maximum value from a test measuring across queues of 1 to 256. RAID 5 measurements with more than 14 drives is for illustrative purposes only. We do not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).

Figure 2. RAID 5 Random write performance: Comparing 4 KiB random write performance with 512 MiB versus 2 GiB FBWC.

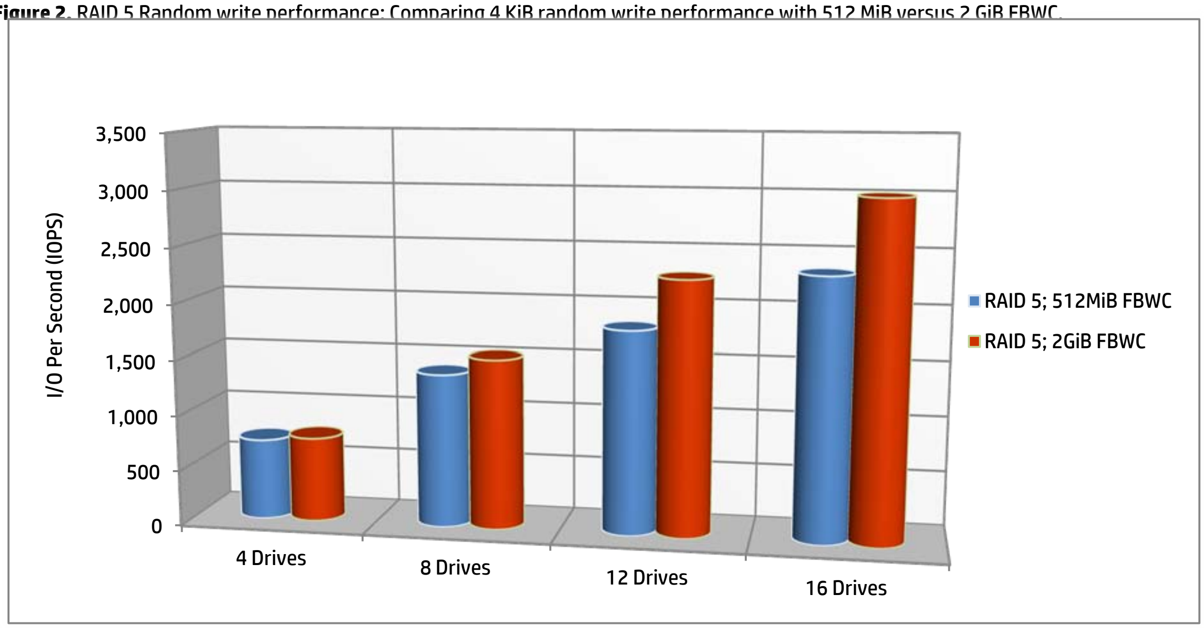


Figure 3. RAID 5 Online Transaction Processing (OLTP) performance (4 KiB aligned, 1/3 random write, 2/3 random read): Comparing 4 KiB OLTP write performance with 512 MiB FBWC versus 2 GiB FBWC.

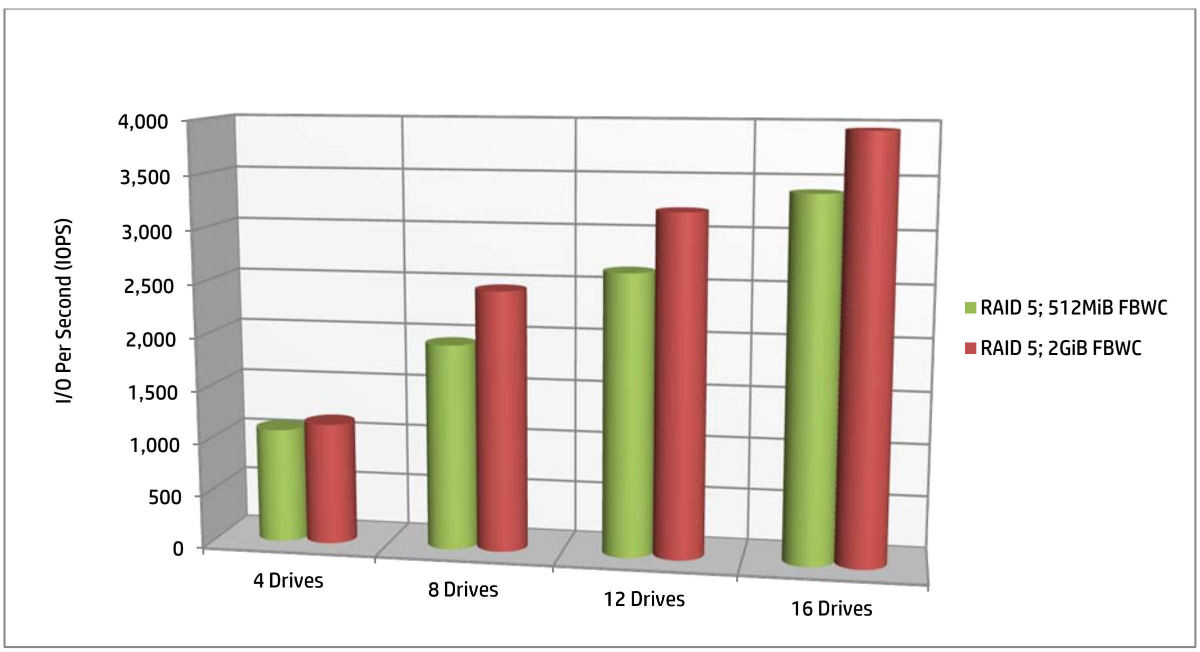
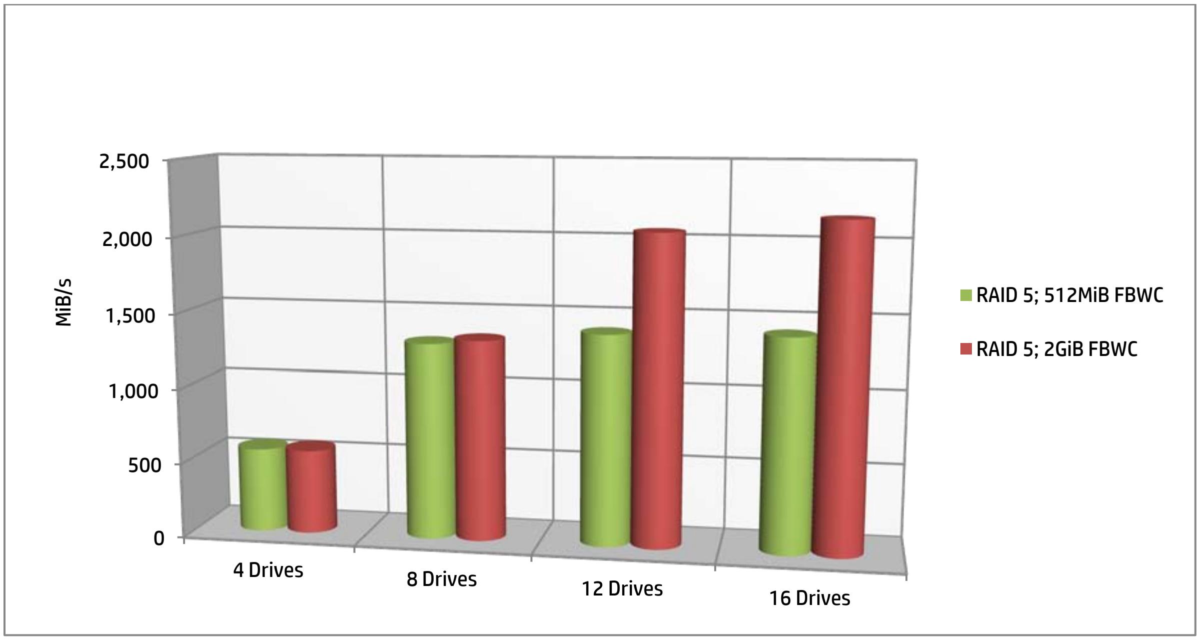


Figure 4. RAID 5 Sequential write performance: Comparing sequential write performance with 512 MiB FBWC versus 2 GiB FBWC.



Optional battery backed or flash backed write cache

In the event of a server or power failure, the data that is in the write cache at the time of the failure could be lost, resulting in possible data corruption. To avoid possible data loss or corruption, we recommend using flash backed write cache (FBWC). FBWC uses onboard power from a capacitor to write the cached data to non-volatile flash memory where it can remain almost indefinitely. FBWC maintains write cache integrity even in the event of a server or power failure.

Earlier versions of the Smart Array controller can use a battery backed write cache (BBWC). BBWC has a battery attached to maintain the contents of cache module. The batteries are capable of maintaining cache data for up to 72 hours.

Smart Array controllers can ship with FBWC; others have BBWC or FBWC available as options. Without BBWC or FBWC, the Smart Array controller will not use any of the cache for posted writes or write-back caching. It is possible to override this default configuration; however, we do not advise doing so. Consequences of using write cache without a BBWC or FBWC include:

- Data loss or corruption if a server or power failure occurs
- Significant degradation of write performance, particularly in RAID 5 and RAID 6 modes and their derivatives

Drives and performance

Smart Array controllers connect to the drives in an array through SAS-2 physical links (SAS ports):

- New generations of Smart Array P222, P420i, P420, P421, and P721m support up to eight primary SAS ports.
- The Smart Array P220i supports up to four primary SAS ports.
- The Smart Array P822 supports up to twenty-four primary SAS ports.

Each physical link can support a maximum bandwidth of the following:

- 6 GiB/s (600 MiB/s) when 6 GiB/s SAS drives are attached
- 6 GiB/s (600 MiB/s) when 6GiB SATA drives are attached
- 3 GiB/s (300 MiB/s) when 3GiB SATA drives are attached

The IOPS that a drive can sustain influences the random read and write performance of drive arrays.

When using spinning media drives, the SAS-2 physical link's bandwidth does not affect performance in application environments that rely heavily on random read and write operations. Consider the following points:

- Currently, the fastest spinning media drives can deliver about 470 random IOPS using 4 KiB reads and writes. This translates to a throughput of 1.8 MiB/s, which is less than one percent of a SAS channel's bandwidth.
- Even in a larger RAID configuration using a SAS expander with six drives behind a single SAS channel, the aggregate throughput is less than 15 MiB/s with six drives, which is still less than the SAS-2 physical link's bandwidth.
- With Solid State Disks (SSDs) becoming universally available, IOPS levels have increased greatly. A 6 GiB SAS SSD can deliver over 50,000 IOPS.

As Table 1 shows, no single drive can sustain a throughput that will saturate a 3 GiB/s SAS-2 physical link. With sequential operations, particularly sequential reads, the SAS-2 physical link's bandwidth can become a factor in overall array performance.

Larger drive arrays can have multiple drives sharing the bandwidth of a single SAS channel. When more than two drives share a single 3 GiB/s SAS channel, the performance for sequential operations will start to be limited by the bandwidth of the SAS-2 physical link. With 6 GiB/s SAS drives attached to 6 GiB/s SAS-2 channels on the newer Smart Array controllers, sequential performance should continue to scale until more than three drives are sharing each channel.

Table 1 lists typical performance numbers; other devices may have slight variations in performance.

Table 1. Maximum sustained throughput and random IOPS capabilities for HP HDD drives Sustained rate can be sensitive to the drive model (even with the same rpm). For example, the newer 15K 6 GiB SSF drives support approximately 190 MiB/s.

Drive RPM	**Form Factor and Interface**	**Maximum throughput** 64 KiB sequential read Queue depth > 4*	**Typical IOPS** 4 KB random read Queue depth = 16
15,000	SFF 6 GiB/s SAS	195 MiB/s	385
10,000	SFF 6 GiB/s SAS	180 MiB/s	285
7,200	LFF 6 GiB/s SAS	150 MiB/s	135
7,200	SFF 6 GiB/s SAS	115 MiB/s	170
7,200	LFF 6 GiB/s SATA	150 MiB/s	135
7,200	SFF (3/6) GiB/s SATA	110 MiB/s	170

*These the numbers are for operations distributed randomly across the whole disk space.

Table 2. Typical performance capabilities for HP SSD drives.

Performance class	**Form Factor and Interface**	**Random reads***	**Random writes***	**Sequential reads**	**Sequential writes**
Enterprise Performance	SFF 6GiB/s SAS	40,000 IOPS	14,500 IOPS	415 MiB/s	180 MiB/s
Enterprise Mainstream	SFF 6GiB/s SAS	46,200 IOPS	9,800 IOPS	370 MiB/s	150 MiB/s
Enterprise Mainstream	SFF 3GiB/s SATA	32,000 IOPS	6,000 IOPS	257 MiB/s	129 MiB/s

*These numbers were Block size of 4096 bytes and queue of 16. The devices were tested at full capacity and preconditioned.

RAID performance

Most RAID levels improve read performance by distributing, or “striping,” data across a set of physical drives that have been configured as a single logical drive. Striping places a pre-determined amount of data onto a different physical drive in the array on a rotating basis. A “strip” is the amount of data written to each drive. A “stripe” is one complete row of data strips across all of the drives in an array.

The strip size for an array is configurable, and can be set from 16 KiB up to 1024 KiB. In general, using a larger strip size delivers higher performance for a drive array. The Array Configuration Utility (ACU) determines the largest strip size that can be set for a given logical array based on the RAID level of the array and the number of physical drives that it contains.

Read performance

Drive arrays increase the read performance by distributing (striping) data across multiple drives. Because the data is distributed, the Smart Array controller can execute read operations in parallel. Neither the Smart Array processor nor the cache size affects the read performance greatly. The drive performance has the greatest impact on the read performance for Smart Array drive arrays.

Random read performance

Two key factors influence drive array read performance, particularly random read performance:

- Data striping
- The number of drives in the array

Data striping distributes data evenly across all the drives in an array. Data striping improves performance because the Smart Array controller executes the read requests in parallel across drives. RAID 0, RAID 5, and RAID 6 have similar overall read performance because they all use data striping.

The number of small (4 KiB to 8 KiB) random IOPS affects the random read performance.

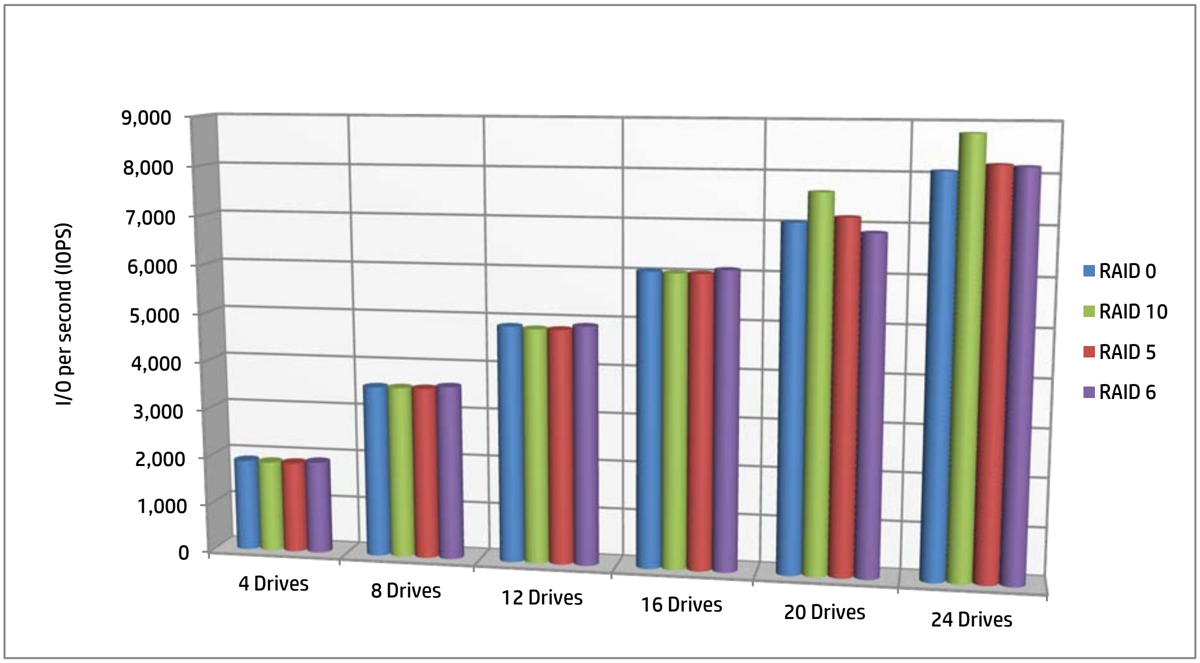
Random read performance for RAID 0, RAID 10, RAID 5, and RAID 6 scales almost directly with drive count. With all other factors being equal, a 12-drive array can deliver approximately four times random IOPS as an array with three drives.

RAID 1+0 (RAID 10) uses striping and mirroring. Its performance scales linearly with the drive count. Because it is mirrored as well as striped, RAID 1+0 requires two physical drives to achieve the same net increase in data storage capacity as a single additional drive does for RAID 0, 5, or 6.

Figure 5. Scaling of 4 KiB Random Read Performance (4 KiB aligned) 8x 6 GiB; drive limited performance

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; maximum IOPS measured for queue depth from 1 to 256

Note: RAID 5 measurements with more than 14 drives are for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).



Sequential read performance

The maximum throughput capability of a drive determines upper limit on sequential performance (see Table 1). The sequential read performance of an array tends to scale directly with the number of drives in the array.

With larger drive arrays, either the aggregate bandwidth of the SAS links or the PCIe bandwidth limit the sequential read performance. The smaller bandwidth for either component limits the performance.

RAID 1+0 performance scales more slowly. This occurs because RAID 1+0 mirrors data in addition to striping; therefore, RAID 1+0 stripes data across fewer drives. It is difficult to get sequential performance from all the drives, and therefore RAID 1+0 is a little slower.

Figure 6. Scaling of 256 KiB sequential read performance, drive limited until reaching the 4 x 6 GiB SAS bandwidth limit

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; maximum measured for queue 1 to 256

Note: RAID 5 measurements with more than 14 drives is for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).

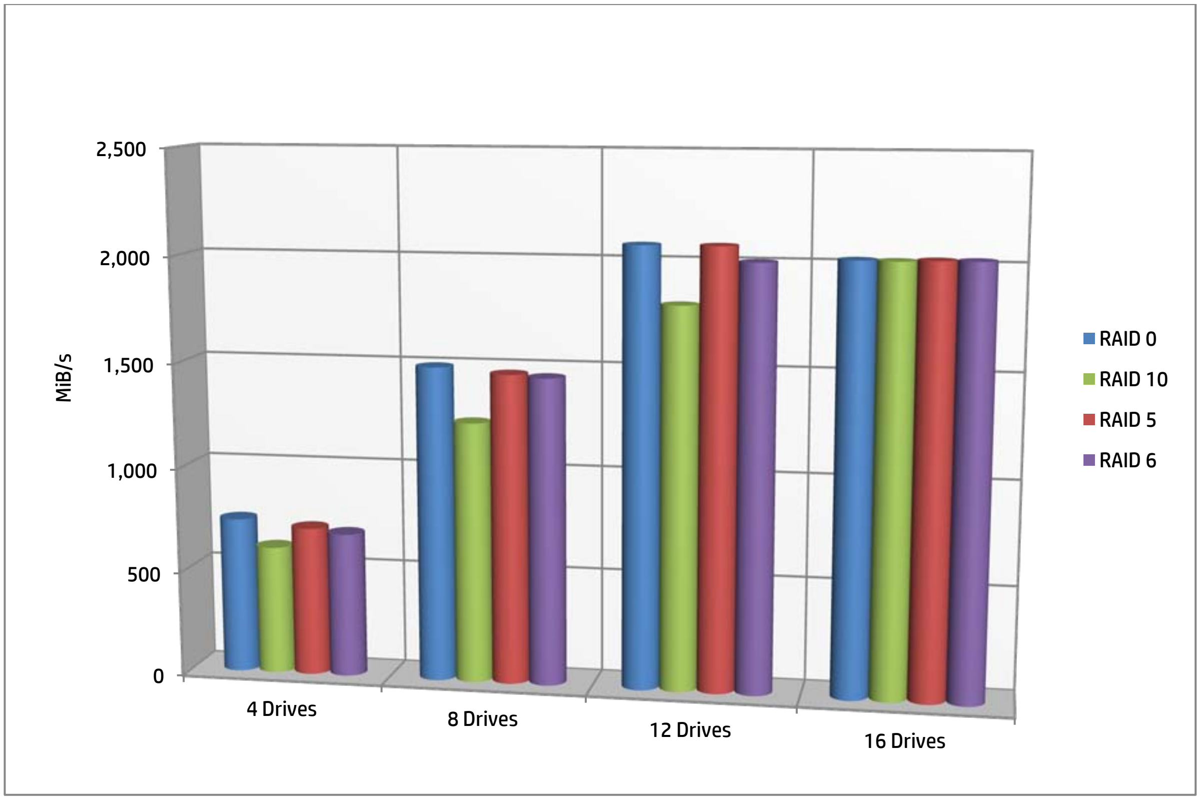
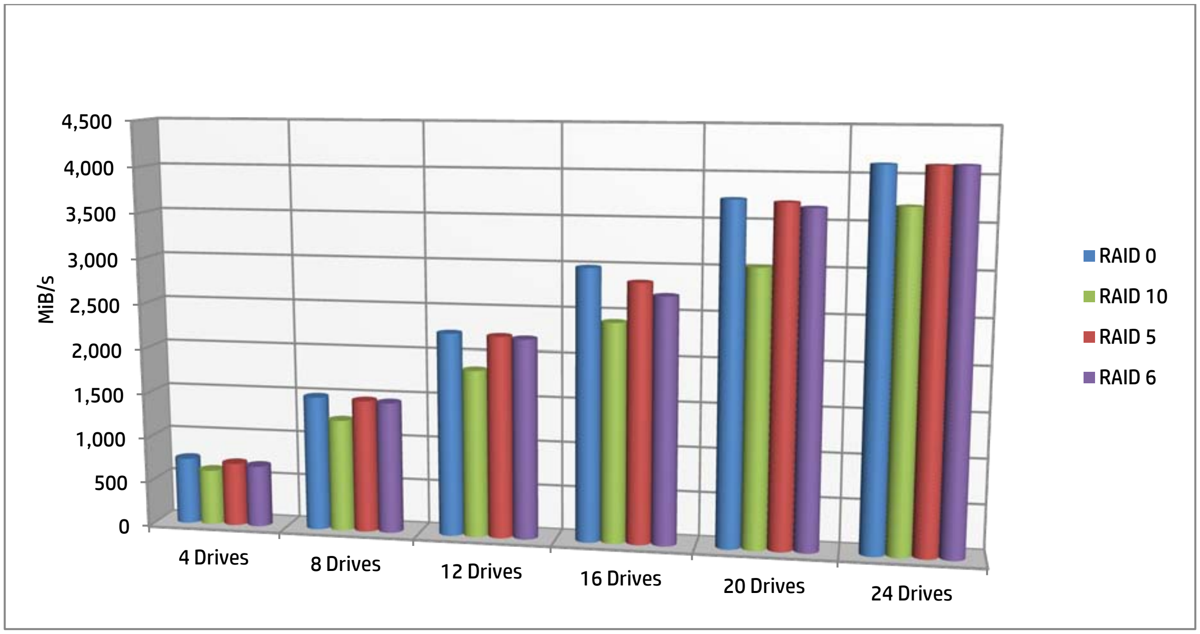


Figure 7. Scaling of 256 KiB Sequential Read Performance, drive limited until reaching the 8 x 6 GiB SAS bandwidth limit

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; maximum measured for queue 1 to 256

Note: RAID 5 measurements with more than 14 drives are for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).



RAID Write performance

With the exception of RAID 0, all RAID levels provide some level of data redundancy and recovery. In the event of a drive failure, this redundancy is necessary for rebuilding the logical drive and recovering data. A tradeoff for redundancy is the potential performance impact of the overhead involved. The Smart Array controller must execute additional low-level commands to establish redundancy. This increases the number low-level reads, writes, and calculations. The Smart Array controller must execute all of these low-level commands before performing a high-level "write" to a logical drive using RAID (with the exception of RAID 0).

Write performance for RAID 0

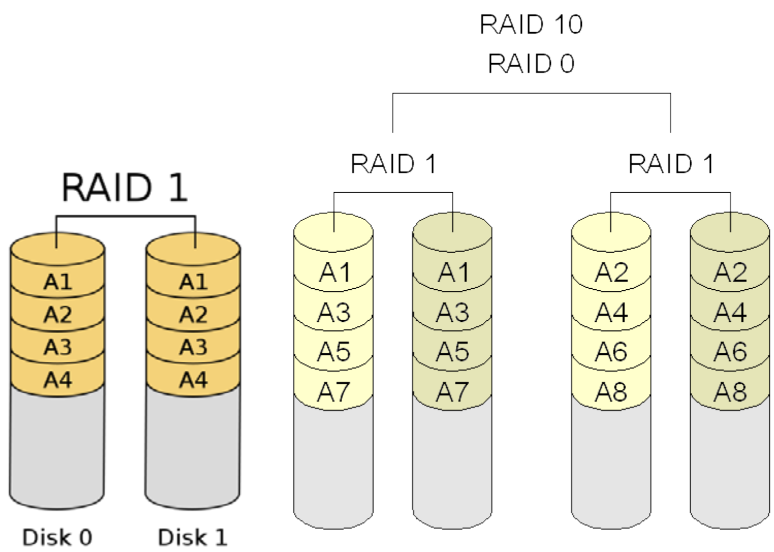
RAID 0 is the only RAID level that does not support any data redundancy. As a result, no extra low-level commands are required to execute a "write" to a logical drive. Because striping distributes data across the physical drives, the Smart Array controller can execute the low-level writes in parallel. For RAID 0, both sequential and random write performance should scale as the number of physical drives increases. RAID 0 provides a useful basis for comparison when evaluating higher RAID level performance.

Write operations for RAID 1 and RAID 1+0

RAID 1 is the simplest example of the additional write overhead associated with redundant RAID levels. RAID 1 mirrors data across a set of two drives. This means that for every "write" of a block of data to a logical drive, the Smart Array controller must execute two low-level writes, one to each of the mirrored drives. In a simple non-cached example, this would mean that in the worst-case scenario, write performance could be one-half that of writing to a non-arrayed physical drive. With RAID 1 there is no striping. This reduces the array controller's ability to execute writes in parallel across multiple physical drives, which results in lower performance than RAID 0; similar to the performance of a single drive.

RAID 1+0 mirrors data; however, it also striped the data across the mirrored drive sets. This distributes the data evenly across the drives and provides better write performance. RAID 1+0 requires executing two low-level writes for each high-level write to the logical drive.

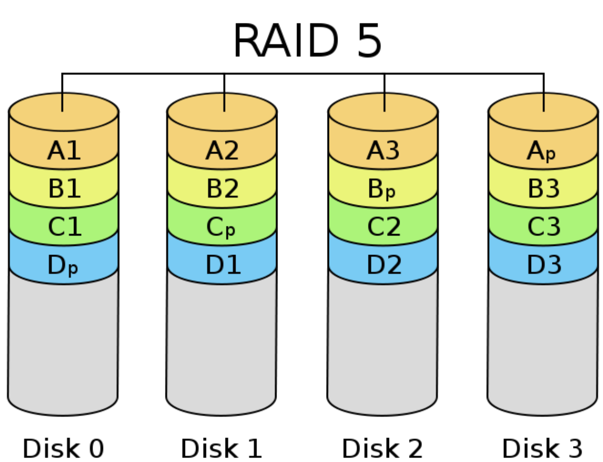
Figure 8. RAID 1 and RAID 1+0 drive arrays



Write operations for RAID 5

RAID 5 provides data protection by creating a "parity strip." If any single drive fails, the data from the other drives can be used to mathematically reconstruct the missing data. Using the data values of the data stripe, the Smart Array uses an XOR calculation to determine parity. The controller then writes the parity to one of the strips; therefore, RAID 5 requires the equivalent of one physical drive for storing the parity information. If your array has N number of drives, the array can store N - 1 drives of data. As shown in Figure 9, the position of the parity strip rotates with each stripe in order to balance overall performance.

Figure 9. Configuration of a RAID 5 drive array



With RAID 5, each high-level write operation to the logical drive takes several lower level operations. As Table 3 shows, each RAID 5 write takes four low level drive operations and a parity calculation. In the worst case, RAID 5 random write performance could be only one-quarter that of a single RAID 0 drive.

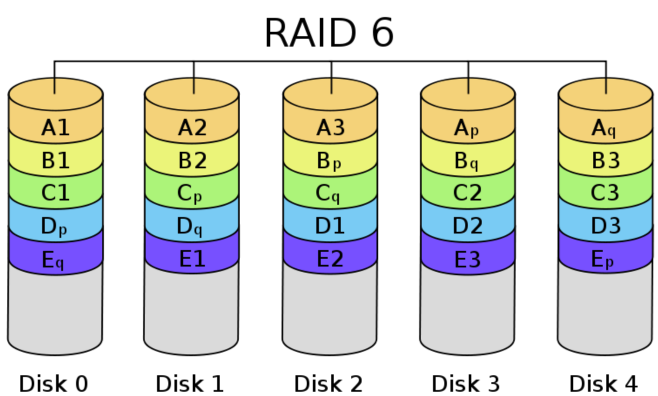
Table 3. Breakdown of a RAID 5 high-level write operation

Low-level operation	Purpose
Read data drive	Retrieve current data
Read parity drive	Retrieve current parity information
Compute new parity	Based on current data and parity plus new data
Write data drive	Write new data values to the data drive
Write parity drive	Write new parity values to parity drive

Write operations for RAID 5 and RAID 6 levels

RAID 6, also known as Advanced Data Guarding (ADG), calculates two independent forms of parity. This creates two parity strips in each data stripe across (see Figure 10). With RAID 6, an array of N drives can store N – 2 drives of data. Any two drives can fail and the data in the array can still be mathematically reconstructed.

Figure10. Configuration of a RAID 6 drive array



With RAID 6, the write performance penalty is greater than with RAID 5. Each high-level write operation to the logical drive requires executing six low-level drive read/write operations and two separate parity calculations. Random write performance for a RAID 6 logical drive could be one-sixth of an equivalent RAID 0 logical drive.

Write operations for Advance Data Mirroring; RAID 1ADM and RAID 1+0ADM

RAID 1 ADM is similar to RAID 1, in that it uses mirrored copies of each drive for data protection. ADM carries this protection one level further than RAID 1 by using two mirror copies of each drive. This means that for every "write" of a block of data to a logical drive, the Smart Array controller must execute three low-level writes, one to each of the mirrored drives. In a simple non-cached example, this would mean that in the worst-case scenario, write performance will be one third that of writing to a non-arrayed physical drive. With RAID 1 ADM there is no striping. This reduces the array controller's ability to execute writes in parallel across multiple physical drives, which results in lower performance than RAID 0; or similar to the performance of a single drive.

RAID 1+0 ADM mirrors data; however, it also striped the data across the mirrored drive sets. This distributes the data evenly across the drives and provides better write performance. RAID 1+0 ADM requires executing three low-level writes for each high-level write to the logical drive.

Smart Array processor and RAID performance

When using any of the parity-based RAID levels, both the write caching and the XOR algorithms for those processes are essential to delivering acceptable write performance for drive arrays.

The significant write performance penalty that occurs without write caching is one of the reasons that the Zero Memory versions of the Smart Array controllers only support RAID 0 and RAID 1.

Write cache allows the Smart Array controller to store pending write commands issued by the OS. The Smart Array processor then analyzes the queue of write commands to determine the most efficient write process. It does this by employing the write coalescing and command reordering techniques discussed in the section on Smart Array write cache.

If the Smart Array controller determines that a full stripe of data is changing—possibly due to write coalescing—it uses a technique known as full stripe writes. RAID 5 and RAID 6 additional read operations to retrieve the current data and parity information is not necessary. All of the information is already in the controller cache. The Smart Array controller calculates the new parity values, and then writes out the new stripe, including the parity strip(s).

Using a larger strip size for an array decreases the number of full stripe writes that the controller accumulates, and therefore may negatively affect write performance to a certain degree. This is because larger strips will naturally result in larger stripes and thus lower the probability that write coalescing will accumulate a full stripe of data in the controller cache. Larger strip sizes do tend to improve read performance.

Random write performance

Figure 10 compares the random write performance of RAID 0, RAID 5, RAID 6 and RAID 1+0 arrays (configured as one logical drive) as the number of physical drives is increased. As predicted, the write performance of RAID 5 and RAID 6 arrays is significantly lower than that of RAID 0 because of the overhead involved with each high level write operation. Performance does scale as the number of drives increases, although not at quite the rate for RAID 6 as for RAID 0.

For the same number drives, RAID 1+0 random write performance is about one half that of RAID 0. This is consistent with the fact that RAID 1+0 requires two low level drive writes for each high level array write, but does not require any extra reads or parity calculations on the part of the Smart Array controller.

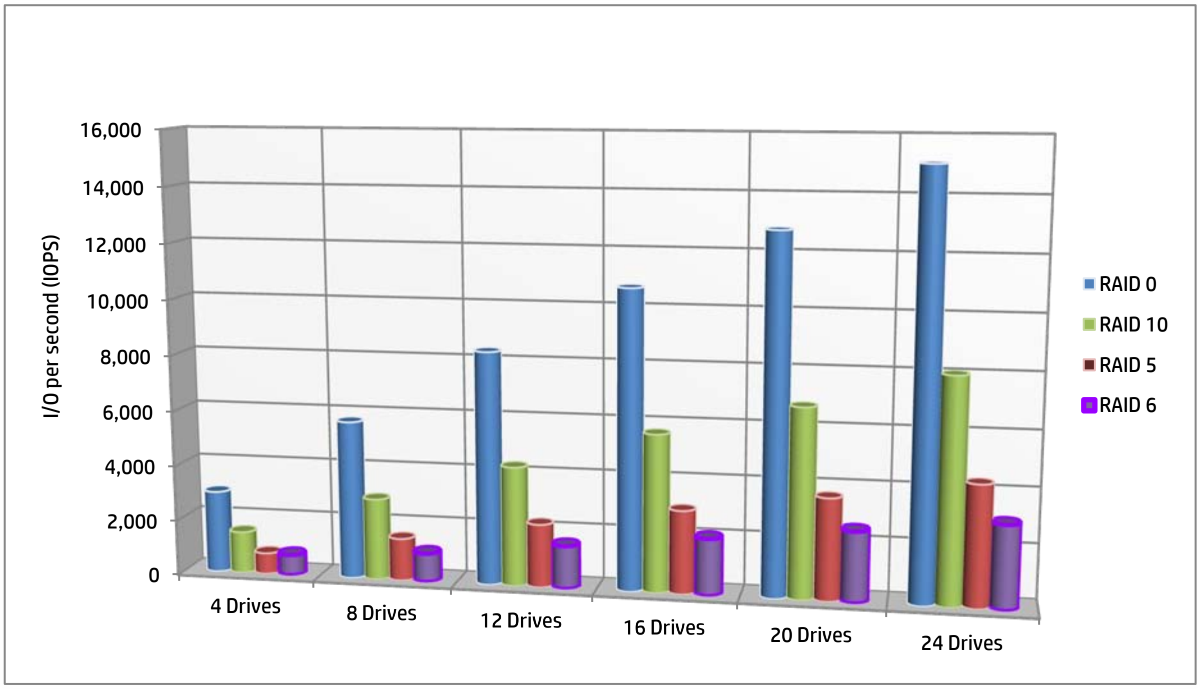
For RAID 5 random write, in addition to writing the data, new parity must be calculated and written. To accomplish this requires two disk reads, data and parity, and two disk writes, data and parity. Therefore, RAID 5 random write performance is about 25% the rate of RAID 0.

RAID 6 requires 3 disk reads and 3 disk writes, data plus two parity drives. Therefore, RAID 6 performance is approximately 1/6 or 16.7% as fast as RAID 0.

Figure 11. Scaling of 4 KiB random write performance (4KiB aligned) 8 x 6 GiB; drive limited performance

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; maximum measurement for queue depth from 1 to 256

Note: RAID 5 measurements with more than 14 drives are for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).



It is important to note that while the relative random *write* performance is impacted significantly by RAID levels than random *read* performance; the write cache does help increase random write performance overall. This is best exemplified by RAID 0, which has no write penalty. A twelve drive RAID 0 logical drive performs 8,350 random writes per second while achieving only 4,850 random reads per second. This difference is primarily attributable to the benefits of the write cache.

Sequential write performance

Figure 12 compares the write performance of the different RAID levels when executing 64 KiB sequential writes. Compared to random writes, there are two noticeable differences in the performance curves. With sequential writes, the difference in performance between RAID 0 and RAID 5 or RAID 6 is not nearly as great as it was for random writes. This can be attributed to the write cache, and more particularly to write coalescing. Sequential writes allow the Smart Array controller to coalesce them into full stripe writes. For RAID 5 and RAID 6, this eliminates the additional read operations normally required and therefore increases their performance relative to RAID 0. Secondly, sequential write performance does not tend to scale as the number of physical drives in the logical array increases past a certain point. When connection bandwidth is not the limiting factor (such as, x8 connection as shown in Figure 13) RAID 5 and RAID 6, sequential write performance plateaus when the controller processor reaches the limits of its ability to perform the required XOR computations. When the aggregated drive throughput approaches the connection bandwidth RAID 0 performance plateaus when the maximum throughput that the drives can maintain is reached.

Figure 12. Scaling of 256 KiB sequential write performance, drive limited until reaching the 4 x 6 GiB SAS bandwidth limit

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; queue fixed at 64 or maximum IOPS measured for queue depth from 1 to 256.

Note: RAID 5 measurements with more than 14 drives are for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).

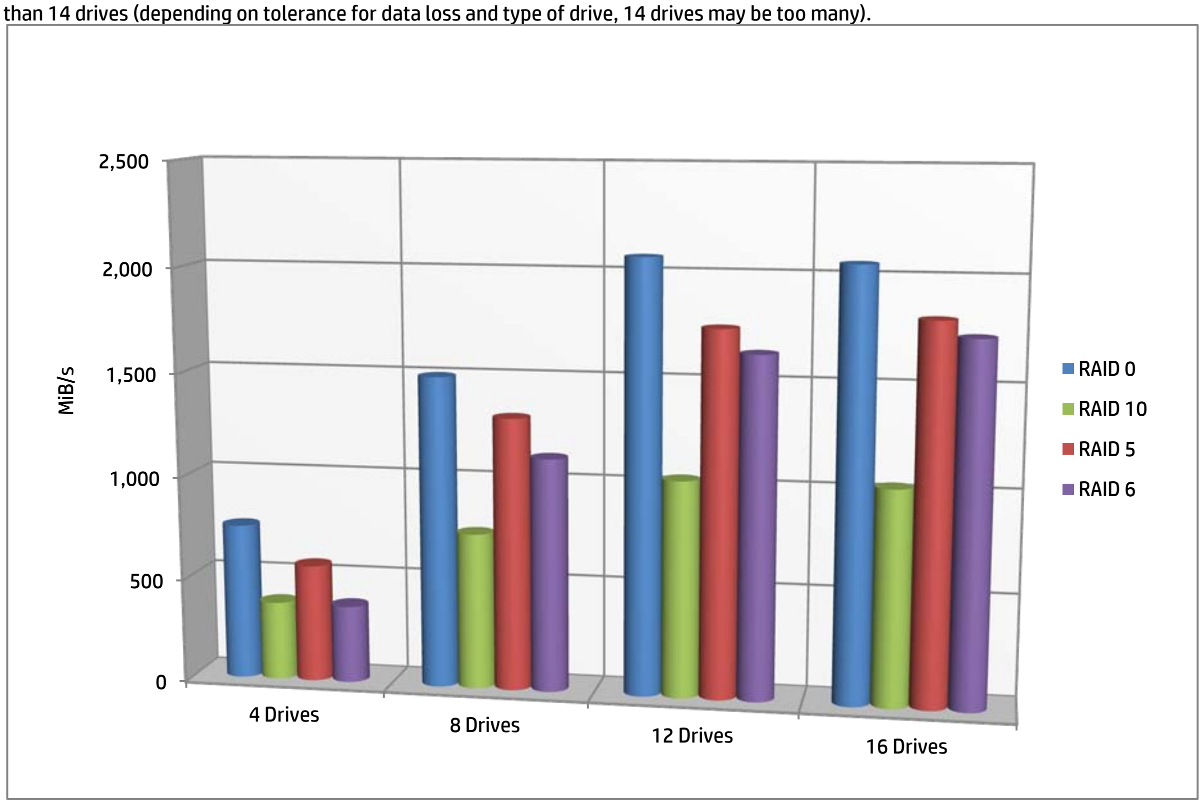
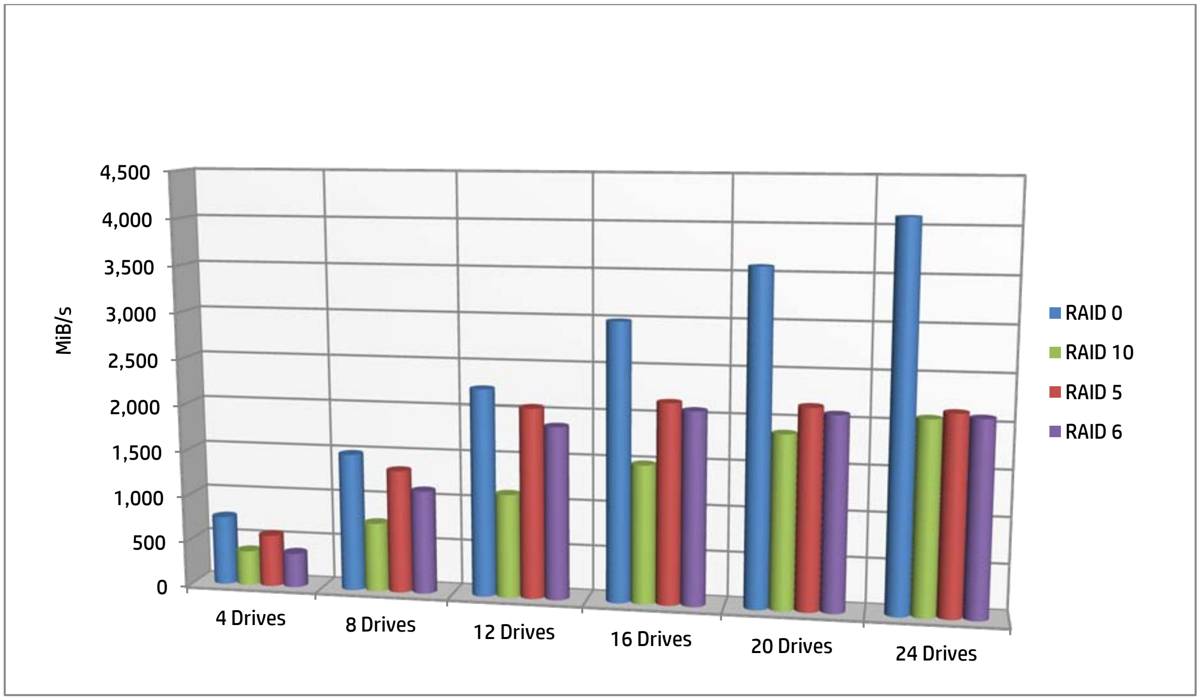


Figure 13. Scaling of 256 KiB sequential write performance; RAID 0/10, drive limited until reaching the 8 x 6 GiB SAS bandwidth limit; RAID 5/6, drive limited until reaching the controller RAID 5/6 sequential write limit

Configuration: Smart Array P421 controller, 2 GiB cache; 15K SAS drives; queue depth 64; maximum IOPS measured for queue depth from 1 to 256

Note: RAID 5 measurements with more than 14 drives are for illustrative purposes only. HP does not recommend RAID 5 arrays larger than 14 drives (depending on tolerance for data loss and type of drive, 14 drives may be too many).



Additional RAID performance characteristics

Many different terms and metrics characterize the performance of Smart Array RAID logical drives. RAID benchmarking tests often refer to queue depth, throughput, and latency. Understanding how these terms relate will help you understand them.

Queue depth

Array performance benchmarks are often run at varying queue depths. It is important to understand that in normal use, queue depth is not a configurable parameter. RAID benchmarking tests can artificially control the queue depth in order to simulate the effects of controller queue depths growing or shrinking under an application load.

In actual operating environments, a request cached by the write cache is considered completed even though necessary disk accesses have been deferred and not yet completed. The controller can analyze the commands in the queue to find more efficient ways to execute them and increase overall throughput for the Smart Array controller.

Throughput versus latency

The Smart Array controller uses various techniques to increase data throughput as queue depth increases. However, increasing queue depths are an indication that the Smart Array controller is falling behind in processing the drive commands from the OS and applications. As queue depth increases, latency—the time the OS or application sees it take to complete a drive request—tends to increase. Applications requiring lower and/or consistent latencies need environments where queue depths remain low. In general, large queue depths against the Smart Array controller can indicate a potential controller and drive I/O bottleneck. Adding more drives to the drive array may resolve the issue.

For more information

Visit the URLs listed below if you need additional information.

Resource description	Web address
HP Smart Array Controller technology – Technology Brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00687518/c00687518.pdf
Performance factors for HP ProLiant Serial Attached Storage (SAS) – Technology Brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01460725/c01460725.pdf
RAID 6 with HP Advanced Data Guarding technology – Technology Brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00386950/c00386950.pdf

Get connected

hp.com/go/getconnected

Current HP driver, support, and security alerts delivered directly to your desktop

© Copyright 2012 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

MIPS is a trademark of MIPS Technologies, Inc. in the United States and other countries. Microsoft and Windows are U.S. registered trademarks of Microsoft Corporation.

TC1205816, Created in December 2012