Best Practices for ATTO ExpressSAS® RAID Controllers

Tips, Tricks and Recommendations for a Better User Experience

Updated October 6, 2016

ATTO Technology, Inc. 155 CrossPoint Parkway Amherst, New York 14068 USA
Please visit us at www.attotech.com. For technical support, reach us at https://www.attotech.com/support/, via e-mail at techsupp@attotech.com, or call us at (716) 691-1999 ext. 242

Contents

Best Practices for ATTO ExpressSAS® RAID Controllers

Abstract

ATTO ExpressSAS® RAID controllers provide high-speed access and data protection for SAS and SATA connected RAID or JBOD storage arrays. These products are easy to install and set up and they feature several system and configuration settings that should be considered in order to obtain optimal performance in specific operating environments. This paper provides guidelines and best practices to help users better understand configuration decisions for optimal performance user experience.

Overview

ExpressSAS RAID controllers are compatible with a wide variety of third-party storage devices. Various configuration options will affect the controller's performance, useable capacity, availability and support costs. Therefore, it is important to understand the configuration options available and define according to the needs of your application(s). There are four key areas to consider: storage and connectivity; creating RAID groups; host and RAID hardware, and troubleshooting.

Let's Get Right to the Point

We know you are busy, so we'll get right to the point. Here are our top tips and best practices for using ExpressSAS RAID adapters:

- Consider doing performance tests with a single drive of the make/model you intend to use. This will provide a baseline performance number that will be useful to understand prior to creating a RAID group.
- Do not mix SAS and SATA drives within the same RAID group.
- Establish a replacement plan for your cables. Cables are typically rated for a certain number of pulls before the electrical connectivity is compromised. Evaluate your application environment and establish a plan to replace cables regularly.
- Keep spare cables on hand - always. You don't want your RAID group out of commission because of a cable.
- Include hot spares in your RAID group, especially when using SATA drives. Capacity utilization will be reduced slightly, but it reduces recovery time from drive failures, as well as ensuring that rebuild operations can automatically begin without human intervention.
- In addition to using hot spares have one or more spare drives of the same make and model standing by to minimize the time that data is at risk.
- Give each RAID group a unique and meaningful name. It's not uncommon to occasionally move RAID groups between adapters. Naming everything "RAID Group 1" only invites trouble.
- Use ATTO's Advanced Initialization feature when creating RAID groups. A media scan will be performed as part of the RAID group creation, ensuring the drive is suitable for use in a RAID group.
- Ensure that your RAID adapters have sufficient cooling airflow and that other cards aren't blocking your RAID adapter.
- Maintain a current backup of your data. RAID protection is *not* a substitute for a good backup strategy.

- Maintain a current backup of your data. RAID protection is *not* a substitute for a good backup strategy (Yeah, it's that important – it's worth saying twice.)
- Use the ATTO ConfigTool™ to configure alerts properly so that you are notified promptly of adverse events. Often alerts will allow you to take action on a small problem before it becomes catastrophic.
- Run a media scan at least weekly. This ATTO ConfigTool feature identifies and corrects media parity errors, ensuring that drives run more reliably and that data is accessible when you need it. A media scan does not consume I/O bandwidth, allowing you to find and resolve media issues before they impact your business work flow.

Want to learn more? Keep reading …

Storage Media and Connectivity

General Rules of Thumb

Drives play one of the most significant factors in determining overall system performance.

- Do not mix SAS and SATA drives within the same RAID group.
- Generally, SAS drives perform significantly faster than SATA in both throughput and IOPs. However, they usually lack the capacity of SATA drives and are more expensive.
    - When running in environments where there is significant random I/O and IOPS performance is important, select a SAS drive with as high a rotational speed (RPM) as possible.
    - For large sequential transfers, SATA drives have reasonable throughput numbers, but the IOPs performance will be slow compared to SAS drives.
    - The amount of cache on the hard drive can help increase performance as well.
- Consider doing performance tests with a single drive of the make/model you intend to use. This will provide a baseline performance number that will be useful to understand prior to creating a RAID group.
- SAS and enterprise class SATA drives are usually designed and manufactured with much higher MTBF ratings as compared to their SATA desktop model counterparts. They will simply run longer before they need to be replaced. The more robust drives are also designed with higher duty cycles in mind. The failure rate of Desktop class drives tend to increase dramatically when they are used for more than eight hours a day.
- With large-block, sequential IO applications such as data warehousing or video, disk speed has little effect when a large number of drives (>16) are connected.
- With small-block, random I/O applications such as interactive databases or mail servers, higher speed disk drives offer a substantial performance advantage.

Drive Class

ATTO strongly recommends using enterprise-class drives with ExpressSAS RAID controllers. These drives have better error-handling algorithms and work with the RAID adapter to ensure the fastest, most efficient recovery possible.

Desktop drives will work with ATTO RAID adapters; however, they are much more susceptible to failure and possible data loss. ATTO does not recommend using desktop, or consumer, grade drives with our ExpressSAS RAID adapters.

- Desktop-class drives wait longer to respond when an error occurs. This can cause the RAID controller to prematurely fail a drive instead of trying alternative methods to recover.
- Desktop-class drives do not balance I/O across the drive, resulting in time-outs and slower response times.
- Many desktop class drives tend to have more media defects out of the factory.
- Ensure that your drives are SAM-3 compliant. These drives support a more robust, optimized SCSI command set that accelerates media scan operations

Qualifying Storage Hardware

"Anything goes" is not a support strategy. Carefully selecting and qualifying drives, storage chassis, firmware versions and other key components *before* putting them into service is a highly recommended best practice to ensure all components operate properly together and that they are able to be easily supported post-sale.

ATTO maintains an interoperability list that may be useful in establishing a list of supported configurations. This list is available at https://www.attotech.com/support/interoperability/.

Cables and Connectivity

A recent study demonstrated that almost two-thirds of connectivity issues are due to cables.

- You get what you pay for. Inexpensive cable usually invite headaches down the road. Qualify and use only well-constructed cables. Insist on seeing specification sheets from your cable vendor to ensure that they meet your standards.
- Establish a replacement plan. Cables are typically rated for a certain number of pulls before the electrical connectivity is compromised. Evaluate your application environment and establish a plan to replace cables regularly.
- It's impossible to determine whether a cable is good by looking at it. Crimps, compromised shielding and loose connections are often impossible to see. When in doubt, throw it out.
- Keep spares. You don't want your RAID group out of commission because of a cable.
- Buy from a vendor you trust. ATTO qualifies and sells cables that have been tested extensively with our RAID adapters.

Creating RAID Groups

Hardware RAID can offer protection from one or more drive failures and will improve your performance. However, there are some common misconceptions:

- I have RAID protection, so I don't have to back up my data.
    - FALSE. There are failure modes that cannot be eliminated using RAID controllers. If data is valuable, it should be backed up and verified regularly.
- RAID protects my data from corruption.

- FALSE. RAID controllers cannot protect against data corruption. The file system and/or data on the RAID group can become corrupt due to a failure in a drive, application, connectivity or OS. Performing regular backup of the data is one way to protect against this type of data loss.

- More drives mean better performance.
  - SOMETIMES. With random I/O, adding more drives will increase performance. With large sequential I/O, more drives does not always mean more performance, and there is rate of diminishing returns when too many drives are used. It may sometimes be better to create multiple RAID 5 groups and stripe them together (RAID 50) or concatenate multiple RAID groups. Test a RAID configuration with anticipated applications to determine the most efficient set-up.

Good RAID Group Practices

- Include hot spares in your RAID group, especially when using SATA drives. Capacity utilization will be reduced slightly, but it reduces recovery time from drive failures, as well as ensuring that rebuild operations automatically begin without human intervention.
- In addition to using hot spares have one or more spare drives of the same make and model standing by to minimize the time that data is at risk.
- Give each RAID group a unique and meaningful name. It's not uncommon to occasionally move RAID groups between adapters. Naming everything "RAID Group 1" only invites trouble.
- Use ATTO's Advanced Initialization feature when creating RAID groups. A media scrub will be performed as part of the RAID group creation, ensuring that the drive is suitable for use in a RAID group.

Application I/O Profile

Before creating a RAID group, it's important to understand how the applications you are using move data so the RAID group can be configured for the highest possible performance. Application testing should be performed in order to determine the best workflow. There are four factors to consider when evaluating the I/O Profile:

- I/O size: Depends on the application or operating system. Some operating systems allow the user to adjust the maximum I/O size. Generally the larger the I/O size equals better throughput.
- Read and Write mix: Performance will typically decrease when mixing reads and writes compared to read-only or write-only environments. The hard drives are usually the primary factor for this performance decrease, but other contributing factors include the selected RAID interleave size, RAID level, IO size and where the data is located on the disk and PCIe slot.
- Sequential, Random or Mixed I/O: Sequential I/O always has higher throughput when compared to Random I/O. For Random I/O, performance is a factor of the number of drives attached.
- Outstanding I/O: Performance will increase if commands can remain queued. This assures the pipeline between the host and the RAID controller is always busy. Performance may begin to drop off with larger I/O sizes if the number of outstanding commands grows too large. Application testing should be performed in order to determine the best workflow.

RAID Group Availability

ATTO ExpressSAS RAID adapters are able to create the following types of RAID groups: RAID 0, 1, 4, 5, 6, 10, 50, 60, JBOD and DVRAID. Select the type that provides the protection and performance appropriate for your application:

- Know your risk tolerance.
- How important is your data? Will the RAID group store unique, irreplaceable data? How good is your backup and archiving plan? Be honest. Is the data on the RAID group simply a working copy? If you lost all of the data on your RAID group right now, would you be OK with that? Factor all of this information into selection of RAID level and backup strategy.
- How important is it to have uninterrupted access to your data? Consider RAID 6 and Adaptive Path Optimization™ for highest up time.
- RAID offers different levels of fault tolerance and performance:
    - **JBOD: Just a Bunch of Disks**
      JBOD configuration allows many individual drives to be available for normal storage operations with no special data protection by combining several drives into one large drive. A special case of a RAID Group, multiple physical drives are assigned to a JBOD RAID Group and their storage areas appear as a single spanned area of storage.
    - **DVRAID: Digital Video RAID**
      Digital Video RAID provides parity redundancy for your data. Optimized for performance for the high data transfer rates required in digital video environments, DVRAID is ATTO Technology proprietary technology which supports the editing of uncompressed High Definition (HD) video and multiple streams of real-time, uncompressed Standard Definition (SD) video.
    - **RAID Level 0: striping, no redundancy**
      RAID Level 0 (striping) is based on the fact that increased performance can be achieved by simultaneously accessing data across multiple drives, increasing data transfer rates while reducing average access time by overlapping drive seeks. Drives are accessed alternately, as if stacked one on top of the other. RAID Level 0 provides no data protection. If one drive fails, all data within that stripe set is lost.
    - **RAID Level 1: mirroring (duplicate drives)**
      RAID Level 1 ensures the security of data by writing the exact same data simultaneously to two different drives. With RAID Level 1, the host sees what it believes to be a single physical drive of a specific size: it does not know about the mirrored pair. This application is used for critical data which cannot be at risk to be lost or corrupted due to the failure of a single drive.
    - **RAID Level 1+0: striping, mirror spans two drives**
      RAID Level 1+0 increases data transfer rates while ensuring security by writing the exact same data simultaneously to two or more different drives. RAID Level 1+0 is used in applications requiring high performance and redundancy, combining the attributes of RAID Levels 1 and 0.
    - **RAID Level 4: striping, one parity drive**
      RAID 4 writes data across multiple drives or devices (striping) with parity blocks written to a single drive in the RAID Group. This increases reliability while using fewer drives than mirroring. RAID Level 4 is best suited for applications that perform mostly sequential access such as video applications. You must have at least three drives to set up RAID Level 4.

- **RAID Level 5: striping, parity distributed among drives**
  RAID Level 5 increases reliability while using fewer drives than mirroring by using parity redundancy: parity is distributed across multiple drives.
- **RAID Level 6: striping, two parity blocks distributed among drives**
  RAID Level 6 increases reliability for mission critical applications by striping both data and dual parity across multiple drives, writing data and parity blocks across all the drives in a RAID Group. RAID 6 can tolerate failure of two drives and provides redundancy during rebuilds.
- **RAID Level 50: striping, multiple parity RAID 5 axles**
  RAID Level 50 increases reliability and allows larger size RAID groups. Each RAID 5 AXLE can sustain 1 drive failure without losing data.
- **RAID Level 60: striping, multiple parity RAID 6 axles**
  RAID Level 60 is one of the most reliable data protection schemes available. Each RAID 6 axle can sustain 2 drive failures without losing data.

RAID Group Size

RAID group usable capacity is affected by the number of drives included in the RAID group, capacity of the drives used, type of RAID being used and use of hot spares.

- Do not mix drives of different capacities in the same RAID group. Drive capacity will be limited to the smallest drive within the RAID group (i.e. mixed drives of 1TB, 2TB and 4TB, etc., all drives will be treated as 1TB drives).
- RAID 0, 1, 4, 5 and 6 groups are limited to 16 members.
- RAID 10 groups are limited to 32 members
- RAID 50 and 60 groups are limited to 128 members.
- Capacity efficiency is defined as the useable capacity divided by the total capacity. The tables below define capacity efficiency for the different RAID configurations with and without a hot spare.
- Just because you can doesn't mean you should. Today's drives can exceed 10TB each. A relatively small RAID 5 group of 8 drives can hold over 70TB of data. That's a massive number – one that brings significant risk with it. In the case of a single drive failure, rebuild times can be days, possibly weeks – putting your data at risk until the RAID group is fully functional. Consider several smaller RAID groups to minimize risk of losing data.

No. of Drives	Capacity Efficiency for various RAID Configurations (no Hot Spares)								
	RAID 1 (n-way)	RAID 1 (concatenate)	RAID 10	RAID 4	DVRAID	RAID 5	RAID 6	RAID 50	RAID 60
2	50%	50%							
3	33%			67%	67%	67%			
4	25%	50%	50%	75%	75%	75%	50%		
5	20%			80%	80%	80%	60%		
6	17%	50%	50%	83%	83%	83%	67%	67%	
7	14%			86%	86%	86%	71%		
8	13%	50%	50%	88%	88%	88%	75%	75%	50%
9	11%			89%	89%	89%	78%		
10	10%	50%	50%	90%	90%	90%	80%	80%	60%
11	9%			91%	91%	91%	82%		
12	8%	50%	50%	92%	92%	92%	83%	83%	67%
13	8%			92%	92%	92%	85%		
14	7%	50%	50%	93%	93%	93%	86%	86%	71%
15	7%			93%	93%	93%	87%		
16	6%	50%	50%	94%	94%	94%	88%	88%	75%

Notes: RAID 0 efficiency is always 100%; RAID 1 with concatenation, RAID 10, RAID 50 and RAID 60 cannot have an odd number of drives

No. of Drives	Capacity Efficiency for various RAID Configurations (1 Hot Spare)								
	RAID 1 (n-way)	RAID 1 (concatenate)	RAID 10	RAID 4	DVRAID	RAID 5	RAID 6	RAID 50	RAID 60
2									
3	50%	33%							
4	33%			50%	50%	50%			
5	25%	40%	40%	60%	60%	60%	40%		
6	20%			67%	67%	67%	50%		
7	17%	43%	43%	71%	71%	71%	57%	57%	
8	14%			75%	75%	75%	63%		
9	13%	44%	44%	78%	78%	78%	67%	67%	44%
10	11%			80%	80%	80%	70%		
11	10%	45%	45%	82%	82%	82%	73%	73%	60%
12	9%			83%	83%	83%	75%		
13	8%	46%	46%	85%	85%	85%	77%	77%	62%
14	8%			86%	86%	86%	79%		
15	7%	47%	46%	87%	87%	87%	80%	80%	67%
16	6%			88%	88%	88%	81%		

Notes: With a hot spare RAID 1 with concatenation, RAID 10, RAID 50 and RAID 60 cannot have an even number of drives

Host and RAID Hardware Considerations

The host system needs to have the proper requirements for the type of application. In general, the faster the processor and the more available RAM in the host system, faster performance can be obtained. Also, various operating systems have different options for maximizing host performance.

Adaptive Path Optimization

Often users need redundancy and improved performance. This can be achieved using dual SAS domains via Adaptive Path Optimization.

- In this configuration, multiple physical connections to drives are created to eliminate pathway failure concerns (i.e. external cable failure, expander failure, RAID controller failure, or failure in a spanned JBOD configuration), as well as to distribute data movement over all available connections to maximize performance.

- Adaptive Path Optimization occurs when multiple paths to a drive are available. It automatically configures primary and alternate paths for highest levels of redundancy, as well as highest data transfer rates.
- Using event-driven algorithms, it evaluates and adaptively reconfigures data path assignments for optimal performance. No user intervention is required for this feature.
- For the cost of a cable, you get: automatic configuration, redundancy and improved performance, path matching that is RAID Group aware, automatic and balanced failover/failback capabilities, as well as guaranteed stable configurations.

Partitions

- When creating multiple partitions from a single RAID group, a section of each physical disk is used. When I/O is spread across all partitions, the I/O profile becomes randomized. The consequence is significant disk head movement which increases seek time and latency, resulting in lower performance.
- When creating a single large partition, the performance is highly dependent upon the workflow. If the multiple clients or multiple applications are working on files that spread across the disks, you will run into the same latency issues discussed above.

PCIe Slots

- The ExpressSAS RAID adapter should be in the fastest available slot (typically a x8 PCIe slot).
- Different PCIe slots, even of the same speed, may perform differently. Host chip sets control multiple PCIe slots and may be bridged into other chipsets before hitting system memory. Placing the host controller in the PCIe slot, where there is high traffic on the chipset or multiple hops to the system memory, may degrade performance of the host controller.
- Ensure that your RAID cards have sufficient cooling airflow (>150 lfm). Refer to the system specifications to ensure that enough airflow is provided. Use supplemental cooling if the system is not capable of meeting this minimum level.
- Pay attention to airflow available in a particular PCIe slot. Some PCIe slots may be located in dead spots (i.e. corners, blocked by other components) and receive insufficient airflow. Whenever possible, place your RAID adapter in a slot with the highest available airflow.
- Evaluate placement of other PCIe adapters. Some add-in cards have very high heat signatures (i.e. graphics cards, PCIe SSDs). Whenever possible, arrange PCIe adapters so that the RAID card is not near another PCIe card that gives off a lot of heat.

File System

Running benchmarks to drives with a file system installed on them will result in slower performance as compared to “raw” drives. File systems considerations to make:

- Sector alignment: With many file systems, the start of a partition is not aligned with an interleave boundary. For smaller, random I/O this misalignment can have a significant impact.
- Fragmentation: When file systems encounter heavy traffic, the files can become fragmented causing poor performance.
- Refer to file system documentation to determine how to resolve interleave alignment and fragmentation issues.

RAID Controller Variables

Interleave

- For large sequential I/O, interleave of 128K or 256K typically gives the best performance.
- For smaller random I/O, interleave of 128K or less typically provides the best performance.
- Proper interleave size depends on the number of drives attached. For parity-protected RAID groups, a good starting point for determining the best interleave is to take the average application transfer size divided by the number of drives minus one (avg. transfer size/n-1).

Number of RAID Groups

The more RAID groups that are created, the more cache memory must be shared between the RAID groups. But sometimes it is better to create multiple RAID groups and stripe them together. Testing should be performed to determine the best configuration for a given application.

Write Cache

By default, the ATTO RAID controllers are configured to automatically manage the size of the write cache to maximize the number of Write commands that can be simultaneously cached. This setting should generally offer the best overall cache performance, but applications that require more performance for reads than for writes may benefit from manually setting the RAIDSpeedWriteLimit to a number between 0 and 8.

In configurations with multiple RAID Groups, the RAID Adapter cache memory resources are divided amongst the RAID groups proportionally to the number of RAID members that they have.

For example, in a system with one two-member RAID1 group and one eight-member RAID5 group, the RAID5 group would by default be allocated 80% (8 / (8+2)) of the cache memory. The customer can adjust this balance with the RGCachePriority command, which assigns a weighting factor to each RAID Group.

By default each RAID group has a cache priority of 1, but this can be set from 0 to 10. For example, if the previously mentioned RAID1 group is not performance-critical the cache priority could be set to 0, and then the RAID5 group would be allocated close to 100% of the cache resources.

Cache Assure

- RAID adapters use onboard caching to improve performance and eliminate bottlenecks as data is being written to the storage media. ATTO's CacheAssure™ protects this cached data in the case of an unexpected system failure or power loss.
- CacheAssure technology is a maintenance-free, non-volatile cache data protection module. Where an end-to-end data protection plan is needed to protect unique, valuable data including in-flight transactions, CacheAssure provides another layer of protection above and beyond basic RAID.
- It is maintenance-free and can be purchased as an add-on option for all ExpressSAS RAID adapters.

Background Processes

Background processes such as rebuilds, RAID level migrations and media scans can all negatively impact performance. Changing the priority of the background processes can help alleviate performance problems but at the expense having these processes run longer.

Thunderbolt RAID Support

Thunderbolt™ enabled storage continues to grow as more workstations and portable computing devices add Thunderbolt connectivity to allow a variety of different storage and networking protocols to take advantage of this high-performance, flexible interface.

For optimal performance in Thunderbolt environments, ATTO offers a full portfolio of ThunderStream® RAID and ThunderLink® products that enable Thunderbolt devices to connect into traditional SAS/SATA, Fibre Channel and 10Gb Ethernet environments. All of ATTO's Thunderbolt connectivity products support daisy-chaining multiple devices, full hot-plug capabilities and on-the-fly reconfiguration as devices are added or removed from the link.

For customers using ATTO ExpressSAS host and RAID adapters in a Thunderbolt expansion chassis, we make the following suggestions:

- Active power management (i.e. sleep, hibernate, wake) must be disabled on the host machine for ExpressSAS adapters to work properly in a Thunderbolt expansion chassis.
- ExpressSAS adapters do not support Thunderbolt cable hot plug – the host system and expansion chassis must be fully powered down to connect, disconnect or reconfigure cables.
- Daisy-chaining of devices is not supported when using ExpressSAS adapters in a Thunderbolt expansion chassis.
- Some Thunderbolt expansion chassis do not automatically reboot when the host is rebooted (and vice versa). This may cause problems with synchronization of devices connected to the ExpressSAS adapter in the chassis. Users must ensure the Thunderbolt chassis and host system are always rebooted simultaneously.

If your storage environment requires any of these features, we recommend selecting a ThunderLink or ThunderStream connectivity solution to provide advanced Thunderbolt performance.

Support and Troubleshooting

Preparing for trouble is the best way to minimize the impact and quickly recover. We recommend taking following steps *before* you experience trouble:

- Maintain a current backup of your data. RAID protection is *not* a substitute for a good backup strategy.
- Use the ATTO ConfigTool to configure alerts properly so that you are notified promptly of adverse events. Often alerts will allow you to take action on a small problem before it becomes catastrophic.
- The ATTO ConfigTool has two components: the application and service daemon. Ensure that *both* are installed on the host computer. The service daemon allows ATTO to collect significantly more logging data, allowing us to provide better technical support in the unfortunate event of a RAID group failure.
- Regularly run a Media Scan. This ATTO ConfigTool feature identifies and corrects media parity errors, ensuring drives run more reliably and data is accessible when you need it. A Media Scan can be scheduled during down-time to eliminate the impact on business functions.
- ExpressSAS RAID adapters offer many different ways to deal with a RAID group becoming degraded or going offline, including a recovery mode to save data from a RAID group that is having difficulty completing a rebuild. Read the ATTO utilities manual, available for download from the ATTO website (www.attotech.com) *before* you experience a problem so you are able to take the best course of action to preserve your data in the case of a failure.
- ATTO technical support is available Monday-Friday, 8:00am to 8:00pm EST. They can be reached via telephone at (716) 691-1999 x242, or via e-mail at techsupp@attotech.com

Conclusion

There are many configuration options that alter the performance of the ATTO RAID controller. Careful consideration and/or testing per the recommendations provided should be explored to determine the best settings for your applications. Proof of concept testing is critical in assuring a RAID configuration will meet system requirements to obtain optimal performance from your ATTO RAID controller.