

JetStor RAID

Quick Questions & Answers

Revision 1.0 03/29/2006

Part 1. Installing and Configuring

- **Q.** “I powered on the raid connected to my network but the LCD shows IP address 192.168.1.100. How can I change it ?”

A. **JetStor** is shipped with DHCP option **enabled**. If network engine doesn't get any respond from DHCP server the **default** IP address is established. To change it use serial port connection to any computer with terminal emulator program like *HyperTerminal*. There is *<Ethernet Config>* option in the terminal menu. See more details in the Manual of your **JetStor** model.

- **Q.** “What is the best interface to configure my **JetStor** RAID ?”

A. JETSTOR can be configured using alphabetical terminal menu or Web browsed RAID Manager. Alphabetical terminal style menu is available through the front panel LCD & touch-control key pad as well as through serial interface and Ethernet/telnet port. RAID Manager is available through Ethernet/ http port and easier for navigation and management. There is only one Ethernet connection is allowed at a time: http OR telnet.

There are some configuration options which only available when using GUI interface like *<Alert by Mail Config>*, *<SNMP Config>* and *<NTP config>* With terminal menu there are additional management options as well but under special password protection. These options can be used under very specific circumstances and it's up to your technical support contact to decide if you really need these options available.

- **Q.** “Web browsed RAID Manager by default selects RAID 6, with a RAID set consisting of 8* 500GB drives is selected it correctly reports the size as 3.0 TB. If RAID level 5 is selected the RAID size displayed remains the same at 3.0TB when it should report 3.5TB.”

- **A.** You probably use MS Internet Explorer To fix the problem select

- **Tools**
- Internet Options
- Security
- Custom level

Go down to Scripting
Enable Active Scripting

Press OK button to update the settings. The RAID Manager should now report the size correctly when the RAID level is changed

- **Q.** “What is the recommended configuration for better data protection ?”

A. When configuring your storage system it’s always a trade-off between data protection, performance and capacity.

1. If data protection is your major priority Volume Set with **RAID level 6** should be chosen even if it imposes 2 disk capacity overhead and additional write penalty. RAID level 6 protects your data against failure of ANY two disks in the RAID set.
2. Configuring RAID Set with **Hot Spare Disks** reduces Mean Time to Repair (MTTR) thus reducing the probability of a second disk failure.
3. It’s always the best practice to setup the highest priority for disk rebuild which may sacrifice the performance but significantly reduces rebuild time. The chance for the second drive failure during disk rebuild process is higher compare to the regular operations.
4. If many Volume Sets needs to be created it’s better to place them to the different RAID Sets which reduces rebuild time in a case of disk failure and localizes disk failure affect to only one Volume Set.
5. To protect data in volatile cache from power failure Battery Backup Module should be installed if <Write Back> cache option is chosen.
6. Consider to turn off <Disk Write Cache> to protect your data from irrecoverable corruption. “Write Back Cache” implemented on hard drives enhances write performance but puts your data on risk in the event of a failure (such as power failure, hardware failure, etc.) if the data on the disk cache has not been flushed out to the disk media.

- **Q.** “Can I daisy chain several **JetStor** units connected to the same host ?”

A. If you have the model with Ultra320 SCSI host interface we don’t recommend to daisy chain them even if the SCSI standard allows you to do it. The whole configuration may turn out unstable. That’s why many storage vendors don’t even put the second SCSI connector for the channel. It’s better to use dual channel SCSI HBA or install the second adapter to the host. For the model with 2Gbit Fiber Channel host interface you can put several units in one FC LOOP connected to host or switch port . Make sure if one of the units goes down the whole FC LOOP goes down as well and needs to be reconfigured.

- **Q.** “Can I share one **JetStor** chassis between 2 or more hosts ?”

A. If you have the model with Ultra320 SCSI host interface you can share it between 2 hosts each connected to the dedicated SCSI channel. The model with Fiber Channel interfaces can be used the same way in point-to-point topology or connected to the FC switch in switched topology. In the last case the RAID can be shared between many hosts on fabric.

Generally there are two different ways of sharing storage device: sharing the capacity or sharing the data.

1. **When sharing the capacity** each host sees it's own **separate** Volume(s) on the RAID. In a case of switch topology SW Zonning or LUN masking mechanism should be used to separate RAID Volumes access from different hosts.
2. **When sharing the data** the same Volume (s) can be seen from different hosts. Special cluster or multipathing software should be installed on the hosts, fabric or special appliance to arbitrate access to the data.

- **Q.** "Does larger cache on **JetStor** controller help to increase performance ?"

A. It completely depends on the type of workloads and I/O patterns. For the best results, the caching algorithm that places data in cache should match the access pattern of the application. In many cases cache really helps to improve I/O performance but increasing total cache size **beyond** certain point not necessary leads to performance boost. **JetStor** units are shipped with default cache size which is practical for the most of workloads. See more details on this issue in the white paper "Disk Array Cache".

- **Q.** "If I install additional disks in my **JetStor** unit can I integrate them in the current Volume set without moving out the data ?"

A. Yes, you can if the new drives are the same capacity or larger than the old ones. First you need to expand the existed RAID Set migrating data to the new disks. Then the Volume Set which belongs to the already expanded RAID Set should be modified with the new size. Everything can be done without compromising your data and application downtime. See more details on this issue in the white paper "Expanding Storage Capacity".

- **Q.** "How can I convert RAID 5 Volume to RAID level 6 ?"

A. Because RAID level 6 uses twice much disk space for parity information you cannot migrate Volume Set from level 5 to level 6 without expanding the RAID Set which this Volume Set belongs to. The reverse operation from level 6 to level 5 is possible allowing the Volume Set size increasing.

- **Q.** "Is a Hot Spare Disk global for the chassis ?"

A. Yes, **all** the Hot Spare Disks are global to **all** RAID sets in **JetStor** unit if the **same capacity** disks are used. In a case of disk failure rebuild process will start to the standby disk of the same or larger capacity staying in the slot with lesser number. If there are RAID sets with disks of different capacity the larger disk can be Hot Spare for all of them.

- **Q.** “What is the difference between RAID 0+1 and RAID 10 ? What RAID level is implemented in **JetStor** controller ?”

A. Both levels are referred as multi-level, combine, hybrid or secondary with two levels of data mapping. Both provide very good overall performance by combining the speed of RAID 0 with the redundancy of RAID 1 without requiring parity calculations. Both levels impose the same capacity overhead of 50%. But there is also difference between two levels which lays in the location of the first RAID level. For RAID 0+1 or “Striped Mirrors” the first level is 0 (Striping) and second level is 1 (Mirror). For RAID level 10 (1+0 or “Mirror Stripes”) it vice versa. With minimum number of 4 disks both levels provide the same reliability but when the number of drives increases RAID 1+0 looks more robust in terms of MTDDL (Mean Time To Data Loss):

1. With RAID 1+0 if one disk fails only second disk in the same mirror became single point of failure. With RAID 0+1 all the disks in the same stripe (2, 3, ...) became SPF.
2. Rebuild process for RAID 0+1 takes more controller resources because half of the disks needs to be rebuild which involves all the disks in the array. With RAID 1+0 only the second disk in the mirror takes part in the rebuild process which also decreases the chance of second disk failure.

Some new RAID controllers like **JetStor** combine the striping and mirroring into a single operation making layout of the blocks for RAID 0+1 and RAID 10 identical. It gains all the benefits of RAID 10 regardless of the level labeling

- **Q.** “Can I have disk drives of different capacity in one **JetStor** unit ?”

A. Yes, you can but it’s better to put them in different RAID Sets. If they stay in one RAID Set the firmware uses a least-common disk size and don’t use all the available space.

- **Q.** “What is the difference in performance between RAID 6 and RAID 5 levels ?”

A. As well as RAID level 5 RAID level 6 does not have any performance penalty for read operations, but it does have an additional performance penalty on write operations due to the overhead associated with the additional parity calculations and blocks recording. **JetStor** controller ASIC based RAID 6

implementation (ADS) offloads the controller CPU from intensive parity computations which minimizes overall negative performance impact. With very intensive 100% sequential write operations RAID 6 Volume shows average 17% performance degradation compare to the identical RAID 5 Volume, with real numbers lays between 6% and 24% depending on the transferring block size.

Part 2. Maintenance and Troubleshooting

- **Q.** “What should I do if there are ‘Read error’ messages in **JetStor** event log and on the front panel LCD ?”

A. If there is “Disk N Read error” message in the eventlog it means:

- a. the raid controller couldn’t read some block from disk N,
- b. tried it again for a second time,
- c. still couldn’t read this block ,
- d. the hard disk controller remapped bad block to a spare one
- e. the controller recovered the data from parity and rewrote the block.

This doesn’t mean the disk is bad but it needs your attention. Monitor this disk if more “Read Error” messages appear in the eventlog. After 10-15 messages you can consider “preemptive swapping” of this drives with “cold spare” one before the spare blocks area is exhausted. The number came from experience and actually depends on disk model. A lot of hard remaps can affect an application causing timeouts of read operations. A lot of remaps within short period of time may indicate of serious problems with the disk. It would be helpful to check the SMART Attribute page for this disk if any value is close to threshold.

- **Q.** “One disk failed in my **JetStor** but when I removed it and put back in the same slot it was recognized without errors. Does it mean the disk is good ?”

A. The answer is No and Yes. It depends on the reason the raid controller failed the disk. Hard Disk can fail in different ways. It can be mechanical failure like read / write head failure and motor problems OR electronic failure which usually relates to problems on the controller board of the disk. In both cases the raid controller doesn’t take the failed disk back. But more often raid controller rejects the disk because of unrecoverable read/write disk errors when some block(s) cannot be accessed or remapped to the spare ones. In this case disk can be initialized if all critical SMART attribute values are larger than thresholds but as soon as the controller tries to access the bad block again it will fail the disk again.

To make sure that the disk is really good use vendor's test utility which can be downloaded from the appropriate Web site.

Sometimes raid controller fails the disk because of the timeout issue: it cannot get the response from the disk within certain period of time. It can happen because of the disk controller firmware flaw or compatibility issue with raid controller firmware. Many similar events with different disks definitely indicates the last case which is the reason to call Technical Support.

- **Q.** "What should I do if there is "H/W Monitor DRAM 1-Bit ECC" message in **JetStor** event log and on the front panel LCD?"

A. You probably have faulty DRAM module which needs to be replaced as soon as possible even though ECC correction still works. Contact Technical Support for assistance. After replacing the module run "Volume Set" check on the raid as well as file system checking utility on your host.

- **Q.** "In the web Raid Manager under "Upgrade Firmware" it says "Upgrade Raid system firmware or Boot ROM" Which one should go first ? Does it matter ? "

A. Yes, it matters. Upload "Boot ROM" first, system firmware after that, then restart the raid. If the wrong file was chosen on any first two steps the appropriate error message would be displayed.

- **Q.** "I'm going to run a Check for RAID 5 Volume and want to confirm I can do this while the system up and running without danger. "

A. *Volume Set Check* **can** be run on the production system. This utility reads all the blocks for all the stripes, recalculates the parity and compares the result with already recorded parity. It doesn't correct or rewrites the parity or data in a case if the difference is found. It just informs the administrator about Volume Set inconsistency with event log message. The administrator in his term should immediately backup the data, recreate the Volume Set and inform Technical Support about the case.

Volume Set Check minimum affects the total system performance because it runs on background with very low priority. Anyway it's better to run it out of peak hours. There is slight risk of disk failure during the checking. If the disk has developed bad blocks out of regular usable area it would be failed during the check which reads **all** the block of the disk. Checking also increases disk heads movement if there is active host load to the RAID Set.

If you cannot find the answer in this document please send your question to sergey@acnc.com and you will get the direct reply. Your question will be included in the next revision of the document.

