

# **Overclocking Data Storage Subsystems: One Approach to Variable Channel Bandwidth**

by

Paul A. Mitchell, B.A., M.S.  
Inventor and Systems Development Consultant

Version 2 -- July 29, 2010 A.D.

## **Introduction**

In recent years, the performance of personal computer systems has seen leaps and bounds by successfully overclocking CPU and memory bus speeds. By comparison, the widespread adoption of storage interface standards has limited the raw speeds of data transmission channels to relatively few options.

In particular, the current SATA "6G" standard still suffers from the [10/8 protocol](#) invented for a bygone era when 9600-baud dial-up modems were state-of-the-art. Thus, even with SATA channels now oscillating at 6.0 Gigabits per second ("6G"), there is still one start bit and one stop bit for every 8-bit byte transmitted.

This article briefly introduces the concept of variable channel frequencies by demonstrating a few of its expected benefits in one peer-to-peer computer hardware application using both internal and external data channels.

## **Quad-Channel Memory Architectures: Coming Soon!**

One of the reasons why Gigabit Ethernet (GbE) adapters remain so popular is their compatibility with legacy PCI slots. Transmitting 32 bits at 33 MHz, one legacy PCI slot supports a raw bandwidth of 1,056 Megabits per second, or just enough for one GbE card. Dividing by 8 bits per byte, a PCI slot performs at ~133 Megabytes per second *i.e.* the "133" in "ATA-133" for Parallel ATA data channels. Ramping up to 10 Gigabit Ethernet ("10GbE") then results in a raw bandwidth approaching ten times that of GbE, or ~1.33 Gigabytes ("GB") per second (+/-).

Neither of these alternatives is very exciting, when compared to the raw speeds now possible with conventional DDR3 RAM DIMMs operating in triple-channel mode. Early specifications confirmed raw bandwidths exceeding 25,000 MB per second at stock clock and latency settings. By overclocking the same RAM, assuming it supports the higher settings without any damage, raw bandwidths have been measured to exceed 30,000 MB per second (30 GBps) with triple-channel chipsets.

A truly exciting future should happen before long, when CPUs are designed to integrate quad-channel memory controllers. This development will permit computer motherboards to achieve very high memory bandwidths with only four DIMM slots, instead of six. To illustrate, by using a scaling factor of 4-to-3, then four such DIMM slots running in quad-channel mode should yield raw bandwidths approaching 40,000 MB/second (30,000 x 4/3) at stock settings. *Constant change is here to stay* (or so reads my bumper sticker).

## **Quad-Channel Designs for Data Storage**

The question then arises: how about using quad-channel designs to accelerate storage subsystems too? This approach is already happening with RAID 0 subsystems that combine four storage devices like rotating hard disk drives ("HDD"), and solid-state drives ("SSD") utilizing common variants of popular Nand Flash memory technology.

Quite apart from the controller overheads that are a necessary feature of such RAID subsystems, its adherence to popular standards means that the computer industry is "stuck" with accepting fixed channel frequencies and outdated transmission protocols like the 10/8 overhead required of all Serial ATA ("SATA") devices.

### **Variable Channel Frequencies: Also Coming Soon?**

As a result of using a little imagination, this author recently theorized what could happen if quad-channel storage subsystems were designed and implemented to remove the extra overhead in the 10/8 SATA protocol, and to permit the frequency of each data channel to be varied upwards to some practical engineering limit.

Specifically, the current SATA/6G protocol is modified to use a minimum number of error correction code ("ECC") bits on a 4,096-byte "jumbo frame" transmitted initially at 6 Gigabits per second: this should result in an effective bandwidth of ~750 MBps (6 Gbps / 8).

Also, the 6G transmission rate is increased to 8 Gigabits per second ("8G"), with a goal of balancing the aggregate bandwidth of quad channels with the use of a specialized controller that utilizes x8 or x16 PCI-Express 2.0 lanes ("PCI-E Gen2").

Thus, by oscillating at 8 Gbps, each data channel now has a raw bandwidth of  $8G / 8 = 1.0$  GB per second. In quad-channel mode, we predict four times that rate, or 4.0 GB per second. This latter number also conveniently corresponds to the raw bandwidth of x8 PCI-E Gen2 lanes *i.e.* x8 lanes @ 500 MB/second = 4.0 GB/second.

Now we are talking very fast storage subsystems, the raw bandwidth of which is in the neighborhood of measured DDR3 RAM speeds. Compare 4.0 GB/second to the maximum bandwidth of 10GbE above, which is much slower by comparison.

### **A Quad-Channel Example for High-Speed Peer-to-Peer File Transfers**

Figure 1 below illustrates the practical potential of quad-channel storage architectures that also vary the data channel frequency upwards to some higher rate that balances other engineering objectives.



PRELIMINARY DESIGN FOR HIGH-SPEED PEER-TO-PEER QUAD-CHANNEL ("QC") BACKUP PROTOCOL

- \* SATA/6G protocol modified for 4K "jumbo frames", no 10/8 protocol overhead
- \* Device drivers connect to target file system seamlessly (cf. RamDisk Plus logic)
- \* Bandwidth is superior to 10 GbE / 8 = ~1.33 GB/sec. at much lower cost
- \* Secondary SFF-8088 ports are available for standard RAID arrays -or- second "QC"
- \* Raw bandwidth of SFF-8088 "QC" cable is comparable to measured speed of DDR2 RAM
- \* Ramdisks at source and destination file systems exploit max QC bandwidth (3.0 GB/sec., 750 MB/sec. per channel)
- \* Custom controller can push channel frequency > "6G" standard e.g. "8G" / 8 = 1.0 GB/sec. per channel x 4

All Rights Reserved: Paul A. Mitchell, B.A., M.S.  
Instructor, Inventor and Systems Development Consultant

**Figure 1: A Quad-Channel Example for High-Speed Peer-to-Peer File Transfers**

In this Figure, we show a standard SFF-8088 "multi-lane" cable connecting two [Highpoint RocketRAID 2722](#) controllers installed in two separate workstations. A custom protocol eliminates the 10/8 transmission overhead by using 4K jumbo frames that replace the start and stop bits on each byte with far fewer ECC bits.

Also, the raw clock frequency of all four data channels is increased from 6G to 8G. In principle, then, each of those four data channels should be able to transit raw data at 1.0 GB/second, for a total of 4.0 GB per second *in quad-channel mode*.

By using a controller with an x8 edge connector, the latter bandwidth of 4.0 GB per second corresponds *exactly* to the one-way bandwidth of the PCI-E bus, because x8 lanes transmit data in each direction at 500 MBps per lane (*i.e.* 8 x 500 = 4,000 MB/second).

Additionally, because the [2722 controller](#) has two external multi-lane ports, Figure 1 allows for *other* physical and logical applications of that second SFF-8088 port, such as conventional RAID arrays or a second quad-channel connection for each workstation.

Assuming for now that all necessary device drivers can be implemented to support all of the capabilities described above, in addition to new controller circuits to oscillate data channels at 8G instead of 6G, we are then in a position to contemplate the possibilities which such high-speed storage brings much closer to reality.

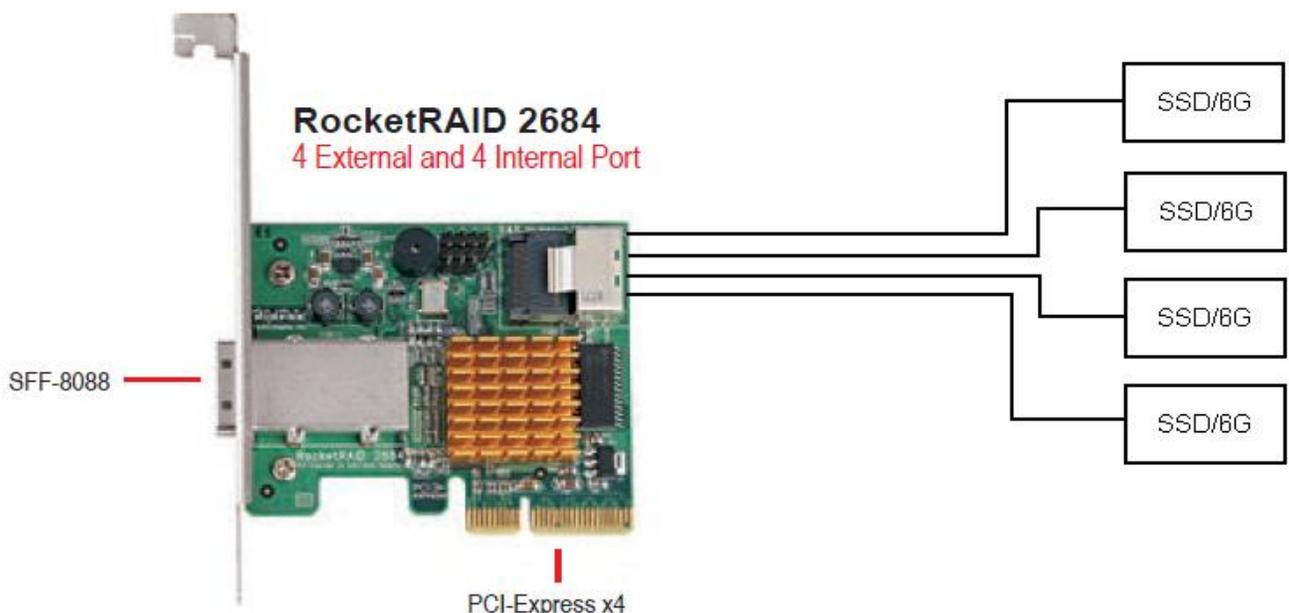
For example, with a raw bandwidth of 4.0 GB/second, system and application software can be written to perform memory-to-memory transfers, when the RAM involved in such transfers is resident in 2 *entirely different* workstations.

Similarly, storing entire file systems in ramdisks is a popular option for many computer enthusiasts at present. For example, see this author's [published review](#) of [RamDisk Plus](#) by [SuperSpeed LLC](#). Likewise, writing drive images of an OS partition will finish much faster when output rates are not hampered by the slow buffer-to-disk speeds of rotating hard drives.

And, this author can foresee a not-too-distant future in which an OS is loaded entirely into RAM, including all of its program and data files. When that capability becomes a proven reality, the quad-channel example illustrated in Figure 1 should result in significantly accelerating the speed with which routine drive images of such an OS are created.

#### **A Hybrid Example Using Internal and External "QC" Connections**

Figure 2 below illustrates a Highpoint RocketRAID 2684 controller wired to four standard 2.5" 6G SSDs. At this writing, an image of the [RocketRAID 2721](#) controller was not available. Key differences include an x4 edge connector instead of an x8 edge connector, and the 2721's data channels are capable of oscillating at 6G instead of 3G. Nevertheless, the cabling "topology" is the main point here.



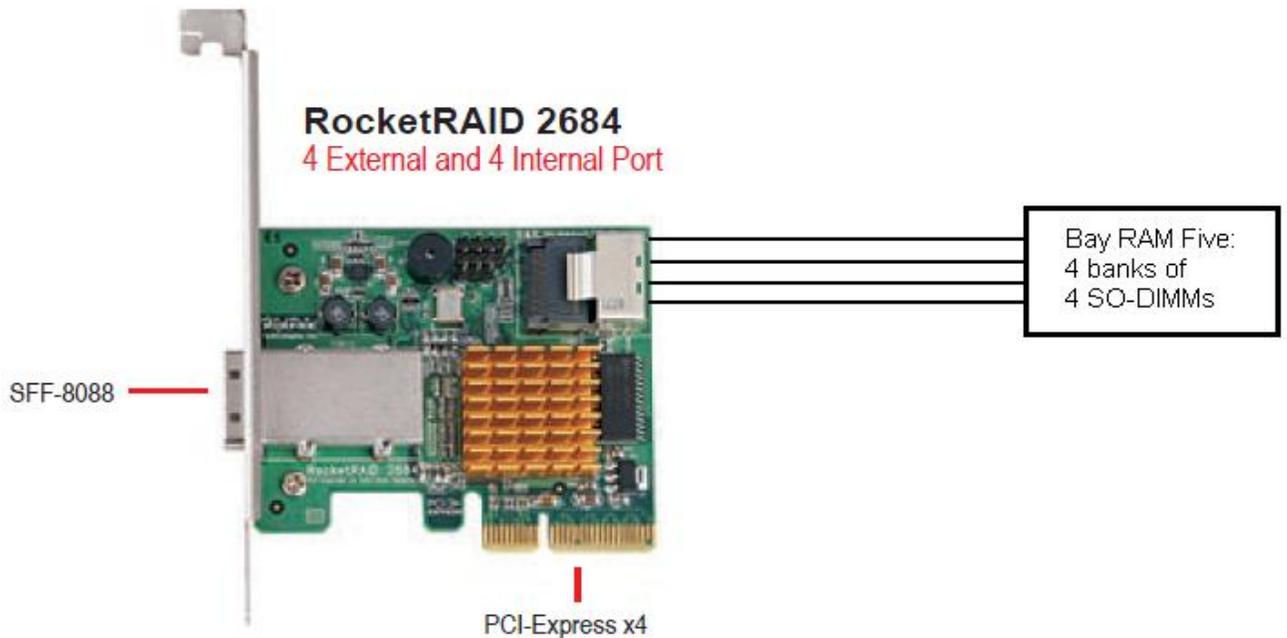
**Figure 2: A Hybrid Example Using Internal and External "QC" Connectors**

By combining the concepts above with the typical wiring layout in Figure 2, it is easy to anticipate the need for internal channels to transfer data at one speed, and for the same controller to operate its external channels at a *different* speed. A variety of methods are available to make this possible, including jumper blocks, auto detection and/or manual changes in an Option ROM launched during power-on self-test ("POST").

For purposes of Figure 2, we replace the 2684 with the [RocketRAID 2721](#), or an entirely different controller having an x16 edge connector. And, we keep the SFF-8088 cable connecting the primary workstation to the secondary workstation, as shown in Figure 1 above. The internal SFF-8087 port operates at the current standard speed of 6G, while the external SFF-8088 operates at the faster speed of 8G. If system requirements call for *concurrent* input/output to both the internal and external channels, then the analysis above would necessarily call for an x16 Gen2 edge connector, in order to supply sufficient bandwidth.

### A Hybrid Example Using Internal "QC" Connections to DDR3 SDRAM

Figure 3 below illustrates the same RocketRAID 2684 controller as shown in Figure 2 above. The major difference in Figure 3 is the use of an internal storage device invented by this author: "Bay RAM Five" is a printed circuit board for 5.25" drive bays that is designed to support four banks of four SO-DIMM sockets (16 SO-DIMMs total). Each bank of four sockets is independently controlled and connected to the host via a separate dedicated data channel -- four (4) data channels total.



**Figure 3: A Hybrid Example Using Internal "QC" Connections to DDR3 SDRAM**

Likewise for purposes of Figure 3, we replace the 2684 with the [RocketRAID 2721](#) and its x8 edge connector, or with an entirely *different* controller having an x16 edge connector. And, we keep the SFF-8088 cable connecting the primary workstation to the secondary workstation, as shown in Figure 1 above.

The major difference now is that the internal SFF-8087 port has four channels that may need to operate at a variety of *different* speeds, depending on the channel speeds of the storage device we are calling *Bay RAM Five* in Figure 3. For example, SATA history has seen standards evolving from 150 to 300 and now 600 MB/second. If such a storage device is to be backwards-compatible, it will need a jumper block or other mechanism to select one of these speeds.

Depending on the design decisions made for that storage device, each of its four data channels may all oscillate at the *same* frequency "in unison", or it is theoretically possible for each channel to oscillate at a frequency that is *different* from the frequencies of the other three channels.

Even when all four channels are designed to oscillate *in unison* for purposes of engineering simplicity, the memory technology populating its 16 SO-DIMM sockets will benefit enormously by increasing oscillation rates above the 6G standard, in order to exploit the internal bandwidth capabilities of DDR3-800, DDR3-1066, DDR3-1333 etc.

Clearly, all of the latter will easily saturate standard 6G channels, even if the 10/8 protocol is replaced with jumbo frames moving at ~750 MB/second instead of 600 MB/second. DDR3-800 alone has an internal bandwidth of 6,400 MB/second (8 bytes @ 800 MHz).

Once again, we see the need for variable channel bandwidth that is capable of exceeding the 6G standard, particularly when the storage device utilizes very fast SDRAM in its memory technology.

## **Conclusion**

An example of an overclocked quad-channel storage subsystem is presented which also modifies the standard SATA protocol with 4K jumbo frames and far fewer ECC bits. A specific combination of PCI-Express hardware and standard SFF-8088 multilane cables is briefly described for its potential to increase raw data bandwidth to 4.0 Gigabytes per second between two PCI-Express workstations. Similarly, a specific combination of PCI-Express hardware and internal SFF-8087 multilane cables is also described for the benefits expected from "tuning" channel bandwidth to the known capabilities of DDR3 SDRAM technology. Overclocking the raw transmission rates of data storage channels is an idea whose time has arrived, in order to achieve *significant* improvements in the performance of data storage subsystems, *particularly* when solid-state memory technology is utilized as the storage medium in those subsystems, instead of rotating platters.

### About the Author:

Paul A. Mitchell, B.A., M.S., is an instructor, inventor and systems development consultant, now living in Seattle, Washington State. He volunteers technical advice frequently at *Tom's Hardware* as MRFS e.g. search for site:www.tomshardware.com +"Best answer from MRFS".

### Addendum

In Figures 2 and 3 above, the Highpoint RocketRAID 2684 was used in those illustrations, because a photo of the RocketRAID 2721 was not available at the time this article was first written. Figure 4 now illustrates the RocketRAID 2721:



**Figure 4: Highpoint RocketRAID 2721 with x8 edge connector, 4 external ports (SFF-8088) and 4 internal ports (SFF-8087)**