GigaRing Disks Update

Tom Hotle, Silicon Graphics, Inc., Chippewa Falls, Wisconsin

ABSTRACT: *This paper discusses the current disk products supported on Silicon Graphics' Cray GigaRing I/O systems and recent enhancements to these products. Fibre Channel Node architecture is briefly discussed as it relates to performance and device characteristics. Finally, plans for new Fibre Channel disk products are discussed along with performance measurements and expectations.*

NOTE: *This is a public version of the paper presented at the 1998 Cray Users Group conference. Proprietary information has been removed, but can be viewed with a CRInform® password at http://crinform.cray.com/cug_papers/GigaRing_0698.html*

Current GigaRing Disk Products

Disk products that are currently supported on Silicon Graphics' Cray GigaRing I/O systems include several interfaces, disk sizes, and rotation speeds. This variety has evolved from the need to provide cost effective and high performance solutions, and from the evolution of disk technology over the time frame supported by GigaRing I/O systems. Table 1 lists the disk drives that are supported on GigaRing I/O systems, including new products and products that are supported in migration from previous generation Model E IOS systems.

Three interface types are supported. SCSI interface devices are provided to allow the attachment of cost effective devices. The Fibre Channel interface was chosen to provide high-performance connections. The IPI interface is included to allow migration of disks from previous generation Model E IOS systems.

Disk sizes in Table 1 vary from 8-inch DD-60 and DD-62 devices to 3.5-inch disks on DD-314, DD-318, DD-308, and DD-301 and DD-302. The 8-inch disk products are older devices that were introduced on Model E IOS systems as early as 1991, but many are still in service on GigaRing systems. Disk capacities vary from 1.4 GB on DD-301s to 9 GB on DD-318s and DD-308s. Disk transfer rates range from 7 MB/s on the DD-6s drives to 20 MB/s on DD-60s.

Disk Arrays

The Fibre Channel and IPI interface disk drives in Table 1 can be attached in either JBOD (Just a Bunch of Disks) or arrays. Arrays are chosen to achieve higher bandwidth and data reliability. Table 2 lists the arrays that can be configured on GigaRing systems.

The DA-308 is the new disk array shipped on GigaRing systems. The Fibre Channel interface, although relatively new,

Table 1. Cray GigaRing Disk Products

		RPM	Disk Size	Capacity	Disk BW	FCS
IPI	DD-60	3600	8”	1.9 GB	20 MB/s	1991
	DD-62	4365	8”	2.7 GB	8 MB/s	1992
	DD-301	5400	3.5”	1.4 GB	8 MB/s	1994
	DD-302	5400	3.5”	1.8 GB	9 MB/s	1995
SCSI	DD-6s	5400	5.25”	9 GB	4.5 – 7 MB/s	1995
	DD-314	7200	3.5”	4.5 GB	5 – 7.5 MB/s	1996
	DD-318	7200	3.5”	9.5 GB	7 – 12 MB/s	1997
Fibre Channel	DD-308	7200	3.5”	9.5 GB	7 – 12 MB/s	1997

Table 2. Cray GigaRing Disk Arrays

	Interface	Array Capacity	Peak Array Bandwidth (Read)	Array Configuration	Peak Node Bandwidth (Read)
DA-60	IPI	7.8 GB	80 MB/s	4+P RAID-3	80 MB/s
DA-62	IPI	10.9 GB	32 MB/s	4+P RAID-3	32 MB/s
DA-301	IPI	5.5 GB	32 MB/s	4+P RAID-3	32 MB/s
DA-302	IPI	7.2 GB	37 MB/s	4+P RAID-3	37 MB/s
DA-308	Fibre Channel	38 GB	48 MB/s	4+P RAID-3	240 MB/s
ND-40E	HIPPI	Up to 200 GB (per bank)	186 MB/s	User definable RAID-1, -5	260 MB/s (4 banks)

has become the interface of choice for high-performance disks. The SCSI interface continues to be popular within the industry, but GigaRing SCSI disk products are not available in array configurations and therefore do not offer the performance and resiliency features that are available in the Fibre Channel disk subsystem. The Fibre Channel interface and the GigaRing Fibre Channel Node (FCN) will be the primary disk interface and controller through the year 2000 on GigaRing systems.

From an external view, the FCN appears similar to the IPI interface node. However, architecturally there are several differences related to Fibre Channel and new disk technology. The key differences can be outlined as follows:

- Serial array rather than parallel array: The FCN implements the array by connecting the drives serially on one Fibre Channel loop. This is possible because the loop has much more bandwidth than each drive and therefore the data from all five drives can be transferred efficiently to the buffers on the FCN and treated as one array.
- Drive buffering: Each Fibre Channel drive has a 1 MByte buffer internal to the drive. This buffer is used to store read data and is enabled to start reading data prior to the actual sector being requested. This has the effect of hiding some of the rotational latency on read transfers.
- Drive instruction queue: Each Fibre Channel drive has an instruction queue that holds multiple commands and performs read-aheads that are optimized for performance.

The Fibre Channel interface is a serial interface that operates at approximately 1Gbit per second. Fibre Channel Arbitrated Loop protocol is used to connect many drives on the same loop and realize the bandwidth of many more drives than is possible with SCSI or IPI. The use of the serial array enables the FCN to run five arrays in parallel to achieve higher bandwidth. An good example of the advantage is comparing the GigaRing IPI RAID controller (IPN) to the FCN. The IPN with DD-60 drives can achieve 80 MB/s, whether connected to one or more arrays, but the FCN with five DA-308 arrays can achieve five times the single array bandwidth, or 240 MB/s.

Roll Out of Fibre Channel Products

The first Fibre Channel disks were shipped on GigaRing systems in April 1997. Only JBOD configurations were supported at that time. Fibre Channel RAID Software was released in a Beta release in October of 1997. This was followed by a production release in December 1997, which was SWS-ION 3.7. This first release delivered RAID functionality and the performance levels referred to in Table 3. In addition to the SWS-ION 3.7 release, mods were required in UNICOS® or UNICOS/mk to support the alternate path feature and configuration and installation. UNICOS 10.0 and UNICOS/mk 2.0.0.50 included these mods. Additional bug fixes and a new disk drive micro-code were released in April 1998, which addressed issues with drive initialization and hot swap.

One hardware enhancement has been completed since first introduction of Fibre Channel disks: the capability to connect drives via fiber optic cables at lengths up to 300 meters. This feature will be available in the third quarter of 1998.

Work on software enhancements continues. Investigations into further optimizing use of buffers and queueing of requests have shown that additional performance can be achieved. These enhancements will be included in subsequent releases of SWS/ION.

Current Performance of Arrays

The maximum measured performance levels of the Fibre Channel Node and IPI nodes are shown in Table 3 (CRInform version only). The table illustrates the point that the IPI GigaRing Node can utilize the bandwidth of a single array while the Fibre Channel Node can support the bandwidth of up to five arrays.

Table 3. Performance

Table 3 is available in the CRInform version of this paper.

New Disk Drive Introduction: DD-309

DD-309 disk drives will be introduced on Cray GigaRing systems in the third quarter of 1998. The product chosen as the follow-on to the DD-308 is the 10,000 RPM version of the 18GB disk drive (20 GB when formatted with 4K byte sectors). Table 4 shows the characteristics of the DD-309 disk drive.

DD-309s will be packaged in the DSF-2, a 3-Standard Unit 19” rack compatible chassis similar to the DSF-1 used on DD308s. Both copper and fiber optic cabling will be supported and can extend the distance from the FCN to DSF-2 to 300 meters.

In-house tests to fully characterize performance have not been completed. Preliminary single drive and array tests have been run and the performance numbers are shown in Table 5, (CRInform version only) with DD-308 numbers shown for comparison.

Conclusion

Fibre Channel disks and arrays have shown continual increases in performance since their introduction in 1997. These increases will continue with improvements in software and new disk devices at least through the year 2000. Table 6 (CRInform version only) shows the characteristics of the drives and subsystems of 1997, second half of 1998, and possible characteristics of year 2000 subsystems.

Table 4. DD-309 vs. DD-308 Disk Drive

Product	DD-308	DD-309
Capacity	9.5 GB	20 GB
RPM	7200	10,000
Transfer rate (outer zone)	12 MB/s	21 MB/s
Transfer rate (inner zone)	7 MB/s	14 MB/s
TPI	5555	8962
BPI (max)	124,000	183,000
Average seek time	9.5 ms	6.5 ms
Seek time, single track	1.1 ms	0.9 ms
Seek time, full stroke	<20 ms	<13 ms

Table 5. Measured Peak Read Transfer Rates, DD-308 vs. DD-309, Single FCN

Table 5 is available in the CRInform version of this paper.

Table 6. Fibre Channel Subsystem Possibilities Through Year 2000

Table 6 is available in the CRInform version of this paper.