SCSI-2 Terms, Concepts, & Implications Technical Brief

Abstract

With the ANSI SCSI-2 standard likely to become official in late 1992, products which implement various SCSI-2 options are becoming widely available in the mass storage market. Unfortunately, there is considerable confusion about exactly what SCSI-2 means and what the user benefits are of its features. This technical brief explains some salient features of SCSI-2, and discusses their implications.

The Need for SCSI-2

The first SCSI standard (adopted by ANSI, ISO, FIPS, and ECMA) has been highly successful, especially as a low cost interface to workstations and personal computers. However, the SCSI-1 standard has some ambiguities in its definition, and does not allow the level of performance required by today's high performance workstations and servers.

SCSI-2 Features

Besides a better defined set of required features, the SCSI-2 standard defines several optional features which can have a meaningful impact on users. This paper discusses a few of the more important options available in the SCSI-2 environment, namely fast SCSI, wide SCSI, differential SCSI, and tagged queueing.

Note that these are all *optional* features. An implementation may be fully SCSI-2 compliant yet implement *none* of these four features. In fact, all current Sun™ SCSI disk and CD-ROM products, as well as the 8mm tape products, are fully compliant with SCSI-2.

Also note that there is much more to the SCSI-2 standard than these options. However, this brief discusses these four options because they are the most commonly discussed, and most widely misunderstood, features of SCSI-2.

Fast SCSI - Higher Bus Speed

The SCSI-1 standard defined two modes of data transfer: asynchronous (handshaked) and synchronous (streamed mode). SCSI-1 synchronous transfer rates are limited to 5-Mbytes/sec. In many environments, this is acceptable. But in configurations with multiple of today's high performance devices on a SCSI bus, 5-MB/s can make the bus a bottleneck.

To alleviate this problem, the SCSI-2 standard defines an option known as "fast SCSI", in which the synchronous transfer rate is doubled to 10-MB/s. The terms "fast SCSI" and "10-MB/s SCSI" are synonymous, and are used interchangeably. The term "SCSI-2" is often incorrectly used to mean "fast SCSI". As noted above, the fast option is just that: an *option* of SCSI-2.

10-MB/s, 5-MB/s, and asynchronous devices may be freely mixed on a SCSI bus. Transfer rates are negotiated on an individual basis between the host and each device.

Fast SCSI requires the proper protocol chips in both the host adapter and device controller, as well as a modified software driver. Solaris® 1.1 Version A (SunOS™ 4.1.3) is the first Sun operating system revision to support fast SCSI. Sun's 2.1-Gbyte and 1.05-Gbyte SCSI disks announced in 1992 are both fast SCSI capable.

The SPARCstation™ 10 and the updated versions of the SPARCserver™ 600MP series are the industry's first UNIX® workstations and servers to offer fast SCSI built into the system board. Sun's FSBE/S and DSBE/S SBus cards also offer fast SCSI-2 interfaces for additional connectivity beyond the built-in SCSI ports.

Wide SCSI - Wider Data Paths

In SCSI-1, all data transfer paths are parallel and 8 bits wide. The SCSI-2 standard defines two options which widen the bus to 16 or 32 bits. Each of these options are referred to as "wide SCSI". Neither the 16 bit nor 32 bit versions of wide SCSI are common today. The implementations that do exist are almost always 16 bits wide, and most also implement the fast option, thus yielding burst transfer rates of 20-Mbytes/sec. The most common application for wide SCSI presently is in disk arrays, particularly in RAID 0 (striping) and RAID 3 where sequential throughput is a driving objective.

It is too early to tell whether wide SCSI will ever become common. Although it offers the level of bus bandwidth that will definitely be needed with future high speed I/O devices, a key disadvantage is that the connector requirements of a 16 or 32 bit wide parallel interfaces can become problematic in 3.5-inch and smaller device form factors. Given that the performance, as well as price/performance, of 3.5-inch disk drives are already approaching that of 5.25-inch drives, it may be that a serial physical interface (with much smaller connectors) will be required to meet high bandwidth requirements in the future. There are several possible high performance serial interfaces which could conceivably supplant wide SCSI.

Differential SCSI - Longer Cabling

Differential SCSI is actually not new in SCSI-2. The original SCSI standard also defined the differential interface. However, differential SCSI was very rare until recently, so most differential SCSI implementations are compliant with the SCSI-2 standard.

The SCSI standards define two types of electrical interfaces: single ended and differential. Single ended is the type Sun has offered for several years. It uses a 50-pin high density external connector. Differential SCSI uses special drivers and receivers that reference the signals to each other, rather than to ground. Sun's differential SCSI implementation uses a slightly larger, industry standard, 68-pin high density connector.

There is no performance benefit to differential SCSI. But it does allow considerably longer SCSI bus lengths than single ended. Differential SCSI busses may be up to 25 meters (82 feet) in length.

Single ended SCSI is limited to 6 meters (19.7 feet) total bus length. In fact, the SCSI-2 standard recommends that busses with fast SCSI devices be limited to 3 meters. However, with high quality shielded cables and proper active (regulated) bus termination as provided by Sun, 6 meter fast busses are quite reliable. Though the ink is not yet dry on the SCSI-2 standard, there is already a draft SCSI-3 standard which allows 6 meters for fast SCSI busses, so this aspect of Sun's SCSI implementation is SCSI-3, not SCSI-2.

In large server configurations, the cable length limit of single ended SCSI can severely restrict configurability. This is why Sun's data center SCSI disk products are differential.

All onboard SCSI ports on Sun systems are single ended, as are the FSBE/S and SBus SCSI Host Adapter cards. The DSBE/S SBus card is the only differential interface presently offered by Sun.

Tagged Queueing - Multiple Outstanding Commands

The options described above pertain to the physical interface aspects of SCSI. The other major part of the SCSI standards is the protocol. Tagged queueing is a new feature of the protocol in SCSI-2.

In SCSI-1, target controllers may accept only one command at a time from an initiator. The I/O mix of UNIX systems, in which multiple processes often issue essentially simultaneous I/O requests, can make this a severe limitation.

Tagged queueing can mitigate this performance roadblock. In tagged queueing, target controllers can accept and queue multiple commands from multiple host processes without having to wait for a pending I/O request to complete.

One benefit of tagged queueing is that it affords more efficient utilization of the SCSI bus. For example, after sending an I/O request to a disk drive, the host can send more commands to the same drive while the drive is performing the relatively slow cylinder seek and rotation operations. Immediately after the data has been read from or written to the media, the device can begin executing one of the queued commands at the same time as the data and/or status from the previous command is sent back up the bus to the host.

A second benefit results from the fact that, unless otherwise specified by the host, the device controller is allowed to execute commands in a different order than received from the host. A powerful and intelligently programmed controller (like the one in Sun's 2.1-Gbyte fast differential SCSI-2 disk drive, which includes a 16 bit 80186 processor and 256 Kbyte segmented RAM data buffer) can use this facility to improve overall performance.

The SCSI-2 standard defines the queueing protocol which is used to communicate about tagged commands between an initiator and a target. But the standard does not say anything about how the target controller should use the capability to improve performance. Thus, implementations of tagged queueing can vary from doing nothing to optimize performance, to simultaneous utilization of miltiple optimization algorithms. The performance benefits of different tagged queueing implementations vary significantly. All tagged queueing products are *not* the same.

There are a number of techniques which a device controller might use to optimize performance using tagged queueing. A common goal of all optimization algorithms is to minimize the time-costly mechanical latencies of I/O. For disk drives, tagged queueing algorithms may be designed to minimize seek times, rotational latencies, or both.

Tagged queueing requires considerable sophistication in both the device controller and the host software driver. SMCC™ Engineering works extremely closely with drive manufacturers to optimize drive tagged queueing algorithms in concert with the Solaris device drivers. Sun's 2.1-Gbyte 5.25 inch and 1.05-Gbyte 3.5 inch SCSI-2 drives are tagged queueing capable. The first Solaris release to offer tagged queueing support is 2.1.

End Results - Performance and Configurability

Differential SCSI has no performance advantages. Its sole benefit is longer cable lengths, and thus more configuration flexibility.

The other options discussed in this brief are all aimed at improved performance. Fast and wide SCSI are two methods for attacking the same issue: bus bandwidth. Tagged queueing allows intelligent device controllers to use information about data placement on the media (which only the controller knows) to optimize command execution independently of the host processor.

The increased bus bandwidth of the fast and wide options can result in significant benefits for certain applications. However, even today's fastest 5.25-inch and 3.5-inch SCSI disks have internal data rates no higher than 5-Mbytes/sec, so a single disk cannot usually saturate even a 5-MB/s bus. While it is true that bus transfers are twice as fast with 10-MB/s SCSI, bus transfer is only a small part of the total time required to service a typical I/O request. So doubling the bus speed does not double I/O performance.

For this reason, there is usually little measurable performance benefit to fast or wide SCSI in single disk configurations. The real benefits of the high bandwidth options are obtained by applications in which multiple drives on a bus are being accessed at the same time, especially in random-access patterns. Applications which perform a lot of large-block sequential I/O can also benefit significantly from the additional bandwidth, since bus time is a larger proportion of total I/O time in that situation.

The benefits of fast or wide SCSI can be significant if the other parts of the I/O subsystem are very high performance. However, fundamental disk performance characteristics like seek time, rotational latency, internal transfer rate, and SCSI command overhead are far more important than higher bus bandwidth. For example, the Sun 1.3-Gbyte SCSI-2 drive, while not fast SCSI capable, will outperform many fast SCSI third party drives. Fast SCSI has real benefits, but cannot make up for low performance in other parts of the I/O subsystem.

The nature of tagged queueing leads to its benefits being most evident when there are many outstanding I/O requests for a particular device. Since most tagged queueing algorithms in disk controllers are designed to minimize total seek times, I/O patterns which tend to cluster requests in one or more “hot spots” of cylinder groups are most likely to benefit. Filesystem inodes in UFS/NFS® environments, and indexes in database applications, tend to create hot spots.

If the disk drive controller tagged queueing firmware is capable of optimizing rotational latency by servicing multiple sequential I/O requests in a single rotation of the platters, small-block sequential I/O requests may be significantly speeded up. This can be especially beneficial in the case of so-called asynchronous writes, in which a process does not have to wait for a command complete message from the device before continuing execution. Note that Sun’s implementation of UFS performs clustering of sequential writes before SCSI commands are sent to the device, so the beneficiary of this type of drive controller optimization is raw sequential asynchronous writes. Sequential reads, both filesystem and raw, can obviously benefit from this type of rotational optimization as well, especially when the drive also performs read-aheads, as Sun’s tagged queueing drives do.

Note that tagged queueing introduces additional overhead into the SCSI command protocol, so can conceivably result in slightly lower performance if the I/O patterns do not allow optimization of command servicing and if the I/O load on the device is light.

The Bottom Line

The optional features of the SCSI-2 standard can have tangible benefits for workstation and server users. None of these features is a panacea, yet each can improve the performance or configurability of Sun systems.

As Sun has often demonstrated, great overall application performance results from balanced system designs. I/O subsystems are key components of overall system performance, and are especially important in applications which move a lot of data or which are reliant upon rapid data access.

Just as great application and system performance are derived from close attention to *all* the components of a workstation or server, so is excellent I/O performance a result of zealous pursuit of high performance in *all* aspects of I/O subsystems, from the intricacies of device drivers to the fine details of device firmware. SMCC's advanced implementations of higher SCSI bus bandwidth, tagged queueing, and differential SCSI are key reasons for selecting Sun workstations, servers, and peripherals.

Sun Microsystems, Inc.
2550 Garcia Avenue
Mountain View, CA 94043
415 960-1300
FAX 415 969-9131

For U.S. Sales Office locations, call:
800 821-4643
In California:
800 821-4642

Australia: (02) 413 2666
Belgium: 32-2-759 5925
Canada: 416 477-6745
Finland: 358-0-502 27 00
France: (1) 30 67 50 00
Germany: (0) 89-46 00 8-0
Hong Kong: 852 802 4188
Italy: 039 60551
Japan: (03) 221-7021
Korea: 822-563-8700
Latin America: 415 688-9464
The Netherlands: 033 501234
New Zealand: (04) 499 2344
Nordic Countries: +46 (0) 8 623 90 00
PRC: 861-831-5568
Singapore: 224 3388
Spain: (91) 5551648
Switzerland: (1) 825 71 11
Taiwan: 2-514-0567
UK: 0276 20444

Elsewhere in the world,
call Corporate Headquarters:
415 960-1300
Intercontinental Sales: 415 688-9000