T10/99-160r0

To: T10 Technical Committee
From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
                 Gerry Houlder, Seagate Technology (Gerry_Houlder@notes.seagate.com)
Date: 1 April 1999
Subject: Arbitration Fairness Control (in SPC-2/SPI-3/FCP-2)

SPI-2 and SPI-3 (spi3r03) include Annex C "SCSI Bus Fairness" which defines an optional arbitration fairness algorithm. To date, we have not found any targets implementing the arbitration fairness algorithm. It has not been proven in actual systems. However, targets are starting to implement it, and the SPI-3 QAS algorithm requires it. Unlike most SPI features, there is no standard way to identify targets which support the fairness algorithm, or to enable or disable the fairness algorithm in those that do.

We propose controlling the algorithm with the Disconnect-Reconnect mode page (2h) by making these changes:

- Redefine the three FA bits as a ARBITRATION FAIRNESS field, to be defined by the protocol standards
- In FCP-2, define the field as three bits: FARd, FAWrt, FASel (same as before)
- In SPI-3, for normal arbitration, define 000b as unfair and 001b – 111b as fair
- In SPI-3, for quick arbitration, define the field as "drop down to normal arbitration every Nth time the device wins quick arbitration." This will help QAS performance when non-QAS devices are not expected to be present.

<u>Existing text</u>
SPC (spc-r11a) and SPC-2 (spc2r08) (section 8.3.5) define three bits in the Disconnect-Reconnect page (mode page 2h) to control arbitration fairness in Fibre Channel (SPC-2 PDF page 153):

> "The Fair Arbitration Read (FARD), Fair Arbitration Write (FAWRT), and Fair Arbitration Status (FASTAT) bits indicate whether the target should use fair or unfair (e.g., priority) arbitration when requesting an interconnect tenancy for a read data transfer, write data transfer, and status or control message transfer respectively. An FA bit of one indicates that the target should use fair arbitration. An FA bit of zero indicates that the target should use unfair (e.g., priority) arbitration."

FCP did not refer to these bits. In the first draft of FCP-2, fcp2r00, section 8.1.1 "Disconnect-reconnect page" contains this text (PDF page 57):

> "The FARd, FAWrt, and FAStat bits indicate whether a target in a loop configuration shall use the access fairness algorithm when beginning the interconnect tenancy. An FA bit of one indicates that the target shall use the access fairness algorithm. An FA bit of zero indicates that the target may choose to not use the access fairness algorithm. The FARd bit controls arbitration when the target wishes to send one or more FCP_DATA frames to an initiator. The FAWrt bit controls arbitration when the target wishes to send one or more FCP_XFER_RDY frames to an initiator. The FAStat bit controls arbitration when the target wishes to send one or more FCP_RSP frames to an initiator or FCP_CMND frames to another target. If the target intends to send multiple frame types, it may choose to not use the access fairness algorithm if any applicable FA bit is zero. Devices attached to links that do not have the concept of link tenancy shall ignore the FA bits. The FA bits are optional for all FCP-2 devices."

SPI-3 (spi3r03) prohibits using any bits in this page that it does not mention in section 11.8 "Use of disconnect-reconnect page parameters" (PDF page 181):

> "SCSI parallel devices shall only use disconnect-reconnect page parameter fields defined below."

<u>Recommended changes</u>

SPC-2 section 8.3.5:
Change the bits in table 109 to a three-bit field named "FAIR ARBITRATION".
Replace the paragraph with:
> "The FAIR ARBITRATION field may be defined by a protocol standard (e.g. SPI-3, FCP-2) to indicate when the target uses fair arbitration."

FCP-2 section 8.1.1:
Add a small table before the FA bits paragraph:
> "The FAIR ARBITRATION field shall be interpreted in this manner:

Bit	Name	Description
2	FARd	Read transfers
1	FAWrt	Write transfers
0	FAStat	Status or control transfers"

SPI-3 section 11.8:
Add this text:
> "If the FAIR ARBITRATION field is set to 000b, the target shall not use arbitration fairness during normal arbitration. If this field is set to a nonzero value, the target shall use arbitration fairness during normal arbitration. The target shall always use arbitration fairness during quick arbitration. The FAIR ARBITRATION field also indicates the maximum number of times the target wins quick arbitration before disconnecting using a BUS FREE phase, letting normal arbitration occur. This count shall be reset every time a BUS FREE occurs.

Value	Normal arbitration	Quick arbitration
		Target drops to BUS FREE after winning
000b	Unfair	1 time
001b	Fair	1 time
010b	Fair	2 times
011b	Fair	3 times
100b	Fair	4 times
101b	Fair	5 times
110b	Fair	6 times
111b	Fair	Never"

The QAS-capable target must add a counter to implement this. The target only has to count when it wins – it doesn't have to observe all the other arbitrations for this counter. An alternative would be: let the fairness register clear N times.

There may not be much value in values like 4-6 times; it may be better to reserve those for future use.

Devices using 000 in the ARBITRATION FAIRNESS field will be unfair if using normal arbitration and very generous to non-QAS devices if using quick arbitration. A QAS-only system can program all devices to 111b to avoid unnecessary BUS FREEs.

SPI-3 Section C.1:
Add this paragraph:
> "Arbitration fairness in targets is controlled with the Disconnect-reconnect mode page (see. section 11.8)."

This annex may also deserve some description of quick arbitration and how the fairness registers are shared between normal and quick arbitrations. We don't propose any text for that purpose.

SPI-3 Section 11.1.2.1 (NORMAL ARBITRATION phase):
This section may need updates, depending on how the editor's note is handled.

SPI-3 Section 11.1.2.2 (Quick arbitration method protocol):

The quick arbitration method allows a target that has quick arbitration enabled that is currently connected to an initiator that has quick arbitration enabled to transfer control of the bus to another SCSI device that has quick arbitration enabled without an intervening BUS FREE phase. SCSI devices that support the quick arbitration method shall report that capability in the INQUIRY command.

An initiator that supports the quick arbitration method shall negotiate the use of the QUICK ARBITRATION phase with each target that has indicated support of the quick arbitration method any time the data transfer agreement is in an indeterminate state, using the PPR message, in order to enable the quick arbitration method. SCSI devices that support the quick arbitration method shall ~~implement~~ support the fairness algorithm (see Annex C) in all normal arbitrations and quick arbitrations. Targets that support the quick arbitration method shall support the FAIR ARBITRATION field in the disconnect-reconnect mode page. SCSI devices shall negotiate the use of the quick arbitration method with a particular SCSI device before using quick arbitration to select or reselect that SCSI device. Also, targets shall have negotiated the use of the quick arbitration method with a particular initiator before using QA REQUEST message to disconnect from that initiator, and initiators shall have negotiated the use of the quick arbitration method with a particular target before accepting a QA REQUEST message from that target. If an initiator receives a QA REQUEST message from a target with which it that has not negotiated the use of quick arbitration, then the initiator shall assert ATN before the ACK signal is released for the QA REQUEST message, and shall report MESSAGE REJECT on the following MESSAGE OUT phase.

In a~~n~~ environment where some SCSI devices have the quick arbitration method enabled and other SCSI devices do not, it is possible for the SCSI devices that have quick arbitration enabled to prevent SCSI devices that do not have quick arbitration enabled from arbitrating for the bus. This occurs when SCSI devices that have quick arbitration enabled never go to a BUS FREE phase. To prevent this from occurring, targets with quick arbitration enabled, when disconnecting following an initial connection, should disconnect using a BUS FREE phase ~~if their fairness register is empty~~after having won the number of times indicated in the ARBITRATION FAIRNESS field.

~~If a target has quick arbitration enabled and the fairness register is not empty or if the connection is not the initial connection (i.e., a reselection) the target should use the quick arbitration method to arbitration for the bus.~~

~~In this mixed environment SCSI devices that do not have the quick arbitration method enabled should disable their fairness algorithm, and those SCSI devices should be assigned a higher priority SCSI ID than SCSI devices with the quick arbitration method enabled.~~ [This seems a bit overzealous. First, a mixed environment cannot be reliably detected. An initiator can tell if itself and all the targets support QA, but doesn't know about other initiators. Second, who controls how devices are assigned their SCSI IDs? We suggest removing this sentence]