adaptec, inc.

X3T9.2/87-54

Memo to: John Lohmeyer, ANSI X3T9.2 Chairman
             NCR Corp.
             3718 N. Rock Road
             Wichita, KS 67226

Memo from: Robert N. Snively
             Adaptec
             580 Cottonwood
             Milpitas, CA 95035

Date: March 13, 1987

Subject: Full Command Queueing Proposal for SCSI-2

Dear Mr. Lohmeyer:

The SCSI architecturally supports very complex multi-tasking
environments. The architecture already supports the queueing of
commands from different initiators against a single SCSI device
and LU. With minor additions, the SCSI-2 is capable of
supporting the full queueing of commands from the same initiator
against a single SCSI device and LU.

This document defines the extensions that are required to provide
that support.

Conceptually, each command transmitted to a target is transmitted
either with a non-synchronizing queue tag, with a synchronizing
queue tag, or without a queue tag. If the command is transmitted
to the target without a tag, it indicates that the command must
be executed in the order received from the transmitting
initiator. No time relationship between the execution of
untagged commands from other initiators is guaranteed except
through use of the reservation instructions. Only one untagged
command can be pending from a particular initiator for each LUN.
If the command is transmitted to the target with a non-
synchronizing queue tag, the command can be executed in any order
on the selected LUN with respect to other non-synchronizing
tagged commands from any initiator. If the command is
transmitted to the target with a synchronizing queue tag, the
command must be placed in the queue after all non-synchronizing
commands previously received from any initiator and before any
non-synchronizing tagged commands received later. Synchronizing

queue tagged commands are executed in the strict order received,
regardless of initiator. In a stream of tagged commands, an
untagged command serves as a synchronizing command.

Only one command with a given tag can be active against an LU
from an initiator at a time. The number of queue tag bits
assures that an unused tag can easily be found.

If the command queue is nearly full, an indication must be given
to an initiator that no more commands can be enqueued until one
or more of the already queued commands is complete. The command
queue full indication is given as a message response to any of
the queue tag messages. The command queue full indication must
be given while there is still space in the queue for operations
from all initiators that have not received a queue full
indication.

The command queue tagging protocol is chosen to be compatible
with SCSI X3.131-1986 as well as functional with the SCSI-2
standard. The command queue tagging protocol is further chosen
to allow disconnection any time after the IDENTIFY message and
the QUEUE TAG message have been transmitted.

Both deferred errors and normal errors are handled in the normal
manner. Normal errors are indicated by CHECK CONDITION status.
REQUEST SENSE commands are executed against the same queue tag to
obtain the information stored for the failing operation.
Deferred errors are normally related to a command that has long
since completed. As such, there is no attempt to point back to
the queue tag assigned to the original failing command. Since
command queueing with disconnection performs much the same
function as the Immediate bit, it is expected that few fully
queued commands will choose to use the Immediate bit. This
limits the set of errors that can become deferred errors.

This proposal is added to the end of section 6 of the SCSI-2
document. Since a previous proposal adds a section 6.5, it is
proposed that this section be added as section 6.6. The section
will include both a description of simple command queueing and
full command queueing. In addition to the inclusion of section
6, section 5 will pick up a new message in table 5-2 and in
section 5.5.2. The details of the modifications are explained
below.

Table 5-2 on page 5-14 of the 1/9/87 document will be modified in
the following manner. The line describing codes 0Dh - 7Fh shall
be replaced with five lines as follows:

0Dh	O	O	QUEUE FULL	In	
0Eh - 0Fh			Reserved Codes		
10h	O	O	NON-SYNC QUEUE TAG	In	Out
11h	O	O	SYNCHRONIZING QUEUE TAG	In	Out
12h - 7Fh			Reserved Codes		

Section 5.5.2 will have new text added between the text

2

describing BUS DEVICE RESET and the text describing IDENTIFY. The text will be as follows:

QUEUE FULL 0Dh.   If the optional full queueing function is implemented, the QUEUE FULL message in is presented to the selecting initiator when a NON-SYNCHRONIZING QUEUE TAG or a SYNCHRONIZING QUEUE TAG message out is generated and when the command queue is so full that no other commands can be accepted from that initiator for that LUN.   Aside from this extra message phase, the command is executed normally. The Initiator is obligated to send no new commands to that LUN until one or more of the previously queued commands has been complete.

NON-SYNCHRONIZING QUEUE TAG 10h.   If the optional full queueing function described in section 6.6.2 is supported and if a non-synchronizing tagged command is required, the NON-SYNC QUEUE TAG extended message is transmitted immediately after the IDENTIFY message to provide a logical identifier for that particular command or link of commands as it is being executed by the initiator and the Logical Unit.   The NON-SYNC QUEUE TAG message is a two byte message with the following structure:

        Byte 0   NON-SYNC QUEUE TAG   10h

        Byte 1   QUEUE TAG logical identifier

The use of the NON-SYNCHRONIZING QUEUE TAG implies that the command or link of commands may be queued by the target SCSI device and executed against the LU in an arbitrary order. The transmission of a command without any tagging message implies that the untagged command or link of commands must be executed in the order transmitted relative to other commands to that LUN from that initiator.   The transmission of a command with the SYNC QUEUE TAG message implies that tagged synchronizing command or link of commands must be executed in the strict order received to that LUN regardless of the initiator.

The NON-SYNC QUEUE TAG message out phase is requested by the initiator immediately after the IDENTIFY message by leaving ATN active after the time that ACK drops in response to the REQ for the IDENTIFY message.   Target devices supporting full queueing will not disconnect until after both the IDENTIFY and the NON-SYNC QUEUE TAG messages out have been received.

The NON-SYNC QUEUE TAG message in is provided immediately after the reselection IDENTIFY message in has been presented to complete the identification process of the operation being continued.   If no type of tagging message in is provided, it is assumed that the reconnection is being made on behalf of an active untagged command.

SYNCHRONIZING QUEUE TAG 11h.   If the optional full queueing function described in section 6.6.2 is supported and if a synchronizing tagged command is required, the SYNC QUEUE TAG extended message is transmitted immediately after the IDENTIFY message to provide a logical identifier for that particular command or link of commands as it is being executed by the initiator and the Logical Unit.   The SYNC QUEUE TAG message is a two byte message with the following structure:

        Byte 0   SYNC QUEUE TAG   11h

        Byte 1   QUEUE TAG logical identifier

The use of the SYNCHRONIZING QUEUE TAG implies that the command or link of commands must be queued by the target SCSI device and executed against the LU in the strict order received.   All previous NON-SYNC QUEUE TAG labeled commands must be executed before the command labeled by the SYNC QUEUE TAG.   All subsequent NON-SYNC QUEUE TAG labeled commands must be executed after the command labeled by the SYNC QUEUE TAG, regardless of initiator. The transmission of a command without any tagging message implies that the untagged command or link of commands must be executed in the order transmitted relative to other commands to that LUN from that initiator.

The SYNC QUEUE TAG message out phase is requested by the initiator immediately after the IDENTIFY message by leaving ATN active after the time that ACK drops in response to the REQ for the IDENTIFY message.   Target devices supporting full queueing will not disconnect until after both the IDENTIFY and the SYNC QUEUE TAG messages out have been received.

The SYNC QUEUE TAG message in is provided immediately after the reselection IDENTIFY message in has been presented to complete the identification process of the operation being continued.   If no type of tagging message in is provided, it is assumed that the reconnection is being made on behalf of an active untagged command.

Section 6.6 will describe the queueing capability of the SCSI with the following text:

6.6 Command Queueing

Command queueing allows more than one command to be in process under control of an LU at the same time.   Simple command queueing allows an LU to simultaneously accept commands from more than one initiator at a time.   Full command queueing allows an LU to simultaneously accept multiple commands from each initiator to each LU.   By using command queueing, a target avoids the overhead of presenting BUSY status to a command when it is actually busy processing

3

4

describing BUS DEVICE RESET and the text describing IDENTIFY. The text will be as follows:

QUEUE FULL 0Dh.   If the optional full queueing function is implemented, the QUEUE FULL message in is presented to the selecting initiator when a NON-SYNCHRONIZING QUEUE TAG or a SYNCHRONIZING QUEUE TAG message out is generated and when the command queue is so full that no other commands can be accepted from that initiator for that LUN.   Aside from this extra message phase, the command is executed normally. The Initiator is obligated to send no new commands to that LUN until one or more of the previously queued commands has been complete.

NON-SYNCHRONIZING QUEUE TAG 10h.   If the optional full queueing function described in section 6.6.2 is supported and if a non-synchronizing tagged command is required, the NON-SYNC QUEUE TAG extended message is transmitted immediately after the IDENTIFY message to provide a logical identifier for that particular command or link of commands as it is being executed by the initiator and the Logical Unit.   The NON-SYNC QUEUE TAG message is a two byte message with the following structure:

    Byte 0   NON-SYNC QUEUE TAG   10h

    Byte 1   QUEUE TAG logical identifier

The use of the NON-SYNCHRONIZING QUEUE TAG implies that the command or link of commands may be queued by the target SCSI device and executed against the LU in an arbitrary order. The transmission of a command without any tagging message implies that the untagged command or link of commands must be executed in the order transmitted relative to other commands to that LUN from that initiator.   The transmission of a command with the SYNC QUEUE TAG message implies that tagged synchronizing command or link of commands must be executed in the strict order received to that LUN regardless of the initiator.

The NON-SYNC QUEUE TAG message out phase is requested by the initiator immediately after the IDENTIFY message by leaving ATN active after the time that ACK drops in response to the REQ for the IDENTIFY message. Target devices supporting full queueing will not disconnect until after both the IDENTIFY and the NON-SYNC QUEUE TAG messages out have been received.

The NON-SYNC QUEUE TAG message in is provided immediately after the reselection IDENTIFY message in has been presented to complete the identification process of the operation being continued.   If no type of tagging message in is provided, it is assumed that the reconnection is being made on behalf of an active untagged command.

3

SYNCHRONIZING QUEUE TAG 11h.   If the optional full queueing function described in section 6.6.2 is supported and if a synchronizing tagged command is required, the SYNC QUEUE TAG extended message is transmitted immediately after the IDENTIFY message to provide a logical identifier for that particular command or link of commands as it is being executed by the initiator and the Logical Unit.   The SYNC QUEUE TAG message is a two byte message with the following structure:

    Byte 0   SYNC QUEUE TAG   11h

    Byte 1   QUEUE TAG logical identifier

The use of the SYNCHRONIZING QUEUE TAG implies that the command or link of commands must be queued by the target SCSI device and executed against the LU in the strict order received.   All previous NON-SYNC QUEUE TAG labeled commands must be executed before the command labeled by the SYNC QUEUE TAG.   All subsequent NON-SYNC QUEUE TAG labeled commands must be executed after the command labeled by the SYNC QUEUE TAG, regardless of initiator. The transmission of a command without any tagging message implies that the untagged command or link of commands must be executed in the order transmitted relative to other commands to that LUN from that initiator.

The SYNC QUEUE TAG message out phase is requested by the initiator immediately after the IDENTIFY message by leaving ATN active after the time that ACK drops in response to the REQ for the IDENTIFY message. Target devices supporting full queueing will not disconnect until after both the IDENTIFY and the SYNC QUEUE TAG messages out have been received.

The SYNC QUEUE TAG message in is provided immediately after the reselection IDENTIFY message in has been presented to complete the identification process of the operation being continued.   If no type of tagging message in is provided, it is assumed that the reconnection is being made on behalf of an active untagged command.

Section 6.6 will describe the queueing capability of the SCSI with the following text:

6.6 Command Queueing

Command queueing allows more than one command to be in process under control of an LU at the same time.   Simple command queueing allows an LU to simultaneously accept commands from more than one initiator at a time.   Full command queueing allows an LU to simultaneously accept multiple commands from each initiator to each LU.   By using command queueing, a target avoids the overhead of presenting BUSY status to a command when it is actually busy processing

4

other commands. In addition, command transmission overheads are decreased because command transmission can be overlapped with mechanical delays in the target LU.

6.6.1 Simple Command Queueing

Simple command queueing allows an LU to accept a new command from an initiator while the LU is actually processing commands for another initiator. Only one command may be active for each LU from each initiator at any time.

A new command may be accepted for the LU at any time that the SCSI bus is free whether or not another command from a different initiator is active. The target may force the command to disconnect before or after the Command Descriptor Block has been received. It is preferable for the target to accept the Command Descriptor Block so that processing may begin immediately upon termination of the active command and other previously scheduled queued commands.

The command is labeled implicitly by the known initiator and target addresses and the LUN. As long as only one command is active from each initiator, the LU can always reconnect to the correct initiator and pointer set with that information. It is the responsiblity of the initiator to assure that no more than one command is active at any time. Section 6.5.1 describes the actions taken if more than one command is activated from an initiator to the same LUN.

Simple command queueing can be supported by SCSI-2 devices or by any SCSI X3.131-1986 device meeting conformance level 2.

It is assumed that the initiator will support the required number of active subchannel functions. Each subchannel is effectively a storage area for the pointers associated with each ongoing tagged queued command.

6.6.2 Full Command Queueing

Full command queueing allows a Logical Unit to continue accepting commands until its command queue is full, regardless of how many commands may already be active from each initiator.

Two messages, NON-SYNC QUEUE TAG and SYNC QUEUE TAG are defined to allow the initiator to uniquely label each command or set of linked commands with a distinctive QUEUE TAG logical identifier immediately after the IDENTIFY message is transferred. This allows the initiator to explicitly expand the implicit labeling of the commands so that a reconnection can always be properly identified by the combination of the initiator address, the target address, the LUN, and the QUEUE TAG logical identifier. Each initiator must assure that all its outstanding QUEUE TAG

logical identifiers are unique.

If only commands labeled by NON-SYNCHRONIZING QUEUE TAGS are being transmitted, the commands may be executed in an arbitrary order selected by the target device. Commands from other initiators are also executed in an arbitrary order. The command ordering is done by the target to meet the performance and functional goals desired for that target and Logical Unit.

For a particular initiator, commands labeled by NON-SYNCHRONIZING QUEUE TAGS, commands labeled by SYNCHRONIZING QUEUE TAGS, and commands without logical identifiers may be executed to the same LU. Commands with SYNCHRONIZING QUEUE TAGS must be executed in the exact order received with respect to other SYNC QUEUE TAG labeled commands and with respect to commands without logical identifiers. In addition, all NON-SYNC QUEUE TAG labeled commands, regardless of initiator, received before a particular SYNC QUEUE TAG labeled command must be executed before that SYNC QUEUE TAG labeled command. All NON-SYNC QUEUE TAG labeled commands, regardless of initiator, received after a particular SYNC QUEUE TAG labeled command must be executed after that command.

Commands without logical identifiers must be managed according to the rules for simple queueing. Only one such command may be active at a time. Because of this restriction, such commands are always executed for each initiator in the order in which they are received from the initiator. No time relationship is determinable between activities on one initiator and activities on another except through the reservation process.

A target may elect to disallow the queueing of certain commands or execute the command without actually queueing it. In particular, the Reserve command and the Release command must not be queued unless the system provides external dead lock protection through software restrictions or a separate communication port. If the queueing of a particular command is not supported, the command may be rejected with an INVALID REQUEST sense key and an additional sense code of INVALID SEQUENCE [CODE TO BE DETERMINED]. The TEST UNIT READY, INQUIRY, and REQUEST SENSE commands are typical of those that will normally be executed immediately without actually queueing the function.

Error information from errors that occur during execution of a command using any QUEUE TAG logical identifier is recovered by execution of a REQUEST SENSE command to the same LUN using the same logical identifier value. The REQUEST SENSE command may be labeled by the initiator as synchronizing or non-synchronizing, since the QUEUE TAG logical identifier will uniquely identify the necessary sense information.

Devices not supporting either QUEUE TAG message, either because they do not support command queueing or because they meet X3.131-1986, reject the QUEUE TAG message. The command is expected to continue from that time on in the normal manner without making use of the QUEUE TAG logical identifier. Command queueing may also be switched off by the device during certain initialization periods or to control internal resource utilization by responding immediately to either type of QUEUE TAG message out with a QUEUE FULL message in. The command will continue normally except that the initiator will be warned that no new command will be accepted for that LUN until at least one command has been completed back to that initiator from that LUN.

The above text should completely define the queueing function for SCSI-2 devices. Thank you for you acceptance of this proposal.

Robert N. Snively

165

7