X3T9.2/92-135r0

Proposed draft working document
American National Standard
for Information systems

SCSI-3 Serial Bus Protocol
(SBP)

July 20, 1992

Secretariat

Computer and Business Equipment Manufacturers Association

Abstract: This document is a proposed working draft of the SCSI-3 Serial Bus Protocol. The purpose of the protocol is to define how SCSI-3 functions are transported on IEEE P1394, High Speed Serial Bus.

This is an internal working document of X3T9.2, a Task Group of Accredited Standards Committee X3. This is not a completed standard and has not been approved by Task Group X3T9.2. This document is made available for review and comment only.

[This is an extremely preliminary draft, with little attention given to formatting and making it look pretty. The goal of this draft is to see if there's enough agreement with the fundamental approach to warrant developing a "real" document.]

Point of Contact:

Edward A. Gardner
Digital Equipment Corporation
301 Rockrimmon Blvd. South
Colorado Springs, CO 80919

gardner@ssag.enet.dec.com

1 Forward

[to be supplied]

2 Scope

[to be supplied]

3 Referenced Standards and Organizations

[to be supplied]

4 Glossary and Conventions

4.1 Glossary

4.1.1 function - Any operation that one device may request of another using a SCSI-3 protocol. Functions include commands as well as certain control functions that were conveyed using the message system in SCSI-2.

Function Control Block (FCB) - A data structure that describes a SCSI-3 function.

initiator - A SCSI device that requests functions to be performed by another SCSI device (a target). SBP initiators are identified by a 16 bit field in the Function Control Block.

Logical Unit Number (LUN) - A 32 bit identifier for a logical unit.

queue tag - A value associated with an I/O process that uniquely identifies it from other I/O processes in the same target. Represented by a 64 bit field in a Function Control Block.

target - A SCSI device that performs functions requested by initiators. SBP targets are primarily identified by the 64 bit address of their function request CSR.

[other entries to be supplied]

4.2 Editorial Conventions

[to be supplied]

92-135R0

5 Overview of IEEE P1394, High Speed Serial Bus

[Include a general overview of P1394, what it is, does, functions it provides, etc. The following informally describes some of the information that will be included.]

P1394 provides an address mapped bus model. That is, SBP and other users of P1394 appear to have direct access to a 64 bit address space. The principal P1394 operations used by SBP are:

1. READ(length, P1394-address, local-buffer, P1394-bus-control)

   Attempt to read length number of bytes from P1394-address into the local-buffer. The operation may succeed, may be rejected due to the remote node or address being busy, or may fail with one of several P1394 errors.

2. WRITE(length, P1394-address, local-buffer, P1394-bus-control)

   Attempt to write length number of bytes from the local-buffer to P1394-address. The operation may succeed, may be rejected due to the remote node or address being busy, or may fail with one of several P1394 errors.

These operations entail sending a request to the remote node containing the P1394 address, and the remote node subsequently returning a response. P1394 nodes have a limited number of buffers capable of receiving requests. If all such buffers are full (e.g., internal or external congestion prevents the buffers being emptied as fast as requests arrive), the P1394 node rejects the request with a busy response. P1394 defines an optional retry protocol for requests that are rejected with a busy response. If that retry protocol is used and it fails, the outcome is one of the "several P1394 errors" referred to above. "Busy" as used in the READ and WRITE operation descriptions above and elsewhere in this document only occurs when the P1394 retry protocol is not used.

A P1394-address identifies either memory or a CSR. Memory has standard memory semantics. That is, read and write are free of side-effects and reads return the last data written to the address. The node containing memory is typically unaware (e.g., it is not interrupted) when a memory address is accessed.

As used in SBP, CSR writes have side effects but CSR reads do not. A CSR write typically places the address and data in a register or FIFO, then interrupts the node's processor. Multiple writes of the same data to the same CSR are distinguished. CSR reads do not necessarily return the data that was written to the CSR. However, a node's internal state (including the values read from CSRs) is not changed by reading a CSR. Thus a failed CSR read may be retried.

6 SBP Logical Characteristics

6.1 Target Identification

[This section will describe the algorithm for scanning a P1394 bus and locating all SBP targets. The result of locating an SBP target is the following information.]

An SBP target is identified by the following information:

1. A 64 bit P1394 address of the target's function request CSR. This is the primary identification of a target for a particular initiator. Depending on the design of the target, multiple initiators sharing a target may share a single function request CSR address or use different addresses.

2. A 64 bit P1394 address of the target's reset CSR and a 32 bit reset data pattern. Writing the reset data pattern to the reset CSR will cause a hard reset of the target. Other targets may or may not be reset as well.

6.2 Function Control Blocks (FCBs)

[There are some open questions in SAM regarding linked commands, which will ultimately affect the format of FCBs. For simplicity, this discussion will ignore linked commands until those questions are resolved. In the context of SBP, the essential issue is whether a linked command sequence re-uses the same FCB for each command or is represented by a linked list of FCBs.]

A Function Control Block describes a SCSI-3 function. Several formats are used, depending on the type of function. Functions are primarily described by the first 64 bytes of an FCB, called the primary region. An optional extension for command FCBs is used to return autosense data.

A P1394-address used to identify an FCB shall always be 64 byte aligned. That is, the low 6 bits of an FCB's P1394-address shall always be zero.

227

2

92-135R0

6.2.1 Common FCB Header -

```
msb <----------------------- 32 bits -----------------------> lsb
+-----------------------------------+-----------+------------------+
|        initiator identifier       |   flags   |  function code   |
+-----------------------------------+-----------+------------------+
|                          LUN identifier                          |
+------------------------------------------------------------------+
|                                                                  |
+---                         queue tag                          ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
+---                    completion address                      ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
/                          function code                           /
/                            dependent                             /
|                                                                  |
+------------------------------------------------------------------+
```

function code - 1 byte. Identifies the FCB format and the function to be performed. [probably the identical encoding as CAM function codes]

initiator identifier - 2 bytes. Identifies the initiator that requested this function among the initiators sharing a function request CSR address. Targets shall consider FCBs sent to the same function request CSR address and containing the same initiator identifier to be from the same initiator. FCBs sent to different CSR addresses or containing different initiator identifiers shall be considered to be from different initiators.

LUN identifier - 4 bytes. Identifies which LUN within the target is being accessed. Not present in all FCB formats.

queue tag - 8 bytes. Identifies an I/O Process.

completion address - 8 bytes. A P1394-address used to report completion of this function.

72

3

6.2.2 Command FCBs -

```
msb <----------------------- 32 bits -----------------------> lsb
+-----------------------------------+-----------+------------------+
|        initiator identifier       |   flags   |  function code   |
+-----------------------------------+-----------+------------------+
|                          LUN identifier                          |
+------------------------------------------------------------------+
|                                                                  |
+---                         queue tag                          ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
+---                    completion address                      ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
+---          reserved for handling linked commands             ---+
|         (e.g., to construct a linked list of FCBs)               |
+------------------------------------------------------------------+
|                       data transfer length                       |
+------------------------------------------------------------------+
|                                                                  |
+---                                                            ---+
|                               CDB                                |
+---                                                            ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
+---                                                            ---+
|                                                                  |
+---                 data transfer identifier                   ---+
|                                                                  |
+---                                                            ---+
|                                                                  |
+------------------------------------------------------------------+
|                                                                  |
/                            autosense                             /
/                            extension                             /
|                                                                  |
+------------------------------------------------------------------+
```

function code - 1 byte. One of several values, which encode the type of queue tag (simple, ordered, etc.). [just like CAM]

queue tag - 8 bytes. The P1394-address of this FCB, used to access the autosense extension when autosense data is stored.

data transfer length - 4 bytes. The total number of data bytes expected to be transferred in the SCSI data transfer.

CDB - 12 bytes. The SCSI-3 CDB, padded with zeros if necessary.

data transfer identifier - 16 bytes. Describes how the SCSI data transfer should be performed on the P1394 bus.

autosense extension - variable length. Used to store any autosense data returned by this command.

92-135 R0

6.3 Function Delivery

An initiator requests that a target perform a function by allocating and initializing an FCB, then writing the FCB's primary region (first 64 bytes) to the target's function request CSR.

[Note that this requires that every SBP capable node have sufficient high speed memory to send and receive 64 byte data packets. I believe this is a reasonable restriction and substantially simplifies the protocol. If I'm wrong and this is objectionable, we can define options where the initiator writes varying amounts of the FCB to the target's CSR and the target later fetches the rest with P1394 read operations. I'd prefer to avoid the complication of such options.]

If the target LUN is incapable of accepting further commands from an initiator, it returns Queue Full status for the first command it cannot accept. Subsequent commands for the same I_T_L nexus are discarded until the initiator acknowledges the queue full status.

[This is some slight hand-waving until queuing is stable and we agree on terminology. This is very similar to ACA, and may in fact be identical, but I can't tell until queuing is stable. If it is not identical to some form of ACA, then it will be defined as an SBP specific mechanism. Note that returning Queue Full status for just one command rather than every command avoids a possible deadlock condition.]

[We could also define a queue structure that was device resident until it was full and then switched to memory resident. I think that's unnecessarily complicated.]

239

6.4 Function Completion

A target reports the completion of a command by writing a completion indication to the completion address specified in the FCB. A completion indication is 16 bytes formatted as follows:

```
msb <------------------------ 32 bits --------------------------> lsb
+---------------+---------------+---------------+---------------+
|  queue avail  | autosense len |    status     | function code |
+---------------+---------------+---------------+---------------+
|                           reserved                            |
+---------------------------------------------------------------+
|                                                               |
+---                                                            |
|                           queue tag                        ---+
|                                                               |
+---------------------------------------------------------------+
```

function code - 1 byte. Indicates the type of completion indication.

[Present in case we need to invent other completion indication formats.]

status - 1 byte. The SCSI-3 status.

autosense len - 1 byte. The number of bytes of autosense data returned.

queue avail - 1 byte. The current number of commands that the initiator should be able to issue without receiving Queue Full status. This is advisory, not a guarantee; the actual number will vary depending on command completions and actions by other initiators.

[At present this is not even half-baked, my purpose is to determine if there is committee sentiment for considering some function for avoiding Queue Full status. If there is, I will develop a well thought out proposal.]

queue tag - 8 bytes. The queue tag of the FCB that has completed.

92-135R0

6.5 Data Transfers

6.5.1 Contiguous Data Transfers - A contiguous data transfer is performed when the data transfer identifier has the following format:

```
msb <----------------------- 32 bits ----------------------> lsb
+---------------------------+-----------------+---------------+
|       P1394 bus control   |      flags      | format = 0    |
+---------------------------+-----------------+---------------+
|                      data buffer length                     |
+-------------------------------------------------------------+
|                                                             |
+---              P1394-address of data buffer             ---+
|                                                             |
+-------------------------------------------------------------+
```

format - 1 byte. The value zero.

flags - 1 byte. Low order bit set if the data transfer must be in ascending address order, clear if data transfer may be arbitrarily re-ordered. Other bits reserved.

P1394 bus control - 2 bytes. Passed through to P1394 bus interface.

data buffer length - 4 bytes. The allocated length of the data buffer.

P1394-address of data buffer - 8 bytes. The base P1394-address to use for data transfers. The address field in data transfer requests is computed by adding this base address and the relative byte offset within the transfer of the data transferred by the request.

6.5.2 Scatter-Gather Data Transfers -

[Is there any sentiment for defining data transfers to/from a scatter-gather list? I think this is needed to support PCs and workstations with virtual memory --- the lower end of such systems cannot afford an adapter that remaps P1394 addresses to internal addresses. The data transfer identifier would point to a list of individual extent identifiers, each formatted identically to the contiguous data transfer identifier shown above.]

6.5.3 Isochronous Data Transfers -

[to be supplied]

240

92-135r0