DocNum: T10/01-101r0
Author: Eric Anderson
Title: Minutes of SBP-3 Working Group
          ftp://ftp.t10.org/t10/document.01/01-101r0.pdf

Minutes of the SBP-3 Working Group meeting, March 6-7, 2001.
Crown Plaza Suites Hotel, Dallas

Attendees:

Eric Anderson	Apple	ewa@apple.com
Tim Bradshaw	Iomega	bradshat@iomega.com
Firooz Farhoomand	Panasonic	firoozf@ix.netcom.com
Lance Flake	Maxtor	lance_flake@maxtor.com
John Fuller	Sony	jfuller@computer.org
Andy Green	Oxford Semiconductor	andy.green@oxsemi.com
David Hunter	Sony	david.hunter@am.sony.com
Peter Johansson	Congruent Software	PJohansson@ACM.org
Randy Lawson	Texas Instruments	rlawson@ti.com
Steve Powers	Microsoft	spowers@microsoft.com
Ron Roberts	Adaptec	ron_roberts@adaptec.com
Bill Russell	Canon	bill.russell@cis.canon.com
David Wooten	Cypress	david.wooten@IEEE.org

The following agenda was presented by Johansson. In the minutes that follow,
the start of discussion of items listed below is denoted by the index number listed
within square brackets, such as [4.1]. Note that these references do not always
appear in order, and may not signify the conclusion of discussion of a previous
agenda item.

1. Introductions and procedures
2. Call for patents
3. Informal liaison
    3.1 IEEE P1394.1 [Johansson]
    3.2 IEEE P1394.3 [Johansson]
    3.3 1394 Printer Working Group [Russell]
    3.4 1394 Trade Association [Fuller]
4. Prior action items
    4.1 Patent inquiry [Johansson]
    4.2 Modify draft in accordance with Project Proposal
        items b), e), f) and g) [Johansson]

[1] Johansson called the meeting to order and added several items of New Business to the agenda, as reflected above.

Johansson reviewed general T10 policies and procedures. In general, attendance and participation at T10 ad hoc meetings (such as this one) is open to both visitors and T10 members. When formal votes are taken, either in an ad hoc meeting or in the T10 plenary, one vote is permitted each organization, to be cast by its principal representative or designated alternative. A two-week rule is in effect: No matter may be voted on unless notice was given at least two weeks prior. Documents to be voted on must have been posted two weeks prior to the vote. The two-week rule can be waived if nobody objects. Announcements of new documents and meetings must be posted to the T10 email reflector; all other business can be conducted on the working group reflector.

Reflector information (SBP3@isg.apple.com):

  Subscribing:
  email requests@isg.apple.com w/subject "subscribe sbp3"

  Help?:
  email requests@isg.apple.com w/subject "help"

[2] The following paragraph about ANSI/T10 patent policy is copied from past T10 Plenary minutes:

A document is available from ANSI, "Procedures for the Development and Coordination of American National Standards", at no charge. This document is also on the web at http://www.ncits.org/help/ansi_sdo.html. Section 1.2.11 contains the ANSI patent policy. Amy Marasco manages patent issues for ANSI and can be contacted at amarasco@ansi.org or 212-642-4954. Gene Milligan prepared a useful "Handy dandy Technical Committee's Patents Guide", which is available at ftp://ftp.t10.org/t10/document.99/99-291r0.pdf.

[3.1] Johansson gave the informal liaison report for IEEE P1394.1, noting that the working group voted that the editor prepare IEEE P1394.1 Draft 1.0 and forward it to the IEEE for Sponsor Ballot. Johansson added that the IEEE is, at present, forming a ballot group by sending invitations to ballot to members of the MSC ballot pool.

[3.2] Johansson noted that he would lead the Ballot Review activity for IEEE P1394.3 at the request of the Chair, Greg LeClair.

[3.3] Russell and [3.4] Fuller reported no developments in the Printer Working Group or 1394 Trade Association since the previous meeting of SBP-3.

Hunter moved to approve the minutes of SBP-3 from January 24-25, 2001:

  ftp://ftp.t10.org/t10/document.01/01-060r0.pdf

Motion passed with none opposed.

[4.1] Johansson reported that one patent possibly related to SBP had been reported. The patent is believed to pertain to passwords and to belong to Maxtor, and to be available as T10 document number 97-271r0.

[4.2] Johansson noted that he had modified the SBP-3 draft in accordance with Project Proposal items b), f), and g), adding that item e) remained to be completed.

[4.3] Johansson also noted that a bibliography had been added to the SBP-3 draft.

[4.4] Anderson reported posting a request for additional bibliography entries on the SBP-3 mailing list, and that no suggestions had yet been received.

[4.5] Johansson noted that the previously discussed topic regarding adding a P1212 Revision entry for the unit directory had become a new business item that would be discussed later in the meeting.

[5] Johansson led a review of changes in the working draft of SBP-3 made since the previous meeting. This document is available here:

 ftp://ftp.t10.org/t10/drafts/sbp3/sbp3r01a.pdf

Johansson noted that the Abort Task function in Table 2 (when performed by a management ORB) is optional, and briefly explained the history thereof.

Johansson pointed out a clarification made on page 43 that the target may store status only once for each ORB that it is given.

In section 6.4.6 (FAST_START register) Johansson agreed to add a requirement that only complete (8-byte) page table entries may be included in a Fast Start packet.

Anderson noted that section 6.4.6 required the target to interpret unused bytes in the FAST_START register as zero, which in the absence of an explicit requirement to examine the length of the received Fast Start packet could cause a target to interpret some page table entries to be zero. Johansson agreed to add text requiring the target to examine the length of the Fast Start packet to determine how many page table entries it contains.

Johansson observed that the text in section 7.4.10 requires a Logical Unit that supports Fast Start to support it for all fetch agents in that Logical Unit, including those associated with streams. Nobody objected to this requirement.

Fuller commented that if a particular fetch agent within a Logical Unit did not support Fast Start, it could interpret the first eight bytes of any Fast Start packet as a write to the ORB_POINTER register and ignore the remainder of the packet. Johansson agreed provided that later study confirmed that the effects would be identical.

Lawson asked if the previously proposed use of the Login Response to indicate support for Fast Start had been dropped, and Johansson confirmed that this usage had been removed.

Fuller noted that in section 9.1.4 the rules for when writing to the FAST_START register is safe were the same as those for when writing to the ORB_POINTER is safe, which had not previously been documented to this extent. Fuller suggested that this equivalence be pointed out for clarity.

An extended discussion of the revised fetch agent state machine diagram in section 9.1.5 was held.

Anderson asked if the f0:f3 transition should clear the doorbell like the f1:f3 transition does. After discussion, the group decided that it should not, to avoid a race condition (or the appearance of a possible race condition).

Anderson suggested that state f4 ("Wait for doorbell") should be renamed, since two of its three exits did not involve the doorbell. The group agreed to rename it "Suspended" or something similar.

The group discussed whether "target resources available", as shown on transition f1:f2, should also be indicated on transition f4:f2. The group decided to add this label to the f4:f2 transition.

Fuller proposed further revision to the state diagram to explicitly show a Fast Start variable corresponding to the register. Johansson suggested that the matter be given further consideration before changes were made. Johansson proposed adding supplemental text to describe the impact of Fast Start, and Fuller agreed. Johansson noted that the new text might go in the f4:f3 transition notes.

Someone requested additional text describing Fast Start be added to section E.2 (ORB_POINTER error recovery).

[7.7] Reviewing section F.3.1, Johansson noted that changes to SAM terminology might impact SBP. Johansson suggested that SBP participants should review recent SAM changes and attend the CAP meeting later in the day. Johansson also updated Annex F accordingly.

[6.1] Johansson noted that at the last SBP-3 meeting, he was asked to propose on the IEEE 1212r reflector that a simple revision entry be defined by 1212 to enable a configuration ROM to identify support for SBP-3 vs. SBP-2. Johansson reported that Ats Nakamura of Canon had objected to this request on procedural grounds, and had noted that the same request had been made by the Direct Print Protocol group, and rejected. Johansson commented that multiple groups wanting this feature was likely evidence that it should be included in 1212 anyway. Wooten stated a preference for placing multiple software_version keys in a single SBP unit directory, but the impact of this on legacy software and firmware enumerators was unknown. Fuller expressed concern about non-computer enumerators whose firmware might not be updateable.

Johansson listed four solutions for distinguishing between SBP-2 and SBP-3:

1. Create a new generic revision key in IEEE 1212r, with sequential ordering (i.e. revision 2 is a backwards-compatible superset of revision 1)

2. Create a revision key specific to SBP, such as the DPP group
did for their work

3. Specify the use of multiple version keys in unit directory

4. Don't modify the unit directory, instead add an SBP-specific
way to discover the SBP version, such as a new Query Version
management ORB.

In discussion, Fuller proposed solution 1, with solution 2 as a fallback, and
Anderson endorsed this proposal. Fuller reiterated that proposal 3 could impact
many existing devices. Johansson agreed to pursue solution 1, then 2, at the
ballot review committee meeting of IEEE 1212 on March 23.

[7.1] Johansson noted that IEEE 1394.3 ("Peer to Peer Data Transport"/PPDT)
does not perfectly comply with SBP-2, upon which it is based, as is explained in
1394.3, regarding the meaning of "unordered". Fuller and Anderson agreed that
it would be helpful to note in SBP-3 that devices reporting the "unordered"
execution model may have higher-level ordering requirements, and that in
general software was obligated to preserve the order of messages generated by
a target, even if the target claimed to be "unordered". Johansson agreed to add
explanatory text. Anderson volunteered to propose a glossary entry for
"unordered", or suitable editorial corrections where the term is used in SBP-3, to
clarify that "unordered" indicates only the possible absence of perfect end-to-end
linear order in the system, not the total absence of requirements to preserve
ordering at some points, such as in the delivery of ORB status notification.

[7.6] The group agreed that SBP-3 should recommend the use of Instance
Directories in addition to traditional Unit Directories, in accordance with IEEE
1212r.

Johansson noted that IEEE 1212r strongly recommends the use of only one Unit
Directory in the Root Directory, and asked which this would be for an
isochronous device: SBP or AVC? Nobody present knew if AVC devices could
handle instance directories. All agreed that legacy operating systems would
need an SBP Unit Directory to discover and access an SBP-3 device as a
storage device.

[7.2] Flake presented his proposal for RBC-driven access to AV/C data:

 ftp://ftp.t10.org/t10/document.01/01-067r0.pdf

The proposal describes new commands transported by RBC to create, destroy, and discover AV tracks, plus read and write data asynchronously to/from such tracks. In discussion, the issue of how to select a track format (computer or AV, or more detailed selections) was raised. Fuller stated a desire to be able to create a track with RBC, populate it with data using RBC, and then designate it as an AV track so that it became available to AV devices only after its contents were complete and coherent. Fuller also noted a desire to describe real-time performance requirements to the drive so that it could allocate blocks appropriately. Fuller added that at least one bit per track would be needed to designate tracks containing content-protected data that could not be read or written via RBC.

Flake offered to add more parameters to the create track command, such as expanding the length argument to six bytes. Johansson noted that Roberts had offered to put new commands into RBC-2 and suggested that the proposed commands be reformatted into drop-in text to facilitate this.

Anderson noted that the AV/C disk model has the ability to grow and shrink existing tracks, and asked if this capability should be made accessible through RBC-style commands. Flake answered that his proposal could be extended to accommodate this feature.

Anderson stated that he wanted the capabilities offered by the new commands, but questioned if Flake had really solved the partitioning problem as claimed in the proposal. After discussion, Johansson concluded that the partitioning problem had not been solved. Fuller and Anderson each stated a desire for the existence of hard drives that could have multiple RBC and AV partitions, as well as a desire for the existence of drives capable of isochronous access to media data in an RBC partition (track) through an SBP/ORB mechanism. Anderson noted that both kinds of drives would share most or all of their firmware and hardware resources. Fuller posited that both kinds of drive would actually be the same drive operating in one of two possible configurations. Fuller noted that data protected by 5C cannot be accessed through an SBP/ORB mechanism.

Green proposed that the problem of merging ORB-driven and AV-driven access could be pushed down into the disk if a computer would use the disk's filesystem instead of imposing one of its own. The computer could store individual files in individual tracks, of which there might be thousands or even millions, causing AV tracks and traditional files to coexist in a filesystem managed by the disk. Green added that traditional filesystem metadata such as creation date or file type could be stored in block zero of non-AV tracks, provided the computer driver adjusted LBA's accordingly. Johansson noted that without complete control over the filesystem layout, a computer might face performance issues in some operations,

such as searching and catalog maintenance. Anderson expressed interest in the idea but noted that considerable further study would be required.

[7.3] Johansson gave a presentation regarding changes to support 1394.1 bridges in SBP-3:

  ftp://ftp.t10.org/t10/document.01/01-070r0.pdf

Johansson described the concept of a Node Handle, which replaces the 16 bit bus ID/node ID field in some data structures. A node handle is never directly used on the 1394 bus as an address, but is communicated within data structures such as ORBs and page tables, to provide a long-term reference to a node that is invariant across bus resets and net generation changes. A target translates the node handle into an actual bus ID/node ID value when sending packets on the bus. In some events the target must revalidate a node handle to learn the currently corresponding bus ID/node ID values. Johansson's proposal described a new Get Node Handle ORB that can be sent to the target to cause it to synthesize a node handle. Johansson proposed that support for node handles be optional in SBP-3, but added that node handles had desirable properties even in the absence of bridges, because they could speed up the resumption of processing after a (local) bus reset by eliminating the need to abort the entire task set.

Johansson asked how a target could indicate that it supported bridge-aware logins (with node handles). Fuller suggested that a new bit in the login ORB could be used to request a bridge-aware login. Johansson noted that the target could generate and return a node handle referring to the initiator, placing this value in the first 16 bits of the 4th quadlet of the login response, which is presently reserved. Johansson offered to look for other uses of 48-bit addresses in SBP to see if further consideration of node IDs and node handles is needed. Fuller observed that the bridge awareness of a target could be discovered from the "b" bit in the target's Bus Info Block.

[7.5] Farhoomand gave a brief presentation regarding DVD playback over 1394 using SBP-2. Farhoomand commented that Panasonic has found this playback to be problematic because there are no real-time guarantees with SBP-2, and bus congestion can disrupt the video playback. Consequently, Panasonic would like to see isochronous services in SBP-3 that could solve this problem. Farhoomand added that Panasonic did not plan to use the AV/C command set or FCP protocol in such an application. The group endorsed this goal and invited a more detailed proposal.

[7.9] Johansson led a discussion regarding which keys from the Unit Directory should also be valid in Logical Unit subdirectories. The group agreed to allow the

Unit_Characteristics key (section 7.4.8) to exist in Logical Unit directories, provided the mgt_ORB_timeout field is ignored. The group also agreed to allow the Fast_Start key (section 7.4.10) in Logical Unit directories. Anderson suggested that hints, examples, or suggestions should be provided regarding the use of text leaves within Logical Unit directories to help identify the function of each logical unit.

[7.8] Fuller presented a proposal prepared by Scott Smyers of Sony regarding access to AV media in an SBP disk drive:

  ftp://ftp.t10.org/t10/document.01/01-102r0.pdf

Fuller noted that the presentation considerably overlapped the presentation made earlier by Flake. Johansson led a discussion of command sets and command parameters, especially those pertaining to track attributes.

Anderson asked about protecting 5C data from asynchronous access. Fuller explained that a drive supporting 5C would decrypt data received from 1394 according to 5C, and re-encrypt the data before storing it on the media. When transmitted back onto 1394, the drive would decrypt the data on the media, then encrypt it according to 5C for transmission on 1394. Whether or not the drive actually needed to re-encrypt the data before recording it was debated, without clear resolution. Anderson noted that drive customers who had no need to store 5C-protected data might be reluctant to pay for hardware supporting two independent encryption mechanisms that they would not use.

Fuller commented that the previously discussed concept of "chunkiness" for storing real-time data on an SBP/AV drive was naive, because it did not take into account the physical geometry of a drive, in which some sectors can be accessed more quickly than others. Fuller stated that sustained bandwidth was a better way to characterize performance requirements than chunkiness. After discussion of how these requirements should be expressed, the group concluded that some kind of metadata would be needed. Johansson noted that considerable effort would be required to collect all the information needed to characterize metadata from the appropriate specifications. Johansson suggested that this metadata could be stored in the track attributes data, allowing a disk to decide where to actually record this information.

Johansson led a discussion of the commands required for SBP-driven access to AV data. The group identified the following commands:

  Read/Write track relative (location expressed as an LBA relative to the start of the track)

Get/Set track info (metadata, perhaps including format data)

Create Track (with a desired length argument)

Delete Track

The group identified the following possible track attributes:

Track ID (an opaque (no specific meaning) 32-bit value)

Track Length (expressed in blocks)

Track Format ID (such as MPEG-2, DV, 61883-6, etc.)

Minimum bandwidth

Permissions (5C protected, etc.)

Owner

Fuller suggested that the Minimum Bandwidth attribute would not be adjustable after a track was created, because the drive would allocate physical sectors based on the bandwidth requirement. Anderson noted that this value might be safely reduced after a track was created to give a drive more flexibility in maintaining the track or scheduling access to it.

[7.4] Wooten led a discussion regarding how physical IO (reads and writes to FireWire memory addresses below four gigabytes) could work in the presence of (and over) 1394.1 bridges. Wooten observed that a host computer on a different bus from an SBP target might reboot. The target would be unaware of this, and might continue to perform physical IO to the host - possibly corrupting data in either the host or the target. Wooten noted that if the host could obtain a new global node ID before enabling its physical unit after a reboot, the host would be safe. Wooten also suggested that the Open HCI interface in the host could be made more selective about allowing physical access, though this would be a somewhat complex addition to OHCI.

Anderson suggested consideration of a brute force solution in which a newly-booted host computer performs an exhaustive search of every device on every bus, and tells each device invalidate any old physical memory references. The host would not enable physical I/O until this process was complete. Johansson noted that no generic mechanism exists to learn that a remote node holds a physical reference to another node. Anderson suggested that a broadcast (or exhaustive unicast) message to invalidate GUID-based references could be

invented. Johansson noted that a broadcast mechanism does exist, but is unreliable because broadcasts are not acknowledged. Johansson added that the existing message does not convey that references must be invalidated, but it could be modified to enable this. Anderson posited that a host computer has adequate processing and memory resources to perform an exhaustive traversal of all buses, and that the performance reduction from avoiding physical I/O during this search would not be noticed by end users. Johansson noted that the existing message is allowed to be sent either as a broadcast or directed to a single node's Message Request register.

Fuller claimed that Open HCI 1.1 could work correctly in a 1394.1 environment, and be bridge-aware, provided that remote physical access is always disabled. Wooten described a future Open HCI that could have a table of off-bus nodes that are allowed to use physical I/O. Wooten suggested that table-driven access control and exhaustive-search invalidation of stale references should both be allowed as viable ways for Open HCI to work with 1394.1. Anderson commented that four to eight slots in the remote access table would probably be plenty, and that a host could manage this resource like a translation lookaside buffer in order to talk to more devices at once. Wooten noted that usage counters would be needed in order to get reasonably efficient use of the slots by host software.

Fuller proposed that any change in a global node ID should cause a target to be obligated to reconnect with its initiator(s) before continuing I/O, rather than just requiring the target to refresh old node IDs based on GUID checks. Wooten noted that when requesting a new global ID, a host computer has no way to know for certain that the new ID is different from one it recently held, so using a newly requested ID does not guarantee security against stale physical I/O.

Johansson asked if annex H should become normative.

Action: Hunter will find an AV/C expert to review annex H for correctness, and to glean track metadata parameters from AV/C specifications for incorporation into Flake's proposal.

[8] Johansson proposed the following meeting schedule, with each meeting taking two consecutive days within the range indicated:

 April 26 - 27, Portland OR - Printer Working Group can help with logistics

 June 4 - 8, Toronto - Also PWG; SBP would probably be June 4-5

 July 17 - 18, Colorado Springs - T10 can help with logistics

 August 27 - 29

October 2 - 5

November 6 - 7, Monterey - T10 can help with logistics

December 10 - 14

The group agreed to hold the next meeting in Portland on April 26 and 27, as proposed. Anderson volunteered to host the proposed August meeting at Apple in Cupertino. Johansson suggested that the proposed October meeting be held in New England.

Meeting adjourned