Accredited Standards Committee
X3, Information Processing Systems

Doc. No.: X3T10/94-100 R0
Date: April 19, 1994
Project:
Ref. Doc.:
Reply to: G. Penokie

To: Membership of X3T10
From: George Penokie and Ralph Weber
Subject: Minutes of RAID Study Group Meeting - April 19, 1994

Agenda

1. Opening Remarks
2. Attendance and Membership
3. Approval of Agenda
4. Report on last RAID Working Group
5. SDA States and Types (94-041r3)
6. SCSI Disk Array Model (94-040r3)
7. Error Handling for SCSI Controllers
8. RAID5 Support on SCSI Disk Drives (Gene Milligan)
9. Dual Controllers
10. Action Items
11. Meeting Schedule
12. Adjournment

Results of Meeting

1. Opening Remarks

George Penokie the RAID Study Group Chair, called the meeting to order at 9:00 am, Tuesday, 19 April 1994. He thanked Boeing for hosting the meeting.

As is customary, the people attending introduced themselves. A copy of the attendance list was circulated for attendance and corrections.

It was stated that the meeting had been authorized by X3T10 and would be conducted under the X3 rules. Ad hoc meetings take no final actions, but prepare recommendations for approval by the X3T10 task group. The voting rules for the meeting are those of the parent committee, X3T10. These rules are: one vote per company; and any participating company member may vote.

The minutes of this meeting will be posted to the SCSI BBS and the SCSI Reflector and will be included in the next committee mailing.

George stated that this is the 18th meeting of the RAID study group and the 5th joint RAID Advisory Board/RAID study group meeting. The purpose of the group is to deal with interface issues related to using RAIDs. The study group will assess the issues and then formulate a strategy for dealing with them.

2. Attendance and Membership

Attendance at working group meetings does not count toward minimum attendance requirements for X3T10 membership. Working group meetings are open to any person or company to attend and to express their opinion on the subjects being discussed.

The following people attended the meeting:

RAID Study Group Meeting Attendees

Name	S	Organization	Electronic Mail Address
Mr. Randy Hall		Array Technology	303-938-6035
Mr. M.W. Jibbe	A	AT&T/GIS	316-636-8810
Mr. George Su		CMD Technology	714-454-0800
Mr. Edward Vegdahl		Computing Devices Int.	612-921-6381
Mr. Allan Cole		CORE International	407-997-6044 x505
Mr. Doug Hagerman	A	Digital Equipment Corp.	hagerman@starch.enet.dec.com
Mr. Paul Massiglia	O	Digital Equipment Corp.	massiglia@genral.enet.dec.com
Mr. Mike Oliver		Dynatek	902-832-3000
Mr. Mike Gerak		Dynatek	902-832-3000
Mr. Howard Grill		Formation	609-234-5020
Mr. Bill Hutchison		Hewlett Packard Co.	hutch@.boi.hp.com
Mr. Greg Twerberg	O	IBM Corp.	512-253-6018
Mr. George Penokie	P	IBM Corp.	gop@rchvmp3.vnet.ibm.com
Mr. John Baudrexl		Intellistor/Fujitsu CPA	303-682-6512
Mr. Dave Fellinger		Mega Drive Systems	310-247-0006
Mr. Mike Glass		Microsoft	206-936-4116
Mr. Dick McCormick		Peer Protocols	714-476-1016
Mr. Gene Milligan	A	Seagate	405-324-3070
Mr. Joe Argento		Seek Systems	206-822-7400
Mr. Frank Ruthofford		Storage Concepts	714-852-8511
Mr. Joe Molina		Technology Forums	507-931-0967
Mr. Gary Watson		Trimm Industries	trimm@netcom.com
Mr. Ken Nelson		Trimm Industries	818-764-9500

23 people present

Status Key: P - Principal ** X3T10 status still being defined
A - Alternate
O - Observer
L - Liaison
S,V - Visitor

3. Approval of Agenda

The agenda developed at the meeting was approved.

4. Report on last RAID Working Group

Last RAID Working Group was held in Newport Beach, CA. Several members of the RAID Advisory Board were present. Minutes of the Newport Beach meeting were published on the SCSI Reflector.

Status of this committee's work: A letter ballot on the project proposal for SCSI-3 Controller Commands project submitted by Doug Hagerman has been circulated. The X3 vote closes on 5/2. Passage of this vote is a gate for submitting and voting on a content proposal.

Assuming a favorable vote on the project proposal, the next step is for this committee to "vote to forward"; that puts the proposed standard out for a 4-month public review (not restricted to members of X3T10).

Final editorial note: The decision has been reached to use Framemaker to be used for documents; George will no longer distribute ASCII documents.

5. SDA States and Types (94-041r3) Penokie

George presented the latest revision of the document defining states of an SDA and the types of devices needed within Disk Arrays (X3T10/94-041r3).

The following clarification will be made: SDA States are intended to be reported as bits.

There was a discussion about the "component failure" bit. Doug Hagerman proposed that it be changed to mean "non-addressable component failure" (since addressable components are called by C-LUIs).

There was a discussion about adding a mechanism for indicating what exactly is broken when non-addressable component failure is indicated. The expectation was that request sense would convey more information about state of SDA (ref: page 2; component failure bit).

Resolution: Nonaddressable component failure will mean "something is broken and you need to ask me (the SDA) more." Abnormal bit means "Something is broken and you need to poll the devices to find out what it is. " In addition, it was agreed to add a new status bit that indicates whether data is protected or not, and to remove the statement about whether data is protected or not from the nonaddressable component failure bit.

There was discussion on the number of protectedness states that should be indicated. There is a difference between protected and fully protected (e.g., for double-parity implementations).

The issues were:

- Protected plus some abnormal bit might be used to indicate something is broken but data is still partially protected.
- Nonprotected should be set under normal circumstances for a JBOD or a RAID0.
- Available can no longer be all zeros; there needs to be an available bit since a JBOD or RAID0 would have nonprotected set normally.

Resolution: George Penokie to look at the possibilities & rewrite. May completely delete the protected/non-protected bit at the SDA level.

Meaning of 'limited' as in limited availability will be clarified to mean 'can't do reads & writes to' the entity.

A new state: volume partially exposed—device has failed but data is still valid & partially protected will be added to the volume states.

It was decided that when the 'available' state is reported is up to the vendor. (V-LUI)

The redundancy group states also need a 'partial exposure' state. There was also discussion on whether it is appropriate to try to encapsulate information on rebuild state and on protection state in a single datum. No resolution.

The exposed state under spare states will be deleted.

A state of "not supported" state (like PLUI's) will be added to all objects.

Question from Trim on adding states to describe CLUIs (e.g., LED, horn, etc. on or off) was discussed.

The consensus was that it doesn't belong in a report of operational state.

6. SCSI Disk Array Model (94-040r3)

George presented a revised version of the SCSI-3 Disk Array Model (X3T10/94-040r3).

Section 4.1.1 (Physical address mode) needs to be updated—The problem is it refers to CLUIs, which is inappropriate in this context.

Section 5.1.2: Service should be called “Attach to C-LUI”. “Deattach” should read “detach” . Also, a clarification: detach implicitly means “in same order as attach”. (i.e. detach a from b is not same as detach b from a). Attach is not reflexive.

Comment: Wording on definition of attach to C-LUI service—C-LUI is frequently used where the word ‘attachment’ is intended.

Section 5.1.9: Service should be “report C-LUI attachment’” service (not associations, as written). Note: The meaning of attachment is and will remain vendor-specific.

Comment from Doug Hagerman: The term attachment is ill-defined; probably better not to clarify wording without understanding the meaning of attachment better.

7. Error Handling for SCSI Controllers (Hagerman)

Doug Hagerman gave a talk on error codes using DEC’s list for one SCSI product as background. George Penokie had proposed a shortening of Doug’s list essentially by folding multiples into one code.
Discussion: Reasons for error codes:
to permit an initiator to take independent action.
to precisely define what is going on so that humans can take action.

Group needs to figure out the purpose of error codes and then it can settle on whether to do a short list or a long one.
MK Jibbe agreed to send a list of error codes to George.
Lengthy discussion on what a ‘long’ list means.
Microsoft argued in favor of a list long enough to allow them to take meaningful driver action for a variety of products.
Note: Vendor-unique 2nd bytes can be used with ‘standard’ first bytes (e.g., configuration failure is a standard; there can be many vendor- unique reasons for a configuration failure).
Straw vote heavily favored short list plus vendor-unique.
Resolution: 30-40 standard codes. Rest would be vendor-unique, with vendor identified from identify message.
Vendor is responsible for supplying the text that goes with each vendor unique code.
Doug agreed to send out George’s condensation of his long list; other members are requested to respond to George at GOP@RCHVMP3.VNET.IBM.COM with their requirements to add to the short list.

There followed a detailed discussion of George’s condensation of Doug’s list of error codes. The following were resolved:
Implementation-specific codes:
“Drive returned unknown sense data” should be part of short list.

- vendor-unique
- reserved
- unsupported

next 13 deleted from George’s condensation of Doug’s list.
Keeping “informational; refer to log to find cause of state change.” “State change” will disappear.
delete next 2.
Keep “redundancy level got better/worse”
>From this point in the list down, codes should be added to scsi-3 (not specific to SDAs).
Kill “non-scsi bus parity”—scsi2 already has “internal
communication failure”.
Kill -battery failure
Kill -unexpected message,disconnect, and tag messages
Keep “state change has occurred”
Kill “check data error”

8. RAID5 Support on SCSI Disk Drives (Gene Milligan)

Gene Milligan of Seagate distributed a white paper on xor functions going into a fibrechannel copper implementation design entitled “RAID5 support on SCSI Disk Drives”.

Seagate had previously concluded that this function was not attractive for parallel implementations. They are now revisiting that decision.

Status of document: This is an explanation of an implementation that Seagate is working on that they are willing to make part of the SCSI standard. Seagate is looking to see if other ANSI members are interested in making this a standard. Address for comments is on inside front page of the document.

Doug Hagerman posed a hypothetical question: “Where would a tape array fit into this model” George asserted that it fits into the current SDA model.

9. Dual Controllers

George asked Jibbe (ATT) for input from Jerry Fredin on dual controller; other members may also wish contribute based on their experiences. George believes dual controller is incorporated; if someone objects, he should make his objections known soon.

10. Action Items

a) Penokie Revise the SCSI Disk Array Model (94-040) document based on input received and toward making it cover dual controller configurations.
b) Penokie Revise the SDA States and Types (94-041) document.
c) Penokie Revise the SDA Commands and Mode Pages (94-042) document to account for the “report attachments” service.

11. Meeting Schedule

The next meeting of the RAID Study Group is planned for May 17, 1994 at the Sheraton Inn in Harrisburg, PA. The meeting is expected to run from 9:00am-5:00pm. This meeting will be a joint RAID Advisory Board Host Interface Group and X3T10 RAID Study Group meeting.

12. Adjournment

The meeting was adjourned at 3:00 pm. on Tuesday, 19 April, 1994.