To: Membership of X3T10 X3T10/94-219r0

From: George Penokie/Tom Battle

Subject: Minutes of X3T10 PFA Study Group Meeting -October 29, 1994

AGENDA

1. Opening Remarks
2. Attendance and Membership
3. Approval of Agenda
4. Clarification of Study Group Objectives
5. Discussion of Exception Handling Selection Mode Page
6. Open Discussion of Alternative Proposals
7. Action Items
8. Meeting Schedule
9. Adjournment

Results of Meeting

1. Opening Remarks

George Penokie convened the meeting at 8:00 am, Friday, October 29, 1994.

He thanked Tom Battle of Adaptec for hosting the meeting.

This is a meeting of the X3T10 PFA study group. The purpose of the group is to deal with PFA reporting issues for SCSI-3. The study group will assess the issues and then formulate a strategy for dealing with them.

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 the X3T10 Chair and would be conducted under the X3 rules. Working group 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.

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.

PFA Study Group Meeting Attendees

Attendee	Company	Email Address
Tom Battle	Adaptec	tom_battle@corp.adaptec
Terry Braun	Adaptec	tab@talking.com
Norm Harris	Adaptec	nharris@eng.adaptec.com
Larry Lamers	Adaptec	ljlamers@aol.com
Jamie Odell	Adaptec	jodell@corp.adaptec.com
Steven Fairchild	Compaq	sfairchild@bangate.compaq.com
Tom Treadway	DPT	treadway@dpt.com
Pat Edsall	HP	edsall@hpdund48.boi.hp.com

Gary Lin	HP	glin@ppg01.sc.hp.com
Jitendra Singh	HP	jsingh@ppg01.sc.hp.com
John Lohmeyer	NCR	john.lohmeyer@ftcollinsco.ncr.com
James McGrath	Quantum	jmcgrath@qntm.com
Kevin Tso	Quantum	ktso@qntm.com
John Lingo	Seagate	john_lingo@notes.seagate.com
Gene Milligan	Seagate	gene_milligan@notes.seagate.com
Vit Novak	Sun	vit.novak@sun.com

3. Approval of Agenda

The agenda developed at the meeting was approved.

4. Clarification of Study Group Objectives

Jim McGrath asked for a clarification of objectives for the Study Group. George Penokie offered the following: 'Develop a SCSI method to report asynchronous events, specifically predictive failures of SCSI devices, but open to reporting events of interest from within storage subsystems.'

5. Discussion of Exception Handling Selection Mode Page

Discussion of Exception Handling Selection Mode Page Prior to the meeting, George had prepared a new Informational Exceptions Control Page proposal. That document formed the basis for most of the discussion.

Jitendra Singh raised the issue of how this proposal would relate to DMI, and whether this might be viewed as an alternative to a DMTF MIF. The Server Group within DMTF is defining the means by which components would report events upward to management software. Several attendees felt George•s proposal would not constitute an alternative method, but could in fact complement the DMI methods. There was also general feeling that any coupling with DMI would be better handled at the driver level. Gene Milligan suggested either CAM or ASPI as the proper interface for DMI.

There was insufficient understanding within the group to thoroughly discuss the relationship to DMI. Larry Lamers proposed that communications be established with that task force activity to discuss possible dependencies.

Jim McGrath questioned why asynchronous event reporting would not be better handled by a log sense enhancement. George replied that log sense is primarily intended to log events.

George read through each section of the proposal, with added comments and background. He noted that the methods for reporting events, defined in Table 2, were the result of previous discussions on reporting asynchronous events. Four methods were required to satisfy the requirements of various developers. Only one method may be used at a time.

George was asked to better define 'informational exception conditions.'

Steve Fairchild asked for clarification on why method-0 differed from 5D00. George replied it differs in that it doesn't rely on Unit Attention, which is often ignored by many device drivers. Multiple methods accommodate those who want Asynchronous Event Notification (AEN), or polling, or Unit Attention.

Jim McGrath suggested George's technique might be used even more broadly, to notify the system of any asynchronous event not related to the current command. He felt this interpretation would help clarify the difference between this method and AEN. George deferred, however, since this could have significant ripple effects on the rest of the standard

which would be very difficult to deal with. A narrower scope, focused on reporting predicted failures, would have greater likelihood of acceptance.

Jim asked whether it would be proper to use the Recovered Error mechanism to report a change in mode page. George replied that systems typically use either Recovered Error or Unit Attention as a trigger. Steve Fairchild said Compaq uses Recovered Error, since it's difficult to report an event unless an IRQ is outstanding. Compaq's other alternative would have been to poll. George pointed out that drivers often ignore Unit Attention, agreeing that this might not be the best choice.

John Lingo asked that George's proposal be made specific to reporting a 'failure prediction' in order to narrow the scope of the document and implementation. George agreed to use his redefinition of 'informational exception conditions' to limit the scope appropriately.

Jim McGrath requested a rewording of Code 0 on Table 2. It says '...not reporting information...' but then '...to find out about information...' Instead it should simply say that the target will respond to a poll.

Gene Milligan asked whether the DPFR switch would override all other related driver settings, e.g., Post Error bit, Log Sense / Log Select? John Lohmeyer suggested the proposal could be either broad or narrow in scope, but must be specific. Jim McGrath suggested that if narrow, then the exception conditions must be specified.

George Penokie felt the wording should limit the scope to asynchronous event coverage only. The RPF bit would be treated as an override bit to related items such as Post Error.

The interval field was designed in a manner similar to other SCSI-3 proposals. No exception values are allowed since the field is a generous 4 Bytes wide. If set to zero, the event is only reported once. Otherwise, the event is reported at the earliest opportunity after the specified time has elapsed.

Terry Braun questioned whether the reporting intervals and mechanisms shouldn't be handled by the driver instead of the device. George stated he thought the systems designers would prefer to keep dependencies isolated in one place, namely, the device.

Jim McGrath felt there needed to be a way to stop the reporting after awhile since, for example, RAID systems can essentially repair themselves. George suggested the addition of a count-down counter to stop the event reporting after it reaches zero.

George also questioned whether events of higher importance/severity should be able to define more frequent report intervals or a longer count-down range. Steve Fairchild felt it would be impossible to agree on defined levels of severity. Jim McGrath stated that mode page information should be kept as generic as possible, and we shouldn't open the door to customization.

Jim suggested that a new event should restart the down-counter. John Lingo agreed with this, since the driver has the ability to turn off reporting if desired. He also expressed a desire to keep the specification simple. Jim added that a predicted failure is more accurately defined as a state than as an event. Once a device enters that state, it doesn't recover. Steve Fairchild asked whether temperature effects would be an exception; George replied that temperature is typically treated as a separate issue from predicting failure.

John Lingo asked whether DPFR disables 'operations' or 'reporting.' He suggested there be two bits, one for each. Even if reporting is turned off to the host, the device manufacturer may want the data for warranty analysis, etc. George agreed to add this to the proposal.

Jim Kahn questioned how video servers might handle predicted failure events. Data delivery rates must be guaranteed, yet he wouldn't want to totally turn off notification of predicted failures. George suggested that Jim's application might lie outside the scope of the proposal. Larry Lamers suggested he'd have to rely on the polling mechanism, and choose this poll event

at a known non-critical interval. Most failures predicted by this mechanism wouldn't occur except over a period of hours or days, so the server shouldn't be at undue risk.

John Lingo raised the question of predicted failures which might be more urgent, especially as algorithmic techniques become more advanced. George agreed the proposal should be reworded to clarify when the opportunity for reporting should start. For example, the first report should be made immediately, not after the timer hits its first timeout. The current SCSI command will be completed before Check Condition is reported. A Unit Attention will kill the command. Larry Lamers argued that a predicted failure caused by recovered errors check condition would always hide the predicted failure check condition. Gene Milligan added that a Read Continuous (RC) setting might do the same. Jim McGrath suggested the driver must poll if these situations are of concern. Or, George suggests a vendor-unique solution of rotating a queue of reported errors.

Jim McGrath suggested some customers might want predicted failure analysis to be done offline. Eventual algorithms could require complex analysis of data available on reserved cylinders, and/or video servers might not be able to afford any time for analysis. This would be considered an 'offline' operation in that the media would be active but could not be accessed. It would be analogous to a thermal recalibration procedure today. Gene Milligan pointed out that RAID subsystems might want to do this analysis at scheduled intervals for the entire subsystem. Through discussion, several suggestions were offered for command options:

a) A 'Send Diagnostic' code could be embedded in a command to specifically initiate analysis. Some degradation of performance would be accepted during this time.

b) The device would be taken logically offline until analysis is complete.

This option wouldn't be acceptable for video servers, but might be OK for RAID systems with a spare pool.

c) A bit could be set indicating that reporting is silent, for video applications requiring guaranteed delivery. This option would specifically state that degradation of performance would not be accepted. Overall, it was felt that this topic was not necessarily part of George Penokie's proposal, and might better be handled separately. Jim McGrath agreed to develop a new proposal for this.

Gene Milligan suggested the need for a 'third party failure report' scheme. This concept stems from discussions in the Small Form factor Single Connector Attach meetings. The idea is to utilize sense pins as a means of reporting failure of other hardware components such as fans. George Penokie suggested this might be better handled by creation of something like a 5D03 code. Vit Novak agreed to follow the topic in the SCA meetings, and try to couple it back to this Study Group if warranted.

Steve Fairchild suggested the need to log errors, events, exceeded thresholds, etc. This information would perhaps be used offline to perform statistical failure analysis for long-term quality improvement, warranty info, etc. He stated that Compaq is currently circulating an internal proposal for this. Once initial feedback is incorporated, he'll make the proposal public on the SCSI Reflector.

Steve suggested that drive vendors should make the thresholds open-ended so that they could be adjusted if needed. Jim McGrath (and others) felt strongly that this would be impossible to deal with. Even metrics generally used by industry are very open to interpretation, and are always hard to specify. The whole issue is very vendor-specific.

Terry Braun and John Lohmeyer suggested 'failure sensitivity' inputs, but this was countered by the premise that a predicted failure is a predicted failure. There shouldn't be gradations put on this, only the imperative that the drive should be replaced as soon as possible. Larry Lamers pointed out that Post Error could be enabled if the user really wants to get all the details. Jim

McGrath reminded us that the drive vendor can always predict failure better than the user or system integrator.

George gave a few details of IBM's predictive failure techniques. They primarily monitor soft errors (prior to ECC on the fly), and head flying height changes. Flying height changes account for 80% of the prediction sensitivity. They feel that 40% of failures are predictable today. Of those, 95% are related to head issues. IBM's false alarm rate is very low.

Jim stated that DEC had pursued this technology for 20 years, achieving about a 50% failure prediction rate.

Gene Milligan cautioned those present to respect potential legal issues regarding predictive failure technology. Care must be taken that this isn't perceived as a scheme to sell more drives. Jim McGrath also cautioned about warranty ramifications if the false prediction rate is too high. John Lohmeyer suggested a name change to 'Probable' or 'Potential Failure Analysis' to clarify expectations for the feature. George agreed to accept and share name suggestions over the SCSI Reflector.

6. Open Discussion of Alternative Proposals

No alternative proposals, other than the enhancements and clarifications noted above, were presented.

7. Action Items

1. Jitendra Singh to provide George Penokie with the name of an IBM member in DMTF.
2. George Penokie to better define 'informational exception conditions,' in part to define the scope of this proposal.
3. George Penokie to reword Code 0 of Table 2.
4. George Penokie to add a down-counter mechanism to the proposal.
5. George Penokie to expand DPFR to two bits, one to disable reporting and one to disable the feature.
6. George Penokie to reword the proposal to clarify when the first opportunity for reporting should start.
7. Jim McGrath to develop a 'Send Diagnostic' proposal, or some other scheme, for controlling how predictive analysis is handled by the device, i.e., online, offline, etc.
8. Vit Novak to follow the 'third party failure report' in the SCA meetings, and try to couple it back to this effort if warranted.
9. Steve Fairchild to publish Compaq proposal for logging errors, events, exceeded thresholds, etc.
10. George Penokie to share candidate name suggestions, e.g., 'Probable Failure Analysis' over the reflector.

8. Meeting Schedule

It was decided that the calendar was too full to conveniently schedule another meeting for this year. Refinement of George's proposal will, for the most part, be done via email on the reflector. If sufficiently complex issues remain after his refinement, then a meeting will be called at that time.

9. Adjournment

The meeting adjourned at 11:45 AM on October 28, 1994.