

APPROVED MINUTES OF THE 12TH FAIS AD HOC MEETING

The twelfth meeting of the T11.5 FAIS Ad Hoc group took place at the Hard Rock Hotel in Chicago, IL, on June 8-9, 2004.

The following people were present:

AAROHI	Parag Bhide	parag@aarohi.net
ASTUTE NETWORKS	Tanjore Suresh	tsuresh@astutenetworks.com
BROCADE	Ed McClanahan	edmc@brocade.com
CISCO	Claudio DeSanti	cds@cisco.com
CISCO	Silvano Gai	sgai@cisco.com
CISCO	Maurilio Cometto	maurilio@cisco.com
CISCO	Ronak Desai	rodesai@cisco.com
CNT	Harry Paul	harry.paul@cnt.com
CNT	David Peterson	dap@cnt.com
DOT HILL	Elizabeth Rodriguez	Elizabeth.Rodriguez@dothill.com
EMC	David Black	black_david@emc.com
EMC	John Kemeny	kemeny_john@emc.com
EMC	Fred Oliveira	oliveira_fred@emc.com
EMULEX	Ken Hirata	ken.hirata@emulex.com
EMULEX	Bob Nixon	bob.nixon@emulex.com
ENDL	Ralph Weber	roweber@ieee.org
HDS	Shoji Kodama	shoji.kodama@hds.com
HEWLETT PACKARD	Vinod Bhat	vinod.bhat@hp.com
IVIVITY	Eddy Quicksall	eddyquicksall@ivivity.com
LSI LOGIC	Louis Odenwald	louis.odenwald@lsil.com
LUCENT	Richard DiPasquale	rdipasquale@lucent.com
MCDATA	Larry Hofer	larry.hofer@mcddata.com
MCDATA	Michael O'Donnell	mike.o'donnell@mcddata.com
QLOGIC	Craig W. Carlson	craig.carlson@qlogic.com

TROIKA NETWORKS	William Chow	wchow@troikanetworks.com
TRUE FOCUS	Horst Truedstedt	hotrues@attglobal.net
VERITAS	Roger Cummings	roger.cummings@veritas.com
VERITAS	Amitara Guha	amitara.guha@veritas.com (?)

A total of 28 people from 19 organizations.

1. Opening remarks and introductions

The Chairman of FAIS, Claudio DeSanti of Cisco, called the meeting to order and welcomed the attenders, and led a round of introductions.

2. Administrivia

2.1 Approval of minutes

The minutes of the May 3 FAIS Ad Hoc meeting had been uploaded to www.t11.org by William Chow of Troika Networks as 04-359v0.

The following changes were noted:

- Ed McClanahan noted that an action item was missing for incorporating operational model into the spec.
- Claudio noted that some abbreviations for people's names needed to be expanded.

Pursuant to these changes, the minutes were approved in the absence of any objections.

2.2 Approval of this agenda

Claudio had distributed a draft agenda for this meeting as 04-410v0 via the T11 Web site, and he now asked for changes to that agenda. No modifications were offered.

Claudio moved to approve the agenda; Ed McClanahan seconded. No objections.

3. Call for Patents

Claudio stated that amongst rules and policies under which this Working Group operates are the T11, INCITS, and ANSI patent policies. He directed persons wishing to make statements relevant to those policies to do so at the T11.5 or T11 plenary meetings.

4. Review of old FAIS action items

Old action items were reviewed. Their status was updated as follows:

- AI 1: Ed McClanahan to review the required operation if an i/o crosses a boundary and the second extent is "held". (Opened 7/9/2003)
Open.
- AI 2: Tanjore Suresh to investigate methods of generating WWNs. (Opened 7/9/2003)
Open.
- AI 3: Robert Mulcahy (ER?) to propose support for journaling. (Opened 7/9/2003)
Open.
- AI 4: Roger Cummings to provide requirements for statistics gathering in FAIS and present proposals (Opened 9/10/2003)
Tabled until object model finalized.
- AI 5: Ed McClanahan and William Chow will work on a draft of Chapter 5. (Opened 11/9/2003)
Open.
- AI 6: Roger Cummings to check on possible restrictions for upcall context. (Opened 12/9/2003)
Closed - OBE.
- AI 7: Tanjore Suresh will cross-reference names/types of capabilities against existing standards/prior work. (Opened 1/13/2004)
Open.
- AI 8: Roger Cummings/Fred Oliveira/Bob Mulcahy will each provide a list of capabilities that they must be able to query support for. (Opened 1/13/2004)
Open.
- AI 9: Ed McClanahan will propose an approach for supporting failover in the event of CPP failure. (Opened 1/13/2004)
Open.
- AI 10: George Penokie will produce a UML conventions section for inclusion in the standard. (Opened 2/3/2004)
Open.
- AI 11: Roger Cummings will provide specific requirements for DPC pass-thru support. (Opened 2/3/2004)
Open.
- AI 12: Elizabeth Rodriguez will propose support for journaling. (Opened 4/6/2004)
Open.
- AI 13: Mike O'Donnell will update the glossary per any overlap with the terms defined in the requirements document. (Opened 5/8/2004)
Open.
- AI 14: William Chow will post an updated requirements document per feedback from May meeting. (Opened 5/8/2004)
Closed by 04-037v2.
- AI 15: Ed McClanahan and Maurilio Cometo will collaborate on a new proposal for the execution model. (Opened 5/8/2004)
Open.
- AI 16: Silvano Gai will update the chapter 6 draft per feedback from May meeting. (Opened 5/8/2004)

Closed by 04-325v1.

AI 17: Fred Oliviera will submit his proposal for xmap layering. (Opened 5/8/2004)

Closed by 04-354v0.

AI 18: David Black and Silvano Gai will propose modifications to the object model to directly support the I_T nexus. (Opened 5/8/2004)

Open.

AI 19: Ed McClanahan will propose chapter 6 text for separate client and provider-specified object handles. (Opened 5/8/2004)

Open.

AI 20: Tanjore Suresh will propose chapter 6 text for iSCSI-specific port attributes. (Opened 5/8/2004)

Open.

AI 21: Ed McClanahan and Silvano Gai will unify their object models. (Opened 5/8/2004)

Open.

AI 22: Ed McClanahan will propose a draft for chapter 1 and 2. (Opened 5/8/2004)

Open.

AI 23: Mike O'Donnell will propose a draft for chapter 3. (Opened 5/8/2004)

Open.

AI 24: Ed McClanahan will incorporate the operational model (04-340) to the draft specification as chapter 1. (Opened 5/8/2004)

Open.

5. Old business

There was no old business at this meeting.

6. Scheduled FAIS business

6.1 Cisco Object Model

04-025v3

Silvano Gai stated two unresolved issues: recursive Xmap and LUN map. He also didn't want to add an xmap as an attribute to a VDEV, as that adds overhead per VDEV processing. He proposes allowing each specific VDEV type to reference an Xmap. Ed asked how does this differ from having the Xmap derive from the VDEV. Silvano responded that he'll defer this issue until EMC's presentation.

Silvano noted that he added a "default LUN map" for pre-login, and "LUN map" for post-login. William asked how they differ. Silvano responded that he wasn't sure. Roger Cummings asked if you could layer an Xmap over an Xmap. Ed noted that this is allowed in Fred Oliveira's proposal but not in Silvano's. Ed said he thought it was necessary to keep the notion of an ITL since we need to support ITL-specific properties, such as access permissions for reservations.

6.2 EMC object model

04-401v0

Ed noted that there still isn't an ITL, and his LUN map doesn't support a permissions for I/Os. David Black said that the reservation is a property of the FLU. Ed asked how task sets would be handled, such as aborting I/Os for a specific initiator. David responded that this still needs to be an attribute of the FLU since you need to abort I/Os for an initiator across multiple TL's. Ed noted that Brocade supports operations on the ITL, and depend on the Client to perform the operation on all necessary ITL objects.

Ed asked about the value of having a FLU; he said this is only required if you don't have a FITL. Fred responded that this simplifies actions that operate on a LU level. William asked if a BITL_set was required. Fred responded no. Ed said this could be a solid arrow (i.e. attribute). William then indicated that this makes it impossible to adjust the BITL, unless some kind of reference is provided; if so, then this becomes a first class object.

Ed noted that a difference from Silvano's proposal is that there is no default LUN map. William noted that the EMC API proposal still supports the notion of a default LUN for every FLU, so this is still implied. Ed said that the LUN map didn't need to be an object, could be handled as a property. He noted that, for the same reason, we should support a "legend" in the Xmap.

William asked about supporting read/write attributes to the BlockRange. He noted that Xmaps are often used with a Concatenated Vdev, since the Client typically uses them together (e.g. snapshot). Fred and Ed noted that this increases the size of the BR. William responded that a single larger BR is no more costly than having a separate BR and Xmap entry.

Fred indicated that he will take the comments and propose modifications as appropriate.

6.3 Chapter 6 - Cisco

04-325v1

1.2.1

Ed asked whether (a) is "durable". Ronak Desai responded "no"; (b) is the "durable" one. William noted that this is incorrect, as (a) is for the initiator while (b) is for the target. Ronak said he will modify this to be FAIS_names for both.

1.2.2

Ronak said that the pathweight is not well defined yet. William noted that this belongs in BITL_set only if BITL is not a member of it. Ronak will add this as a property into the list of BITLs within the set. William asked what "passive" meant (e.g. whether it implied trespass-handling by the Provider). Ronak responded "no".

Ed said that FAIS_name wasn't noted in the previous minutes. Mike O'donnell asked if FAIS_name was even defined. Ed proposed calling this FAIS_Port_name. Dave Peterson asked about multiple ports with the same node name. Ed said that this would be different names for each port. William said that there are other attributes involved for logins. Claudio suggested that this is a separate discussion; if there is anyone interested in this, someone should propose something.

Ed asked about path failures and how they would be detected or handled. Ronak responded that this is implementation dependent so he will remove text on this.

1.2.3

Ed noted that (c) is in the ParentToChild, so shouldn't be in the VDEV. Fred suggested that "RAID type" should be "geometry type".

1.2.4

William noted that the diagram should use names of objects from the object model. Ronak will update it.

Ed noted that the "block range" and the "concat" should be switched. William noted that there should be separate diagrams for the mapping tree from the logical layout. Ronak said he will propose a new diagram, possibly two (i.e., separated).

1.2.5

William said that it is burdensome for providers to check for consistency and completeness. Fred said that Clients don't necessarily need it since problems indicate a s/w error in the client. Ed noted that the alternative is to have some default behavior in the case of inconsistency/incompleteness. Ronak indicated that he will remove the last line. Ed suggested that we need to ensure that the behavior (i.e. fault) in the event of error is identical.

Parag Bhide asked about the difference b/t Control and Fault_*. Amitara Guha suggests that the Client primarily cares about when they need to handle data transfer directly (i.e. thru buffers instead of mappings). Parag suggested that the direct data transfer capability should be supported for faulted I/Os too. William responds that some providers may not be able to support that. Ed noted that there is a requirement to support a converting a fault to a control-path I/O.

Ed suggested renaming Fault to Fault_first. Ronak indicated he will change this.

Ed suggested making quiesce separate from fault permissions. John Kemeny responded that this allowed two different apps to modify the same Xmap for different reasons (one Client controls read/write while the other does quiesce/unquiesce). Ed suggested that this isn't useful and layered Xmaps obviates this. Maurilio noted that quiescing a range means updating all entries within an Xmap. Ed responded that this is expensive and unnecessary since an upper layer Xmap would have much fewer entries. William suggests that the Cisco proposal isn't incompatible with the layered approach. Ed says supporting this requires additional API (i.e. must support the revert) and provider burden (i.e. expensive to modify and remember).

6.4 EMC FAIS API

04-402v0

John said this is based on Cisco object model.

Tanjore Suresh objected to the number of functions necessary to set the callback; he suggested a callback specified in the call parameter or request structure. Ed noted that other options are: 1) central registration function for multiple callback, 2) single callback for all with type-specific information. William suggested that supporting async requests already imposes the bulk of the overhead (i.e. supporting per-function callback is an incremental cost). Eddy Quicksall said that

the per-call callback setup is expensive; he noted that he wants to minimize cycles. Fred noted that these operations are not performance sensitive since they are control-path. If so, Eddy prefers per-call callback. Straw poll for a callback in the invocation: y=6, n=1, a=3.

Ed asked whether people liked callback specification in the function parameters versus in a request structure. William noted that a previously suggested convention that he supports is to pass a request as a structure instead of as parameters. He noted that this could support the callback function. Fred clarified that every callback structure has multiple common attributes. He will update their API to propose something along these lines.

Tanjore asked about timeouts. Amitara noted that the issue is whether the client or the provider implements the timeout. Tanjore suggested that the timeout could be part of the request structure. He also said that the transport timeouts need to be handled. Mike noted that the transport timeout may be dynamic, so the client cannot readily determine the appropriate value. Eddy noted that the client knows the reason for the cancellation, including if it is a timeout. Claudio asked if there was any consensus. Ed said he could support this only if the client could cancel a request. William responded that the provider needs to be able to reject a cancellation attempt. Mike noted that the target has its own timers that can't be controlled by the client. William asked if the cancellation is async; Fred responded yes. Ed asked if the cancellation has a callback; Fred responded yes. Ed recapped the options are: 1) timeout in request enforced by provider, and 2) cancellation supported by the provider w/o a timeout in the request. Claudio asked if there was any objections for #2; none noted.

Tanjore said that the higher-level abstractions (e.g. specific fields called out for the sense data) is unnecessary and should be generic. John indicated that these can be removed.

Tanjore said that the term "path" is inappropriate, since they only identify the target port. Fred said they will replace this with BITL.

Tanjore said the "new" event is unnecessary since there is a "log" event. He said that one need not know about a remote port until it logs in.

William asked if they could walk-thru an example for a mapping tree update. Fred indicated that they will add an example.

Tanjore said that probing and discovery puts SCSI logic in the provider (e.g. probe). Ed added that some of this functionality is already covered in other APIs, such as name service queries (e.g. discoverBE). Parag noted that that the specification should identify which API to use as an alternative.

Ed noted that it was previously suggested that the login accept/reject response could be provided by the client.

Tanjore noted that the SCE (2.0.1) and SCB (2.0.2) structure doesn't support over/under-runs. He also noted that task mgmt functions are not supported.

Fred indicated that he will take the comments and propose modifications as appropriate.

6.5 Chapter 6 - Cisco

04-325v1

Ronak said he thought there was consensus for sync/commit. William said he didn't think this was true since he wanted the option to not support it. Ed also noted that he prefers a pre-commit step to identify the start of the transaction. John suggested that it is necessary to still support quiesce, such as for platforms that don't support a transaction model. Ed asked what type of operations require a transaction. For example, he would not need to do this if he was transitioning an xmap entry from quiesce to allow. Ronak said it would be more consistent to require a transaction for all updates. William said that this is undesirable because it requires all platforms to support a scratch pad, and requires all clients to wrap all operations into a transaction. Fred said that a client doesn't need a transaction model because: 1) the consistency check is no longer required, and 2) the client still needs to quiesce around a transaction (e.g. updates around multiple providers). William suggested that there seems to be consensus that we must support updates w/o requiring a transaction. Straw poll for mappings updates w/o requiring batching: no objections.

Ed asked about where a Client can quiesce. He proposes that we must support quiescing a Vdev that is not an Xmap. William agreed that this helps reduce the resource and I/O processing overhead since it eliminates an extra Xmap. Maurilio asked if every Vdev must support quiescence. William said yes; Ed noted that the API could support it while a provider need not support. Ed moved for quiescing a Vdev; William seconded. Silvano disagreed and noted that this is already covered by using an Xmap. David said that this is not a valid motion since it doesn't address a specific item. Ed moved to change EMC's FLU quiesce to be a Vdev quiesce; Elizabeth Rodriguez seconded. y=5, n=1, a=5; motion carried.

6.6 FAIS Conventions - McData

04-393v0

Larry Hofer said this was copied from the HBA and SWAPI APIs. John asked about how examples are handled. Mike responded that examples should be in an informative annex. John wants to know if any examples are required (i.e. can the spec standalone unmodified w/o them). William said we had a previous discussion on chapter 6 for putting mapping updates examples into an informative annex. Mike noted that in GS-4, they had examples in introductory text as well as in an informative annex.

Larry noted that it doesn't directly indicate inclusion of a header file.

Fred noted that it should be FAIS, not FC-FAIS.

Larry was actioned to note any differences from HBA or SWAPI API specs.

7. Unscheduled FAIS business

Ed proposed adding his proposal for FAIS_port_names. Time expired before this could be discussed.

8. Review of FAIS action items

Old action items were reviewed. Their status was updated as follows:

- AI 1: Ed McClanahan to review the required operation if an i/o crosses a boundary and the second extent is "held". (Opened 7/9/2003)
Open.
- AI 2: Tanjore Suresh to investigate methods of generating WWNs. (Opened 7/9/2003)
Open.
- AI 3: Robert Mulcahy (ER?) to propose support for journaling. (Opened 7/9/2003)
Open.
- AI 4: Roger Cummings to provide requirements for statistics gathering in FAIS and present proposals (Opened 9/10/2003)
Tabled until object model finalized.
- AI 5: Ed McClanahan and William Chow will work on a draft of Chapter 5. (Opened 11/9/2003)
Open.
- AI 7: Tanjore Suresh will cross-reference names/types of capabilities against existing standards/prior work. (Opened 1/13/2004)
Open.
- AI 8: Roger Cummings/Fred Oliveira/Bob Mulcahy will each provide a list of capabilities that they must be able to query support for. (Opened 1/13/2004)
Open.
- AI 9: Ed McClanahan will propose an approach for supporting failover in the event of CPP failure. (Opened 1/13/2004)
Open.
- AI 10: George Penokie will produce a UML conventions section for inclusion in the standard. (Opened 2/3/2004)
Open.
- AI 11: Roger Cummings will provide specific requirements for DPC pass-thru support. (Opened 2/3/2004)
Open.
- AI 12: Elizabeth Rodriguez will propose support for journaling. (Opened 4/6/2004)
Open.
- AI 13: Mike O'Donnell will update the glossary per any overlap with the terms defined in the requirements document. (Opened 5/8/2004)
Open.
- AI 15: Ed McClanahan and Maurilio Cometo will collaborate on a new proposal for the execution model. (Opened 5/8/2004)
Open.
- AI 18: David Black and Silvano Gai will propose modifications to the object model to directly support the I_T nexus. (Opened 5/8/2004)
Open.
- AI 19: Ed McClanahan will propose chapter 6 text for separate client and provider-specified object handles. (Opened 5/8/2004)
Open.

AI 20: Tanjore Suresh will propose chapter 6 text for iSCSI-specific port attributes. (Opened 5/8/2004)

Open.

AI 21: Ed McClanahan and Silvano Gai will unify their object models. (Opened 5/8/2004)

Open.

AI 22: Ed McClanahan will propose a draft for chapter 1 and 2. (Opened 5/8/2004)

Open.

AI 23: Mike O'Donnell will propose a draft for chapter 3. (Opened 5/8/2004)

Open.

AI 24: Ed McClanahan will incorporate the operational model (04-340) to the draft specification as chapter 1. (Opened 5/8/2004)

Open.

AI 25: Fred Oliveira will update EMC's LU layout proposal per feedback from June meeting.

New.

AI 26: Fred Oliveira will update EMC's FAIS API proposal per feedback from June meeting.

New.

9. Next meeting schedule

The next meeting of the FAIS group will be at the T10 Plenary week in Colorado Springs, CO on the week of 7/12/2004. Claudio asked if people wanted to meet there since the only T11 meeting there is for FAIS. Fred and William indicated that there is sufficient work to justify it. Claudio responded that he would then ask for 6 hours.

10. Adjournment

A motion to adjourn the meeting was passed unanimously.