

Minutes of T11.1 HIPPI Ad Hoc Working Group
August 11-12, 1998
Portsmouth, UK

1. Opening remarks and introductions

The Chairman, Don Tolmie of Los Alamos National Laboratory, opened this at 1 PM and thanked Paul Levin and Xyratex for hosting this meeting. This group is constituted as both the HIPPI Working Group under T11.1, and the HIPPI Networking Forum (HNF) - Technical Committee (TC). Don lead a round of introductions. The list of attendees is at the end of these minutes. Don volunteered to produce the minutes.

2. Review / modify the draft agenda

Draft agendas were distributed via e-mail before the meeting. No changes were made at the meeting. These minutes reflect the approved agenda.

3. Review minutes of previous meeting

The minutes of the July 20-21, 1998, working meeting in Albuquerque, NM, were reviewed. Jeff Young moved, and Don Woelz seconded, to approve the July 20-21, 1998 working meeting minutes as written. Motion passed unanimously.

4. Review old action items

The action items from the July meeting were reviewed.

1. Von Welch to contact HIPPI-6400 MIB users and developers for comments on the current draft, and to prepare a presentation on the MIB for a future meeting. (Carryover)
2. Von Welch to look at developing a HIPPI-6400 host system MIB (for a NIC), to be done now as an annex of the present MIB with the possibility of splitting it out as a separate document at a later date. (Carryover)
3. Everyone to review the HIPPI-6400 MIB. (Carryover)

4. Jean-Michel Pittet to update "ARP and IP Broadcast over HIPPI-800" Rev 00 with the changes agreed to at the June meeting. (Carryover)
5. Jean-Michel Pittet to request another HIPPI-800 Hardware Type from IANA. (Carryover)
6. Jean-Michel Pittet to forward the request for an IETF Working Group to the IETF. (Carryover)
7. Jean-Michel Pittet to generate an "ARP and IP over HIPPI-6400" document. (Carryover)
8. Marck Doppke to process the HIPPI-800 switch MIB through the IETF. (Carryover)
9. Greg Chesson to draft an IEEE tutorial on HIPPI-6400 ULA usage and the ULAs special to HIPPI-6400. (Carryover)
10. Jean-Michel Pittet to inform Don Tolmie of the ARP text in HIPPI-LE that should be removed. (Carryover)
11. Don Tolmie to update HIPPI-LE document. (In process)
12. Greg Chesson and Jeffrey Chung to consider developing "reason codes" to explain why a particular ST Operation was rejected. (Carryover)
13. Greg Chesson to send e-mail detailing reasons for not doing a queue for client/server applications, and suggesting how they could be done in ST. (Carryover)
14. Bob Willard to write up something on big/little endian issues for inclusion in the document. (Carryover)
15. Greg Chesson to collect text for a "folklore" annex in the document. (Carryover)
16. Greg Chesson to draft text describing how you differentiate duplicate operations from legal operations. (Carryover)
17. Jim Pinkerton to draft a timestamp option for the optional payload for inclusion in the next ST revision. (Carryover)
18. Craig Warner to update the Put, Get, and FetchOp FSMs and send them to Don Tolmie for inclusion in the next revision of ST. (Done)

19. Ian Philp to draft a C code tiling example for inclusion as an annex in ST. (Carryover)
20. Don Tolmie to update ST Rev 2.0 with the changes agreed to at the July meeting. (Done)

21. Michael McGowen to collect, tabulate, and document everyone's requirements for HIPPI-800 and HIPPI-6400 translation environments. (Carryover)
22. Don Tolmie to update the June 1998 HIPPI Working Group minutes with the corrections identified at the July meeting. (Done)

5. HIPPI-6400-PH (ref: Rev 2.3)

5.1 SuMAC and cable test results

Roger Ronald reported that they have seen the error rate vary as a function of training sequence repetition rate, on the order of 5 errors per minute. Roger did not feel that we have a major risk

Ed Cady said that Berg is working on a cast backshell to reduce the EMI and decrease the cable clamping pressure. Cable management in the field may be a problem with the pressure on the cable when going around corners and people stepping on the cables.

Henry Brandt noted that normal manufacturing practice is to use a longer cable for testing, e.g., plus 10% or 15%. He asked if we are too close to the edge now. Roger said that the eye opening spec is what they see, and as long as they see it things work.

It was noted that the system errors are cable distance dependent, not SuMAC dependent. This was shown by Roger's experience with the Raytheon E-Systems HIPPI-6400 32-port switch. They cabled the switch outputs to inputs and then configured it to send through all ports, i.e., 64 SuMACs in series. The cables used were on the order of 12 inches. No errors were seen, i.e., the SuMACs and short cables are bullet-proof. It was noted that as for HIPPI-800, optical links will eventually be the default, but today optics are about \$25k more expensive.

Henry recommended that we error on the side of being conservative. Ed Cady echoed that he would like some margin too, for ease of manufacturer. Bob Hyer asked if we had more cables that we could test; the answer was no. We haven't tried 35 or 38 or 30 meters so don't know what will work now. We also

don't know of any systems that are hide-bound at 40 meters.

We expect cables to improve with time and experience. Ed Cady noted that Fibre Channel is working on 2 Gbit/s and allowing more headroom; 33 meters on 1 Gbit/s today. For the Fibre Channel cables they started with 2 suppliers for cable, now there are 4 to 6 suppliers; they have invested in engineering. Ed felt that Fibre Channel would be lucky to get 25 for 2 Gbit/s; aiming for 20 m now.

5.2 Review document changes

The changes between Rev 2.2a (the public review document) and Rev 2.3 were reviewed. They consisted of correcting the HIPPI web page URL, adding in the ANSI editor's comments, correcting the acronym for "mega", and adding a note about the single-ended vs. differential receiver input specifications. It was agreed that none of these changes were "technical", all were "editorial".

5.3 Public Review status

A vote to forward is planned for the T11.1 Plenary. Don said that Joel Darnauer, who was doing the cable testing at SGI, was scheduled to call him with the latest results that evening. David Hyer asked what our definition of "correct operation was shown" was. The reply was that it was in the eye of the beholder. Since errors are automatically corrected by retransmission, the user would never see them. The only effect would be to slightly reduce the throughput, e.g., less than 1% degradation for most cases shown. So one view was that corrected errors were acceptable, another view was that there should not be any errors, even corrected ones since that may indicate the possibility of uncorrected errors.

The proposed response to the public review comment, assuming no technical changes were made, was reviewed and modified in preparation for the T11.1 Plenary.

6. HIPPI-6400-SC (reference Rev 2.4)

6.1 Review document changes

Roger Ronald reviewed the changes made since the Albuquerque meeting. All were editorial in nature.

A few additional editorial changes were made at the meeting, e.g., in the Introduction changing "Universal LAN Address (ULA) assignment" to "ULA" and "...16-bit subset..." to "...subset...". Struck the last sentence of the second paragraph of 6.1.

Roger made the changes overnight and gave electronic copies to those attendees with computers the next day. This final document was Rev 2.5.

6.2 Public Review status

A vote to forward is planned for the T11.1 Plenary. The group agreed that HIPPI-6400-SC was ready for a second public review. It was agreed that forwarding HIPPI-6400-SC without forwarding HIPPI-6400-PH was reasonable since 6400-SC required a second public review and 6400-PH would not unless we made technical changes.

The proposed response to the public review comment was reviewed in preparation for the T11.1 Plenary.

7. Scheduled Transfer (reference Rev 2.1)

7.1 Review general document changes

The changes were mostly accepted as written with additional changes in some places. The wording change to the Abstract, Introduction, Foreword, and Scope was accepted. The new definitions for "exposed" and "Schedule Header queue" were reviewed and accepted.

The new sentence and note in 6.1.4.4 for FetchOp sequences were reviewed. The sentence said that FetchOp sequences are done in F-id order. The note said that if a location was accessed through multiple PMRs then the ordering was indeterminate. Another sentence referenced the new FetchOp buffer management text in annex A.6.

In 6.2.1 on Sequence identifiers, the new sentence about unique identifiers within a PMR, and placing the F-id and G-id text to follow, was reviewed and accepted. The open issue about overlapping F-id and G-id was covered in agenda item 7.5.

New section 10.8 on LLP errors was reviewed and accepted as written. There were more document

changes that were essentially editorial in nature that were accepted as written.

7.2 Annex A.5, ST over Fibre Channel

Don did a fair amount of re-write of this section. It now references "IP and ARP over Fibre Channel" (called the "RFC" in the rest of these minutes) instead of FC-LE, and includes many more references to specific parameters, e.g., usage of the other R_CTL bits as defined in the RFC.

In A.5.2.2, Don felt that he had done the EtherType incorrectly, but was convinced by the group that it was OK as is, i.e., specifying the format of the message. In A.5.2.3, discussing the ST header, Don said that he would add a sentence noting that the Op field is the first 5 bits in the ST Header.

Rather than having a separate FC Exchange for each ST Virtual Connection, Don re-wrote it to have a single FC Exchange shared by multiple Virtual Connections. Separate Exchanges were specified for Control and Data based on the fact that the RFC specified that only a single FC Sequence can be active on an Exchange. Hence, if Control and Data shared an Exchange, then a Control message could be blocked by a large Data message. Don raised the question in T11.4 of whether an Exchange could support two active Sequences, i.e., what will common FC hardware support? The answer was that only one active Sequence per Exchange is allowed.

With having multiple Virtual Connections sharing an Exchange, the Exchange setup and teardown also changed. Now a dummy ST Header (with Op = x'00' - a reserved value) is sent to setup or teardown the Exchange. Teardown is now based on a timeout after no more Virtual Connections were active. Roger Ronald felt that using a dummy ST operation with Op = x'00' was not the best way to do it; there may be other native ways in Fibre Channel that support this function. We need a Fibre Channel expert to check this out for us. If nothing else, we should probably define Op=x'00' as "No-op" rather than reserved in ST clause 8.1.

Don noted that he would add text about how the Max_STU could be modified by the FC-ST Mapping entity based on its knowledge of the underlying FC fabric.

7.3 Annex D.6, FetchOp buffer management

This text was based on what Bob Willard presented at the Albuquerque meeting in July. Don made some additional editorial changes as he included it in the document. The text was reviewed. The word "idempotency" was viewed as not helping the understanding. Don said that he would take a look at rewriting the section to help improve clarity.

7.4 Annex E.7, Put, Get, FetchOp FSMs

The changes in annex E were to include the Put, Get, and FetchOp FSMs developed by Craig Warner. Craig had updated the material presented in Albuquerque, and Don made further editorial changes to make it consistent with the rest of the document. The text and tables were reviewed in detail. A few changes were made, e.g., adding the word "Acknowledged" to the title of table E.15 and adding an output action for "invalidate FetchOp cache entry" in table E.18.

One thing that people objected to was the "SS" parameter implying "SS = 1" unless stated otherwise. It was noted that this was consistent with the syntax in the other FSMs and Table E.1. Don will take a look at doing a global change to change all instances to "SS = 0" and "SS = 1" as appropriate.

7.5 F-id and G-id overlap

Ian Philp and some other people at Los Alamos are writing code for an ST implementation. In the process, they noted inconsistencies in the Get and FetchOp sequences. One problem is that there is no clean way to differentiate a Data operation resulting from a Get or FetchOp that use the same value for F-id and G-id. It looks like the F-id and G-id should be combined into a single monotonically increasing parameter to avoid the potential overlap.

Another problem is that the persistent memory region is identified to the Initiator by R-id, a parameter assigned by the Responder. For consistency, it was felt that I-id should be used instead.

Ian has sent this information to Greg Chesson and Jim Pinkerton for their review and assessment. This will be discussed at the September meeting.

7.6 Other proposed changes

None

8. HIPPI-LE (reference Rev 3.4)

8.1 Removal of address resolution text

Jean-Michel Pittet has an action item to recommend the stuff that should be removed from HIPPI-LE that is also covered in the "ARP and IP Broadcast over HIPPI-800" RFC. This action item is still pending. Jean-Michel was not present at this meeting.

Don Tolmie reported that he is in the process of moving the document from his Mac to his PC in anticipation of making the changes. Don plans to highlight sections of the text that he feels could be removed; these will be reviewed at a future meeting for approval by the group.

8.2 Other changes

No other changes were proposed at this meeting. As stated previously, the deadline for other proposed changes is the October meeting.

9. ST-API Mappings

The Project Proposal for ST-API is in process at NCITS. It is unlikely that we will get much done on the project until October or November due to Jim Pinkerton's schedule.

10. Other HIPPI items

10.1 ARP and IP Broadcast over HIPPI-800

Jean-Michel Pittet was not present at this meeting and nothing was done on the project. Jeff Young reported that Jean-Michel has taken on some other duties in SGI, but is committed to completing the document. If Jean-Michel cannot do it, then Jeff will probably step in and finish it.

10.2 HIPPI end-point MIB

Don has a copy of the 1995 document which he put on his web page. It is an Internet Draft by John Renwick that expired in March 1996, and has been

removed from the IETF Internet Draft repositories. If we want to pursue this MIB further then we need a champion for it. Nothing new at this meeting.

10.3 HIPPI switch MIB

Marck Doppke of Essential Communications has a draft document out for comment. Michael McGowen previously offered Marck's services for processing the HIPPI-800 switch MIB through IETF, and Marck has acknowledged this assignment.

10.4 HIPPI-6400 MIB

Von Welch of NCSA has a draft document, based on HIPPI-6400-PH Rev 1.4, out for comment. Von was not at this meeting and nothing new was reported.

10.5 HIPPI-6400 ARP and IP RFC

Jean-Michel Pittet previously reported that the HIPPI-6400 ARP and IP RFC would essentially be a cut-and-paste of the ARP over HIPPI-800 document. He is concentrating on the -800 document now since it is the harder one; the -6400 version should be a subset. Nothing new at this meeting.

10.6 Tutorial for HIPPI-6400 ULA use

Greg Chesson has obtained the format material from Bob Snively of Sun. Drafting of the actual text is pending.

11. Future meeting schedule

11.1 Interim meeting, September 1-2, 1998

The next interim working meeting will be hosted by Mike Andrewartha at the Hewlett-Packard facility: 3000 Warterview Parkway, Richardson, TX. Several hotels close to the facility were identified. (See the meeting announcement on the web page at <http://www.cic-5.lanl.gov/lanp/ANSI/> for further details and a road map of the area.)

Tuesday - September 1 : 1 PM - 9 PM
Wednesday - September 2 : 8 AM - 9 PM

11.2 Plenary week, October 6-7, 1998

The October Plenary week meetings will be on Wednesday, October 7, from 6 PM to 8 PM. The location is the Fort Lauderdale Marriott North, 6650

North Andrews Ave., Ft. Lauderdale, FL 33309, phone: (954)771-0440 or (800)343-2459, fax: (954)771-7519. The room rate is \$125.00 US/night-all inclusive. Use the group name T11/Adaptec when making reservations; reservation cutoff date is September 5, 1998. The host is Adaptec and Norm Harris, nharris@corp.adaptec.com.

Tuesday 10/6 9 AM - 6 PM HIPPI ad hoc
Tuesday 10/6 6 PM - 9 PM HIPPI-6400-OPT
Wednesday 10/7 9 PM - 6 PM HIPPI ad hoc
Wednesday 10/7 6 PM - 8 PM T11.1 Plenary

11.3 Future meeting dates and locations

The T11.1 (i.e., HIPPI), Plenary meeting will be on Wednesday evening of the T11 Plenary week, following the HIPPI working meetings.

Note that T11 schedules the plenary meetings and that the 1998 schedule is firm. Hopefully HIPPI-6400 will be far enough along that we will not continue to need interim working meetings after September. Recent additions and changes are underlined and bold.

The November 17-18, at the SGI facility in Mountain View, CA, is tentative. It was felt that ST would be winding down, but ST-API and the RFC items would be heating up.

1998 -

Sep 1-2 Interim Richardson, TX HP
Oct 6-7 Plenary Ft. Lauderdale, FL Adaptec
Nov 17-18 Interim Mountain View, CA SGI
Dec 14-18 Plenary Tucson FSI

The 1999 and 2000 schedules just include the Plenary weeks; no interim working meetings are scheduled yet. Meeting locations and hosts marked with (?) are tentative at this time. The meetings in bold underline without a (?) have been firmed up. Note that the HIPPI and T11.1 meeting days are not specified; they will be somewhere within the Plenary week.

1999 -

Feb 8-12 Plenary Huntington Beach, CA Qlogic
Apr 12-16 Plenary Palm Springs, CA Brocade
Jun 7-11 Plenary Minneapolis, MN Ancor
Aug 2-6 Plenary Minneapolis, MN ENDL
Oct 4-8 Plenary Ft. Lauderdale, FL Adaptec
Dec 6-10 Plenary Reno, NV Solution

2000 (dates approved, locations and hosts open) -

Feb 7-11	Plenary	San Diego, CA (?)	QLogic
		Austin, TX (?)	Crossroads
Apr 3-7	Plenary	Palm Springs, CA (?)	Brocade
Jun 5-9	Plenary	Boise, ID	HP
Aug 7-11	Plenary	New Hampshire (?)	Hitachi (?)
Oct 2-6	Plenary	San Diego, CA (?)	QLogic
		Austin, TX (?)	Crossroads
Dec 4-8	Plenary	(?)	(?)

10. Jean-Michel Pittet to inform Don Tolmie of the ARP text in HIPPI-LE that should be removed.

11. Don Tolmie to update HIPPI-LE document.

12. Greg Chesson and Jeffrey Chung to consider developing "reason codes" to explain why a particular ST Operation was rejected.

13. Greg Chesson to send e-mail detailing reasons for not doing a queue for client/server applications, and suggesting how they could be done in ST.

14. Bob Willard to write up something on big/little endian issues for inclusion in the document.

15. Greg Chesson to collect text for a "folklore" annex in the document.

16. Greg Chesson to draft text describing how you differentiate duplicate operations from legal operations.

17. Jim Pinkerton to draft a timestamp option for the optional payload for inclusion in the next ST revision.

18. Ian Philp to draft a C code tiling example for inclusion as an annex in ST.

19. Don Tolmie to update ST Rev 2.1 with the changes agreed to at the August meeting.

20. Michael McGowen to collect, tabulate, and document everyone's requirements for HIPPI-800 and HIPPI-6400 translation environments.

21. Don Tolmie to put HIPPI-6400-SC Rev 2.5 on the T11 web page.

12. Review action items

(The action items are grouped by project or category to hopefully make them easier to find.)

1. Von Welch to contact HIPPI-6400 MIB users and developers for comments on the current draft, and to prepare a presentation on the MIB for a future meeting.
2. Von Welch to look at developing a HIPPI-6400 host system MIB (for a NIC), to be done now as an annex of the present MIB with the possibility of splitting it out as a separate document at a later date.
3. Everyone to review the HIPPI-6400 MIB.
4. Jean-Michel Pittet to update "ARP and IP Broadcast over HIPPI-800" Rev 00 with the changes agreed to at the June meeting.
5. Jean-Michel Pittet to request another HIPPI-800 Hardware Type from IANA.
6. Jean-Michel Pittet to forward the request for an IETF Working Group to the IETF.
7. Jean-Michel Pittet to generate an "ARP and IP over HIPPI-6400" document.
8. Marck Doppke to process the HIPPI-800 switch MIB through the IETF.
9. Greg Chesson to draft an IEEE tutorial on HIPPI-6400 ULA usage and the ULAs special to HIPPI-6400.

13. Adjournment

The meeting adjourned at 5 PM on August 11; a HIPPI-6400 Optical meeting followed from 6 PM to 9PM. The planned HIPPI working meeting for August 12 was canceled (mainly due to lack of critical people not in attendance).

Attendance

Ed Cady	Berg	503-359-4556	edcady@aol.com
Arie Van Praag	CERN	41-22-7675034	arie@ch.msm.cern.ch
David Hyer	Compaq Computer Corp.	978-493-6139	david.hyer@digital.com
Donald Woelz	Genroco, Inc.	414-644-2505	don@genroco.com
Henry Brandt	IBM Poughkeepsie	914-433-7684	hrap@kgnvmz.vnet.ibm.com
Don Tolmie	Los Alamos National Lab	505-667-5502	det@lanl.gov
Roger Ronald	Raytheon Systems Co.	972-205-8043	rronald@esy.com
Jeff Young	SGI/Cray Research	651-683-5536	jsy@cray.com
Arun Agarwal	Siemens Microelectronics	408-725-3407	arun.agarwal@smi.siemens.com