

To: T10 Technical Committee
 From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
 Date: 19-20 June 2001
 Subject: Minutes of the SRP teleconference – 19-20 June 2001

Revision History

Revision 0: (19 June 2001) first revision

Agenda

1. Opening remarks and introduction
2. Attendance
3. Approve agenda
4. Review Minutes of SRP teleconference 25 May 2001 (01-178r0) (Rob Elliott)
5. Review old action items from SRP teleconference 25 May 2001
6. SRP LOGIN REJECT proposal (01-171r1) (Cris Simpson)
7. SRP to SAM-2 protocol (01-172r0) (George Penokie)
8. SRP Model for RDMA communication services (01-177r0) (Ed Gardner)
9. SRP communications model proposal (01-186r0) (Cris Simpson)
10. OS Considerations for SRP on IB (01-189r0) (Rob Haydt)
11. SRP InfiniBand annex (01-028r5) (Rob Elliott)
12. Access Controls TransportIDs for SBP, SRP and iSCSI (01-181r0) (Jim Hafner and Rob Elliott)
13. SPC-3 Extended Copy target descriptor for SRP (01-192r0) (Rob Elliott)
14. SRP alias entry designation format (01-193r0) (Rob Elliott)
15. SRP Multichannel proposal (01-085r1) (Cris Simpson)
16. Document review of SRP Revision 6 (srp-r06) (Ed Gardner)
 - 16.1 Editor's notes
 - 16.1.1 Editor's note 1: (Section 4.1) RDMA model rewrite/terminology
 - 16.1.2 Editor's note 2: (Section 4.4.1) "wholly contained"
 - 16.1.3 Editor's note 3: (Section 4.4.2.3) no limit to segments in an indirect buffer
 - 16.1.4 Editor's note 4: (Section 5.1) length for SRP_AER_REQ
 - 16.2 Other issues
 - 16.2.1 (Table 10) Rules for MULTI-CHANNEL ACTION of 01h (independent operation)
 - 16.2.2 SRP_LOGIN_REJ also when initiator refuses the target's LOGIN ACK?
 - 16.2.3 Target port identifier 8 vs. 16 bytes
 - 16.2.4 Reserved fields for multichannel
 - 16.3 General discussion
17. Review new/open action items
18. Meeting schedule
19. Adjourn

Topics

1 Opening Remarks and introduction

The meeting started at 9am Pacific time. Rob Haydt was thanked for hosting.

2 Attendance and membership

Name	S	Organization	Electronic Mail Address
Mr. John Carrier	V	Adaptec	jcarrier@corp.adaptec.com
Mr. Chuck Gibson	V	Adaptec	cgibson@corp.adaptec.com
Mr. David Tanaka	V	Adaptec	dtanaka@corp.adaptec.com

Mr. Kamran Tavakoli	V	Adaptec	ktavakoli@corp.adaptec.com
Mr. Robert C. Elliott	P	Compaq Computer Corp.	Robert.Elliott@compaq.com
Mr. Robert Griswold	V	Crossroads	rgriswold@crossroads.com
Mr. Ralph O. Weber	P	ENDL Texas	roweber@acm.org
Mr. Cris Simpson	P	Intel Corp.	cris.simpson@intel.com
Mr. Rob Haydt	P	Microsoft	robhay@microsoft.com
Mr. Nathan Obr	V	Microsoft	natobr@microsoft.com
Mr. Edward Gardner	P	Ophidian Designs	eag@ophidian.org
Mr. Gerald Maurer	V	QLogic Corp.	gerald.maurer@qlogic.com
Mr. George Penokie	P	Tivoli	gpenokie@tivoli.com

13 People Present

Status Key: P - Principal
A,A# - Alternate
AV - Advisory Member
L - Liaison
V - Visitor

3 Approve agenda

The agenda was approved as listed above.

4 Review Minutes of SRP teleconference 25 May 2001 (01-178r0) (Rob Elliott)

The minutes from the previous meeting were approved.

5 Review old action items from 25 May 2001 teleconference

5.1 Cris Simpson will work with the InfiniBand Trade Association to fix the IOController attributes page in AWG (perhaps replacing it with a new page).

Cris reported that AWG agreed that the fields in question are useless and will recommend to MWG that the fields be made reserved in future releases of InfiniBand specification (after version 1.0a).

5.2 The SRP editor should clarify that the initiator port identifier is also a name that needs to be persistent.

The SAM multiport definition should imply all behavior of import. No change needed.

5.3 Cris Simpson will create a proposed rewrite for section 4.1 (the model section). He will ask Ed for any material he has so far and create a combined or competing proposals for that section.

Still open (later abandoned).

5.4 Cris Simpson will ensure there are enough reserved fields for the future multichannel proposal.

01-085r1 if accepted has the values needed, if not shows what IUs need to grow (LOGIN and CMD).

5.5 Ed Gardner will write a description of the MULTICHANNEL=01h state, both with a formal rule and with examples.

Still open.

5.6 Ed Gardner will incorporate 01-171r1 and 01-173r2, add “per responder” text, add “restricted”, and add additional reason codes in SRP revision 6.

Done.

5.7 Ed Gardner will release SRP revision 6 well in advance of the Redmond meeting.

Done (Not well in advance).

6 SRP LOGIN REJECT proposal (01-171r1) (Cris Simpson)

Ed has already started incorporating this in srp-r06. The incorporated version adds reserved fields to make the IU 32 bytes rather than 16 bytes and increases the REASON field 32 bits. Cris agreed with these changes.

SRP_LOGIN_REJ can only be sent by the target to the initiator. An initiator will not include it in a InfiniBand CM_REJ message rejecting an SRP_LOGIN_RSP. The initiator should do a logout rather than a reject if any problems occur at the SRP level. CM_REJ can still be used for transport reasons but will not include SRP_LOGIN_REJ as PrivateData in this direction.

Cris will prepare 01-171r2 that refers to SRP for the final format.

7 SRP to SAM-2 protocol (01-172r0) (George Penokie)

This annex, adapted from SPI-4, maps the SAM object model onto SRP objects. Revision 1 was started with a few changes before the meeting and was reviewed.

The first paragraph was rewritten.

The table of mappings was removed since SRP uses “initiator port identifier” and “target port identifier” directly; there are no objects left to map.

The procedure objects table was reviewed and some rows were removed. All the remaining objects are defined in SAM-2, raising the issue of whether this table and even the annex belongs in SAM-2 itself rather than in a protocol standard.

One function of the annex is evident in the description of functions like Send-SCSI Command. SAM-2 has more arguments for Send-SCSI Command than SRP implements (namely Command Reference Number). By listing its own table, SRP documents which arguments it supports. The group noted that it needs to show that Autosense_request is required for SRP by not putting it in brackets (SAM-2 lists it as optional). Ralph Weber and George will work on the tables to improve the links to SRP.

George will post rev 1 with the changes made during the meeting and work with Ralph on rev 2 (this week).

8 SRP Model for RDMA communication services (01-177r0) (Ed Gardner)

The text is still written in annex format, although revision 1 will redirect it back into section 4.1.

The “node” in figure A.1 will be removed (both the term and the dotted lines in the picture). No need to provide this level of abstraction here. The InfiniBand annex is free to add that level.

The “channels” in figure A.1 will be labeled.

Should SRP just include a 16 byte opaque memory descriptor and lets the annex for each interconnect define the format (e.g. virtual address, memory handle, and data length)? This might be necessary for mapping SRP over the WARP proposal being discussed in IETF. The group agreed not to make this change at this time. It could be done in SRP-2 if needed without breaking SRP-1.

(whenever done, SRP must still require that each descriptor include a DATA LENGTH field so its definition of TOTAL LENGTH makes sense. The TOTAL LENGTH field is fixed at 32 bits and might also need to be interconnect specific)

Aside: Ed will remove “n+” from the indirect buffer descriptor table.

“can” and “will” will be removed everywhere. “which” will be changed to “that” where appropriate.

QOS will be removed from figure A.2 and following text. Ed will change “channel type” to “channel transport attributes.”

Unordered lists will be formatted as: a) xxx; b) xxx; and c) xxx.

Ordering rule assumptions will be rewritten. RDMA write to RDMA write ordering will be mandated, and the fact that no RDMA read ordering is needed will be mentioned. Rob Haydt would like to ask the IBTA LWG about how RDMA Writes in flight are handled.

Ed will prepare revision 1 for the next teleconference.

9 SRP communications model proposal (01-186r0) (Cris Simpson)

Withdrawn. Cris will provide comments on Ed’s proposal.

10 OS Considerations for SRP on IB (01-189r0) (Rob Haydt)

Rob expressed concerned with plug and play/booting on InfiniBand, particularly with 1000s of nodes, 1000s of LUNs, and bridges to FC and iSCSI all coexisting. He would like to see virtual IO controllers created in the storage devices to represent different views of storage for different initiators. He will work on a proposed addition to the SRP InfiniBand annex for managing IB storage devices.

11 SRP InfiniBand annex (01-028r5) (Rob Elliott)

The group made editorial changes.

The group voted unanimously to incorporate 01-028r5 as revised (into r6) into SRP.

12 Access Controls TransportIDs for SBP, SRP and iSCSI (01-181r0) (Jim Hafner and Rob Elliott)

Part of this proposal defines the Access Control TransportID for SRP as the initiator port identifier.

The group agreed that this is the correct identifier. This proposal will be reviewed in its entirety at the next CAP meeting, as it covers other protocols and other changes and is headed for SPC-3.

13 SPC-3 Extended Copy target descriptor for SRP (01-192r0) (Rob Elliott)

This proposal defines an Extended Copy target descriptor for SRP that contains the target port identifier. The code was corrected from E3h to E5h in the table and the target port identifier was expanded to 16 bytes.

The group recommended at 01-192r0 as revised (to r1) be included in SPC-3. This proposal will be briefly reviewed at the next CAP meeting too, as it is headed for SPC-3.

14 SRP alias entry designation format (01-193r0) (Rob Elliott)

This proposal defines two alias designation formats for SRP: 1) target port identifier only and 2) InfiniBand GID plus target port identifier. These are used by the proposed new ALIAS commands in 00-425.

The group voted 4-4 that 01-193r0 as revised (to r1) be included in SRP (as an annex) if 00-425 passes in July. This means the motion failed. The group would like to see 00-425 approved and resolve the question of whether this material should go into individual protocol standards or SPC-3 before approving this for SRP.

15 SRP Multichannel proposal (01-085r1) (Cris Simpson)

This revision removed support for multiple control channels, only allowing multiple data channels.

The group discussed ordering issues. If the status returned via the control channel reaches the application before RDMA Write data reaches memory over the data channel(s), the application may access the data before it is ready. Rob Elliott suggested a token passing scheme for cases where the RDMA memory server contained the application which actually initiated the transfer; it would pass a token to the initiator which sends it along with the command on the control channel. The target would then use a SEND message after completing the RDMA write to notify the consumer that the data has been sent. Ordering rules would guarantee the data is in memory before the token reaches the receiving application.

Rob Haydt asked that the RDMA to fetch the scatter gather list be limited to the control channel and that the entries in the SG table be allowed to point to different data channels.

Rob Haydt asked about some error scenarios like how data channel disappearance is reported to the initiator over the control channel and what happens if a CA dies after sending an RDMA Ack.

Rob Haydt asked about security and suggested that a target report on its control channel group when an initiator tries to join a data channel group. Rob Haydt suggested that the target not assign session numbers and control channel numbers, instead letting software manage them. Bob Griswold suggested making the data channel logins passive – just indicate data channel login but don't pass an ID. The target would return an identifier that the data-channel initiator then passes to the control-channel initiator, which sends it to the target to indicate which data channel logins are allowed. This lets the control-channel initiator control everything and prevents other clusters from intruding. Cris suggested a mix; let the data login include an identifier but also require the target return an identifier to the control-channel initiator to get the channel activated. Rob Haydt suggested a simple incrementing number from the target that need not be per-initiator based. John Carrier suggested letting the target ask the control-channel initiator for permission, delaying the data channel login until permission has been obtained.

Cris will revise the proposal for the next meeting.

16 Document review of SRP Revision 6 (srp-r06) (Ed Gardner)

16.1 Editor's notes

16.1.1 Editor's note 1: (Section 4.1) RDMA model rewrite/terminology

This is being covered with 01-177.

16.1.2 Editor's note 2: (Section 4.4.1) "wholly contained"

This section will be reworded.

16.1.3 Editor's note 3: (Section 4.4.2.3) no limit to segments in an indirect buffer

Ed will remove the note.

16.1.4 Editor's note 4: (Section 5.1) length for SRP_AER_REQ

Ed will choose a resolution for this. Probably the shortest possible IU length will be listed, ignoring the fact that 18 bytes of sense data is always required.

16.2 Other issues

16.2.1 (Table 10) Rules for MULTI-CHANNEL ACTION of 01h (independent operation)

Ed has an action item to write this description.

16.2.2 SRP_LOGIN_REJ also when initiator refuses the target's LOGIN ACK?

See discussion of 01-171.

16.2.3 Target port identifier 8 vs. 16 bytes

The group offered no complaints to the increase in target port identifier size to 16 bits introduced in srp-r04.

16.2.4 Reserved fields for multichannel

From discussion of Cris' multichannel proposal, the complete set of fields needed is not yet settled. Keeping fields reserved in the current IUs may not matter if multichannel goes with SRP-2 and the driver recognizes SRP-1 vs. SRP-2.

16.3 General discussion

Ed will prepare a comprehensive list of comments and resolutions for srp-r07. Some are listed here:

The only italics that should appear are the ISO number in the references section (not the name, just the number).

Boxes around figures are prohibited by ISO.

Add SRP before every request and response.

Control, data, and control/data channel terminology will be introduced (with notes on their restrictions).

Cris Simpson, Rob Haydt, and Bob Griswold requested an initiator LOGOUT IU to notify the target that a logout is going to occur. Rob wants the target to flush outstanding activity and return a LOGOUT_RSP. Others in the group felt the target should just disconnect; the initiator should let its commands finish before issuing the logout if it cares. Cris will write a proposal detailing this.

The group agreed not to add an SRP version number to LOGIN_REQ and LOGIN_RSP. For InfiniBand the IOC profile include a version that can be used to load the correct driver. Other protocols may have similar features. The group agreed that future SRP-n releases must always support the basic SRP-1 IUs so an INQUIRY command can be sent to retrieve version descriptors.

17 Review new/open action items

Ed Gardner will write a description of the MULTICHANNEL=01h state, both with a formal rule and with examples.

Ed Gardner will release SRP revision 7 well in advance of the 7 July teleconference.

Rob Haydt will write a proposal about detecting storage IOCs in InfiniBand.

Cris Simpson will write a proposal for a LOGOUT IU from the initiator to the target.

18 Meeting schedule

No letter ballot is possible until Monday 9 July since John Lohmeyer, chair of T10, is on vacation. If issued before July T10 week the meeting time would not be usable, so the group agreed to defer the letter ballot until after T10 week. A teleconference call is set up one week prior (after 4th of July week) to make sure work is proceeding.

Interim teleconference:

9 July 2001 Mon 8:00 am-10:00am Pacific teleconference (Rob Elliott, Compaq hosting)

Topics: review revision 7 and updated proposals

T10 Week:

18 July 2001 Wed Colorado Springs CAP

19 July 2001 Thu Colorado Springs SRP (begins 1 hour after T10 plenary, usually around 1:00 pm

20 July 2001 Fri Colorado Springs SRP

Topics: review revision 8 for release to letter ballot

19 Adjournment

The meeting adjourned at 2:00pm Wednesday.