To: T10 Technical Committee
From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
Date: 4-5 January 2001
Subject: Minutes of the SRP WG – January 4-5, 2001 – Houston, TX

Revision History
Revision 0: first revision

Attendance

```
        Name           S       Organization        Electronic Mail Address
---------------------- -- ------------------------- ------------------------
Mr. William Lynn       V  Adaptec, Inc.             bill_lynn@corp.adaptec.com
Mr. Robert C. Elliott  P  Compaq Computer Corp.     Robert.Elliott@compaq.com
Mr. Scott Dawson       V  Compaq Computer Corp.     scott.dawson@compaq.com
Mr. Joe Foster         V  Compaq Computer Corp.     Joe.Foster@compaq.com
Mr. Greg Pellegrino    V  Compaq Computer Corp.     greg.pellegrino@compaq.com
Mr. Bob Nixon          O  Emulex                    bob.nixon@emulex.com
Mr. Ralph O. Weber     P  ENDL Texas                roweber@acm.org
Mr. George O. Penokie  P  IBM / Tivoli Systems      gpenokie@tivoli.com
Mr. Cris Simpson       V  Intel Corp.               cris.simpson@intel.com
Mr. John Lohmeyer      P  LSI Logic Corp.           lohmeyer@t10.org
Mr. Tim Hoglund        V  LSI Logic Corp.           tim.hoglund@lsil.com
Mr. Keith Holt         V  LSI Logic Corp.           keith.holt@lsil.com
Mr. Konrad Lei         V  McData Corp.              klie@mcdata.com
Mr. Edward A. Gardner  P  Ophidian Designs          eag@ophidian.com
Mr. Frank Campbell     V  QLogic                    frank.campbell@qlc.com
Mr. Bill Terrell       V  Troika Networks           terrell@troikanetworks.com


16 People Present
Status Key:  P   -  Principal
             A,A# -  Alternate
             O   -  Observer
             L   -  Liaison
             V   -  Visitor
```

Results of meeting
John Lohmeyer opened the SCSI over RDMA protocol meeting at 9am Thursday 4 January 2001 and thanked Compaq for hosting. This protocol standard maps SCSI over InfiniBand™ Architecture, Virtual Interface (VI) Architecture, and similar transports supporting RDMA (remote direct memory access).

Agenda
The agenda was created at the meeting.

1. Scatter/Gather lists (part of SRP revision 2 by Ed Gardner)
2. Identifying devices/addressing (part of SRP revision 2 by Ed Gardner)
3. Proposed InfiniBand annex (01-028r0 by Greg Pellegrino and Rob Elliott)
4. Document review (all other topics; SRP revision 2 by Ed Gardner)
5. Meeting schedule
6. Adjournment

Topics
1. Scatter/Gather lists (part of SRP revision 2 by Ed Gardner)
Ed Gardner merged the scatter-gather proposal (00-410r0) with the changes discussed at the last SRP meeting into SRP revision 2 (distributed at this meeting). He presented the SG contents of SRP revision 2.

Cris Simpson suggested renaming Region Handle to Memory Handle.

Discussion of the text stating that the initiator shall keep its RDMA requests within the virtual address + data length window.   Should the target be required to only issue addresses within the window for its RDMAs? The group consensus was yes.   Should the initiator be required to check the inbound addresses? The group consensus was no.

Ed will replace "virtually contiguous" with "contiguous within the region's virtual address space" throughout the document.

Ed will label some additional components in figure 5 so they can be referenced from the text. Ed will add another figure showing the direct mapping case.

Several editorial issues were discussed.
• parenthesis on MSB, LSB
• small caps mixed with caps
• RESERVED being small caps or not
• MSB/LSB on the TAG field?
• MSB/LSB on the CDB field?
Ed will try to match SPC-2's editorial style as much as possible.

For each IU payload table, Ed will add the value next to the word TYPE (e.g. TYPE (02h)). This makes the style similar to that used for opcodes in SPC-2.

This phrase will be used to describe the DOIND and DOIND bits: "DOIND specifies that data out buffer descriptor, if present, defines a direct data buffer."

Ed will make a table of the DOIND, DIIND, DATA OUT MEMORY DESCRIPTOR COUNT, and DATA IN MEMORY DESCRIPTOR COUNT fields in section 4.7, so one set of descriptions is shared by all the command IUs.

Rev 2 does not require that initiators use the smallest possible CDB IU as agreed in the last meeting. A 6-byte CDB can be placed in the 32-byte CMD IU. The group accepted this change. Text will be added noting that the target may accept or reject (at the SRP level) an IU that contains a small CDB.

There was discussion whether the 16-byte command IU should be dropped, leaving just 32-byte and variable length command IUs. The group decided to keep them all. The 32-byte IU needs 80 bytes, assuming one data pointer. The 16 byte IU fits in 64 bytes.

Should SRP_LOGIN_REQ and SRP_LOGIN_RSP use different type values? Revision 2 has them both using 00h. Since they travel in different directions, there should be no confusion about which IU is being transmitted. However, Ed agreed to make them different for clarity.

Ed noted that, for InfiniBand, he envisioned private data on the connection establishment being used to communicate the login IUs.

2. Identifying devices/addressing   (discussion led by George Penokie)
At the last meeting, Ed Gardner led discussion of this topic using 00-357r1. This meeting, no document was used.

George Penokie started discussion by providing a background on the hierarchical LUN structure defined in SCC-2 and SAM-2. If the HiSup bit is set in INQUIRY the 8-byte LUN field is split into four 2-byte sub-fields, each corresponding to a different level of nested busses. These sub-fields map into target and LUN identifiers for the nested busses. Each sub-field uses either a

Peripheral, Logical Unit, and Volume Set Addressing mode. As the LUN field is passed from level to level, the contents are shifted up by two bytes with zeros shifted in at the end.

For peripheral addressing, the lowest level bus has to contain only LUN 0s due to the zero-fill when shifting.

The rule in SAM-2 and SCC-2 that says "The received command shall be relayed to LUN zero" appears incorrect. That is only required when relaying to an SCC device, not a non-SCC device. George will consider this problem further.

SCSI targets include VPD page 83 device identifiers which includes a 64 or 128-bit unique identifier. 128-bit comes from the FC-FS-2 registered extended format; the other formats are all 64 bits. See T10/97-101r2 and T11/FC-FS-2 revision 0.50 for more information on the formats.

The group discussed how Fibre Channel and InfiniBand addressing work.
* Each Fibre Channel port has a 24-bit address. Each frame contains both the source and destination address (Source_ID and Destination_ID) which are used to route within the fabric.

* Each InfiniBand port has a 16-bit Local ID (LID) assigned by a local subnet manager. Each packet contains a Source LID and a Destination LID which are used to route within subnets. Additionally, each InfiniBand port has a 128-bit Global ID (GID) that is an IPv6 address. Packets optionally contain a global route header which specifies the Source GID and Destination GID that are used to route packets between subnets. InfiniBand packets also identify a 24-bit Destination Queue Pair Number.

The group discussed InfiniBand SRP to parallel SCSI, InfiniBand SRP to Fibre Channel FCP, and InfiniBand SRP to iSCSI bridges, discussing whether the hierarchical LUN structure could be used. A LUN mapping bridge needs to intercept REPORT LUNS commands, TARGET RESET task management functions, and other commands that affect the target rather than the logical unit.

George suggested resurrecting an old SCSI concept called target routines to avoid overloading LUN 0 with functions like REPORT LUNS. This would signal, outside of the CDB, that a command (or task management request) is destined for the target rather than the logical unit. George may bring a proposal to CAP to handle this. SRP would be the first protocol to support it.

3. Proposed InfiniBand annex   (01-028r0 by Greg Pellegrino and Rob Elliott)
The group reviewed a proposed annex detailing the SRP mapping over InfiniBand. Changes discussed at the meeting will be contained in 01-028r1.

There is currently no SRP logout IU; a device just issues an InfiniBand level disconnect to release a connection. The group discussed whether an SRP logout IU should be added. Since it travels over the reliable connection, it would be reliable. The DREQ used to disconnect in InfiniBand is not. The logout would be guaranteed in-order, which DREQ is not. Should the logout (request) IU be accompanied by a logout response IU? Ed Gardner and George Penokie felt it was not necessary. The group also discussed documenting a general purpose IU reject for targets to send to initiators that are not logged in. Ed will write a proposal for target-initiated SRP logout and document the general purpose rejection functionality.

Ed had originally planned to write a VI annex (e.g. for SRP over VI). He will probably not do so unless requested. It will not make it into SRP-1 unless requested soon.

SRP needs another annex showing the mapping of SRP terms to SAM terms. Ed will create this based on FCP-2.

The definitions of initiator and target identifiers is still unclear. For example, an initiator identifier is used by persistent reservations. If the initiator identifier includes the queue pair number, then it won't be persistent through power loss. If the initiator identifier includes a port GUID, it won't be persistent through CA replacements. If the initiator identifier includes a port LID, it won't persist through subnet resets. If the initiator identifier includes a port GUID, InfiniBand's alternate path failover features will not work correctly. The initiator's Node GUID (apparently being renamed the xCA GUID by the InfiniBand Trade Association) may be the most stable identifier to use.

4. Document review   (SRP Revision 2 by Ed Gardner)
No document review was done. Revision 2 should be posted on the T10 web site soon. It includes the scatter-gather list details, but does not yet include 00-426r1 (Mode pages) or 01-010r1 (AER) which were approved at the previous meeting.

5. Meeting schedule
The next T10 week in Orlando, FL at the Grosvenor Resort, hosted by Adaptec, will have special SRP meetings in addition to the usual CAP meeting.

17 January 2001 Wed CAP (SRP-specific items NOT to be discussed)
18 January 2001 Thu SRP (begins 1 hour after T10 plenary, usually around 2pm)
19 January 2001 Fri SRP 9am – 1pm

Additionally, InfiniBand Application Working Group members are meeting in San Francisco on these dates:
24 January 2001 Wed IBTA AWG
25 January 2001 Thu IBTA AWG
26 January 2001 Fri IBTA AWG

Ed Gardner originally intended to request a letter ballot after the 19 Jan meeting.

6. Adjournment
The meeting adjourned around 1pm Friday.
.