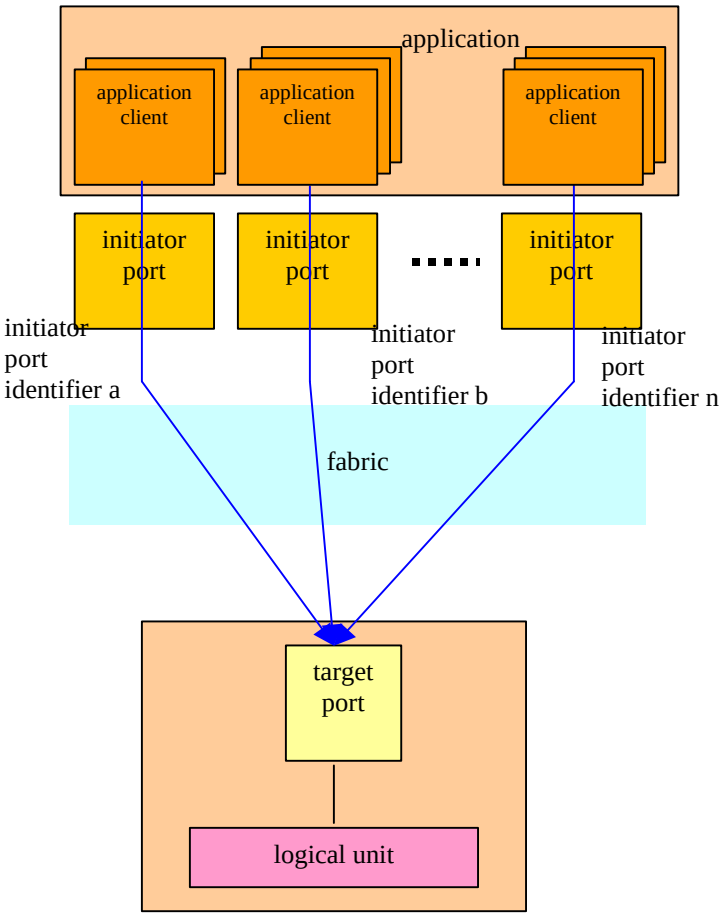
To: T10 Technical Committee
From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
Date: 3 January 2002
Subject: T10/01-100r3 Letting persistent reservations ignore initiator ports

Revision History
Revision 0 (7 March 2001). First revision. Not complete, but as presented in CAP working group.
Revision 1 (24 April 2001). Reflects input from the March CAP working group. Settles on separate PR OUT service action to register an initiator ID.
Revision 2 (18 October 2001): Incorporated input from May CAP working group. Changed "initiator device name" to "application name." Brought up to date with 01-099r4.
Revision 3 (3 January 2002): Incorporated input from November CAP working group. Changed to an explicit list of initiator port identifiers rather than a higher level identifier.

Related Documents
T10/spc3r02 – SCSI Primary Commands 3 revision 2 (by Ralph Weber)
T10/00-232r9 – Asymmetric target behavior (by Ken Moe)
T10/01-099r5 - SPC-3 Letting persistent reservations ignore target ports (by Rob Elliott)
T10/01-268r2 - Access Controls rewrite (by Ralph Weber and Jim Hafner)

Overview
SPC-2 requires that logical units remember the initiator port through which a reservation was made. An application wishing to register needs to run the PERSISTENT RESERVE OUT command through each initiator port that can route to the logical unit.



In many cases, the logical unit doesn't care which initiator port the command came through. It just wants to distinguish between applications. Requiring reservation commands for each initiator port burdens the applications with issuing extra commands and burdens the logical unit with extra non-volatile storage. Cases include a single host with multiple ports and multiple hosts in a cluster.

In some clusters of high-end systems, there may be hundreds of initiator ports per node (these machines have lots of PCI busses each holding HBAs), dozens of initiator nodes (8-100 node clusters), all talking to a logical unit through a handful of target ports (4 port targets are common). For example, there may be 256 * 64 * 4 = 65,536 different I_T paths between the same application and logical unit. For an all-registrants type of reservation, this results in 65,536 registration commands, each of which may be slow and require media access.

This proposal suggests passing initiator port identifiers explicitly in the command to register many initiator ports at one time. It borrows the TransportID data structure from access controls to identify the initiators.

Suggested Changes

Changes are applied to SPC-3 revision 2 as modified by 01-099r5 and 01-268r2. Only directly changed areas are excerpted.

[Editor's note: these sections are from 01-099r5]

Table 73. REPORT CAPABILITIES parameter data

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)							
1				LENGTH (0008h)				(LSB)
2				Reserved	SPECIFY INITIATORS CAPABLE	ALL TARGET PORTS CAPABLE	ELEMENT SCOPE CAPABLE	APTPL CAPABLE
3				Reserved			Rsvd	PTPL
4				Reserved				
5				Reserved				
6				Reserved				
7				Reserved				

A SPECIFY INITIATORS CAPABLE bit of one indicates that the device server supports the SPECIFY INITIATORS bit in the PERSISTENT RESERVE OUT command. A SPECIFY INITIATORS CAPABLE bit of zero indicates that the device server does not support the SPECIFY INITIATORS bit in the PERSISTENT RESERVE OUT command.

7.11.3 PERSISTENT RESERVE OUT parameter list

The parameter list required to perform the PERSISTENT RESERVE OUT command is defined in table 75. All fields shall be sent on all PERSISTENT RESERVE OUT commands, even if the field is not required for the specified service action and scope values.

Table 75 — PERSISTENT RESERVE OUT parameter list

Byte\Bit	7	6	5	4	3	2	1	0
0	RESERVATION KEY							
7								
8	SERVICE ACTION RESERVATION KEY							
15								
16	SCOPE-SPECIFIC ADDRESS							
19								
20	Rsvd	Rsvd	Rsvd	Rsvd	SPECIFY INITIATORS	ALL TARGET PORTS	Rsvd	APTL
21	Reserved							
22	Obsolete							
23								
24	Additional parameter data							
n								

[Editor's note: Add this paragraph among the other field descriptions:]
If the SPECIFY INITIATORS bit is set to one for the REGISTER or REGISTER AND IGNORE EXISTING KEY service actions, the parameter data shall include a list of transport IDs as specified in table xx and the device server shall apply the registration to each initiator specified by a TransportID. If the SPECIFY INITIATORS bit is set to zero, the device server shall not include additional parameter data and shall apply the registration only to the initiator port that sent the PERSISTENT RESERVE OUT command.

Table xx — PERSISTENT RESERVE OUT additional parameter data

Byte\Bit	7	6	5	4	3	2	1	0
24	TRANSPORTID PARAMETER DATA LENGTH (n - 24)							
27								
28	One or more TransportIDs							
n								

The TRANSPORTID PARAMETER DATA LENGTH indicates how many bytes are consumed by TransportIDs.

See 8.99.99 for definitions of TransportIDs.

Table 76 summarizes which fields are set by the application client and interpreted by the device server for each service action and scope value.

[Editor's note: add another column to the table:]

Table 76 — PERSISTENT RESERVE OUT service actions and valid parameters

Service action	Allowed SCOPE	SPECIFY INITIATORS
REGISTER	ignored	V
REGISTER AND IGNORE EXISTING KEY	ignored	V
RESERVE	LU_SCOPE	I
	ELEMENT_SCOPE	I
RELEASE	LU_SCOPE	I
	ELEMENT_SCOPE	I
CLEAR	ignored	I
PREEMPT	LU_SCOPE	I
	ELEMENT_SCOPE	I
PREEMPT & ABORT	LU_SCOPE	I
	ELEMENT_SCOPE	I

Key: V=valid I=ignored

[Editor's note: this section is from 01-268r2]

9.3.1.3.2 Access identifiers

9.3.1.3.2.1 Access identifiers overview

Initiators are identified in an ACE using one of the following types of access identifiers:

a) AccessID - based on initiator enrollment as described in 9.3.1.3.2.2;

b) TransportID - based on protocol specific identification of initiators as described in ~~9.3.1.3.2.3~~8.99.99; or

c) vendor specific access identifiers.

[Editor's note: change all references to 9.3.1.3.2.3 (TransportID access type identifiers) to 8.99.99 (TransportIDs) since the former is being deleted. I found only five in 01-268r0 including the one changed above.]

9.3.1.3.2.2 AccessID access identifiers

AccessID access identifiers shall have the format shown in table t3.

Table t3 — AccessID access identifier format

…

~~**9.3.1.3.2.3 TransportID access type identifiers**~~

~~Use of the TransportID is protocol specific.~~

~~An initiator is identified by a TransportID if that initiator accesses the SCSI target device with that TransportID. At any given time, an initiator may be identified or associated with at most one TransportID.~~

~~TransportIDs (see table t12) shall be at least 24 bytes long and shall be a multiple of four bytes in length.~~

~~**Table t4 — TransportID format**~~

~~Byte\Bit~~	~~7~~	~~6~~	~~5~~	~~4~~	~~3~~	~~2~~	~~1~~	~~0~~
~~0~~	~~Reserved~~				~~PROTOCOL IDENTIFIER~~			
~~1~~	~~SCSI protocol specific data~~							
~~n>22~~								

~~The PROTOCOL IDENTIFIER field (see table t56 in 8.99.1) identifies the SCSI protocol to which the TransportID applies.~~

~~The format of the SCSI protocol specific data depends on the value in the PROTOCOL IDENTIFIER field. The SCSI protocol specific data in a TransportID shall only include initiator port identifiers or names (see SAM-2) that persist across common reset events in the service delivery subsystem. TransportID formats specific to SCSI protocols are listed in table t5.~~

~~Table t5 — TransportID formats for specific SCSI protocols~~

~~SCSI protocol~~	~~Protocol Standard~~	~~Reference~~
~~Fibre Channel~~	~~FCP-2~~	~~8.99.99.1~~
~~Parallel SCSI~~	~~SPI-4~~	~~8.99.99.2~~
~~IEEE 1394~~	~~SBP-2~~	~~8.99.99.3~~
~~Remote Direct Memory Access (RDMA)~~	~~SRP~~	~~8.99.99.4~~
~~Internet SCSI~~	~~iSCSI~~	~~8.99.99.5~~
~~Reserved~~		

[Editor's note: this section is from 01-268r2]
8.99 Protocol specific parameters
8.99.1 Protocol specific parameters introduction
Some commands use protocol specific information in their CDBs or parameter lists. This subclause describes those protocol specific parameters.

Protocol specific parameters may include a PROTOCOL ~~IDENTIFIER~~ CODE field (see table t56) as a reference for the SCSI protocol to which the protocol specific parameter applies.

Table t56 — PROTOCOL IDENTIFIER values

Protocol ~~Identifier~~Code	Description	Protocol Standard
0h	Fibre Channel	FCP-2
1h	Parallel SCSI	SPI-4
2h	SSA	
3h	IEEE 1394	SBP-2
4h	Remote Direct Memory Access (RDMA)	SRP
5h	Internet SCSI	iSCSI
6h - Fh	Reserved	

8.99.99 ~~Access controls~~ TransportIDs ~~access identifiers~~
8.99.99.0 TransportID overview
An initiator is identified by a TransportID if that initiator accesses the SCSI target device with that TransportID. At any given time, an initiator may be identified or associated with at most one TransportID.

TransportIDs (see table t12) shall be at least 24 bytes long and shall be a multiple of four bytes in length.

Table t4 — TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE			
1	SCSI protocol specific data							
n>22								

The PROTOCOL CODE field (see table t56 in 8.99.1) identifies the SCSI protocol to which the TransportID applies.

The format of the SCSI protocol specific data depends on the value in the PROTOCOL IDENTIFIER field. The SCSI protocol specific data in a TransportID shall only include initiator port identifiers or names (see SAM-2) that persist across common reset events in the service delivery subsystem. TransportID formats specific to SCSI protocols are listed in table t5.

Table t5 — TransportID formats for specific SCSI protocols

SCSI protocol	Protocol Standard	Reference
Fibre Channel	FCP-2	8.99.99.1
Parallel SCSI	SPI-4	8.99.99.2
IEEE 1394	SBP-2	8.99.99.3
Remote Direct Memory Access (RDMA)	SRP	8.99.99.4
Internet SCSI	iSCSI	8.99.99.5
Reserved		

TransportIDs are initiator identifiers used as:
a) access identifiers (see 9.3.1.3.2) used in ACL ACEs to allow access to a logical unit from an initiator; and
b) initiator identifiers used in parameter data for the REGISTER and REGISTER AND IGNORE EXISTING KEY service actions of the PERSISTENT RESERVE OUT command when the SPECIFY INITIATORS bit is set to one.

8.99.99.1 TransportID~~s~~ for initiators using SCSI over Fibre Channel

A Fibre Channel TransportID (see table t57) ~~is a type of access identifier (see 9.3.1.3.2) used in ACL ACEs to allow logical unit access to a~~ identifies an FCP-2 initiator based on the world wide unique initiator port name belonging to that initiator.

Table t57 — Fibre Channel TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE (0h)			
1	~~(MSB)~~	Reserved						
7								~~(LSB)~~
8	(MSB)	WORLD WIDE NAME						
15								(LSB)
16	~~(MSB)~~	Reserved						
23								~~(LSB)~~

The WORLD WIDE NAME field shall contain the port World Wide Name defined by the Physical Log In (PLOGI) extended link service, defined in FC-FS.

~~A Fibre Channel TransportID allows the initiator specified by the world wide name access to the logical units described in an ACE (see 9.3.1.3).~~

8.99.99.2 TransportID~~s~~ for initiators using a parallel SCSI bus

A parallel SCSI bus TransportID~~s~~ (see table t58) ~~is a type of access identifier (see 9.3.1.3.2) used in ACL ACEs to allow logical unit access to~~identifies a SPI-4 initiator based on the SCSI address of an initiator and the SCSI target device relative port through which the initiator accesses the SCSI target device.

Table t58 — Parallel SCSI bus TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE (1h)			
	Reserved							
1	(MSB)	SCSI ADDRESS						
7								(LSB)
8	(MSB)	RELATIVE PORT IDENTIFIER						
15								(LSB)
16	~~(MSB)~~	Reserved						
23								~~(LSB)~~

The SCSI ADDRESS field specifies the SCSI address (see SPI-4) of the initiator.

The RELATIVE PORT IDENTIFIER field specifies the four-byte binary number identifying a specific port in the SCSI target device relative to other ports. The relative port identifier value shall be one of the values returned in the Device Identifier VPD page (see 8.z.z). If the RELATIVE PORT IDENTIFIER does not reference a port in the device, the TransportID is invalid.

~~In order for a parallel SCSI bus TransportID to allow access to the logical units described in an ACE (see 9.3.1.3), an initiator having the specified SCSI address shall access the SCSI target device via the port specified by the relative port identifier.~~

8.99.99.3 TransportID~~s~~ for initiators using SCSI over IEEE 1394

An IEEE 1394 TransportID (see table t59) ~~is a type of access identifier (see 9.3.1.3.2) used in ACL ACEs to allow logical unit access to~~identifies a SBP-2 initiator based on the EUI-64 initiator port name belonging to that initiator.

Table t59 — IEEE 1394 TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE (3h)			
1	~~(MSB)~~	Reserved						
7								(~~LSB~~)
8	(MSB)	EUI-64 NAME						
15								(LSB)
16	~~(MSB)~~	Reserved						
23								~~(LSB)~~

The EUI-64 NAME field shall contain the EUI-64 IEEE 1394 node unique identifier (see SBP-2) for an initiator port.

~~A IEEE 1394 TransportID allows the initiator specified by the EUI-64 node unique identifier access to the logical units described in an ACE (see 9.3.1.3).~~

8.99.99.4 TransportID~~s~~ for initiators using SCSI over an RDMA interface

An RDMA TransportID (see table t60) ~~is a type of access identifier (see 9.3.1.3.2) used in ACL ACEs to allow logical unit access to~~identifies an SRP initiator based on the world wide unique initiator port name belonging to that initiator.

Table t60 — RDMA TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE (4h)			
1	~~(MSB)~~	Reserved						
7								~~(LSB)~~
8	(MSB)	INITIATOR PORT IDENTIFIER						
23								(LSB)

The INITIATOR PORT IDENTIFIER field shall contain an SRP initiator port identifier (see SRP).

~~A RDMA TransportID allows the initiator specified by the initiator port identifier access to the logical units described in an ACE (see 9.3.1.3).~~

8.99.99.5 TransportID~~s~~ for initiators using SCSI over Internet SCSI

A iSCSI TransportID (see table t61) ~~is a type of access identifier (see 9.3.1.3.2) used in ACL ACEs to allow logical unit access to~~identifies an iSCSI initiator based on the world wide unique initiator port name belonging to that initiator.

Table t61 — iSCSI TransportID format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved				PROTOCOL CODE (5h)			
1	~~(MSB)~~	Reserved						
3								~~(LSB)~~
4	(MSB)	ISCSI NAME						
n								(LSB)
n+1		NULL/PAD						
m								

The ISCSI NAME field shall contain the iSCSI name of an iSCSI initiator node (see iSCSI). The ISCSI NAME field shall not contain a byte set to 00h. The first byte containing 00h after byte 4 terminates the ISCSI NAME field without regard for the specified length of the iSCSI TransportID.

The NULL/PAD field shall contain between one and four bytes set to 00h. The length of the NULL/PAD field shall be chosen so that the total length of the iSCSI TransportID (m+1) is a multiple of four.

~~A iSCSI TransportID allows the initiator specified by the world wide name access to the logical units described in an ACE (see 9.3.1.3).~~

NOTE 13 - The maximum length of the iSCSI TransportID is 260 bytes because the iSCSI name length does not exceed 255 bytes.