TO: T10 Membership
FROM: Paul A. Suhler, Seagate Removable Storage Solutions, LLC
DATE: 10 February 2003
SUBJECT: T10/03-077r0, SCSI Surrogate Proposal

Specification of SCSI Surrogate Operation requires changes to:

- ADC model to describe surrogate operation
- ADC automation drive descriptor to select different surrogate modes
- ADT link services clause to add process login and logout
- ADT SCSI Command frame payload to add an I_T nexus identifier

FCP-2 doesn't have an entry in the definitions clause for process login, so I haven't provided one here.

Thanks to Paul Entzel for providing large parts of this text. I wordsmithed some of his text, so blame me for any problems.

I still have a number of questions about how things should be done; see the editor's notes.

The following is to be added to the ADC model section:

Editor's Note: The general feature is "surrogate" operation and the two variants are "bridged" and "passthrough." Are these names sufficiently descriptive or should they be changed? Comments?

4.2.x SCSI surrogate operation

The data transfer device may optionally support one or more forms of SCSI surrogate operation for the automation device. When this operation is enabled, the data transfer device reports a logical unit to its primary interface ports that represents the SMC device server. Commands addressed to that logical unit are passed to the SMC device server through the ADT transport. This is typically used in low-cost automation devices that do not have separate primary interface ports.

A data transfer device that supports the SCSI surrogate feature may implement a device server that can support a subset of the SMC command set. For instance, a REPORT LUNS may be routed to an internal device server instead of being passed to the SMC device. This may be necessary if the internal device server has direct knowledge of the other LUNs supported by the target device. The term “local SMC device server” is used to refer to this partial device server resident in the data transfer device. The Medium Changer descriptor in the ADC Target Device Control mode page allows for several types of SCSI surrogate operation to be enabled in the data transfer device. Support for any particular type of SCSI surrogate operation is optional.

4.2.x.1 Bridged surrogate operation

In bridged surrogate mode, the SMC device server is partitioned between the data transfer device and the automation device. No indication of the original initiator port of the SCSI command is passed between the data transfer and automation devices. In this mode of operation, the data transfer device contains a device server capable of servicing commands and task management functions that require knowledge of the originating initiator port. Effectively, the data transfer device acts as a protocol bridge in this mode.

If any of the following commands are supported, they shall be serviced by the device server in the data transfer device and not passed through to the SMC device:

a) RESERVE and RESERVE (10)
b) RELEASE and RELEASE (10)
c) PERSISTENT RESERVE IN

d) PERSISTENT RESERVE OUT
e) REPORT LUNS
f) REQUEST SENSE

The local SMC device server shall also perform the following actions:

a) Check for reservation conflicts on all commands routed to the SMC device server. Return RESERVATION CONFLICT on all commands that violate a reservation condition and not pass them through to the SMC device.
b) Manage UNIT ATTENTION conditions generated for multiple initiators. Before passing a command through to the SMC device, pending UNIT ATTENTION conditions shall be checked. If the device server detects a UNIT ATTENTION condition is pending for an initiator port when a new command is received from it, the command device server shall return Check Condition for the command without passing it to the SMC device.
c) When the primary interface uses contingent allegiance, sense data must be saved on a per initiator port basis.

4.2.x.2 Passthrough surrogate operation

In passthrough operation, all of the SMC device server functions are implemented in the automation device. The Process Login link service is transmitted by the data transfer device to the automation device to notify the SMC device server that an I_T nexus has been established between an Initiator Port and the data transfer device's Target Port. The Process Login shall be performed prior to transmission of a request from the data transfer device to the automation device if no login exists.

Editor's Note: Is reporting Logical Unit Communication Failure mandatory? If so, then we will need a definition of when it has occurred.

4.2.x.3 Caching SMC data

In some implementations, the local SMC device server may cache some data from the SMC device. For instance, the local device server may save the INQUIRY data from the SMC device and return it to any initiator port that requests it. The ADT standard provides means for the SMC device to notify the local SMC device server when data that may be cached may have changed. The local SMC device server should discontinue using cache data from the SMC device once it is informed of a possible change until such time as it can be refreshed.

The following change is for ADC clause 6.2.2.3.3, Medium Changer descriptor parameters.

- The one-bit ENABLE field is expanded into a three-bit SURROGATE MODE field and supporting text and a table is added.

Table 29 – Medium changer descriptor

Bit / Byte	7	6	5	4	3	2	1	0
6	Reserved					SURROGATE MODE		
7	Reserved							

[Existing paragraph describing the enabled field is replaced by the following.]

The SURROGATE MODE field determines which form, if any, of surrogate mode operation is performed by the data transfer device. When it is set to zero, the logical unit is not reported in response to a REPORT LUNS command and does not respond to commands. If it is set to a

non-zero value, the Logical Unit is reported and supported. Commands received for this logical unit shall be passed on to the automation using the ADI interface. The valid values of the SURROGATE MODE field are shown in Table x:

Table x - Surrogate mode values

Surrogate Mode	Description
0	Disabled
1	Passthrough
2	Bridged
3 – 7	Reserved

Editor's Note: I considered only using one (ENABLE) bit as before, but that allows no way for automation expecting to use Bridged mode to find out that the drive only supports Passthrough mode. This way, the drive can reject the MODE SENSE and point to the SURROGATE MODE field.

If the SURROGATE MODE field is set to Passthrough, then the data transfer device shall perform a process login with the automation device for each initiator from which it receives commands for the medium changer logical unit. Each SCSI command frame shall contain the I_T nexus identifier value established in that initiator's process login.

If the SURROGATE MODE field is set to Bridged, then the data transfer device shall not perform a process login with the automation device. Each SCSI command frame shall contain a value of zero in its I_T NEXUS IDENTIFIER field.

If the SURROGATE MODE field is changed from any other value to Disabled, then the data transfer device shall implicitly abort all commands issued to the automation device and shall report a status of CHECK CONDITION with a sense key of COMMAND ABORTED and an additional sense code of I_T NEXUS LOSS OCCURRED for each command. An implicit process logout shall be performed.

Editor's Note: This last paragraph describes a corner case that shouldn't happen in normal operation. Should we report status on each command as written here, or just let the initiators figure it out when we start reporting logical unit not supported to subsequent commands? In transports like Fibre Channel that have the equivalent of REC, the initiator will first see the problem there.

The following changes are for ADT clause 6.5, Link service frames:

- Add two new values to Table 4.
- Add descriptive text and tables for process login and logout.

Table 4 – Link service frame types

Payload Type	Description
0h	ACK (acknowledge)
1h	NAK (negative acknowledge)
2h	Port login
3h	Port logout
4h	Pause
5h	NOP (no operation)
6h	Initiate recovery
7h	Process login
8h	Process logout
9h – Fh	Reserved

6.5.9 Process login

Process login frames are sent by the data transfer device to the automation device to identify an initiator port to the SMC device server. Process login frames are sent by the automation device to the data transfer device to accept or reject the process login.

The data transfer device shall perform port login with the automation device prior to performing process login. If the automation device transmits a Port Logout frame, then all process logins are implicitly terminated.

Table x defines the payload of the Process Login frame:

Table x – Process login frame payload contents

Bit / Byte	7	6	5	4	3	2	1	0
0	Rsvd	ACCEPT	Reserved					
1	I_T NEXUS IDENTIFIER							
2	INITIATOR PORT IDENTIFIER LENGTH (m)							
3	INITIATOR PORT NAME LENGTH (n)							
4	(MSB)	INITIATOR PORT IDENTIFIER						
3+m								(LSB)
4+m	(MSB)	INITIATOR PORT NAME						
3+m+n								(LSB)

The ACCEPT bit shall be set to zero in the process login payload transmitted by the data transfer device. If the automation device supports surrogate operation, then it shall set the ACCEPT bit in the process login payload it transmits.

The I_T NEXUS IDENTIFIER field contains a value identifying the SCSI Port of an application client sending a request to the SMC device server. This value shall not be zero.

The INITIATOR PORT IDENTIFIER field contains the port identifier of the SCSI Port of the application client. The INITIATOR PORT IDENTIFIER LENGTH field contains the length in bytes of the INITIATOR PORT IDENTIFIER field.

The INITIATOR PORT NAME field contains the port name of the SCSI Port of the application client. The INITIATOR PORT NAME LENGTH field contains the length in bytes of the INITIATOR PORT NAME field.

6.5.10 Process logout

Process logout frames are sent by the data transfer device to the automation. Upon receiving a Process Logout frame (see Table x), the device shall abort all open exchanges associated with that login.

Table x - Process logout frame payload contents

Bit / Byte	7	6	5	4	3	2	1	0
0	Rsvd	ACCEPT	Reserved					
1	I_T NEXUS IDENTIFIER							

The I_T NEXUS IDENTIFIER field contains the I_T nexus identifier for the initiator port being logged out.

Editor's Note: Is an ACCEPT required? Can we delete the Reserved field to make this a one-byte payload?

The following changes are for ADT clause 7.1.2 SCSI Command frame:

- The I_T NEXUS and COMMAND REFERENCE NUMBER fields are added to the SCSI Command frame payload in Table 8.
- The paragraph describing the I_T NEXUS field is added following the paragraph describing the LUN field.

Table 8 – SCSI command frame payload contents

Byte \ Bit	7	6	5	4	3	2	1	0
0	(MSB)			LUN				
1								(LSB)
2	I_T NEXUS IDENTIFIER							
3	COMMAND REFERENCE NUMBER							
4	TASK MANAGEMENT FUNCTION							
5	Reserved							
6	CDB							
21								
22	(MSB)			BUFFER ALLOCATION LENGTH				
25								(LSB)

If the SCSI command frame is being transmitted by the automation device, then the I_T NEXUS IDENTIFIER field shall contain zero. If the SCSI command frame is being transmitted by the data transfer device as part of SCSI Passthrough Surrogate operation, then the I_T NEXUS IDENTIFIER field shall contain the value in the I_T NEXUS field in the Process Login request for the initiator originating the SCSI request. If the SCSI command frame is being transmitted by the data transfer device as part of SCSI Surrogate Bridged Surrogate operation, then the I_T NEXUS IDENTIFIER field shall contain zero.

The COMMAND REFERENCE NUMBER (CRN) field contains the number sent by the initiator to assist in performing precise delivery checking for commands. If precise delivery is enabled, a nonzero value in the CRN field shall be treated as a command reference number in determining the receipt and ordering of commands from a particular initiator to the particular logical unit. If precise delivery is enabled, a zero value in the CRN field indicates that command shall not be verified for precise delivery. If precise delivery checking is not enabled, the COMMAND REFERENCE NUMBER field shall be ignored by the device server. If the SCSI command frame specifies a task management function, the CRN field shall be reserved and set to zero and the SCSI command frame shall not be verified for precise delivery.

Editor's Note: As you see, I opted to put in ADC the normative text describing how a mode descriptor setting requires the use of an ADT feature, e.g., process login. Is this acceptable?