June 27, 2000 T10/00-232 revision 1

To: John Lohmeyer, chairperson, T10
From: Bob Snively
Date: June 27, 2000
Subject: Asymmetrical SCSI behavior

1 Introduction

A significant number of SCSI storage subsystems have the property of asymmetrical accessibility to logical units through various ports. Typically, one target may provide full performance access to a logical unit, while another target, possibly on a different physical controller, provides either lower performance access or supports a subset of the available SCSI commands to the same logical unit. In some cases, the logical unit can be modified to provide full performance access to the limited target if the original full performance target fails. This proposal proposes a set of SCSI tools necessary to properly support such asymmetrical access and failure recovery.

This revision of the document contains the recommendations of the SCSI working group in the meeting of May 17, 2000. No change bars are included in this new revision because of the extent of the changes.

2 Overview

Symmetrical access to logical units is very desirable, since it provides for very rapid recovery from link failures and it provides the infrastructure that supports dynamic load balancing capabilities. Symmetrical access is characteristic of almost all simple disk drives and JBODs. Symmetrical access is managed by the simple SCSI mechanisms already defined by SAM-2, SPC-2 and other documents.

Asymmetrical access is useful, since it can be implemented for very large storage sub-systems with very simple and low-cost storage controller configurations. Asymmetrical access requires additional SCSI mechanisms to indicate which targets are fully accessible and which targets have only partial access to a particular logical unit. Mechanisms are also required to allow controlled transfer of the full performance functionality from one target to another target for a particular logical unit. The proposals in this document provide those mechanisms.

3 Parameters and states for managing asymmetrical access to SCSI logical units

The proposal is designed to be included in SPC-3. The commands indicating that the logical unit implements asymmetrical access are extracted from SPC-2. The commands selected by the working group for target discovery will be carried into SPC-3, but were originally defined for SCC-2.

Brocade Communications Systems, Inc.
1901 Guadalupe Parkway n San Jose, CA 95131
T 408.487.8000 n F 408.487.8101
www.brocade.com

Target state

Each target or symmetrical group of targets must be in one of the following states with respect to the ability to access a particular LUN:

Active:

The target is capable of immediately accessing the logical unit. All commands operate exactly as specified by SPC-2, SBC, and SES today. The target participates in all task management commands as defined in SPC-2, SAM-2 and other documents.

Degraded:

The target is capable of supporting all commands operating exactly as specified by SPC-2, SBC, and SES today. The execution of certain commands, especially those involving data transfer or caching, may operate with lower performance than they would if the target were in active state. The target participates in all task management commands as defined in SPC-2, SAM-2, and other documents.

Inactive:

The target is capable of performing a limited set of SPC-2 commands. Those commands that operate behave precisely as specified by SPC-2, SBC, and SES today. Those commands that do not operate will provide the specified error indication. Commands that operate include those necessary for:

Diagnosing and testing the logical unit and its paths

Identifying the path

Identifying the logical unit

Determining the operational state of the logical unit

Determining the active/inactive state of the unit

Allowed commands include:

INQUIRY,

LOG SELECT and LOG SENSE

MODE SELECT (6/10) and MODE SENSE (6/10)

REPORT LUNS

RECEIVE DIAGNOSTIC RESULTS and SEND DIAGNOSTIC.

The target participates in all task management commands as defined in SPC-2, SAM-2, and other documents.

Unavailable:

The target cannot access the requested logical unit. The only commands that operate normally are INQUIRY, REQUEST SENSE, and, for logical unit 0, REPORT LUNS. Those commands that do not operate provide the specified error indication. The target does not participate in any task management commands relative to the logical unit from other targets.

Management function

The following management operation can be performed for each target:

Make target(s) active:

The target or group of targets indicated is made active. The previously active targets not kept active are made degraded, inactive, or unavailable in a vendor specific manner. (See 4.5 below)

This function can be done explicitly or automatically. See 4.2 below to determine whether the device requires explicit or implicit change. [Alternatively, implicit or explicit behavior can be specified by adding a few bits to item 4.5 below.]

Discovery requirements

The following information must be discoverable by an appropriate mechanism.

Identify asymmetrical access requirement:

A value is provided to indicate that only one target or group of targets at a time is allowed normal access to a logical unit. (See 4.2, below)

Identify target group to logical unit:

A value is provided identifying the target and target group through which a command is being passed. This can be optionally associated with a port world-wide name through either the SCSI command set or the Fibre Channel command set. (See 4.3, below)

Report active target group:

A value is provided identifying the target or group of targets that is active. (See 4.3, below)

Report target groupings:

A list of available groups of targets is provided for the logical unit. The software driver may be required to assist in the identification of targets and target groupings and their relationship to other identification parameters.

4 PROPOSAL FOR CHANGES TO SPC-2 or SBC-3 DOCUMENT

4.1 Section 5.5, model for asymmetric target behavior

The following text is provided immediately after section 5.5 to define the behavioral model for asymmetric target behavior, pushing clause 5.6 and subsequent functions one number higher.

5.6 Asymmetric target behavior

5.6.1 Introduction

Asymmetric target behavior is defined as SCSI operation such that logical units may have different performance properties and allowable command sets when accessed through different targets. Devices having the same access to all logical units through all targets exhibit symmetric target behavior.

Symmetric access to logical units is often desirable, since it provides for very rapid access to a logical unit through alternate targets when the connection to one target is lost. Symmetrical access is characteristic of many simple SCSI devices such as disk drives and disk enclosures containing providing direct access to a set of disk drives. Symmetrical access is managed by the simple mechanisms already defined for SCSI.

Asymmetric access to logical units is useful for other reasons. Devices with targets implemented in separate physical units may need to designate one or more active targets for each logical unit. While access to the logical unit may be possible through any target, the performance or accepted command set may be less complete through targets that are not designated as active. When a failure on the path to an active target is detected, the target can perform automatic internal reconfiguration to make another target active or may be instructed by the initiator to make a different set of targets active.

If more than one target has the same asymmetric target behavior, the set of targets having the same behavior is called a target group. All members of a target group have exactly the same behavior with respect to a logical unit. Other target groups may have different and asymmetric target behavior with respect to the same logical unit. An example of such a device is a RAID device with two separated controllers. All targets on one controller will have the same behavior with respect to a logical unit and will be members of the same target group. Targets on the other controller will also be members of another target group. The behavior of each target group may be different with respect to the logical unit, but all members of a single target group have the same behavior with respect to the logical unit.

Asymmetric target behavior may be managed explicitly by an application client using the REPORT TARGET GROUPS and SET TARGET GROUPS management commands. Alternatively asymmetric behavior may be managed implicitly by the logical unit and target based on the type of transactions being performed across through each target and the internal configuration capabilities of the set of targets through which the logical unit can be accessed. The reactions of logical units when implicit asymmetric target behavior management is performed are outside the scope of this standard. In general, the logical units attempt to maintain full performance across the links that are busiest and which show the most reliable performance, allowing links to other targets to assume one of the lower performance states.

5.6.2 Asymmetric target behavior states.

When a logical unit shows asymmetrical target behavior, the target or symmetrical group of targets must be in one of the following states with respect to the ability to access that logical unit:

Active state:

The target or group of targets is capable of immediately accessing the logical unit. All commands operate exactly as specified by this standard, SBC, SES or other command set standards. The target participates in all task management operations as defined SAM-2.

Degraded:

The target or group of targets is capable of supporting all commands operating exactly as specified by this standard, SBC, SES, or other command set standards. The execution of certain commands, especially those involving data transfer or caching, may operate with lower performance than they would if the target were in active state. The target participates in all task management operations as defined in SAM-2.

Inactive:

The target or group of targets is capable of performing a limited set of the commands supported by this standard. For those commands that do not operate, the device server shall return CHECK CONDITION status with the sense key set to NOT READY and an ASC/ASCQ of LOGICAL UNIT NOT READY, PATH STATE INACTIVE. In the inactive state, the following commands shall operate normally.

INQUIRY

LOG SELECT and LOG SENSE

MODE SELECT (6/10) and MODE SENSE (6/10)

REPORT LUNS

RECEIVE DIAGNOSTIC RESULTS and SEND DIAGNOSTIC

REPORT TARGET GROUPS

SET TARGET GROUPS

The target participates in all task management operations as defined SAM-2.

Unavailable:

The target or group of targets cannot access the requested logical unit. The only commands that operate normally are INQUIRY, REQUEST SENSE, and, for logical unit 0, REPORT LUNS. The INQUIRY command indicates that the device is unavailable by setting the peripheral device qualifier to 001b (see 7.6.2). For those commands that do not operate, the device server shall return CHECK CONDITION status with the sense key set to NOT READY and an ASC/ASCQ of LOGICAL UNIT NOT READY, PATH STATE UNAVAILABLE. The target does not participate in any task management operations from other targets.

Transition:

The target or group of targets is in the process of changing from one state to another. A command may operate according to one of the following four states:

the state from which the target is changing, or

the state to which the target is changing, or

the device server shall return BUSY status, or

the device server shall return CHECK CONDITION status with the sense key set to NOT READY and an ASC/ASCQ of LOGICAL UNIT NOT READY, PATH STATE TRANSITION IN PROGRESS.

5.6.3 Discovery of asymmetric target behavior

SCSI logical units with asymmetric target behavior are identified by performing an INQUIRY command to the logical unit. The value of the asymmetric target behavior (ABT) bit (see 7.5.1) indicates whether the logical unit shows asymmetric target behavior or does not show asymmetric target behavior. If the logical unit does not indicate the support of asymmetric target behavior, the target behavior may be either unspecified or symmetric. When a logical unit supports asymmetric target behavior, the value of the ABTI bit (see 7.5.1) indicates whether the logical unit supports implicit or explicit management of the asymmetric target behavior.

5.6.4 Implicit asymmetric target behavior management

SCSI logical units with implicit asymmetric target behavior management set the state of each target or target group using mechanisms outside the scope of this standard. The following mandatory mechanism is available to determine the asymmetric target behavior state of each target and the membership of a target group.

All logical units with asymmetric target behavior shall implement the INQUIRY command VPD page 83h, the device identifier page. The logical unit shall implement the 04h and 05h identifier types as specified in 8.4.3. The 04h identifier type defines the relative target identifier for the target through which the INQUIRY command was passed. This also provides the necessary information to identify targets defined by the REPORT TARGET GROUPS command. The 05h identifier type defines the target group and the present state of the target group. The status of the target group may change at any time as required by implicit asymmetric target behavior management mechanisms. The REPORT TARGET GROUPS command is optional for implicit asymmetric target behavior management.

5.6.5 Explicit asymmetric target behavior management

SCSI logical units with explicit asymmetric target behavior management identify the relative target identifier, target group, and present status of the target group using the mandatory features of the INQUIRY command as defined in 5.6.4. The REPORT TARGET GROUPS command, which is mandatory for logical units with explicit asymmetric target behavior management, can then be used to provide a complete list of all the target groups, the relative target identifiers of the members of each target group, and the asymmetric target behavior state for each target group.

The SET TARGET GROUPS command, mandatory for all logical units with explicit asymmetric target behavior management, is used to set the asymmetric target behavior state for all target groups. The allowable combination of states for the target groups is vendor specific. If the SET TARGET GROUPS attempts to establish an invalid combination of states, the device server shall return CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and an ASC/ASCQ of INVALID FIELD IN PARAMETER LIST.

4.2 Section 7.5.1, Standard INQUIRY data

Asymmetric target behavior bit

A new bit will be placed in byte 5 of INQUIRY data to indicate that the logical unit has asymmetric targets and supports asymmetric target behavior. The default of 0 indicates symmetric or unspecified target behavior. The value of 1 indicates that asymmetric target behavior is observed by this logical unit and that the asymmetric target management and discovery tools are supported.

Implicit asymmetric target behavior bit

A new bit will be placed in byte 5 of INQUIRY data to indicate that the logical unit supports asymmetric behavior using only implicit target group activation. The default of zero is explicit target group activation. If the asymmetric target behavior bit is zero, the implicit asymmetric target behavior bit is reserved and shall be zero.

The text to be included in this section is defined below:

An asymmetric target behavior (ATB) bit of one indicates that the device has asymmetric target behavior. An ATB bit of zero indicates that the device does not support asymmetric target behavior. An ATB bit of zero may indicate either that the device behavior with respect to asymmetry is undefined or that the device supports symmetric target behavior. If the ATB bit is set to zero, the IATB bit shall be zero.

An implicit asymmetric target behavior (IATB) bit of one indicates that the device performs management of asymmetric target behavior automatically. If this bit is set to one, the ATB bit shall be set to one. An IATB bit of zero and an ATB bit of one indicates that the device does not perform any automatic management of asymmetric behavior and the device shall support the REPORT TARGET GROUPS and SET TARGET GROUPS commands.

At the discretion of the editor, these two paragraphs may also be explained with a table if desired.

4.3 Section 8.4.3, Device Identification Page

The document presently defines a 4-byte “relative target identifier” that, when included, with the association value of 1, indicates the target (relative to some arbitrary internal ordering) through which the command is passed.

The device identification page allows simultaneous presentation of identifiers, including the target WWN, the LUN WWN, the relative target identifier, and any other similar information.

The relative target identifier is used to locate the actual target paths which will be reported and controlled using the REPORT TARGET GROUPS and SET TARGET GROUPS commands.

An additional identifier entry is included for implicit asymmetric behavior state discovery and for implicit asymmetric behavior target group discovery. The tables are updated and added as below.

Table 181 — Identifier type

Value	Description
5h	If the ASSOCIATION value is 1h, the IDENTIFIER value contains the target group and target group status as defined in table ZZ1. For this case, the CODE SET field shall be set to 1h and the IDENTIFIER LENGTH field shall be set to 4. If the ASSOCIATION value is not 1h, use of this identifier type is reserved.
6h - Fh	Reserved

Table ZZ1— Target Group and Target Group Status Identifier

Bit / Byte	7	6	5	4	3	2	1	0
0	reserved							
1	reserved							
2	TARGET GROUP IDENTIFIER (first group)							
3					ACTIVE	DEGRADED	INACTIVE	UNAVAIL

See 7.x for the definitions of the TARGET GROUP IDENTIFIER field and the ACTIVE, DEGRADED, INACTIVE, and UNAVAIL bits.

4.4 Section 7.3.6, new ASC/ASCQ

Proper error indications are defined in this section 7.3.6 of the SPC-2 document. The error indications would include:

LOGICAL UNIT NOT READY, PATH STATE INACTIVE (proposed value = 04/0B)

LOGICAL UNIT NOT READY, PATH STATE UNAVAILABLE (proposed value = 04/0C)

LOGICAL UNIT NOT READY, PATH STATE TRANSITION IN PROGRESS (proposed value = 04/0A)

4.5 Target group management

The committee has elected to place the management functions in the MAINT IN command (operation code A3) and the MAINT OUT command (operation code A4), primarily defined in SCC-2. Service action 0A has been assigned to Asymmetric Target Group Set and Asymmetric Target Group Sense functions.

The text, placed at the same level as a standard command in section 7, will read:

7.x REPORT TARGET GROUPS

The REPORT TARGET GROUPS command (see table XX1) requests that the device server send target group information to the application client. This command is mandatory for all devices that report in the INQUIRY data that they support asymmetric target behavior and do not sup-

port implicit asymmetric target behavior. This command is optional for all devices that report in the INQUIRY data that they support asymmetric target behavior and do support implicit asymmetric target behavior. This service action shall be rejected by all other devices that support the MAINTENANCE IN command with an ASC/ASCQ of INVALID FIELD IN CDB.

As defined in the SCC-2 standard, the REPORT TARGET GROUPS command is the REPORT TARGET GROUPS service action of the MAINTENANCE IN command. Additional MAINTENANCE IN service actions (that apply to SCC-2 devices and devices that set the SCCS bit in their Standard Inquiry data) are defined in SCC-2. Only those service actions of MAINTENANCE IN that are defined in this standard concern all SCSI devices. SCC-2 defines specific usages for bytes 4 and 5, and bit 1 in byte 10, however these fields are reserved for the REPORT TARGET GROUPS command defined by this standard.

Table XX1 — REPORT TARGET GROUPS command

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (A3h)							
1	Reserved			SERVICE ACTION (0Ah)				
2	Reserved							
3	Reserved							
4	See SCC-2 reserved							
5								
6	(MSB)							
7	ALLOCATION LENGTH							
8								
9								(LSB)
10	Reserved						See SCC-2	Reserved
11	CONTROL							

The ALLOCATION LENGTH field indicates how much space has been allocated for the returned parameter data. If the length is not sufficient to contain all the parameter data, the first portion of the data shall be returned. This shall not be considered an error. The actual length of the parameter data is available in the IDENTIFIER LENGTH field in the parameter data. If the remainder of the parameter data is required, the application client should send a new REPORT TARGET GROUP command with an ALLOCATION LENGTH field large enough to contain all the data.

The REPORT DEVICE IDENTIFIER parameter list (see table XX2) contains a four-byte field that contains the length in bytes of the parameter list and the logical unit's identifier.

Table XX2 — REPORT TARGET GROUPS parameters

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1	REPORT FIELD LENGTH							
2								
3								(LSB)
4	TARGET GROUP IDENTIFIER (first group)							
5					ACTIVE	DEGRADED	INACTIVE	UNAVAIL
6	TARGET COUNT IN GROUP (first group, =n)							
0	(MSB)							
1	RELATIVE TARGET IDENTIFIER (first target of first group)							
2								
3								(LSB)
8	...							
0	(MSB)							
1	RELATIVE TARGET IDENTIFIER (last target of first group)							
2								
3								(LSB)
4	TARGET GROUP IDENTIFIER (second group)							
5					ACTIVE	DEGRADED	INACTIVE	UNAVAIL
6	TARGET COUNT IN GROUP (second group, =n)							
0	(MSB)							
1	RELATIVE TARGET IDENTIFIER (first target of second group)							
2								
3								(LSB)
8	...							

Table XX2 — REPORT TARGET GROUPS parameters

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1	RELATIVE TARGET IDENTIFIER (last target of second group)							
2								
3								(LSB)

The REPORT FIELD LENGTH field specifies the length in bytes of the list of target groups. If the ALLOCATIONLENGTH field in the CDB is too small to transfer all of the identifier, the length shall not be adjusted to reflect the truncation.

The TARGET GROUP IDENTIFIER field specifies a one-byte identification of the target group. A particular logical unit may be served by up to 256 target groups.

The state bits, ACTIVE, DEGRADED, INACTIVE, and UNAVAIL define the state of the target group as defined by 5.6. Exactly one of these bits shall be set.

The TARGET COUNT IN GROUP field specifies the number of targets that have access to the logical unit in that group. Not all targets may be connected or available, but all are always listed in the list and counted by this field.

The RELATIVE TARGET IDENTIFIER field specifies a four-byte identification of a target. This is the same value provided for relative target identifier in the vital product page 83h, the Device Identification Page.

The execution of a REPORT TARGET GROUPS command may require the enabling of a nonvolatile memory within the logical unit. If the nonvolatile memory is not ready, the device server shall return CHECK CONDITION status, rather than wait for the device to become ready. The sense key shall be set to NOT READY and the additional sense data shall be set as described in the TEST UNIT READY command (see 7.27). This information should allow the application client to determine the action required to cause the device server to become ready.

7.y SET TARGET GROUPS

The SET TARGET GROUPS command (see table YY1) requests that the device server set the state of the target groups for the specified logical unit. This command is mandatory for all devices that report in the INQUIRY data that they support asymmetric target behavior and do not support implicit asymmetric target behavior. This command is optional for all devices that report in the INQUIRY data that they support asymmetric target behavior and do support implicit asymmetric target behavior. This service action shall be rejected by all other devices that support the MAINTENANCE OUT command with an ASC/ASCQ of INVALID FIELD IN CDB.

As defined in the SCC-2 standard, the SET TARGET GROUPS command is the SET TARGET GROUPS service action of the MAINTENANCE OUT command. Additional MAINTENANCE OUT service actions (that apply to SCC-2 devices and devices that set the SCCS bit in their Standard Inquiry data) are defined in SCC-2. Only those service actions of MAINTENANCE OUT that are defined in this standard concern all SCSI devices. SCC-2 defines specific usages for bytes 4 and 5, and bit 1 in byte 10, however these fields are reserved for the SET TARGET GROUPS command defined by this standard.

Table YY1 — SET TARGET GROUPS command

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (A4h)							
1	Reserved			SERVICE ACTION (0Ah)				
2	Reserved							
3	Reserved							
4	See SCC-2 reserved							
5								
6	(MSB)							
7	PARAMETER LIST LENGTH							
8								
9								(LSB)
10	Reserved						See SCC-2	Reserved
11	CONTROL							

The PARAMETER LIST LENGTH field specifies the length in bytes of the target group management parameters that shall be transferred from the application client to the device server. A parameter list length of zero indicates that no data shall be transferred, and that no change shall be made in the state of any target groups. The number of target groups supported by a logical unit is vendor specific. The number of target groups that must be provided in the parameter list and the allowable values to which their states may be set is vendor specific. If the parameter list length violates the vendor specific length requirements, then the device server shall return CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and an additional sense code of INVALID FIELD IN CDB.

The SET TARGET GROUPS parameter list (see table YY2) contains a four-byte field that contains the length in bytes of the parameter list.

Table YY2 — SET TARGET GROUPS parameters

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1	SET FIELD LENGTH (=n)							
2								
3								(LSB)
4	TARGET GROUP IDENTIFIER (first group)							
5					ACTIVE	DEGRADED	INACTIVE	UNAVAIL
	...							
n+2	TARGET GROUP IDENTIFIER (second group)							
n+3					ACTIVE	DEGRADED	INACTIVE	UNAVAIL

The SET FIELD LENGTH field specifies the length in bytes of the list of target groups.

The TARGET GROUP IDENTIFIER field specifies a one-byte identification of the target group. A particular logical unit may be served by up to 256 target groups.

The state bits, ACTIVE, DEGRADED, INACTIVE, and UNAVAIL are set to request the state of the target group as defined by 5.6. Exactly one of these bits shall be set.

The execution of a SET TARGET GROUPS command may require the enabling of a nonvolatile memory within the logical unit. If the nonvolatile memory is not ready, the device server shall return CHECK CONDITION status, rather than wait for the device to become ready. The sense key shall be set to NOT READY and the additional sense data shall be set as described in the TEST UNIT READY command (see 7.27). This information should allow the application client to determine the action required to cause the device server to become ready.