

Date: Feb,05 1994

X3T10/94-040 rev 1

To: X3T9.2 Committee (SCSI)

From: George Penokie (IBM)

Subject: SCSI Disk Array Model

| - Denotes changes from 93-140r3 to 94-040r1

0.0 References

This proposal uses many definitions and ideas contained within the SCSI-3 Architecture Model (SAM). It is assumed a reader of this proposal has knowledge of SAM definitions.

The SAM draft standard number X3T9.2/994D is under development within the X3T9.2 committee. Copies of the SAM document may be purchased from:

Global Engineering Documents, 15 Inverness Way East,
Englewood, Co. 80112-5704, (800)854-7179 or (303) 792-2181
FAX: (303) 792-2192

1.0 Glossary

1. Check Data - Information contained within a redundancy group that allows lost or destroyed user data to be recreated. Any overlapping redundancy groups shall not have check data within the overlapping P-extents.
1. Check Data Mapping - The distribution of the check data within a redundancy group.
1. Disk Array Conversion Layer (DACL) - Converts input logical unit identifiers to output logical unit identifiers and may convert input LBAs to output LBAs.
1. Group - Objects that are independent from one another and may overlay one another. Groups may span more than one device. A single device may contain more than one group or may contain the entire group.
- | 1. P-extent - All or part of the host addressable space
| within a single P-LUI. The parameters that define a fixed block P-extent are:

- Physical Logical Unit Identifier (P-LUI)
- Start P-LBA
- Number of P-LBAs
- Number of Bytes per P-LBA

The parameters which define a variable block P-extent

are:

- Physical Logical Unit Identifier (P-LUI)
- Start P-LBA
- Number of P-LBAs

The P-LUI defines the address of the physical device that contains the P-extent.

The start P-LBA is the first addressable P-LBA of the P-extent.

The number of P-LBAs provides the capacity of the P-extent in blocks.

The number of bytes per P-LBA provides the size, in bytes, of the blocks in the P-extent.

1. P-LUI - Physical Device Logical Unit Identifier
1. P-LBA - Physical Device Logical Block Address
1. Protected Space - The portion of a redundancy group that does not contain check data.
1. PS-extent - All or part of the protected space within a single P-LUI configured as a redundancy group.

PS-extents are created on the successful completion of a create/modify redundancy group service request. PS-extents are reported within the parameter list returned by a report unassigned redundancy group service or a report volume set service. PS-extents are used to define volume sets. Volume sets are defined using a create/modify volume set service.

The parameters that define a fixed block PS-extent are:

- Physical Logical Unit Identifier (P-LUI)
- Start PS-LBA
- Number of PS-LBAs
- Number of Bytes per PS-LBA

The parameters which define a variable block PS-extent are:

- Physical Logical Unit Identifier (P-LUI)
- Start PS-LBA
- Number of PS-LBAs

The P-LUI defines the address of the physical device that contains the PS-extent.

The start PS-LBA is the first unallocated addressable

PS-LBA of protected space within the PS-extent.

The number of PS-LBAs provides the capacity of the protected space PS-extent in blocks.

Note: In PS-extents the number of PS-LBAs does not include any LBAs that have been assigned to contain check data in contrast to P-extents in which the number of P-LBAs includes all addressable LBAs on a P-LUI.

The number of bytes per PS-LBA provides the size, in bytes, of the blocks in the PS-extent.

1. PS-LBA - Protected space Logical Block Address
1. R-LUI - Redundancy group Logical Unit Identifier
1. Rebuild Operation - Recreation of protected space contents or any check data within a P-extent using check data and protected space contents from the remaining P-extents within the redundancy group. The regenerated protected space contents or any recalculated check data shall be written to the P-extent being rebuilt.
1. Recalculate Operation - Recreation of check data from protected space contents. The recreated check data shall be written to the check data location being recalculated.
1. Regenerate Operation - Recreation of inaccessible protected space contents from accessible check data and protected space contents. The recreated protected space contents is not saved.
1. Redundancy Group - A grouping of protected space and associated Check Data (Check Data may be null) into a single Logical Unit that shall only have a single type of redundancy. P-extents that contain no check data may be associated with more than one redundancy groups.
1. Set - Objects that do not intersect and are independent from one another. Sets may span more than one device. A single device may contain more than one set or may contain the entire set.
1. S-LUI - Spare Logical Unit Identifier
1. Spare - A range of LBAs or a P-LUI associated with one or more redundancy groups or P-LUIs that can be used to replace a portion of a redundancy group or a P-LUI.
1. SCSI Disk Array (SDA) - A device that processes SCSI command descriptor blocks and performs the services of a disk array conversion layer. A single SDA may contain multiple DACLs.

- 1. User Data - The addressable continuum of logical blocks that are input to the Disk Array Conversion Layer. Check Data is not part of this addressable continuum of logical blocks.
- 1. User Data Mapping - The distribution of user data within a volume set.
- 1. Verify Operation - Recreate check data from protected space contents and compare the recreated check data with the current check data. If the recreated check data does not match the current check data an exception condition shall be created.
- 1. V-LUI - Volume set Logical Unit Identifier
- 1. V-LBA - Volume set Logical Block Address
- 1. Volume Set - One or more PS-extents grouped into a single Logical Unit.

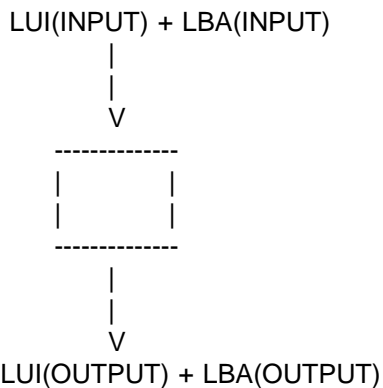
2.0 SCSI Disk Array Layering Model

A system is typically composed of many protocol conversion layers, and these layers may exist in hardware or software. Each of these layers has input(s) and output(s). The next layer accessed is determined by the preceding layers output.

These protocol conversion layers include, but are not limited to: transport modules, host adapter drivers, SIMs, HBAs, bridge controllers, and disk drives. Each of these layers will be represented by a simple block that has an input and output.

All requests to or from a DACL contain LUIs but not all requests contain LBAs.

Note: The logical unit identifier is fully defined in SAM.



2.1 Generic Layer

Generic Layers do not modify the logical unit identifier or the LBA.

Note: There are types of layers other than generic and disk array conversion layers, however, these are not covered in this model.

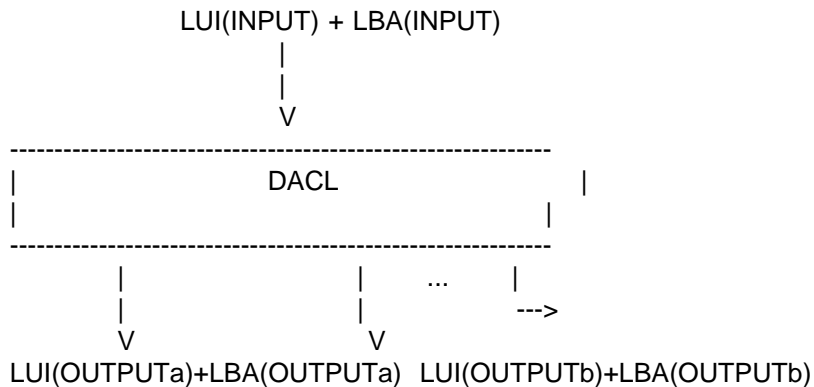
.pa

2.2 Disk Array Conversion Layer (DACL)

The disk array conversion layer is capable of extensive manipulation on the logical unit identifier and the LBA, based upon a consistent algorithm that follows the defined configuration. It is possible that a single LUI(INPUT) or LBA(INPUT) may be converted to multiple different LUI(OUTPUT)s and LBA(OUTPUT)s. Disk array conversion layers will be shown with 'DACL' in the block.

The LUI(OUTPUT)s may not be SCSI devices.

Disk array conversion layers provide the functions listed in section 4.0.



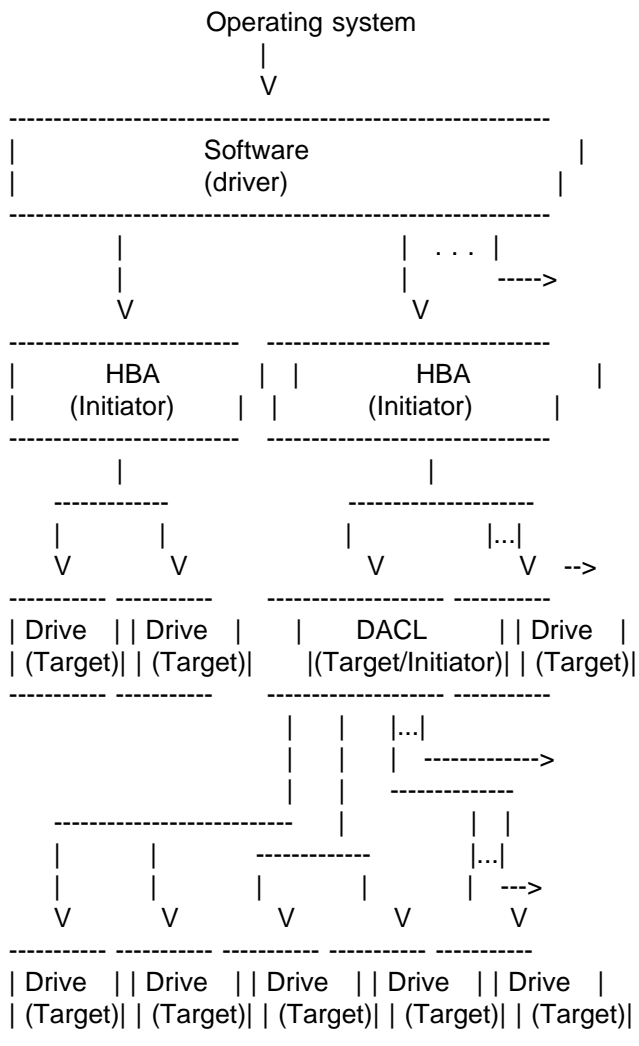
.pa

3.0 Examples of SCSI Disk Array Variations

Typically a system diagram will be composed of many layers combined into a tree. For example, a driver may connect to multiple HBAs, which in turn may connect to multiple Targets, etc. The example below shows a system that consists of:

- One initiator has two SCSI targets attached on a single SCSI bus and this SCSI bus is not expandable.
- One initiator has two targets attached on a single SCSI bus and this SCSI bus is expandable. One of the targets is a DACL.
- The DACL initiator has three SCSI buses with targets attached and is capable of driving more SCSI buses. Two of the SCSI buses contain two targets each and these SCSI buses

are not expandable. One of the SCSI buses contains one target and this SCSI bus is expandable.

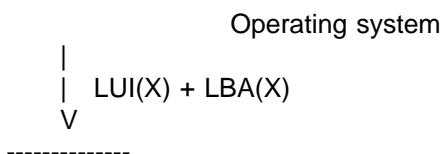


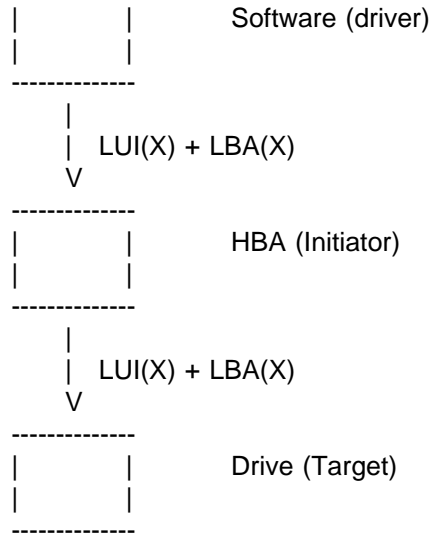
In the following diagrams, only a single branch of this tree will be shown.

.pa

3.1 Branch of Generic Layers

This diagram shows that all layers pass the logical unit identifier and LBA directly through.

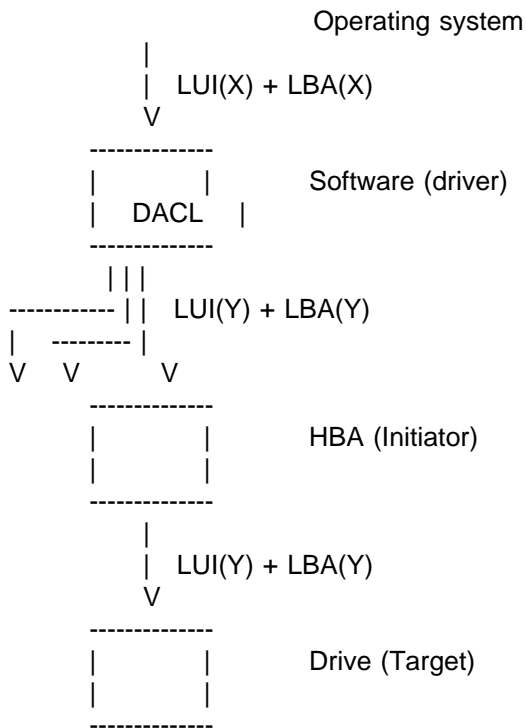




.pa

3.2 Software Array with a Branch of SCSI Disks

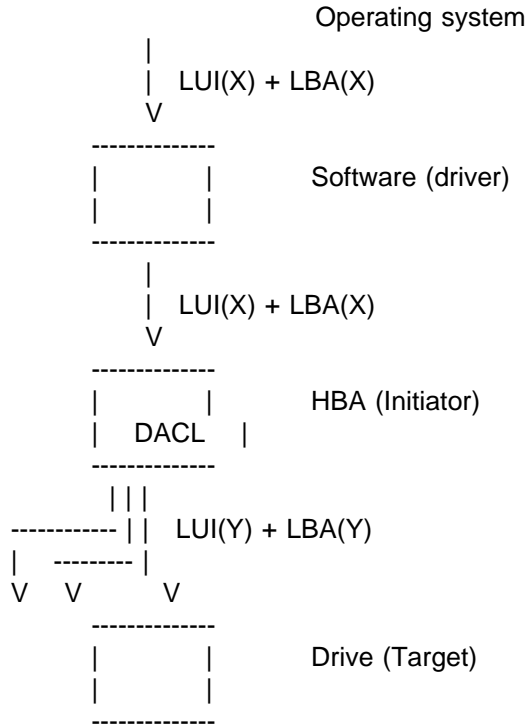
This diagram shows software performing SCSI Disk Array functions. These functions convert the input logical unit identifier (LUI(X)) and the input LBA (LBA(X)) to the output logical unit identifier (LUI(Y)) and output LBA (LBA(Y)). All other layers pass the logical unit identifier and LBA through.



.pa

3.3 Branch with HBA SCSI Disk Array

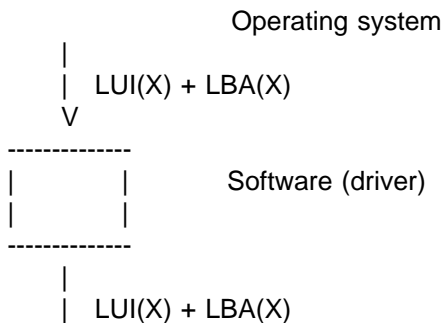
This diagram shows a HBA performing SCSI Disk Array functions. These functions convert the input logical unit identifier (LUI(X)) and the input LBA (LBA(X)) to the output logical unit identifier (LUI(Y)) and output LBA (LBA(Y)). All other layers pass the logical unit identifier and LBA through.

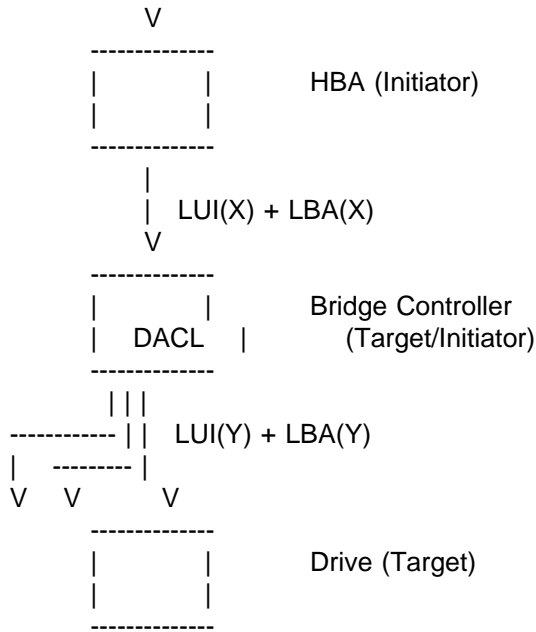


.pa

3.4 Branch with Bridge Controller SCSI Disk Array

This diagram shows a bridge controller performing SCSI Disk Array functions. These functions convert the input logical unit identifier (LUI(X)) and the input LBA (LBA(X)) to the output logical unit identifier (LUI(Y)) and output LBA (LBA(Y)). All other layers pass the logical unit identifier and LBA through.



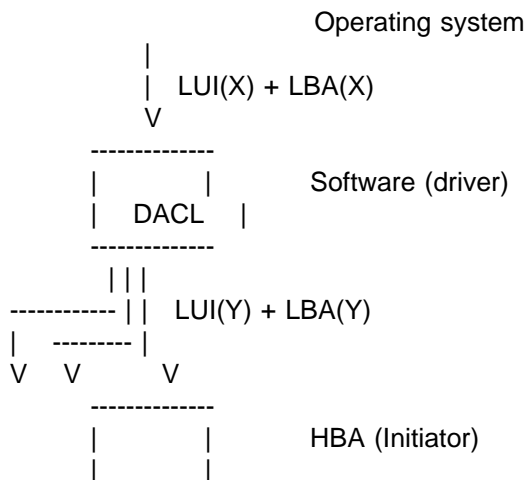


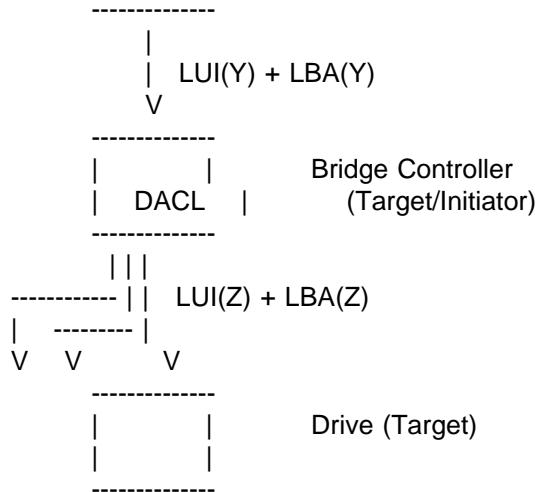
.pa

3.5 Branch with Multiple Layers of SCSI Disk Arrays

This diagram shows one example with two disk array controller layers. This example shows a software layer performing SCSI disk array functions and a bridge controller performing SCSI disk array functions.

The software functions convert the input logical unit identifier (LUI(X)) and the input LBA (LBA(X)) to the output logical unit identifier (LUI(Y)) and output LBA (LBA(Y)). The bridge controller functions then convert the input logical unit identifier (LUI(Y)) and the input LBA (LBA(Y)) to the output logical unit identifier (LUI(Z)) and output LBA (LBA(Z)). All other layers pass the logical unit identifier and LBA through.





.pa

4.0 Model for SCSI disk arrays

The following is a model that defines SCSI disk arrays.

4.1 SCSI Disk Array Addressing

When addressing a SCSI Disk Array (SDA) one of the following methods is used:

METHOD OF ADDRESSING	ADDRESSABLE USING IDENTIFY MESSAGE	ADDRESSABLE WITHIN COMMANDS	PREFIX	SPACE ADDRESSED
Volume Set	yes	yes	V-	user data
Physical	yes	yes	P-	physical
Redundancy Group	no	yes	R-	protected
Spare	no	yes	S-	physical

Redundancy groups and spares shall be accessed by addressing service requests to the P-LUI of the SDA that controls the redundancy groups and spares.

<<Editors note: Parts of document 93-161rx will go here. That document describes the addressing used when talking to a SDA - GOP>>

4.1.1 Physical Address Mode

<<Editors note: Parts of document 93-161rx will go here. That document describes the addressing used when talking to a SDA - GOP>>

All SCSI operations are allowed in physical address mode, however P-LUIs and SDAs are not required to support all SCSI commands. SDAs are not required to honor pass-through

requests from the Application Client. Any command that is not supported or passed-through shall terminate in error.

If the P-LUI is contained within the SDA and is not controlled by a DACL (e.g. power supplies, etc., that are under the control of the SDA itself) then the SDA executes the command.

4.1.2 Redundancy Group Address Mode

A redundancy group shall only be accessed by addressing the SDAs' P-LUI that controls or will control the redundancy group.

4.1.3 Volume Set Address Mode

The volume set address mode points to the DACL that executes the command using the algorithms defined by the configuration.

The algorithms may instruct the DACL to translate the input V-LUI to output P-LUI(s) and the command into zero or more commands.

Note: The volume set may not be under the control of the addressed SDA. It may be in a SDA lower in the tree structure.

All SCSI operations are allowed in volume set address mode, however V-LUIs are not required to support all SCSI commands. Any command that is not supported shall terminate in error.

<<Editors note: Parts of document 94-031rx will go here. That document describes the addressing used when talking to a SDA - GOP>>

4.1.4 Spare Address Mode

A spare shall only be accessed by addressing the SDAs' P-LUI that controls or will control the spare.

5.0 SCSI Disk Array Requests

5.1 Maintenance Services

The SDA provides the following maintenance services:

- Report P-LUI Service
- Report P-LUI Geographic Location Service
- Set P-LUI Geographic Location Service
- | -Report P-extents Service
- | -Report Unassigned P-extents Service
- Exchange P-LUI Service
- Remove P-LUI Service
- Exchange P-extent Service

-Set method of reporting informational exception conditions
Service

5.1.1 Report P-LUI Service

The report P-LUI service outputs parameters that contain the following information for each P-LUI addressed by the SDA:

Note:The P-LUI of the SDA is included as a parameter.

-Physical Logical Unit Identifier	(P-LUI) (R)
-Peripheral Device Type	(1 per P-LUI)
-Replaceable Unit	(1 per P-LUI)
-State of the P-LUI	(1 per P-LUI)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-LUI defines the address of the physical logical unit.

The peripheral device type indicates the type of logical unit is associated with the P-LUI. See the SCSI-3 Primary Commands Standard of a list of peripheral device types.

The replaceable unit indicates if the P-LUI is a replaceable unit and if it is a Customer Replaceable Unit (CRU) or a Field Replaceable Unit (FRU).

The state of the P-LUI indicates the current state of the the P-LUI. See x.x.x for a list of states.

5.1.2 Report P-LUI Geographic Location Service

The report P-LUI geographic location service output parameters that contains the following information for the selected P-LUI:

-Physical Logical Unit Identifier	(P-LUI) (S)
-Geographic Location	(S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-LUI defines the address of the physical logical unit.

The geographic location is a vendor specific ASCII indication as to the position of the P-LUI within the SDA.

5.1.3 Set P-LUI Geographic Location Service

The set P-LUI geographic location service receives input parameters that contain the following information for the selected P-LUI that the Application Client wants to set the

geographic location:

- Physical Logical Unit Identifier (P-LUI) (S)
- Geographic Location (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-LUI indicates the address of the physical logical unit.

The geographic location is a vendor specific ASCII field that indicates where the P-LUI is located.

| 5.1.4 Report P-extents Service

| The report P-extents service outputs parameters that contain
| the following information for each device addressable by the
| SDA that may be used for user data or check data:

- P-extent (R)
- Peripheral Device Type (1 per P-extent)
- State of the P-extent (1 per P-extent)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-extent defines the location and range of the P-LBAs.

The peripheral device type indicates the type of device associated with the P-extent. See the SCSI-3 Primary Commands Standard of a list of peripheral device types.

The state of the P-extent indicates the current state of the the P-extent. See x.x.x for a list of states.

| 5.1.5 Report Unassigned P-extents Service

| The report unassigned P-extents service outputs parameters
| that contain the following information for each P-LUI
| addressable by the SDA that may be configured but has not yet
| been configured into a set, group, or spare:

- P-extent (R)
- Peripheral Device Type (1 per P-extent)
- State of the P-extent (1 per P-extent)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-extent defines the location and range of the unassigned P-LBAs.

The peripheral device type indicates the type of device associated with the P-extent. See the SCSI-3 Primary Commands Standard of a list of peripheral device types.

The state of the P-extent indicates the current state of the P-extent. See x.x.x for a list of states.

5.1.6 Exchange P-LUI Service

The exchange P-LUI service provides for the replacement of an old device with new device. When an old device is replaced the new device takes on all the logical characteristics of the old device (e.g. redundancy group and volume set mappings). An exact copy of any protected space contents and/or check data on the old device shall be made on the new device.

The exchange P-LUI service receives input parameters that contain the following information:

-Old P-LUI (S)
-New P-LUI (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

Old P-LUI is the address of the device to be replaced. The old P-LUI is removed from any configurations it is currently associated with.

New P-LUI is the address of the device that will replace the old P-LUI.

| An exchange P-LUI service request shall be rejected if an
| exchange is attempted with a new P-LUI that has less capacity
| than the old P-LUI. An exchange P-LUI service request shall
| be rejected if different peripheral device types are selected.

5.1.7 Remove P-LUI Service

The remove P-LUI service provides parameters that indicate the P-LUI to remove. The addressed P-LUI is removed from any configurations it is currently associated with. The P-LUI to remove is defined by the following output parameter:

-P-LUI (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-LUI defines the address of the device to be removed.

The remove P-LUI service request shall be rejected if any

volume sets are associated with the P-LUI to be removed or if a redundancy group has check data contained on any part of the P-LUI to be removed.

5.1.8 Exchange P-extent Service

The exchange P-extent service provides for the replacement of an old P-extents with new P-extents. When an old P-extent is replaced the new P-extent takes on all the characteristics of the old P-extent (e.g. redundancy group and volume set mappings). An exact copy of any protected space contents and/or check data on the old P-extent shall be made on the new P-extent.

The exchange P-extent service receives input parameters that contain the following information:

- Old P-extent (S)
- New P-extent (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

Old P-extent is the address of the P-extent to be replaced. The old P-extent is removed from any configurations it is currently associated with.

New P-extent is the address of the P-extent that will replace the old P-extent.

An exchange P-extents service request shall be rejected if an exchange is attempted with P-extents of different sizes or peripheral device types.

5.1.9 Set Method of Reporting Informational Exception Conditions Service

The set method of reporting informational exception conditions service instructs the SDA how to report informational exception conditions. Informational exception conditions are exception conditions that provide the application client information on condition of the SDA. (e.g. the SDA is predicting a failure, something is broken but the SDA is still operational, etc.)

The set method of reporting informational exception conditions service receives input parameters that contain the following information:

- Method of Reporting Informational Exceptions (S)
- Log Informational Exceptions (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per

service request.

The method of reporting informational exceptions indicates the SDA shall report informational exceptions by one of the following methods:

- Periodic Check Condition Method
- Generate Unit Attention Method
- Asynchronous Event Reporting Method
- No indication of exception condition Method

The periodic check condition method instructs the SDA to report informational exception conditions by periodically returning a sense of CHECK CONDITION on a command issued to the SDA. In response to a REQUEST SENSE command the SDA shall return sense data. The sense key shall be set to RECOVERED ERROR and the additional sense code shall be indicate the the cause of the informational exception condition.

Note:The command that has the CHECK CONDITION shall complete without error before any informational exception condition may be reported.

The generate unit attention method instructs the SDA to report informational exception conditions by issuing a unit attention condition on a command issued to the SDA. In response to a REQUEST SENSE command the SDA shall return sense data. The sense key shall be set to UNIT ATTENTION and the additional sense code shall be indicate the the cause of the informational exception condition.

Note:The command that has the CHECK CONDITION shall not be executed before the informational exception condition is reported.

The asynchronous event reporting method instructs the SDA to report informational exception conditions by using the rules for asynchronous event reporting as described in the SCSI-3 Architecture Model Standard and the relevant Protocol Standard.

The no indication of exception condition method instructs the SDA to not report information exception conditions. To find out about information exception conditions the Application Client shall poll the SDA.

| The log informational exceptions indicates whether or not
| informational exceptions shall be recorded in a log.

5.2 SCSI Disk Array Services

When a create/modify service is requested and the volume set, redundancy group, or spare being addressed was created in a previous create/modify service request the SDA shall add the requested extents to the addressed volume set, redundancy

group, or spare. Extents shall only be removed by deleting the entire volume set, redundancy group, or spare.

The SDA performs the following services to a redundancy group:

- Create/Modify Redundancy Group Service
- Report Redundancy Groups Service
- Report Unassigned Redundancy Group Space Service
- Delete Redundancy Group Service
- Recalculate Check Data Service
- Verify Check Data Service
- Control Generation of Check Data Service
- Rebuild Service

The SDA performs the following services to a volume set:

- Create/Modify Volume Set Service
- Report Volume Sets Service
- Delete Volume Set Service
- Recalculate Check Data Service
- Verify Check Data Service
- Control Generation of Check Data Service
- Control Write Operations Service

The SDA performs the following services to a spare:

- Create/Modify P-extent Spare Service
- Report P-extent Spare Service
- Create/Modify P-LUI Spare Service
- Report P-LUI Spare Service
- Delete Spare Service

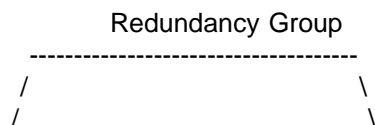
5.2.1 Redundancy Group Services

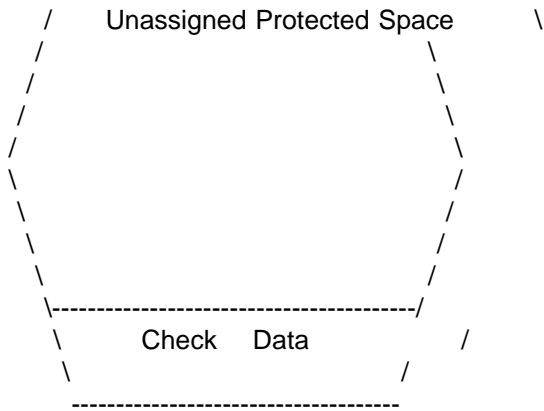
The redundancy group services provide for the creation, modification, reporting, deletion, and control of redundancy groups.

A redundancy group defines the check data and the protected space covered by the check data.

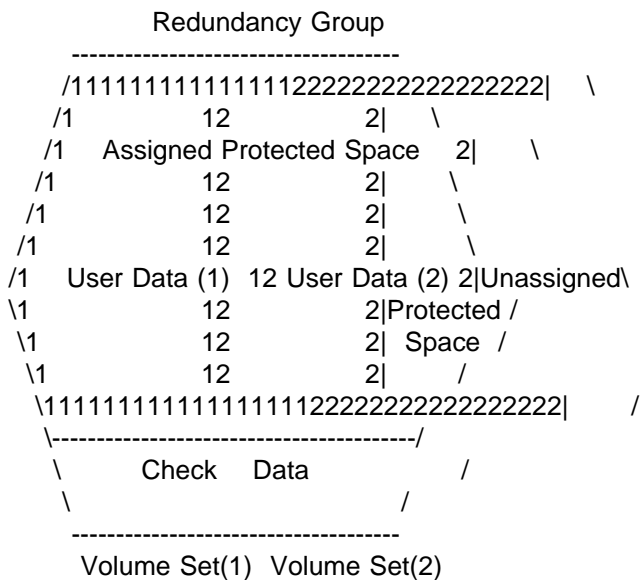
After successful completion of a create redundancy group service the application client shall request a report unassigned redundancy group service to determine the PS-extents that may be used for the creation of volume sets.

The following figure shows the relationship between the check data and protected space before any volume sets have been defined.





The following figure shows the relationship between the check data and protected space after two volume sets have been defined within the redundancy group.



5.2.1.1 Create/Modify Redundancy Group Service

The create/modify redundancy group service provides for the creation of a new redundancy group and the modification of an existing redundancy group.

The create/modify redundancy group service provides parameters that define a redundancy group and the check data mapping within that redundancy group. The redundancy group is defined by the following input parameters:

- Redundancy Group Logical Unit Identifier (R-LUI) (S)
- Granularity of Units (bit, byte, logical block, etc.) (S)
- Redundancy Type Identifier (S)
- P-Extent (R)
- | -Recalculate check data (1 per P-extent)
- | -Preserve protected space (1 per P-extent)

- | -Set protected space to known value (1 per P-extent)
- Start check data interleave P-LBA (1 per P-extent)
- Number of units of check data (1 per P-extent)
- Number of units of user data (1 per P-extent)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The parameter list that defines the redundancy group may contain one or more P-extents. For each P-extent the parameter list shall contain a recalculate check data, preserve protected space, set protected space to known value, start check data interleave P-LBA, number of units of check data, and number of units of user data.

The R-LUI defines the address to be assigned to the redundancy group being created.

The redundancy type identifier parameter defines the check data mapping within the redundancy group as one of the following:

- No Redundancy
- Copy Redundancy
- XOR Redundancy
- P+Q Redundancy
- Vendor Specific Redundancy

In a redundancy group that has a redundancy type identifier of no redundancy user data is not protected. If this type identifier is requested then the following parameters shall all be set to zero:

- Granularity of Units (bit, byte, logical block, etc.)
- Start Check Data Interleave P-LBA
- Number of units of check data
- Number of units of user data

In a redundancy group that has a redundancy type identifier of copy redundancy all user data is replicated on all the P-extents listed in the parameter list. If this type identifier is requested then the following parameters shall be set as indicated:

- Granularity of Units = logical block
- Start Check Data Interleave P-LBA = 0
- Number of units of check data = 0
- Number of units of user data = set to desired value

Note: For copy redundancy the number of units of user data shall be equal for the P-extents within a single create/modify redundancy group service request.

In a redundancy group that has a redundancy type identifier of XOR redundancy or P+Q redundancy the user data is protected by use of check data located within the check data areas. The mapping of the check data is defined by the check data interleave parameters. If this type identifier is requested then the following parameters shall be set as indicated:

- Granularity of Units = set to desired value
- Start Interleave P-LBA = set to desired value
- Number of units of check data = set to desired value
- Number of units of user data = set to desired value

| Note:For XOR redundancy and P+Q redundancy the number of units
| of user data shall be equal for the P-extents within a single
| create/modify redundancy group service request.

In a redundancy group that has a redundancy type identifier of vendor specific redundancy the user data is protected in a vendor specific manor.

The granularity of units indicates if bits, bytes, words, or logical blocks shall be used as the units when mapping the check data.

The P-extent defines the position and range of the redundancy group P-LBAs within a P-LUI. The extent includes check data space if there is any. There may be one or more P-extents per create/modify redundancy group request.

| The recalculate check data indicates whether or not the check
| data for the P-extent being created or modified shall be
| recalculated before a service completion is confirmed.

| The preserve protected space indicates if the information
| contained within the protected space of the P-extent shall be
| preserved during the creation or modification of a redundancy
| group.

| The set protected space indicates the protected space of the
| P-extent shall be set to a specified value and the value to
| set the protected space to.

Start check data interleave P-LBA is the beginning LBA of the check data interleave ratio on the P-LUI.

Note:It is not required that the start check data interleave P-LBA be equal to the start P-LBA.

Number of units of check data is the number of consecutive units to be reserved for check data.

Number of units of user data is the number of consecutive units to be reserved for user data.

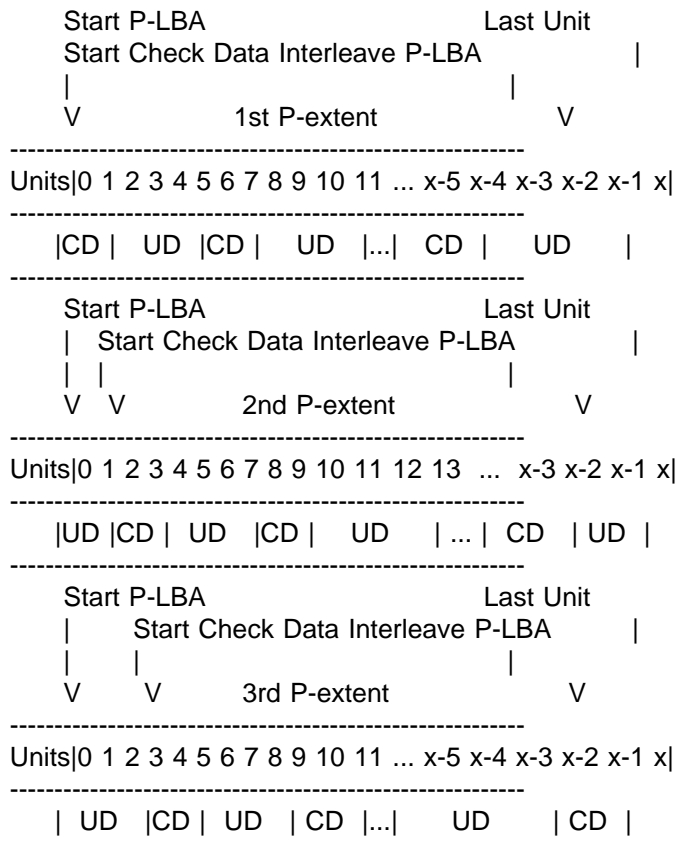
The check data mapping within the P-extent shall repeat until the end of the range defined by the extent is reached. The check data interleave shall begin with check data.

| All space between the start P-LBA and the start check data
 | interleave P-LBA shall be assigned as user data.
 .pa

5.2.1.1.1 Examples of different redundancy types

An example of how XOR redundancy or P+Q redundancy check data would be interleaved with user data in follows:

Number or units of check data = 2
 Number or units of user data = 4



CD = Check Data
 UD = User Data

5.2.1.2 Report Redundancy Groups Service

The report redundancy group service reports information on the configured redundancy groups under control of the addressed SDA.

Note: The information returned by this service may be used to determine the boundaries of all the current redundancy groups to allow more redundancy groups be configured or to allow a redundancy group may be expanded.

| Note: The report redundancy group service may contain more | P-extents than the application client has configured for the | selected redundancy group. The SDA may have added P-extents | based on internal requirements.

The report redundancy groups service provides parameters that define the information to report. The parameters are:

- | -Report all or Selected Redundancy Groups (S)
- Redundancy Group Logical Unit Identifier (R-LUI) (S)

| The report all or selected redundancy groups indicates if all redundancy groups shall be reported or if just the selected redundancy group shall be reported.

The R-LUI defines the address of the redundancy group to report if the report single redundancy group is selected. The R-LUI is not used if the report all redundancy groups is selected.

The report redundancy groups service outputs parameters that contain the following information for each redundancy group within the SDA or, optionally, for a single redundancy group.

- Redundancy Group Logical Unit Identifier (R-LUI) (R or S)
- Redundancy Type Identifier (1 per R-LUI)
- Granularity of Units (bit, byte, etc.) (1 per R-LUI)
- State of the Redundancy Group (1 per R-LUI)
- P-Extent (x per R-LUI)
- Start Check Data Interleave P-LBA (1 per P-extent)
- Number of units of check data (1 per P-extent)
- Number of units of user data (1 per P-extent)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The R-LUI defines the address of the redundancy group.

The redundancy type identifier parameter defines the check data mapping within the redundancy group. See 5.2.1.1 for a description of the redundancy type identifiers.

The granularity of units indicates if bits, bytes, words, or logical blocks were used as the units when mapping the check data.

The state of the redundancy group indicates the current state of the redundancy group. See x.x.x for a list of states.

The P-extent defines the position and range of the redundancy group P-LBAs within a P-LUI. The extent includes check data space if there is any. There may be one or more P-extents per redundancy group.

Start check data interleave P-LBA is the beginning LBA of the check data interleave ratio on the P-LUI.

Note: It is not required that the start check data interleave P-LBA be equal to the start P-LBA.

Number of units of check data is the number of consecutive units reserved for check data.

Number of units of user data is the number of consecutive units reserved for user data.

5.2.1.3 Report Unassigned Redundancy Group Space Service

The report unassigned redundancy group service reports information on the configured redundancy groups under control of the addressed SDA that have protected space not yet assigned to any volume sets.

Note: The information returned by this service is used to determine the boundaries of any unassigned protected space. The unassigned protected space may be used to create new volume sets or to expand existing volume sets.

| Note: The report unassigned redundancy group service may contain more P-extents than the application client has configured for the selected redundancy group. The SDA may have added P-extents based on internal requirements.

The report unassigned redundancy groups service provides parameters that define the information to report. The parameters are:

| -Report all or selected Unassigned Redundancy Groups (S)
| -Redundancy Group Logical Unit Identifier (R-LUI) (S)

| The report all or selected unassigned redundancy groups indicates if all unassigned redundancy groups shall be reported on or if just the selected unassigned redundancy group shall be reported on.

The R-LUI defines the address of the unassigned redundancy group to report on if the report single unassigned redundancy group is selected. The R-LUI is not used if the report all unassigned redundancy groups is selected.

The report unassigned redundancy group space service outputs parameters that contain the following information for portions of all the configured redundancy groups within the SDA that

have not been assigned to a volume set or, optionally, for portions of a single redundancy group that have not been assigned to a volume set:

- Redundancy Group Logical Unit Identifier (R-LUI) (R or S)
- Redundancy Type Identifier (1 per R-LUI)
- State of the Redundancy Group (1 per R-LUI)
- PS-extent (x per R-LUI)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

NOTE: The information provided by this service provides information to the application client so it can place volume set groups on specific devices.

The R-LUI defines the address of the redundancy group.

The redundancy type identifier parameter defines the check data mapping within the redundancy group. See 5.2.1.1 for a description of the redundancy type identifiers.

The state of the redundancy group indicates the current state of the redundancy group. See x.x.x for a list of states.

The PS-extent defines the location and range of unassigned protected space within the R-LUI. The extent excludes check data space. There may be one or more PS-extents per redundancy group.

5.2.1.4 Delete Redundancy Group Service

The delete redundancy group service provides parameters that indicate the redundancy group to delete. The redundancy group to delete is defined by the following output parameter:

- Redundancy Group Logical Unit Identifier (R-LUI) (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The R-LUI defines the address of the redundancy group to be deleted.

The delete redundancy group service request shall be rejected if any volume sets are associated with the R-LUI to be deleted.

5.2.1.5 Recalculate Check Data Service

The recalculate check data service requests the recalculation of check data within a redundancy group.

The recalculate check data service receives input parameters that contain the following information:

-Redundancy Group Logical Unit Identifier (R-LUI) (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The R-LUI defines the address of the redundancy group to be recalculated.

5.2.1.6 Verify Check Data Service

The verify check data service requests a verification that check data is consistent with the protected space within a redundancy group.

The verify check data service receives input parameters that contain the following information:

-Redundancy Group Logical Unit Identifier (R-LUI) (S)
-Enable/Disable Continuous Verification (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The R-LUI defines the address of the redundancy group to be verified.

The enable/disable continuous verification indicates if the check data within the redundancy group shall be continuously verified or only verified one time.

Note:If continuous verification is requested and the SDA supports that option than the service shall be confirmed with no exception conditions. The continuous verification shall execute as an SDA background operation in a user specific manner. Continuous may be defined as only verifying check data associated with the LBA range of any write commands that occur to volume sets that overlay a redundancy group that has continuous verification enabled.

5.2.1.7 Control Generation of Check Data Service

The control generation of check data service requests a that the generation of check data within redundancy group be enabled or disabled.

The control generation of check data service receives input parameters that contain the following information:

-Redundancy Group Logical Unit Identifier (R-LUI) (S)

-Enable/Disable Check Data (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The R-LUI defines the address of the redundancy group that shall have the generation of check data enabled or disabled.

The enable/disable check data indicates if the generation of check data shall be enabled or disabled.

5.2.1.8 Rebuild P-LUI Service

The rebuild P-LUI service requests the rebuild of a portion of a redundancy group.

The rebuild P-LUI service receives input parameters that contain the following information:

-P-LUI (S)
-Redundancy Group Logical Unit Identifier (R-LUI) (R)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-LUI defines the device to be rebuilt.

The R-LUI defines the address of the redundancy group(s) that contain the P-LUI to be rebuilt.

Note:If the P-LUI contains multiple redundancy groups only the range of P-LBAs contained within the R-LUIs from the parameter list for this service shall be rebuilt.

5.2.1.9 Rebuild P-extent Service

The rebuild P-extent service requests the rebuild of a portion of a redundancy group.

The rebuild P-extent service receives input parameters that contain the following information:

-P-extent (S)
-Redundancy Group Logical Unit Identifier (R-LUI) (R)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The P-extent defines the position and range of the P-LBAs to be rebuilt.

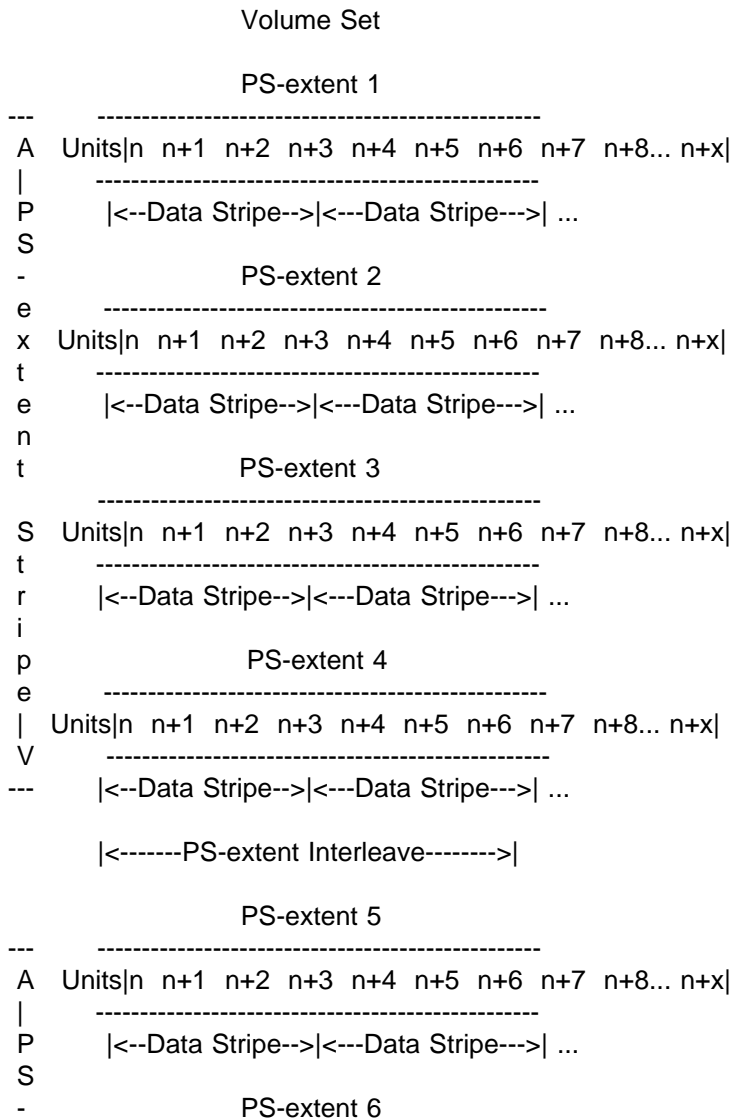
| stripe length to be an even multiple of the number of
 | PS-extents.

PS-extent interleave depth is used to determine the point at which the mapping shall continue onto the next consecutive PS-extent beyond the current stripe. When the last PS-extent is reached, the mapping continues at the first PS-extent. This mapping continues until all the units of the PS-extents are mapped.

Note: The PS-extent interleave depth is only used if the PS-extent stripe length is not equal to the number of PS-extents.

The following figure shows the most general implementation of the three parameters used to map the user data.

.pa



```

e      -----
x Units|n n+1 n+2 n+3 n+4 n+5 n+6 n+7 n+8... n+x|
t      -----
e      |<--Data Stripe-->|<---Data Stripe--->| ...
n
t      PS-extent 7
      -----
S Units|n n+1 n+2 n+3 n+4 n+5 n+6 n+7 n+8... n+x|
t      -----
r      |<--Data Stripe-->|<---Data Stripe--->| ...
i
p      PS-extent 8
e      -----
| Units|n n+1 n+2 n+3 n+4 n+5 n+6 n+7 n+8... n+x|
V      -----
---    |<--Data Stripe-->|<---Data Stripe--->| ...

      |<-----PS-extent Interleave----->|

```

The granularity of units indicates if bits, bytes, words, or logical blocks shall be used as the units when mapping the check data.

The PS-extent defines the position and range of PS-LBAs within a P-LUI that shall be mapped into the volume set's V-LBAs. The order of the PS-extents within the parameter list shall be the order V-LBAs are mapped into the volume set.

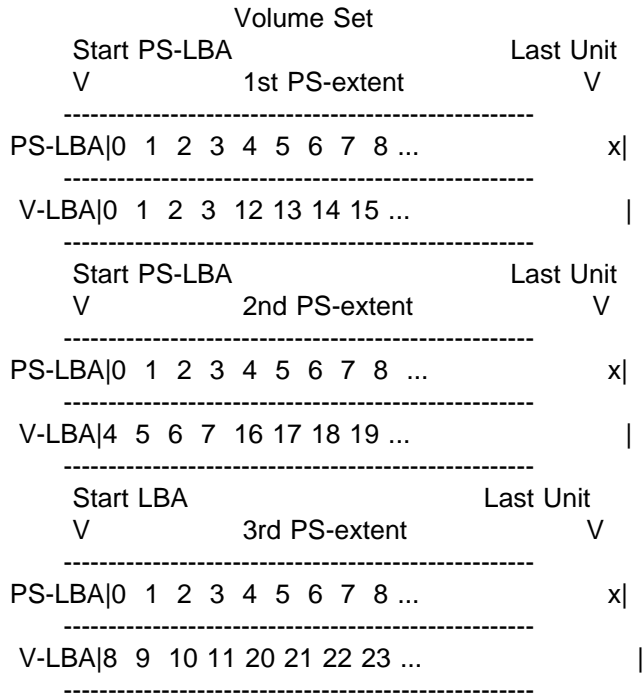
| All the PS-extents within a single create/modify volume set service shall be equal unless the PS-extent strip length equals one.

User data stripe depth is the number of contiguous units to count within a PS-extent before proceeding to the next PS-extent. Any units defined as check data shall not be counted to determine when the user data stripe depth is reached.

The increment/decrement PS-LBA count determines if the PS-LBAs shall be mapping in ascending or descending order within the PS-extent. If increment is selected as the increment/decrement PS-LBA count then PS-LBAs after the start PS-LBA shall be assigned in ascending order. If decrement is selected as the increment/decrement PS-LBA count then PS-LBAs after the start PS-LBA shall be assigned in descending order.

EXAMPLE 1

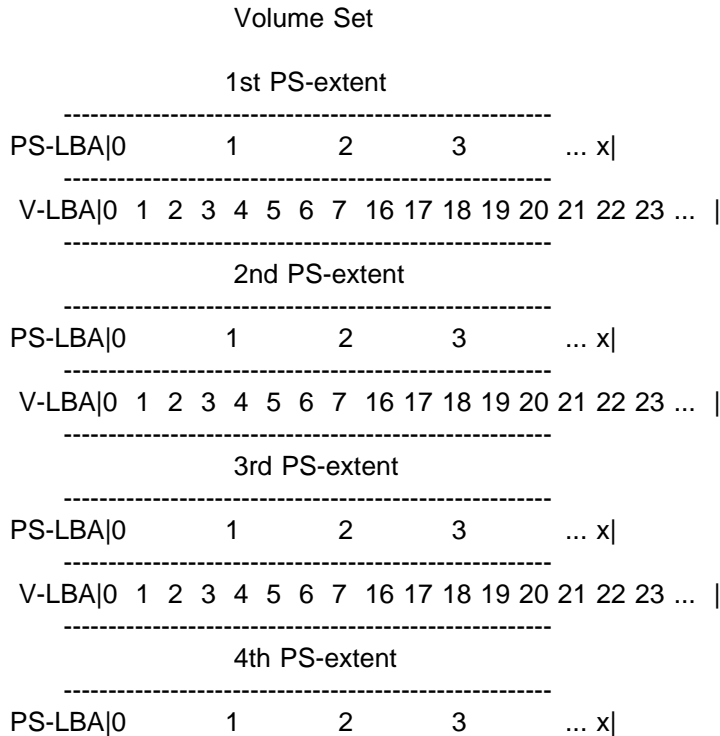
| PS-extent stripe length = 3 extents
 | PS-extent interleave depth = Don't Care
 | Granularity of Units = Logical Blocks
 | User Data stripe depth = 4 units

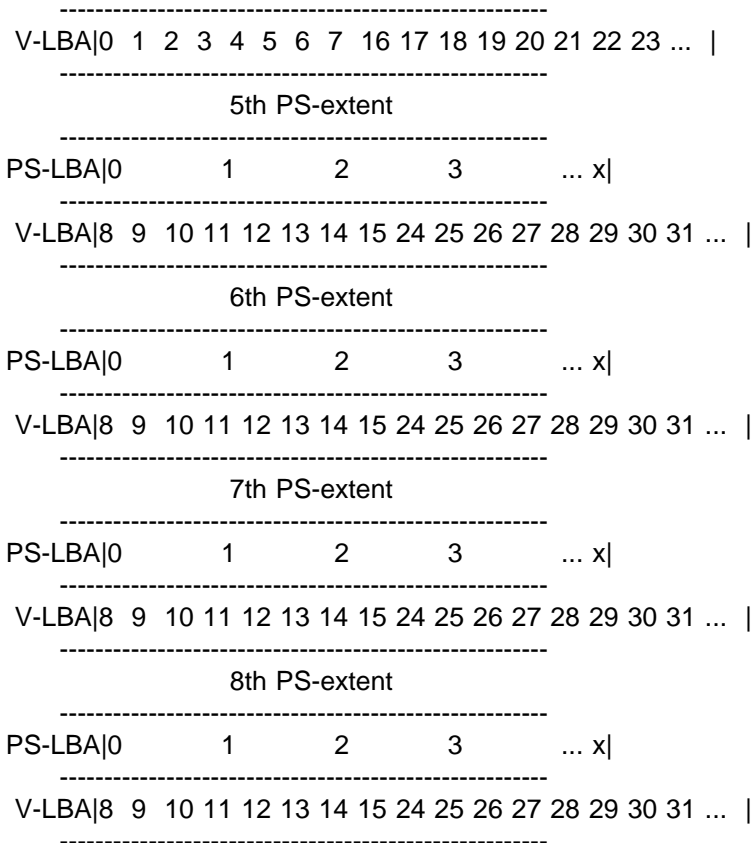


.pa

EXAMPLE 2

- | PS-extent stripe length = 4 extents
- | Granularity of Units = Bytes
- | User Data stripe depth = 128 Units
- | PS-extent interleave depth = 8 stripe depths





5.2.2.2 Report Volume Sets Service

The report volume set service reports information on the configured volume sets under control of the addressed SDA.

| Note: The report volume sets service may contain more volume sets than the application client has configured. The SDA may have added volume sets based on internal requirements.

The report volume sets service provides parameters that define the information to report. The parameters are:

- | -Report all or selected Volume Sets (S)
- | -Volume Set Logical Unit Identifier (V-LUI) (S)

| The report all or selected volume sets indicates if all volume sets shall be reported or if just the selected volume set shall be reported.

The V-LUI defines the address of the volume set to report if the report single volume set is selected. The V-LUI is not used if the report all volume sets is selected.

The report volume sets service outputs parameters that contain the following information for each volume set within the SDA or, optionally, for a single volume set.

- Volume Set Logical Unit Identifier (V-LUI) (R or S)
- PS-extent Stripe Length (1 per V-LUI)
- PS-extent Interleave Depth (1 per V-LUI)
- Granularity of Units (bit, byte, etc.) (1 per V-LUI)
- State of the Volume Set (1 per V-LUI)
- PS-extent (x per V-LUI)
- User Data Stripe Depth (1 per PS-extent)
- Increment/Decrement PS-LBA Count (1 per PS-extent)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set.

PS-extent stripe length indicates the number of contiguous PS-extents counted before looping back to the first PS-extent of the current stripe.

PS-extent interleave depth indicates the point at which the mapping shall continue onto the next consecutive PS-extent beyond the current stripe.

The granularity of units indicates if bits, bytes, words, or logical blocks were used as the units when mapping the check data.

The state of the volume set indicates the current state of the volume set. See x.x.x for a list of states.

The PS-extent defines the position and range of PS-LBAs within a P-LUI that are mapped into the volume set's V-LBAs. The order of the PS-extents within the parameter list indicates the order V-LBAs are mapped into the volume set.

User data stripe depth is the number of contiguous units counted within a PS-extent before proceeding to the next PS-extent. Any units defined as check data are not counted to determine when the user data stripe depth is reached.

The increment/decrement PS-LBA count determines if the PS-LBAs are mapped in ascending or descending order within the PS-extent.

5.2.2.3 Delete Volume Set Service

The delete volume set service provides parameters that indicate the volume set to delete. The volume set to delete is defined by the following input parameter:

- Volume Set Logical Unit Identifier (V-LUI) (S)

Note: Any parameters indicated by a (R) may be repeated. Any

parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set to be deleted.

5.2.2.4 Recalculate V-LUI Check Data Service

The recalculate V-LUI check data service requests the recalculation of any redundancy groups overlaid by a range of V-LBAs.

The recalculate V-LUI check data service receives input parameters that contain the following information:

- Volume Set Logical Unit Identifier (V-LUI) (S)
- Start V-LBA (S)
- Number of V-LBAs (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set that contains the range of V-LBAs to recalculate.

The start V-LBA is the V-LBA to begin the recalculation.

The number of V-LBAs provides the range of the recalculation.

| 5.2.2.5 Verify V-LUI Check Data Service

| The verify V-LUI check data service requests a verification that check data is consistent with the protected space within any redundancy groups overlaid by a range of V-LBAs.

| The verify V-LUI check data service receives input parameters that contain the following information:

- Volume Set Logical Unit Identifier (V-LUI) (S)
- Enable/Disable Continuous Verification (S)
- Start V-LBA (S)
- Number of V-LBAs (S)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set that contains the range of V-LBAs to verify.

The enable/disable continuous verification indicates if the check data within the redundancy groups that overlay the range of V-LBAs shall be continuously verified or only

verified one time.

Note:If continuous verification is requested and the SDA supports that option than the service shall be confirmed with no exception conditions. The continuous verification shall execute as an SDA background operation in a user specific manner. Continuous may be defined as only verifying check data associated with the LBA range of any write commands that occur within the range of V-LBAs that overlay a redundancy group that has continuous verification enabled.

The start V-LBA is the V-LBA to begin the verification.

The number of V-LBAs provides the range of the verification.

5.2.2.6 Control Generation of Check Data Service

The control generation of check data service requests that the generation of check data within redundancy groups overlaid by a volume set be enabled or disabled.

The control generation of check data service receives input parameters that contain the following information:

-Volume Set Logical Unit Identifier (V-LUI) (S)
-Enable/Disable Check Data (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set that shall have the generation of check data enabled or disabled.

The enable/disable check data indicates if the generation of check data shall be enabled or disabled.

5.2.2.6 Control Write Operations Service

The control write operations service requests that write operations to a volume set be enabled or disabled.

The control write operations service receives input parameters that contain the following information:

-Volume Set Logical Unit Identifier (V-LUI) (S)
-Enable/Disable Write Operations (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The V-LUI defines the address of the volume set that shall have write operations enabled or disabled.

The enable/disable write operations indicates if the write operations shall be enabled or disabled.

5.2.3 Spare Services

The spare services provide for the creation, modification, and deletion of spares.

The spare is a P-extent or P-LUI that will replace a failed P-extent or P-LUI. The redundancy group or P-LUI that contains the failed range of P-LBAs or the failed P-LUI must be associated with a spare before the replacement can occur. A spare is associated with a redundancy group or P-LUI using the create/modify spare services.

Note:P-extents shall only replace P-extents and P-LUIs shall only replace P-LUIs.

5.2.3.1 Create/Modify P-extent Spare Service

The create/modify P-extent spare service provides for the creation of new P-extent spares and the modification of existing P-extent spares.

The create/modify P-extent spare service provides parameters that define the boundaries of a P-extent spare and the redundancy group(s) and/or the P-extents it is associated with. The P-extent spare is defined by the following input parameters:

- Spare Logical Unit Identifier (S-LUI) (S)
- P-Extent (S)
- Redundancy Group Logical Unit Identifier (R-LUI) (R)
- Associated P-Extent (R)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The parameter list that defines the P-extent spare may contain zero or more R-LUIs and/or zero or more associated P-extents.

The S-LUI defines the address to be assigned to the P-extent spare being created.

The P-Extent defines the position and range of the P-extent spare. There shall be only one P-Extent per create P-Extent spare service request.

The R-LUI parameter is the address of a redundancy group that shall be associated with this P-extent spare. There may be zero or more redundancy group(s) associated with a single P-extent spare.

The associated P-extent defines the position and range of a P-extent that shall be associated with this P-extent spare. There may be zero or more P-extents associated with a single P-extent spare.

5.2.3.2 Report P-extent Spare Service

The report P-extent spare service reports information on the configured P-extent spares under control of the addressed SDA.

| Note: The report P-extent spare service may contain more P-extents than the application client has configured. The SDA may have added P-extent spares based on internal requirements.

The report P-extent spare service provides parameters that define the information to report. The parameters are:

| -Report all or selected P-extent Spares (S)
-Spare Logical Unit Identifier (S-LUI) (S)

| The report all or selected P-extent spares indicates if all spares shall be reported or if just the selected P-extent spare shall be reported.

The S-LUI defines the address of the P-extent spare to report if the report single P-extent spare is selected. The S-LUI is not used if the report all P-extent spares is selected.

The report P-extent spare service output parameters that contain the following information for each P-extent spare within the SDA or, optionally, for a single P-extent spare.

-Spare Logical Unit Identifier (S-LUI) (S or R)
-P-Extent (1 per S-LUI)
-State of the spare (1 per S-LUI)
-Redundancy Group Logical Unit Identifier (R-LUI)(x per S-LUI)
-Associated P-Extent (x per S-LUI)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The S-LUI defines the address of the P-extent spare.

The P-Extent defines the position and range of the P-extent spare.

The state of the spare indicates the current state of the P-extent spare. See x.x.x for a list of states.

The R-LUI defines the address of any Redundancy group that the P-extent spare has been assigned. There may be zero or more R-LUIs associated with each P-extent spare.

The associated P-extent defines the position and range of any P-extent that has been assigned to this P-extent spare. There may be zero or more P-extents associated with each P-extent spare.

5.2.3.3 Create/Modify P-LUI Spare Service

The create/modify P-LUI spare service provides for the creation of new P-LUI spares and the modification of existing P-LUI spares.

The create/modify P-LUI spare service provide parameters that define a P-LUI spare and the redundancy group(s) and/or the P-LUI(s) it is associated with. The P-LUI spare is defined by the following input parameters:

- Spare Logical Unit Identifier (S-LUI) (S)
- P-LUI (S)
- Redundancy Group Logical Unit Identifier (R-LUI) (R)
- Associated P-LUI (R)

Note: Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The parameter list that defines the P-LUI spare may contain zero or more R-LUIs and/or zero or more associated P-LUIs.

The S-LUI defines the address to be assigned to the P-LUI spare being created.

The P-LUI defines the P-LUI spare. There shall be only one P-LUI per create P-LUI spare service request.

The R-LUI parameter is the address of a redundancy group that shall be associated with this P-LUI spare. There may be zero or more redundancy group(s) associated with a single P-LUI spare.

The associated P-LUI defines a P-LUI that shall be associated with this P-LUI spare. There may be zero or more P-LUIs associated with a single P-LUI spare.

5.2.3.4 Report P-LUI Spare Service

The report P-LUI spare service reports information on the configured P-LUI spares under control of the addressed SDA.

| Note: The report P-LUI spare service may contain more P-LUI
| than the application client has configured. The SDA may have
| added P-LUI spares based on internal requirements.

The report P-LUI spare service provides parameters that define the information to report. The parameters are:

- | -Report all or selected P-LUI Spares (S)
- Spare Logical Unit Identifier (S-LUI) (S)

| The report all or selected P-LUI spares indicates if all spares shall be reported or if just the selected P-LUI spare shall be reported.

The S-LUI defines the address of the P-LUI spare to report if the report single P-LUI spare is selected. The S-LUI is not used if the report all P-LUI spares is selected.

The report P-LUI spare service outputs parameters that contain the following information for each P-LUI spare within the SDA or, optionally, for a single P-LUI spare.

- Spare Logical Unit Identifier (S-LUI) (S or R)
- P-LUI (1 per S-LUI)
- State of the spare (1 per S-LUI)
- Redundancy Group Logical Unit Identifier (R-LUI)(x per S-LUI)
- Associated P-LUI (x per S-LUI)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The S-LUI defines the address of the P-LUI spare.

The P-extent defines the P-LUI spare.

The state of the spare indicates the current state of the P-LUI spare. See x.x.x for a list of states.

The R-LUI defines the address of any Redundancy group that the P-LUI spare has been assigned. There may be zero or more R-LUIs associated with each P-LUI spare.

The associated P-LUI defines any P-LUI that has been assigned to this P-LUI spare. There may be zero or more P-LUIs associated with each P-LUI spare.

5.2.3.5 Delete Spare Service

The delete spare service provide parameters that indicate the spare to delete. The spare to delete is defined by the following input parameter:

- Spare Logical Unit Identifier (S-LUI) (S)

Note:Any parameters indicated by a (R) may be repeated. Any parameters indicated by a (S) shall only occur once per service request.

The S-LUI defines the address of the spare to be deleted.

Spares that have been assigned to any redundancy groups, P-extents, or P-LUIs shall not be deleted.

6.0 Disk Array Conversion Layer Functions

A disk array conversion layer initiates several functions when a media access service is requested. The type of media access and the translation in effect determine which functions are used. The following are the functions available within a disk array conversion layer:

- Translation of input Logical Unit Identifiers to output Logical Unit Identifiers.
- Translation of input LBAs to output LBAs.
- Reading data from and writing data to locations based on the translation in effect for the addressed volume set.
- Calculates and updates the check data (if any).
- Regeneration of protected space contents within the volume set using check data information or duplicate data.
- Rebuild of protected space contents associated with the redundancy group and(or) check data within the redundancy group using the contents of the redundancy group.
- Recalculation of the check data within a redundancy group.
- Recalculation of the check data associated with a volume set.
- Determines when a P-extent should be disabled and/or replaced.
- Returns a confirmation to the higher level as to the success or failure of a service request and, in the case of a failure, gives possible corrective actions.

7.0 Exception Conditions

Exception conditions inform an application client that:

- a change occurred in the physical configuration,
- a change occurred in a volume set configuration,
- a change occurred in a redundancy group configuration
- a change occurred in a spare
- a change occurred in the operation state of the DACL
- a repair action is requested (e.g. disk is predicting failure),
- a repair action is required to restore the volume sets availability (e.g. power supply failure),
- a repair action is required to restore the volume sets level of integrity (e.g. disk fails), or
- an error occurred.

Which exception conditions are returned in confirmations and how often they are returned is based on configuration requests received from an application client.