To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 18 January 2005
Subject: 05-018r1 SAS-1.1 Discover process disable unused entries

Revision history
Revision 0 (24 December 2004) First revision
Revision 1 (18 January 2005) Incorporated feedback from SAS protocol WG

Related documents
sas1r07 - Serial Attached SCSI 1.1 revision 7

Overview
1. For table-routing phys that are running, the discover process needs to disable all route table entries above the ones it fills in (based on discover process optimization rules). Some of the entries may already be unused based on the current detected topology. If the topology gets smaller, there would be stale enabled entries left at the end causing routing problems.

2. For table-routing phys that are not running (no physical link rate, e.g. because nothing is currently attached), the discover process needs to disable all the entries in the column. If it doesn't clear them, when the phy comes back online, the expander starts routing based on stale information.

3. There should be a rule that an expander shall not use the routing table entries for a phy that is not running. If a table-routing phy temporarily goes down, it should not continue to claim addresses for routing (and trigger timeouts rather than OPEN_REJECT (NO DESTINATION)s).

4. The term "end device" is defined as a "SAS device that is not contained within an expander device." However, a SAS device that is so contained (e.g. an SSP target port in an expander for enclosure management) is reported as an "end device" via DISCOVER for the virtual phy attached to that SAS device, and the purpose of the discover process is to learn the addresses of all SAS devices and expander devices (not all end devices and expander devices). So, in several places "end device" needs to be changed to "SAS device."

5. The "Discover process" clause is buried in the expander device model. Since it is still run (at one level) for direct attached end devices, it should really stand alone from the expander device model.

Suggested changes

3.1.58 end device: A SAS device that is not contained within an expander device.

4.6 Expander device model

4.6.7 Expander device routing

4.6.7.1 Routing attributes and routing methods [not shown]

4.6.7.2 Connection request routing

The ECM shall determine how to route a connection request from a source expander phy to a destination expander phy in a different expander port <u>if the destination expander phy is enabled and operating at a valid physical link rate (i.e., the DISCOVER function reports a NEGOTIATED PHYSICAL LINK RATE field set to 8h or 9h)</u> using the following precedence:

1) route to an expander phy with the direct routing attribute or table routing attribute when the destination SAS address matches the attached SAS address;
2) route to an expander phy with the table routing attribute when the destination SAS address matches an enabled routed SAS address in the expander route table;
3) route to an expander phy with the subtractive routing attribute; or
4) return an Arb Reject confirmation (see 4.6.6.3) to the source expander phy.

If the destination expander phy only matches an expander phy in the same expander port from which the connection request originated, then the ECM shall return an Arb Reject confirmation.

1

If the destination SAS address of a connection request matches a disabled routed SAS address in an expander route table, then the ECM shall ignore the match.

4.6.7.3 Expander route table [not shown]

~~**4.6.7.4 Discover process**~~

4.7 Discover process

4.7.1 Discover process overview

Management application clients direct an SMP initiator port to request SMP functions from an SMP target port. Management application clients are located in every SAS initiator device and every self-configuring expander device. The discover process is the process of:

    a)   discovering all the SAS devices and expander devices in the SAS domain (i.e., determining their device types, SAS addresses, and supported protocols); and
    b)   configuring routing tables in the expander devices as needed.

~~The~~ A management application client performing the discover process shall perform a level-order (i.e., breadth-first) traversal of the SAS domain ~~to identify the end devices and expander devices in the SAS domain~~. The order of traversal shall be to discover:

    1)   the device(s) to which the device containing the management application client is attached;
    2)   if ~~the~~an attached device is an expander device, every device attached to that expander device; and
    3)   for each expander device found, every device attached to that expander device.

This order is repeated until all expander devices have been traversed. If the management application client is inside an expander device, then the process shall be repeated on each expander port.

Figure 1 shows an example of level-order traversal.
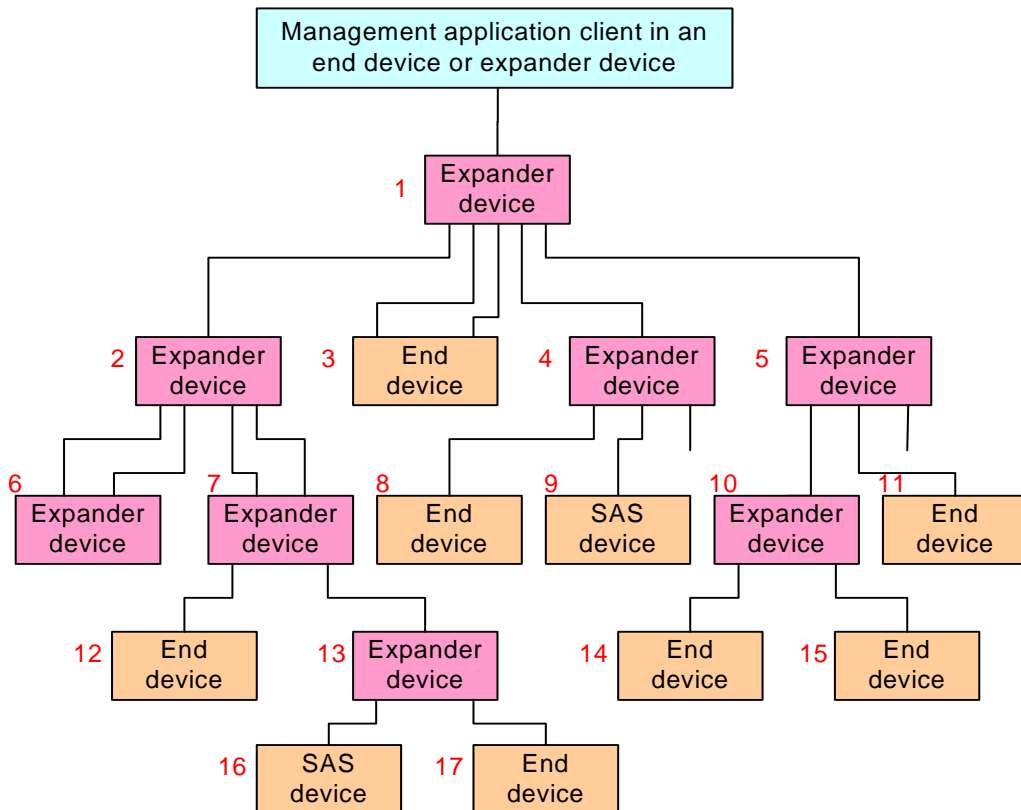
Assume that the lowest phy identifier in each expander device is on the top
right, and the remaining phys are numbered counter-clockwise

Figure 1 — Level-order traversal example

~~The discover process begins with the management application client determining that an expander device is attached.~~ The management application client determines whether an expander device or SAS device is attached at each point in the traversal. For the first device (i.e., the device directly attached), this is determined from the DEVICE TYPE field in the IDENTIFY address frame information received by the phy that the management application client is using. For other devices (i.e., devices not directly attached), this is determined from ATTACHED DEVICE TYPE field the SMP DISCOVER function response.

If an expander device is ~~detected~~attached, the management application client shall use the SMP REPORT GENERAL function (see 10.4.3.3) to determine how many phys are in the expander device, and then use the SMP DISCOVER function (see 10.4.3.5) to determine what is attached to each expander phy (e.g., the device type, SAS address, and supported protocol(s)). If ~~an expander device is detected and its~~the expander device's CONFIGURABLE ROUTE TABLE bit is set to one in the SMP REPORT GENERAL function response, the management application client shall configure its expander route table as described in ~~4.6.7.5~~4.7.3 and 4.7.4.

If ~~an end device is detected~~ SAS device is attached, the discover process need not obtain any more information about the SAS device. Additional discovery software may access that SAS device, however:

a) if the SAS device supports an SMP target port, the management application client may ~~attempt to open an SMP connection to determine if it contains an SMP target port, and use the SMP REPORT GENERAL function to determine additional information about the end device. End devices are not required to support SMP.~~use SMP functions (e.g., REPORT GENERAL and REPORT MANUFAC-TURER INFORMATION) to determine additional information about the SAS target device;

b) If the SAS device supports an SSP target port, a SCSI application client may transmit SCSI commands (e.g., INQUIRY and REPORT LUNS) to determine additional information about the SCSI target device; and

c) If the end device supports an STP target port, an ATA application client may transmit ATA commands (e.g., IDENTIFY DEVICE and IDENTIFY PACKET DEVICE) to determine additional information about the ATA device.

The result of the discover process is that the management application client has the necessary information (e.g., the device type, SAS address, and supported protocols) to communicate with each SAS device and expander device in the SAS domain and each configurable expander device is configured with the necessary expander route entries to allow routing of connection requests through the SAS domain.

Annex K contains an example algorithm used by an application client to perform the discover process.

The discover process may be aborted prior to completion if there is an indication that it may be based on incorrect information (e.g., arrival of a BROADCAST (CHANGE)).

4.7.2 Allowed topologies

The management application client shall allow the following attachments between expander phys:

a) edge expander phy with the subtractive routing attribute attached to an edge expander phy with the subtractive routing attribute;

b) edge expander phy with the subtractive routing attribute attached to an edge expander phy with the table routing attribute;

c) edge expander phy with the subtractive routing attribute attached to a fanout expander phy with the table routing attribute.

If the management application client detects any other combination of expander phy attachment (e.g., expander phy with table routing attached to expander phy with table routing, or an expander phy with direct routing attached to an expander phy), it shall report an error in a vendor-specific manner.

If the management application client detects an overflow of the edge expander route index, it shall report an error in a vendor-specific manner.

If the discover process optimization (see 4.7.3) is disabled and the management application client detects an expander route entry that references the SAS address of the expander device itself (i.e., self-reference), it shall disable the expander route entry by setting the DISABLE EXPANDER ROUTE ENTRY bit to one in the SMP CONFIGURE ROUTE INFORMATION function (see 10.4.3.9). [join paragraphs] The management application client shall disable each expander route entry in the route table by setting the DISABLE EXPANDER ROUTE ENTRY

bit to one in the SMP CONFIGURE ROUTE INFORMATION function (see 10.4.3.9) for each expander phy that has its attached device type set to 000b (i.e., no device attached).

If the management application client detects a port with a SAS address it has already found attached to another expander device, it has found a routing loop and may break the loop by using the SMP PHY CONTROL function (see 10.4.3.10). ~~by disabling~~ to disable all the expander phys attached to that SAS address except for phys in the expander device with the lowest SAS address ~~by using the SMP PHY CONTROL function (see 10.4.3.10)~~.

4.7.3 Discover process optimization

The management application client shall support a discover process optimization that reduces the number of entries required in an expander route table. The method used to enable and disable the discover process optimization is vendor specific.

If the discover process optimization is enabled, then discovered SAS addresses shall be qualified for insertion into the expander device route table by the following:

   a) Discovered SAS addresses referencing the expander device being configured (i.e., self-referencing addresses) shall be excluded from the expander device route table;
   b) Discovered SAS addresses directly attached to the expander device being configured shall be excluded from the expander device route table;
   c) The first instance of a discovered SAS address shall be inserted in the configured expander device route table. Subsequent instances of a discovered SAS address shall be excluded from the configured expander device route table; and
   d) Discovered phys shall be inserted in the expander device route table even if the end device is currently not attached and a SAS address is not available. If a SAS address is not available an entry shall be inserted with the ROUTED SAS ADDRESS field set to 00000000 00000000h and the DISABLE EXPANDER ROUTE ENTRY bit set to one (see 10.4.3.9).

If the discover process optimization is disabled, then all SAS addresses shall be considered qualified for insertion in the expander device route table.

If the management application client supports the discover process optimization, then the management application client should provide a vendor-specific method for initiating a check of the resulting expander route tables. The check should be performed under the following situations:

   a) when an OPEN_REJECT (NO DESTINATION) is received for a connection request to a SAS address that is expected to be in an existing expander device route table;
   b) when a discover process has been completed;
   c) when another SMP initiator port is discovered in the SAS domain; or
   d) when a self-configuring expander device is discovered in the SAS domain.

If the management application client detects an inconsistency in the expander route tables when the discover process optimization is enabled, then the management application client shall report an error in a vendor-specific manner and shall disable the discover process optimization. The management application client should then re-initiate a discover process with the discover process optimization disabled.

~~Annex K contains an example algorithm used by an application client to perform the discover process.~~

~~**4.6.7.5 Expander route index order**~~

4.7.4 Expander route index order

The expander route table shall be configured for each expander phy (in either a fanout expander device or an edge expander device) that has a table routing attribute.~~:~~

   ~~a)   has a table routing attribute; and~~
   ~~b)   is attached to an edge expander device.~~

If the phy is not attached to an edge expander device, every expander route entry for that phy shall be disabled (i.e., the ROUTED SAS ADDRESS field shall be set to 00000000 00000000h and the DISABLE EXPANDER ROUTE ENTRY bit shall be set to one).

If the phy is attached to an edge expander device, the expander route table shall be configured for that phy as follows. For purposes of configuring the expander route table for that phy, the edge expander devices attached to the expander phy are assigned levels:

1) the expander device in which the expander route table is being configured is level 0;
2) the attached edge expander device is considered level 1;
3) devices attached to the level 1 edge expander device, except for the level 0 expander device, are considered level 2;
4) devices attached to level 2 edge expander devices, except for level 1 edge expander devices, are considered level 3; and
5) for each n greater than 3, devices attached to level n-1 edge expander devices, except for level n-2 edge expander devices, are considered level n.

The expander route table for each expander phy shall be configured starting from expander route index 0 by level (i.e., all level 1 entries first, then all level 2 entries, then all level 3 entries, etc.) up to the value of the EXPANDER ROUTE INDEXES field reported by the SMP REPORT GENERAL function (see 10.4.3.3).

Assuming the level 1 edge expander device has expander phys attached to N phys with qualified SAS addresses (see 4.7.3), the first N entries shall be used for those SAS addresses in expander phy order (i.e., the addresses attached to lower expander phy numbers first).

For each of the level 2 devices that:

a) is an edge expander device attached to M phys with qualified SAS addresses; and
b) is attached to an expander phy in the level 1 edge expander device with the table routing attribute,

the next M entries shall be used for the level 2 edge expander device's qualified SAS addresses in expander phy order (i.e., lower phy numbers first).

This process shall repeat for all levels of edge expander devices in the edge expander device set.

After the expander route table has been configured with entries for all levels of edge expander devices, all remaining expander route table entries, if any, shall be disabled (i.e., the ROUTED SAS ADDRESS field shall be set to 00000000 00000000h and the DISABLE EXPANDER ROUTE ENTRY bit shall be set to one). The management application client is not required to disable entries if the topology of expander devices has not changed.

Figure 2 shows a portion of an edge expander device set, where phy A in the root edge expander device is being configured.
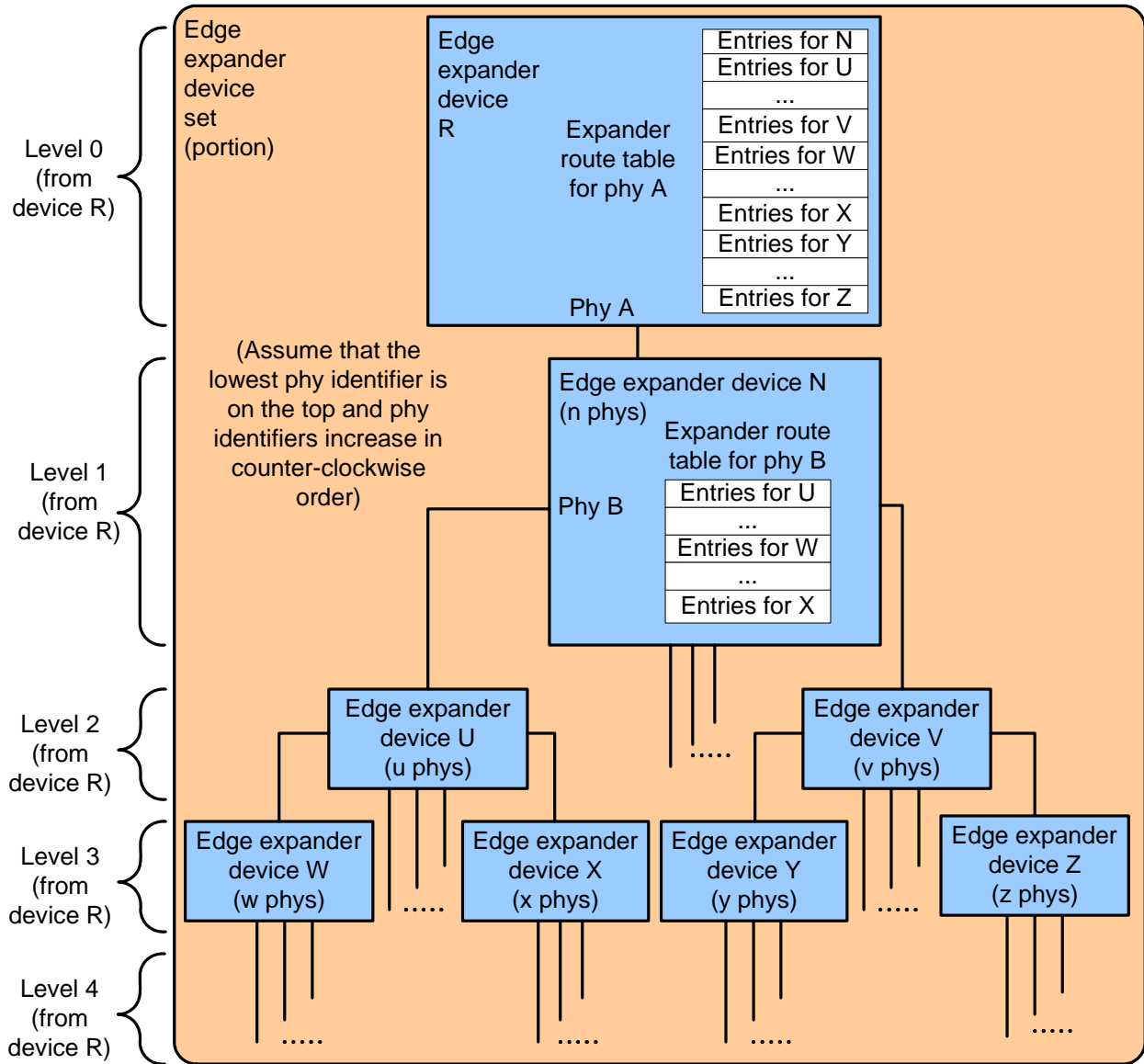
Edge expander device set (portion)

Level 0 (from device R)

Edge expander device R

Expander route table for phy A

| Entries for N |
| Entries for U |
| ... |
| Entries for V |
| Entries for W |
| ... |
| Entries for X |
| Entries for Y |
| ... |
| Entries for Z |

Phy A

Level 1 (from device R)

(Assume that the lowest phy identifier is on the top and phy identifiers increase in counter-clockwise order)

Edge expander device N (n phys)

Expander route table for phy B

| Entries for U |
| ... |
| Entries for W |
| ... |
| Entries for X |

Phy B

Level 2 (from device R)

Edge expander device U (u phys)

Edge expander device V (v phys)

.....

Level 3 (from device R)

Edge expander device W (w phys)

Edge expander device X (x phys)

.....

Edge expander device Y (y phys)

Edge expander device Z (z phys)

.....

Level 4 (from device R)

....

Figure 2 — Expander route index levels example

Figure 3 shows a fanout expander device and an edge expander device set, where phy A in the fanout expander device is being configured.
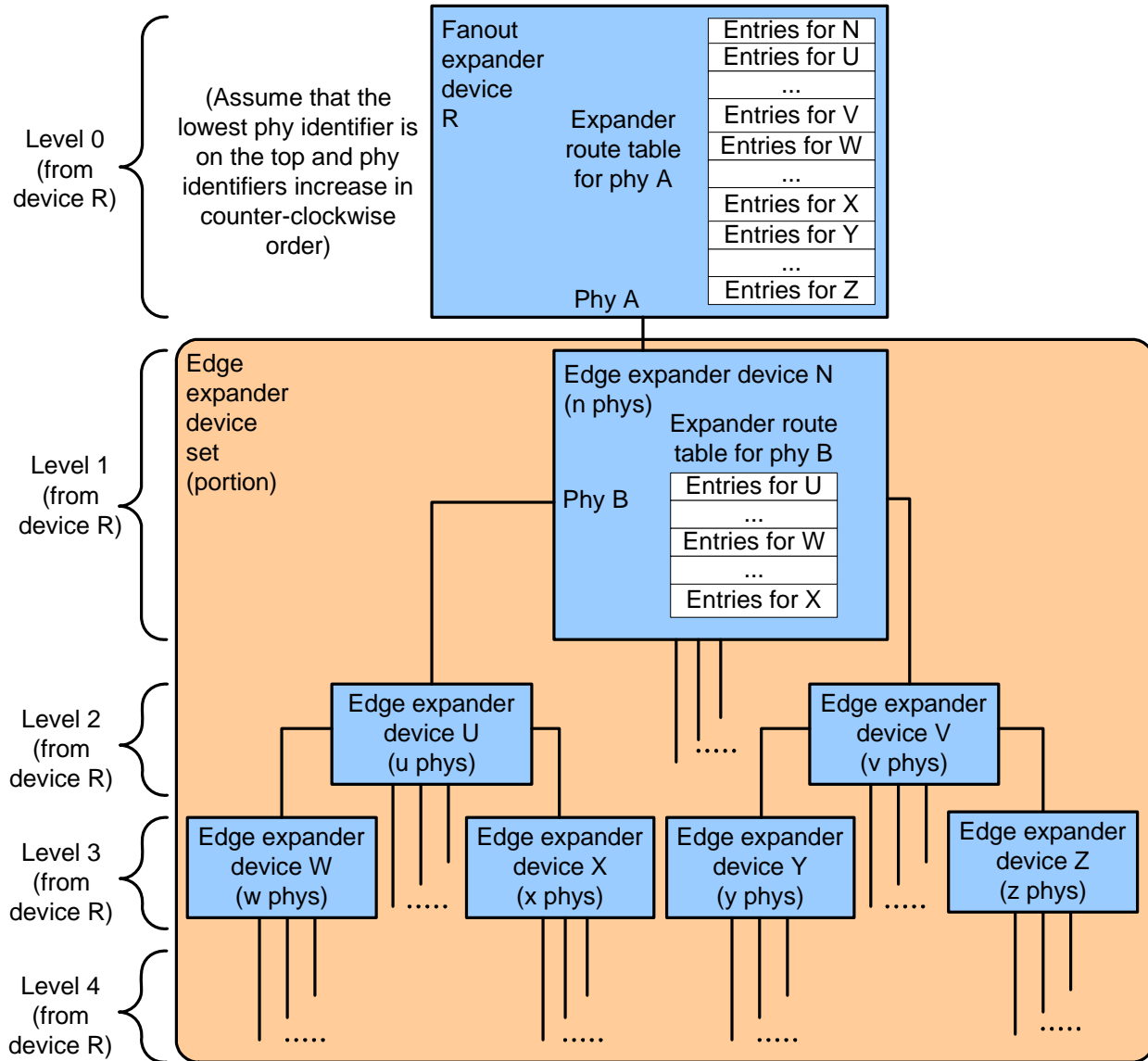

Figure 3 — Expander route index levels example with fanout expander device

Table 1 shows how the expander route table is configured for edge expander device R phy A in figure 2 and the fanout expander device R phy A in figure 3.

Table 1 — Expander route table levels for edge expander device R or fanout expander device R

Expander route index	Expander route entry contents
Level 1 (from device R) entries	
0 .. ($\leq$ n entries)	Qualified SAS addresses attached to edge expander device N
Level 2 (from device R) entries	
($\leq$ u entries)	Qualified SAS addresses attached to edge expander device U
...	...additional qualified SAS addresses for expander devices at level 2...
($\leq$ ventries)	Qualified SAS addresses attached to edge expander device V
Level 3 (from device R) entries	
($\leq$ w entries)	Qualified SAS addresses attached to edge expander device W
...	...additional qualified SAS addresses for expander devices at level 3...
($\leq$ x entries)	Qualified SAS addresses attached to edge expander device X
($\leq$ y entries)	Qualified SAS addresses attached to edge expander device Y
...	...additional qualified SAS addresses for expander devices at level 3...
($\leq$ z entries)	Qualified SAS addresses attached to edge expander device Z
Entries for additional levels	
...	...
Disabled entries	
...	...

Table 2 shows how the expander route table is configured for edge expander device N phy B in figure 2 and figure 3.

Table 2 — Expander route table levels for edge expander device N

Expander route index	Expander route entry contents
Level 1 (from device N) entries	
($\leq$ u entries)	Qualified SAS addresses attached to edge expander device U
Level 2 (from device N) entries	
($\leq$ w entries)	Qualified SAS addresses attached to edge expander device W
...	...additional qualified SAS addresses for expander devices at level 2...
($\leq$ x entries)	Qualified SAS addresses attached to edge expander device X
Entries for additional levels	
...	...
Disabled entries	
...	...

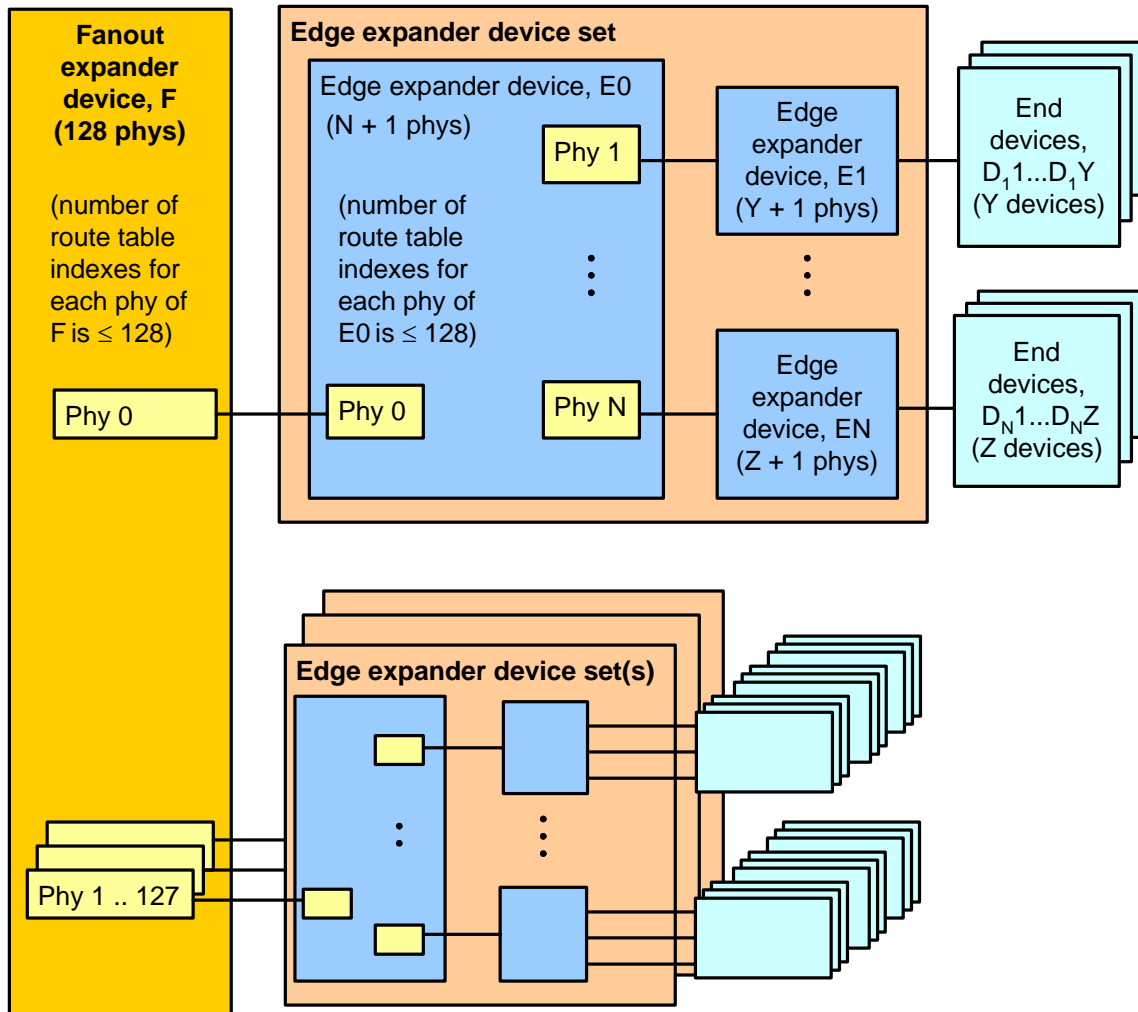Figure 4 shows an example topology with a fanout expander device.

Figure 4 — Expander route index order example

Table 3 shows the expander route index order for the edge expander E0 phy 1 in figure 4.

Table 3 — Expander route entries for edge expander E0 phy 1

Expander route index	Expander route entry contents
Level 1 entries	
0	SAS address (e.g., $D_1 1$) of the device attached to phy 1 of edge expander device E1
1	SAS address (e.g., $D_1 2$) of the device attached to phy 2 of edge expander device E1
...	...
Y - 1	SAS address (e.g., $D_1 Y$) of the device attached to phy Y of edge expander device E1
Level 2 and beyond: no entries	
Disabled entries	

Table 4 shows the expander route index order for the fanout expander F phy 0 in figure 4.

Table 4 — Expander route entries for fanout expander device F phy 0

Expander route index	Expander route entry contents
Level 1 entries	
0	SAS address (e.g., E1) of the device attached to phy 1 of edge expander device E0
...	...additional qualified SAS addresses for edge expander device E0...
N - 1	SAS address (e.g., EN) of the device attached to phy N of edge expander device E0
Level 2 entries	
N	SAS address (e.g., $D_1 1$) of the device attached to phy 1 of edge expander device E1
...	...additional qualified SAS addresses for edge expander device E1...
	SAS address (e.g., $D_1 Y$) of the device attached to phy Y of edge expander device E1
...	...additional qualified SAS addresses for edge expander devices E2 through EN-1...
	SAS address (e.g., $D_N 1$) of the device attached to phy 1 of edge expander device EN
...	...additional qualified SAS addresses for edge expander device EN...
	SAS address (e.g., $D_N Z$) of the device attached to phy Z of edge expander device EN
Level 3 and beyond: no entries since all devices attached to E1 through EN, except for E0, are end devices	
Disabled entries	