To: T10 Technical Committee
From: Steven Fairchild, HP (steve.Fairchild@hp.com)
Date: 9 September 2002
Subject: Expander Configuration Details

The purpose for this document is to define the topology addressing details for the edge and fanout expanders. Most of the wording is to clarify existing concepts, with the exception of a new section on configuration of expander route tables within edge and fanout devices.

Also, the current SAS specification imposes rules about the number of edge devices or fanout devices that may be connected to each other, implying that topologies will fail if the rules are broken. The recommendation is to not restrict the interconnection of any end, edge or fanout device. During the Discovery process, when an initiator determines that an illegal topology has been created, it may take appropriate action. Disabling illegal links and reporting un-reachable SAS addresses as necessary.

Change 1: Definitions added to 3.1

expander route table: A table of SAS addresses internal to an expander that is referenced by the expander to resolve SAS OPEN Frames to a specific outgoing Phy.

expander route entry: A single SAS address reference within an expander route table. The expander route entry is referenced by using the expander route slot and expander route index variable.

expander route slot: A variable used in combination with the expander route index to select a single SAS address reference within an expander route table.

expander route index: A variable used in combination with the expander route slot to select a single SAS address reference within an expander route table.

configurable expander component: An expander component that contains an expander route table that needs to be populated with expander route entries to function within an expander device.

expander component: One or more discrete physical devices that interconnect cooperatively to provide the phy count of an expander device.

subtractive routing: The simple routing capability of an edge expander device. The edge expander device contains a single expander port used to route any SAS address that is not resolved within the edge expander device.

table routing: The routing capability of a fanout expander device, or an edge expander device with configurable expander components. The fanout expander device or edge expander device contains one or more expander ports used to route SAS addresses that are resolved by an expander route table within the expander device.

Change 2: Expander devices

4.1.7 Expander devices
Expander devices are part of the service delivery subsystem. Expander devices contain two or more external expander ports. Expander devices may also include initiator ports and target ports (e.g., an expander device may include an embedded SCSI enclosure services target port) attached to internal expander ports. Expander ports may support being attached to SAS initiator ports, SAS and/or SATA target ports, and to other expander ports.

Figure 9 shows an expander device and the type of ports it may support.

<figure from sas spec>

There are two classes of expander devices:
   a) fanout expander devices: expanders that contain complex routing capability; and
   b) edge expander devices: expanders that only contain simple routing capability.

4.1.7.1 Expander components
Expander devices may be constructed from one or more expander components. Expander components contain two or more external expander ports. Expander components may also include initiator ports and target ports attached to internal expander ports.

Figure X shows expander components used to construct an expander device.
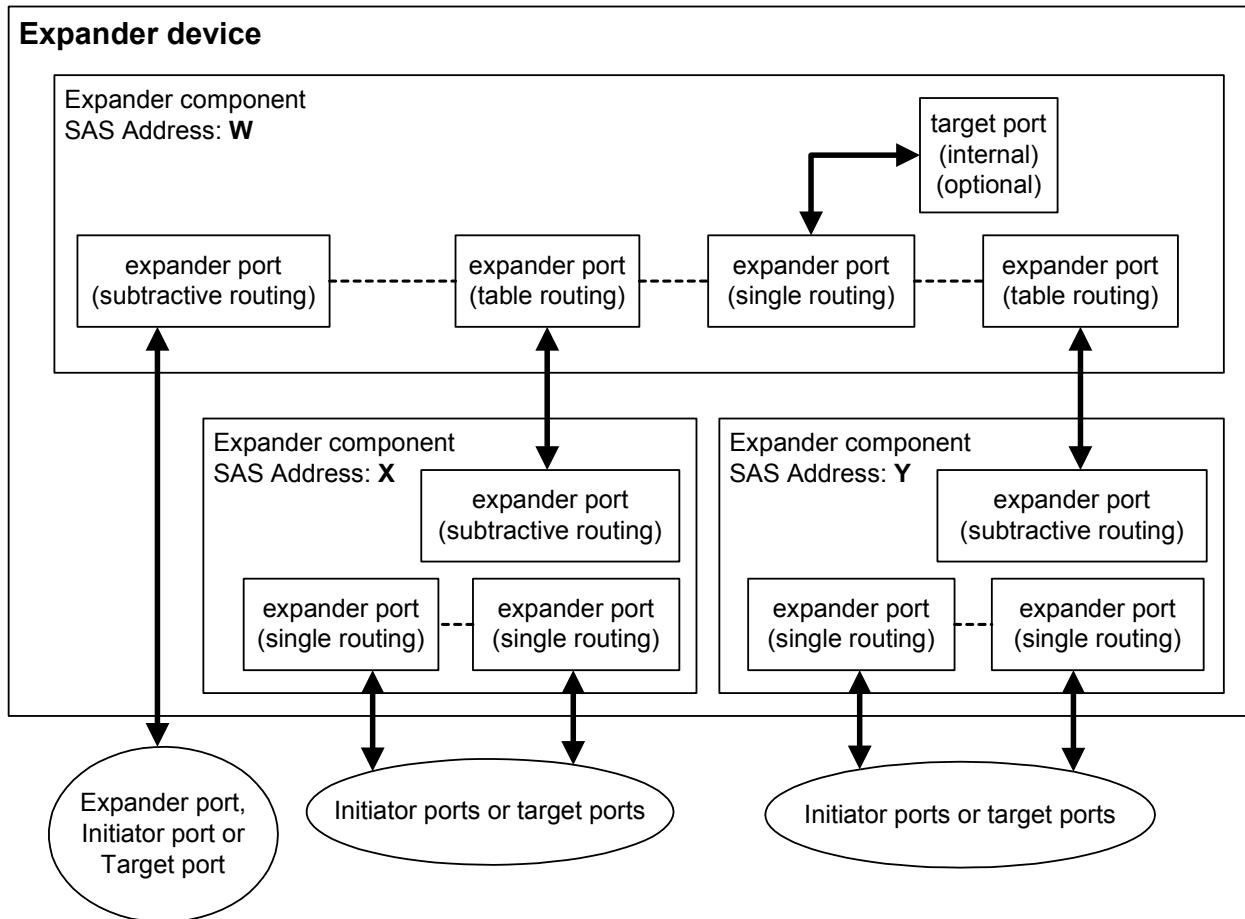

Figure X – Expander device constructed with expander components

An expander device that is constructed of expander components shall have a SAS address identifying each of the expander components. The expander components used to construct expander devices shall be detectable and addressable by their respective SAS addresses during the discovery process (see 4.1.9.2).

4.1.7.1.1 Configurable expander components

Expander components may have configurable routing capability. Expander components with configurable routing capability require an initiator device to use the discovery process (see 4.1.9.2) to populate an internal expander route table (see 4.1.9.1) prior to the expander component being used to route destination SAS addresses. Expander devices may be constructed from expander components with and without configurable routing capability.

Change 3: Expander topologies

4.1.9 Expander topologies
The domain consists of initiator devices, target devices, and expander devices.

No more than one fanout expander device shall be included in a SAS domain. The fanout expander device may be attached to up to 64 edge expanders, initiator port or target ports.

Each edge expander device may be attached to no more than one fanout expander device. Each edge expander device may be attached to up to 64 initiator port or target ports.

An edge expander device may be attached to another edge expander device only if there are no other expanders in the SAS domain.

Figure 12 shows a SAS domain with the maximum number of expander devices.

<figure from sas spec>

Initiator devices and target devices may be attached directly to the fanout expander device, as shown in figure 13.

<figure from sas spec>

4.1.9.1 Expander device SAS address routing
Expander devices may use the following route methods to route SAS addresses from expander ports to attached ports within the expander topology:
  a) Single Routing: expander ports defined as routing only the SAS address attached from the expander port. An expander port with an initiator or target port attached shall use this mechanism to route the SAS address of the attached initiator or target device.
  b) Subtractive Routing: expander port defined as routing any SAS address that is not resolved using the single routing or table routing methods within an expander device. An expander device shall have at most one subtractive routing port. An expander port to support being attached to expander ports predefines subtractive routing. An expander port defined as supporting subtractive routing shall also support attachment of initiator or target ports by single routing.
  c) Table Routing: expander port defined as routing any SAS address that is not resolved using the single routing method by referencing an expander route table to determine the expander port and phy identifier used to route the SAS address. An expander device may have one or more table routing ports. An expander port to support being attached to expander ports predefines table routing. An expander port defined as supporting table routing shall also support attachment of initiator or target ports.

The route method supported by an attached expander port is represented in the ADDRESS DECODE field of the IDENTIFY frame and SMP DISCOVER response.

An expander device with ports that support differing route methods shall determine how to route a destination SAS address with the following precedence:
  a) check for a port with single routing for the destination SAS address, route to the port if found.
  b) check for a port with table routing where the expander routing table contains an expander route entry for the destination SAS address, route to port identified by the expander route entry if found.
  c) check for a port with subtractive routing, route the destination SAS address to the port if found.
  d) if none of these checks result in a successful route of the SAS address, then return an OPEN-REJECT (NO DESTINATION).

4.1.9.1.1 Expander route table

An expander device that supports the table routing method (see 4.1.9.1) shall contain an expander route table. The expander route table is an organized structure that provides an association between destination SAS addresses and expander ports. Each association is an expander route entry.

An expander device may indicate that it is capable of maintaining the expander route entries without initiator device support by setting the CONFIGURABLE ROUTE TABLE field in the SMP REPORT GENERAL response to zero. If the expander device requires an initiator device to provide the expander route entries for the expander route table, then the CONFIGURABLE ROUTE TABLE field in the SMP REPORT GENERAL response is set to one.

The expander route table shall have the expander route entries organized in the same manner; independent of whether the expander device or an initiator device provides them (see 4.1.9.2).

The initiator device references a specific expander route entry within an expander route table by using the EXPANDER ROUTE SLOT and EXPANDER ROUTE INDEX fields in the SMP REPORT ROUTE INFORMATION or SMP CONFIGURE ROUTE INFORMATION functions.

Figure X shows a logical representation of an expander route table with the elements identified.
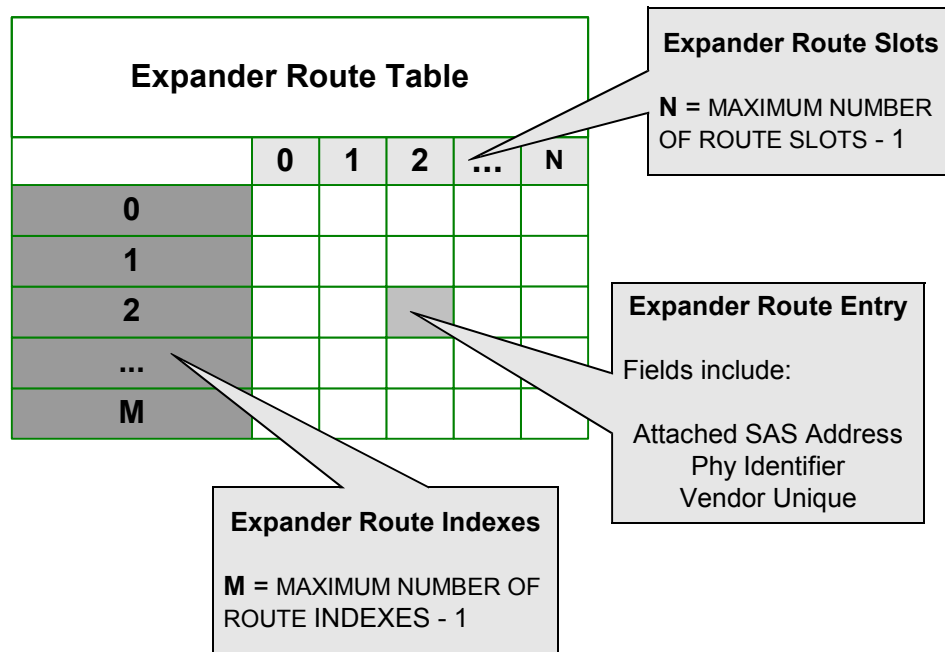
Figure X – Expander route table example

4.1.9.2 Discovery process

The discovery process is the method an initiator device uses to complete a level order traversal of the topology to identify the initiator devices, target devices and expander devices within the expander topology. The order of traversal should be:

1) expander device to which the initiator port is attached;
2) every device attached to that expander device; and

3) if another expander device is found, every device attached to that expander device.

Figures X, Y and Z show an example topology, which is used to describe the discovery process.

SAS Edge Expander, "E1"

Initiator I1

C1

T1 C1

Expander Component C1

0	I1
T1	3
1$_S$	C21
C2	4$_T$
2$_S$	C21
C3	5$_T$

C21-1

C21-2

					T2	C1
					T3	----
					C1	----
					T4	T7
					T5	----
					T6	T9

Single Routing	
Subtractive Routing	
Table Routing	

T2 C2

T3 C2

Expander Component C2

0	T2	T4	3
1	T3	T5	4
2$_S$	C1	T6	5

T4 C2

T5 C2

T6 C2

Expander Component C3

0$_S$	C1	T7	3
1	?	?	4
2	?	T9	5

T7 C3

T9 C3
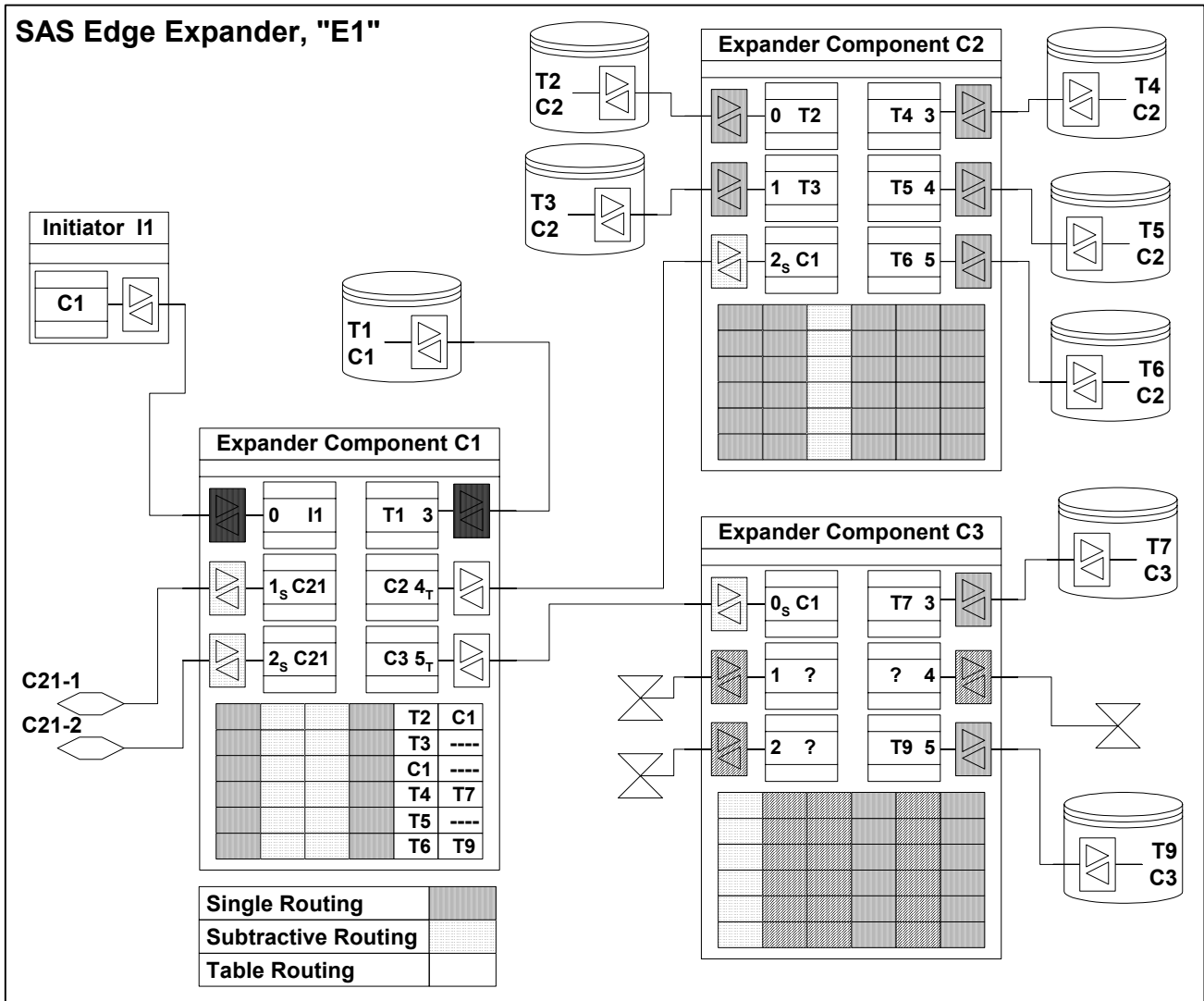
Figure X – Edge Expander, "E1"

Figure Y – Edge Expander, "E2"

SAS Fanout Expander, "F21"

Initiator I24

Expander Component C21

C21

0$_T$ I24   T26 3$_T$ **T26 C21**

1$_T$ C1   C11 4$_T$ **C11-1**

C1-1

2$_T$ C1   C11 5$_T$ **C11-2**

C1-2

MAXIMUM NUMBER OF INDEXES >= 64

MULTIPLE OF PHY COUNT OF EXPANDER COMPONENT

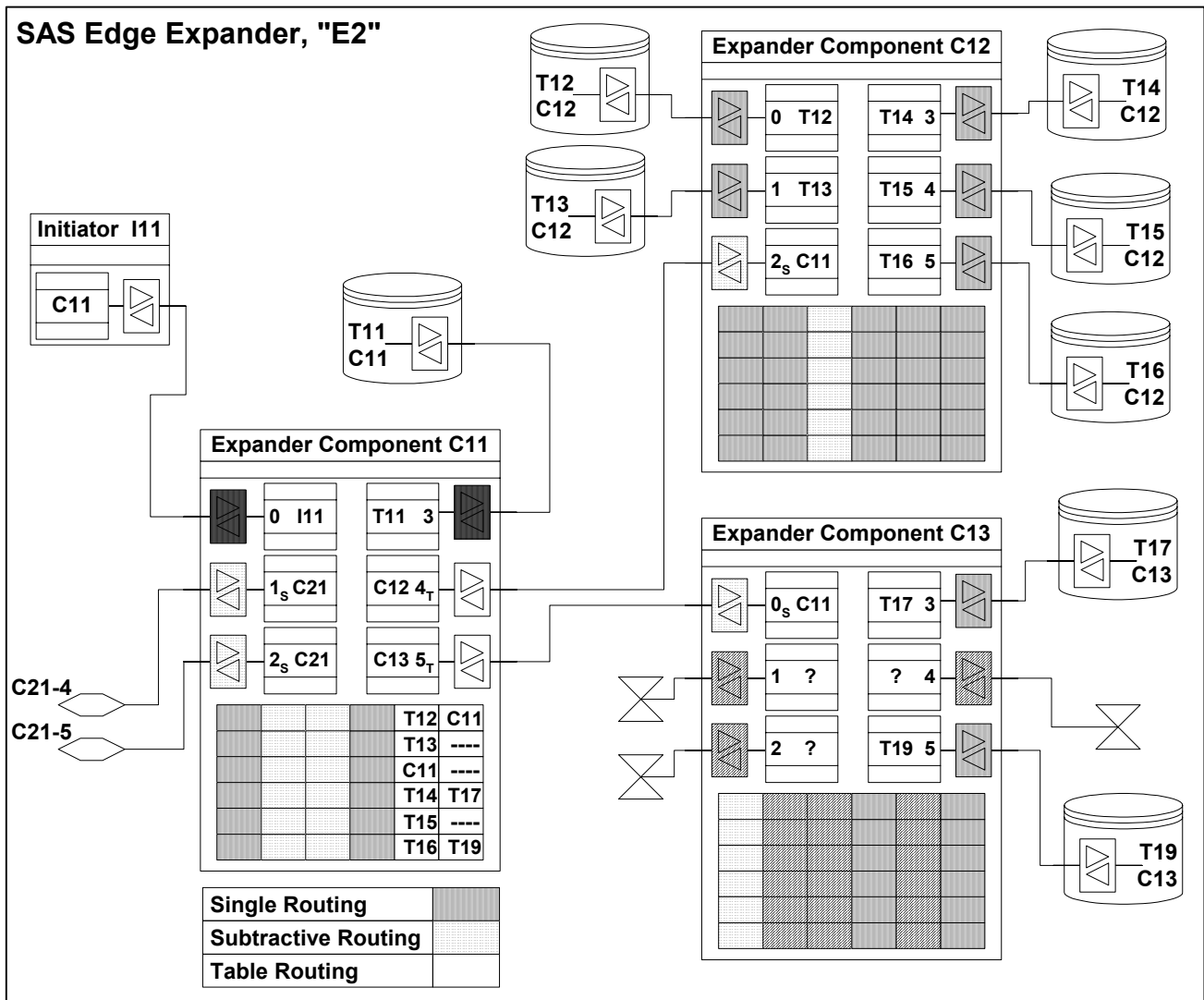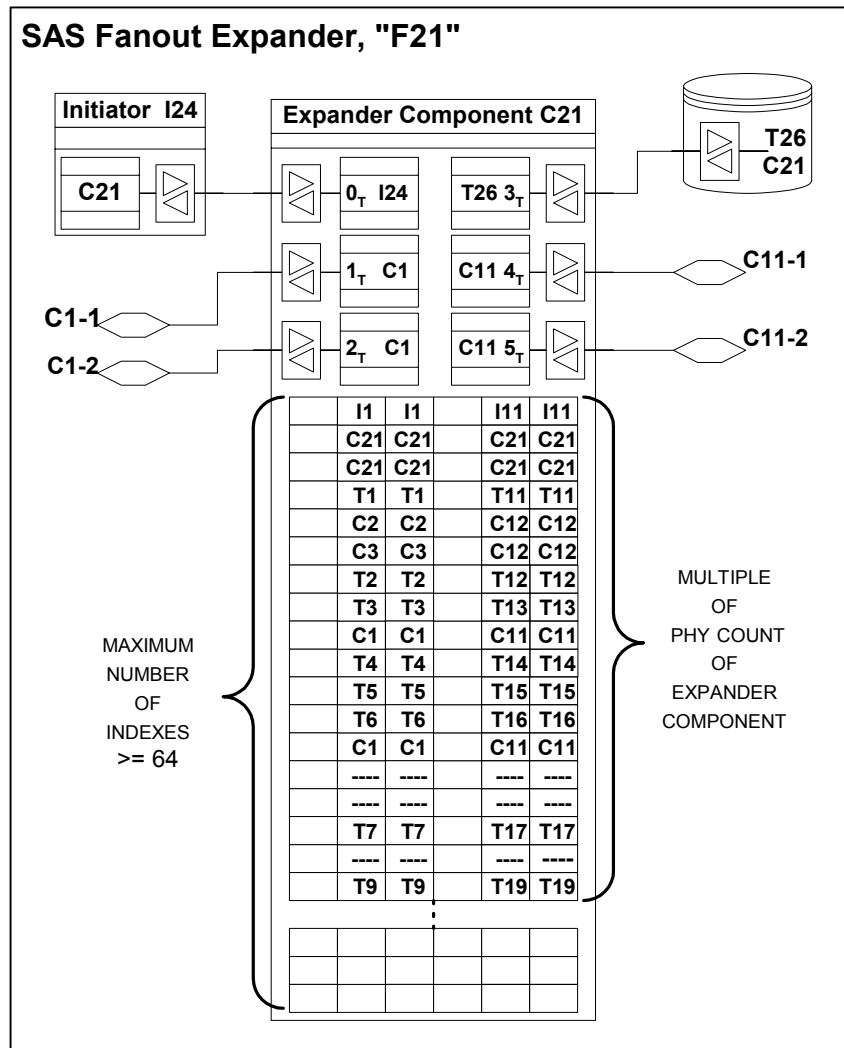|				
I1	I1		I11	I11
C21	C21		C21	C21
C21	C21		C21	C21
T1	T1		T11	T11
C2	C2		C12	C12
C3	C3		C12	C12
T2	T2		T12	T12
T3	T3		T13	T13
C1	C1		C11	C11
T4	T4		T14	T14
T5	T5		T15	T15
T6	T6		T16	T16
C1	C1		C11	C11
----	----		----	----
----	----		----	----
T7	T7		T17	T17
----	----		----	----
T9	T9		T19	T19

Figure Z – Fanout Expander, "F1"

The discovery process begins with the initiator device determining that an expander device is attached. The initiator device makes this determination by checking the DEVICE TYPE field in the IDENTIFY frame from the attached device. If the DEVICE TYPE is either an edge device or fanout device, then the initiator shall determine whether the expander device requires configuration. The initiator device makes this determination by opening an SMP connection to the attached expander device and doing a REPORT GENERAL request and examining the CONFIGURABLE ROUTE EXPANDER bit in the response. If the expander device is configurable, CONFIGURABLE ROUTE EXPANDER bit set to one; then the initiator device also determines the size of the expander route table to configure from the EXPANDER ROUTE SLOTS and EXPANDER ROUTE INDEXES fields.

The initiator device begins traversing the topology by opening an SMP connection to the attached expander device and querying each phy in ascending order (from 0 to NUMBER OF PHYS – 1) using the SMP DISCOVER function. The initiator device shall construct a topology table containing at least the information from the SMP DISCOVER response for each phy queried.

When the initiator device completes the querying of each phy in the attached expander device, it will repeat the query process for each successive expander device phy with a table routing method encountered.

Expander device phys that do not support the table routing method are not traversed during this phase of the process. The result is to group end devices by expander components within an edge expander device.

Using Figure X as an example, initiator (I1) shall have completed the equivalent of Table X after traversing the edge expander (E1).

Expander Route Entry	Expander SAS Address	Phy Identifier	Attached SAS Address	Device Type	Address Decode	Phy Identifier
0	C1	0	I1	end		0
1	C1	1	C21	edge	subtractive	1
2	C1	2	C21	edge	subtractive	2
3	C1	3	T1	end		3
4	C1	4	C2	edge	table (configurable)	4
5	C1	5	C3	edge	table (configurable)	5
6	C2	0	T2	end		0
7	C2	1	T3	end		1
8	C2	2	C1	edge	subtractive	2
9	C2	3	T4	end		3
10	C2	4	T5	end		4
11	C2	5	T6	end		5
12	C3	0	C1	edge	subtractive	0
13	C3	1		none		1
14	C3	2		none		2
15	C3	3	T7	end		3
16	C3	4		none		4
17	C3	5	T9	end		5

Table X – Initiator topology table for example edge expander "E1"

For configurable expander devices, if the successive expander device is also determined to be configurable, then the current configurable expander device shall be configured with the successive expander attached devices before descending beyond the current expander level in the traversal. This step is required so that the configurable expander devices are supplied with sufficient information in the local expander route table to be capable of routing subsequent destination SAS address to the successive expander devices.

Using Figure Z as an example, initiator (I24) shall have completed the equivalent of Table Y after traversing the fanout expander (F1) and encountering the configurable edge expanders (C1 and C11).

Expander Route Entry	Expander SAS Address	Phy Identifier	Attached SAS Address	Device Type	Address Decode	Phy Identifier
0	C21	0	I24	end		0
1	C21	1	C1	fanout	table (configurable)	1
2	C21	2	C1	fanout	table (configurable)	2
3	C21	3	T26	end		3
4	C21	4	C11	fanout	table (configurable)	4
5	C21	5	C11	fanout	table (configurable)	5
6	C1	0	I1	end		0
7	C1	1	C21	edge	subtractive	1
8	C1	2	C21	edge	subtractive	2
9	C1	3	T1	end		3
10	C1	4	C2	edge	table (configurable)	4
11	C1	5	C3	edge	table (configurable)	5
12	C1	0	I1	end		0
13	C1	1	C21	edge	subtractive	1
14	C1	2	C21	edge	subtractive	2
15	C1	3	T1	end		3
16	C1	4	C2	edge	table (configurable)	4
17	C1	5	C3	edge	table (configurable)	5
18	C11	0	I11	end		0
19	C11	1	C21	edge	subtractive	1
20	C11	2	C21	edge	subtractive	2
21	C11	3	T11	end		3
22	C11	4	C12	edge	table (configurable)	4
23	C11	5	C13	edge	table (configurable)	5
24	C11	0	I11	end		0
25	C11	1	C21	edge	subtractive	1
26	C11	2	C21	edge	subtractive	2
27	C11	3	T11	end		3
28	C11	4	C12	edge	table (configurable)	4
29	C11	5	C13	edge	table (configurable)	5

Table Y – Initiator topology table for example fanout expander "F1"

Once the initiator device has completed the construction of the topology table, it must configure the route tables for configurable expanders.

The initiator device begins the configuration of the attached expander route table by opening an SMP connection to the attached expander device and using the SMP CONFIGURE ROUTE function for each expander route slot and expander route index in the table. In the case of a configurable expander device without successive configurable expander devices, a single iteration through the expander route slots and expander route indexes is required to complete the configuration. In the case of a configurable expander

device with success configurable expander device, multiple iterations are required, with additional topology discovery between configuration iterations.

The algorithm for expander route table configuration based on the results in the topology table is shown in Figure W.

Proposal note: Figure W is not available yet. Algorithm will be provided in r1 of this document.

Change 4: Phy identifier

4.2.6 Phy identifier

Each phy in a device shall be assigned a unique 8 bit identifier. The phy identifier is used for management functions.

Phy identifiers shall be numbered sequentially without values skipped and be greater than or equal to 00h and less than 40h.

An expander device constructed of expander components shall not guarantee phy identifier uniqueness across expander components. Each expander component shall guarantee phy identifier uniqueness within the component.

Change 5: Expander device model

4.8.4.1 SAS address mapping

The Expander Function shall provide SAS address to phy identifier mapping using the SAS address and device type, which are exchanged during phy initialization.

Fanout expander devices must perform additional discovery (see 7.6.3) for each attached edge expander device in order to establish an expander route table.

An initiator device must perform the discovery process (see 4.1.9.2) to identify and configure any expander devices with a configurable expander route table.

Change 6: Discover rules

7.6.2 Initiator device specific rules
After identifying that it is attached to an expander device after a link reset sequence, or after receiving a CHANGE primitive sequence, an initiator port should perform a discovery process (see 4.1.9.2).

When this is done after a link reset sequence, this lets the initiator discover information about all the devices in the domain. When this is done after a CHANGE, this lets the initiator port determine what changed in the domain.

This information may be used to select link rates for connection requests.

If the initiator device detects a port with a SAS address it has already encountered, it has found a loop. It should disable routing of destination SAS addresses through the expander port used to reach this previously encountered port to break the loop. The disable expander route field in a SMP Configure Route Function request is used to disable the expander port of an expander device.

7.6.3 Fanout expander device specific rules
After identifying that it is attached to an edge expander device after a link reset sequence, or after receiving a CHANGE primitive sequence, a fanout expander device that does not require initiator device configuration shall follow the initiator device specific rules (see 7.6.2) to discover the topology. Then the fanout device shall follow the edge expander device specific rules (see 7.6.4).

A fanout expander device that requires initiator device configuration shall follow the edge expander device specific rules (see 7.6.4).

7.6.4 Edge expander device specific rules
After completing the identify sequence, an edge expander device is capable of routing destination SAS addresses that are attached directly to it's ports.

An edge expander device that requires initiator device configuration is dependent on at least one initiator device completing the discovery process before it is capable of routing destination SAS address beyond it's ports.

Change 6: Domain changes

7.9 Domain changes

SAS initiator ports scan the domain during a discovery process (see 4.1.9.2) to search for initiator devices, expander devices and target devices after power on or receiving a CHANGE primitive sequence.

The CHANGE primitive sequence shall only be sent outside of a connection. The expander device shall transmit CHANGE to one physical link attached to each expander port. The expander device should not transmit CHANGE to more than one physical link per expander port.

> *Proposal note: After examining the startup process, the recommendation is to remove the restriction on expander devices to not transmit CHANGE to more than one physical link per expander port (statement in red above). The restriction requires complexity in the expander that is unnecessary, since the startup process will typically require the initiator devices to handle multiple CHANGE requests from the same expander phy within short durations as attached devices come ready.*

CHANGE shall not be transmitted to any phy that is part of the expander port that is the cause for sending CHANGE.

Expander devices shall transmit CHANGE for the following reasons:
 a) after an expander phy has lost bit sychronization;
 b) after the link reset sequence completes; and
 c) after the expander device receives CHANGE.

CHANGE shall only be sent outside of connections. CHANGE may be sent by initiator ports to force expander devices to exchange SAS addresses, but should not be sent by target ports.

An expander device is not required to queue multiple CHANGE indications for the same expander port. If a second CHANGE indication is requested before the first indication has been transmitted, the second indication may be dropped.

An edge expander device that detects CHANGE shall follow the edge device specific rules (see 7.6.4).

A fanout expander device that detects CHANGE shall follow the fanout device specific rules (see 7.6.3) to discover the topology.

An initiator port that detects CHANGE shall follow the initiator device specific rules (see 7.6.2) to discover the topology.

Change 7: Connections

7.11.4.2 Edge expander devices

When a edge expander device receives a connection request, it compares the destination SAS address to the SAS addresses of the devices to which each of its phys is attached.

If they match, then the expander device shall arbitrate for access to one of the matching physical links and pass along the connection request.

If they do not match, but an expander device is attached and the request did not come from that expander device, the connection request shall be forwarded to the expander device.

If they do not match, and no expander device is attached or an expander device is attached and the request came from that expander device, the edge expander device shall reply with OPEN_REJECT (NO DESTINATION).

If the output port indicated by the routing table (see table 55) matches the input port, it shall reply with OPEN_REJECT (BAD DESTINATION).

When two edge expander devices are attached, this means requests to non-existent devices return OPEN_REJECT (BAD DESTINATION) rather than OPEN_REJECT (NO DESTINATION). When a fanout expander device is involved, an OPEN_REJECT (NO DESTINATION) is sent.

Table 55 shows the routing table maintained by an edge expander device.

<table in SAS spec>

7.11.4.3 Fanout expander devices

When a fanout expander device receives a connection request, it compares the destination SAS address to the SAS addresses of the devices to which each of its phys is attached. For each port, which is attached to an edge expander, it also compares the SAS addresses to which edge expander device reported being attached.

If it finds a match, it shall arbitrate for access to one of the matching physical links and pass along the connection request.

If it does not find a match, it shall reply with OPEN_REJECT (NO DESTINATION). If the output port indicated by the routing table matches the input port, it shall reply with OPEN_REJECT (BAD DESTINATION).

Table 56 shows the routing table maintained by a fanout expander device.

<table in SAS spec>