

Accredited Standards Committee
X3, Information Processing Systems

Doc: X3T10.1/96a155r0
Date: 21 August, 1996
Project: TL2, S3P
Ref Doc.: X3T10.1/
Reply to: Mark DeWilde

To: X3T10.1 Membership
From: Mark DeWilde
Subject: Comments on TL2 and S3P

BACKGROUND

This proposal is a collection of comments and recommended changes to Rev 2 of TL2 and Rev 2 of S3P. Many of these changes are to clarify and correct the additions from earlier proposals.

PROPOSALS:

Modify Section 3.1 as follows: (numbering deleted)

Dual port node: A node with two ports and a ~~cut through~~ router between them.

Assuming Master: A Pending Master that has received all of the Formal Resignations from the Resigning Masters is the Assuming Master until it has completed the web configuration and is then the web Master. The Assuming Master may receive a Resignation Request and not become the web Master.

Master: A function performed by A single configurator unique node within a web that coordinates error recovery, configures the ports in the web, monitors the web for blocked communications and clears them with web resets, and manages SAT regions. A switch may be the web Master if it contains a configurator function and wins the election of the Master.

Formal Resignation: A Resigning Master Formally Resigns to a node with which it has completed a Resignation Request transaction at an earlier time when it sends a QUERY NODE SMS with the MA bit set to that node.

Pending Master: The current Master of a sub-web that, after walking a newly attached sub web containing one or more active Masters, has discovered that it has won mastership of that new sub web. It waits to assume mastership of the new sub-web until all of the resigning masters formally resign.

Pending Resignation: A Resignation is Pending whenever a QUERY NODE REPLY SMS is sent with the CM bit set. The Pending Resignation is satisfied when the Formal Resignation is performed.

Resigning Master: The (temporary) Master node of a Sub-Web that has received one or more Resignation Requests. It no longer accepts ASYNC ALERT SMS's, but has not yet finished processing those received prior to the QUERY NODE SMS.

Resignation Request: Received by a Master or Resigning master of a sub-web, this is a QUERY NODE SMS that contains either a higher master priority than the receiving node's, or the same Master priority and a higher Unique ID. A resignation request is answered with a QUERY NODE REPLY with the CM bit set.

store-and-forward routing: A process where a frame is completely received and ack'ed prior to routing it to the next node on the path to its destination. Store-and-forward routing introduces a one frame delay per node transversed, and is usually not implemented except when routing between nodes with different speeds.

Sub-Web: A string of nodes with target functions, initiator functions, or a combination of both that are added to another web as a unit.

switch: A node with 4 or more ports (an even number) capable of routing frames from any port to any other port based on the address field of the frame. A switch node shall contain a configurator or responder function in addition to the switch function.

Modify Section 8.4.1, paragraph b1 to read as follows:

After incrementing the WAITING FOR ACK FLAG a port shall wait to receive the ACK character pair for a minimum of 25 μ s and a maximum of 50 μ s before incrementing the ACK bit in the Link Status Byte and invoking the Link ERP process (this range allows a simple binary counter to be used for the timer) for the default SSA link behavior. Ports capable of running the Extended Distance Option shall use a 100 μ s to 200 μ s timer, and reset this timer each time an ACK-ACK character pair is received while the WAITING FOR ACK FLAG is non zero. Extended distance nodes come out of reset with their window size -1 defaulting to "1", and their ACK timeouts defaulting to 100 μ s to 200 μ s. If an extended distance capable port is connected to a port not capable of extended distance, then when the Master sets the window size to "1", the extended distance capable node shall reduce its ack timeout to a minimum of 25 μ s and a maximum of 50 μ s.

Modify Section 8.4.1, paragraph b2 to read as follows: (for extended distance option)

If a port's WAITING FOR ACK FLAG is equal to the window size set in the last configure port SMS still non-zero when it finishes transmitting the CRC field of the next frame then it shall not transmit the trailing FLAG character. Instead it sends NUL characters until either the ACK character pair is received or an ACK time-out occurs. If an ACK time-out occurs in this state, then the port shall abort the frame (see Error! Reference source not found.), increment the ACK bit in the Link Status Byte and invoke the Link ERP process.

Modify Section 8.4.3, to read as follows: (for extended distance option)

Receiver Ready Pacing

Each port has two associated flags, the WAITING FOR RR FLAG and the RR PENDING FLAG, that control pacing as follows:

- b) When a port enters the Enabled state it increments the WAITING FOR RR FLAG and RR PENDING FLAG.

- c) A port may start sending a frame when either of the following conditions is satisfied:
 - 1) the port is in the Ready state and it's WAITING FOR RR FLAG is less than 1 zero;
 - 2) the transmitted frame is a Control frame and the port is in the Enabled, Check or Ready states.
- d) The WAITING FOR RR FLAG is incremented when a port transmits the CONTROL field of a Privileged, Reserved or Application frame and is decremented when the port receives an RR character pair.
- e) When all of the following conditions are satisfied, a port shall transmit an RR character pair after the current character:
 - 1) the port is in the Ready state and the RR PENDING FLAG is non-zero;
 - 2) buffer space is available to receive at least one frame on any active Channel, or shall be available before an incoming frame is received.
- f) The RR PENDING FLAG is incremented non-zero when a port receives the CONTROL field of Privileged, Reserved or Application frame. The RR PENDING FLAG is decremented when the port transmits an RR character pair.
- g) When a port is enabled to operate in extended distance mode by the receipt of a configure port SMS with a window size >1, the WAITING FOR RR FLAG is initialized to the negative of the window size. Each time a frame is sent, the flag is incremented until it equals 1, at which time no additional frames are sent. Each time an RR-RR character pair is received, the flag is decremented.

Change the note at the end of section 8.6 to read as follows:

Note that each item in this list has associated pre-requisite conditions before it is eligible for transmission (e.g., Application and Privileged frames cannot be transmitted if the WAITING FOR RR FLAG is equal to 1 non-zero).

Modify section 9.1.1 to read as follows: (for extended distance option)

Frame buffers

Since the unit of pacing is a frame each port shall provide buffering to receive at least 1 frame. The buffer shall accommodate the maximum-length CONTROL, ADDRESS and DATA fields (i.e., $1+6+128 = 135$ bytes total).

In certain occurrences of a transmission error the Link ERP process retransmits the last ~~one or two~~ frames sent. In nodes implementing the extended distance option, this can be up to 65535 frames, limited by the number of buffers implemented up to the 16 bit extended FSN passed during the link ERP. Therefore a port shall also retain each frame it transmits until it receives the corresponding ACK character pair.

To achieve continuous communication at the full bandwidth of the link it is necessary for each port to have additional buffering. If a port has a pair of transmit buffers available and a pair of receive buffers to provide A/B buffering, one buffer of each pair is emptied/filled by the link while the other is filled/emptied by the application. If the extended distance option is implemented, then the additional buffers permit the sending of additional frames while waiting for earlier transmitted ones to be acknowledged. The number of buffers (B) required to maintain full speed data transmission depends on the data rate (R_{DATA}), the round trip

propagation delay of the transmission medium(T_{RT}), and the size of the frame(S_{frame}). The relationship is as follows:

$$B \geq (T_{RT} * R_{DATA} / S_{frame}) + 1$$

Modify Section 9.2.4 as follows: (for management of configurator table proposals)

A node makes an entry in its Configurator table for each QUERY NODE SMS received, unless it specifies disable registration (DR bit set) or an identical equivalent entry already exists.

When a node executes an ordinary QUIESCE SMS it sets the NUMBER OF PATHS field in the Configurator table to zero, and releases the resources used by all but one of ~~invalidates~~ the associated return path id(s), and but retains the UNIQUE ID entry. A flag is maintained for each return path ID which is set when the table entry has been reported to the Master in response to a Query Registration SMS, and is cleared when it is released as a result of a quiesce or DELETE RETURN PATH ID SMS. The upper level protocols are notified of the quiesce via an asynchronous event notification. along with any upper level protocol specific information. This quiesce action is typically performed when a previously active path from a that Configurator node has been disconnected, and that configurator will be continuing communications over an alternate path. The Configurator node sends the node a QUERY NODE SMS over each path that it intends to continue using. The first registration is guaranteed to succeed since one return path table entry was reserved. The success of subsequent registrations depends on availability of table space. This ensures that any unused RETURN PATHS are removed from the Configurator table.

When a node executes a QUIESCE SMS with the RCTE bit set, it deletes the entire Configurator table entry for the specified UNIQUE ID. This action is performed when a Configurator node has been disconnected from the Web. The upper level protocols are notified of the event through an asynchronous event notification.

The Configurator table is cleared when a node executes a Total Reset or Absolute Reset Control frame. When one of the links to the node becomes Operational again each Configurator node receives an Asynchronous Alert and walks the link with QUERY NODE SMSs. The Master node issues a CONFIGURE PORT SMSs to each port. These actions rebuild the Configurator table.

Modify Section 9.4.2 as follows: (required for extended distance option)

The Speed Negotiation process for dual speed devices attempts to synchronize at 400 MHz for 400 μ sec, then at 200 MHz for 200 μ sec, then at 400 MHz for 100 μ sec, then at 200 MHz for 300 μ sec then repeats the sequence indefinitely. At each speed, the process sends at least 200 DIS characters to allow the remote port to synchronize prior to possibly forwarding a Control frame. There is a 10% tolerance to the numbers to allow firmware implementation rather than forcing hardware implementation. Ports capable of extended distance operation come out of reset behaving as single speed 200 MHz nodes. When the Master configures the two ports on a link to operate using the extended distance option, it sets their NEGOTIATE SPEED fields to the highest common supported speed of the two ports. The two ports then behave as single speed ports at the highest common speed.(see 9.4.1)

Modify Section 9.5.2 as follows: note: this attempts to answer the questions of what happens when a node does not return a query node reply, and how long should you wait for it.

The Base node shall be set to the Configurator node performing the Web Walking process. All ports on the Configurator node shall be marked as unexplored.

The term Base node is defined as the node from which that the Configurator walks a string or loop. The Web Walking process is defined in the following steps in the order shown, and shall be performed until all ports are explored:

- b) The Configurator node selects a port (the "current port") on the base node and marks it as explored. The Web walking process is repeated for each unexplored operational port on the base node.
- c) The Configurator node shall send a QUERY NODE SMS with the DR bit set out the current port to the next unexplored node. destination node incrementing the path address by one. ~~The configurator node starts a 60 millisecond timer waiting for a until a~~ QUERY NODE REPLY. If no query node reply is received before the timer expires, the query node SMS is resent and the timer restarted. If no reply is received, a total reset is issued to the node. The configurator then waits 100 milliseconds and retries the query node with timeout one more time. If the timer expires before receiving the query node reply, then the node is assumed to be inoperative and the walk continues on the next unexplored port. If a query node reply is received before the timer expires on any query node reporting:
 - 1) only one port operational indicating the end of a string has been reached;
 - 2) a Unique ID that already exists in the Topology table indicating a loop has been established;
 - 3) or the TOTAL OTHER PORTS field value exceeds two indicating that a switch has been discovered.
- d) The Configurator node may send a QUERY PROTOCOL SMS and QUERY PORT SMSs as desired.
- e) If a switch is discovered the Configurator node shall send a QUERY SWITCH SMS to determine the number of operational ports on the switch. The base node shall be set to that switch node and the Web Walking process shall be recursively called. The port indicated in the QUERY NODE REPLY SMS shall be marked as explored.

Modify Section 9.6As Follows (if 96a154r0 is accepted):

Resets

This clause defines the various types of reset that are associated with SSA-TL2. Upper level protocols that use TL2 may have reset functions of their own. These reset functions shall not affect TL2 data structures or TL2 managed hardware, and shall not affect other upper level protocols that may be supported by the same TL2 layer. Upper level protocol resets shall not affect any hardware or data structures in such a way as to alter the functioning of the TL2 layer or other upper level protocols. Resets of the TL2 layer do not in general affect the system in which the node resides, but rather are limited in scope to the node itself.

Modify Section 9.6.2 As Follows (if 96a154r0 clarification of resets proposal is accepted):

Total Reset or Absolute Reset

Total Reset is used to reset the transport layer software and hardware of SSA, and leaves other hardware and protocol layers unaffected directly. Absolute Reset is virtually the same in effect as a power-up reset, causing the reset of all node hardware, ULPs, and invoking POST, if applicable. The extent of the reset is limited to the SSA Device only, not the entire system in which the SSA device may reside. Absolute Reset shall invoke at least the same functions as

~~Total Reset. Optionally, some implementations treat Absolute Reset as a stronger form of Total Reset, invoke additional internal resets and invoke the POST. If no additional resets are done, the destination node processes an Absolute Reset frame the same as a Total Reset frame.~~
Storage requirements for handling Absolute Reset frames may be shared with Total Reset frames.

A Configurator node may reset a destination node by originating either a Total Reset or an Absolute Reset Control frame that is addressed to the node. All Configurator nodes shall be capable of originating Total Reset, Reserved or Absolute Reset Control frames. The destination node shall process a Total Reset Control frame as follows:

- b) Perform optional resets and optional POST;
- c) Clear the OPERATIONAL FLAG;
- d) All ports are set to the Disabled state and Privileged mode;
- e) Any upper level protocols running on the node are notified of the reset by means of an Asynchronous Event Notification. Other node functions are optionally re-initialized (e.g., the interface processor in a node) by an Absolute Reset frame;
- f) The frame buffers shall be initialized;
- g) The Transport Layer's pool of Tags will be initialized
- h) Any hardware such as DMA engines under the control of the Transport Layer will be initialized, reservations of hardware by ULPs canceled.
- i) The TRANSMIT POINTER and RETRY POINTER shall be cleared;
- j)

Modify Section 9.6.4 as indicated below: (for clarification of resets proposal)

The process to reset a loop or all of the strings connected to a particular Base node is as follows in the order shown:

- a) The Master selects an Unselected port on the Base node as the new Previous port and marks it as selected.
- b) The Master issues a CONFIGURE PORT SMS (or the internal equivalent) to Previous port. This ensures that the Master receives subsequent Asynchronous Alerts from the Previous port.
- c) The Master sends a QUERY NODE SMS to the Current node and starts a 5 ms timer.
- d) If the Master does not receive a QUERY NODE REPLY SMS before the timer expires, it continues with step e5. Otherwise, if the Master receives a QUERY NODE REPLY SMS, the Master shall compare the UNIQUE ID field in the QUERY NODE REPLY SMS with each entry in its Configuration table. If the Unique ID is in the Configuration table then the Master shall mark the port indicated by the Current port field in QUERY NODE REPLY SMS as Selected and shall skip to step i9; the Master has returned on a loop or a cyclic path to itself or a switch that it has already reset. If the Unique ID is not in the Configuration table, the Master shall continue with step e5; since the Current node has not yet been reset.
- e) The Master sends a Total Reset to the Current node, starts a 1055 ms timer, and waits to receive an ASYNC ALERT SMS with an ALERT CODE value of PORT NOW OPERATIONAL from Previous port. Except for issuing an ASYNC REPLY SMS, the Master shall ignore any other ASYNC ALERT SMSs (for example the Total Reset may cause the Previous port to initially invoke the Asynchronous Alert process with an ALERT CODE value of REMOTE PORT DISABLED).
- f) If the Master does not receive an ASYNC ALERT SMS with an ALERT CODE value of PORT NOW OPERATIONAL from Previous port before the 5 ms timer expires then the Master

- shall send an Absolute Reset to the Current Node, starts a 5 second timer, and waits to receive an ASYNC ALERT SMS with an ALERT CODE value of PORT NOW OPERATIONAL from Previous port. If the Master does not receive an ASYNC ALERT SMS with an ALERT CODE value of PORT NOW OPERATIONAL from Previous port before the 5 second timer expires then the Master shall skip to step i9; the Current node shall be assumed to be not operational. Otherwise the Master shall send a QUERY NODE SMS to Current node and starts a 5 ms timer.
- g) If the Master does not receive a QUERY NODE REPLY SMS before the timer expires, then the Master shall skip to step i9; the Current node shall be assumed to be not operational. Otherwise the Master shall make an entry in its Configuration table for the Current node.
 - h) If the Current node is a single port node, the Master shall continue with step i9. If the Current node is a dual port node, the Master updates the Previous port to refer to the port on Current node that is further from the Master and returns to step b2. If the Current node is a switch, the Master shall mark the port indicated by the PORT field in QUERY NODE REPLY SMS as Selected and shall continue with step i9.
 - i) If there is an Unselected port remaining on the Base node, the Master returns to step 1. Otherwise the walk from the current Base node is complete and the Master selects a new Base node if one exists.

Modify Section 10.1.3 as follows: (for extended distance option)

- b) A Link Reset frame is transmitted. If the link is operating in extended distance mode, then an EXTENDED LINK RESET frame is transmitted instead of the LINK reset frame. If no acknowledgment is received the Link Reset or EXTENDED LINK RESET frame is retransmitted once. If no acknowledgment is received to the retransmitted Link Reset or EXTENDED LINK RESET frame, then the Link ERP process waits 25 ms to ensure that the remote port recognizes an unrecoverable error, sets the port to the Disabled state, and exits with an ALERT CODE value of LINK RESET FAILED.

Modify section 10.1.3 as follows: Note: where are the terms “transmit pointer” and “retry pointer” defined? Depending on the definitions, the following section may need some rework....

- h) The number of outbound frames with an acknowledgment outstanding is calculated. When using strict cyclic buffers this is obtained as follows:

$$Q = (\text{TRANSMIT POINTER} - \text{RETRY POINTER}) \bmod N$$

where N is the window size in use by the transmitter. number of transmit buffers provided. Q is 0 or 1 frames for nodes not using the extended distance option, or an integer number of frames up to the window size for nodes using the extended distance option.

The number of outbound frames that the local port is expecting an acknowledgment for but have not been received by the remote port is also calculated, where local refers to the ports TRANSMIT SEQUENCE NUMBER and remote refers to the RSN field received in the Link Status Byte.

$$P = (\text{local TRANSMIT SEQUENCE NUMBER} - \text{remote RSN}) \bmod N_4$$

where P is less than or equal to Q.

If the node is not using the extended distance option, and either $Q > 1$ or $P > Q$, or if the node is using the extended distance option and $Q > \text{Window Size}$ or the difference

of the extended local TRANSMIT SEQUENCE NUMBER - extended remote RSN is greater than the window size. -then the Link ERP process waits 25 ms to ensure that the remote port recognizes an unrecoverable error, sets the port to the Disabled state and it exits with an ALERT CODE value of INVALID RETRY STATUS.

- i) The Link ERP process arranges to resend the number of frames indicated by subtracting P from its TRANSMIT POINTER, modulo N. Those outbound buffers that do not need to be retransmitted shall be discarded using the following algorithm.

```
Do while RETRY POINTER ≠ TRANSMIT POINTER;
    Discard buffer at RETRY POINTER;
    Increment RETRY POINTER modulo N;
End;
```

Modify section 10.1.3.2 as follows: (for extended distance option)

If the Link ERP process exits unsuccessfully then each node shall take the following actions:

- 1) Clear the port's OPERATIONAL FLAG.
- 2) If the port was performing extended distance operation, return the port to normal distance operation.
- 3) Set the port to Privileged mode. The transmitter discards Application frames. This avoids frames backing up and blocking other traffic. It also prevents any further data from being transferred through the affected port.
- 4) Invoke the Asynchronous Alert process (see Error! Reference source not found.) with the ALERT CODE specified in the Link ERP exit.
- 5) Restart the affected port, as described in Error! Reference source not found..

Modify section 10.6.1.6 as follows: (for extended distance option)

- b) a Privileged, Reserved or Application frame and no buffer is available, (i.e., when the RR PENDING FLAG is = 1 non-zero);

Modify section 10.5 as follows: (clarification)

- h) A node detects a recovered Link ERP process threshold is exceeded.
- i) A change in the list of supported Upper level Protocols has occurred since the last query node/query protocol received.

Modify Section 11 as follows: (for clarification only)

An SMS has a maximum length of 32 bytes. Only the bytes defined in the appropriate SMS structure are required, however if the SMS is padded, it ~~may~~ be padded with any number of bytes up to a total length of 32 bytes. If the source node pads the SMS, the pad bytes shall be zero. The destination node shall ignore any pad bytes.

Modify table 21 as follows: (adding SMSs to support configurator table management)

Table 1 - SSA-TL2 messages supported

SMS	SMS CODE	SMS FRAME TYPE	Node type support	
			Sent from	Received by
QUERY NODE	00h	Privileged	Configurator	all
QUERY NODE REPLY	01h	Privileged	all	Configurator
CONFIGURE PORT	02h	Privileged	Master	all
RESPONSE	03h	Privileged	all	Configurator
ASYNCR ALERT	04h	Privileged	all	Master
MASTER ALERT	05h	Privileged	Master	Configurator
QUIESCE	06h	Privileged	Configurator	all
ASYNCR REPLY	07h	Privileged	Master	all
QUERY PROTOCOL	08h	Privileged	Configurator	all
QUERY PROTOCOL REPLY	09h	Privileged	all	Configurator
QUERY PORT	0Ah	Privileged	Configurator	all
QUERY PORT REPLY	0Bh	Privileged	all	Configurator
QUERY SWITCH	0Ch	Privileged	Configurator	all
QUERY SWITCH REPLY	0Dh	Privileged	all	all
REQUEST SAT REGION	0Eh	Privileged	Configurator	Master
REQUEST SAT REGION REPLY	0Fh	Privileged	Master	Configurator
<u>QUERY REGISTRATION</u>	<u>10h</u>	<u>Privileged</u>	<u>Master</u>	<u>Responder</u>
<u>DELETE RETURN PATH ID</u>	<u>11h</u>	<u>Privileged</u>	<u>Master</u>	<u>Responder</u>
reserved for Privileged frames	12h-1Fh	Privileged	reserved	reserved
DATA READY	20h	Application	all	all
DATA REPLY	21h	Application	all	all
DATA REQUEST	22h	Application	all	all
DATA REDIRECT	23h	Application	all	all
reserved for Application frames	24h-7Fh	Application	reserved	reserved
Defined by ULP	see 0	defined in ULP		

Note: The first byte of all SMSs defined by Upper Level Protocols shall be the SMS Code assigned to that Upper Level Protocol, and is used by the Transport Layer to Route the SMS to that Protocol. The maximum length of Upper Level Protocol SMSs shall be 32 bytes.

Modify Section 11.2.4 as follows:

Table 2 - MASTER ALERT SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (05h)							
1	PORT							
2	TAG							
3	TAG							
4	RETURN PATH							
5	RETURN PATH							
6	RETURN PATH							
7	RETURN PATH							
8	UNIQUE ID							
9	UNIQUE ID							
10	UNIQUE ID							
11	UNIQUE ID							
12	UNIQUE ID							
13	UNIQUE ID							
14	UNIQUE ID							
15	UNIQUE ID							
16	ALERT CODE							
17	ALERT CODE							
18	ALERT CODE							
19	reserved							
20	CONTROL							
21	CHANNEL							
22	CHANNEL							
23	FRAME DATA							
24	FRAME DATA							
25	FRAME DATA							
26	FRAME DATA							
27	FRAME DATA							
28	FRAME DATA							
29	FRAME DATA							
30	FRAME DATA							

The PORT field identifies the port that invoked the Asynchronous Alert.

The TAG field is assigned by the Master and returned in the response from the destination Configurator node. It shall be unique among all the tags currently active from the Master.

The RETURN PATH field specifies the path component that shall be placed in the ADDRESS field of the RESPONSE SMS.

The UNIQUE ID field identifies the node that invoked the Asynchronous Alert.

The ALERT CODE field consists of three parts as shown in Error! Reference source not found., and is copied from the same field of the associated ASYNC ALERT SMS.

The CONTROL field is copied from the same field in the associated ASYNC ALERT SMS.

The CHANNEL field is copied from the same field in the associated ASYNC ALERT SMS.

The FRAME DATA field is copied from the first ~~eight~~^{two} bytes of the same field in the associated ASYNC ALERT SMS.

Modify section 11.2.6 as follows: (Interlocked election changes)

QUERY NODE SMS

This SMS is sent from a Configurator node to every other Operational node during the Configuration process. QUERY NODE is also used as a remote wrap test to verify the integrity of the Path.

The QUERY NODE SMS is defined in Table 3.

Table 3 - QUERY NODE SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (00h)							
1	SSA-TL VERSION (10h)							
2	TAG							
3	TAG							
4	RETURN PATH							
5	RETURN PATH							
6	RETURN PATH							
7	RETURN PATH							
8	UNIQUE ID							
9	UNIQUE ID							
10	UNIQUE ID							
11	UNIQUE ID							
12	UNIQUE ID							
13	UNIQUE ID							
14	UNIQUE ID							
15	UNIQUE ID							
16	DR	MA	<u>MY MASTER PRIORITY</u>			reserved		

The SSA-TL VERSION field defined in Table 4 identifies the version of SSA-TL being used by the sender. If multiple SSA-TL versions are supported, the node shall report the highest level supported.

Table 4 - SSA-TL VERSION field values

Version	Description
00h	SSA-TL1 implementations before standardization
01h	SSA-IA/95PH
02h	SSA-TL1 standard
03h	reserved
04h	SSA-IA/96PH
05h-0Fh	reserved
10h	This (SSA-TL2) standard
11h-FFh	reserved

The TAG field is returned in the QUERY NODE REPLY SMS. The TAG is assigned by the Configurator node and it shall be unique among the TAG values that are currently active from that Configurator node.

The RETURN PATH field specifies the path component that shall be placed in the Configurator table entry created in response to this QUERY NODE SMS, if an entry is created. This value is used

for the address field of the resulting QUERY NODE REPLY SMS, and is used for the address field of any future application SMS that utilizes this Configurator table entry.

The UNIQUE ID field contains the Unique ID of the Configurator node that issued QUERY NODE SMS.

The DR bit (Disable Registration) controls the updating of the Configurator table in the node. If the DR bit is cleared then the node shall enter the specified RETURN PATH and UNIQUE ID into its Configurator table. If a Configurator node intends to use several alternative paths to the same node then it shall issue QUERY NODE SMS with the DR bit is cleared once over each path. If the DR bit is set the Configurator table shall not be updated.

The MA bit (Master Alive) is set when the Master invokes the Master Alive process (see Error! Reference source not found.), and is cleared all other times. When a node receives a QUERY NODE SMS with the MA bit set, it restarts the 10 second Master Alive timer (see Error! Reference source not found.). The MA bit is also used during the Master Election process to indicate that a resigning master has resigned.

The MY MASTER PRIORITY field is valid for Master Capable nodes and contains the Master Priority of the node sending the QUERY NODE SMS.

Modify section 11.2.6 as follows: (for interlocked election of master)

QUERY NODE REPLY SMS

The QUERY NODE REPLY SMS as defined in Table 5 is returned when a QUERY NODE SMS is received. The QUERY NODE REPLY SMS is returned on the same port that received the corresponding QUERY NODE SMS.

Table 5 - QUERY NODE REPLY SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (01h)							
1	PORT							
2	TAG							
3	TAG							
4	UPPER LEVEL PROTOCOL							
5	ITF	MASTER PRIORITY				reserved		
6	TOTAL OTHER PORTS							
7	SSA-TL VERSION (10h)							
8	UNIQUE ID							
9	UNIQUE ID							
10	UNIQUE ID							
11	UNIQUE ID							
12	UNIQUE ID							
13	UNIQUE ID							
14	UNIQUE ID							
15	UNIQUE ID							
16	RETURN PATH ID							
17	RETURN PATH ID							
18	RETURN PATH ID							
19	RETURN PATH ID							
20	P10	P20	LONG	<u>CM</u>	reserved			

The PORT field indicates the number of the port currently being used.

The TAG field is copied from the QUERY NODE SMS. It identifies the QUERY NODE SMS that this reply is being generated for.

The UPPER LEVEL PROTOCOL field identifies the upper-level protocol that the node shall respond to. The UPPER LEVEL PROTOCOL field shall contain a value from Table 6.

Table 6 - Upper-level protocol code values

ULP code	Protocol	Notes
00h	USE THE QUERY PROTOCOL SMS TO LIST UPPER LEVEL PROTOCOLS SUPPORTED	1
01h	SHALL RESPOND TO NO UPPER-LEVEL PROTOCOL	1
02h	Vendor specific	1,2
03h-7Fh	reserved	
80h	SSA-IA / 95SP	1,2
81h	reserved	
82h	SSA-S2P	1,2
83h	SSA-S3P	1,2
84h-FAh	reserved	
FBh-FFh	Vendor specific	2
NOTES 1) Valid in QUERY NODE REPLY. 2) Valid in QUERY PROTOCOL REPLY		

The Configurator table full (ITF) bit is set when there is no space left in the Configurator table to make an entry for the Configurator node that sent QUERY NODE SMS.

The MASTER PRIORITY field defines the priority of the node for becoming the Web Master. A value of 000b indicates that the node is a Responder node and is not capable of functioning as a Master. Any value greater than 000b indicates the node’s priority for becoming the Master. A value of 001b is the lowest priority and 111b the highest

NOTE 1 - A Configurator node optionally fixes its MASTER PRIORITY at 4 (the default priority) or provides some mechanism outside of SSA to change its MASTER PRIORITY dynamically.

The TOTAL OTHER PORTS field contains a value that is one less than the number of ports implemented. If this value exceeds two, the QUERY SWITCH SMS is used to retrieve a port mask for the switch.

NOTE 2 - The TOTAL OTHER PORTS field value shall be 0 (single port), 1 (dual port) or an odd number greater than 1 (switch) as per clause Error! Reference source not found..

The SSA-TL field defined in Table 4 identifies the version of SSA-TL being used by the sender. If multiple levels of SSA-TL are supported, then highest value shall be reported that is equal to or less than the SSA-TL field in the associated QUERY NODE SMS

The UNIQUE ID field contains the node’s Unique ID.

The RETURN PATH ID contains a value created by the node and returned to the Configurator node. This field shall be used by the Configurator node in any future application SMSs that are utilizing the same path as used by the QUERY NODE SMS that caused this QUERY NODE REPLY SMS to be generated.

The P10 bit is only valid for single port and dual port nodes, Switch nodes shall clear the P10 bit. For single port or dual port nodes, the P10 bit is set if port 1 is operational, and is cleared if port 1 is not operational.

The P20 bit is only valid for single port and dual port nodes, Switch nodes shall clear the P20 bit. For single port or dual port nodes, the P20 bit is set if port 2 is operational, and is cleared if port 2 is not operational.

The LONG bit is set if any port in the node is equipped with the Extended Distance Option, and cleared otherwise. If the long bit is set, a QUERY PORT SMS may be used to determine the window size of each port.

The CM bit is set when the node sending the NODE REPLY SMS is currently the Master of it's sub-web.

The Path component of the ADDRESS field in a QUERY NODE REPLY SMS frame is copied from the RETURN PATH field in the corresponding QUERY NODE SMS. All padding bytes shall be discarded.

Modify section 11.2.13 as follows: (to improve layering of protocol)

QUIESCE SMS

The QUIESCE SMS defined Table 7 is sent from a Configurator node to another node for the purpose of invalidating return path Ids and releasing the resources associated with them in the transport layer for a given unique ID. This is used during error recovery to abort all transport layer data transport I/O processes that were originated/received from the Configurator node specified by the UNIQUE ID. The node aborts the associated data transfer I/O processes, updates the Configurator table entry, notifies the upper level protocols of the event using an asynchronous event notification, and returns a RESPONSE SMS. Upper level protocols will differ in their response to the notification. Protocols such as S2P and S3P will abort I/O processes associated with the quiesced node. The node does not return status for an aborted I/O process unless it has already completed.

Table 7 - QUIESCE SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (06h)							
1	RCTE	reserved						
2	TAG							
3	TAG							
4	RETURN PATH							
5	RETURN PATH							
6	RETURN PATH							
7	RETURN PATH							
8	UNIQUE ID							
9	UNIQUE ID							
10	UNIQUE ID							
11	UNIQUE ID							
12	UNIQUE ID							
13	UNIQUE ID							
14	UNIQUE ID							
15	UNIQUE ID							

When the RCTE bit (remove Configurator table entry) is set, the node shall remove the Configurator table entry associated with the UNIQUE ID field. When the RCTE bit is cleared, the node shall delete all but one of the the return paths in the Configurator table entry associated with UNIQUE ID field, and mark the remaining one as invalid, thus reserving the table space. This guarantees at

least one table entry will be available when the node issuing the quiesce re-registers with the node.-

NOTE 3 - It is recommended that the RCTE bit only be set by a Configurator node quiescing it's own I/O processes, by the Master or by the resigning Master node.

The TAG field is assigned by the Configurator node and returned in the RESPONSE SMS from the node. The TAG is assigned by the Configurator node and it shall be unique among the TAG values that are currently active from that Configurator node.

The RETURN PATH field specifies the path component that shall be placed in the ADDRESS field of the RESPONSE SMS.

The UNIQUE ID field specifies the Unique ID of the Configurator node whose return path ID associated resources are to be released. ~~I/O processes are to be aborted.~~

Specifying the Configurator node with a Unique ID rather than a return path allows the Master to issue a third party Quiesce on behalf of a missing Configurator node.

The node shall search its Configurator table to convert the Unique ID to a return path before selecting the data transport! ~~processes to be aborted.~~ If the Configurator node is using multiple paths to the same node then there is more than one return path for the given Unique ID and the node shall Quiesce the data transport! ~~processes for all paths.~~

If a Configurator node is using multiple paths to the same node then it shall issue a QUIESCE SMS over each path that remains connected. When it receives the RESPONSE SMS for a particular path the Configurator node safely assumes that no frames remain in the Web for I/O processes issued over that path.

In order to release configurator table space on the other node, an~~An ordinary QUIESCE SMS shall be issued by configurators that have lost a path to another node, as indicated by a Master Alert. The node will then re-register over any paths it intends to continue using. A QUIESCE SMS with the RCTE bit set is sent by a node to release all of the configurator table space on the other node when no more additional I/O to that node is needed. This policy of registering, performing I/O and un-registering in a multi-initiator system maximizes the number of nodes that may share access to a given node. also causes the node to delete all return paths from the selected Configurator table entry. The Configurator node issues a QUERY NODE SMS with the DN bit cleared over each path that it intends to continue using. This protocol ensures that there are no stale return paths left in the Configurator table when the Web configuration changes. A third party QUIESCE SMS causes the node to delete the entire Configurator table entry for the specified Unique ID.~~

The acknowledgment for the QUIESCE SMS is the receipt of a RESPONSE SMS with the same tag value as was used for the QUIESCE SMS. The RETURN CODE field of the RESPONSE SMS shall indicate whether or not the quiesce was processed successfully (see Error! Reference source not found.).

Modify section 11.2.14 as follows: Note: This is modified to include a new field indicating a region identifier. The region identifier is returned with return code and is used by the Master to identify the Region. Other configurators will be given the region identifier if their create request fails due to a region overlap. They then use this region identifier in the extend request to identify the region to be extended. This is necessary since an extension request may extend only one endpoint, both endpoints, or neither endpoint so that both endpoints of the new region must be supplied with the extension request. Similarly, when deleting a region, the use of a region identifier is simpler than providing the endpoints.

REQUEST SAT REGION SMS

The REQUEST SAT REGION SMS defined in Table 8 is used to by a Configurator to request SAT region control from the Master.

Table 8 - REQUEST SAT REGION SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (0Eh)							
1	SSA-TL VERSION (10h)							
2	TAG							
3	TAG							
4	RETURN PATH							
5	RETURN PATH							
6	RETURN PATH							
7	RETURN PATH							
8	ENDPOINT #1 UNIQUE ID							
...	ENDPOINT #1 UNIQUE ID							
15	ENDPOINT #1 UNIQUE ID							
16	ENDPOINT #2 UNIQUE ID							
...	ENDPOINT #2 UNIQUE ID							
23	ENDPOINT #2 UNIQUE ID							
24	reserved						SAT OPCODE	
25	<u>REGION IDENTIFIER</u>							
26	ENDPOINT #1 PORT NUMBER							
27	ENDPOINT #2 PORT NUMBER							
28	A QUOTA							
29	B QUOTA							

The SSA-TL VERSION field is defined in Table 4, and is the version of the transport layer currently being used between the node and the Master.

The TAG field is returned by the Master in the REQUEST SAT REGION REPLY SMS in response to the REQUEST SAT REGION SMS. The TAG value is assigned by the Configurator node and it shall be unique among the TAG values that are currently active from that Configurator node.

The RETURN PATH field specifies the PATH component that shall be placed in the ADDRESS field of the REQUEST SAT REGION REPLY SMS frame.

The ENDPOINT #1 UNIQUE ID field is the Unique ID of the node that is the first endpoint of the requested SAT region when used with an SAT OPCODE field value of CREATE SAT REGION. ~~The ENDPOINT #1 UNIQUE ID field is the Unique ID of the SAT Region Creator when used or~~ with an SAT OPCODE field value of EXTEND SAT REGION. ~~The field is ignored with a SAT OPCODE field value of~~ -DELETE SAT REGION.

The ENDPOINT #2 UNIQUE ID field is the Unique ID of the node that is the second endpoint of the requested SAT region for a SAT OPCODE value of either CREATE SAT REGION or EXTEND SAT REGION. The ENDPOINT #2 UNIQUE ID field is ignored for a SAT OPCODE value of DELETE SAT REGION.

The SAT OPCODE field specified in Table 9 identifies the operation requested by the REQUEST SAT REGION SMS.

The REGION IDENTIFIER field specified in Table 9 identifies the region being extended or deleted if the SAT OPCODE field is 01b or 10b. If the SAT OPCODE field is 00b, this field is ignored.

Table 9 - SAT OPCODE field

Value	Description
00b	CREATE SAT REGION
01b	EXTEND SAT REGION
10b	DELETE SAT REGION
11b	reserved

THE CREATE SAT REGION opcode requests a new SAT region to be created.

The EXTEND SAT REGION opcode requests that the union of the existing region and a new region be created, and the ownership of the new region be given to the node extending the region. Note that the union of the two regions could be the same region as the original region if the extension request region was completely contained in the old region. by moving the endpoint of the region specified by the ENDPOINT #1 UNIQUE ID field returned in a REQUEST SAT REGION REPLY SMS in response to an earlier CREATE SAT REGION SAT OPCODE value to the ENDPOINT #2 and port shown in the respective fields.

The DELETE SAT REGION opcode causes the region identified by the REGION IDENTIFIER field specified in Table 9 to be deleted. in which this node resides, the region shall have been created by a node that resides in the same region as this node. The nodes in the deleted region return to the general pool of nodes that existed prior to the creation of the SAT region. The ENDPOINT #1 PORT NUMBER field is a number that represents port numbers from 1 to 128 on the node indicated by the ENDPOINT #1 UNIQUE ID field for a SAT OPCODE value of CREATE SAT REGION or EXTEND SAT REGION. Since SAT regions may end on Switch ports, an 8-bit field is required. The ENDPOINT #1 PORT NUMBER field is ignored for a SAT OPCODE value of DELETE SAT REGION.

The ENDPOINT #2 PORT NUMBER field is a number that represents port numbers from 1 to 128 on the node indicated by the ENDPOINT #2 UNIQUE ID field for a SAT OPCODE value of CREATE SAT REGION or EXTEND SAT REGION. Since SAT regions may end on Switch ports, an 8-bit field is required. The ENDPOINT #2 PORT NUMBER field is ignored for a SAT OPCODE value of DELETE SAT REGION.

The A QUOTA field is the value the Master shall use in the A QUOTA field of the CONFIGURE PORT SMSs sent to the nodes that are being grouped into the SAT Region. All nodes in the region have their A QUOTA values set to this value upon the successful creation or extension of a SAT region.

The B QUOTA field is the value the Master shall use in the B QUOTA field of the CONFIGURE PORT SMSs sent to the nodes that are being grouped into the SAT Region. All nodes in the region have their B QUOTA values set to this value upon the successful creation or extension of a SAT region.

Several rules apply to SAT regions:

- a) SAT Regions do not “nest”. No SAT region can overlap or be enclosed by another.
- b) If the Master detects that a web change has caused a SAT region to contain no Configutors, it deletes the SAT region and return the nodes to the pre-region configuration.
- c) When Configutors membership in a region changes due to a request they did not originate, the Master notifies them of the change using a Master Alert (see Error! Reference source not found.) indicating addition to or deletion from a region.
- d) A SAT region shall contain more than one node.

Modify Section 11.2.15 as follows to include Region Identifiers:

REQUEST SAT REGION REPLY SMS

The REQUEST SAT REGION REPLY SMS defined in Table 10 is originated by a Master in response to a REQUEST SAT REGION SMS.

Table 10 - REQUEST SAT REGION REPLY SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (0Fh)							
1	SSA-TL VERSION (10h)							
2	TAG							
3	TAG							
4	REGION IDENTIFIER CREATOR UNIQUE ID							
125	RETURN CODE							

The SSA-TL VERSION field is defined in Table 4, and is the version of the transport layer currently being used between the Master and destination node.

The TAG field is copied from the REQUEST SAT REGION SMS. It identified the REQUEST SAT REGION SMS this reply is being generated for,

The ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field is returned by the Master and is used by the master to identify the region. is the Unique ID of the node that the Master acknowledges as the creator of the region. If the node receiving the reply was the successful creator, this serves as a verification. If the node was attempting a create and failed, it can be used in a subsequent extend command to add nodes to the region. If the node performed a successful extend operation, the ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field contains the region identifier of Unique ID of the node that created the original region. If the extend operation failed, the ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field is invalid and ignored. contains the Unique ID of the creator node of the SAT region preventing the extend operation, if any.

The ~~RETURNSTATUS~~ CODE field indicates the result of the REQUEST SAT REGION SMS as defined in Table 11.

Table 11 - RETURN CODE values

Value	Description
00h	SUCCESSFUL OPERATION
01h	SUCCESS WITH WARNING: CREATED OR EXTENDED REGION DOES NOT INCLUDE REQUESTER.
02h	SUCCESS WITH WARNING: DELETED REGION DID NOT INCLUDE REQUESTER.
03h	FAILURE. SAT REGIONS NOT SUPPORTED
04h	FAILURE. REQUESTED CREATE OR EXTEND OVERLAPS EXISTING REGION
05h	FAILURE NO PATH BETWEEN ENDPOINTS FOUND.
06h	SMS REJECTED, INVALID ENDPOINT, REGION IDENTIFIER CREATOR UNIQUE ID OR PORT NUMBER

Modify section 12.3 as follows:

SAT Region process

~~The optional process that a Configurator uses to manage a SAT region is as follows consists of three parts, Create SAT region, Extend SAT region and Delete SAT region. While it is optional for configurators to request, extend, or delete SAT regions, it is not optional for any configurator capable of becoming master in a multi-configurator web to reply to a REQUEST SAT REGION SMS. If the configurator does not support the creation and management of SAT regions, it shall return a REQUEST SAT REGION REPLY SMS with a return code of 03h, indicating failure, SAT regions not supported. The optional process that a Configurator uses to manage a SAT region is as follows and consists of three parts, Create SAT region, Extend SAT region and Delete SAT region.~~

Create SAT region

To create a SAT region, a Configurator shall send a REQUEST SAT REGION SMS to the Master indicating the endpoints and SAT quotas of the requested region, as well as a SAT OPCODE field value of CREATE A SAT REGION. The Master shall check its internal data structures for possible overlap of SAT regions, and violation of SAT region rules.

~~If an overlap exists, the Master shall return REGION IDENTIFIER of the overlapping region ~~THE UNIQUE ID OF THE REGION'S CREATOR~~ in the REGION IDENTIFIERCREATOR UNIQUE ID field along with a RETURN CODE field value of FAILURE. REQUESTED CREATE OR EXTEND OVERLAPS EXISTING REGION in the REQUEST SAT REGION REPLY SMS. The requesting node may use this information to extend the existing region. ~~If a Create Region request resulted in a failure due to overlapped regions, then the requester may issue a REQUEST SAT REGION SMS with an opcode of EXTEND REGION. The ENDPOINT #1 UNIQUE ID field is set to the CREATOR UNIQUE ID field value that was returned with the REQUEST SAT REGION REPLY SMS indicating failure on the create region request. The ENDPOINT #1 PORT NUMBER field is ignored. The ENDPOINT #2 UNIQUE ID and ENDPOINT #2 PORT NUMBER fields contain the new endpoint identifier for the extended region, which is extended from the endpoint with the lowest address (shortest path) from the new endpoint. The quotas remain as set by the original creator, as does the value for the creator's UNIQUE ID as acknowledged by the Master. The Master checks that the requested extend does not overlap a SAT region already created between the old endpoint and new endpoint of the region to be extended.~~~~

If there is no overlap, then the Master shall enter the new region's ENDPOINT #1 UNIQUE ID, ENDPOINT #2 UNIQUE ID, A QUOTA and B QUOTA values into its internal data structures. The Master shall return the REGION IDENTIFIER of Unique ID of the node requesting the region in the REGION IDENTIFIERCREATOR UNIQUE ID field of the REQUEST SAT REGION REPLY SMS with a RETURN CODE field value of SUCCESSFUL COMPLETION.

If the creation was successful, the Master shall configure the ports in the region using the quotas supplied in the REQUEST SAT REGION SMS. The Master shall send a Master Alert with an ALERT CODE value of SAT MEMBERSHIP ESTABLISHED. to each Configurator in the region. The endpoints are set to reflect SAT and SAT' as necessary to form the region.

The endpoints of the region requested are defined by the ENDPOINT #1 UNIQUE ID and ENDPOINT #2 UNIQUE ID fields in conjunction with the ENDPOINT #1 PORT NUMBER and ENDPOINT #2 PORT NUMBER fields in the REQUEST SAT REGION SMS. If the region defined by the endpoints does not contain the requester, and the SAT OPCODE field in the REQUEST SAT REGION SMS is CREATE A SAT REGION or EXTEND A SAT REGION, then the Master returns a REQUEST SAT REGION REPLY SMS with a RETURN CODE value of WARNING THE REQUESTER THAT THE CREATED OR EXTENDED REGION DOES NOT INCLUDE THE REQUESTER. The REGION IDENTIFIERCREATOR UNIQUE ID field is valid.

If the Master does not support the creation of SAT regions, then it returns a REQUEST SAT REGION REPLY SMS to the requester with a RETURN CODE of SAT REGIONS NOT SUPPORTED. The ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field shall be ignored.

Extend SAT region

If a Create Region request resulted in a failure due to overlapped regions, then the requester may issue a REQUEST SAT REGION SMS with an opcode of EXTEND REGION. The ENDPOINT #1 UNIQUE ID, ENDPOINT #1 PORT NUMBER, ENDPOINT #2 UNIQUE ID, and ENDPOINT #2 PORT NUMBER fields contain the new endpoint identifiers for the extended region. The quotas remain as set by the original creator, as does the value for the REGION IDENTIFIER as assigned by the Master. The Master checks that the requested extend does not overlap another SAT region. The endpoints of the new region are the result of the union of the old set of nodes in the region, and the new set of nodes. Therefore the new region could have neither, one, or both endpoints moved from the original region. If there is no overlap caused by extending the region, then the Master returns a REQUEST SAT REGION REPLY SMS with a RETURN CODE of SUCCESSFUL OPERATION. The Master then configures the ports in the region using the quotas defined in the original create request. The Master also sends a Master Alert with an ALERT CODE value of SAT REGION MEMBERSHIP ESTABLISHED to all the Configutors that have been added to the region.

If overlap exists, the Master returns the REGION ID of the overlapped region Unique ID of the blocking region's creator along with a RETURN CODE of FAILURE, REQUESTED CREATE OR EXTEND OVERLAPS EXISTING REGION. The ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field of the REQUEST SAT REGION REPLY SMS contains the REGION IDENTIFIER of the region the UNIQUE ID of the node that created the region that caused the conflict.

Delete SAT region

A Configutor may request the deletion of a region as follows in order:

- a) The Configutor creates a REQUEST SAT REGION SMS with an SAT OPCODE field of DELETE A SAT REGION. The node shall supply the region's ~~region identifier creator Unique ID~~ in the REGION IDENTIFIER field. The ENDPOINT #1 UNIQUE ID, field. ~~The ENDPOINT #2 UNIQUE ID, ENDPOINT #1 PORT NUMBER and ENDPOINT #2 PORT NUMBER fields are ignored.~~
- b) If the deletion is successful and if the region defined does not contain the requester, then the Master returns a REQUEST SAT REGION REPLY SMS with a status of SUCCESSFUL WITH WARNING: DELETED REGION DID NOT INCLUDE REQUESTER. The ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ field is valid.
- c) If the region defined does contain the requester, the Master returns a REQUEST SAT REGION REPLY SMS with a status of SUCCESSFUL OPERATION. The REGION IDENTIFIER field is valid.
- d) ~~if the delete request was successful,~~ the Master configures the ports to the system default state. The Master also sends a Master Alert with an ALERT CODE value of SAT REGION MEMBERSHIP DELETED to all the Configutors that resided in the region. The Master deletes the region information from its internal data structures. ~~Finally, the Master returns a REQUEST SAT REGION REPLY SMS with a status of SUCCESSFUL OPERATION. The CREATOR UNIQUE ID field is valid.~~
- e) If a new Master is elected, the SAT region information held by the old Master is lost. The new Master issues CONFIGURE PORT SMSs to all the ports in the system, using it's system default values which effectively clears out all existing SAT regions. When the creators of SAT regions receive the CONFIGURE PORT SMS indicating a new Master, they may repeat the process of creating SAT regions, thus rebuilding the Master's SAT region database if the new Master supports SAT regions.

- f) If the endpoints, ~~region identifier creator Unique ID~~ or port number fields in the REQUEST SAT REGION SMS are invalid, then the Master returns a REQUEST SAT REGION REPLY SMS with a status code of SMS REJECTED, INVALID ENDPOINT, ~~REGION IDENTIFIER CREATOR UNIQUE ID~~ OR PORT NUMBER. Byte 4 contains the first byte number of the offending field in the REQUEST ~~SAT REGION SMS~~.

Modify section 12.1, to read as follows:

Master Negotiation process

In a web with a stable topology after power-up and After the configuration process, the Configurator node with the highest MASTER PRIORITY is elected to be the Master (see 11.2.6). If more than one Configurator node is set at the same highest MASTER PRIORITY, then the Configurator node with the highest Unique ID among them is elected to be the Master.

Each Configurator node compares its own MASTER PRIORITY with the MASTER PRIORITY it receives from each other Configurator node in the QUERY NODE REPLY SMS, or the QUERY NODE SMS from the other Configurator, whichever arrives first. Each Configurator node in the group with the highest MASTER PRIORITY then compares its own Unique ID with those of the other Configurator nodes in the group. The Configurator node within the group with the highest Unique ID becomes the Master.

If two sub-Webs that both contain a Configurator node are joined by a new link, then the Master node of each sub-web each Configurator node invokes the web walking process for the new link. When each Master receives the QUERY NODE SMS from the other, they compare their Master Priorities and Unique IDs. The Master node with the lower MASTER PRIORITY or Unique ID becomes a Resigning Master and notifies the other Master of this change in status by setting the CM bit in the QUERY NODE REPLY SMS. The Resigning Master no longer responds to new ASYNC ALERT SMSs. The resigning master then completes the processing of all pending ASYNC ALERT SMSs, and then performs a Formal Resignation to the Pending Master by sending it a QUERY NODE SMS with the MA bit set. During the resignation procedure, the Master Configurator node with the higher MASTER PRIORITY or Unique ID becomes the Pending Master of the sub web containing the Resigning Master, sends a QUERY NODE REPLY with the CM bit set to the Resigning Master and continues acting as the Master of it's own sub-web. When the Pending Master receives the Formal Resignation from the Resigning Master it becomes the Assuming Master of the Resigning Master's Sub-Web and invokes the Configuration Process for the new link. issues a CONFIGURE PORT SMS to each new port to indicate that it is assuming the role of Master. It then receives any pending or future ASYNC ALERT SMSs. A Pending Master shall wait for 5 seconds to receive a Formal Resignation. If one has not been received in 5 Seconds, then the Pending Master shall send a Query Node with the CM bit set to the Resigning Master. If a Query Node Reply is not received within 5 seconds, then the Assuming Master shall issue a Total Reset to the node, and invoke the configuration Process for the Sub Web. If a QUERY NODE REPLY is received with the CM bit clear, the Pending Master treats this as a Formal Resignation. If the CM bit is set, the Pending Master re-starts the Timer. Should the second attempt also fail, then the resigning Master node is reset and configuration continues.

If a Resigning Master receives one or more additional QUERY NODE SMSs from links added after it became a Resigning Master, then it shall return QUERY NODE REPLY SMSs with the CM bit set, and perform a Formal Resignation to each that would have won Mastership. This ensures that whatever node finally becomes Web Master will wait until the Resigning Master has completed it's processing of pending ASYNC ALERT SMSs prior to configuration.

If a Pending Master receives one or more QUERY NODE SMSs that would cause it to lose the Master Election, it becomes a Resigning Master. It stops acknowledging ASYNC ALERT SMSs, completes processing of all pending ASYNC ALERT SMSs, and formally resigns to all winning Masters. Any Formal Resignations Received are answered with QUERY NODE REPLY SMSs with the CM bit clear.

If a sub-web containing no configutors is attached to a sub-web containing a Master, the Master shall invoke the Web Reset process on the newly attached Sub-web prior to notifying other configutors of the attachment.

The Pending Master node ~~that~~ determines that it now has the highest MASTER PRIORITY and Unique ID of the remaining Configutor nodes as described above, ~~then it~~ becomes the ~~new~~Pending Master. The ~~Pending~~new Master then issues a CONFIGURE PORT SMS to each remaining registered Configutor node to indicate that it is the Master.

{ the remaining text is unchanged}...

The Following Changes Apply to S3P, Rev 2.

Modify Section 7 as follows:

All SMSs contain a 2-byte TAG field that is used to relate replies to the original request. The TAG field value is assigned by the initiator and it shall be unique among all the TAG field values that are currently active from that initiator on any target or Logical Unit throughout the Web. Effectively the TAG field identifies the target, LUN and Queue-tag components of the SCSI-2 nexus. (The initiator and logical path are identified by the RETURN PATH ID field in the SMS.) If the command associated with the TAG field uses the Transport Layer's Data Transport SMSs, then the TAG must be obtained from the Transport Layer's pool of tags, and must be unique among all outstanding tags from the Transport Layer.

Modify Section 7.3 as follows:

SMS Buffer Full condition

The S3P protocol layer of the destination node shall enter an SMS Buffer Full condition for a given initiator when an SCSI COMMAND SMS is received from that initiator and the command cannot be accepted due to a TASK SET FULL or BUSY. Upon entering an SMS Buffer Full condition, the destination node shall ~~acknowledge~~~~complete~~ the SCSI COMMAND SMS by generating a SCSI STATUS SMS indicating a TASK SET FULL status. When in an SMS Buffer Full condition, the destination node shall discard SCSI COMMAND SMSs with a RESUME bit cleared from the same Initiator even if the condition which caused the SMS Buffer Full condition no longer exists.

Modify section 7.9 as follows:

ENABLE AER SMS

The ENABLE AER SMS is sent from an initiator to a target to enable or ~~disable~~ Asynchronous Event Reporting by supplying a TAG and RETURN PATH ID field to be used by the target for any future AER REPORT SMS between the target and that particular initiator. If multiple initiators send ENABLE AER SMSs to a target, the target shall report Asynchronous Events to all initiators the enabled reporting. If an event applies to multiple targets, it shall be reported to the

multiple targets. The target shall respond with a SCSI RESPONSE SMS using the TAG value specified in the ENABLE AER SMS. If a target S3P layer receives notification of a quiesce in a given Unique Id/Initiator, then it treats this notification as an implied disable of Asynchronous Event Reporting.

Table 12 defines the contents of the DATA field in the ENABLE AER SMS.

Table 12 - ENABLE AER SMS

Byte	Bit 7	6	5	4	3	2	1	Bit 0
0	SMS CODE (83h)							
1	S3P CODE (20h)							
2	TAG							
3	TAG							
4	RETURN PATH ID							
5	RETURN PATH ID							
6	RETURN PATH ID							
7	RETURN PATH ID							
8	LUN							
9	reserved							AER-ON
10	AER TAG							
11	AER TAG							

The TAG field is used to associate the AER SMS with a particular target.

The RETURN PATH ID field identifies the logical path that shall be used by the target to return the associated AER SMS.

The AER-ON field is set to enable Asynchronous Event Reporting, or to “0” to disable Asynchronous Event Reporting.

Question on Asynchronous events: How are things like changes in mode pages, spin up/down, logical unit reset, target reset and clear task set reported to other initiators sharing the device with AER enabled?

Modify section 7.14 and 7.16 as follows:

The SSA-TL2 constructs of Initiator Tables, pending asynchronous alerts, and unprocessed SMSs shall be unaffected. There shall be no effect on other upper level protocols running in the node.

Modify section 8.3 as follows:

Tag interaction between SSA-TL2 and SSA-S3P

If a command is not using the transport layer’s data movement application SMSs, then there~~Since there~~ is no overlap between the responses to SSA-S3P SMS, and SSA-TL2 SMSs (including ASYNC ALERT SMSs), and hence there is no requirement that the TAG field values be unique across outstanding SSA-S3P and SSA-TL2 SMSs. However, if the transport layer’s data movement services SMSs are used by a command in an Upper Layer Protocol, tags used by the Upper Layer Protocols -SSA-S3P shall be unique across all ULPs that use the SSA-TL2 layer’s data transport services.

Modify section 8.4 as follows:

Effects of reset conditions

SSA-TL2 defines the following types of reset: Link Reset, Total Reset, Absolute Reset, and Power On Reset.

Link Reset has no effect on the SCSI constructs.

In response to an ~~an Total Reset~~, Absolute Reset, or Power On Reset, the S3P layer shall clear SCSI constructs (equivalent to an SCSI-3 Hard reset). In addition to the normal clearing of all tasks, reservations and operating conditions shall be reset to their last saved states.

In response to a notification of a Total Reset, the S3P layer shall clear all tasks, queues, reservations and statuses to their initialized state. Mode parameters and operating parameters shall remain in their previous states.

Neither SSA-TL2 nor SSA-S3P provide a global reset equivalent to the RST signal in the parallel SCSI bus.

Modify Section 9.2 as follows: (clarification and more detail on usage, causes TL2 Master Alert Change also to extend the frame data returned to 8 bytes)

Confirmed status

The Confirmed Status feature of SSA-S3P is enabled with the CONFIRM bit in the SCSI COMMAND SMS, and is therefore settable on a command by command basis. When the confirmed Status feature is enabled, it allows three additional features of SSA-S3P:

- a) Confirm status;
- b) Restart Command;
- c) Redirect Command.

Confirm status

The Confirm Status feature provides an interlock between the target and the initiator. The receipt of the CONFIRM STATUS SMS by the target with the COMPLETE bit set confirms that the initiator has received the SCSI STATUS SMS sent by the target. The target is now free to discard the status information and perform the next task. If no CONFIRM STATUS SMS is received within 5 seconds, the target shall generate an ASYNC ALERT SMS with an ALERT CODE of UNCONFIRMED STATUS. The Initiators in the web shall not discard command information on Confirmed Status Commands until 10 seconds after the most recent CONFIRM STATUS SMS for the command has been sent. This allows for an opportunity for the target to notify the initiators of a lost CONFIRM STATUS SMS. If a target receives a MASTER ALERT indicating UNCONFIRMED STATUS, it will use the FRAME DATA field to discover the TAG and RETURN PATH ID of the unconfirmed command. If the initiator indicated has already sent a CONFIRM STATUS SMS, then the initiator re-sends the CONFIRM STATUS SMS. If the Initiator has not completed the command, then the RESTART COMMAND PROCESS is invoked.

Restart command

The Restart Command feature provides a method of the initiator instructing the active task to be restarted from a particular point. The receipt of the CONFIRM STATUS SMS by the target with the COMPLETE bit cleared instructs the target to restart the active task from the point specified in the BYTE OFFSET field. Typically, this is used when a physical path has been lost and the logical path is being associated to a new physical path. In the case of a data transfer to the target from the initiator, the target shall send a DATA REQUEST SMS to the initiator using the offset and NEW RETURN PATH ID provided in the CONFIRM STATUS SMS. The initiator will send the data and confirm the status when received from the target. In the case of a data transfer to the initiator from the target, the target shall send a DATA READY SMS to the initiator using the offset and NEW RETURN PATH ID from the CONFIRM STATUS SMS. The

Initiator shall send a DATA REPLY SMS to begin the data transfer. Following the receipt of the status, the initiator shall send the CONFIRM STATUS SMS with the COMPLETE bit set to indicate the receipt of the status. Since the old path and channel number used by the command may not be flushed of old data and SMSs, it must be held reserved until the command completes on the new path. Data and SMSs arriving on that port and channel shall be discarded during this period.

Redirect command

The Redirect Command feature provides a method of the initiator instructing the active or queued task to use another path. Typically, this is used when a physical path has not been lost but, for load balancing and efficiency reasons, the logical path is being associated to a new physical path -The receipt of the CONFIRM STATUS SMS by the target with the COMPLETE bit cleared instructs the target to redirect the active or queued task using the path specified in the NEW RETURN PATH ID field. The protocol follows that of the restart command.

Sincerely,
Mark A. DeWilde
Principal System Architect
Pathlight Technologies
Voice: (607)266-4000 X-403
FAX: (607)266-0352
Email: mark@pathlight.com