

Network Working Group
Request for Comments: 2297
Updates: 1987
Category: Informational

P. Newman, Nokia
W. Edwards, Sprint
R. Hinden, Nokia
E. Hoffman, Nokia
F. Ching Liaw
T. Lyon, Nokia
G. Minshall, Fiberlane
March 1998

Ipsilon's General Switch Management Protocol Specification
Version 2.0

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1998). All Rights Reserved.

Abstract

This memo specifies enhancements to the General Switch Management Protocol (GSMP) [RFC1987]. The major enhancement is the addition of Quality of Service (QoS) messages. Other improvements have been made to the protocol resulting from operational experience. GSMP is a general purpose protocol to control an ATM switch. It allows a controller to establish and release connections across the switch; add and delete leaves on a multicast connection; manage switch ports; request configuration information; and request statistics.

Table of Contents

1. Introduction.....	3
2. GSMP Packet Encapsulation.....	4
2.1 ATM Encapsulation.....	4
2.2 Ethernet Encapsulation.....	6
3. Common Definitions and Procedures.....	7
3.1 GSMP Packet Format.....	8
3.2 Failure Response Messages.....	11
4. Connection Management Messages.....	16
4.1 Add Branch Message.....	21
4.2 Delete Tree Message.....	23
4.3 Verify Tree Message.....	24
4.4 Delete All Message.....	24
4.5 Delete Branches Message.....	25
4.6 Move Branch Message.....	27
5. Port Management Messages.....	29
5.1 Port Management Message.....	29
5.2 Label Range Message.....	34
6. State and Statistics Messages.....	37
6.1 Connection Activity Message.....	38
6.2 Statistics Messages.....	40
6.2.1 Port Statistics Message.....	44
6.2.2 Connection Statistics Message.....	44
6.2.3 QoS Class Statistics Message.....	44
6.3 Report Connection State Message.....	45
7. Configuration Messages.....	49
7.1 Switch Configuration Message.....	50
7.2 Port Configuration Message.....	51
7.3 All Ports Configuration Message.....	57
8. Event Messages.....	59
8.1 Port Up Message.....	60
8.2 Port Down Message.....	60
8.3 Invalid VPI/VCI Message.....	61
8.4 New Port Message.....	61
8.5 Dead Port Message.....	61
9. Quality of Service Messages.....	61
9.1 Abstract Switch Model.....	62
9.2 QoS Configuration Message.....	66
9.3 Scheduler Establishment Message.....	74

9.4 QoS Class Establishment Message.....	78
9.5 QoS Release Message.....	85
9.6 QoS Connection Management Message.....	86
9.7 QoS Failure Response Codes.....	97
10. Adjacency Protocol.....	97
10.1 Packet Format.....	98
10.2 Procedure.....	101
10.3 Loss of Synchronization.....	103
11. Summary of Failure Response Codes.....	104
12. Summary of Message Set.....	105
References.....	107
Security Considerations.....	107
Authors' Addresses.....	107
Full Copyright Statement.....	109

1. Introduction

The General Switch Management Protocol (GSMP), is a general purpose protocol to control an ATM switch. GSMP allows a controller to establish and release connections across the switch; add and delete leaves on a multicast connection; manage switch ports; request configuration information; and request statistics. It also allows the switch to inform the controller of asynchronous events such as a link going down. GSMP runs across an ATM link connecting the controller to the switch, on a control connection (virtual channel) established at initialization. GSMP operation across an Ethernet link is also specified. The GSMP protocol is asymmetric, the controller being the master and the switch being the slave. Multiple switches may be controlled by a single controller using multiple instantiations of the protocol over separate control connections.

A switch is assumed to contain multiple "ports". Each port is a combination of one "input port" and one "output port". Some GSMP requests refer to the port as a whole whereas other requests are specific to the input port or the output port. ATM cells arrive at the switch from an external communication link on incoming virtual paths or virtual channels at an input port. ATM cells depart from the switch to an external communication link on outgoing virtual paths or virtual channels from an output port. Virtual paths on a port or link are referenced by their virtual path identifier (VPI). Virtual channels on a port or link are referenced by their virtual path and virtual channel identifiers (VPI/VCI).

A virtual channel connection across a switch is formed by connecting an incoming virtual channel to one or more outgoing virtual channels. Virtual channel connections are referenced by the input port on which they arrive and the virtual path and virtual channel identifiers (VPI/VCI) of their incoming virtual channel. A virtual path connection across a switch is formed by connecting an incoming virtual path to one or more outgoing virtual paths. Virtual path connections are referenced by the input port on which they arrive and their virtual path identifier (VPI). In a virtual path connection the value of the VCI in each cell on that, connection is not used by the switch and remains unchanged by the switch.

GSMP supports point-to-point and point-to-multipoint connections. A multipoint-to-point connection is specified by establishing multiple point-to-point connections each of them specifying the same output branch. A multipoint-to-multipoint connection is specified by establishing multiple point-to-multipoint trees each of them specifying the same output branches.

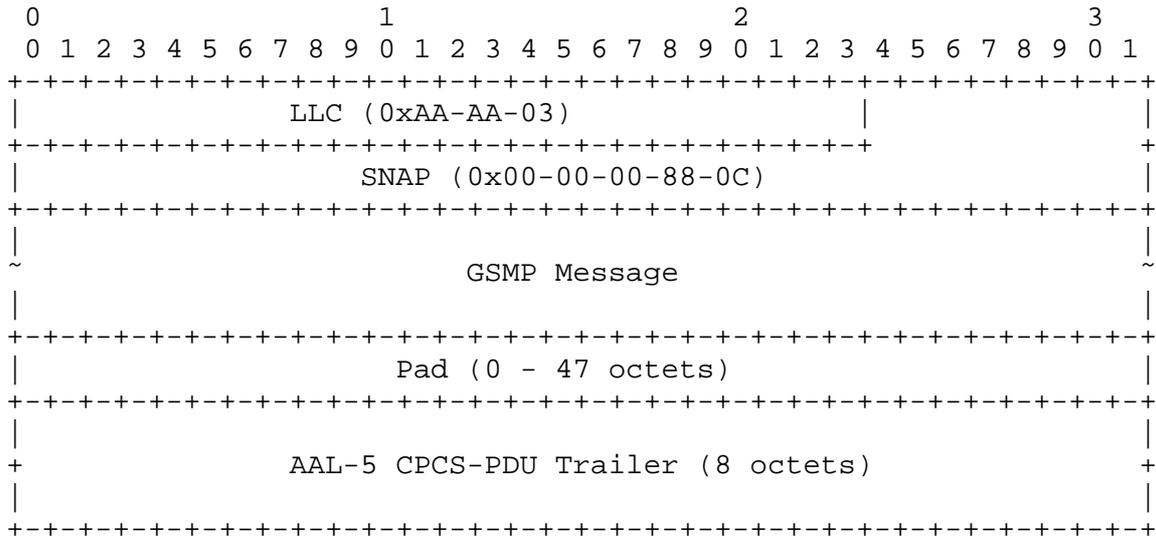
In general a virtual channel is established with a certain quality of service (QoS). A rich set of QoS messages is introduced in this version of the protocol. However, implementation or operation of GSMP without any of the messages defined in Section 9, "Quality of service messages," is permitted. In this case each virtual channel connection or virtual path connection may be assigned a priority when it is established. It may be assumed that for virtual connections that share the same output port, an ATM cell on a connection with a higher priority is much more likely to exit the switch before an ATM cell on a connection with a lower priority if they are both in the switch at the same time. The number of priorities that each port of the switch supports may be obtained from the port configuration message.

GSMP contains an adjacency protocol. The adjacency protocol is used to synchronize state across the link, to negotiate which version of the GSMP protocol to use, to discover the identity of the entity at the other end of a link, and to detect when it changes.

2. GSMP Packet Encapsulation

2.1 ATM Encapsulation

GSMP packets are variable length and for an ATM data link layer they are encapsulated directly in an AAL-5 CPCS-PDU [I.363] with an LLC/SNAP header as illustrated:



(The convention in the documentation of Internet Protocols [RFC1700] is to express numbers in decimal. Numbers in hexadecimal format are specified by prefacing them with the characters "0x". Data is pictured in "big-endian" order. That is, fields are described left to right, with the most significant octet on the left and the least significant octet on the right. Whenever a diagram shows a group of octets, the order of transmission of those octets is the normal order in which they are read in English. Whenever an octet represents a numeric quantity the left most bit in the diagram is the high order or most significant bit. That is, the bit labeled 0 is the most significant bit. Similarly, whenever a multi-octet field represents a numeric quantity the left most bit of the whole field is the most significant bit. When a multi-octet quantity is transmitted, the most significant octet is transmitted first. This is the same coding convention as is used in the ATM layer [I.361] and AAL-5 [I.363].)

The LLC/SNAP header contains the octets: 0xAA 0xAA 0x03 0x00 0x00 0x00 0x88 0x0C. (0x880C is the assigned Ethertype for GSMP.)

The maximum transmission unit (MTU) of the GSMP Message field is 1492 octets.

The virtual channel over which a GSMP session is established between a controller and the switch it is controlling is called the GSMP control channel. The default VPI and VCI of the GSMP control channel for LLC/SNAP encapsulated GSMP messages on an ATM data link layer is:

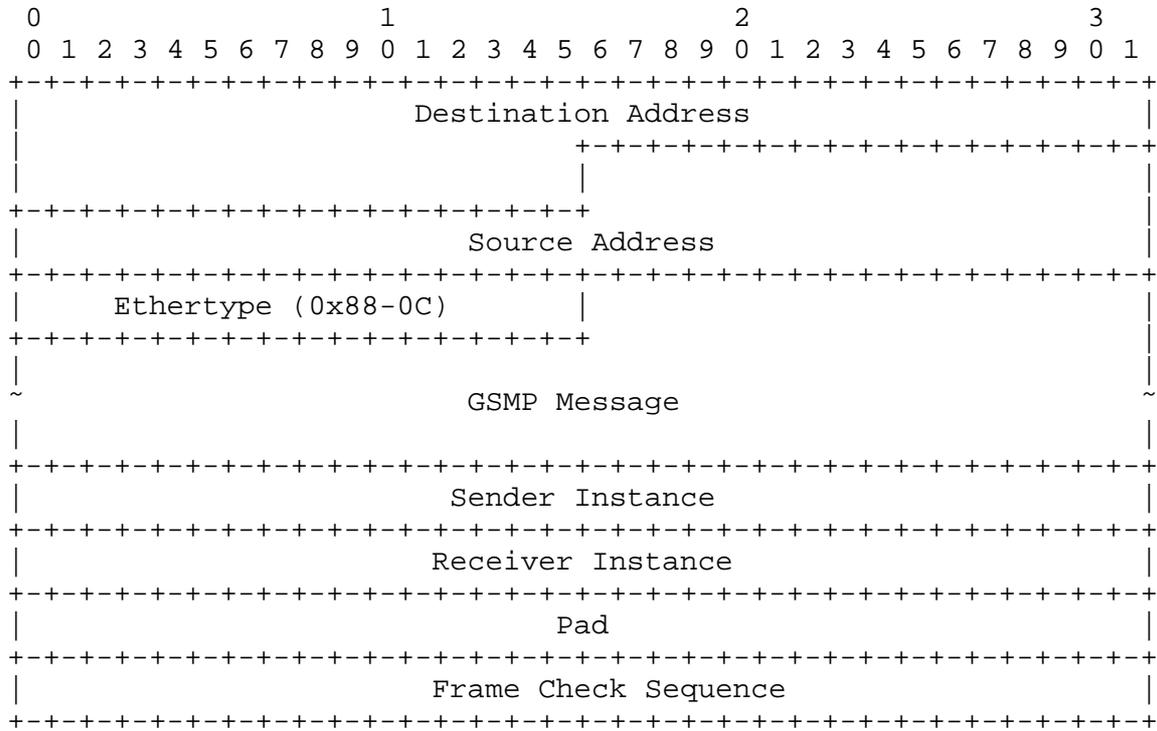
```

VPI = 0
VCI = 15.

```

2.2 Ethernet Encapsulation

GSMP packets may be encapsulated on an Ethernet data link as illustrated:



Destination Address

For the SYN message of the adjacency protocol the Destination Address is the broadcast address 0xFFFFFFFF. (Alternatively, it is also valid to configure the node with the unicast 48-bit IEEE MAC address of the destination. In this case the configured unicast Destination Address is used in the SYN message.) For all other messages the Destination Address is the unicast 48-bit IEEE MAC address of the destination. This address may be discovered from the Source Address field of messages received during synchronization of the adjacency protocol.

Source Address

For all messages the Source Address is the 48-bit IEEE MAC address of the sender.

Ethertype

The assigned Ethertype for GSMP is 0x880C.

GSMP Message

The maximum transmission unit (MTU) of the GSMP Message field is 1492 octets.

Sender Instance

The Sender Instance number for the link obtained from the adjacency protocol. This field is already present in the adjacency protocol message. It is appended to all non-adjacency GSMP messages in the Ethernet encapsulation to offer additional protection against the introduction of corrupt state.

Receiver Instance

The Receiver Instance number is what the sender believes is the current instance number for the link, allocated by the entity at the far end of the link. This field is already present in the adjacency protocol message. It is appended to all non-adjacency GSMP messages in the Ethernet encapsulation to offer additional protection against the introduction of corrupt state.

Pad

The minimum length of the data field of an Ethernet packet is 46 octets. If necessary, padding should be added such that it meets the minimum Ethernet frame size. This padding should be octets of zero and it is not considered to be part of the GSMP message.

After the adjacency protocol has achieved synchronization, for every GSMP message received with an Ethernet encapsulation, the receiver must check the Source Address from the Ethernet MAC header, the Sender Instance, and the Receiver Instance. The incoming GSMP message must be discarded if the Sender Instance and the Source Address do not match the values of Sender Instance and Sender Name stored by the "Update Peer Verifier" operation of the GSMP adjacency protocol. The incoming GSMP message must also be discarded if it arrives over any port other than the port over which the adjacency protocol has achieved synchronization. In addition, the incoming message must also be discarded if the Receiver Instance field does not match the current value for the Sender Instance of the GSMP adjacency protocol.

3. Common Definitions and Procedures

GSMP is a master-slave protocol. The controller issues request messages to the switch. Each request message indicates whether a response is required from the switch and contains a transaction

identifier to enable the response to be associated with the request. The switch replies with a response message indicating either a successful result or a failure. There are five classes of GSMP request-response message: Connection Management, Port Management, State and Statistics, Configuration, and Quality of Service. The switch may also generate asynchronous Event messages to inform the controller of asynchronous events. Event messages are not acknowledged by the controller. There is also an adjacency protocol message used to establish synchronization across the link and maintain a handshake.

For the request-response messages, each message type has a format for the request message and a format for the success response. Unless otherwise specified a failure response message is identical to the request message that caused the failure, with the Code field indicating the nature of the failure. Event messages have only a single format defined as they are not acknowledged by the controller.

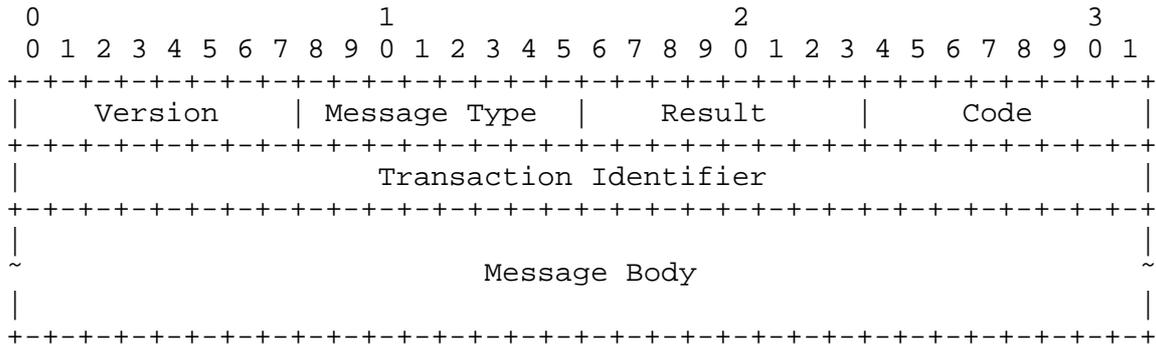
Switch ports are described by a 32-bit port number. The switch assigns port numbers and it may typically choose to structure the 32 bits into subfields that have meaning to the physical structure of the switch (e.g. slot, port). In general, a port in the same physical location on the switch will always have the same port number, even across power cycles. The internal structure of the port number is opaque to the GSMP protocol. However, for the purposes of network management such as logging, port naming, and graphical representation, a switch may declare the physical location (physical slot and port) of each port. Alternatively, this information may be obtained by looking up the product identity in a database.

Each switch port also maintains a port session number assigned by the switch. A message, with an incorrect port session number must be rejected. This allows the controller to detect a link failure and to keep state synchronized.

Except for the adjacency protocol message, no GSMP messages may be sent across the link until the adjacency protocol has achieved synchronization, and all GSMP messages received on a link that does not currently have state synchronization must be discarded.

3.1 GSMP Packet Format

All GSMP messages, except the adjacency protocol message, have the following format:



Version

The version number of the GSMP protocol being used in this session. It should be set by the sender of the message to the GSMP protocol version negotiated by the adjacency protocol.

Message Type

The GSMP message type. GSMP messages fall into six classes: Connection Management, Port Management, State and Statistics, Configuration, Quality of Service, and Events. Each class has a number of different message types. In addition, one Message Type is allocated to the adjacency protocol.

Result

Field in a Connection Management request message, a Port Management request message, or a Quality of Service request message is used to indicate whether a response is required to the request message if the outcome is successful. A value of "NoSuccessAck" indicates that the request message does not expect a response if the outcome is successful, and a value of "AckAll" indicates that a response is expected if the outcome is successful. In both cases a failure response must be generated if the request fails. For State and Statistics, and Configuration request messages, a value of "NoSuccessAck" in the request message is ignored and the request message is handled as if the field were set to "AckAll". (This facility was added to reduce the control traffic in the case where the controller periodically checks that the state in the switch is correct. If the controller does not use this capability, all request messages should be sent with a value of "AckAll.")

In a response message the result field can have three values: "Success," "More," and "Failure". The "Success" and "More" results both indicate a success response. The "More" result indicates that the success response exceeds the maximum transmission unit of the data link and that one or more further messages will be sent to complete the success response. All messages that belong to the same success response will have the same Transaction Identifier. The "Success" result indicates a success response that may be contained in a single message or the final message of a success response spanning multiple messages.

The encoding of the result field is:

NoSuccessAck:	Result = 1
AckAll:	Result = 2
Success:	Result = 3
Failure:	Result = 4
More:	Result = 5.

The Result field is not used in an adjacency protocol message.

Code

Field gives further information concerning the result in a response message. It is mostly used to pass an error code in a failure response but can also be used to give further information in a success response message or an event message. In a request message the code field is not used and is set to zero. In an adjacency protocol message the Code field is used to determine the function of the message.

Transaction Identifier

Used to associate a request message with its response message. For request messages the controller may select any transaction identifier. For response messages the transaction identifier is set to the value of the transaction identifier from the message to which it is a response. For event messages the transaction identifier should be set to zero. The Transaction Identifier is not used, and the field is not present, in the adjacency protocol.

The following fields are frequently found in GSMP messages. They are defined here to avoid repetition.

Port

Gives the port number of the switch port to which the message applies.

Port Session Number

Each switch port maintains a Port Session Number assigned by the switch. The port session number of a port remains unchanged while the port is continuously in the Available state and the link status is continuously Up. When a port returns to the Available state after it has been Unavailable or in any of the Loopback states, or when the line status returns to the Up state after it has been Down or in Test, or after a power cycle, a new Port Session Number must be generated. Port session numbers should be assigned using some form of random number.

If the Port Session Number in a request message does not match the current Port Session Number for the specified port, a failure response message must be returned with the Code field indicating, "Invalid port session number." The current port session number for a port may be obtained using a Port Configuration or an All Ports Configuration message.

Any field in a GSMP message that is unused or defined as "reserved" must be set to zero by the sender and ignored by the receiver.

It is not an error for a GSMP message to contain additional data after the end of the Message Body. This is to support development and experimental purposes. However, the maximum transmission unit of the GSMP message, as defined by the data link layer encapsulation, must not be exceeded.

A success response message must not be sent until the requested operation has been successfully completed.

3.2 Failure Response Messages

A failure response message is formed by returning the request message that caused the failure with the Result field in the header indicating failure (Result = 4) and the Code field giving the failure code. The failure code specifies the reason for the switch being unable to satisfy the request message.

If the switch issues a failure response in reply to a request message, no change should be made to the state of the switch as a result of the message causing the failure. (For request messages that contain multiple requests, such as the Delete Branches message, the

failure response message will specify which requests were successful and which failed. The successful requests may result in changed state.)

If the switch issues a failure response it must choose the most specific failure code according to the following precedence:

Invalid Message

Failure specific to the particular message type (failure code 16). (The meaning of this failure is dependent upon the particular message type and is specified in the text defining the message.)

A failure response specified in the text defining the message type.

Connection Failures

Virtual Path Connection Failures

Multicast Failures

QoS Failures (QoS failures are specified in Section 9.7.)

General Failures

If multiple failures match in any of the following categories, the one that is listed first should be returned. The following failure response messages and failure codes are defined:

Invalid Message

- 3: The specified request is not implemented on this switch. The Message Type field specifies a message that is not implemented on the switch or contains a value that is not defined in the version of the protocol running in this session of GSMP.
- 5: One or more of the specified ports does not exist. At least one of the ports specified in the message is invalid. A port is invalid if it does not exist or if it has been removed from the switch.
- 4: Invalid Port Session Number. The value given in the Port Session Number field does not match the current Port Session Number for the specified port.

Connection Failures

- 8: The specified connection does not exist.
An operation that expects a connection to be specified, either a virtual channel or a virtual path connection, cannot locate the specified connection. A virtual channel connection is specified by the input port, input VPI, and input VCI on which it arrives. A virtual path connection is specified by the input port and input VPI on which it arrives.
- 9: The specified branch does not exist.
An operation that expects a branch of an existing connection to be specified, either a virtual channel or a virtual path connection, cannot locate the specified branch. A branch of a virtual channel connection is specified by the virtual channel connection it belongs to and the output port, output VPI, and output VCI on which it departs. A branch of a virtual path connection is specified by the virtual path connection it belongs to and the output port and output VPI on which it departs.
- 18: One or more of the specified input VPIs is invalid.
- 19: One or more of the specified input VCIs is invalid.
- 20: One or more of the specified output VPIs is invalid.
- 21: One or more of the specified output VCIs is invalid.
- 22: Invalid Class of Service field in a Connection Management message.
The value of the Class of Service field is invalid.
- 23: Insufficient resources for QoS Profile.
The resources requested by the QoS Profile in the Class of service field are not available.

Virtual Path Connections

- 24: Virtual path switching is not supported on this input port.
- 25: Point-to-multipoint virtual path connections are not supported on either the requested input port or the requested output port.
One or both of the requested input and output ports is unable to support point-to-multipoint virtual path connections.

- 26: Attempt to add a virtual path connection branch to an existing virtual channel connection.
It is invalid to mix branches switched as virtual channel connections with branches switched as virtual path connections on the same point-to-multipoint connection.
- 27: Attempt to add a virtual channel connection branch to an existing virtual path connection.
It is invalid to mix branches switched as virtual channel connections with branches switched as virtual path connections on the same point-to-multipoint connection.

Multicast Failures

- 10: A branch belonging to the specified point-to-multipoint connection is already established on the specified output port and the switch cannot support more than a single branch of any point-to-multipoint connection on the same output port.
- 11: The limit on the maximum number of point-to-multipoint connections that the switch can support has been reached.
- 12: The limit on the maximum number of branches that the specified point-to-multipoint connection can support has been reached.
- 17: Cannot label each output branch of a point-to-multipoint tree with a different label.
Some early designs, and some low-cost ATM switch designs, require all output branches of a multicast connection to use the same value of VPI/VCI.
- 28: Only point-to-point bidirectional connections may be established.
It is an error to attempt to add an additional output branch to an existing connection with the bidirectional flag set.
- 13: Unable to assign the requested VPI/VCI value to the requested branch on the specified point-to-multipoint connection. Although the requested VPI and VCI are valid, the switch is unable to support the request using the specified values of VPI and VCI for some reason not covered by the above failure responses. This message implies that a valid value of VPI or VCI exists that the switch could

support. For example, some switch designs restrict the number of distinct VPI/VCI values available to a point-to-multipoint connection. (Most switch designs will not require this message.)

- 14: General problem related to the manner in which point-to-multipoint is supported by the switch.
Use this message if none of the more specific multicast failure messages apply. (Most switch designs will not require this message.)

General Failures

- 2: Invalid request message.
There is an error in one of the fields of the message not covered by a more specific failure message.
- 6: One or more of the specified ports is down.
A port is down if its Port Status is Unavailable. Connection Management, Connection State, Port Management, and Configuration operations are permitted on a port that is Unavailable. Connection Activity and Statistics operations are not permitted on a port that is Unavailable and will generate this failure response. A Port Management message specifying a Take Down function on a port already in the Unavailable state will also generate this failure response.
- 15: Out of resources.
The switch has exhausted a resource not covered by a more specific failure message, for example, running out of memory.
- 1: Unspecified reason not covered by other failure codes.
The failure message of last resort.

The following failure response messages are only used by the Label Range message.

- 29: Cannot support requested VPI range.
- 30: Cannot support requested VCI range on all requested VPIs.

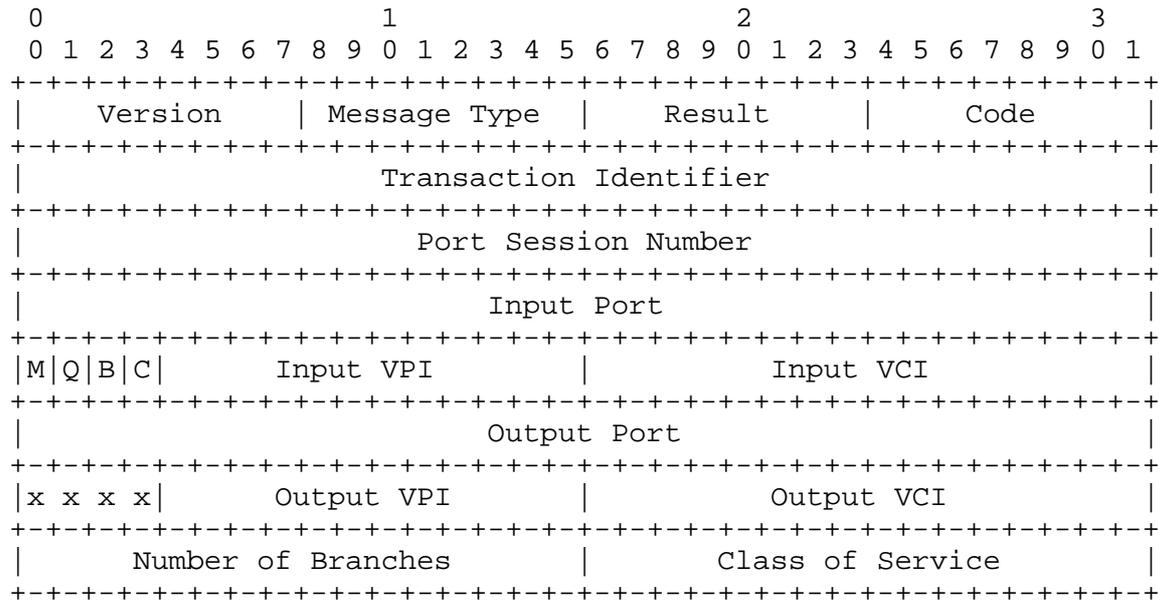
The following failure response messages are only used by the Set Transmit Cell Rate function of the Port Management message.

- 31: The transmit cell rate of this output port cannot be changed.

32: Requested transmit cell rate out of range for this output port.

4. Connection Management Messages

Connection management messages are used by the controller to establish, delete, modify and verify virtual channel connections and virtual path connections across the switch. The Add Branch, Delete Tree, and Delete All connection management messages have the following format for both request and response messages:



Input Port Identifies a switch input port.

Flags

M: Multicast

The Multicast flag is used as a hint for point-to-multipoint connections in the Add Branch message. It is not used in any other connection management messages and in these messages it should be set to zero. If set, it indicates that the virtual channel connection or the virtual path connection is very likely to be a point-to-multipoint connection. If zero, it indicates that this connection is very likely to be a point-to-point connection or is unknown.

The Multicast flag is only used in the Add Branch message when establishing the first branch of a new connection. It is not required to be set when establishing subsequent branches of a point-to-multipoint connection and on such connections it should be ignored by the receiver. (On receipt of the second and subsequent Add Branch messages the receiver knows that this is a point-to-multipoint connection.) If it is known that this is the first branch of a point-to-multipoint connection this flag should be set. If it is unknown, or if it is known that the connection is point-to-point this flag should be zero. The use of this flag is not mandatory. It may be ignored by the switch. If unused the flag should be set to zero. Some switches use a different data structure for point-to-multipoint connections than for point-to-point connections. This flag avoids the switch setting up a point-to-point structure for the first branch of a point-to-multipoint connection which must immediately be deleted and reconfigured as point-to-multipoint when the second branch is established.

Q: QoS Profile

The QoS Profile flag, if set, indicates that the Class of Service field contains a QoS Profile Identifier. If this flag is zero, it indicates that the Class of Service field contains a Priority or a Scheduler Identifier.

B: Bidirectional

The Bidirectional flag applies only to the Add Branch message. In all other Connection Management messages it is not used. It may only be used when establishing a point-to-point connection. The Bidirectional flag in an Add Branch message, if set, requests that two unidirectional virtual channels or virtual paths be established, one in the forward direction, and one in the reverse direction. It is equivalent to two Add Branch messages, one specifying the forward direction, and one specifying the reverse direction. The forward direction uses the values of Input Port, Input VPI, Input VCI, Output Port, Output VPI, and Output VCI as specified in the Add Branch message. The reverse direction is derived by exchanging the values specified in the Input Port, Input VPI, and Input VCI fields, with those of the Output Port, Output VPI, and Output VCI fields respectively. Thus, a virtual connection in the reverse direction arrives at the input port specified by the Output Port field, on the VPI/VCI specified by the Output VPI and Output VCI fields. It departs from the output port specified by the Input Port

field, on the VPI/VCI specified by the Input VPI and Input VCI fields.

The Bidirectional flag is simply a convenience to establish two unidirectional virtual connections in opposite directions between the same two ports, with identical VPI/VCIs, using a single Add Branch message. In all future messages the two unidirectional virtual connections must be handled separately. There is no bidirectional delete message. However, a single Delete Branches message with two Delete Branch Elements, one for the forward connection and one for the reverse, may be used.

C: Congestion Indication

The Congestion Indication flag, if set, requests that cells on this connection be marked if congestion is experienced. If this connection passes through a queue that the switch considers to be congested, the Congestion Experienced bit will be set in the Payload Type field of the cell header of all cells on the connection. GSMP does not specify the algorithm or any threshold by which the switch decides when a queue is congested.

Input VPI

Identifies an ATM virtual path arriving at the switch input port indicated by the Input Port field.

Input VCI

Identifies an ATM virtual channel arriving on the virtual path indicated by the Input VPI field at the switch input port indicated by the Input Port field. For virtual path connections the Input VCI field is not used.

Output Port

Identifies a switch output port.

x: Unused

Output VPI

Identifies an outgoing virtual path departing from the switch output port indicated in the Output Port field.

Output VCI

Identifies an outgoing virtual channel departing on the virtual path indicated by the Output VPI field from the switch output port indicated in the Output Port field. For virtual path connections the Output VCI field is not used.

Number of Branches

In a success response message and a failure response message, gives the number of output branches on a virtual channel connection or a virtual path connection after completion of the requested operation. (A point-to-point connection will have one branch, a point-to-multipoint connection will have two or more branches.) If the switch is unable to keep track of the number of branches on a virtual path connection or a virtual channel connection it must respond with the value 0xFFFF meaning: "number of branches unknown". This field is not used in the request message.

Class of Service

This field can contain either a QoS Profile Identifier, a Priority, or a Scheduler Identifier. If the QoS Profile flag in the Flags field is set, the Class of Service field contains a QoS Profile. If the QoS Profile flag in the Flags field is zero, and the value of the Class of Service field is greater than or equal to 0x100, the Class of Service field contains a Scheduler Identifier. If the QoS Profile flag in the Flags field is zero, and the value of the Class of Service field is less than 0x100, the Class of Service field contains a Priority. (Values of Scheduler Identifier less than 0x100 are interpreted as priorities.) The Class of Service field is only used in the Add Branch and Move Branch messages.

A QoS Profile Identifier is an opaque 16-bit value. It is used to identify a QoS profile in the switch which specifies the Quality of Service required by the connection. QoS profiles are established by a mechanism external to GSMP.

A Scheduler Identifier is an alternative method of communicating the QoS requirements of a connection. The Scheduler Identifier is defined in Section 9, "Quality of Service Messages."

A Priority specifies the priority of the connection for Add Branch and Move Branch messages that choose not to use a QoS profile, or the QoS capabilities defined in Section 9, "Quality of Service Messages." The highest priority is numbered zero and the lowest priority is numbered "Q-1" where "Q" is the number of priorities that the output port can support. The ability to offer different qualities of service to different connections based upon their priority is assumed to be a property of the output port of the

switch. It is assumed that for virtual path connections or virtual channel connections that share the same output port, an ATM cell on a connection with a higher priority is much more likely to exit the switch before an ATM cell on a connection with a lower priority, if they are both in the switch at the same time. The number of priorities that each output port can support is given in the Port Configuration message.

For all connection management messages, except the Delete Branches message, the success response message is a copy of the request message returned with the Result field indicating success and the Number of Branches field indicating the number of branches on the connection after completion of the operation. The Code field is not used in a connection management success response message.

The failure response message is a copy of the request message returned with a Result field indicating failure and the Number of Branches field indicating the number of branches on the connection.

Fundamentally, no distinction is made between point-to-point and point-to-multipoint connections. By default, the first Add Branch message for a particular Input Port, Input VPI, and Input VCI will establish a point-to-point virtual connection. The second Add Branch message with the same Input Port, Input VPI, and Input VCI fields will convert the connection to a point-to-multipoint virtual connection with two branches. (For virtual path connections the Input VCI is not required.) However, to avoid possible inefficiency with some switch designs, the Multicast Flag is provided. If the controller knows that a new connection is point-to-multipoint when establishing the first branch, it may indicate this in the Multicast Flag. Subsequent Add Branch messages with the same Input Port, Input VPI, and Input VCI fields will add further branches to the point-to-multipoint connection. Use of the Delete Branch message on a point-to-multipoint connection with two branches will result in a point-to-point connection. However, the switch may structure this connection as a point-to-multipoint connection with a single output branch if it chooses. (For some switch designs this structure may be more convenient.) Use of the Delete Branch message on a point-to-point connection will delete the point-to-point connection. There is no concept of a connection with zero output branches. All connections are unidirectional, one input virtual path or virtual channel to one or more output virtual paths or virtual channels.

GSMP supports point-to-point and point-to-multipoint connections. A multipoint-to-point connection is specified by establishing multiple point-to-point connections each of them specifying the same output branch. (An output branch is specified by an output port and output

VPI for a virtual path connection and by an output port, output VPI, and output VCI for a virtual channel connection.) A multipoint-to-multipoint connection is specified by establishing multiple point-to-multipoint trees each of them specifying the same output branches.

The connection management messages apply both to virtual channel connections and virtual path connections. The Add Branch and Move Branch connection management messages have two Message Types. One Message Type indicates that a virtual channel connection is required, and the other Message Type indicates that a virtual path connection is required. The Delete Branches, Delete Tree, and Delete All connection management messages have only a single Message Type because they do not need to distinguish between virtual channel connections and virtual path connections. For virtual path connections, neither Input VCI fields nor Output VCI fields are required. They should be set to zero by the sender and ignored by the receiver. Virtual channel branches may not be added to an existing virtual path connection. Conversely, virtual path branches may not be added to an existing virtual channel connection. In the Port Configuration message each switch input port may declare whether it is capable of supporting virtual path switching (i.e. accepting connection management messages requesting virtual path connections).

The connection management messages may be issued regardless of the Port Status of the switch port. Connections may be established or deleted when a switch port is in the Available, Unavailable, or any of the Loopback states. However, all connection state on an input port will be deleted when the port returns to the Available state from any other state, i.e. when a Port Management message is received for that port with the Function field indicating either Bring Up, or Reset Input Port.

4.1 Add Branch Message

The Add Branch message is a connection management message used to establish a virtual channel connection or a virtual path connection or to add an additional branch to an existing virtual channel connection or virtual path connection. It may also be used to check the connection state stored in the switch. The connection is specified by the Input Port, Input VPI, and Input VCI fields. The output branch is specified by the Output Port, Output VPI, and Output VCI fields. The quality of service requirements of the connection are specified by the Class of Service field. To request a virtual channel connection the Virtual Channel Connection (VCC) Add Branch message is:

Message Type = 16

To request a virtual path connection the Virtual Path Connection (VPC) Add Branch message is:

Message Type = 26

If a VPC Add Branch message is received and the switch input port specified by the Input Port field does not support virtual path switching, a failure response message must be returned indicating, "Virtual path switching is not supported on this input port."

If the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields; or the virtual path connection specified by the Input Port and Input VPI fields; does not already exist, it must be established with the single output branch specified in the request message. If the Bidirectional Flag in the Flags field is set, the reverse connection must also be established. The output branch should have the QoS attributes specified by the Class of Service field.

For the VCC Add Branch message, if a virtual path connection already exists on the virtual path specified by the Input Port and Input VPI fields, a failure response message must be returned indicating, "Attempt to add a virtual channel connection branch to an existing virtual path connection." For the VPC Add Branch message, if a virtual channel connection already exists on any of the virtual channels within the virtual path specified by the Input Port and Input VPI fields, a failure response message must be returned indicating, "Attempt to add a virtual path connection branch to an existing virtual channel connection."

If the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields; or the virtual path connection specified by the Input Port and Input VPI fields; already exists, but the specified output branch does not, the new output branch must be added. The new output branch should have the QoS attributes specified by the Class of Service field.

If the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields; or the virtual path connection specified by the Input Port and Input VPI fields; already exists and the specified output branch also already exists, the QoS attributes of the connection, specified by the Class of Service field, if different from the request message, should be changed to that in the request message. A success response message must be sent if the Result field of the request message is "AckAll". This allows the controller to periodically reassert the state of a connection or to change its priority. If the result field of the request message is "NoSuccessAck" a success response message should not be returned.

This may be used to reduce the traffic on the control link for messages that are reasserting previously established state. For messages that are reasserting previously established state, the switch must always check that this state is correctly established in the switch hardware (i.e. the actual connection tables used to forward cells).

If the output branch specified by the Output Port, Output VPI, and Output VCI fields for a virtual channel connection; or the output branch specified by the Output Port and Output VPI fields for a virtual path connection; is already in use by any connection other than that specified by the Input Port, Input VPI, and Input VCI fields, then the resulting output branch will have multiple input branches. If multiple point-to-point connections share the same output branch the result will be a multipoint-to-point connection. If multiple point-to-multipoint trees share the same output branches the result will be a multipoint-to-multipoint connection.

If the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields, or the virtual path connection specified by the Input Port and Input VPI fields, already exists, and the Bidirectional Flag in the Flags field is set, a failure response must be returned indicating: "Only point-to-point bidirectional connections may be established."

It should be noted that different switches support multicast in different ways. There will be a limit to the total number of point-to-multipoint connections any switch can support, and possibly a limit on the maximum number of branches that a point-to-multipoint connection may specify. Some switches also impose a limit on the number of different VPI/VCI values that may be assigned to the output branches of a point-to-multipoint connection. Many switches are incapable of supporting more than a single branch of any particular point-to-multipoint connection on the same output port. Specific failure codes are defined for some of these conditions.

4.2 Delete Tree Message

The Delete Tree message is a connection management message used to delete an entire virtual channel connection or an entire virtual path connection. All remaining branches of the connection are deleted. A virtual channel connection is specified by the Input Port, Input VPI, and Input VCI fields. A virtual path connection is specified by the Input Port and Input VPI fields. The Output Port, Output VPI, and Output VCI fields are not used in this message. The Delete Tree message is:

Message Type = 18

If the Result field of the request message is "AckAll" a success response message must be sent upon successful deletion of the specified connection. The success message must not be sent until the delete operation has been completed and if possible, not until all data on the connection, queued for transmission, has been transmitted. The Number of Branches field is not used in either the request or response messages of the Delete Tree message.

4.3 Verify Tree Message

The Verify Tree message has been removed from this version of GSMP. Its function has been replaced by the Number of Branches field in the success response to the Add Branch message which contains the number of branches on a virtual channel connection after successful completion of an add branch operation.

Message Type = 19 is reserved.

If a request message is received with Message Type = 19 a failure response must be returned with the Code field indicating: "The specified request is not implemented in this version of the protocol."

4.4 Delete All Message

The Delete All message is a connection management message used to delete all connections on a switch input port. All connections that arrive at the specified input port must be deleted. On completion of the operation all dynamically assigned VPI/VCI values for the specified port must be unassigned, i.e. there must be no virtual connections established in the VPI/VCI space that GSMP controls on this port. The Input VPI, Input VCI, Output Port, Output VPI, and Output VCI fields are not used in this message. The Delete All message is:

Message Type = 20

If the Result field of the request message is "AckAll" a success response message must be sent upon completion of the operation. The Number of Branches field is not used in either the request or response messages of the Delete All message. The success response message must not be sent until the operation has been completed.

The following failure response messages may be returned to a Delete All request.

The specified request is not implemented on this switch.

One or more of the specified ports does not exist.

Invalid Port Session Number.

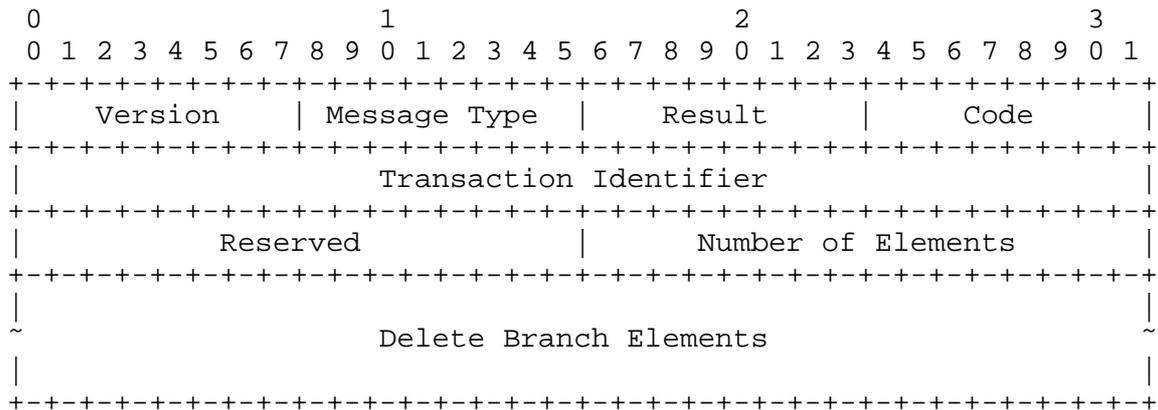
If any field in a Delete All message not covered by the above failure codes is invalid, a failure response must be returned indicating: "Invalid request message." Else, the delete all operation must be completed successfully and a success message returned. No other failure messages are permitted.

4.5 Delete Branches Message

The Delete Branches message is a connection management message used to request one or more delete branch operations. Each delete branch operation deletes a branch of a virtual channel connection or a virtual path connection, or in the case of the last branch of a connection, it deletes the connection. The Delete Branches message is:

Message Type = 17

The request message has the following format:



Number of Elements

Specifies the number of Delete Branch Elements to follow in the message. The number of Delete Branch Elements in a Delete Branches message must not cause the packet length to exceed the maximum transmission unit defined by the encapsulation.

Each Delete Branch Element specifies an output branch to be deleted and has the following structure:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Port Session Number                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Input Port                                         |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Error |      Input VPI      |      Input VCI      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Output Port                                       |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|x x x x|      Output VPI   |      Output VCI   |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

Error

Is used to return a failure code indicating the reason for the failure of a specific Delete Branch Element in a Delete Branches failure response message. The Error field is not used in the request message and must be set to zero. A value of zero is used to indicate that the delete operation specified by this Delete Branch Element was successful. Values for the other failure codes are specified in Section 3.2, "Failure Response Messages."

All other fields of the Delete Branch Element have the same definition as specified for the other connection management messages.

In each Delete Branch Element, either a virtual channel connection is specified by the Input Port, Input VPI, and Input VCI fields; or a virtual path connection is specified by the Input Port and Input VPI fields. The specific branch to be deleted is indicated by the Output Port, Output VPI, and Output VCI fields for virtual channel connections and by the Output Port and Output VPI for virtual path connections.

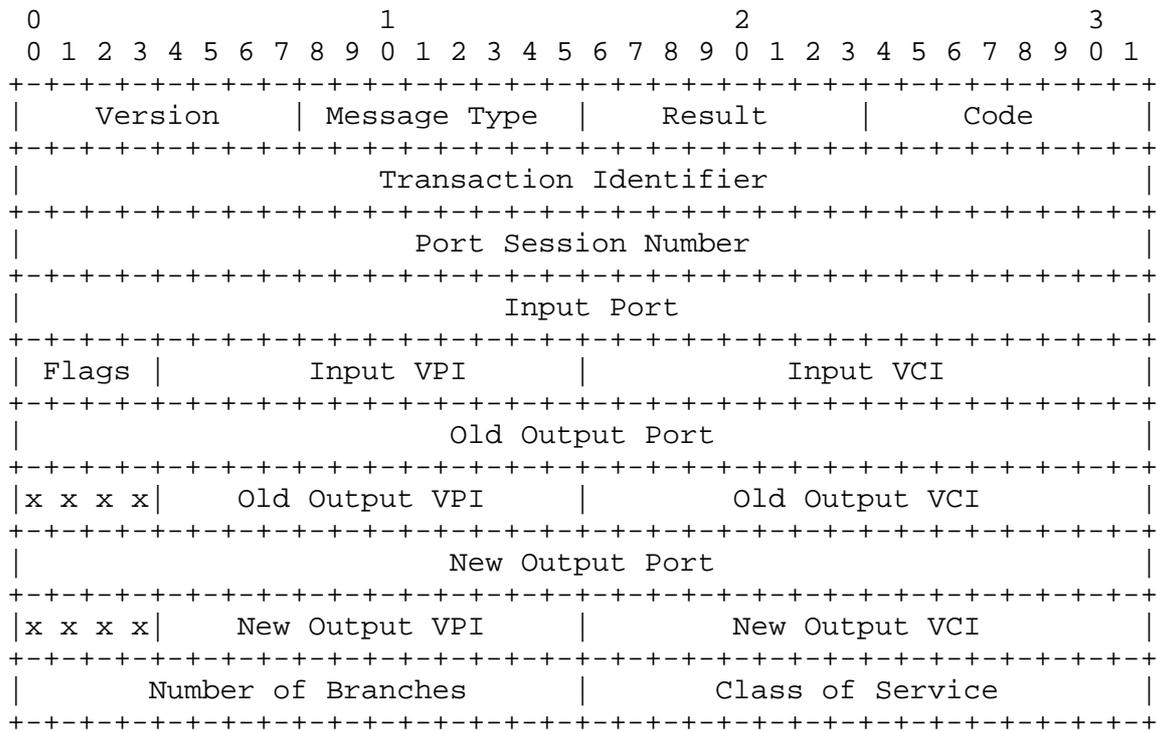
If the Result field of the Delete Branches request message is "AckAll" a success response message must be sent upon successful deletion of the branches specified by all of the Delete Branch Elements. The success response message must not be sent until all of the delete branch operations have been completed. The success response message is only sent if all of the requested delete branch operations were successful. No Delete Branch Elements are returned in a Delete Branches success response message and the Number of Elements field must be set to zero.

If there is a failure in any of the Delete Branch Elements a Delete Branches failure response message must be returned. The Delete Branches failure response message is a copy of the request message with the Code field of the entire message set to, "Failure specific

to the particular message type," and the Error field of each Delete Branch Element indicating the result of each requested delete operation. A failure in any of the Delete Branch Elements must not interfere with the processing of any other Delete Branch Elements.

4.6 Move Branch Message

The Move Branch message is used to move a branch of an existing connection from its current output port VPI/VCI to a new output port VPI/VCI in a single atomic transaction. This operation occurs frequently in IP switching, every time a flow is switched from hop-by-hop forwarding to a dedicated virtual channel. The Move Branch connection management message has the following format for both request and response messages:



The VCC Move Branch message is a connection management message used to move a single output branch of a virtual channel connection from its current output port, output VPI, and output VCI, to a new output port, output VPI, and output VCI on the same virtual channel connection. None of the other output branches are modified. When the operation is complete the original output VPI/VCI on the original output port will be deleted from the connection. The VCC Move Branch message is:

Message Type = 22

For the VCC Move Branch message, if the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields already exists, and the output branch specified by the Old Output Port, Old Output VPI, and Old Output VCI fields exists as a branch on that connection, the output branch specified by the New Output Port, New Output VPI, and New Output VCI fields is added to the connection and the branch specified by the Old Output Port, Old Output VPI, and Old Output VCI fields is deleted. If the Result field of the request message is "AckAll" a success response message must be sent upon successful completion of the operation. The success response message must not be sent until the Move Branch operation has been completed.

For the VCC Move Branch message, if the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields already exists, but the output branch specified by the Old Output Port, Old Output VPI, and Old Output VCI fields does not exist as a branch on that connection, a failure response must be returned with the Code field indicating, "The specified branch does not exist."

The VPC Move Branch message is a connection management message used to move a single output branch of a virtual path connection from its current output port and output VPI, to a new output port and output VPI on the same virtual channel connection. None of the other output branches are modified. When the operation is complete the original output VPI on the original output port will be deleted from the connection. The VPC Move Branch message is:

Message Type = 27

For the VPC Move Branch message, if the virtual path connection specified by the Input Port and Input VPI fields already exists, and the output branch specified by the Old Output Port and Old Output VPI fields exists as a branch on that connection, the output branch specified by the New Output Port and New Output VPI fields is added to the connection and the branch specified by the Old Output Port and Old Output VPI fields is deleted. If the Result field of the request message is "AckAll" a success response message must be sent upon successful completion of the operation. The success response message must not be sent until the Move Branch operation has been completed.

For the VPC Move Branch message, if the virtual path connection specified by the Input Port and Input VPI fields already exists, but the output branch specified by the Old Output Port and Old Output VPI fields does not exist as a branch on that connection, a failure response must be returned with the Code field indicating, "The specified branch does not exist."

If the virtual channel connection specified by the Input Port, Input VPI, and Input VCI fields; or the virtual path connection specified by the Input Port and Input VPI fields; does not exist, a failure response must be returned with the Code field indicating, "The specified connection does not exist."

If the output branch specified by the New Output Port, New Output VPI, and New Output VCI fields for a virtual channel connection; or the output branch specified by the New Output Port and New Output VPI fields for a virtual path connection; is already in use by any connection other than that specified by the Input Port, Input VPI, and Input VCI fields then the resulting output branch will have multiple input branches. If multiple point-to-point connections share the same output branch the result will be a multipoint-to-point connection. If multiple point-to-multipoint trees share the same output branches the result will be a multipoint-to-multipoint connection.

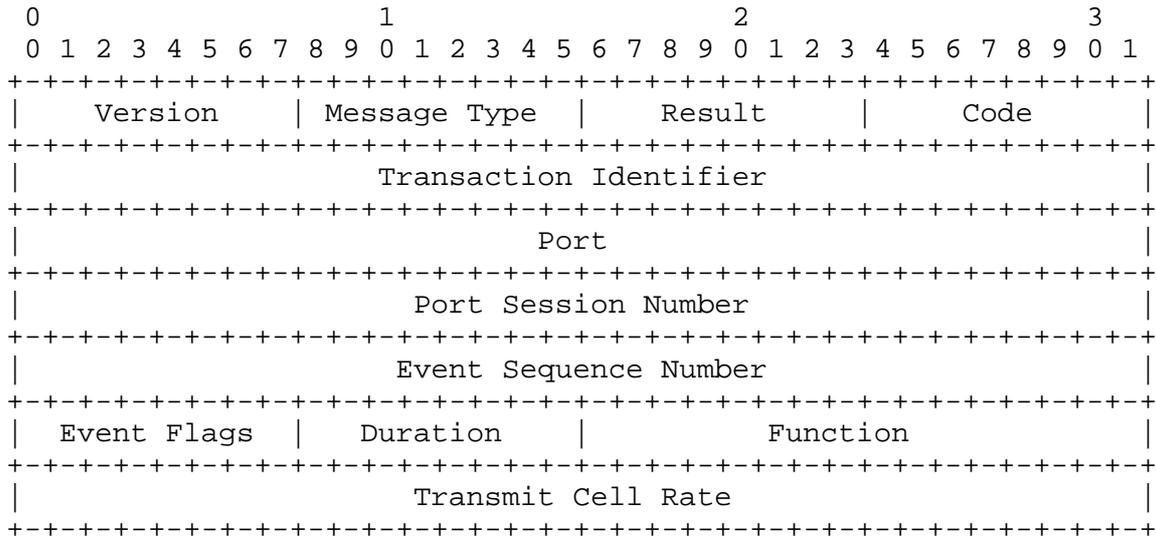
5. Port Management Messages

5.1 Port Management Message

The Port Management message allows a port to be brought into service, taken out of service, looped back, reset, or the transmit cell rate changed. Only the Bring Up and the Reset Input Port functions change the connection state (established connections) on the input port. Only the Bring Up function changes the value of the Port Session Number. If the Result field of the request message is "AckAll" a success response message must be sent upon successful completion of the operation. The success response message must not be sent until the operation has been completed. The Port Management Message is:

Message Type = 32

The Port Management message has the following format for the request and success response messages:



Event Sequence Number

In the success response message gives the current value of the Event Sequence Number of the switch port indicated by the Port field. The Event Sequence Number is set to zero when the port is initialized. It is incremented by one each time the port detects an asynchronous event that the switch would normally report via an Event message. If the Event Sequence Number in the success response differs from the Event Sequence Number of the most recent Event message received for that port, events have occurred that were not reported via an Event message. This is most likely to be due to the flow control that restricts the rate at which a switch can send Event messages for each port. In the request message this field is not used.

Event Flags

Field in the request message is used to reset the Event Flags in the switch port indicated by the Port field. Each Event Flag in a switch port corresponds to a type of Event message. When a switch port sends an Event message it sets the corresponding Event Flag on that port. The port is not permitted to send another Event message of the same type until the Event Flag has been reset. If the Function field in the request message is set to "Reset Event Flags," for each bit that is set in the Event Flags field, the corresponding Event Flag in the switch port is reset.

The Event Flags field is only used in a request message with the Function field set to "Reset Event Flags." For all other values of the Function field, the Event Flags field

is not used. In the success response message the Event Flags field must be set to the current value of the Event Flags for the port, after the completion of the operation specified by the request message, for all values of the Function field. Setting the Event Flags field to all zeros in a "Reset Event Flags" request message allows the controller to obtain the current state of the Event Flags and the current Event Sequence Number of the port without changing the state of the Event Flags.

The correspondence between the types of Event message and the bits of the Event Flags field is as follows:

```

  0 1 2 3 4 5 6 7
+---+---+---+---+
|U|D|I|N|Z|x x x|
+---+---+---+---+

```

```

U: Port Up           Bit 0, (most significant bit)
D: Port Down        Bit 1,
I: Invalid VPI/VCI Bit 2,
N: New Port         Bit 3,
Z: Dead Port        Bit 4,
x: Unused           Bits 5--7.

```

Duration

Is the length of time, in seconds, that any of the loopback states remain in operation. When the duration has expired the port will automatically be returned to service. If another Port Management message is received for the same port before the duration has expired, the loopback will continue to remain in operation for the length of time specified by the Duration field in the new message. The Duration field is only used in request messages with the Function field set to Internal Loopback, External Loopback, or Bothway Loopback.

Function

Specifies the action to be taken. The specified action will be taken regardless of the current status of the port (Available, Unavailable, or any Loopback state). If the specified function requires a new Port Session Number to be generated, the new Port Session Number must be returned in the success response message. The defined values of the Function field are:

```

Bring Up:
  Function = 1. Bring the port into service. All

```

connections that arrive at the specified input port must be deleted and a new Port Session Number must be selected using some form of random number. On completion of the operation all dynamically assigned VPI/VCI values for the specified input port must be unassigned, i.e. no virtual connections will be established in the VPI/VCI space that GSMP controls on this input port. The Port Status of the port afterwards will be Available.

Take Down:

Function = 2. Take the port out of service. Any cells received at this port will be discarded. No cells will be transmitted from this port. The Port Status of the port afterwards will be Unavailable.

The behavior is undefined if the port is taken down over which the GSMP session that controls the switch is running. (In this case the most probable behavior would be for the switch either to ignore the message or to terminate the current GSMP session and to initiate another session, possibly with the backup controller, if any.) The correct method to reset the link over which GSMP is running is to issue an RSTACK message in the adjacency protocol.

Internal Loopback:

Function = 3. Cells arriving at the output port from the switch fabric are looped through to the input port to return to the switch fabric. All of the ATM functions of the input port above the physical layer, e.g. header translation, are performed upon the looped back cells. The Port Status of the port afterwards will be Internal Loopback.

External Loopback:

Function = 4. Cells arriving at the input port from the external communications link are immediately looped back to the communications link at the physical layer without entering the input port. None of the ATM functions of the input port above the physical layer are performed upon the looped back cells. The Port Status of the port afterwards will be External Loopback.

Bothway Loopback:

Function = 5. Both internal and external loopback are

performed. The Port Status of the port afterwards will be Bothway Loopback.

Reset Input Port:

Function = 6. All connections that arrive at the specified input port must be deleted and the input and output port hardware re-initialized. On completion of the operation all dynamically assigned VPI/VCI values for the specified input port must be unassigned, i.e. no virtual connections will be established in the VPI/VCI space that GSMP controls on this input port. The range of VPIs and VCIs that may be controlled by GSMP on this port will be set to the default values specified in the Port Configuration message. The transmit cell rate of the output port must be set to its default value. The Port Session Number is not changed by the Reset Input Port function. The Port Status of the port afterwards will be Unavailable.

Reset Event Flags:

Function = 7. For each bit that is set in the Event Flags field, the corresponding Event Flag in the switch port must be reset. The Port Status of the port is not changed by this function.

Set Transmit Cell Rate:

Function = 8. Sets the transmit cell rate of the output port as close as possible to the rate specified in the Transmit Cell Rate field. In the success response message the Transmit Cell Rate must indicate the actual transmit cell rate of the output port. If the transmit cell rate of the requested output port cannot be changed, a failure response must be returned with the Code field indicating: "The transmit cell rate of this output port cannot be changed." If the transmit cell rate of the requested output port can be changed, but the value of the Transmit Cell Rate field is beyond the range of acceptable values, a failure response must be returned with the Code field indicating: "Requested transmit cell rate out of range for this output port." In the failure response message the Transmit Cell Rate must contain the same value as contained in the request message that caused the failure. The transmit cell rate of the output port is not changed by the Bring Up, Take Down, or any of the Loopback functions. It is returned to the default value by the Reset Input Port function.

Transmit Cell Rate

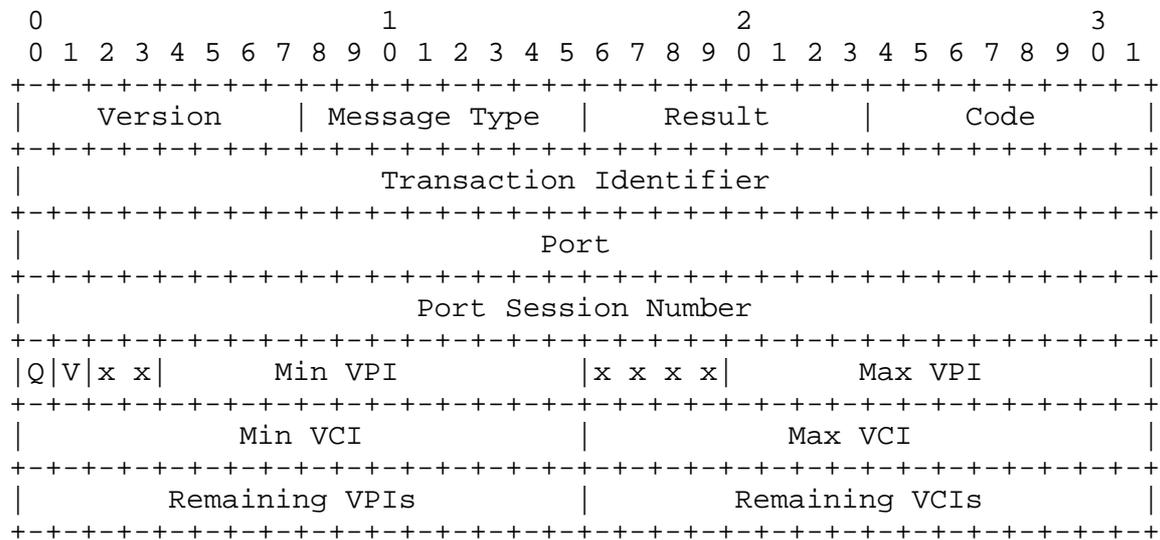
This field is only used in request and success response messages with the Function field set to "Set Transmit Cell Rate." It is used to set the output cell rate of the output port. It is specified in cells/s. If the Transmit Cell Rate field contains the value 0xFFFFFFFF the transmit cell rate of the output port should be set to the highest valid value.

5.2. Label Range Message

The default label range, Min VPI to Max VPI and Min VCI to Max VCI, is specified for each port by the Port Configuration or the All Ports Configuration messages. When the protocol is initialized, before the transmission of any Label Range messages, the label range of each port will be set to the default label range. (The default label range is dependent upon the switch design and configuration and is not specified by the GSMP protocol.) The Label Range message allows the range of VPIs supported by a specified port, or the range of VCIs supported by a specified VPI on a specified port, to be changed. Each switch port must declare whether it supports the Label Range message in the Port Configuration or the All Ports Configuration messages. The Label Range message is:

Message Type = 33

The Label Range message has the following format for the request and success response messages:



Flags

Q: Query

If the Query flag is set in a request message, the switch must respond with the current range of valid VPIs, or the current range of valid VCIs on a specified VPI, according to the VPI/VCI flag. The current label range is not changed by a request message with the Query flag set. If the Query flag is zero, the message is requesting a label change operation.

V: VPI/VCI

If the VPI/VCI flag is set, the message refers to a range of VPIs only. The Min VCI and Max VCI fields are unused. If the VPI/VCI flag is zero the message refers to a range of VCIs on either one VPI or on a range of VPIs.

x: Unused

Min VPI

Max VPI

Specify a range of VPI values, Min VPI to Max VPI inclusive. A single VPI may be specified with a Min VPI and a Max VPI having the same value. In a request message, if the value of the Max VPI field is less than or equal to the value of the Min VPI field, the requested range is a single VPI with a value equal to the Min VPI field. Zero is a valid value. In a request message, if the Query flag is set, and the VPI/VCI flag is zero, the Max VPI field specifies a single VPI and the Min VPI field is not used. The maximum valid value of these fields for both request and response messages is 0xFFFF.

Min VCI

Max VCI

Specify a range of VCI values, Min VCI to Max VCI inclusive. A single VCI may be specified with a Min VCI and a Max VCI having the same value. In a request message, if the value of the Max VCI field is less than or equal to the value of the Min VCI field, the requested range is a single VCI with a value equal to the Min VCI field. Zero is a valid value. (However, VPI=0, VCI=0 is not available as a virtual channel connection as it is used as a special value in ATM to indicate an unassigned cell.)

Remaining VPIs

Remaining VCIs

These fields are unused in the request message. In the

success response message and in the failure response message these fields give the maximum number of remaining VPIs and VCIs that could be requested for allocation on the specified port (after completion of the requested operation in the case of the success response). It gives the switch controller an idea of how many VPIs and VCIs it could request. The number given is the maximum possible given the constraints of the switch hardware. There is no implication that this number of VPIs and VCIs is available to every switch port.

If the Query flag and the VPI/VCI flag are set in the request message, the switch must reply with a success response message containing the current range of valid VPIs that are supported by the port. The Min VPI and Max VPI fields are not used in the request message.

If the Query flag is set and the VPI/VCI flag is zero in the request message, the switch must reply with a success response message containing the current range of valid VCIs that are supported by the VPI specified by the Max VPI field. If the requested VPI is invalid, a failure response must be returned indicating: "One or more of the specified input VPIs is invalid." The Min VPI field is not used in either the request or success response messages.

If the Query flag is zero and the VPI/VCI flag is set in the request message, the Min VPI and Max VPI fields specify the new range of VPIs to be allocated to the input port specified by the Port field. Whatever the range of VPIs previously allocated to this port it should be increased or decreased to the specified value.

If the Query flag and the VPI/VCI flag are zero in the request message, the Min VCI and Max VCI fields specify the range of VCIs to be allocated to each of the VPIs specified by the VPI range. Whatever the range of VCIs previously allocated to each of the VPIs within the specified VPI range on this port, it should be increased or decreased to the specified value. The allocated VCI range must be the same on each of the VPIs within the specified VPI range.

The success response to a Label Range message requesting a change of label range is a copy of the request message with the Remaining VPIs and Remaining VCIs fields updated to the new values after the Label Range operation.

If the switch is unable to satisfy a request to change the VPI range, it must return a failure response message with the Code field set to "Cannot support requested VPI range." In this failure response

message the switch must use the Min VPI and Max VPI fields to suggest a VPI range that it would be able to satisfy.

If the switch is unable to satisfy a request to change the VCI range on all VPIs within the requested VPI range, it must return a failure response message with the Code field set to "Cannot support requested VCI range on all requested VPIs." In this failure response message the switch must use the Min VPI, Max VPI, Min VCI, and Max VCI fields to suggest a VPI and VCI range that it would be able to satisfy.

In all other failure response messages for the label range operation the switch must return the values of Min VPI, Max VPI, Min VCI, and Max VCI from the request message.

While switches can typically support all 256 or 4096 VPIs the VCI range that can be supported is often more constrained. Often the Min VCI must be 0 or 32. Typically all VCIs within a particular VPI must be contiguous. The hint in the failure response message allows the switch to suggest a label range that it could satisfy in view of its particular architecture.

While the Label Range message is defined to specify both a range of VPIs and a range of VCIs within each VPI, the most likely use is to change either the VPI range or the range of VCIs within a single VPI. It is possible for a VPI to be valid but to be allocated no valid VCIs. Such a VPI could be used for a virtual path connection but to support virtual channel connections it would need to be allocated a range of VCIs.

A Label Range request message may be issued regardless of the Port Status or the Line Status of the target switch port. If the Port field of the request message contains an invalid port (a port that does not exist or a port that has been removed from the switch) a failure response message must be returned with the Code field set to, "One or more of the specified ports does not exist."

6. State and Statistics Messages

The state and statistics messages permit the controller to request the values of various hardware counters associated with the switch input and output ports, virtual path connections, virtual channel connections, and QoS Classes. They also permit the controller to request the connection state of a switch input port. The Connection Activity message is used to determine whether one or more specific virtual channel connections or virtual path connections have recently been carrying traffic. The Statistics message is used to query the various port, connection, and QoS class traffic and error counters.

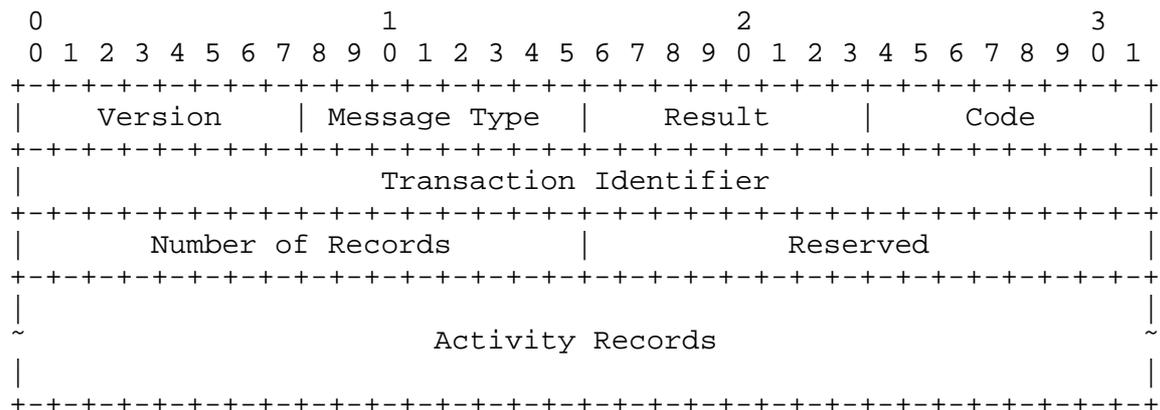
The Report Connection State message is used to request an input port to report the connection state for a single virtual channel connection, a single virtual path connection, or for the entire input port.

6.1 Connection Activity Message

The Connection Activity message is used to determine whether one or more specific virtual channel connections or virtual path connections have recently been carrying traffic. The Connection Activity message contains one or more Activity Records. Each Activity Record is used to request and return activity information concerning a single virtual channel connection or virtual path connection. Each virtual channel connection is specified by its input port, input VPI, and input VCI. Each virtual path connection is specified by its input port and input VPI. These are specified in the Input Port, Input VPI, and Input VCI fields of each Activity Record. Two forms of activity detection are supported. If the switch supports per connection traffic accounting, the current value of the traffic counter for each specified virtual channel connection or virtual path connection must be returned. The units of traffic counted are not specified but will typically be either cells or frames. The controller must compare the traffic counts returned in the message with previous values for each of the specified connections to determine whether each connection has been active in the intervening period. If the switch does not support per connection traffic accounting, but is capable of detecting per connection activity by some other unspecified means, the result may be indicated for each connection using the Flags field. The Connection Activity message is:

Message Type = 48

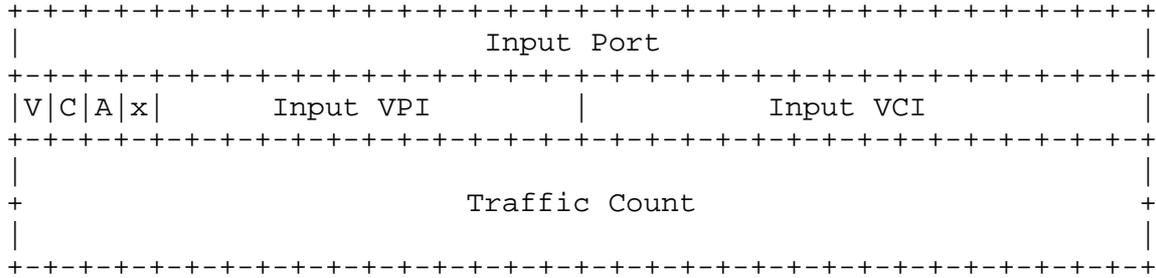
The Connection Activity request and success response messages have the following format:



Number of Records

Field specifies the number of Activity Records to follow. The number of Connection Activity records in a single Connection Activity message must not cause the packet length to exceed the maximum transmission unit defined by the encapsulation.

Each Activity Record has the following format:



Input Port

Identifies the port number of the input port on which the connection of interest arrives in order to identify the connection (regardless of whether the traffic count for the connection is maintained on the input port or the output port).

Input VPI

Input VCI

Fields identify the specific virtual path connection or virtual channel connection for which statistics are being requested. For a virtual path connection the Input VCI field is not used.

Flags

V: Valid Record

In the success response message the Valid Record flag is used to indicate an invalid Activity Record. The flag must be zero if any of the fields in this Activity Record are invalid, if the input port specified by the Input Port field does not exist, or if the specified connection does not exist. If the Valid Record flag is zero in a success response message, the Counter flag, the Activity flag, and the VC Traffic Count field are undefined. If the Valid Record flag is set, the Activity Record is valid, and the Counter and Activity flags are valid. The Valid Record flag is not used in the request message.

C: Counter

In a success response message, if the Valid Record flag is set, the Counter flag, if zero, indicates that the value in the VC Traffic Count field is valid. If set, it indicates that the value in the Activity flag is valid. The Counter flag is not used in the request message.

A: Activity

In a success response message, if the Valid Record and Counter flags are set, the Activity flag, if set, indicates that there has been some activity on this connection since the last Connection Activity message for this connection. If zero, it indicates that there has been no activity on this connection since the last Connection Activity message for this connection. The Activity flag is not used in the request message.

x: Unused**Traffic Count**

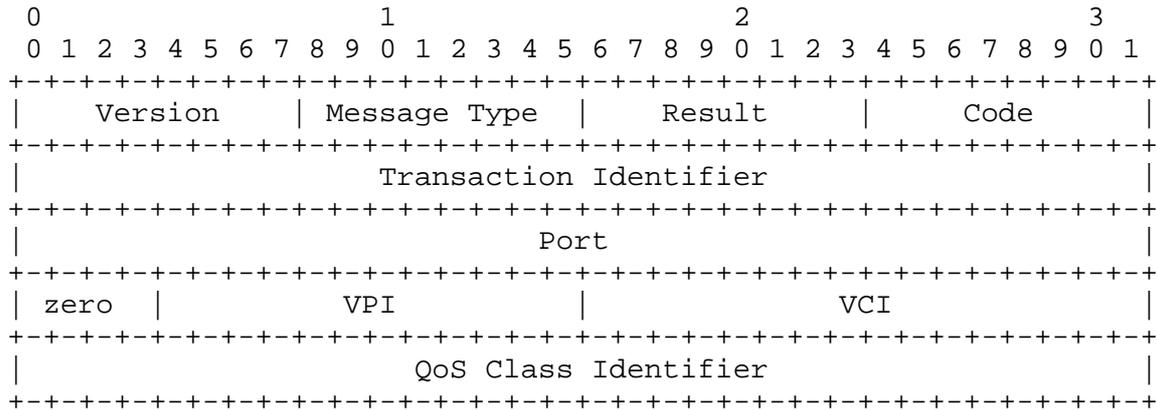
Field is not used in the request message. In the success response message, if the switch supports per connection traffic counting, the Traffic Count field must be set to the value of a free running, connection specific, 64-bit traffic counter counting traffic flowing across the specified connection. The value of the traffic counter is not modified by reading it. If per connection traffic counting is supported, the switch must report the Connection Activity result using the traffic count rather than using the Activity flag.

The format of the failure response is the same as the request message with the Number of Records field set to zero and no VC Activity records returned in the message. If the switch is incapable of detecting per connection activity, a failure response must be returned indicating, "The specified request is not implemented on this switch."

6.2 Statistics Messages

The Statistics messages are used to query the various port, connection, and QoS class traffic and error counters.

The Statistics request messages have the following format:



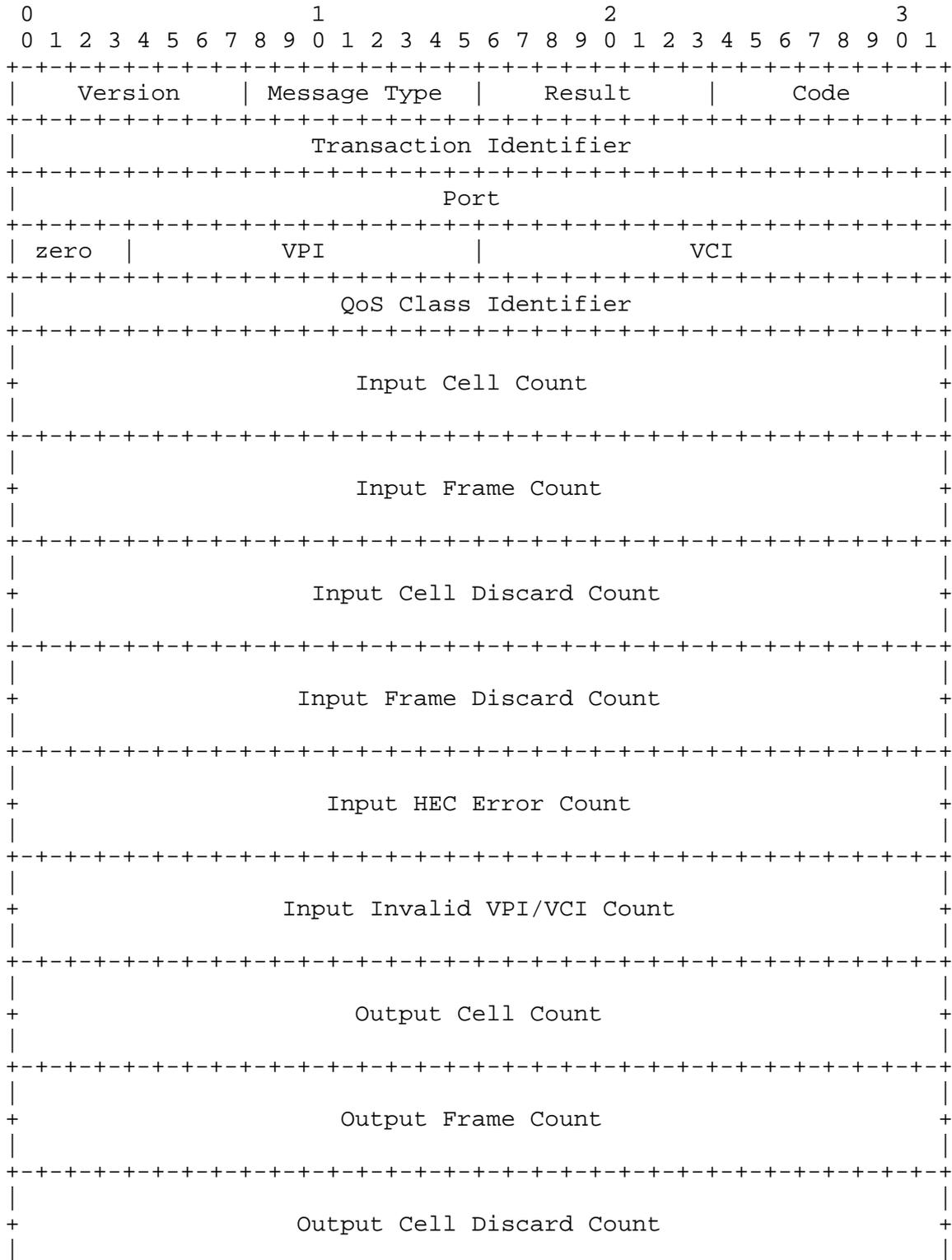
VPI
VCI

Fields identify the specific virtual path connection or virtual channel connection for which statistics are being requested. For a virtual path connection the Input VCI field is not used. For requests that do not require a virtual path connection or virtual channel connection to be specified, the VPI and VCI fields are not used.

QoS Class Identifier

Field identifies the QoS class for which statistics are being requested. This field is only used if the QoS Class Establishment message defined in section 9.4 is implemented.

The success response for the Statistics message has the following format:



```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|
+           Output Frame Discard Count           +
|
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

Port
VPI/VCI
QoS Class Identifier
Fields are the same as those of the request message.

Input Cell Count
Output Cell Count
Give the value of a free running 64-bit counter counting cells arriving at the input or departing from the output respectively.

Input Frame Count
Output Frame Count
Give the value of a free running 64-bit counter counting frames (packets) arriving at the input or departing from the output respectively.

Input Cell Discard Count
Output Cell Discard Count
Give the value of a free running 64-bit counter counting cells discarded due to queue overflow on an input port or on an output port respectively.

Input Frame Discard Count
Output Frame Discard Count
Give the value of a free running 64-bit counter counting frames discarded due to congestion on an input port or on an output port respectively.

HEC Error Count
Gives the value of a free running 64-bit counter counting cells discarded due to header checksum errors on arrival at an input port.

Invalid VPI/VCI Count
Gives the value of a free running 64-bit counter counting cells discarded because their VPI/VCI is invalid on arrival at an input port. For a virtual channel connection an incoming VPI/VCI is invalid if no connection is currently established having that value of VPI/VCI. For a virtual path connection an incoming VPI is invalid if no connection is currently established having that value of VPI.

6.2.1 Port Statistics Message

The Port Statistics message requests the statistics for the switch port specified in the Port field. The contents of the VPI/VCI and the QoS Class Identifier fields in the Port Statistics request message are ignored. All of the count fields in the success response message refer to per-port counts regardless of the connection or QoS class to which the cells belong. Any of the count fields in the success response message not supported by the port must be set to zero. The Port Statistics message is:

Message Type = 49

6.2.2 Connection Statistics Message

The Connection Statistics message requests the statistics for the virtual channel connection specified in the VPI/VCI field, or the virtual path connection specified in the VPI field, that arrives on the switch input port specified in the Port field, regardless of the QoS class to which the cells belong. All of the count fields in the success response message refer only to the specified connection. The HEC Error Count and Invalid VPI/VCI Count fields are not connection specific and must be set to zero. Any of the other count fields not supported on a per connection basis must be set to zero in the success response message. The Connection Statistics message is:

Message Type = 50

6.2.3 QoS Class Statistics Message

The QoS Class Statistics message requests the statistics for the QoS class specified by the QoS Class Identifier field that arrives on the switch input port specified in the Port field, regardless of the connection to which the cells belong. The QoS Statistics message is only used if the QoS Class Establishment message defined in section 9.4 is implemented. The contents of the VPI/VCI fields in the QoS Class Statistics request message are ignored. All of the count fields in the success response message refer only to the specified QoS class. The HEC Error Count and Invalid VPI/VCI Count fields are not specific to a QoS class and must be set to zero. Any of the other count fields not supported on a per QoS class basis must be set to zero in the success response message. The QoS Class Statistics message is:

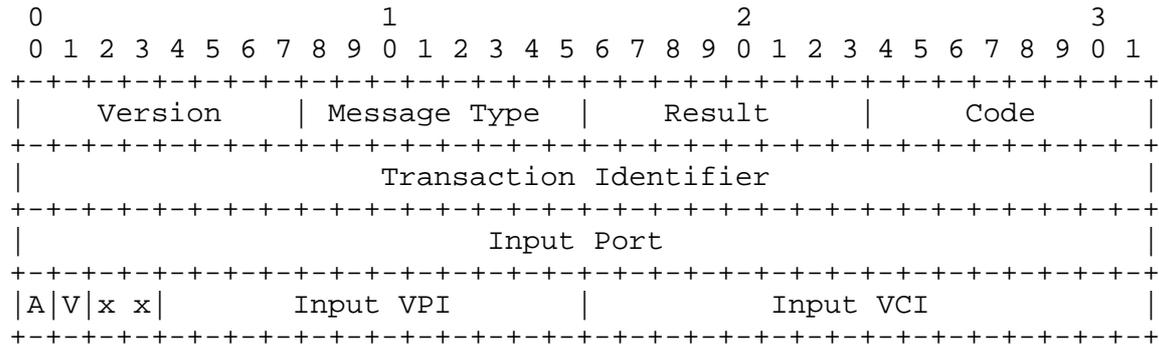
Message Type = 51

6.3 Report Connection State Message

The Report Connection State message is used to request an input port to report the connection state for a single virtual channel connection, a single virtual path connection, or for the entire input port. The Report Connection State message is:

Message Type = 52

The Report Connection State request message has the following format:



Input Port

Identifies the port number of the input port for which the connection state is being requested.

Flags

A: All Connections

If the All Connections flag is set, the message requests the connection state for all virtual path connections and virtual channel connections that arrive at the input port specified by the Input Port field. In this case the Input VPI and Input VCI fields and the VPI/VCI flag are unused.

V: VPI/VCI

If the All Connections flag is zero and the VPI/VCI flag is set, the message requests the connection state for the virtual path connection that arrives at the input port specified by the Input Port and Input VPI fields. If the specified Input VPI identifies a virtual path connection (i.e. a single switched virtual path) the state for that connection is requested. If the specified Input VPI identifies a virtual path containing virtual channel connections, the message requests the connection state for all virtual channel connections that belong to the specified virtual path. The Input VCI field is not used.

If the All Connections flag is zero and the VPI/VCI flag is also zero, the message requests the connection state for the virtual channel connection that arrives at the input port specified by the Port, Input VPI and Input VCI fields.

x: Unused.

Input VPI
Input VCI

Fields identify the specific virtual path connection, the specific virtual path, or the specific virtual channel connection for which connection state is being requested. For a virtual path connection (switched as a single virtual path connection) or a virtual path (switched as one or more virtual channel connections within the virtual path) the Input VCI field is not used. For requests that do not require a virtual path connection or virtual channel connection to be specified, the Input VPI and Input VCI fields are not used.

The Report Connection State success response message has the following format:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Version										Message Type										Result										Code									
Transaction Identifier																																							
Input Port																																							
Sequence Number																																							
~ Connection Records ~																																							

Input Port

Is the same as the Input Port field in the request message. It identifies the port number of the input port for which the connection state is being reported.

Sequence Number

In the case that the requested connection state cannot be reported in a single success response message, each successive success response message in reply to the same

that the Connection Record refers to a virtual channel connection.

M: More

If the More flag is set, it indicates that another Connection Record, in response to the same request message, will follow either in the same success response message or in a successive success response message. If the More flag is zero it indicates that this is the last Connection record in this success response message and that no further success response messages will be sent in response to the current request message. It indicates that the response to the request message is now complete.

Input VPI

Input VCI

The input VPI and VCI of the connection specified in this Connection Record. If this Connection Record specifies a virtual path connection (the VPC flag is set) the Input VCI field is unused.

Output Branch Records

Each Connection Record must contain one or more Output Branch Records. Each Output Branch Record specifies a single output branch belonging to the connection identified by the Input VPI and Input VCI fields of the Connection Record. A point-to-point connection will require only a single Output Branch Record. A point-to-multipoint connection will require multiple Output Branch Records. The last Output Branch Record of each Connection Record is indicated by the Last Branch flag of the Output Branch Record. If a point-to-multipoint connection has more output branches than can fit in a single Connection Record contained within a single success response message, that connection may be reported using multiple Connection Records in multiple success response messages.

Each Output Branch Record has the following format:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Output Port                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|L|x x x|      Output VPI      |      Output VCI      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

Output Port

The output port of the switch to which this output branch is routed.

Flags

L: Last Branch

The Last Branch flag, if set, indicates that this is the last Output Branch Record of this Connection Record. If zero, it indicates that one or more further Output Branch Records are to follow. If this is the last Output Branch Record in the message and the Last Branch flag is zero, further output branches belonging to the same connection will be given in another Connection Record. This Connection Record will be the first Connection Record in the next success response message. This Connection Record must have the same Input VPI and Input VCI values as the current Connection Record.

x: Unused.

Output VPI

Output VCI

The output VPI and VCI of the output branch specified in this Output Branch Record. If this Output Branch Record is part of a Connection Record that specifies a virtual path connection (the VPC flag is set) the Output VCI field is unused.

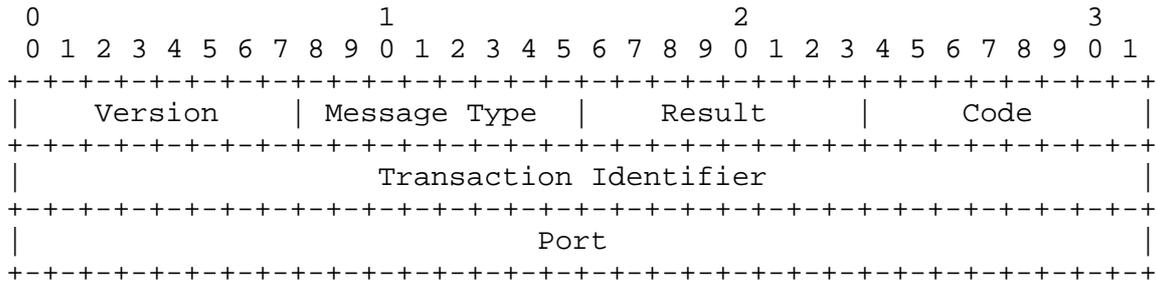
A Report Connection State request message may be issued regardless of the Port Status or the Line Status of the target switch port.

If the Input Port of the request message is valid, and the All Connections flag is set, but there are no connections established on that port, a failure response message must be returned with the code field set to, "Failure specific to the particular message type." For the Report Connection State message, this failure code indicates that no connections matching the request message were found. This failure message should also be returned if the Input Port of the request message is valid, the All Connections flag is zero, and no connections are found on that port matching the specified virtual path connection, virtual path, or virtual channel connection.

7. Configuration Messages

The configuration messages permit the controller to discover the capabilities of the switch. Three configuration request messages have been defined: Switch, Port, and All Ports.

All configuration request messages have the following format:



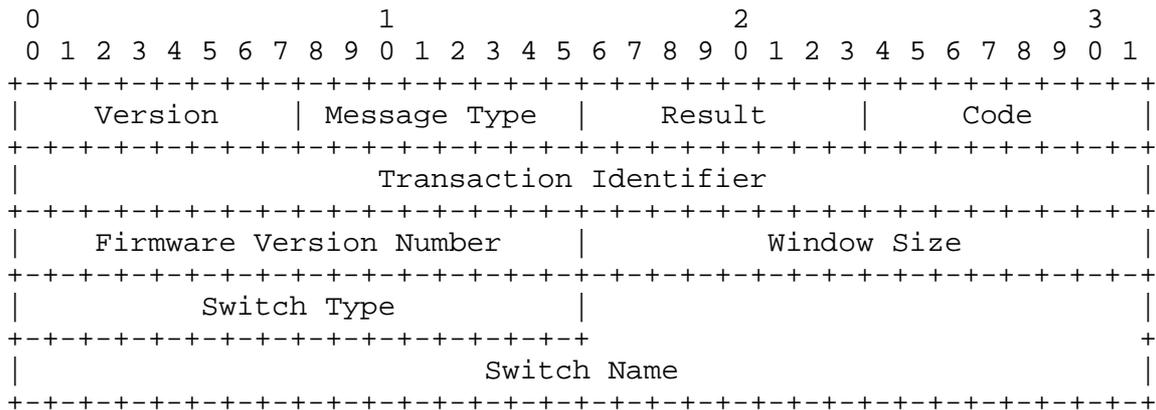
7.1 Switch Configuration Message

The Switch Configuration message requests the global (non port-specific) configuration for the switch. The Switch Configuration message is:

Message Type = 64

The Port field is not used in the request message.

The Switch Configuration success response message has the following format:



Firmware Version Number

The version number of the switch control firmware installed.

Window Size

The maximum number of unacknowledged request messages that may be transmitted by the controller without the possibility of loss. This field is used to prevent request messages being lost in the switch because of overflow in the receive buffer. The field is a hint to the controller. If desired, the controller may experiment with higher and

lower window sizes to determine heuristically the best window size.

Switch Type

A 16-bit field allocated by the manufacturer of the switch. (For these purposes the manufacturer of the switch is assumed to be the organization identified by the OUI in the Switch Name field.) The Switch Type identifies the product. When the Switch Type is combined with the OUI from the Switch Name the product is uniquely identified. Network Management may use this identification to obtain product related information from a database.

Switch Name

A 48-bit quantity that is unique within the operational context of the device. A 48-bit IEEE 802 MAC address, if available, may be used as the Switch Name. The most significant 24 bits of the Switch Name must be an Organizationally Unique Identifier (OUI) that identifies the manufacturer of the switch.

7.2 Port Configuration Message

The Port Configuration message requests the switch for the configuration information of a single switch port. The Port field in the request message specifies the port for which the configuration is requested. The Port Configuration message is:

Message Type = 65.

The Port Configuration success response message has the following format:

0								1								2								3															
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Version								Message Type								Result								Code															
Transaction Identifier																																							
Port																																							
Port Session Number																																							
V M L R				Min VPI								Q x x x				Max VPI																							
Min VCI																Max VCI																							
Receive Cell Rate																																							
Transmit Cell Rate																																							
Port Status								Port Type								Line Status								Priorities															
Physical Slot Number																Physical Port Number																							

Port

The switch port to which the configuration information refers. Configuration information relating to both the input and the output sides of the switch port is given. Port numbers are 32 bits wide and allocated by the switch. The switch may choose to structure the 32 bits into subfields that have meaning to the physical structure of the switch hardware (e.g. physical slot and port). This structure may be indicated in the Physical Slot Number and Physical Port Number fields.

Flags

V: VP Switching

The VP Switching flag, if set, indicates that this input port is capable of supporting virtual path switching. Else, if zero, it indicates that this input port is only capable of virtual channel switching.

M: Multicast Labels

The Multicast Labels flag, if set, indicates that this output port is capable of labelling each output branch of a point-to-multipoint tree with a different label. If zero, it indicates that this output port is not able to label

each output branch of a point-to-multipoint tree with a different label.

L: Logical Multicast

The Logical Multicast flag, if set, indicates that this output port is capable of supporting more than a single branch from any point-to-multipoint connection. This capability is often referred to as logical multicast. If zero, it indicates that this output port can only support a single output branch from each point-to-multipoint connection.

R: Label Range

The Label Range flag, if set, indicates that this switch port is capable of reallocating its VPI label range or its VCI label range and therefore accepts the Label Range message. Else, if zero, it indicates that this port does not accept Label Range messages.

Q: QoS

The QoS flag, if set, indicates that this switch port is capable of handling the Quality of Service messages defined in section 9 of this specification. Else, if zero, it indicates that this port does not accept the Quality of Service messages.

x: Unused

Min VPI

The default minimum value of dynamically assigned incoming VPI that the connection table on the input port supports and that may be controlled by GSMP. This value is not changed as a result of the Label Range message.

Max VPI

The default maximum value of dynamically assigned incoming VPI that the connection table on the input port supports and that may be controlled by GSMP. This value is not changed as a result of the Label Range message.

At power-on, after a hardware reset, and after the Reset Input Port function of the Port Management message, the input port must handle all values of VPI within the range Min VPI to Max VPI inclusive and GSMP must be able to control all values within this range. It should be noted that the range Min VPI to Max VPI refers only to the incoming VPI range that can be supported by the associated port. No restriction is placed on the values of outgoing

VPIs that may be written into the cell header. If the switch does not support virtual paths it is acceptable for both Min VPI and Max VPI to specify the same value, most likely zero.

Use of the Label Range message allows the range of VPIs supported by the port to be changed. However, the Min VPI and Max VPI fields in the Port Configuration and All Ports Configuration messages always report the same default values regardless of the operation of the Label Range message.

Min VCI

The default minimum value of dynamically assigned incoming VCI that the connection table on the input port can support and may be controlled by GSMP. This value is not changed as a result of the Label Range message.

Max VCI

The default maximum value of dynamically assigned incoming VCI that the connection table on the input port can support and may be controlled by GSMP. This value is not changed as a result of the Label Range message.

At power-on, after a hardware reset, and after the Reset Input Port function of the Port Management message, the input port must handle all values of VCI within the range Min VCI to Max VCI inclusive, for each of the virtual paths in the range Min VPI to Max VPI inclusive, and GSMP must be able to control all values within this range. It should be noted that the range Min VCI to Max VCI refers only to the incoming VCI range that can be supported by the associated port on each of the virtual paths in the range Min VPI to Max VPI. No restriction is placed on the values of outgoing VCIs that may be written into the cell header.

Use of the Label Range message allows the range of VCIs to be changed on each VPI supported by the port. However, the Min VCI and Max VCI fields in the Port Configuration and All Ports Configuration messages always report the same default values regardless of the operation of the Label Range message.

For a port over which the GSMP protocol is operating, the VCI of the GSMP control channel may or may not be reported as lying within the range Min VCI to Max VCI. A switch should honor a connection request message that specifies

the VCI value of the GSMP control channel even if it lies outside the range Min VCI to Max VCI.

Receive Cell Rate

The maximum rate of cells that may arrive at the input port in cells/s.

Transmit Cell Rate

The maximum rate of cells that may depart from the output port in cells/s. (The transmit cell rate of the output port may be changed by the Set Transmit Cell Rate function of the Port Management message.)

Port Status

Gives the administrative state of the port. The defined values of the Port Status field are:

Available:

Port Status = 1. The port is available to both send and receive cells. When a port changes to the Available state from any other administrative state, all dynamically assigned virtual connections must be cleared and a new Port Session Number must be generated.

Unavailable:

Port Status = 2. The port has intentionally been taken out of service. No cells will be transmitted from this port. No cells will be received by this port.

Internal Loopback:

Port Status = 3. The port has intentionally been taken out of service and is in internal loopback: cells arriving at the output port from the switch fabric are looped through to the input port to return to the switch fabric. All of the ATM functions of the input port above the physical layer, e.g. header translation, are performed upon the looped back cells.

External Loopback:

Port Status = 4. The port has intentionally been taken out of service and is in external loopback: cells arriving at the input port from the external communications link are immediately looped back to the communications link at the physical layer without entering the input port. None of the ATM functions of the input port above the physical layer are performed upon the looped back cells.

Bothway Loopback:

Port Status = 5. The port has intentionally been taken out of service and is in both internal and external loopback.

The Port Status of the port over which the GSMP session controlling the switch is running, must be declared Available. The controller will ignore any other Port status for this port. The Port Status of switch ports after power-on initialization is not defined by GSMP.

Port Type

The type of physical transmission interface for this port. The values for this field are defined by the atmIfType object specified in the Ipsilon IP Switch MIB [IpsilonMIB].

Line Status

The status of the physical transmission medium connected to the port. The defined values of the Line Status field are:

Up:

Line Status = 1. The line is able to both send and receive cells. When the Line Status changes to Up from either the Down or Test states, a new Port Session Number must be generated.

Down:

Line Status = 2. The line is unable either to send or receive cells or both.

Test:

Line Status = 3. The port or line is in a test mode, for example, power-on test.

Priorities

The number of different priority levels that this output port can assign to virtual connections. Zero is invalid in this field. If an output port is able to support "Q" priorities, the highest priority is numbered zero and the lowest priority is numbered "Q-1". The ability to offer different qualities of service to different connections based upon their priority is assumed to be a property of the output port of the switch. It may be assumed that for connections that share the same output port, an ATM cell on a connection with a higher priority is much more likely to exit the switch before an ATM cell on a connection with a lower priority if they are both in the switch at the same time.

Physical Slot Number

The physical location of the slot in which the port is located. It is an unsigned 16-bit integer that can take any value except 0xFFFF. The value 0xFFFF is used to indicate "unknown." The Physical Slot Number is not used by the GSMP protocol. It is provided to assist network management in functions such as logging, port naming, and graphical representation.

Physical Port Number

The physical location of the port within the slot in which the port is located. It is an unsigned 16-bit integer that can take any value except 0xFFFF. The value 0xFFFF is used to indicate "unknown." The Physical Port Number is not used by the GSMP protocol. It is provided to assist network management in functions such as logging, port naming, and graphical representation.

There must be a one to one mapping between Port Number and the Physical Slot Number and Physical Port Number combination. Two different Port Numbers must not yield the same Physical Slot Number and Physical Port Number combination. The same Port Number must yield the same Physical Slot Number and Physical Port Number within a single GSMP session. If both Physical Slot Number and Physical Port Number indicate "unknown" the physical location of switch ports may be discovered by looking up the product identity in a database to reveal the physical interpretation of the 32-bit Port Number.

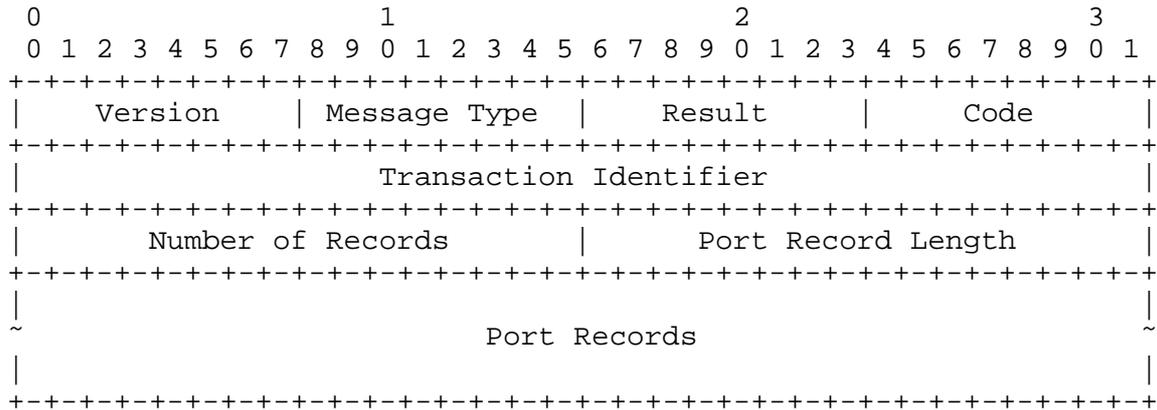
7.3 All Ports Configuration Message

The All Ports Configuration message requests the switch for the configuration information of all of its ports. The All Ports Configuration message is:

Message Type = 66

The Port field is not used in the request message.

The All Ports Configuration success response message has the following format:



Number of Records

Field gives the total number of Port Records to be returned in response to the All Ports Configuration request message. The number of port records in a single All Ports Configuration success response must not cause the packet length to exceed the maximum transmission unit defined by the encapsulation. If a switch has more ports than can be sent in a single success response message it must send multiple success response messages. All success response messages that are sent in response to the same request message must have the same Transaction Identifier as the request message and the same value in the Number of Records field. All success response messages that are sent in response to the same request message, except for the last message, must have the result field set to "More." The last message, or a single success response message, must have the result field set to "Success." All Port records within a success response message must be complete, i.e. a single Port record must not be split across multiple success response messages.

Port Record Length

Field gives the length of each port record in bytes. This is currently 32 but the Port Record Length field allows for the future definition of further fields at the end of the port record while preserving compatibility with earlier versions of the protocol.

Port Records

Follow in the remainder of the message. Each port record has the following format:

Event Sequence Number

The current value of the Event Sequence Number for the specified port. The Event Sequence Number is set to zero when the port is initialized. It is incremented by one each time the port detects an asynchronous event that the switch would normally report via an Event message. The Event Sequence Number must be incremented each time an event occurs even if the switch is prevented from sending an Event message due to the action of the flow control.

VPI/VCI

Field gives the VPI/VCI to which the event message refers. If this field is not required by the event message it is set to zero.

Each switch port must maintain an Event Sequence Number and a set of Event Flags, one Event Flag for each type of Event message. When a switch port sends an Event message it must set the Event Flag on that port corresponding to the type of the event. The port is not permitted to send another Event message of the same type until the Event Flag has been reset. Event Flags are reset by the "Reset Event Flags" function of the Port Management message. This is a simple flow control preventing the switch from flooding the controller with event messages. The Event Sequence Number of the port must be incremented every time an event is detected on that port even if the port is prevented from reporting the event due to the action of the flow control. This allows the controller to detect that it has not been informed of some events that have occurred on the port due to the action of the flow control.

8.1 Port Up Message

The Port Up message informs the controller that the Line Status of a port has changed from either the Down or Test state to the Up state. When the Line Status of a switch port changes to the Up state from either the Down or Test state a new Port Session Number must be generated, preferably using some form of random number. The new Port Session Number is given in the Port Session Number field. The VPI/VCI field is not used and is set to zero. The Port Up message is:

Message Type = 80

8.2 Port Down Message

The Port Down message informs the controller that the Line Status of a port has changed from the Up state to the Down state. This message will be sent to report link failure if the switch is capable of detecting link failure. The port session number that was valid before

the port went down is reported in the Port Session Number field. The VPI/VCI field is not used and is set to zero. The Port Down message is:

Message Type = 81

8.3 Invalid VPI/VCI Message

The Invalid VPI/VCI message is sent to inform the controller that one or more cells have arrived at an input port with a VPI/VCI that is currently not allocated to an assigned connection. The input port is indicated in the Port field, and the VPI/VCI in the VPI/VCI field. The Invalid VPI/VCI message is:

Message Type = 82

8.4 New Port Message

The New Port message informs the controller that a new port has been added to the switch. The port number of the new port is given in the Port field. A new Port Session Number must be assigned, preferably using some form of random number. The new Port Session Number is given in the Port Session Number field. The state of the new port is undefined so the VPI/VCI field is not used and is set to zero. The New Port message is:

Message Type = 83

8.5 Dead Port Message

The Dead Port message informs the controller that a port has been removed from the switch. The port number of the port is given in the Port field. The Port Session Number that was valid before the port was removed is reported in the Port Session Number field. The VPI/VCI fields are not used and are set to zero. The Dead Port message is:

Message Type = 84

9. Quality of Service Messages

The GSMP Quality of Service (QoS) messages allow a controller to group virtual path connections and virtual channel connections into QoS classes, and to allocate QoS resources to both QoS classes and to individual connections. At initialization, the switch describes its QoS capabilities to the controller, in terms of the abstract switch model, using the QoS Configuration message. The controller issues

Scheduler Establishment messages to configure the scheduler on each switch output port. It also issues QoS Class Establishment messages to configure QoS classes. Connections may be added to, or deleted from, a QoS class using the QoS Connection Management message. QoS resources may also be assigned to individual connections using the QoS Connection Management message. Connections that only require the scheduler may use the simple connection management messages defined in Section 3, "Connection Management Messages."

9.1 Abstract Switch Model

The abstract switch model, fig. 1, is the means by which a switch can describe its fundamental QoS capabilities to a controller. It consists of four main functions: a policer, a classifier, a regulator, and a scheduler. The classifier groups multiple connections (VPCs or VCCs) together into a QoS class such that QoS resources may be shared by the QoS class as a whole. Within a QoS class there is no differentiation between members of the class in terms of QoS resources received. However, the ordering of cells within each constituent VPC or VCC must be preserved on exit from the switch. Connections are not required to be aggregated into a QoS class with other connections; they may be allocated individual QoS resources.

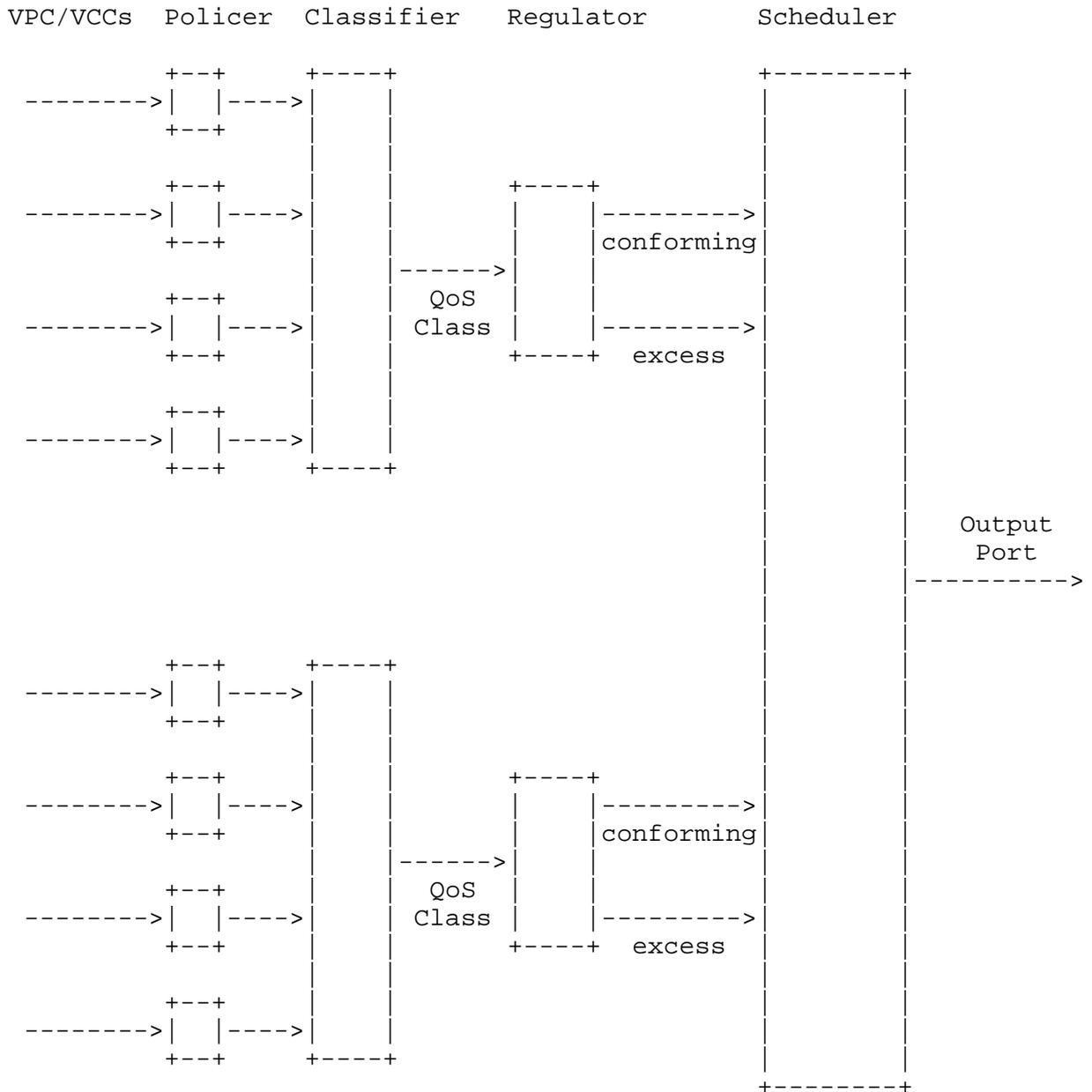


Fig. 1: Abstract Switch Model

The policer is a single input, single output device that can discard or tag cells. A policer may be applied to police each individual connection. A policer may also be applied to police the aggregate traffic of a QoS class. The policer is used to enforce an upper bound on the traffic on a connection or on a QoS class.

The regulator follows the policer and classifier. It offers either a policing function or a shaping function. The policing function evaluates cells as conforming to the rate specified by the regulator parameters or as being in excess of that rate. One of three actions can be specified to be taken for each cell as a result of this evaluation: tagging, discard or differentiated scheduling. Tagging sets the CLP bit of cells deemed to be in excess of the rate defined by the regulator parameters. The discard function discards excess cells. The differentiated scheduling function allows conforming cells and excess cells to be scheduled for service at different points in the scheduler. This would allow conforming cells, for example, to receive service with a QoS guarantee, whereas excess cells receive best-effort service. The implementation of differentiated scheduling, however, is complicated by the requirement not to reorder cells within each connection.

The shaping function of the regulator paces cells out, on each QoS class or individual connection, at the rate specified by the regulator parameters. No jitter requirement may be specified, nor is any specific guarantee of jitter given. If traffic arrives on any QoS class or individual connection at a greater rate than the output rate specified, that traffic will be delayed. If the delayed traffic for any QoS class or individual connection exceeds a bound, discard will occur. Differentiated scheduling is supported by the shaper but its application to shaping is somewhat different than its application to policing. Conforming traffic is that traffic which leaves the shaper as a result of the shaping process. The conforming pointer specifies the point in the scheduler structure where such traffic is scheduled for output. (This is typically the highest priority of the scheduler but the GSMP specification permits other priorities to be specified.) If an excess pointer is also enabled for a particular QoS class or individual connection, traffic in excess of the rate specified by the shaper may also be transmitted. The position of the excess pointer in the scheduler structure determines the undefined amount of additional traffic that will be supported. The excess traffic may be tagged if required, if tagging is supported. The excess pointer will receive the same share of bandwidth that a best-effort class or connection would receive at the same location in the scheduler structure.

The location of the classifier and regulator functions in the switch is important. If the classifier is located on an input port, only virtual connections that arrive at that input port may be aggregated into a QoS class. If the classifier is centralized, or located on an output port, virtual connections that arrive at any input port may be aggregated into the same QoS class. If the regulator is located on an output port all virtual connections within a QoS class passing through that regulator must exit the switch at that output port.

However, if the regulator is centralized, or located on an input port, virtual connections that are part of the same QoS class may be switched to different output ports. Each switch port must specify the location of its classifier and regulator functions.

The scheduler is located on the output port, fig. 2. It distributes the bandwidth of the output link between the QoS classes and individual connections. It is a two-level scheduler: a priority scheduler at one level and a FIFO or a weighted scheduler at the other. Up to 255 strict priority levels may be supported. Traffic in any specific priority level may only be transmitted if no traffic is queued for transmission in any higher priority level. Within each priority level a weighted scheduler may be defined. Each leaf of the scheduler tree is connected to a waiting room. The waiting room has two functions. When it receives service from the scheduler, it must select a QoS class or individual connection for transmission. When it is notified of traffic arrival on a QoS class or connection, it must decide whether there is enough room left in the waiting room to accept the traffic, else that traffic must be discarded. The waiting room has a size parameter indicating how much traffic may be accepted. Other queueing parameters may be attached to the waiting room. Multiple conforming and excess pointers from the regulators may point to each waiting room. Within a waiting room, the scheduling of multiple connections sharing that waiting room may support weighted sharing between the connections.

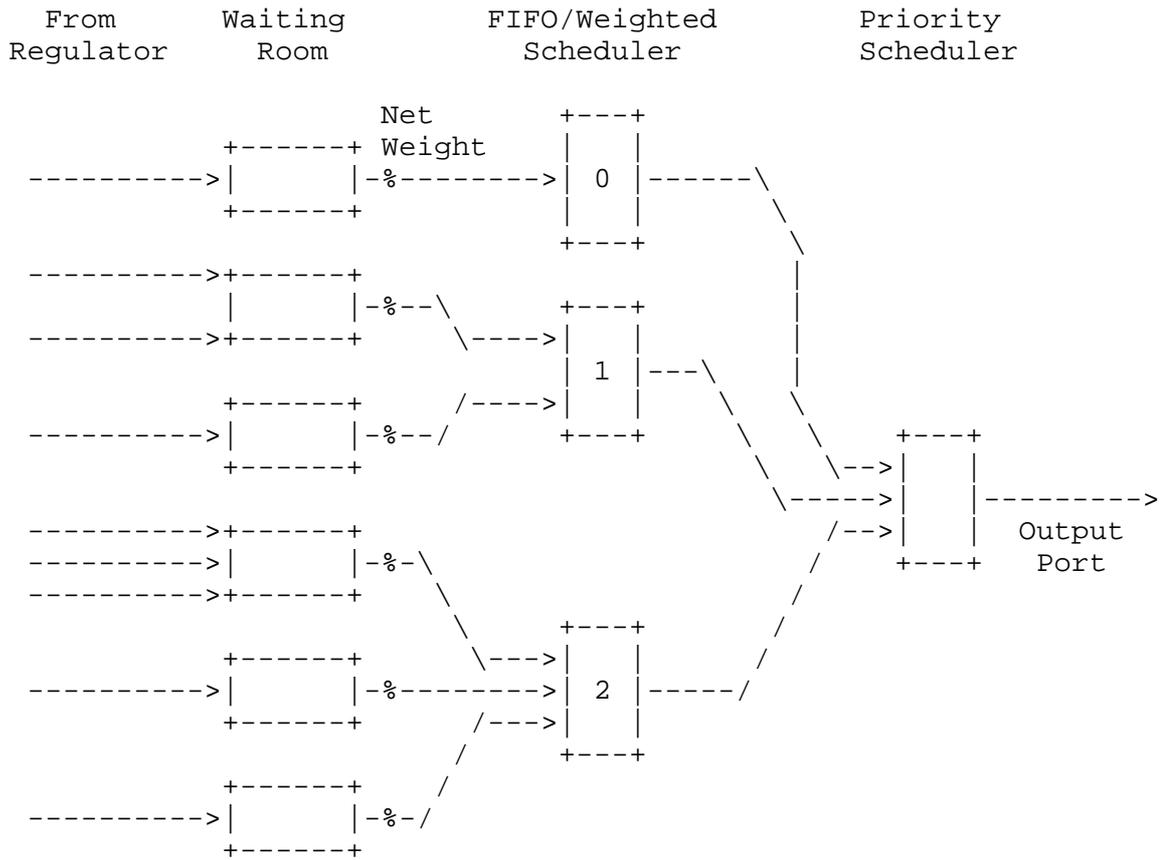


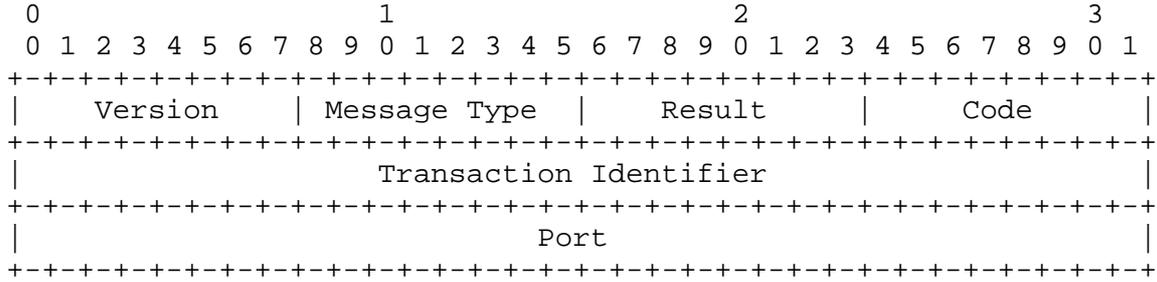
Fig. 2: The Scheduler

9.2 QoS Configuration Message

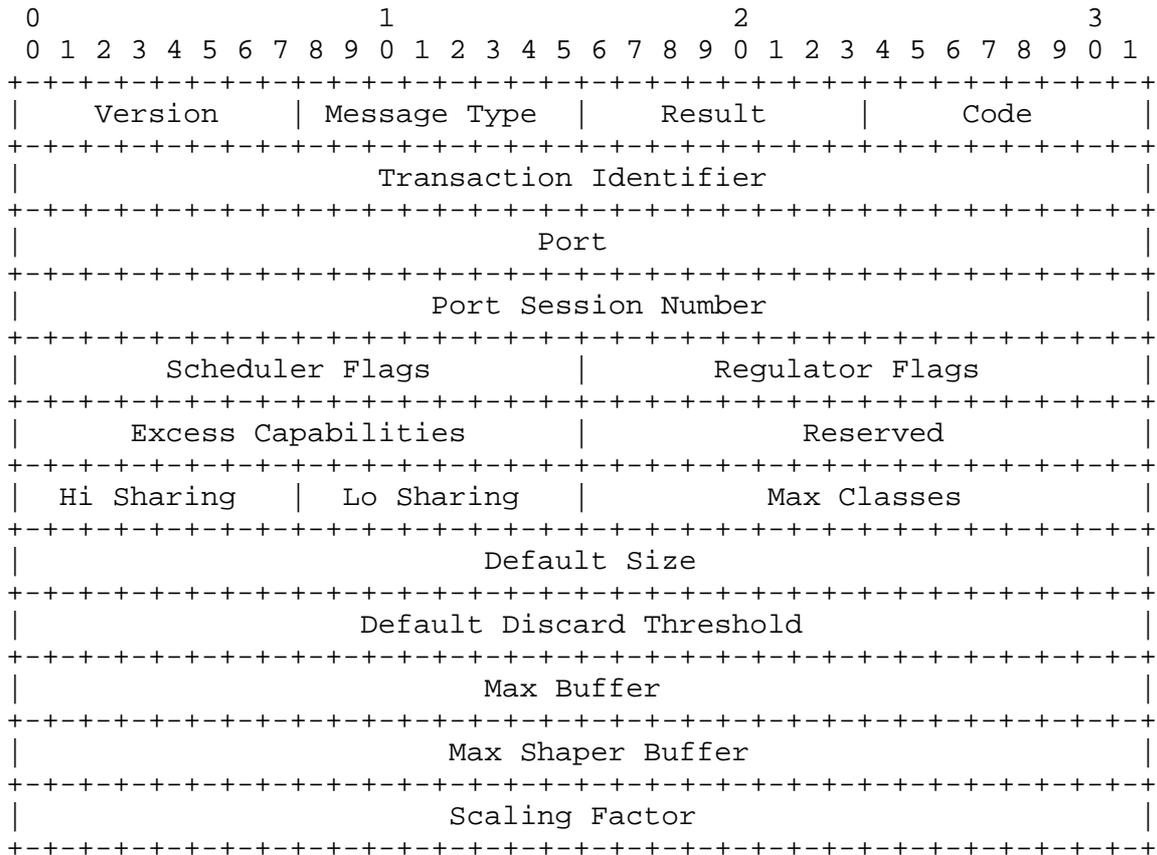
The QoS Configuration message permits the controller to discover the QoS capabilities of each switch port in terms of the abstract switch model. The QoS Configuration message is:

Message Type = 96

The QoS Configuration request message has the following format:



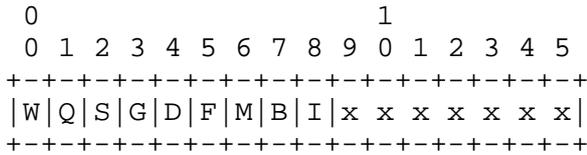
The QoS Configuration success response message has the following format:



Port

The switch port to which the QoS configuration information refers. QoS configuration information relating to both the input and the output sides of the switch port is given.

Scheduler Flags



W: Weighted Connections

Bit 0 of the Scheduler Flags field, if set, indicates that a weighted service algorithm (such as weighted round-robin) is available for allocation of service to individual connections within at least some waiting rooms. It means that a Connection Weight parameter can be attached to a QoS Connection Management message. Not all waiting rooms at all priority levels may be able to support this function. Whether a particular waiting room can support this function will be discovered when a QoS Connection Management message is issued.

Q: Weighted QoS Classes

Bit 1 of the Scheduler Flags field, if set, indicates that a weighted service algorithm (such as weighted round-robin) is available for allocation of service to QoS classes within at least some waiting rooms. It means that a QoS Class Weight parameter can be attached to a QoS Class Establishment message. Not all waiting rooms at all priority levels may be able to support this function. Whether a particular waiting room can support this function will be discovered when a QoS Class Establishment message is issued.

S: Shared Waiting Room

Bit 2 of the Scheduler Flags field, if set, indicates that multiple QoS classes and multiple connections may be scheduled within a single waiting room. This is expected to be the normal case. If Bit 2 of the Scheduler Flags field is zero, it indicates that only a single QoS class or a single connection may be directed to any single waiting room.

G: Global Max Classes

Bit 3 of the Scheduler Flags field, if set, indicates that the Max Classes field gives the maximum number of QoS classes that may be supported by the entire switch. If zero, it indicates that the Max Classes field gives the maximum number of QoS classes that may be supported by this switch port.

D: Packet Discard

Bit 4 of the Scheduler Flags field, if set, indicates that the scheduler on this output port is capable of packet discard. Packet discard indicates a discard algorithm that is aware of AAL-5 packet boundaries and attempts to discard whole packets. No specific algorithm is indicated though Early Packet Discard (EPD) is likely to be the most common. Other algorithms such as "push from front" schemes, dynamic threshold, or Random Early Detection (RED) are also examples of possible packet discard algorithms. The only parameters available to the packet discard algorithm, via GSMP, are the Size and Discard Threshold of the waiting room.

F: Frame-Based Scheduling

Bit 5 of the Scheduler Flags field, if set, indicates that the scheduler on this output port is capable of frame-based scheduling. In frame-based scheduling, a connection is only scheduled for transmission when a complete AAL-5 packet is available. When a connection is scheduled for transmission, all cells belonging to one or more complete packets from that connection will be transmitted without being interleaved with any other cells on that output port (regardless of their priority). Frame-based scheduling is a property of the waiting room and is requested in the Scheduler Establishment message. A QoS class may be routed through a waiting room configured with frame-based scheduling. In this case each component connection of the QoS class will receive frame based scheduling. For correct distribution of bandwidth, each QoS class that requires frame-based scheduling should have its own waiting room.

M: VC Merging

Bit 6 of the Scheduler Flags field, if set, indicates that the scheduler on this output port is capable of VC merging by a mechanism other than frame-based scheduling. VC merging indicates that the switch is capable of the multipoint-to-point merging of two or more incoming virtual connections onto a single outgoing virtual connection without interleaving cells from different AAL-5 packets that bear the same VPI/VCI. VC merging differs from frame-based scheduling in that cells with a different VPI/VCI may be interleaved with those of a multipoint-to-point VC merging connection. Thus, higher priority cells may be interleaved during the transmission of a packet on a lower priority VC merging connection. Most switches achieve VC merging by using frame-based scheduling. VC merging is a property of the waiting room and is requested in the

Scheduler Establishment message. A QoS class may be routed through a waiting room configured with VC merging. In this case each component connection of the QoS class will receive VC merging.

B: Shared Buffer

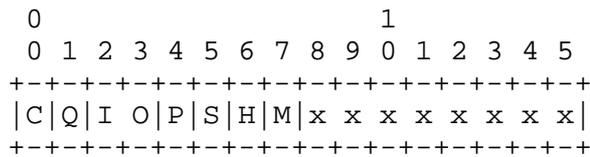
Bit 7 of the Scheduler Flags field, if set, indicates that at least some of the buffer space specified by the Max Buffer field is shared with other ports. If zero, it indicates that the buffer space specified by the Max Buffer field is not shared with other ports.

I: Identical Ports

Bit 8 of the Scheduler Flags field, if set, indicates that all ports of the switch have identical QoS capabilities. If this bit is set the controller does not have to request the QoS configuration of each port individually as all ports have the same capability.

x: Bits 9--15 of the Scheduler Flags field are not used.

Regulator Flags



C: Connection Policing

Bit 0 of the Regulator Flags field indicates that this input port supports the policing of individual incoming connections. The parameters for the policer are specified in the QoS Connection Management message when the connection is established.

Q: QoS Class Policing

If bit 1 of the Regulator Flags field is set, a policer function is available to police each QoS class on output from the classifier. The parameters for this policer are specified in the QoS Class Establishment message. If this bit is zero, no policer function is available to police a QoS class.

IO: QoS Class Location

Bits 2 and 3 of the Regulator Flags field specify the location of the classifier and regulator functions. If both

bits 2 and 3 of the Regulator Flags field are zero, no classifier or regulator function is available to this port.

If bit 2 of the Regulator Flags field is set and bit 3 is zero, the classifier and regulator functions are available on the input port. This implies that only virtual connections arriving at this input port may be grouped into QoS classes by this classifier. However, connections in a QoS class output from this regulator may be switched to any output port.

If bit 2 of the Regulator Flags field is zero and bit 3 is set, the classifier and regulator functions are available on the output port. This implies that virtual connections arriving at any input port may be grouped into QoS classes by this classifier. However, all connections in any QoS class output from this regulator may only be switched to this output port.

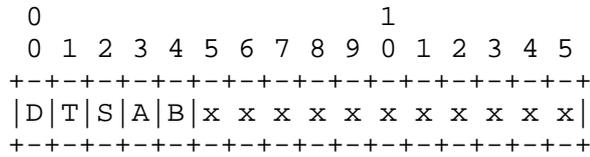
If both bits 2 and 3 of the Regulator Flags field are set, this switch port has access to centralized classifier and regulator functions. This implies that virtual connections arriving at any input port may be grouped into a QoS class by this classifier. Also, connections in a QoS class output from this regulator may be switched to any output port.

Regulator Function

- P: If bit 4 of the Regulator Flags field is set, the regulator is able to support the policing function.
- S: If bit 5 of the Regulator Flags field is set, the regulator is able to support the shaping function on all priority levels of the scheduler.
- H: If bit 5 of the Regulator Flags field is zero and bit 6 is set, the regulator is able to support the shaping function but only on the highest priority level of the scheduler. All connections and QoS classes using this regulator must be routed to a waiting room at the highest priority level of the scheduler.
- M: QoS Multicast
If bit 7 of the Regulator Flags field is set, any point-to-multipoint connection arriving on this input port, with QoS parameters established by the GSMP Quality of Service messages, must use the same QoS parameters for all output branches.

x: Bits 8--15 of the Regulator Flags field are not used.

Excess Capabilities



Policer:

- D: If bit 0 of the Excess Capabilities field is set, the policer function of the regulator is able to support discard.
- T: If bit 1 of the Excess Capabilities field is set, the policer function of the regulator is able to support tagging.
- S: If bit 2 of the Excess Capabilities field is set, the policer function of the regulator is able to support differentiated scheduling.

Shaper:

- A: If bit 3 of the Excess Capabilities field is set, the shaper function of the regulator is able to support tagging.
- B: If bit 4 of the Excess Capabilities field is set, the shaper function of the regulator is able to support differentiated scheduling.

x: Bits 5--15 of the Excess Capabilities field are not used.

Hi Sharing

Lo Sharing

Defines a range of priority levels that support weighted sharing. Each priority level in the range Lo Sharing to Hi Sharing inclusive, supports weighted sharing. A priority level that supports weighted sharing offers a weighted sharing algorithm (for example, weighted round-robin) between waiting rooms within that priority level. This permits the output link bandwidth available at that priority level, to be shared between the waiting rooms allocated to that priority level, according to the Net Weight parameter of each waiting room. The value 0xFF for both parameters indicates that this output port does not support weighted sharing in any priority level.

Max Classes

If bit 3 of the Scheduler Flags field is zero, Max Classes gives the maximum number of QoS classes that may be supported by this switch port. In this case the maximum number of QoS classes that may be supported by this switch port is not affected by the number of QoS classes in use by other switch ports. If bit 3 of the Scheduler Flags field is set, Max Classes gives the maximum number of QoS classes that may be supported by the entire switch. In this case it is assumed that use of these QoS classes may be distributed among the various switch ports.

Default Size

The size of waiting room that this output port allocates by default. The actual size of waiting room may be specified in the Scheduler Establishment message. The size of a waiting room specifies the maximum number of cells permitted to wait for transmission via that waiting room. Any further cells arriving at that waiting room beyond this number will be discarded.

Default Discard Threshold

The value of discard threshold that this output port allocates by default. The actual value of discard threshold may be specified in the Scheduler Establishment message. The discard threshold specifies the number of cells waiting for transmission via a waiting room after which further arriving cells will be subject to a discard mechanism.

Max Buffer

The maximum amount of buffer space, measured in cells, available to this port. If bit 7 of the Scheduler Flags field is zero this, buffer space is not shared with other ports. If bit 7 of the Scheduler Flags field is set, at least some of this buffer space is shared with other ports.

Max Shaper Buffer

The maximum amount of buffer space, measured in cells, available to a QoS connection or a QoS class within the shaper function of the regulator. This shaper buffer space is likely to be shared among all QoS classes and QoS connections using the shaper, so there is no guarantee that the amount of buffer space defined by the Max Shaper Buffer field will be available to any particular QoS class or QoS connection.

Scaling Factor

The QoS Class Establishment and QoS Connection Management messages require parameters that describe cell rates in cells per second or their reciprocal, cell interarrival periods, in seconds per cell. In order that these parameters may be specified with a 32-bit unsigned integer, the switch defines a Scaling Factor to be used in defining such parameters. By appropriate choice of the Scaling Factor the switch can select the range and granularity of rate or time that can be specified with the 32-bit unsigned integer. Further details are given in the discussion of the UPC Parameters field of the QoS Connection Management message.

9.3 Scheduler Establishment Message

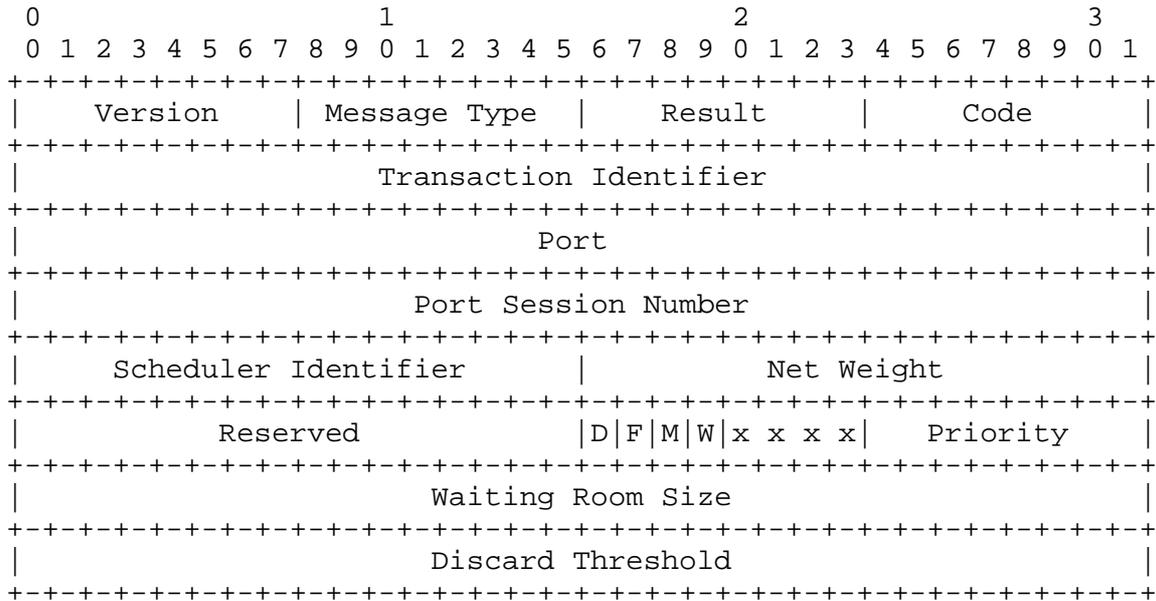
The Scheduler Establishment message is used to configure the scheduler on a specified output port. It is used to configure a waiting room, attach it to a leaf of the scheduler tree, and return a Scheduler Identifier to reference the waiting room. The Scheduler Establishment message may also be used to modify the parameters of an already established waiting room.

Scheduler Identifiers in the range 0--255 represent default values. They are used for the priority levels that may be specified in the Class of Service field of Connection Management messages without requiring explicit establishment via a Scheduler Establishment message. Each of these default values specifies a single waiting room with default parameters, configured as a FIFO queue, on each of the valid scheduler priority levels. (This permits Connection Management messages to continue to specify QoS requirements as a priority without requiring the use of any of the QoS messages.) The number of priority levels available to the scheduler is specified in the Priorities field of the Port Configuration and All Ports Configuration messages.

The Scheduler Establishment Message is:

Message Type = 97

The Scheduler Establishment request and success response messages have the following format:



Scheduler Identifier

The Scheduler Identifier is selected by the controller. It is used to identify the waiting room being established or modified in future messages. The Scheduler Identifier is taken from a namespace that is local to the switch port. A Scheduler Identifier in the Scheduler Establishment message must be greater than 0x00FF but less than 0xFFFF. The values 0 -- 0x00FF are reserved for use as default values. The default values of the Scheduler Identifier are used to specify the default settings for the scheduler. Each of the default values maps directly to one of the scheduler priority levels. The value 0xFFFF is reserved for use in the QoS Connection Management message.

Net Weight

The Net Weight specifies the share of the bandwidth available to the priority level, specified by the Priority field, that should be given to this waiting room. The Net Weight parameter is only valid if the priority level specified by the Priority field supports weighted sharing.

The Net Weight is an unsigned 16-bit field specifying a binary fraction. I.e. the bandwidth share, as a fraction of the bandwidth available to the priority level, is given by:

$$\text{Bandwidth share} = \text{Net Weight} * 2^{*(-16)}$$

A Net Weight of zero indicates equal sharing between all waiting rooms sharing this priority level that request a Net Weight of zero. While a 16-bit field is used to specify the Net Weight it is understood that the accuracy of the bandwidth sharing is hardware dependent and is not specified.

If weighted sharing is not required at a particular priority level, a waiting room with a Net Weight value of 0xFFFF must be specified for that priority level. A priority level that does not support weighted sharing can only support a single waiting room.

Flags

D: Packet Discard

Bit 0 of the Flags field, if set, indicates that packet discard is required on all connections and QoS classes routed through this waiting room.

F: Frame-Based Scheduling

Bit 1 of the Flags field, if set, indicates that frame-based scheduling is required on all connections and QoS classes routed through this waiting room. In frame-based scheduling, a connection is only scheduled for transmission when a complete AAL-5 packet is available. When a connection is scheduled for transmission, all cells belonging to one or more complete packets from that connection will be transmitted without being interleaved with any other cells on that output port. A QoS class may be routed through a waiting room configured with frame-based scheduling. In this case each component connection of the QoS class will receive frame based scheduling. For correct distribution of bandwidth, each QoS class that requires frame-based scheduling should have its own waiting room.

M: VC Merging

Bit 2 of the Scheduler Flags field, if set, indicates that VC merging is required on all connections and QoS classes routed through this waiting room. VC merging enables the multipoint-to-point merging of two or more incoming virtual connections onto a single outgoing virtual connection, without interleaving cells from different AAL-5 packets that bear the same VPI/VCI. VC merging differs from frame-based scheduling in that cells with a different VPI/VCI may be interleaved with those of a multipoint-to-point VC merging connection. Most switches achieve VC merging by

using frame-based scheduling. A QoS class may be routed through a waiting room configured with VC merging. In this case each component connection of the QoS class will receive VC merging.

W: Weighted Scheduling

Bit 3 of the Flags field, if set, indicates that weighted scheduling is required on all connections and QoS classes routed through this waiting room. All connections and QoS classes routed through this waiting room will require a Connection Weight or a QoS Class Weight respectively. The Connection Weight is specified in the QoS Connection Management message. The QoS Class Weight is specified in the QoS Class Establishment message. If weighted scheduling within this waiting room is unavailable, a failure response message must be returned indicating, "Weighted scheduling within this waiting room is unavailable."

Bit 3 of the Flags field, if zero, indicates that this waiting room should be configured as a single FIFO queue. All cells arriving at this waiting room will receive first-in-first-out service. If Frame-Based Scheduling or VC Merging are also selected, the strict first-in-first-out service discipline will be modified by the requirement to support Frame-Based Scheduling or VC Merging.

x: Bits 4--7 of the Flags field are not used.

Priority

Specifies the priority level in the scheduler to which the waiting room should be attached. Priorities are numbered from zero, with priority level zero being the highest priority.

Waiting Room Size

The required size of the waiting room. The size of a waiting room specifies the maximum number of cells permitted to wait for transmission via that waiting room. Any further cells arriving at that waiting room beyond this number will be discarded. If the switch is unable to grant the size requested in the Scheduler Establishment request message it may reply with the actual size allocated to the waiting room in the Waiting Room Size field of the success response message. A value of zero for the Waiting Room Size indicates that the default value should be used.

Discard Threshold

The required value of the discard threshold. The discard threshold specifies the number of cells waiting for transmission via a waiting room after which further arriving cells will be subject to a discard mechanism. The value of the Discard Threshold must be less than or equal to the value of the Waiting Room Size parameter for any given waiting room. If the switch is unable to grant the value of discard threshold requested in the Scheduler Establishment request message it may reply with the actual value of discard threshold allocated to the waiting room in the Discard Threshold field of the success response message. A value of zero for the Discard Threshold indicates that the default value should be used.

9.4 QoS Class Establishment Message

The QoS Class Establishment message is used to configure a QoS class on a specified port or to modify the parameters of an already established QoS class. It configures the classifier and the regulator functions for the QoS class. It also configures the QoS class policer if a policing function is available for QoS classes.

Two styles of QoS class are available. In one style each component connection of the QoS class may be routed independently to an output port and waiting room specified in its connection management message. In this case the Scheduler Identifier, and if required, the Excess Scheduler Id, are specified in the QoS Connection Management message that references this style of QoS class. In the alternative style of QoS class, all component connections in the QoS class are routed to the same waiting room on the same output port. In this case the Output Port, the Scheduler Identifier, and if required, the Excess Scheduler Id, are specified in the QoS Class Establishment message.

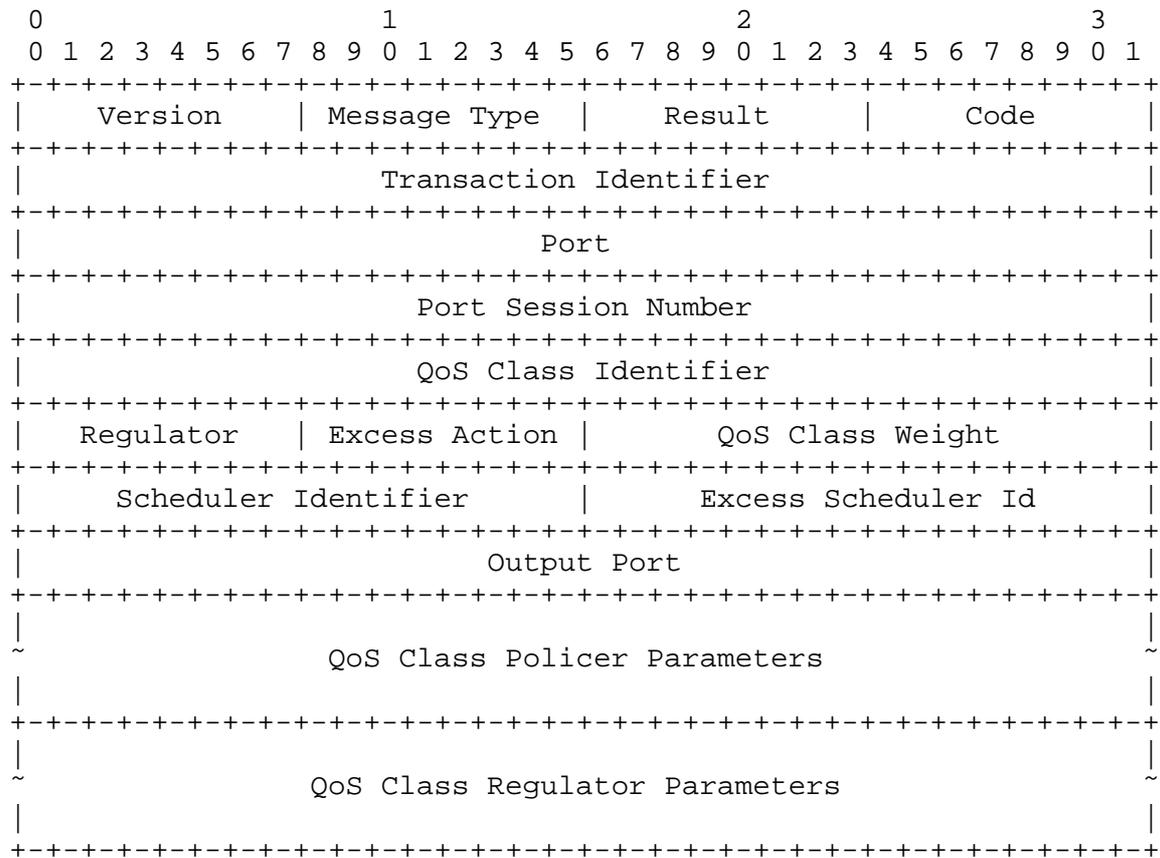
The classifier and regulator functions must be located together, either on an input port, on an output port, or centralized. Each port declares the location of its classifier and regulator functions at initialization using the QoS Configuration message. If the classifier and regulator functions are located on an input port, only connections that arrive at that input port may join a QoS class established on that port. However, each connection that is part of a QoS class established on that port may be switched to a different output port. If the classifier and regulator functions are located on an output port, connections that arrive at any input port may join a QoS class established on that port. However, all connections within a QoS class established on that port must be switched to that output port. For a centralized classifier and regulator function, there is

no restriction on the input ports on which connections in a QoS class must arrive, or on the output ports to which connections in a QoS class must be switched. (For the case of a centralized classifier and regulator the actual port specified in the QoS Class Establishment message is used only for administrative purposes. Any valid value of Port and Port Session Number, that specifies a centralized classifier and regulator function, may be used.)

The QoS Class Establishment message is:

Message Type = 98

The QoS Class Establishment request and success response messages have the following format:



QoS Class Identifier

The QoS Class Identifier is selected by the controller. It is used to identify the QoS class being established or modified, in future QoS Connection Management and QoS Class Establishment messages. It is taken from a namespace that

is global across the entire switch. No two QoS classes may have the same QoS Class Identifier regardless of the switch ports on which they are defined. A QoS Class Identifier in a QoS Class Establishment message must be greater than 0 and less than 0xFFFFFFFF.

Regulator

The Regulator field specifies which function is required of the regulator. Three possible functions are currently defined: none, policing, and shaping.

None: Regulator = 1
Policing: Regulator = 2
Shaping: Regulator = 3

If the Regulator function is specified as none, no operations are performed by the regulator on the cells output from the classifier. Cells output from the classifier are transferred directly to the waiting room specified by the Scheduler Identifier.

If policing is specified, a token bucket policer will be applied to the QoS class. The policer determines which cells conform to the specified policer traffic parameters and which do not. Conforming cells are transferred directly to the waiting room specified by the Scheduler Identifier. The action to be taken by the policer on the excess traffic is specified by the Excess Action field. The policer traffic parameters are specified in the QoS Class Regulator Parameters fields.

If shaping is specified, traffic shaping will be applied to the QoS class. Cells in a QoS class should leave the regulator spaced evenly apart at a rate defined by the QoS Class Regulator Parameters fields. These cells are transferred directly to the waiting room specified by the Scheduler Identifier. The jitter on the conforming cell stream on exit from the shaping function of the regulator is not specified.

Excess Action

0 1 2 3 4 5 6 7
+--+--+--+--+--+--+
|T|D|S|x x x x x|
+--+--+--+--+--+--+

T: Tagging

If bit 0 of the Excess Action field is set, all cells transferred to the waiting room specified by the Excess Scheduler Id will have their CLP bit set. If bit 0 of the Excess Action field is zero, the CLP bit of cells transferred to the waiting room specified by the Excess Scheduler Id will remain unchanged.

D: Discard

This function is only available if policing is selected as the regulator function. If the Regulator field specifies Policing, and bit 1 of the Excess Action field is set, all cells determined by the policer to be in excess of the traffic parameters must be discarded. In this case the Excess Scheduler Id is not used and bit 0 of the Excess Action field should be ignored.

S: Differentiated Scheduling

This function operates differently according to whether policing or shaping is selected as the regulator function.

If the Regulator field specifies Policing, and bit 1 of the Excess Action field is zero, and bit 2 of the Excess Action field is set, all cells determined by the policer to be in excess of the traffic parameters must be transferred to the waiting room specified by the Excess Scheduler Id. In this case care must be taken in the implementation to ensure that within each virtual path connection or virtual channel connection, cells depart in the same order that they arrived. If the Regulator field specifies Policing, and bit 1 of the Excess Action field is zero, and bit 2 of the Excess Action field is zero, all cells determined by the policer to be in excess of the traffic parameters must be transferred to the waiting room specified by the Scheduler Identifier. In this case the Excess Scheduler Id is not used.

If the Regulator field specifies Shaping, and bit 2 of the Excess Action field is zero, cells will be transferred from the QoS class to the waiting room pointed to by the Scheduler Identifier at a rate defined by the QoS Class Regulator Parameters. In this case the Excess Scheduler Id is not used. If the Regulator field specifies Shaping, and bit 2 of the Excess Action field is set, additional cells will be scheduled for transmission by the waiting room pointed to by the Excess Scheduler Id. This permits a minimum cell rate to be allocated to the QoS class using the QoS Class Regulator Parameters and additional bandwidth

to be shared by the QoS class. The additional share of bandwidth is determined according to the parameters of the waiting room pointed to by the Excess Scheduler Id. If the Excess Scheduler Id is specified in the QoS Class Establishment message, the additional bandwidth will be shared by the entire QoS class. If the Excess Scheduler Id is specified in each individual QoS Connection Management message, the additional bandwidth is specific to that connection and not shared by the entire QoS class. Care must be taken in the implementation to ensure that within each virtual path connection or virtual channel connection, cells depart in the same order that they arrived.

x: Bits 3--7 of the Excess Action field are not used.

QoS Class Weight

If bit 1 of the Scheduler Flags field of the QoS Configuration message indicates that weighted service may be applied to a QoS class, the QoS Class Weight parameter specifies the share of the bandwidth available to the waiting room that should be given to this QoS class.

The QoS Class Weight is an unsigned 16-bit field specifying a binary fraction. I.e. the bandwidth share, as a fraction of the bandwidth available to the waiting room, is given by:

$$\text{Bandwidth share} = \text{QoS Class Weight} * 2^{*(-16)}$$

A QoS Class Weight of zero indicates equal sharing between all QoS classes sharing this waiting room that request a QoS Class Weight of zero. While a 16-bit field is used to specify the QoS Class Weight it is understood that the accuracy of the bandwidth sharing is hardware dependent and is not specified.

If the Regulator field of the QoS Class Establishment message indicates None, or Policer, the QoS Class Weight should be applied to the waiting room pointed to by the Scheduler Identifier. If the Regulator field of the QoS Class Establishment message indicates Shaper, the QoS Class Weight should be applied to the waiting room pointed to by the Excess Scheduler Id.

If the specified waiting room is unable to offer weighted sharing for a QoS class, a failure response message should be returned with the failure code indicating: "This waiting room is unable to offer weighted sharing for a QoS class."

Scheduler Identifier

If all conforming traffic from this QoS class is directed to the same waiting room, on the same output port, this field specifies the Scheduler Identifier for the entire QoS class. The Scheduler Identifier points to the waiting room, on the output port specified by the Output Port field, to which all conforming traffic should be sent. If this field is not used it should be set to 0xFFFF. If each component connection of the QoS class specifies its own output port and waiting room, the Scheduler Identifier must be specified in the QoS Connection Management message and this field must be set to 0xFFFF.

Excess Scheduler Id

If all conforming traffic from this QoS class is directed to the same waiting room, on the same output port, this field specifies the Excess Scheduler Id for the entire QoS class. The Excess Scheduler Id points to the waiting room, on the output port specified by the Output Port field, to which all excess traffic should be sent. If this field is not used it should be set to 0xFFFF. If each component connection of the QoS class specifies its own output port and waiting room, the Excess Scheduler Id must be specified in the QoS Connection Management message and this field must be set to 0xFFFF. If the Scheduler Id is specified in the QoS Class Establishment message, the Excess Scheduler Id must also be specified in the QoS Class Establishment message (or not used). If the Scheduler Id is specified in the QoS Connection Management message, the Excess Scheduler Id must also be specified in the QoS Connection Management message (or not used). The Excess Scheduler Id must not point to the same waiting room on the same output port as the Scheduler Identifier.

Output Port

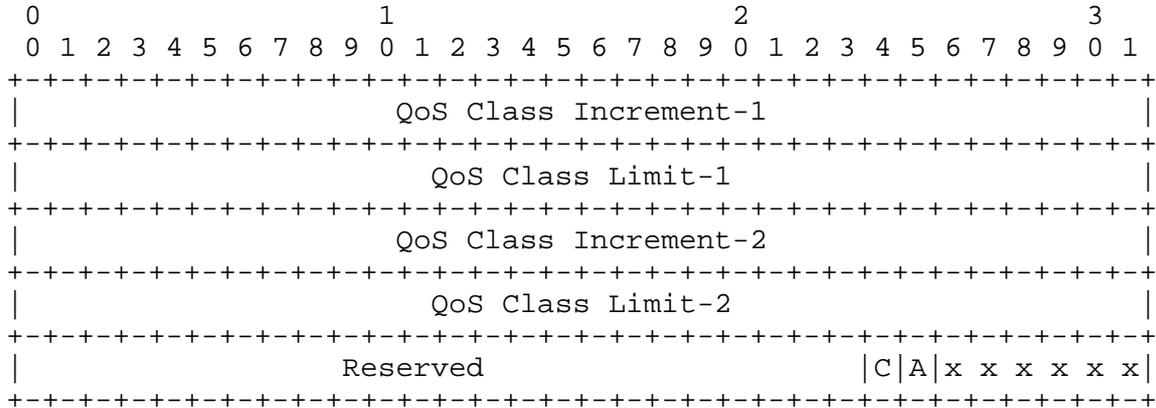
If the Scheduler Identifier field in the QoS Establishment message is not 0xFFFF the Output Port field specifies the Output Port to which traffic from this QoS class should be routed. If the Scheduler Identifier field in the QoS Establishment message is 0xFFFF, this field is not used.

QoS Class Policer Parameters

A policer function may be applied to a QoS class on output from the classifier independently of the regulator function. The QoS class policer function is identical to the connection policer function defined in the QoS

Connection Management message with the exception that it applies to all cells that belong to the QoS class rather than just cells that belong to a single connection.

The QoS Class Policer Parameters have the following format:

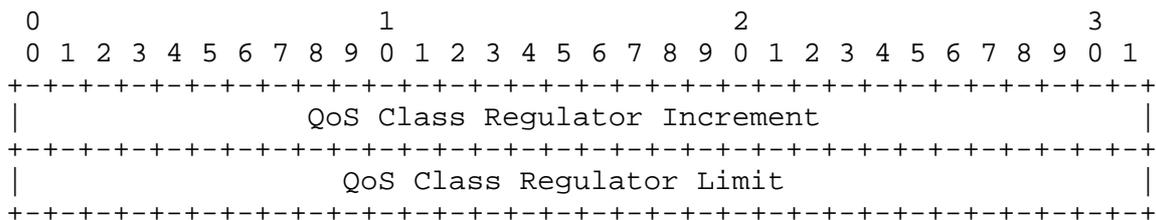


The definition of these fields is given in the UPC Parameters section of the QoS Connection Management message.

QoS Class Regulator Parameters

The QoS class regulator function is identical to the regulator function defined in the QoS Connection Management message with the exception that it applies to all cells that belong to the QoS class rather than just cells that belong to a single connection.

The QoS Class Regulator Parameters have the following format:



The definition of these fields is given in the Regulator Parameters section of the QoS Connection Management message.

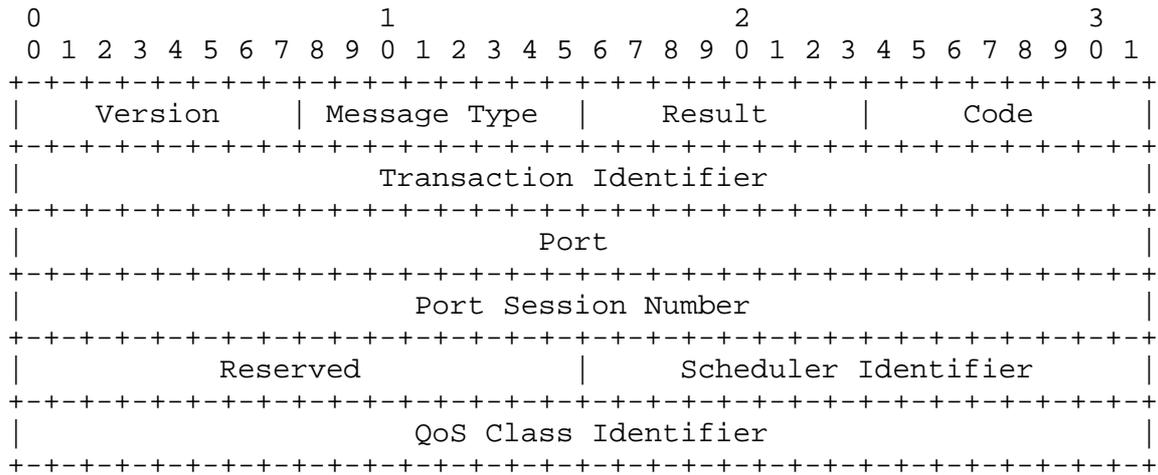
9.5 QoS Release Message

The QoS Release message is used to delete a Scheduler Identifier or a QoS Class Identifier and to release all resources associated with it.

The QoS Release message is:

Message Type = 99

The QoS Release request and success response messages have the following format:



Port

If the QoS Release message contains a Scheduler Identifier, the Port field must contain the Port Number of the switch output port to which the Scheduler Identifier applies. If the QoS Release message contains a QoS Class Identifier, any valid Port number may be used. (The QoS Class Identifier has a global namespace.)

Port Session Number

The current Port Session Number for the port specified in the Port field.

Scheduler Identifier

If the Scheduler Identifier contains the value 0xFFFF the QoS Class Identifier specified in the QoS Class Identifier field should be released. Else, if the value of the Scheduler Identifier lies in the range 0x0100 -- 0xFFFFE inclusive, the Scheduler Identifier specified by the Scheduler Identifier field should be released. A Scheduler

Identifier with a value less than 0x0100 is invalid in a QoS Release message. (It specifies a default value which may not be released.)

QoS Class Identifier

If the Scheduler Identifier contains the value 0xFFFF the QoS Class Identifier field specifies the QoS Class Identifier to be released.

If the QoS Release message requests that a Scheduler Identifier be released, and the Scheduler Identifier is still in use by one or more established connections, a failure response must be returned with the failure code indicating: "Scheduler Identifier still in use." If the QoS Release message requests that a QoS Class Identifier be released, and the QoS Class Identifier is still in use by one or more established connections, a failure response must be returned with the failure code indicating: "QoS Class Identifier still in use."

9.6 QoS Connection Management Message

The QoS Connection Management message is used by the controller to establish and modify virtual channel connections and virtual path connections across the switch which require QoS parameters to be specified. The functionality of the QoS Connection Management message is identical to that of the Add Branch connection management message with the additional specification of QoS parameters. No specific QoS connection release messages are defined. QoS connections may be released with the Delete Tree, Delete All, and Delete Branches messages defined in Section 4, "Connection Management Messages." When a QoS connection is released, all associated QoS resources are released.

There are three styles of connection with specified QoS parameters:

QoS Connection:

This connection style specifies its own individual QoS parameters and is routed independently to the waiting room and output port specified in the QoS Connection Management message. It is not a member of a QoS class. Each output branch of a point-to-multipoint QoS connection may specify its own QoS parameters which may be different from all other output branches of that point-to-multipoint QoS connection, if the switch supports this capability. However, all output branches must specify identical connection policer parameters. A QoS Connection Management message requesting this style of connection is identified by a QoS Class Identifier with the value 0xFFFFFFFF.

QoS Class Connection:

This connection style does not specify its own individual QoS parameters. It is a member of a QoS class, and the QoS parameters are specified by the QoS class. It is, however, routed independently to the waiting room and output port specified in the QoS Connection Management message. Each output branch of a point-to-multipoint QoS Class Connection must use the same QoS parameters. A QoS Connection Management message requesting this style of connection will have a valid QoS Class Identifier and a valid Scheduler Identifier.

QoS Class Member:

This connection style does not specify its own individual QoS parameters. It is a member of a QoS class, and the QoS parameters are specified by the QoS class. The QoS class also specifies the waiting room and output port to which all members of the class are routed. This style of connection does not support point-to-multipoint connections. A QoS Connection Management message requesting this style of connection will have a valid QoS Class Identifier and a Scheduler Identifier with the value 0xFFFF.

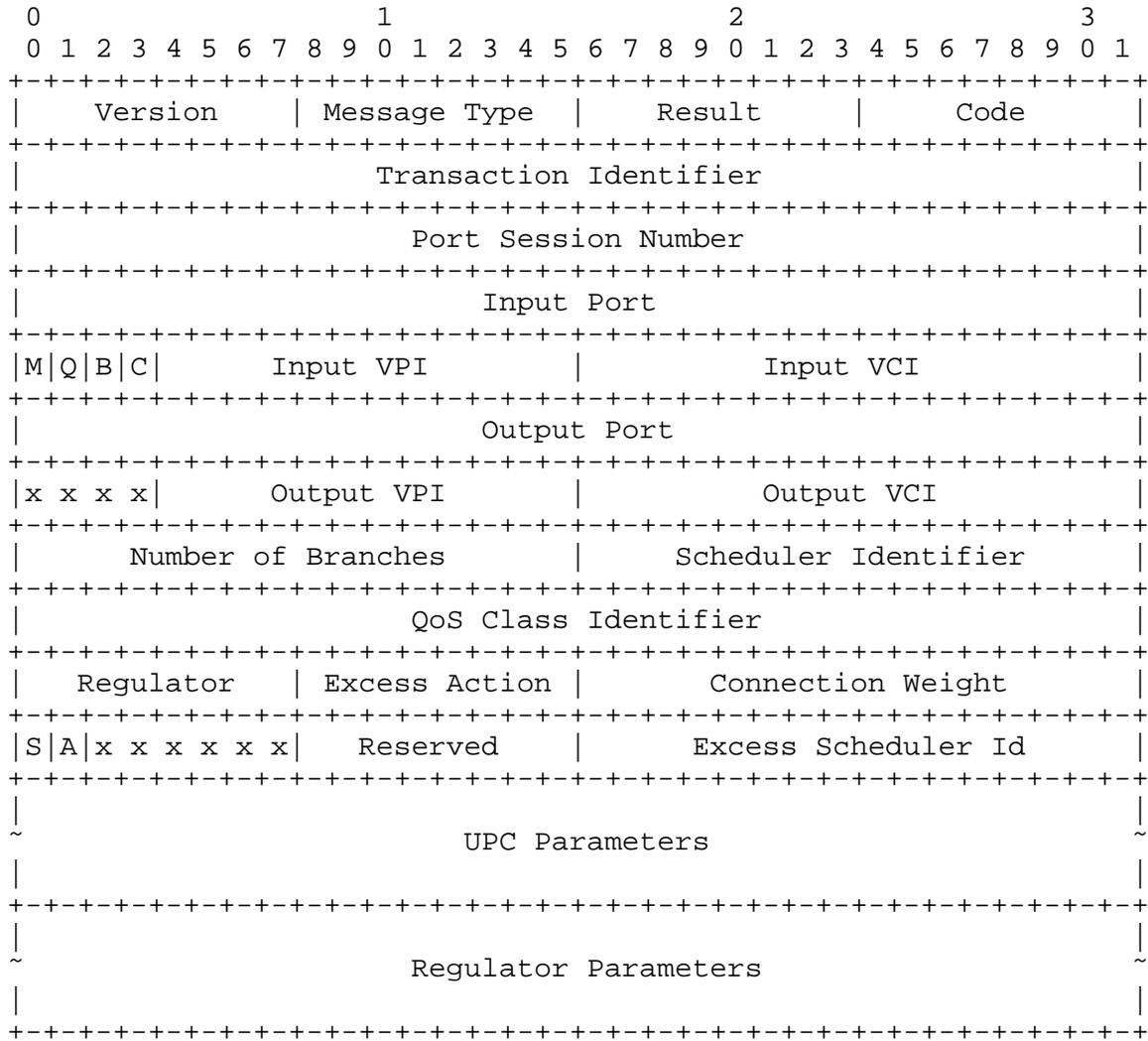
To request a virtual channel connection with specified QoS parameters, the Virtual Channel Connection (VCC) QoS Connection Management message is:

Message Type = 100.

To request a virtual path connection with specified QoS parameters, the Virtual Path Connection (VPC) QoS Connection Management message is:

Message Type = 101.

The QoS Connection Management message has the following format for both request and response messages:



Port Session Number
 Input Port
 Input VPI
 Input VCI
 Output Port
 Output VPI
 Output VCI
 Number of Branches

The definition of these fields is exactly the same as defined for the Add Branch message in Section 4.1, "Connection Management Messages."

M B C Flags

The definition of the M, B, and C flags is exactly the same as defined in Section 4, "Connection Management Messages." They apply to the QoS Connection Management message exactly as defined for the Add Branch message.

Q: QoS Profile Flag The QoS Profile flag is not used in the QoS Connection Management message.

Scheduler Identifier

For QoS Connection and QoS Class Connection styles, the Scheduler Identifier points to the waiting room, on the output port specified by the Output Port field, to which all conforming traffic on the connection should be routed. The values 0 -- 255 specify the default settings for the scheduler. Each of the default values maps directly to one of the scheduler priority levels. A Scheduler Identifier in the range 0 -- 255 may be used without first being established by a Scheduler Establishment message. All Scheduler Identifiers in the range 0x0100 to 0xFFFFE must first be established by a Scheduler Establishment message.

A Scheduler Identifier with a value of 0xFFFF indicates that a QoS Class Member connection style is being requested. In this connection style, the waiting room and output port are specified by reference to the QoS class specified by the QoS Class Identifier field. In this case the QoS Class Identifier field must contain a valid QoS Class Identifier.

QoS Class Identifier

For QoS Class Connection and QoS Class Member connection styles, the QoS Class Identifier specifies the QoS Class to which the connection belongs. It must first be established by a QoS Class Establishment message and must have a value greater than 0 and less than 0xFFFFFFFF.

A QoS Class Identifier with a value of 0xFFFFFFFF indicates that a connection of style "QoS Connection" is being requested. In this connection style, the connection does not belong to a QoS class. All QoS parameters are specified by the QoS Connection Management message and apply only to the specified connection.

Regulator

Excess Action

The Regulator and Excess Action parameters are only used in connection requests of style "QoS Connection." The

definition of these fields in the QoS Connection Management message is exactly the same as defined for the QoS Class Establishment message with the exception that they apply to an individual connection rather than to an entire QoS class.

Connection Weight

This field is only used in connections of style "QoS Connection" and "QoS Class Connection." For QoS Class Member style connections, the QoS Class Weight parameter of the QoS Class Establishment message should be used to assign a weight to the QoS Class.

If bit 0 of the Scheduler Flags field of the QoS Configuration message indicates that weighted service may be applied to a connection, the Connection Weight parameter specifies the share of the bandwidth available to the waiting room that should be given to this connection.

The Connection Weight is an unsigned 16-bit field specifying a binary fraction. I.e. the bandwidth share, as a fraction of the bandwidth available to the waiting room, is given by:

$$\text{Bandwidth share} = \text{Connection Weight} * 2^{(-16)}$$

A Connection Weight of zero indicates equal sharing between all connections in this waiting room that request a Connection Weight of zero. While a 16-bit field is used to specify the Connection Weight it is understood that the accuracy of the bandwidth sharing is hardware dependent and is not specified.

For connections of style "QoS Class Connection," if the Regulator function of the QoS Class is specified as None, or Policer, the Connection Weight should be applied to the waiting room pointed to by the Scheduler Identifier field in the QoS Connection Management message. If the Regulator function of the QoS Class is specified as Shaper, the Connection Weight should be applied to the waiting room pointed to by the Excess Scheduler Id field in the QoS Connection Management message.

For connections of style "QoS Connection," if the Regulator field of the QoS Connection Management message specifies None, or Policer, the Connection Weight should be applied to the waiting room pointed to by the Scheduler Identifier field. If the Regulator field of the QoS Connection

Management message specifies Shaper, the Connection Weight should be applied to the waiting room pointed to by the Excess Scheduler Id field.

If the specified waiting room is unable to offer weighted sharing for a connection, a failure response message should be returned with the failure code indicating: "this waiting room is unable to offer weighted sharing for a connection."

QoS Flags

S: Selective Discard

If the Selective Discard flag is set, only cells with the Cell Loss Priority (CLP) bit set will be subject to the discard mechanism when the number of cells in the waiting room exceeds the Discard Threshold. If the Selective Discard flag is zero, all cells (CLP=0 and CLP=1) will be subject to the discard mechanism when the number of cells in the waiting room exceeds the Discard Threshold. Selective discard can be combined with packet discard. In this case only packets in which at least one cell has the CLP bit set will be subject to the discard mechanism.

A: All Branches

For a QoS Connection Management message that specifies a point-to-multipoint connection, if the All Branches flag is set, all branches of the point-to-multipoint connection must be set to the QoS parameters specified in the message. If the All Branches flag is zero, only the single output branch specified in the message should be set to the QoS parameters specified in the message. For a QoS Connection Management message that specifies a point-to-point connection, the All Branches flag is not used.

x: Unused

Excess Scheduler Id

For connections of style "QoS Connection" and "QoS Class Connection," the Excess Scheduler Id points to the waiting room, on the output port specified by the Output Port field, to which all excess traffic should be routed. The values 0 -- 255 specify the default settings for the scheduler. Each of the default values maps directly to one of the scheduler priority levels. An Excess Scheduler Id in the range 0 -- 255 may be used without first being established by a Scheduler Establishment message. All

values of Excess Scheduler Id in the range 0x0100 to 0xFFFFE must first be established by a Scheduler Establishment message.

If this field is not used it should be set to 0xFFFF. The Excess Scheduler Id must not point to the same waiting room on the same output port as the Scheduler Identifier.

UPC Parameters

All connection styles may be subject to a Usage Parameter Control (UPC) function, also known as a connection policer. The policing function is applied to each individual connection before it is combined with other connections into a QoS class by the classifier function. A policing function applied to an entire QoS class is defined in the QoS Class Establishment message.

The connection policer is defined by reference to the Generic Cell Rate Algorithm (GCRA) defined by the ATM Forum [af-tm-0056], although any equivalent policing algorithm may be used. The GCRA takes two parameters, the increment (I) and the limit (L). The reciprocal of the increment (1/I) specifies the rate being policed. The limit specifies the burst tolerance. (For comparison with the token bucket policer discussed in [Partridge], the size of the token bucket is given by L/I.)

Two policers in series may be specified to permit the policing of both peak rate and average rate (also called sustainable rate). The parameters for the first policer are Increment-1 and Limit-1. For comparison with the ATM Forum specification these would be used to police the Peak Cell Rate (PCR) and Cell Delay Variation Tolerance (CDVT) respectively. The parameters for the second policer are Increment-2 and Limit-2. For comparison with the ATM Forum specification these would be used to police the Sustainable Cell Rate (SCR), and Burst Tolerance. (The Burst Tolerance may be computed from the Maximum Burst Size [af-tm-0056].)

There are two configurations in which the two policers may be connected in series. In the All Cells configuration, all cells (cells with the Cell Loss Priority (CLP) bit set to zero and cells with the CLP bit set to one) are subject to the policing action of both policers in series. In the CLP Selective configuration, all cells, both CLP=0 and CLP=1, are policed by the first policer; but only cells

with CLP=0 are subject to policing by the second policer. Either tagging or discard may be specified for each of the policer configurations.

The values of the parameters Increment and Limit in the UPC Parameters fields are given in terms of a time unit specified by the switch in the QoS Configuration Parameters message. The time unit is specified by the switch as a rate, the Scaling Factor, which gives the rate in cells per second that would result from an Increment parameter value of one. Thus to determine the value of the Increment parameter from the desired policed rate given in cells per second:

$$\text{Increment parameter} = (\text{Scaling_Factor}) / (\text{policed_rate})$$

To determine the value of the Limit parameter from the desired Cell Delay Variation Tolerance (CDVT) given in seconds:

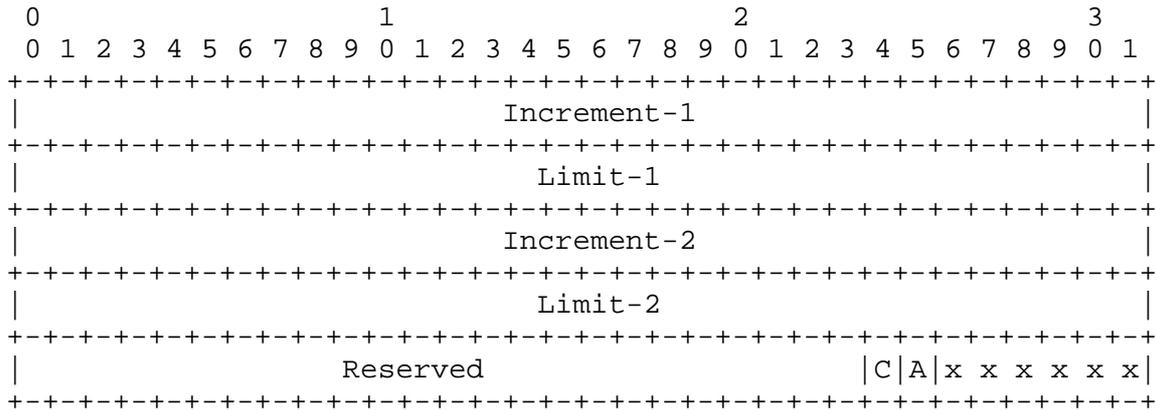
$$\text{Limit parameter} = \text{CDVT} * \text{Scaling_Factor}$$

To determine the value of the Limit parameter from the desired Burst Tolerance (BT) given in seconds:

$$\text{Limit parameter} = \text{BT} * \text{Scaling_Factor}$$

The Increment and Limit parameters are specified as 32-bit unsigned integers; so the choice of the Scaling Factor allows the switch to select the range and granularity of the policer parameters with respect to the line rate of the switch port. For example, a SONET STS-3c (155.52 Mbps) switch port has a line rate of approximately 353 kcells/s. With a Scaling Factor value of 353,000,000 we can specify a policed rate slightly less than the line rate with a granularity of 0.1%. For a policed rate of 1 kbps we can still support a bucket size of 31 cells.

The UPC Parameters have the following format:



Increment-1

The increment parameter for the first policer, specified as a 32-bit unsigned integer. A value of zero for the Increment-1 parameter is used to disable the first policer. In this case all cells will be considered to conform to the traffic parameters of the first policer.

Limit-1

The limit parameter for the first policer, specified as a 32-bit unsigned integer.

Increment-2

The increment parameter for the second policer, specified as a 32-bit unsigned integer. A value of zero for the Increment-2 parameter is used to disable the second policer. In this case all cells will be considered to conform to the traffic parameters of the second policer.

Limit-2

The limit parameter for the second policer, specified as a 32-bit unsigned integer.

Flags

C: Configuration

If the Configuration flag is set, the policer should be set to the All Cells configuration. If the Configuration flag is zero, the policer should be set to the CLP Selective configuration.

In the All Cells configuration, all cells (both CLP=0 and CLP=1) are subject to the policing action of both policers in series. In the CLP Selective configuration, all cells, both CLP=0 and CLP=1, are policed by the first policer; but

only cells with CLP=0 are subject to policing by the second policer. Either tagging or discard may be specified for each of the policer configurations.

A: Action

If the Action flag is zero, discard is required as the policing action. If the Action flag is set, tagging is required as the policing action.

If tagging is selected in the All Cells configuration, any cell with CLP=0 in either policer, that the policer determines to be in excess of the specified policer parameters, will be changed to CLP=1. If discard is selected in the All Cells configuration, any cell (CLP=0 or CLP=1) in either policer, that the policer determines to be in excess of the specified policer parameters, will be discarded.

In the CLP Selective configuration, the first policer is always set to discard any cell (CLP=0 or CLP=1) that it determines to be in excess of its specified policer parameters. If tagging is selected in the CLP Selective configuration, the second policer will change the CLP bit to CLP=1 of any cell that it determines to be in excess of its specified parameters. If discard is selected in the CLP Selective configuration, the second policer will discard any cell that it determines to be in excess of its specified parameters.

To configure the policer for the conformance definitions specified by the ATM Forum [af-tm-0056] the following configurations are suggested:

CBR.1:	One policer,	All Cells,	Discard
VBR.1:	Two policers,	All Cells,	Discard
VBR.2:	Two policers,	CLP Selective,	Discard
VBR.3:	Two policers,	CLP Selective,	Tagging
UBR.1:	One policer,	All Cells,	Discard
UBR.2:	One policer,	All Cells,	Tagging.

x: Unused

Regulator Parameters

Only connections of style "QoS Connection" require the Regulator Parameters to be specified in the QoS Connection Management message. For connections of style "QoS Class Connection" and "QoS Class Member" the Regulator Parameters are specified in the QoS Class Establishment message.

The Regulator Parameters are specified in a similar manner to the UPC parameters. If the regulator function is specified as Policing, a single GCRA policer is applied to all cells (both CLP=0 and CLP=1) on the connection. The policer takes two parameters: an increment, the Regulator Increment, and a limit, the Regulator Limit. The reciprocal of the increment ($1/I$) specifies the rate being policed. The limit (L) specifies the burst tolerance. (For comparison with the token bucket policer discussed in [Partridge], the size of the token bucket is given by L/I .)

The Regulator Increment and Regulator Limit parameters are 32-bit unsigned integers. Their values are determined in terms of the Scaling Factor specified by the switch in the QoS Configuration Parameters message. To determine the value of the Regulator Increment parameter from the desired policed rate given in cells per second:

$$\text{Regulator Increment} = (\text{Scaling_Factor}) / (\text{policed_rate})$$

For a policed rate (r) the GCRA policer guarantees that over any time period T the amount of traffic determined by the policer to be conforming to the traffic parameters does not exceed:

$$rT + L/I$$

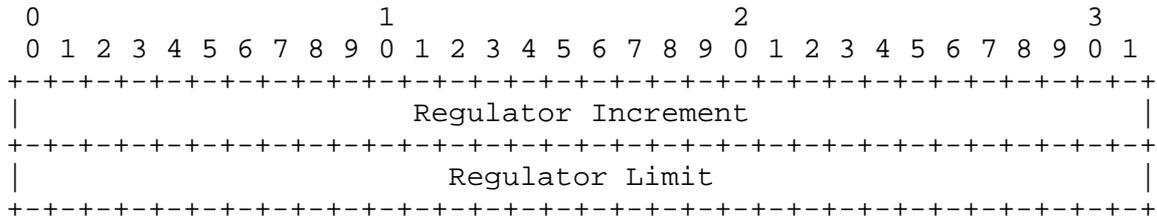
The value of the Regulator Limit may be determined from this relation.

If the regulator function is specified as Shaping, only the Regulator Increment parameter is used. The Regulator Limit parameter is not used. The value of the Regulator Increment parameter is determined in terms of the Scaling Factor specified by the switch in the QoS Configuration Parameters message. To determine the value of the Regulator Increment parameter from the desired shaper rate, given in cells per second, on output from the shaper:

$$\text{Regulator Increment} = (\text{Scaling_Factor}) / (\text{shaper_rate})$$

An Increment value of zero is used to disable the policer. In this case all cells on that connection will be considered to conform to the policer traffic parameters. A shaper given a Regulator Increment parameter of zero will perform no shaping function on that connection.

The Regulator Parameters have the following format:



9.7 QoS Failure Response Codes

A failure response message is formed by returning the request message that caused the failure with the Result field in the header indicating failure (Result = 4) and the Code field giving the failure code. The failure code specifies the reason for the switch being unable to satisfy the request message. The following additional failure codes are defined for use in response to QoS messages. General failure codes are specified in Section 3.2, Failure Response Messages.

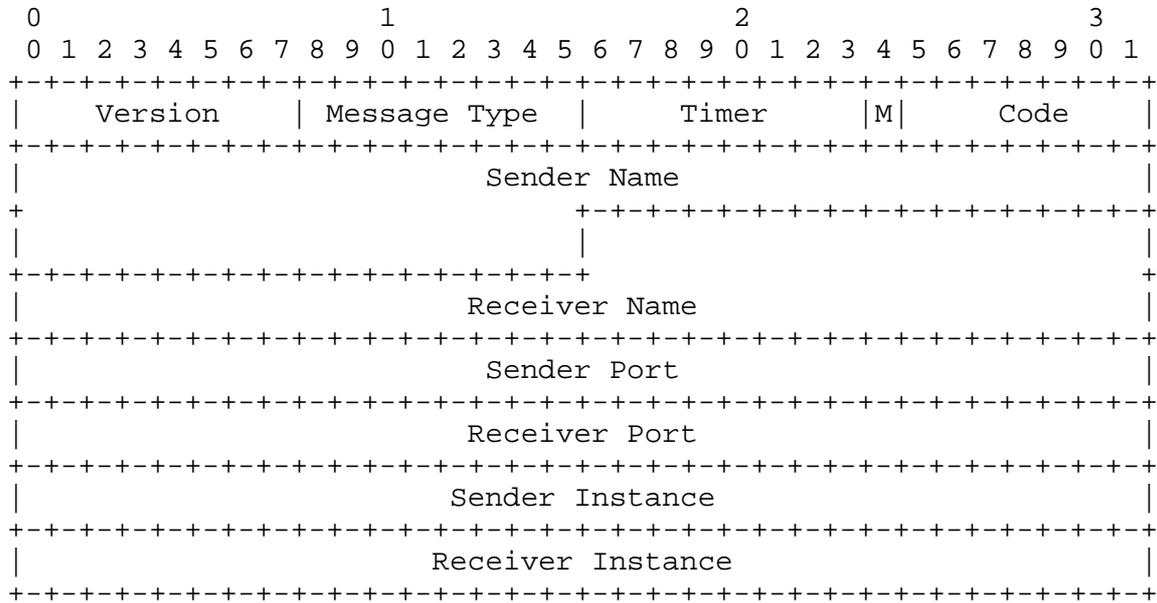
- 128: Weighted scheduling within this waiting room is unavailable.
- 129: This waiting room is unable to offer weighted sharing for a QoS class.
- 130: This waiting room is unable to offer weighted sharing for a connection.
- 131: Scheduler Identifier still in use.
- 132: QoS Class Identifier still in use.
- 133: Invalid QoS parameter.
- 134: Insufficient QoS resources.
- 135: Any point-to-multipoint connection arriving on this input port must use the same QoS parameters for all output branches.

10. Adjacency Protocol

The adjacency protocol is used to synchronize state across the link, to agree on which version of the protocol to use, to discover the identity of the entity at the other end of a link, and to detect when it changes. GSMP is a hard state protocol. It is therefore important to detect loss of contact between switch and controller, and to detect any change of identity of switch or controller. No GSMP messages other than those of the adjacency protocol may be sent across the link until the adjacency protocol has achieved synchronization.

10.1 Packet Format

All GSMP messages belonging to the adjacency protocol have the following structure:



Version

In the adjacency protocol the Version field is used for version negotiation. In a SYN message the Version field always contains the highest version understood by the sender. A receiver receiving a SYN message with a version higher than understood will ignore that message. A receiver receiving a SYN message with a version lower than its own highest version, but a version that it understands, will reply with a SYNACK with the version from the received SYN in its GSMP Version field. This defines the version of the GSMP protocol to be used while the adjacency protocol remains synchronized. All other messages will use the agreed version in the Version field.

The version number for the version of the GSMP protocol defined by this specification is Version = 2.

Message Type

The adjacency protocol is:

Message Type = 10

Timer

The Timer field is used to inform the receiver of the timer value used in the adjacency protocol of the sender. The timer specifies the nominal time between periodic adjacency protocol messages. It is a constant for the duration of a GSMP session. The timer field is specified in units of 100ms.

M-Flag

The M-Flag is used in the SYN message to indicate whether the sender is a master or a slave. If the M-Flag is set in the SYN message, the sender is a master. If zero, the sender is a slave. The GSMP protocol is asymmetric, the controller being the master and the switch being the slave. The M-Flag prevents a master from synchronizing with another master, or a slave with another slave. If a slave receives a SYN message with a zero M-Flag, it must ignore that SYN message. If a master receives a SYN message with the M-Flag set, it must ignore that SYN message. In all other messages the M-Flag is not used.

Code

Field specifies the function of the message. Four Codes are defined for the adjacency protocol:

```
SYN:      Code = 1
SYNACK:   Code = 2
ACK:      Code = 3
RSTACK:   Code = 4.
```

Sender Name

For the SYN, SYNACK, and ACK messages, is the name of the entity sending the message. The Sender Name is a 48-bit quantity that is unique within the operational context of the device. A 48-bit IEEE 802 MAC address, if available, may be used for the Sender Name. If the Ethernet encapsulation is used the Sender Name must be the Source Address from the MAC header. For the RSTACK message, the Sender Name field is set to the value of the Receiver Name field from the incoming message that caused the RSTACK message to be generated.

Receiver Name

For the SYN, SYNACK, and ACK messages, is the name of the entity that the sender of the message believes is at the far end of the link. If the sender of the message does not know the name of the entity at the far end of the link, this field should be set to zero. For the RSTACK message,

the Receiver Name field is set to the value of the Sender Name field from the incoming message that caused the RSTACK message to be generated.

Sender Port

For the SYN, SYNACK, and ACK messages, is the local port number of the link across which the message is being sent. For the RSTACK message, the Sender Port field is set to the value of the Receiver Port field from the incoming message that caused the RSTACK message to be generated.

Receiver Port

For the SYN, SYNACK, and ACK messages, is what the sender believes is the local port number for the link, allocated by the entity at the far end of the link. If the sender of the message does not know the port number at the far end of the link, this field should be set to zero. For the RSTACK message, the Receiver Port field is set to the value of the Sender Port field from the incoming message that caused the RSTACK message to be generated.

Sender Instance

For the SYN, SYNACK, and ACK messages, is the sender's instance number for the link. It is used to detect when the link comes back up after going down or when the identity of the entity at the other end of the link changes. The instance number is a 32-bit number that is guaranteed to be unique within the recent past and to change when the link or node comes back up after going down. Zero is not a valid instance number. For the RSTACK message, the Sender Instance field is set to the value of the Receiver Instance field from the incoming message that caused the RSTACK message to be generated.

Receiver Instance

For the SYN, SYNACK, and ACK messages, is what the sender believes is the current instance number for the link, allocated by the entity at the far end of the link. If the sender of the message does not know the current instance number at the far end of the link, this field should be set to zero. For the RSTACK message, the Receiver Instance field is set to the value of the Sender Instance field from the incoming message that caused the RSTACK message to be generated.

10.2 Procedure

The adjacency protocol is described by the following rules and state tables.

The rules and state tables use the following operations:

- o The "Update Peer Verifier" operation is defined as storing the values of the Sender Instance, Sender Port, and Sender Name fields from a SYN or SYNACK message received from the entity at the far end of the link.
 - o The procedure "Reset the link" is defined as:
 1. Generate a new instance number for the link
 2. Delete the peer verifier (set to zero the values of Sender Instance, Sender Port, and Sender Name previously stored by the Update Peer Verifier operation)
 3. Send a SYN message
 4. Enter the SYNSENT state.
 - o The state tables use the following Boolean terms and operators:
 - A The Sender Instance in the incoming message matches the value stored from a previous message by the "Update Peer Verifier" operation.
 - B The Sender Instance, Sender Port, and Sender Name fields in the incoming message match the values stored from a previous message by the "Update Peer Verifier" operation.
 - C The Receiver Instance, Receiver Port, and Receiver Name fields in the incoming message match the values of the Sender Instance, Sender Port, and Sender Name currently sent in outgoing SYN, SYNACK, and ACK messages.
- "&&" Represents the logical AND operation
- "||" Represents the logical OR operation
- "!" Represents the logical negation (NOT) operation.
- o A timer is required for the periodic generation of SYN, SYNACK, and ACK messages. The value of the timer is announced in the Timer field. The period of the timer is unspecified but a value of one second is suggested.

There are two independent events: the timer expires, and a packet arrives. The processing rules for these events are:

```

Timer Expires:   Reset Timer
                  If state = SYNSENT Send SYN
                  If state = SYNRCVD Send SYNACK
                  If state = ESTAB   Send ACK

```

```

Packet Arrives:
  If incoming message is an RSTACK:
    If (A && C && !SYNSENT) Reset the link
    Else Discard the message.
  If incoming message is a SYN, SYNACK, or ACK:
    Response defined by the following State Tables.
  If incoming message is any other GSMP message and state !=
    ESTAB:
    Discard incoming message.
    If state = SYNSENT Send SYN (Note 1)
    If state = SYNRCVD Send SYNACK (Note 1)

```

Note 1: No more than two SYN or SYNACK messages should be sent within any time period of length defined by the timer.

- o State synchronization across a link is considered to be achieved when the protocol reaches the ESTAB state. All GSMP messages, other than adjacency protocol messages, that are received before synchronization is achieved will be discarded.

State Tables

State: SYNSENT

Condition	Action	New State
SYNACK && C	Update Peer Verifier; Send ACK	ESTAB
SYNACK && !C	Send RSTACK	SYNSENT
SYN	Update Peer Verifier; Send SYNACK	SYNRCVD
ACK	Send RSTACK	SYNSENT

State: SYNRCVD

Condition	Action	New State
SYNACK && C	Update Peer Verifier; Send ACK	ESTAB
SYNACK && !C	Send RSTACK	SYNRCVD
SYN	Update Peer Verifier; Send SYNACK	SYNRCVD
ACK && B && C	Send ACK	ESTAB
ACK && !(B && C)	Send RSTACK	SYNRCVD

State: ESTAB

Condition	Action	New State
SYN SYNACK	Send ACK (note 2)	ESTAB
ACK && B && C	Send ACK (note 3)	ESTAB
ACK && !(B && C)	Send RSTACK	ESTAB

Note 2: No more than two ACKs should be sent within any time period of length defined by the timer. Thus, one ACK must be sent every time the timer expires. In addition, one further ACK may be sent between timer expirations if the incoming message is a SYN or SYNACK. This additional ACK allows the adjacency protocol to reach synchronization more quickly.

Note 3: No more than one ACK should be sent within any time period of length defined by the timer.

10.3 Loss of Synchronization

If after synchronization is achieved, no valid GSMP messages are received in any period of time in excess of three times the value of the Timer field announced in the incoming adjacency protocol messages, loss of synchronization may be declared.

The preferred procedure for a switch to use when it loses synchronization with its active controller is to attempt to establish

synchronization with (one of) its backup controller(s). However, in this preferred approach, it must not reset its state until it achieves synchronization with a backup controller. This means that if, before achieving synchronization with a backup controller, it regains synchronization with its original controller, it may continue the original session (and cease attempting to establish synchronization with a backup controller). If synchronization with the original session is regained it is the responsibility of the controller to ensure consistent state between the switch and controller.

While the above is the preferred procedure, it is also the case that the simplest procedure when declaring loss of synchronization with the active controller is to reset the switch state, and start searching for a controller. This simple procedure is legitimate.

11. Summary of Failure Response Codes

The following list gives a summary of the failure codes defined for failure response messages:

- 1: Unspecified reason not covered by other failure codes.
- 2: Invalid request message.
- 3: The specified request is not implemented on this switch.
- 4: Invalid Port Session Number.
- 5: One or more of the specified ports does not exist.
- 6: One or more of the specified ports is down.
- 7: Unused. (This failure code has been replaced by failure codes 18--21.)
- 8: The specified connection does not exist.
- 9: The specified branch does not exist.
- 10: A branch belonging to the specified point-to-multipoint connection is already established on the specified output port and the switch cannot support more than a single branch of any point-to-multipoint connection on the same output port.
- 11: The limit on the maximum number of point-to-multipoint connections that the switch can support has been reached.
- 12: The limit on the maximum number of branches that the specified point-to-multipoint connection can support has been reached.
- 13: Unable to assign the requested VPI/VCI value to the requested branch on the specified point-to-multipoint connection.
- 14: General problem related to the manner in which point-to-multipoint is supported by the switch.
- 15: Out of resources (e.g. memory exhausted, etc.).
- 16: Failure specific to the particular message type. (The meaning of this failure code depends upon the Message Type. It is

- defined within the description of any message that uses it.)
- 17: Cannot label each output branch of a point-to-multipoint tree with a different label.
 - 18: One or more of the specified input VPIs is invalid.
 - 19: One or more of the specified input VCIs is invalid.
 - 20: One or more of the specified output VPIs is invalid.
 - 21: One or more of the specified output VCIs is invalid.
 - 22: Invalid Class of Service field in a Connection Management message.
 - 23: Insufficient resources for QoS Profile.
 - 24: Virtual path switching is not supported on this input port.
 - 25: Point-to-multipoint virtual path connections are not supported on either the requested input port or the requested output port.
 - 26: Attempt to add a virtual path connection branch to an existing virtual channel connection.
 - 27: Attempt to add a virtual channel connection branch to an existing virtual path connection.
 - 28: Only point-to-point bidirectional connections may be established.
 - 29: Cannot support requested VPI range.
 - 30: Cannot support requested VCI range on all requested VPIs.
 - 31: The transmit cell rate of this output port cannot be changed.
 - 32: Requested transmit cell rate out of range for this output port.
 - 128: Weighted scheduling within this waiting room is unavailable.
 - 129: This waiting room is unable to offer weighted sharing for a QoS class.
 - 130: This waiting room is unable to offer weighted sharing for a connection.
 - 131: Scheduler Identifier still in use.
 - 132: QoS Class Identifier still in use.
 - 133: Invalid QoS parameter.
 - 134: Insufficient QoS resources.
 - 135: Any point-to-multipoint connection arriving on this input port must use the same QoS parameters for all output branches.

12. Summary of Message Set

The following table gives a summary of the messages defined in this version of the specification. It also indicates which messages must be supported in a minimal implementation of the protocol. Those messages marked as "Required" must be supported by the switch for an implementation to be considered to conform to this specification. (While the controller should also implement those messages marked

"Required," conformance cannot be tested for the controller due to the Master-Slave nature of the protocol.)

Message Name	Message Type	Status
Connection Management Messages		
Add Branch VCC.....	16	Required
VPC.....	26	
Delete Tree.....	18	Required
Delete All.....	20	
Delete Branches.....	17	
Move Branch VCC.....	22	
VPC.....	27	
Port Management Messages		
Port Management.....	32	Required
Label Range.....	33	
State and Statistics Messages		
Connection Activity.....	48	Required
Port Statistics.....	49	
Connection Statistics.....	50	
QoS Class Statistics.....	51	
Report Connection State.....	52	
Configuration Messages		
Switch Configuration.....	64	Required
Port Configuration.....	65	Required
All Ports Configuration.....	66	Required
Event Messages		
Port Up.....	80	
Port Down.....	81	
Invalid VPI/VCI.....	82	
New Port.....	83	
Dead Port.....	84	
Quality of Service Messages		
QoS Configuration.....	96	
Scheduler Establishment.....	97	
QoS Class Establishment.....	98	
QoS Release.....	99	
QoS Connection Management VCC....	100	
VPC....	101	
Adjacency Protocol.....	10	Required

REFERENCES

- [af-tm-0056] ATM Forum Traffic Management Specification 4.0, af-tm-0056.000, April 1996.
- [I.361] "B-ISDN ATM Layer Specification," International Telecommunication Union, ITU-T Recommendation I.361, Mar. 1993.
- [I.363] "B-ISDN ATM Adaptation Layer (AAL) Specification," International Telecommunication Union, ITU-T Recommendation I.363, Mar. 1993.
- [IpsilonMIB] Ipsilon IP Switch MIB,
<http://www.ipsilon.com/products/ips.mib>
- [Partridge] C. Partridge, "Gigabit Networking," Addison-Wesley, 1994.
- [RFC1700] Reynolds, J., and J. Postel, "Assigned Numbers," STD 2, RFC 1700, October 1994.
- [RFC1987] Newman, P, Edwards, W., hinden, R., Hoffman, E. Ching Liaw, F., Lyon, T. and G. Minshall, "Ipsilon's General Switch Management Protocol Specification," Version 1.1, RFC 1987, August 1996.

SECURITY CONSIDERATIONS

Physical security on the control link connecting the controller to the switch is assumed. Security issues are not discussed in this document.

AUTHORS' ADDRESSES

Peter Newman Nokia	Phone: +1 (408) 990 2003 EMail: pn@ipsilon.com
W. L. Edwards, Chief Scientist Sprint	Phone: +1 (913) 534 5334 EMail: texas@sprintcorp.com
Robert M. Hinden Nokia	Phone: +1 (408) 990 2004 EMail: hinden@ipsilon.com
Eric Hoffman Nokia	Phone: +1 (408) 990 2010 EMail: hoffman@ipsilon.com

Fong Ching Liaw
Coppercom

Phone: +1 (408) 873 2688
EMail: fong@coppercom.com

Tom Lyon
Nokia

Phone: +1 (408) 990 2001
EMail: pugs@ipsilon.com

Greg Minshall
Fiberlane Communications

Phone: +1 (650) 237 3164
EMail: minshall@fiberlane.com

Nokia (Sunnyvale) is located at:

232 Java Drive
Sunnyvale, CA 94089
USA

Sprint is located at:

Sprint
Sprint Technology Services - Long Distance Division
9300 Metcalf Avenue
Mailstop KSOPKB0802
Overland Park, KS 66212-6333
USA

Fiberlane Communications is located at:

1399 Charleston Road
Mountain View, CA 94043
USA

Full Copyright Statement

Copyright (C) The Internet Society (1998). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

