Recommendation Y.2121 (formerly Y.flowreq)

Requirements for the support of flow state aware transport technology in an NGN

Summary

This Recommendation specifies the requirements for the support of the flow state aware (FSA) transfer capability in a Next Generation Network (NGN). The FSA transfer capability provides QoS controls that operate on a per-flow basis, allowing flows to receive different treatment depending on signalled parameters. These parameters are requested using in-band signalling. The parameters contained in these signals are included in the "flow state" maintained on each flow at each FSA node.

Service options that may be selected include requested support of the highest available end-to-end (or FSA edge-to-edge) rate for data transfer. Another option is immediate transmission, wherein a flow may start or assume a new rate immediately on the understanding that the network will provide a guaranteed rate as soon as possible. This will be provided when network resources permit. Yet another option is for a negotiated guaranteed rate. These services are targeted at access scenarios where media flows may result in temporary congestion and where best effort would not act selectively on the last few flows that had contributed to the onset of congestion. These services may also be applied to flow aggregates, providing the possibility of highest available rate between the aggregation end-points or the option of supporting immediate aggregate rate changes that act in conjunction with per-flow controls.

Table of Contents

1. Scope

This document provides Flow State Aware requirements in support of per-flow service options providing for edge-to-edge QoS and transport resource control (including resource reservation and admission control), in Next Generation Networks (NGNs). The pertinent protocol specifications and measurement requirements will be described in separate Recommendations. Note that network management functionality is outside the scope of this Recommendation. Administrations may require operators and service providers to take into account national regulatory and national policy requirements in implementing this Recommendation.

2. References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.1221]	ITU-T Recommendation Y.1221 Amendment 2 (2005), *Traffic control and congestion control in IP-based networks*
[ITU-T Y.2012]	ITU-T Recommendation Y.2012 (2006), *Functional requirements and architecture of the NGN*
[ITU-T Y.2111]	ITU-T Recommendation Y.2111 (2006), *Resource and Admission Control Functions in Next Generation Networks*

3. Definitions

This Recommendation uses the following terms

3.1 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.1.1. Aggregation End-point. An end-point within the network which attaches or deletes the common Flow Aggregate Identifiers to ensure commencement of/ cessation of common routing and QoS treatment of packets. This end-point also initiates/ terminates in-band signalling to control Flow State information retained for treatment of the flow aggregate.

3.1.2. Available Rate (AR). A FSA (Flow State Aware) transport service primarily for applications that can flexibly adapt to the current available capacity and can quickly adjust their sending rate as the available capacity changes.

3.1.3. Flow. A unidirectional sequence of packets with the property that, along any given network link, a Flow Identifier has the same value for every packet.

3.1.4. Flow Aggregate. A hierarchical flow construct that is associated with a group of flows. The carried flows may extend beyond the flow aggregate. Except for the end nodes, flow aggregate forwarders in general do not know that they are carrying flows within the flow aggregate. All

packets belonging to a given flow aggregate are commonly routed between Aggregation End-points.

3.1.5. Flow admission control. The determination, for authorised requests, of whether or not to accept a given flow.

3.1.6. Flow Identifier. A vector comprising the values of a number of elements taken from the IP, TCP/UDP header fields, encapsulation header, and label fields attached to a packet. The Flow Identifier for a flow within a single FSA network is unique. Appendix II, II.10 describes examples of suitable identifiers.

3.1.7. Flow State. A set of values stored per flow identifier at each Flow State Aware node. This set of values will determine controls applied on a per flow basis, dealing with forwarding rate, delay, and congestion recovery.

3.1.8. Flow State Aware Node. A network node that is capable of maintaining Flow State and applying per-flow QoS controls, based on recognising Flow Identifier and associated signals.

3.1.9. **Flow State Aware Signalling Edge Function.** A function that provides the origin and/ or termination of the Flow State Aware end-to-end signalling path, and participates in requests and responses on behalf of a user application or management action. It may be located, for example, in a user end-system or at a network edge node where it serves as the signalling end-point of multiple users and associated applications. Alternatively, it may be located at an Aggregation End-point where it supports the signalling requirements of flow aggregates.

3.1.10. **Guaranteed Rate (GR).** A FSA transport service for applications that require guaranteed bandwidth for the duration of the flow.

3.1.11. In-band signalling. A mode of signalling where the signalling messages follow a path that is tied to the data packets. Signalling messages are routed only through nodes that are in the data path.

3.1.12. **Maximum Rate (MR).** A FSA transport service for applications that want packet loss characteristics to be sufficient for streamed services as soon as possible but are unwilling to wait or be rejected by network admission control, if network resource for this target QoS is not available immediately.

3.1.13. **Out-of-band signalling**. A mode of signalling where the signalling messages follow a different path to the data packets and are routed to one or more nodes that are not in the data path.

3.1.14. **Preference Priority:** A parameter used to determine whether to admit a flow in case of network overload. In network overload state, flow with the lower preference priority may be rejected while the one with higher preference priority level still admitted.

3.1.15. **Qos Structure.** The block of QoS signalling information in a signalling packet

3.1.16. **Variable Rate (VR).** A FSA transport service for applications that want an AR service and can flexibly take advantage of additional available capacity, but with a minimum capacity consistent with MR transport characteristics, allowing the option of immediate transmission at or above this minimum rate.

4. Abbreviations and Acronyms

This Recommendation uses the following abbreviations:

AAA Authentication, Authorisation and Accounting

ACK	TCP response: acknowledgement number is valid
ANF	Access Node Function
ABR	Available Bit Rate
AR	Available Rate
ATM	Asynchronous Transfer Mode
BA	Behaviour Aggregate
BRAS	Broadband Remote Access Server
CPE	Customer Premises Equipment
Diffserv	Differentiated Services
DSL	Digital Subscriber Line
DOS	Denial of Service
eFSA	egress FSA
FR	Fixed Rate
FSA	Flow State Aware
GR	Guaranteed Rate
GRE	Generic Routing Encapsulation
iFSA	Ingress FSA
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
L2TP	Layer 2 Tunnelling Protocol
LAC	L2TP Access Concentrator
LNS	L2TP Network Server
LSP	Label Switched Path
MPLS	Multi-Protocol Label Switching
MR	Maximum Rate
NACF	Network Attachment Control Functions
NAT	Network Address Translator
NGN	Next Generation Network
NNI	Network-Network Interface
NR	Network Rate
PDA	Personal Digital assistant
PD-FE	Policy Decision Functional Entity
PE-FE	Policy Enforcement Functional Entity
PPP	Point-to-Point Protocol
PT	Payload Type
QoS	Quality of Service
PHB	Per Hop Behaviour
RACF	Resource and Admission Control Functions
RM	Resource Management
RTP	Real-time Transport Protocol
SCF	Service Control Functions
SIP	Session Initiation Protocol
SYN	TCP connection set-up request
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UNI	User-Network Interface
VP	Virtual Path
VR	Variable Rate

5. Overview

To meet specific network performance requirements, a network operator needs to implement capabilities such as those specified in [ITU-T Y.1221], which now includes clause 6.5 describing a Conditionally Dedicated Bandwidth transfer capability.

To implement this transfer capability as defined in [ITU-T Y.1221], a network needs to provide specific user plane functionality at the UNI and NNI. This Recommendation specifies the Flow State Aware (FSA) requirements for such functionality.

Appendix V provides example implementation principles associated with FSA nodes.

In terms of QoS aspects of Flow State Aware transfer, a network may be provisioned to meet performance requirements either statically or dynamically. The provisioning would be applied on a per flow basis. This Recommendation defines the requirements for such provisioning.

Static network provisioning is typically performed by a network management system. Static provisioning takes into account both overall network performance requirements and performance requirements for individual customers based on traffic contracts between the customer and the network operator.

Dynamic network provisioning at a UNI and/or NNI node allows the ability to dynamically request a traffic contract for an IP flow (as defined in [ITU-T Y.1221]) from a specific source node to one or more destination nodes. In response to the request, the network determines if resources are available to satisfy the request and provision the network. Clause 6.5 of [ITU-T Y.1221] describes the case of dynamic provisioning within a Flow State Aware network, initiated by a Flow State Aware source node.

QoS requirements (as would be applied to services supported by Flow State Aware transfer) go beyond just the delay and loss that can occur in the transport of IP packets. The requirements include:

– bandwidth/capacity needed by the application, and

– the priority with which such bandwidth will be maintained during congestion and with which it will be restored after various failure events.

To achieve the required QoS for FSA transfer, networks must incorporate the following functions:

1) Functions supporting the FSA Packet forwarding behaviours that are applied per flow.

2) Flow admission control recognising and processing requests for associated FSA transport services.

3) Functions supporting the signalling for allocating necessary resources for each flow.

Figure 1 shows the main functions which are involved in establishing and ceasing FSA transfer and ensuring the correct provisioning of resources to meet QoS objectives.

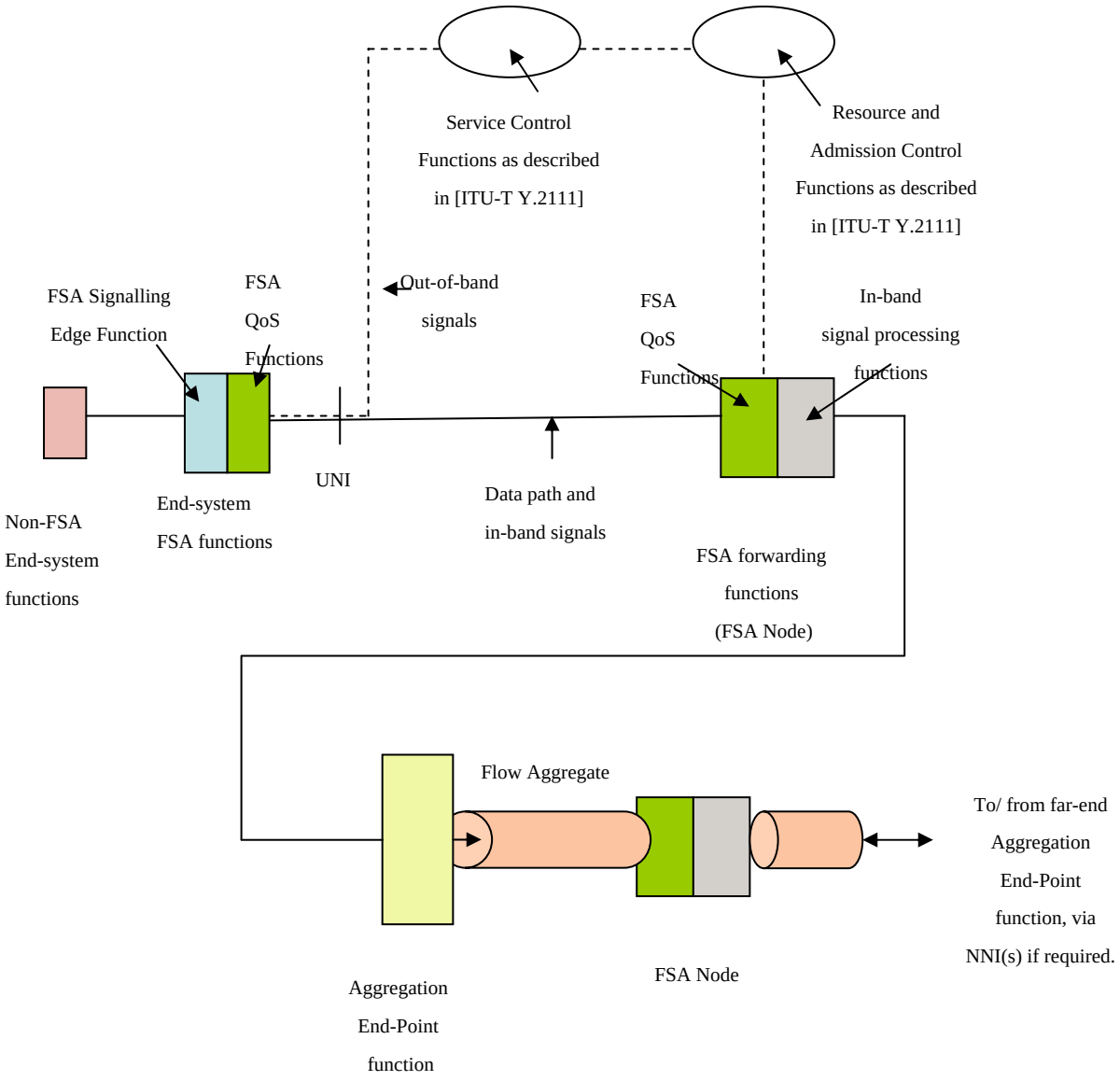


Figure 1 Overview of FSA functions

The requirements in this Recommendation cover the following areas:

- Requirements associated with Flow State Aware dynamic provisioning requests initiated by a Flow State Aware Signalling Edge Function, for example in user end-system. See 6.1.
- Resource modification. See 6.1.
- Response to Flow State Aware requests. See 6.2.
- iFSA Function answer to network response to Flow State Aware requests. See 6.2.
- Form of a verifiable Flow State Aware request. See 6.3
- Request performance requirements (e.g. negotiation delays). See 6.3.
- Release of resources no longer required. See 6.3.
- Error reporting. See 6.3.

- Preference resolution. See 6.4.
- Parameters and values for transport connections. See 6.5.
- Security considerations and requirements. See 6.6.

6. Requirements

All of the requirements are numbered consecutively across all subordinate clauses.

6.1 Dynamic provisioning requests from a FSA Signalling Edge Function

6.1.1 Flow identifier

1. Signalling from the FSA Signalling Edge Function shall always carry flow identifier.
2. The flow identifier shall be used at each Flow State Aware node along the data path to recognise which data packets belong to which flow.
3. Aggregation End Points shall be able to create a flow aggregate identifier and notify next FSA signalling nodes about the flow aggregate identifier.
    - Aggregation End Points shall be able to aggregate selected flows into fewer flow aggregates, based on some criteria such as the FSA transport services parameters (see 6.1.4 and 6.1.7), and the path in the network.
4. Aggregation End Points shall be able to change the flow aggregate identifier to which a flow belongs
    - For example if aggregation is based on Preference Priority (see 6.1.4) and a flow violates the contract, it causes an automatic reallocation of a flow to a different aggregate.

Further information on the flow identifier is provided in Appendix I, I.1.

6.1.2 Signalling negotiations

5. For applications on user end-systems with a FSA Signalling Edge function and wanting service to be supported via a FSA transport service, the FSA Signalling Edge function shall forward an in-band signalling request across the user-network interface whenever a new flow is starting.
    - All other per-flow negotiations involving signals to or from a FSA Signalling Edge Function, such as response, confirmation, and close, may be conducted either in-band or out-of-band. In-band signalling is recommended as the default for all negotiations between a FSA Signalling Edge Function and one or several FSA nodes in the end-to-end data path.
6. For user end systems not capable of, or not willing to support Flow State Aware signalling, an ingress edge node may provide the option that it will generate Flow State Aware signalling requests on behalf of the user end system. In this case the edge node performs FSA Signalling Edge Function.
7. If a CPE without the FSA signalling capability wants this option to be actively supported, it shall register for the FSA transport services with information of its device type if such information has not been pre-configured. The method for registration is for further study.
8. The ingress edge nodes may provide the option of alternative choices of pre-defined parameters to be used in the request signal and may determine whether a flow should belong to a specific choice within such alternatives.

9. The ingress edge nodes may determine the FSA parameters by mapping from the IP level flow descriptor, which is distributed by a central admission entity such as the RACF through the authorization and admission procedure.
10. In order to perform signalling requests on behalf of an end-system, an edge node should maintain the registered users for future active identification, by means suitable for each network.
11. The edge node may aggregate the flows into a flow-aggregate if appropriate.
12. Edge nodes should be able to aggregate signalling messages of end-to-end flows into aggregated signalling messages, and de-aggregate the aggregated signalling messages into signalling messages of the end-to-end flows. In order to do this, edge nodes should be able to:
    - set an indication in every signalling packet that downstream FSA nodes may ignore such packets but transparently forward them.
    - remove this indication at any downstream FSA node where end-to-end signalling packets are to be restored.

See also Annex A, A.2 and Appendix I, I.2.

6.1.3 Rate

13. The new flow request shall specify the Requested Rate.

6.1.4 Preference Priority

Preference Priorities may be used by different Network Operators for the development of different service propositions. The meaning of a Preference Priority depends on the Service Context (see also 6.2.6).

14. The new flow request shall specify a Preference Priority.
15. A default value shall be used in a request signal to signify to downstream FSA nodes that the flow is a flow aggregate containing multiple flows of different Preference Priority.

See also Appendix I, I.3.

6.1.5 Authorisation and Enforcement

16. The new flow request may contain authorisation information.
17. QoS-related transport resources shall not be provided for new Flow State Aware requests unless they can be associated with an authorisation.
18. If the request is admitted, admission decisions are enforced as follows:
    - The ingress edge node, where this node is signalling on behalf of registered users, shall be able to map the traffic descriptors distributed by a central admission entity to the FSA parameters, prior to sending the request.
    - Policy enforcement shall occur for all resource requests related to a Guaranteed Rate (GR) or Maximum Rate (MR) flow request (see 6.1.6).
    - Policy enforcement shall occur for all resource requests related to the MR component of a Variable Rate (VR) flow, i.e. for resource requests related to the minimum rate of a VR flow (see 6.1.6).

See also Annex A, A.2.

6.1.6 FSA transport service requirements

19. Available Rate (AR)

The FSA Signalling Edge Function shall initiate and then frequently repeat a request, consistent with requirements 6.1.1, 6.1.2 and 6.1.3 such that:

- each such request shall include the Requested Rate whose value is pre-configured. For FSA Signalling Edge Functions in the user end-system, the configured value for the Requested Rate is recommended to take account of the highest rate that the source application can sustain, as well as entitlement captured in service profile information;
- the rate of sending data shall be controlled by the source to be always less than the latest offered rate (as provided to the source, utilising response requirement 6.2.4). If the sending rate is sustained at a value above the offered rate, packet discards may be applied.
- the Signalling Edge Function shall always update the source with the latest offered rate.

See also Annex A, A.3.1.

20. Guaranteed Rate (GR)

- In this context the required response to the signal for a new flow with a Requested Rate, as described further in requirements 6.2.3 and 6.2.4 below, indicates either: rejection of the flow or acceptance of the flow with an assignment of the flow to the “discard last” Packet Discard Priority (see 6.2.2).
- If the sending rate is sustained at a value above the guaranteed rate (the Requested Rate), packet discards may be applied.

See also Annex A, A.3.2.

21. Maximum Rate (MR)

- In this service context, the required response to the Requested Rate indicates either rejection of the flow or acceptance of the flow with conditional guarantee of the requested rate in accordance with the principles of 6.5 of [ITU-T Y.1221].
- Even in case where the in-band response is used, MR shall support user applications transmitting immediately after sending the in-band request signal associated with a new flow. This is termed the “immediate transmission” option. However, this is subject to authorisation and also subject to the flow rejection requirement discussed in 6.2.3.
- Where the immediate transmission option is supported, applications shall be able to transmit at any rate up to the Requested Rate specified in the in-band new flow request signal. If the sending rate is sustained at a level above this rate, packet discards may be applied.
- Where the immediate transmission option is not supported, applications are recommended to wait for an in-band response signal (see 6.3) prior to entering the data transmission phase.

See also Annex A, A.3.3.

22. Variable Rate (VR)

- The initial Requested Rate (consistent with requirements 6.1.1, 6.1.2, and 6.1.3) shall be interpreted as the minimum rate. In other words, policing will never reject packets sent at or below this rate, subject to authorisation and subject to the rejection requirement in 6.2.3.
- Subsequent request signals that follow the initial request signal shall be interpreted as requests for bandwidth over and above this minimum rate. The traffic within a flow that corresponds to

this additional portion of the requested bandwidth shall be treated with the Available Rate transport service.

- The policed rate is recommended to be updated following each AR request/ response but shall never be set to a rate less than the rate specified in the minimum rate requested.

See also Annex A, A.3.4.

6.1.7 Resource modification

23. Resource modification request shall be done in-band as a default. It shall further contain the Requested Rate, Preference Priority and service context.

24. For the AR and VR service context, the source shall wait for the network response to the Requested Rate before changing its sending rate upwards.

See also Annex A, A.8.

6.2 Network and FSA Signalling Edge Function responses to Flow State Aware requests

6.2.1 Requiring the flow identifier to reference the per-node flow state

25. Each accepted flow identifier shall be assigned a flow state, including a Packet Discard Priority value, at each FSA node along the data path.

6.2.2 Packet Discard Priority assignment requirements

26. The Packet Discard Priority shall not be part of signalling information.

27. The Packet Discard Priority shall allow for the possibility of distinguishing between at least two values, namely “discard first” and “discard last”.

6.2.3 Flow rejection response

28. The flow rejection response may be conveyed in-band or out-of band. In-band signalling is recommended to be the default.

29. The rejection response shall be used to reject any flow that violates Network Operator policies.

6.2.4 Flow acceptance response to requested rate

30. The acceptance response shall confirm that all FSA nodes along the edge-to-edge route have either accepted, or modified the Requested Rate from the source (or otherwise the rejection response shall be sent).

31. The acceptance response may be an in-band signal or an out-of-band signal. In-band signalling is recommended to be the default.

6.2.5 Flow acceptance response to Preference Priority

32. The acceptance response shall confirm that all FSA nodes along the edge-to-edge route have either accepted, or modified the Preference Priority request from the source (or otherwise the rejection response shall be sent).

6.2.6 Priority of packet discard, including service context use of Preference Priorities

33. Flows in the “discard last” state shall always have priority with respect to the available buffer capacity.

34. When congestion conditions are such that congestion persists even after the removal of packets of all flows with the “discard first” priority, the network should re-mark additional flows as “discard first”, starting in order from the lowest Preference Priority.

35. For any service contexts it may be possible (subject to Network Operator conditions) for a user to establish a pre-assigned Requested Rate for any Preference Priority.

See also Appendix I, I.5.

6.2.7 Congestion notification

36. An FSA node shall send a Congestion Notification signal to the iFSA Function (from any congested network node) as determined by:
    - Whether congestion conditions require that the AR flow rate be immediately reduced to avoid packet discards.
    - Whether congestion conditions are so severe it may be desirable to suppress the sending of Congestion Notifications if they would exacerbate the problem. The mechanisms for achieving this are for further study.
37. The Congestion Notification signal shall be sent to the eFSA Function if the congested network node can only communicate control signals in the direction of the flow forwarding path. After receiving the Congestion Notification, the eFSA function shall forward the message to the iFSA function, either in-band or out of band.
38. This Congestion Notification signal may be sent in-band or out-of-band. In-band signalling is recommended to be the default.
39. The detection of packet discards in end-system equipment shall not require a Congestion Notification signal to be generated within a congested FSA node, for the purposes of forwarding to and notifying an end-system of packet losses. Detection of packet losses in end systems shall be provided for by the applications themselves and/ or protocols operating at the end-to-end level.

See also Appendix I, I.5.

6.3 Signalling Requirements

6.3.1 Form of Flow State Aware signalling packets

40. Signalling packets shall be uniquely marked so that the FSA Signalling Edge Functions and the FSA nodes can easily recognise them.
41. It shall be possible to mark data packets from CPEs expecting Flow State Aware service so that they are recognisable as data packets under FSA forwarding treatment.

See also Appendix II, II.1.

6.3.2 Performance Requirements for Requests and Responses

42. Signalling exchanges shall be completed fast enough so as to meet the needs of frequent AR rate adjustments during the lifetime of a flow and to meet the needs of the MR immediate transmission option.

See also Appendix II, II.3.

6.3.3 Release of resources no longer required for GR

43. For obtaining the exact start time for the utilisation of resources, an in-band confirmation signal may be sent from the FSA Signalling Edge function in the iFSA (i.e the edge function that sent the initial request), so that all nodes in the path know the final capacity that is agreed upon.

44. For obtaining the exact ending time for the utilisation of resources, an in-band “close” signal may be sent from the FSA Signalling Edge Function in the iFSA to insure all nodes know that the capacity can be released.

See also Appendix II, II.4.

6.3.4 Protection against lost signalling packets

Requirements 45 and 46 follow from the observation that, for all services, it is important that the sender know if the initial signalling packet was lost.

45. If a response is required and is not received within predetermined retransmission time-out period, iFSA Function may send a second request.

46. For the GR service, to insure against lost in-band signals (if sent) conveying either a “confirmation” or “close” indication, the FSA Signalling Edge Function in the eFSA (which may be in the end user system) shall resend the response if data arrives before a confirmation is received and shall send an in-band confirmation indication to a received in-band “close” indication .

6.3.5 Service Contexts

47. The service context information is recommended to be included in all request and response and confirmation signals.

48. There are four Service Contexts; GR, MR, VR, and AR. Each Service Context shall have a code.

GR Service

49. When the destination receives a GR request, it shall send a response back to the iFSA Function with the Requested Rate received or, if desired, a lower rate.

50. The initiating FSA Signalling Edge Function shall send a confirmation after receiving the response telling all FSA nodes the final agreement.

51. The eFSA Function shall confirm the close to insure that there are no lost close packets.

52. The iFSA Function shall send an additional close signal (see 6.3.8) after a short timeout has elapsed indicating either the close signal or its confirmation have been lost.

MR Service

53. The sender need not wait for a response and shall be able to send at the Requested Rate immediately after the request.

54. The eFSA Function shall send a response reflecting the rate received, or a lower rate if desired.

55. No confirmation is required.

VR Service

56. The initial minimum rate request and response shall not imply a permanent sending rate, but shall be treated like an MR request.

57. The eFSA Function shall send a response with a rate no higher than received.

58. No confirmation is required.

AR Service

59. The iFSA Function is recommended to send a request with a Requested Rate set to the maximum rate that can be supported by the end equipment or application.

60. The FSA nodes shall forward the request with the rate reduced to a rate they can support.

61. The destination shall be able to reduce the rate to the maximum rate it chooses to support and shall return the rate value received or some lower rate to the iFSA Function in a response packet.

62. The sender shall conform to the offered rate received in the response.

63. Frequently the iFSA Function may send a new request to see if a higher rate is available.

64. The network FSA nodes may send to the iFSA Function signalling packets specifying a new lower rate.

65. No confirmation is required.

See also Appendix II, II.6.

6.3.6 Preference Priority

66. This parameter is recommended to be included in all requests, responses and confirmation signals.

See also Appendix II, II.7.

6.3.7 Delay Priority

67. This parameter shall be included in all requests, responses and confirmations

See also Appendix II, II.8.

6.3.8 Signalling type

68. Every QoS Signalling packet shall specify the signalling type.

69. There are at least five types of in-band signalling packets as follows:

- **Request:** The start of the signalling process is a request packet. There shall be a code for request.
- **Response**: When the eFSA Function receives a request packet which may have been modified by the network, it is recommended that it returns a response packet containing all the requested parameters and how they have been modified. There shall be a code for response.
- **Confirm**: For GR there needs to be a confirm packet. There shall be a code for confirm. This code shall also be used to confirm the GR response packet including setting the Fixed Rate (FR) indication, as further discussed in Appendix II.5.2 and insuring all FSA nodes know the final rate. It shall also be used after the GR close packet to insure that all FSA nodes actually received the close and remove any guarantees.
- **Renegotiate**: There shall be a code for renegotiate. Depending on the application, the iFSA Function may wish to try for a different or higher rate, preference, delay priority or burst tolerance. To do this and minimise confusion for the FSA nodes and the eFSA Function, it is important to use a different code rather than “request”. Thus Renegotiate requests a new QoS for the ongoing flow. It should be noted that increasing the delay priority may cause out of order packets due to the queue change, and thus this type of change would be at the sender’s risk.

- **Close:** For GR service, all the FSA nodes in the path must be informed that the reserved bandwidth can be released. Thus a there shall be a code for close. For all other services the FSA nodes shall time out the flow if no packets are seen for a given period. This period shall be up to the FSA node to determine.

In conjunction with a central admission entity such as the RACF, Response, Confirmation, and Close may be omitted. See also Appendix VI.

6.3.9 Signalling Aggregation Indication

70. **Ignore Indication.** There shall be an Ignore Indication within every signalling packet. This indication enables aggregated signalling, such that the end-to-end signalling messages are hidden from (ignored by) the core nodes. The edge nodes of an aggregation region shall be able to set this indication to show whether to ignore signalling by downstream nodes.

6.3.10 Charging Direction Indication

71. In order to allow services where the flow is paid for by the receiver, there shall be a charging direction indication. This indication could be set by the iFSA Function. If configuration or policy associated with the eFSA Function conflicts with the indicated charging direction, the response from the eFSA Function may clear the indication and such a conflict between requests and responses may be treated as appropriate to network policy (e.g. revert to sender is charged). This means that all flows will be identified as to the paying party.

6.3.11 QoS Structure Extension

72. There shall be an indication in any QoS Signalling packet that enables a single signalling packet to carry both a response to an incoming request and, separately, an outgoing request for a flow in the opposite direction.

See also Annex B, B.1.

6.3.12 Authorisation Information Attached

73. There shall be an indication that authorisation information is present.

See also Annex B, B.2.

6.3.13 Flow Aggregation Request

74. Request signals shall be able to carry an indication for a Flow Aggregation.

See also Annex B, B.3.

6.3.14 Burst Tolerance

75. There shall be some tolerance for short duration rates that exceed the Requested Rate. A parameter defining such a tolerance shall be included in all requests, responses and confirmations.

See also Appendix II, II.9.

6.3.15 QoS Approval Indication

76. A FSA node shall set the QoS Approval Indication if it has approved the request from a flow.

77. A FSA node may clear this indication to inform the iFSA and/or eFSA Functions that the request is not approved.

6.3.16 FSA Signalling Edge Function

78. The eFSA Function shall be able to create and send a response to a request to show their acceptance or rejection of a flow request, or modify the request so that it is accepted.

    - The iFSA Function shall be able to receive the response and if the request has changed, it shall be able to notify to the application.

6.3.17 FSA node requirements

6.3.17.1 FSA node operation

79. The ingress edge FSA node shall be responsible for checking an attached Authorisation Indication and also determining if the Preference Priority requested is authorised for this user.

80. All FSA nodes receiving a flow request shall check their available capacity and resources and adjust downward the Requested Rates, Delay Priority, Preference Priority, and Burst Tolerance requested to what they believe they can support with reasonable assurance.

81. For Available Rate (AR) and Variable Rate (VR) services, any FSA node along the path may send a message indicating that the rate needs to be lowered.

    - It is recommended for this message to go the iFSA Function, but if this is difficult, it shall be sent to the terminating FSA Signalling Edge Function who will then send a response to the iFSA Function.

See also Annex B, B.4.

6.4 Admission Decision

This clause defines minimum requirement for managing the admission decision process.

6.4.1 Admission Decision for Maximum Rate (MR) Flows

82. The decision process shall be fast enough to support the immediate transmission capability, taking account of Preference Priority and available capacity.

See also Appendix III, III.1.

6.4.2 Admission Decision for Available Rate (AR) Flows

83. The decision process shall be capable of supporting frequent rate update requests on each AR flow.

See also Appendix III, III.2.

6.4.3 Admission Decision for Variable Rate (VR) Flows

84. VR flows are expected to be treated like AR flows except for their minimum guaranteed capacity which shall be treated like MR. Thus the admission decision for VR flows will be a combination of the one for the MR and the one for the AR.

6.4.4 Admission Decision for Guaranteed Rate (GR) Flows

85. Total capacity assigned for GR flows on a port shall be checked before the admission decision.

6.5. General architectural requirements on the management of transport connections carrying Flow State Aware traffic and other traffic.

86. Flow State Aware QoS controls shall be agnostic to the underlying transport technology.

87. A given network link (between two network nodes) may be configured so that it is not dedicated to carrying Flow State Aware traffic only.

See also Appendix IV, IV.1.

6.6 Security Considerations and Requirements

88. *Authentication* – User authentication is addressed in 6.1.5. FSA nodes within a domain may authenticate to peer FSA nodes within the domain. FSA nodes communicating as peers across a domain boundary should authenticate with each other.

89. *Authorisation* – Authorisation is addressed in 6.1.5.

90. *Data Confidentiality* – The use of Flow State Aware transport technology shall not impose additional data confidentiality requirements. In the case where user flows with data confidentiality requirements also invoke the AR, MR, VR, or GR service, the parameters describing the in-band service request shall not be encrypted.

91. *Data Integrity* – Flow State Aware parameters may be protected against unauthorized modification while in transit. Flow State Aware parameter requests may be protected against replay attacks, in conjunction with data integrity protection binding a set of Flow State Aware parameters to a specific flow.

92. *Accountability* – It is recommended that Flow State Aware service invocations are logged, including the identity of the entity requesting the service, the actual service request, and actual service granted.

93. *Availability/accessibility* – Flow State Aware services shall respect the priority preference of each authenticated entity in making admission decisions.

94. *Privacy* – It is recommended that Flow State Aware services ensure the privacy of user specific policy profiles defining QoS parameter limits and privileges.

95. *Protection against network attacks, from within or outside* – It is recommended that Flow State Aware services include mechanisms to protect against malformed service invocations and to mitigate Denial of Service (DoS) attacks.

Annex A: Dynamic provisioning requests from an FSA Signalling Edge Function

This annex forms an integral part of this Recommendation.

A.1. Negotiations for a CPE without signalling capability

This clause references requirements in 6.1.2

With respect to requirement 8, as stated there, the edge nodes may have pre-defined parameters. Each new flow may be monitored, and if the flow under observation reaches some threshold that determines which parameter choice to use, the node generates an in-band signalling request. This starts negotiation for the flow.

Thus, if a data packet from a registered user is received, the edge node records the flow identifier as an ‘unidentified flow identifier’ and monitors for the unidentified flows. If conditions are met (Parameters are set a priori), the edge node puts the unidentified flow into the ‘identified flow’ list with proper request items a CPE would request (Requested Rate, Preference Priority, and Service Context), and then performs signalling as a CPE with Flow-State-Aware capability would do. The signalling function further required are performed by the edge node. The edge node may aggregate the flow into a flow-aggregate if appropriate.

A.2. Authorisation

According to the principles of [ITU-T Y.2111], the following requirements may be derived for the authorisation of Flow State Aware in-band QoS requests while recognising the fast processing nature of services created by FSA nodes. It is based on a “Two-Phase Scheme” as described in [ITU-T Y.2111]: Authorisation is performed in one step (see Figure A.1), followed by reservation and commitment in another step (see Figure A.2).

Before the CPE forwards the initial in-band signal associated with a new flow request, it shall first trigger the QoS initial authorisation procedure. This initial procedure (see Figure A.1) may be triggered by a service-establishment signalling message (e.g. a SIP Invite message) that in turn instigates an SCF request to the PD-FE. The PD-FE may, as a result, generate an authorisation token as optional for a given service and send it to the SCF. The SCF may forward the authorisation token in the service signalling message to the CPE (see Figure A.1).

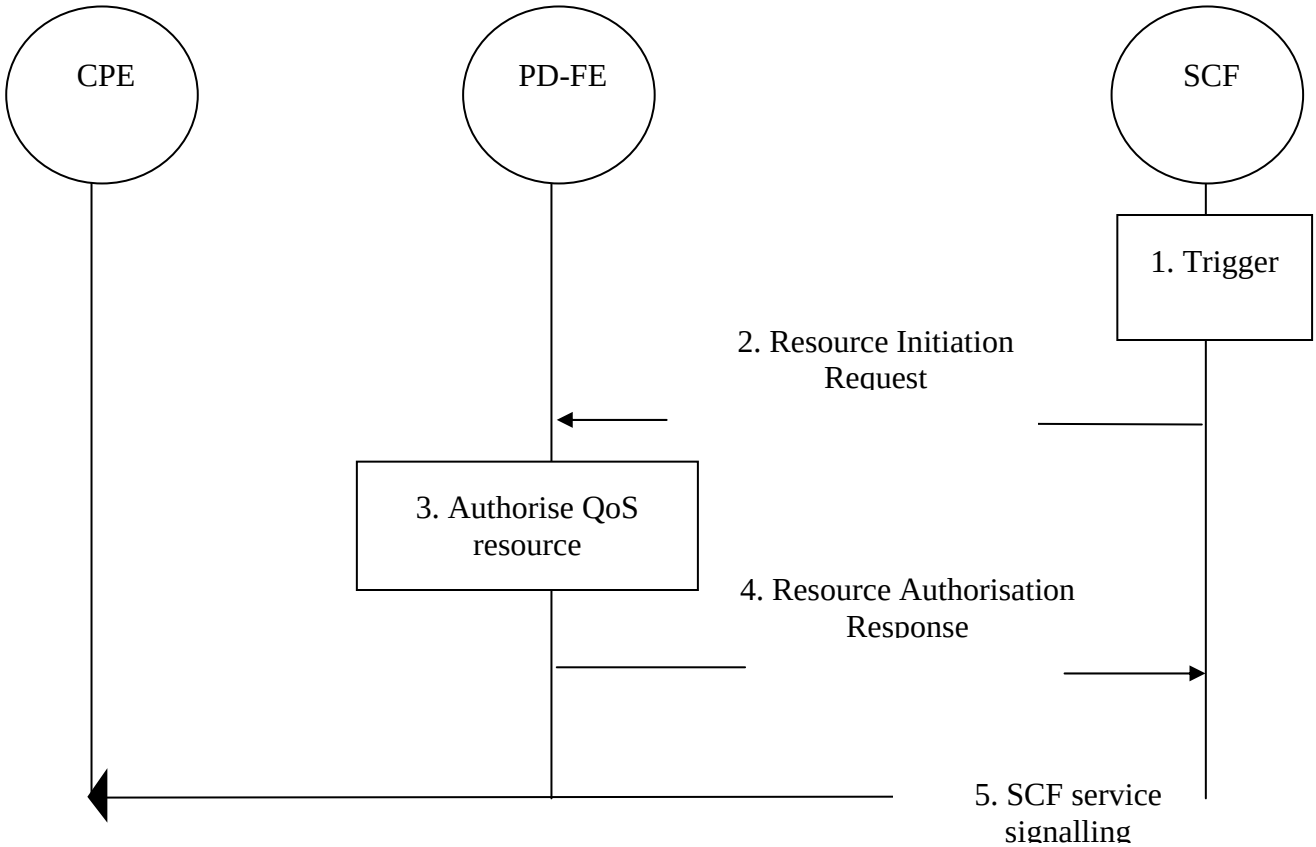


Figure A.1 SCF-requested QoS initial authorisation procedure

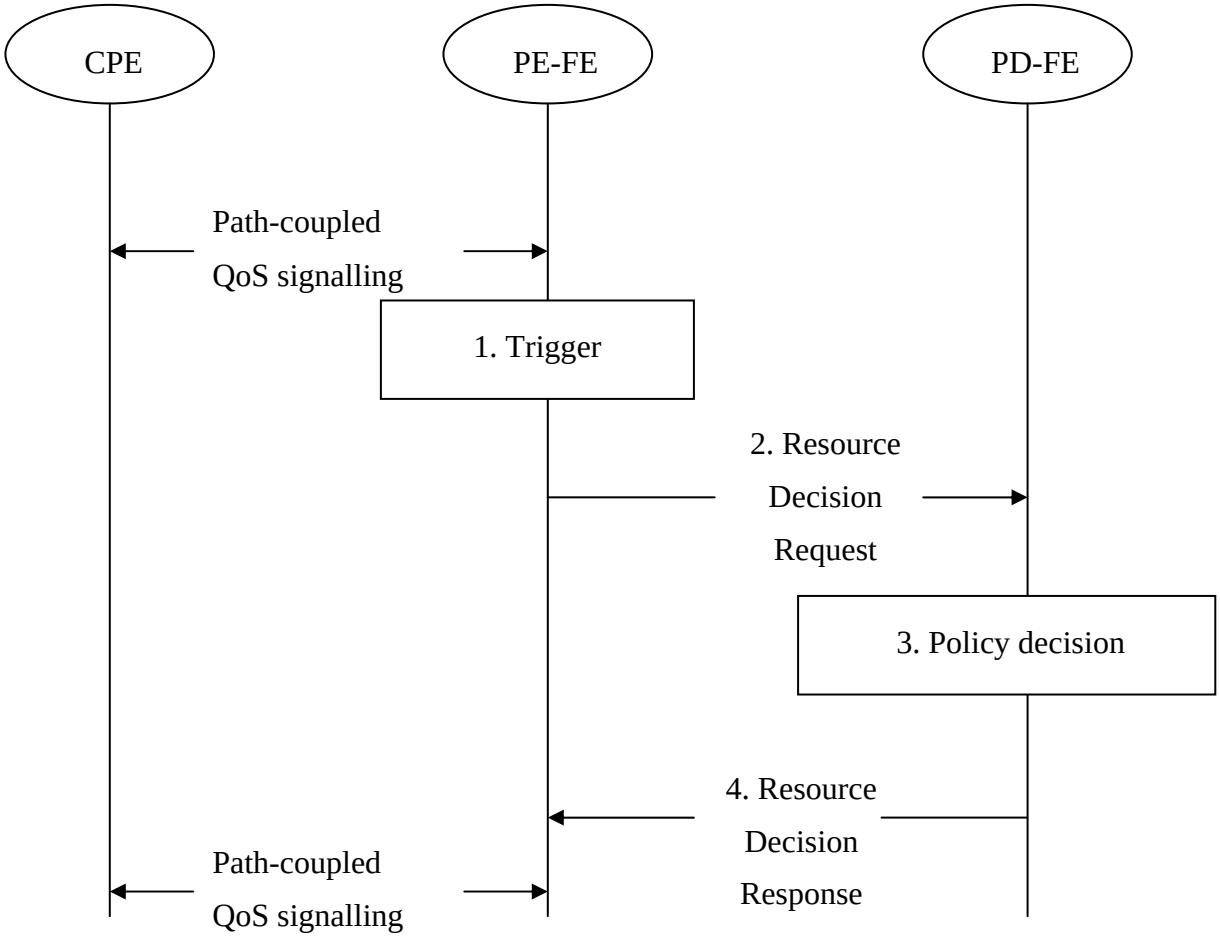


Figure A.2 CPE-requested QoS resource reservation procedure

A.3. Service context

This clause references requirements in 6.1.6

Requests & responses used for AR	Requests & responses used for GR	Requests & responses used for MR	Requests & responses used for VR
Basic requests & responses of Y.1221,clause 6.5, dealing with the association of a flow state with a flow identity, together with a requested Maximum Rate & Preference Priority			

Figure A.3 Service Options

Figure A.3 summarises the grouping of new flow requests and responses into service contexts. This greatly increases the utility of the Flow State Aware capability as supported through the underlying requests and responses related to [ITU-T Y.1221].

A.3.1. Available Rate (AR)

Utilising the service context rules of 6.1.6, the transfer capability described in clause 6.5 of [ITU-T Y.1221] provides support for an Available Rate (AR) service, similar to that described in clause 6.2.4 of [b-ITU-T I.371]. In [b-ITU-T I.371] it is stated that "in ABR, the user regularly polls the network for the currently available bandwidth by sending RM cells conveying a requested rate to the network". RM cells are a form of in-band signal.

A.3.2. Guaranteed Rate (GR)

This service context (termed the Guaranteed Rate or GR service context) may be offered for applications requiring guaranteed bandwidth for the entire duration of the flow. Additional requirements apply to this service to ensure the reservation is aborted if not all nodes can support the request and to ensure that clear down of reservations is completed successfully at the end of the flow.

A.3.3. Maximum Rate (MR)

In effect the service offer is that the requested flow may proceed, with this "as soon as possible" condition on meeting the loss component of the requirement. We term this the Maximum Rate (MR) service. As stated in 6.5.2 of [ITU-T Y.1221], while in "discard first" status, a flow may subject to a higher probability of packet discards during moments of congestion. It is up to the application to choose to terminate the flow or continue and compensate for any bursts of losses as and when they occur.

A.3.4. Variable rate (VR)

The service context rules of 6.1.6 may be offered as a service that supports applications that require a minimum rate and have some degree of elasticity above this minimum.

Annex B: Signalling requirements

This annex forms an integral part of this Recommendation.

B.1 Second QoS Structure Attached

This clause references requirement 72.

When the eFSA Function (which may be in the end system) receives a QoS Signalling Request and desires to establish a return connection, it may proceed as follows. This FSA Signalling Edge Function may send both a response to the request and also a QoS Signalling Request for the opposite direction in the same packet. This fits ideally into the normal TCP start up handshake where the first packet is a SYNC and would include the forward request, the next packet is the SYNC/ACK which would contain the forward response and the reverse request, and the third packet is an ACK which would carry the reverse response.

B.2 Authorisation Information Attached

This clause references requirement 73.

It is important to allow a cryptographic authorisation header to be attached to each flow Request so that the network can verify the user and the user's privileges.

B.3 Flow Aggregation Request

This clause references requirement 74.

There is a strong need in all networks for flow aggregation which groups and maps some of flows into a smaller set of flow aggregates. With the full capabilities of the signalling available to specify the flow aggregate, a flow aggregate may be one of GR, MR, VR or AR, with any rate, preference priority, and delay priority desired. It may utilise encapsulating IP header (IPv4 or IPv6), MPLS label, VLAN ID, or the likes to specify the destination such that, with this header, the flow aggregate could be routed across any transport network to its destination. At the egress FSA node, a flow aggregate would be terminated and all the traffic inside the flow aggregate directed to the specified address. A Signalling Edge Function may request flow aggregation.

B.4 FSA node operation

This clause references requirements in 6.3.17.1.

With reference to these requirements, when a compliant FSA node receives an in-band signalling packet it will need to inspect the QoS Structure to determine what action it needs to take, if any. If this is a new flow, it will need to check the capacity of the output port and determine what rate it can accept.

Note that, with respect to requirement 80, it is not necessary for FSA nodes to guarantee support under all possible conditions, only that they have high confidence that they can support the resulting request under normal operating conditions.

Appendix I: Supplementary information on information exchanges via requests from a FSA Signalling Edge Function and associated responses

This appendix does not form an integral part of this Recommendation.

I.1 Flow identifier

[ITU-T Y.1221] describes the combination of parameters “source IP address”, “destination IP address”, “source and destination port numbers”, “protocol”, and “experimental/ Diffserv value” as a basis for a flow identifier.

Requirement 1 is derived from 6.5.1 of [ITU-T Y.1221], where it can be inferred that signalling from the FSA Signalling End Function always carries flow identifier information that is the same as that carried in the header of each data packet of the flow. This information about the flow identifier shall be used at each Flow State Aware node along the data path to recognise which data packets belong to which flow.

Ingress edge nodes may support the option to aggregate selected flows into fewer flow aggregates, based on some criteria such as the Service Context (see 6.1.6), the Preference Priority (see 6.1.4), and the path in the network.

An alternative method for carrying the flow identifier, or especially for carrying a flow aggregate identifier, may be through the Multi-Protocol Label Switching (MPLS) label. [b-RFC 3270] describes two standard methods to map the Diffserv Behaviour Aggregates (BAs) onto MPLS Label Switched Paths (LSPs).

For networks without MPLS functions, an encapsulating protocol (e.g. IP in IP encapsulation [b-RFC 2003] or Generic Routing Encapsulation (GRE) [b-RFC 2748]) header with IPv4 address may be used as the identifier for flow aggregates within a single administrative network domain.

I.2 In-band negotiations

Requirement 5, while not fully explained in 6.5.1 of [ITU-T Y.1221], supports the signalling performance requirement 42. It also facilitates service options like Available Rate as discussed further in 6.1.6.

Again referring to requirement 5 there is a reference to the same terminology in clause 6.5.1 of [ITU-T Y.1221], i.e. that there is an in-band signal.

Requirement 7 envisages an edge node that is used for flow negotiations on behalf of an end node. To perform this function, the edge node may utilise information about the end nodes that was obtained during registration. The following steps describe this in more detail:

1. An end node registers itself for flow identification with its IP address to an edge node. Or an access node (e.g. Access Node Function (ANF) in Access Transport) registers its attached end nodes for the flow identification with their IP addresses and end nodes information to an edge node.
2. The end node information may be terminal type (if the terminal is a user terminal), or application type (if the terminal is an application server).
3. The edge node stores the IP address and end node information.
4. When this edge node receives a packet with an unknown identifier from a registered end node, it retrieves the end node registration information. It may monitor the flow for

additional information to determine the flow characteristics and perform signalling as necessary.

The followings are a few examples of possible end nodes which would benefit from active flow identification.

- VoIP terminals
- Mobile handheld devices (including cellular phone, PDA, etc.): Traffic from such terminals may be either voice, streaming video, or message (text or multimedia). The voice and streaming video are with distinct encoding rates. If an end node registers as a mobile handheld device, the edge node may utilise this information together with information obtained by monitoring a few packets to determine the characteristics of the flow.
- Media streaming servers: If such a server is the source of consecutive packets with relatively regular interval, the edge node may infer that the flow is a media stream. The encoding rate information may be retrieved from information stored during registration.
- Emergency service related equipments: Light-weight emergency or military equipments may not have the capability to support Flow-state-aware signalling. Registration information may enable a node to assign such equipment the pertinent Service Contexts, rates, and priorities.

One of the possible ways to expedite the identification is to look at the upper layer header (e.g. RTP Payload Type (PT) field in the header), only when the end node of the packet is figured to be registered. The PT field provides detailed information on whether the flow contains audio or video, the type of encoding, and the encoding rate.

I.3 Preference Priority request

The higher Preference Priorities may be reserved for such purposes as emergency services and military commands, etc.

I.4 Authentication

The following service scenarios illustrate authentication aspects of Flow State Aware transport and service.

Scenario 1: mobile user access

With reference to Figure I.1 a mobile user access scenario is shown. For example, a terminal (End-A) could connect to a public WLAN hotspot (Node 1 of Figure I.1). It may establish an IPsec tunnel to an IPsec gateway (Node 2) to access enterprise services. However, assuming IPv4 is used, Flow State Aware QoS controls cannot be applied to the QoS experience obtained along the IPsec tunnel. If IPsec is used, there is no way for the FSA nodes to recognise the flow since the ports and protocol are hidden. Thus with IPsec and IPv4, the protocol cannot be used and only Diffserv is available. IPsec can be used with IPv6 and FSA protocols.

With IPv4, to make use of Flow State Aware QoS, End-A may make use of split IPsec tunnelling, whereby:

- non Flow State Aware traffic destined to the service provider domain is sent via IPsec,
- Flow State Aware traffic bypasses IPsec.

Thus, in Figure I.1, the FSA IPv4 End-A is shown gaining access to a FSA content services platform (e.g. a Content Distribution and Delivery platform) that is accessible from a Flow State Aware Node 1.

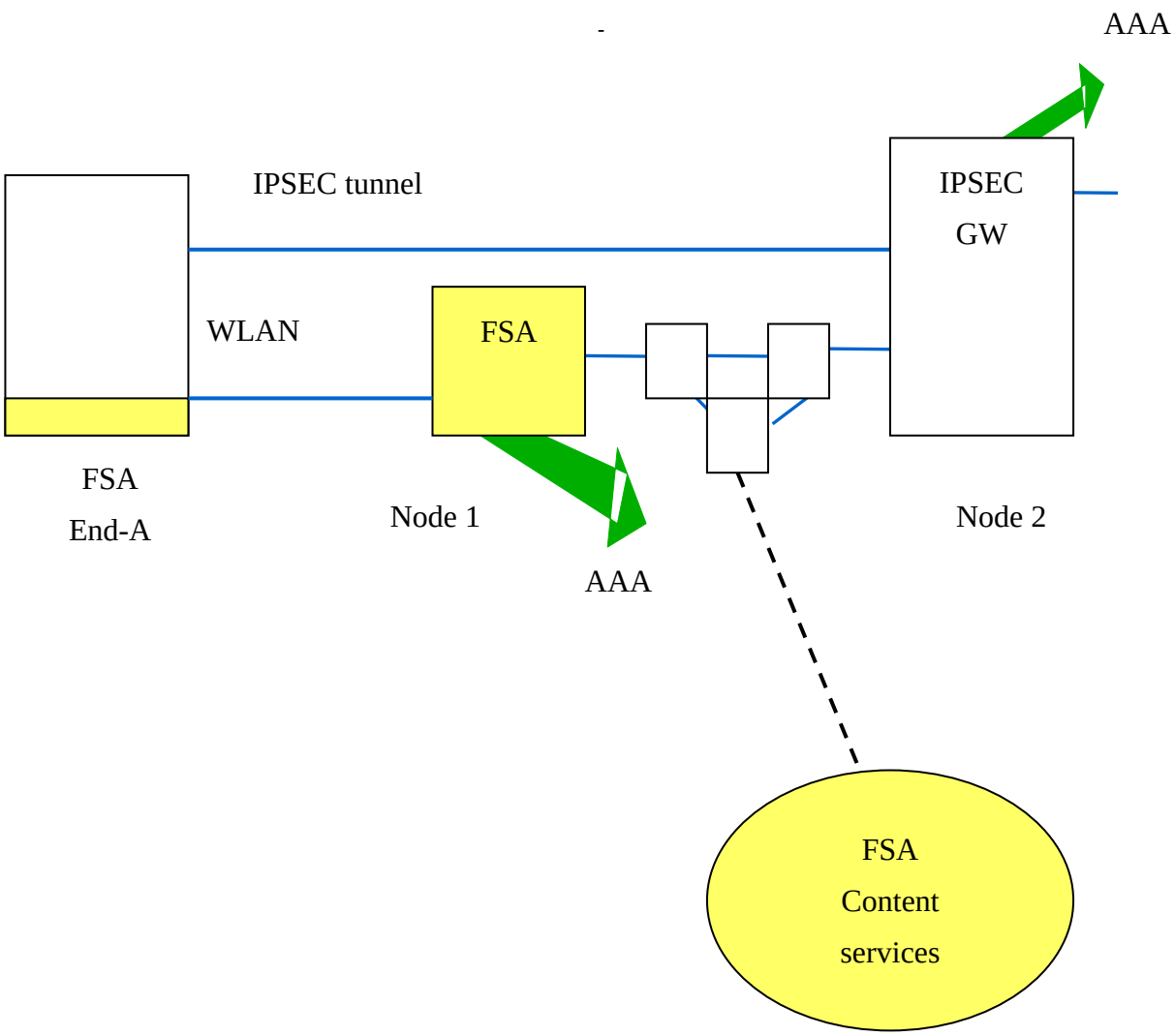


Figure I.1 Scenario 1: Mobile user access.

Scenario 2: Broadband access

Figure I.2 shows End-A connected over DSL to a BRAS. This BRAS acts as a LAC and forwards the traffic by using L2TP to a second BRAS acting as an LNS. Node 1 may issue a Radius request to obtain attributes for the tunnel to be established (e.g. [b-RFC 2868]). The second BRAS terminates the PPP state machine and may issue a Radius request to perform user authentication. Policy enforcement is generally only done in node 2. Part of that policy enforcement could be a Flow Aggregate with associated control that manages downstream traffic (towards End-A) at the point where it is aggregated and manages upstream Requested Rate (GR, MR or VR) requests and AR-controlled back-pressure.

An alternative arrangement is shown in Figure I.3. Here, the L2TP tunnel is replaced by a Flow State Aware flow aggregate. The objective is to handle any packet discards, either upstream or downstream at the flow aggregate entrance. For example, the sudden insertion of an upstream emergency call may cause this.

.

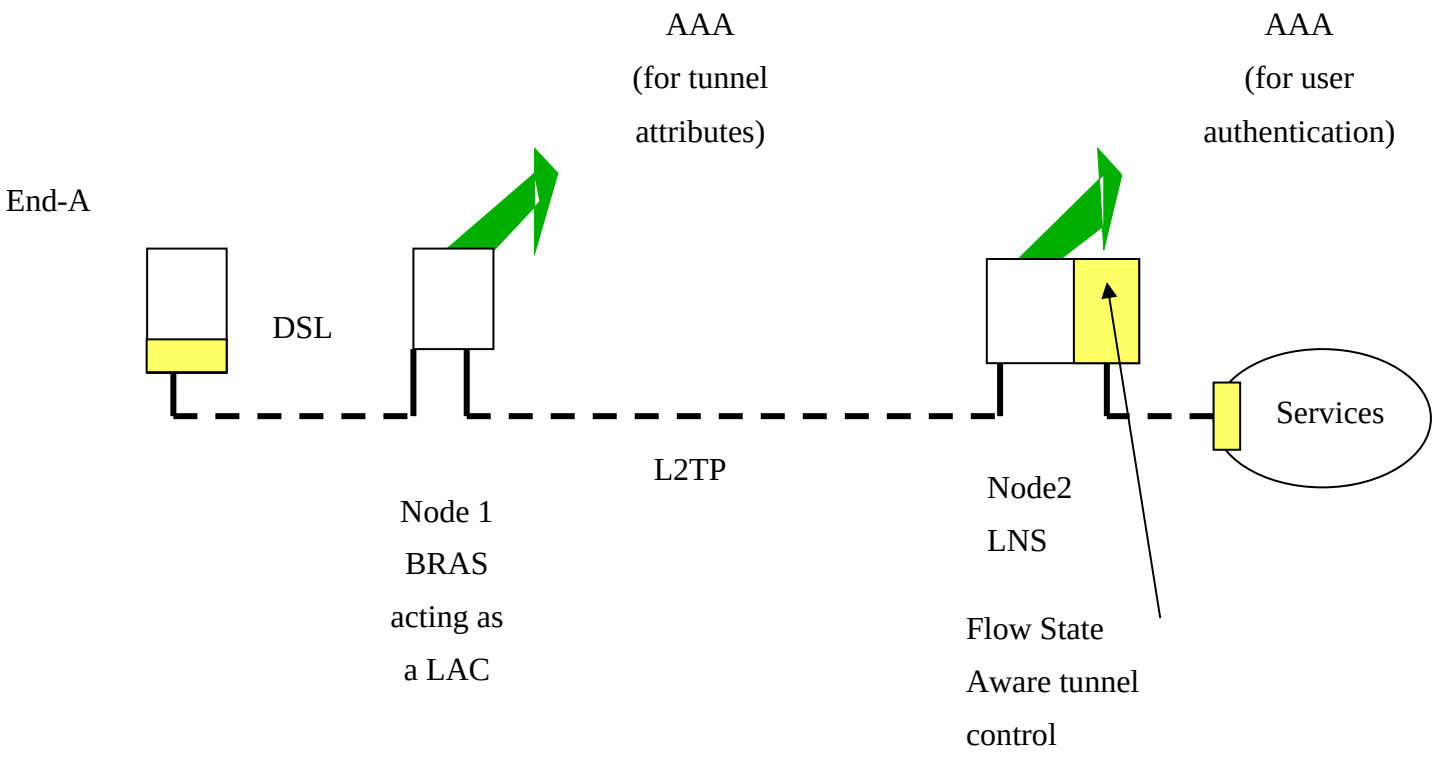


Figure I.2 Broadband services via an L2TP tunnel

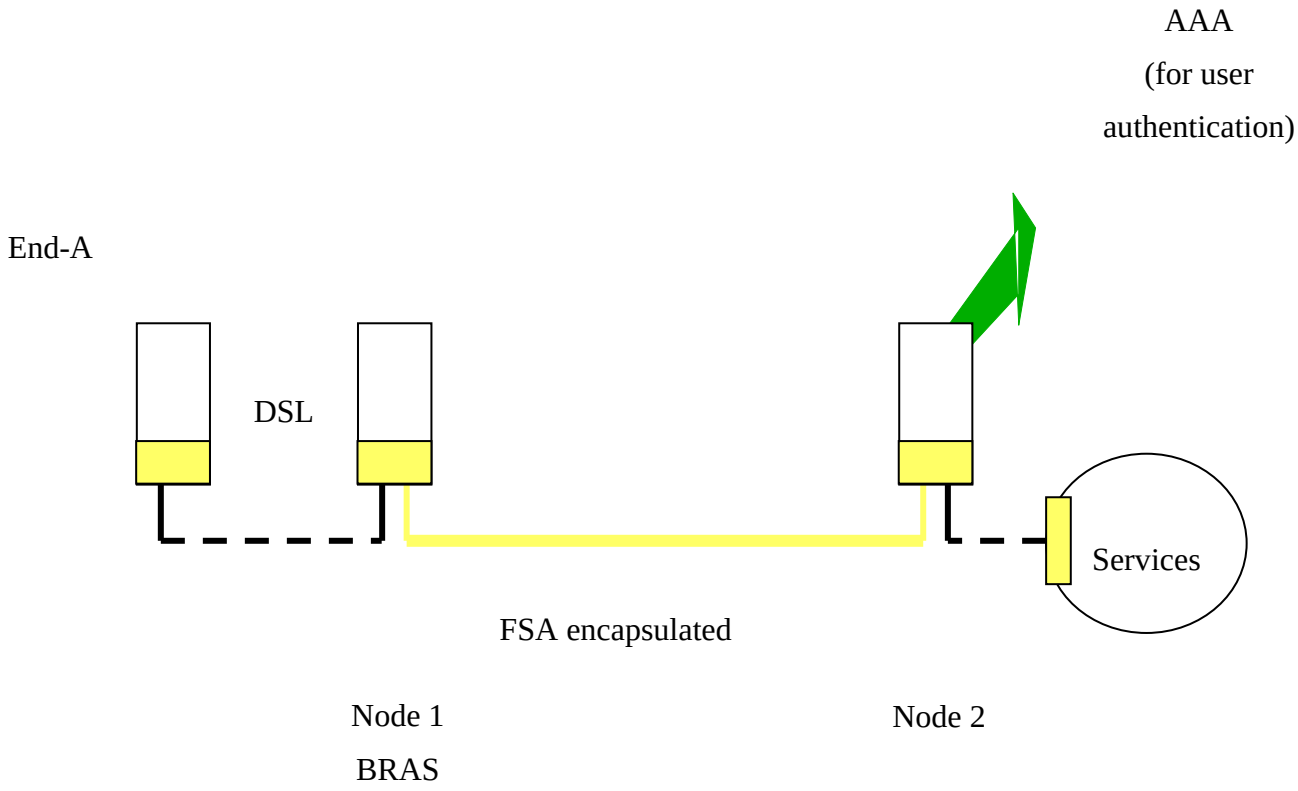


Figure I.3 Flow Aggregate with FSA outer header established by in-band signals between Nodes 1 & 2

Network Management commands are used to trigger Nodes 1 and 2 to exchange in-band Flow State Aware signals. These establish, say, a flow aggregate with GR Service Context including a defined capacity (Requested Rate). All IP packets from End-A's or from "Services" are encapsulated with an outer packet header bearing the flow identifier of the aggregate flow. Flow State Aware packets sent by End-A are QoS-managed along the aggregate flow at Node 1. Similarly, Flow State Aware "services" packets are managed along the aggregate flow at Node 2.

I.5. Priority of packet discard, including service context use of Preference Priorities

This clause references requirements in 6.2.6.

Typically a high Preference Priority flow may be established even when network capacity is already fully loaded with flows of lower Preference Priorities.

For example, a high Preference Priority Emergency Services MR flow could be established where transmission could begin immediately, without even waiting for a network response, provided the authorisation was accepted. This would be one possible service context. Another context may allow parental configuration of content so that certain types of content or certain user identities determine the Preference Priority.

Therefore, two parameters (Preference Priority and Packet Discard Priority) determine the probability of packet discard. The relationship between the nodal behaviours and these parameters in Flow State Aware network nodes are as follows.

- Admission decisions are not governed by Packet Discard Priority. But newly admitted flows are usually marked as "discard first". Depending upon policy, some newly admitted flows

(for example, those with a high Preference Priority) may be immediately marked “discard last”.

- Packet discard (buffer management) decisions are governed by Packet Discard Priority, but this may be influenced by Preference Priority as follows:
  - When packet discard is necessary, if there are packets with “discard first” priority, drop those packets,
  - If buffer congestion persists, remark additional flows of the lowest Preference Priority(s) as “discard first”.

I.6. Congestion notification

This clause references requirements in 6.2.7.

It is desirable that the implementation has queue buffers that are large enough to absorb the remaining round-trip delay prior to the rate reduction.

Appendix II: Supplementary information to Signalling Requirements

This appendix does not form an integral part of this Recommendation.

II.1 Recognition of QoS Signalling Packets

This clause references requirement 40.

There is always a need to identify some in-band signalling packets. However, in-band signalling need only be done at the start of a flow.

The exact method for identification of data packets is for further study.

II.2 Form of QoS information

This clause references requirement 40.

The requirement 40 follows from the fact that data packets, by chance, could include information that looks similar to the information in a signalling packet. Therefore a unique marking for signalling packets is required.

FSA nodes should be able to read and modify a QoS Structure within a signalling request and, if used, a response, confirmation, or close in real time at any port speed. Port speeds are never faster than the logic speeds currently possible for simple operations like reading and/or writing packets into memory, checking sum checks, operating linked lists, reading and interpreting fixed length fields, and the like. Usually the serial input stream is first converted to a parallel word stream so that standard CMOS logic can keep up. Thus, so long as the QoS Structure follows some simple rules, any port speed currently feasible for high speed FSA nodes could be processed.

The following text illustrates an example implementation, not the only implementation, facilitating fast processing..

II.2.1 All information fields are in a fixed location in the QoS Structure.

II.2.2 All numerical value which need to be treated as a number are structured as an integral number of bytes and are byte aligned.

- If the number is 2 bytes it is dual byte aligned and if 4 bytes it is quad byte aligned. This is because the parallel input stream in IP is quad byte aligned and it is more steps to shift and mask to find a value.

II.2.3 IP protocols also works with 8 byte boundaries and thus the total QoS Structure is a multiple of 8 bytes.

- Also, the number of 8 byte groups required is as small as possible to reduce overhead and processing.

II.2.4 Since overhead is of great importance, options within the QoS Structure may be packed to a byte.

II.3 Performance Requirements for Requests and Responses

This clause references requirement 42.

The immediate transmission option relies on the fast establishment of flow state at every FSA node such that, in the extreme, the request packet arrives at an FSA node and is immediately followed by the first data packet of the flow.

II.4 Release of resources no longer required

This clause references requirements in 6.3.3.

An FSA node may need to reject new traffic when too many absolute capacity guarantees are no longer required but not yet released.

II.5 QoS Signalling Parameters

Subsections II.5.1 and II.5.2 provide information for illustrative purposes relating to an example implementation, not the only implementation.

II.5.1 IPv6 Header

The QoS Structure in IPv6 may be chosen to be a hop-by-hop option.

II.5.2 Rates

II.5.2.1 As an example implementation, two rates may be utilised, one that is network selected (Network Rate or NR) and one that is a user requirement (Fixed Rate or FR). NR would be used where the user is sending buffered data where the rate may vary (typically TCP). FR would be used where the user needs a fixed rate available at all times.

II.5.2.2 These may have multiple uses and may be used together so as to support the four types of service (GR, MR, VR, and AR). As an example:

GR Service

II.5.2.2.1 The GR service could use the FR field to specify the Requested Rate with the NR field set to zero.

MR Service

II.5.2.2.2 The MR service could use the FR field to specify the Requested Rate with the NR field set to zero.

AR Service

II.5.2.2.3 The AR service could use the NR rate to specify the Requested Rate, setting the NR rate to the maximum rate that the application or computer can support. The FR field could be set to zero.

VR Service

II.5.2.2.4 The VR service could use the NR plus FR rates to specify the Requested Rate. The FR portion could carry the minimum required rate (lowest value) for the flow. The NR rate could be set as in the AR service, see II.5.2.2.3

II.5.2.3 The lowest rate may be low enough that the FSA node would not find any significant value in reserving or managing a flow to a lower rate.

- The lowest rate could be 1 kbps.

II.5.2.4 Zero may be represented since NR or FR may be zero.

II.6 Service Contexts

This clause references requirements in 6.3.5.

GR Service

Guaranteed Rate has the most stringent requirement and is for the equivalent of a leased line where the rate is permanent until closed.

GR service requires commitments from FSA nodes even in the absence of traffic.

MR Service

The second type of flow is Maximum Rate where a network timeout after a period of no data is sufficient. MR may be used for video, voice, or other streaming media whenever the QoS requirement is for low loss and low delay variance, but where such requirements allow for a network conditional guarantee that facilitates immediate transmission. The conditional guarantee is that the network will support the loss/ delay QoS targets as soon as possible, while allowing the immediate transmission to continue..

AR Service

This service will typically be used for TCP flows but can be used with any protocol. It offers an available rate that can be immediately supported across the network.

VR Service

This service offers a minimum rate with the conditional guarantee and immediate transmission characteristics of MR. Additionally it allows an application to exploit the latest available rate so that it may send at a higher rate but need never send at a lower rate than the MR value. For example, a stock trade may be required to be transacted or reported in some known and acceptable time but faster is better.

II.7 Preference Priority

This clause references requirement 66.

The Preference priority is a parameter used to determine which flows should be admitted in the case of a network overload. The overload could be that on an access line has too many video requests, or it could be due to trunk or network equipment failures. This type of capability is necessary as the network moves toward streaming media flows (GR and MR) and away from available rate flows. It has always been implemented in telephone and military networks. However, since the Internet was mostly TCP to start, it has not been required or standardized before this year. Now it is required. The number of preference priorities needs to be sufficient for the multiple military and civilian emergency systems, corporate priorities, and home priorities. Assigning the preference codes is beyond the scope of these requirements but the number of levels needs to be set.

II.8 Delay Priority

This clause references requirement 67.

Although absolute delay is not controllable in a network due to the speed of light, delay variance or jitter can be controlled and may have different requirements for different services. Typically video and voice require lower delay variance than file transfer but there may be many other services with many different requirements.

II.9 Burst Tolerance

This clause references requirements 75.

The Rates negotiated in GR, MR, VR, and AR are maximum rates and the user is free to send at lower rates. But if the transmission rate momentarily exceeds the agreed rate, it is typical in packet networks to include some burst tolerance.

II.10 Flow Identifier Fields

This clause references requirements in 6.1.1.

IPv6 Flow Identifier Field

The following is an example implementation for illustrative purposes. In IPv6 the 20 bit Sender Flow Label may be used as part of the vector of parameters that identify the correct flow from a response packet. The triplet Source Address, Destination Address, and the Flow Label could together identify a flow. The response has the same Source and Destination address (reversed) but the Senders Flow Label is still required to identify the matching request. Thus, in this example case there would be a 20 bit Source Flow Label field in the Response packet. The eFSA Function fills this in from the request.

IPv4 Flow Identifier Fields

The following is an example implementation for illustrative purposes. In IPv4, if unencrypted, the source address, destination address, source port, destination port, and protocol could define the flow. When the eFSA Function sends a response, these fields are all in the packet although source and destination are reversed. Thus, the sender needs no extra information to identify the flow.. It may be noted that most senders are behind NAT devices and the source address and source port have been changed, perhaps several times by the time the eFSA Function gets the request. However, they are all restored by the time the response gets to the iFSA Function. Thus no extra issues are raised by NAT.

However if the iFSA Function moves (mobile user) the eFSA Function will not recognise the new source address and source port. Nor does the iFSA Function know what the source address and port were that the eFSA Function recorded. Therefore, in order to support mobility of the source, a simple solution is for the eFSA Function to copy the source address and source port it receives in the request into the response. The iFSA Function could save this information as the identifier of the flow. Then when the iFSA Function moves, it makes a new QoS Request and includes the saved original source address and port in the request instead of zero in these fields. The network treats this as a new request and times out the old one. However, the eFSA Function can see that these fields are not zero and match up the original source address and port with ongoing flows and thus determine that this is a continuation of the flow.

An example implementation could utilise two fields for IPv4, the “original source address” (32 bits) and the “original source port” (16 bits). Since IPv6 requires 32 bits of header information, this space can be used for the IPv4 original source address. Then the IPv4 original source port can be placed in the same 20 bit space that IPv6 requires for the source Flow Label.

Appendix III: Preference Resolution

This appendix does not form an integral part of this Recommendation.

III.1 Preference Resolution for Maximum Rate (MR) Flows

This clause references requirement 82.

The preference level of a newly admitted MR flow may alter the QoS support of MR flows of lower preference levels. If necessary, when there is insufficient capacity to support all MR flows, some or all of the lower preference flows may continue to transmit as in the case of the immediate transmission option. For such flows the network will revert to the conditional guarantee that loss/ delay targets will be supported as soon as possible.

III.2 Preference Resolution for Available Rate (AR) Flows

This clause references requirement 83.

With AR traffic (typically TCP) the network decides on the rate it could support for each flow. There may be different classes of AR traffic where some flows are permitted more capacity, but within a class it assumed that the goal would be for rate equality or fairness. When a new AR flow is received, the network equipment would examine the available capacity for the class and assign what it believed was a fair rate to the flow. Typically the process would allow the total AR traffic in the class to utilise as much of the class capacity as is presumed to be safe and thus most of the time there would be a preference decision to be considered.

Requirement 83 does not specify how much more or less capacity is given to flows of different preferences, and that may well be a function that the network operator may wish to control. This requirement however leads to a very simple process, the preference can be converted to a weight by a function or a table lookup so that a weight is determined for each flow. Then each new flow could be assigned a weighted fraction of the total capacity. One measurement is required, the sum of the flow weights in process. The new flow would then receive its weight times the total safe capacity divided by the sum of the active flow weights. This is one solution that illustrates how a large number of preferences could be supported without the computation and memory increasing with the number of preferences.

Appendix IV: Supplementary information relating to requirements on the management of transport connections carrying Flow State Aware traffic and other traffic.

This appendix does not form an integral part of this Recommendation.

IV.1 General architectural assumptions

This clause references requirements in 6.5

The main requirement is that Flow State Aware QoS controls are agnostic to the underlying transport technology. This can be Ethernet, ATM, or any other choice of the Network Provider, and may be a mixture of different transport technologies at different locations.

Additional assumptions:

- A given network link (between two network nodes) may not be dedicated to carrying Flow State Aware traffic only.
- Where a given network link is carrying a mixture of Flow State Aware traffic and other traffic, the FSA node shall assume that part of the capacity of that link is guaranteed to be available for Flow State Aware traffic. The role of FSA QoS control should be to control the utilisation of that capacity among the competing flows (AR, MR, VR, CR). The network may achieve capacity guarantees in various ways, including:
  - IP-layer scheduling functions to manage and limit the capacity available to the non-FSA traffic, together with FSA QoS controls to limit the capacity available to the FSA traffic. Note that this method may allow for the dynamic "borrowing" of unused capacity by either FSA or non-FSA traffic as discussed in the next bullet.
  - FSA management of link capacity does not imply a dedicated FSA management function per link. For example, consider the Broadband access scenario shown in Figure IV.1 and Figure IV.2. These illustrate an arrangement where a single FSA Access QoS functional entity is capable of managing both DSL and Ethernet link capacity limits. Clearly the scale capabilities of this arrangement will be vendor-specific. In the case of Figure IV.2 the FSA Access QoS management functional entity is shown as being combined with an Ethernet VLAN switch. Of course these two functions could also be split. The main difference with Figure IV.1 is that all traffic, FSA and non-FSA, passes through FSA QoS control management. The non-FSA traffic is either passed transparently or managed e.g. throttled (as necessary), if the Network Provider chooses to offer such a service.
  - Figure IV.1 and Figure IV.2 also illustrate the requirement that FSA management of link capacity shall take into account the capacity reserved on that link for non-FSA traffic on both the DSL and Ethernet links. i.e.:
    - the non-FSA traffic within (in the scenario of Figure IV.1 and Figure IV.2) an Ethernet VLAN carrying a mix of FSA and non-FSA traffic, or
    - the additional non-FSA traffic on the physical link towards an Access Node, in dedicated non-FSA VLANs.

    Note that, within a single VLAN, dynamic "borrowing" of capacity between FSA and non-FSA traffics **may** be performed by the FSA Access QoS functional entity.

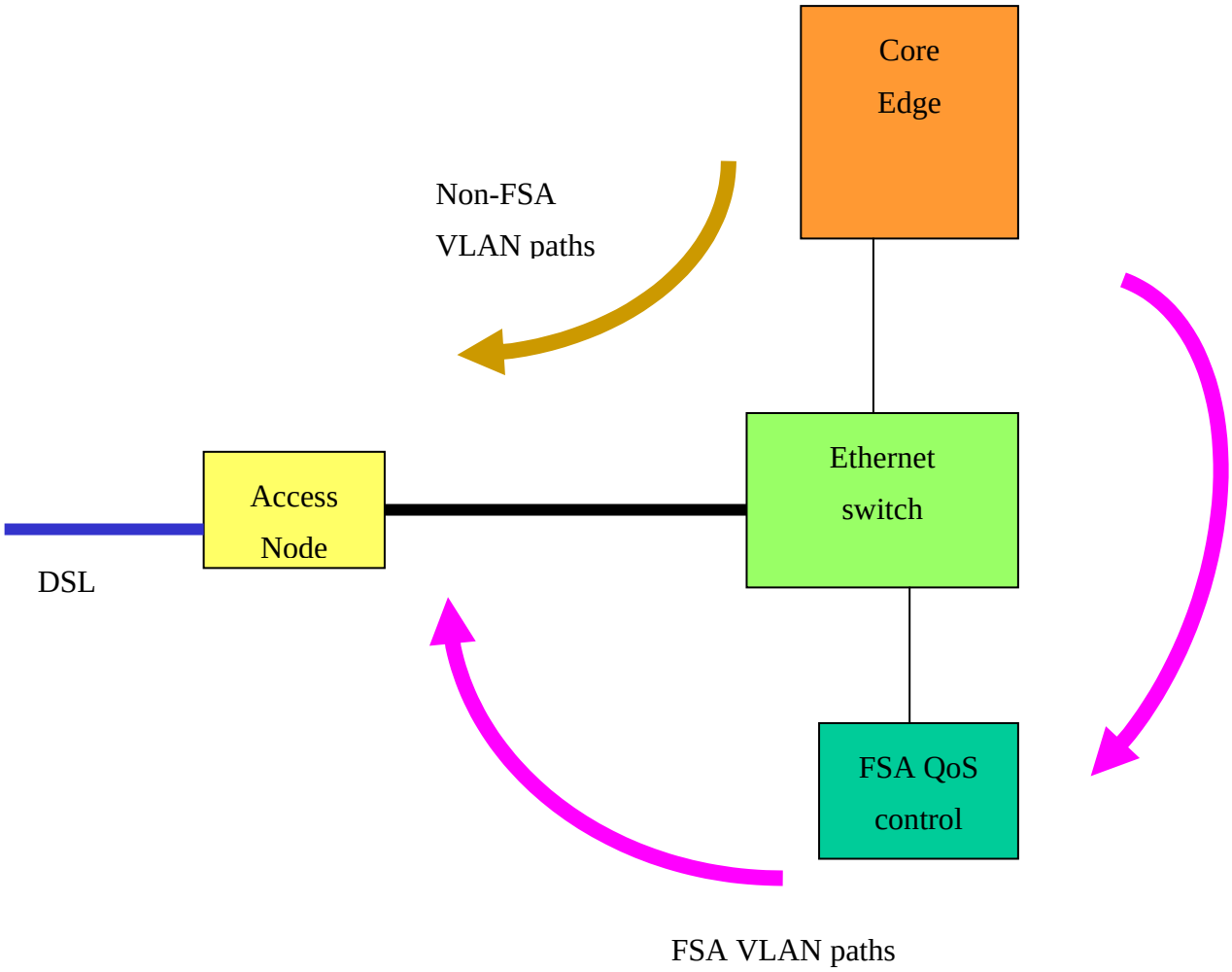
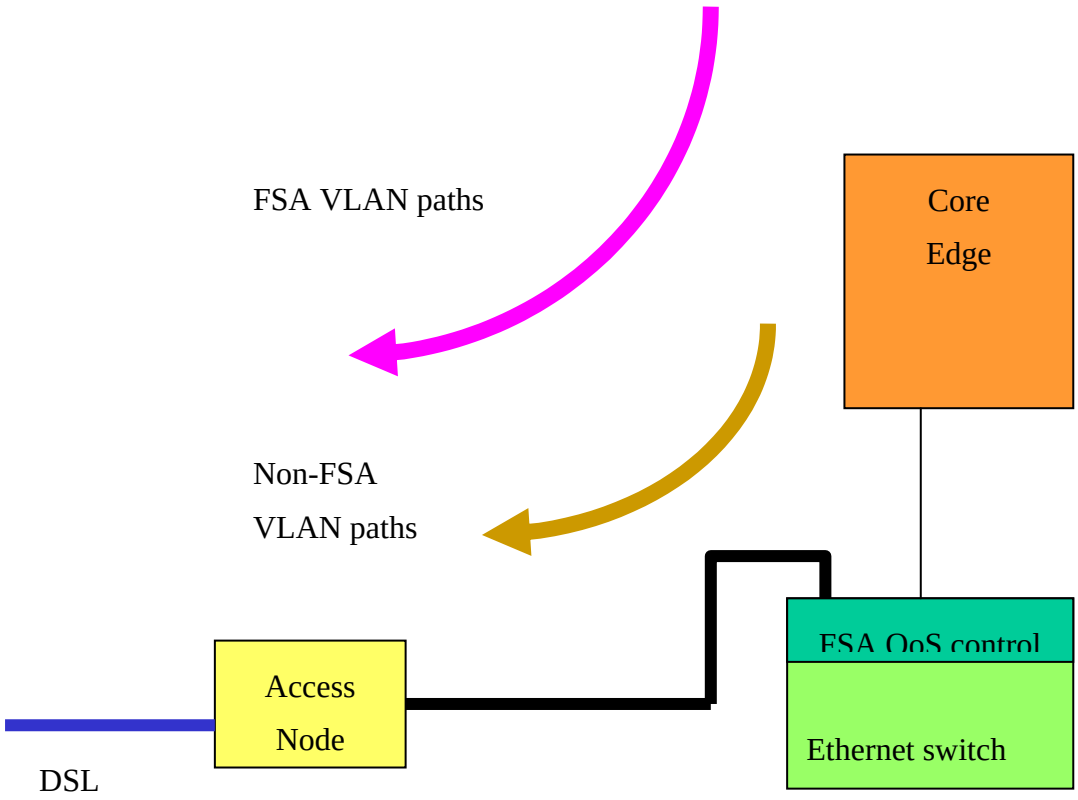


Figure IV.1 FSA control of some of the access VLANs



FigureIV.2 FSA IP-layer inspection and management of all access links

IV.2 General issues on the management of access links shared by FSA and non-FSA traffic

With reference to Figure IV.3 below, an access link is shown with downstream traffic being forwarded on to it from an Edge FSA node. The downstream traffic consists of both FSA and non-FSA components.

Figure IV.3 shows a grooming process that consists of two stages. Stage 1 is the separate grooming of FSA flows (with the non-FSA traffic by-passing this stage). Stage 2 is the grooming of the FSA traffic with the non-FSA traffic.

It will be appreciated that this description of a two-stage process could be realised without actually implementing two physically separate stages. The description has been chosen for clarification of QoS management, and not to suggest any particular implementation.

Similarly, although Figure IV.3 shows both FSA and non-FSA capabilities in an Edge FSA node, this does not suggest an actual implementation. The Stage 1 function could be implemented within the Edge FSA node or external to it.

In more detail, Stage 1 may groom traffic for one or several Ethernet VLANs or ATM VPs according to different policies, including:

- Provider-specific requirements on Preference Priority handling.
- Aggregate maximum rate limit per VLAN or VP and end-user maximum rate limit for downstream traffic.
- A multiple-Provider shared access with an overall maximum rate limit and with equal or preferential policies on AR rate allocation or MR discard first.

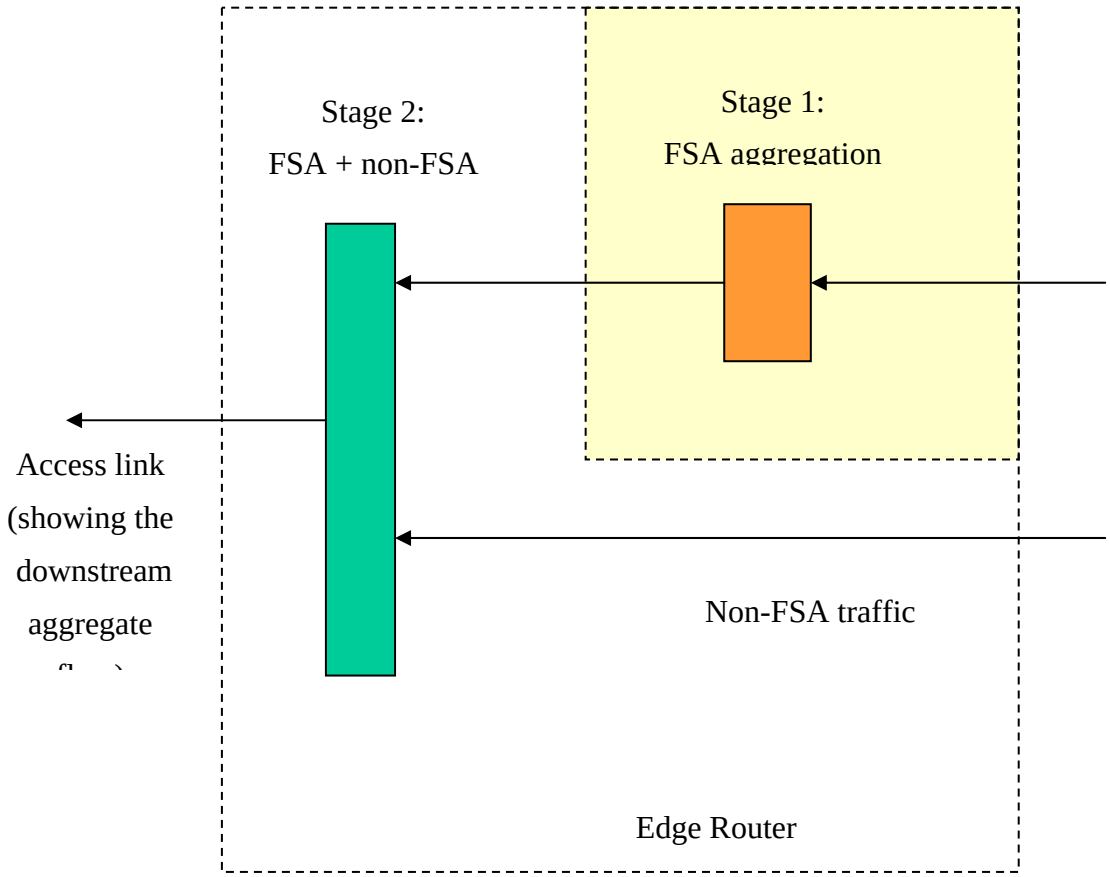


Figure IV.3 Conceptual 2-stage grooming of FSA + non-FSA for a shared access link

As shown in Figure IV.4 below, Stage 1 grooms the FSA traffic into virtual access links corresponding to one or several different access products. In addition, several of these virtual access links may share a single physical access link. The method of sharing may allow:

- Strict adherence to the maximum rate on each of these virtual access links.
- A virtual access link borrows unused physical link capacity (in terms of what is pre-set as the notional physical link capacity assigned for Stage 1). It returns to its guaranteed maximum rate as other virtual links demand more.

In the simplest of these cases (strict adherence to each assigned maximum rate) the Stage 1 grooming process:

- Manages AR rate shares within any virtual link on the basis of the available maximum rate of that virtual link, subtracting any GR or MR discard last rates from the total available.
- Manages MR packet discards to ensure strict adherence to the maximum rate.

In the alternative case, MR packet discards may be reduced if there is some unused physical capacity. An alternative and possibly more complex arrangement would allow AR rates to exploit the unused capacity and rapidly relinquish it as demands from other virtual links increase.

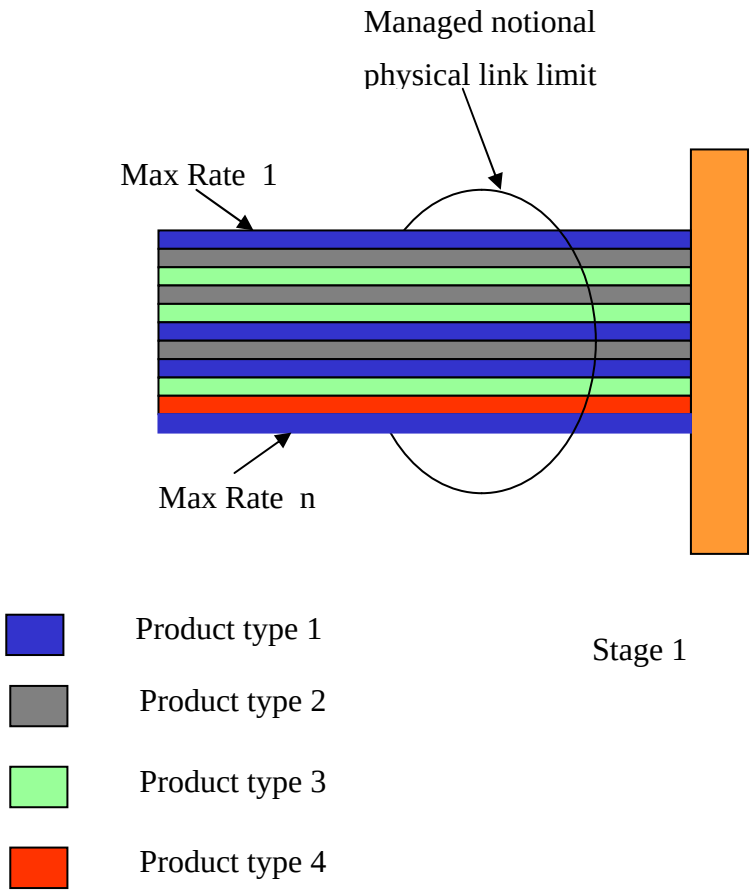


Figure IV.4 FSA traffic groomed into different virtual access links at Stage 1

The Stage 2 grooming process forwards these virtual access links onto a number of physical access links, combining this traffic with the non-FSA traffic. Prior to Stage 2, the non-FSA traffic may also be groomed into one or several virtual access links per physical access link (e.g. associated with one or several Ethernet VLANs).

If the FSA traffic is groomed into Ethernet VLANs at Stage 1 and if the non-FSA traffic has also been groomed into Ethernet VLANs, then Stage 2 grooming could be performed by an Ethernet switch. In this case, for QoS reasons, there should be strict adherence to the maximum rate of each

virtual access link. Where several virtual access links share the same physical access link, the sum of their maximum rates should be managed to be less than the physical link capacity.

On the other hand, it was noted that one type of access product may borrow unused capacity and may temporarily exceed its guaranteed maximum rate. To perform Stage 2 grooming in this case it is necessary to take account of congestion. One option is to perform Stage 2 grooming in the Edge FSA node. Another possibility is to perform this function in an external function that may or may not be FSA capable (if the latter, then this external function may also simultaneously perform Stage 1 – see, for example, Fig IV.2).

IV.2.1. Case 1: FSA-aware Stage 2

Consider Figure IV.5 which shows the case of a FSA-aware Stage 2. It is assumed that any reduction in bandwidth available to a virtual access link at the output of Stage 2, may not immediately cause a reduction in the output load from Stage 1. Therefore, on a FSA virtual access link that is allowed to borrow unused bandwidth above its guaranteed maximum the simplest procedure is:

- Stage 2 allows unused capacity to be used for forwarding packets of this virtual access link, but reduces the aggregate forwarding rate to the guaranteed maximum rate when necessary.
- MR “discard first” packets may be discarded if there are too many packets waiting to be forwarded at Stage 2.
- The sum of the rates of all GR flows added to the sum of MR flows in “discard last” state (including the MR component of VR) should not exceed the guaranteed maximum rate of the virtual access link.
- Stage 2 should reduce AR rates as appropriate when it is necessary to reduce the aggregate rate of the virtual access link to its guaranteed maximum rate.

Therefore Stage 2 supports all guaranteed bandwidth flows (whether in FSA or non-FSA virtual access links). This is provided that the call acceptance process does not allow any guaranteed bandwidth flow to be accepted if it should cause a virtual access maximum rate to be exceeded (and provided the sum of the set of maximum rates does not exceed a physical link rate).

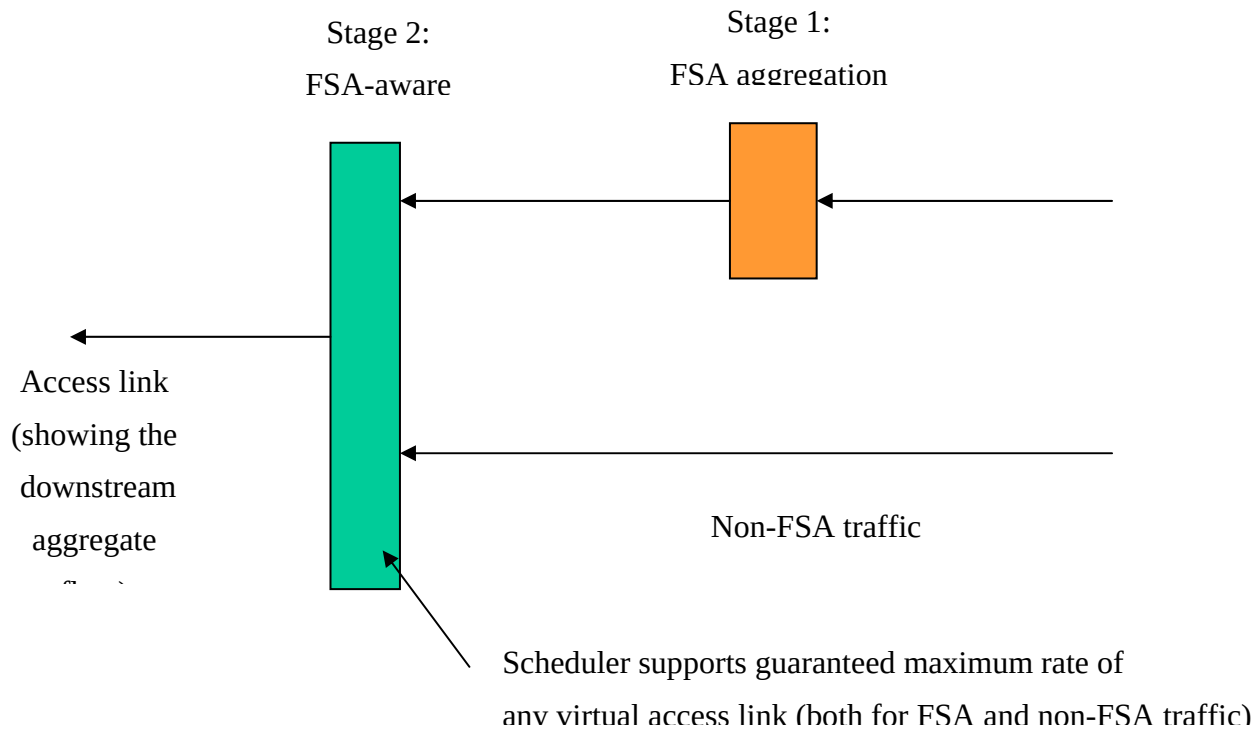


Figure IV.5 Case 1 – FSA-aware Stage

IV.2.2. Case 2: Non FSA-aware Stage 2

This case requires more complexity in the marking of packets so that Stage 2 correctly discards excess packets when it has to limit the virtual access link to its guaranteed maximum rate. One possibility is that MR "discard first" packets are marked differently to MR "discard last" packets (requiring the marking to be changed on a flow if the "discard first" state changes to "discard last"). Then Stage 2 operates as follows:

- Stage 2 allows unused capacity to be used for forwarding packets of this virtual access link, but reduces the aggregate forwarding rate to the guaranteed maximum rate when necessary.
- Packets appropriately marked for discard will be discarded if there are too many packets waiting to be forwarded at Stage 2.
- The sum of the rates of all GR + AR flows added to the sum of VR+MR flows in "discard last" state should not exceed the guaranteed maximum rate of the virtual access link.

This implies that only MR "discard first" packets are exploiting the excess unused capacity on the physical link. With these restrictions the same statement may be made that a non FSA-aware Stage 2 supports all guaranteed bandwidth flows (whether in FSA or non-FSA virtual access links). Again this is provided that the call acceptance process does not allow any guaranteed bandwidth flow to be accepted if it should cause a virtual access maximum rate to be exceeded (and provided the sum of the set of maximum rates does not exceed a physical link rate).

IV.3 Combined flow-level and aggregate flow-level FSA controls

It may be anticipated that Network Providers may want to combine FSA controls at the flow level and aggregate flow level. This creates the opportunity of providing variable-rate aggregation products that are adapted to the needs of the Service Providers and their customers.

The following text describes examples of variable rate aggregation products for the purposes of clarifying requirements only:

- Products that connect between (see Figure IV.6):
  - an access product service edge point, carrying traffic to/from a Service Provider Point of Presence;
  - an access product service edge point, carrying traffic to/ from a given set of end-users..
- Connecting between the service edge points of access products that carry traffic to/ from enterprise sites (see Figure IV.7)
- Assigning capacity to the aggregate based on the latest available rate, where the flow level consists of a set of AR flows only those share this capacity on a basis that recognises Preference Priority (see Figure IV.8).
- Assigning capacity to the aggregate based on a minimum (MR-requested) rate plus an available rate top-up. Here the set of flows can be MR, VR and AR and, again, the capacity sharing of the top-up portion recognises Preference Priority (see Figure IV.9).
- Assigning capacity (perhaps to cover sudden needs) based on a new MR-requested aggregate rate while recognising that this new rate may not be instantly available. However it will be made available as soon as possible and, meanwhile, the policed rate will be adjusted to the new need. Here the flow level consists of a set of MR flows only, and discard control again recognises Preference Priority (see Figure IV.10).
- Assigning capacity based on a new GR-requested aggregate rate, with the possibility that this request may be rejected. Accepted requests will allow flow-level controls to operate and modify, for example, AR rates (again taking Preference Priority into account). This is shown in Figure IV.11.

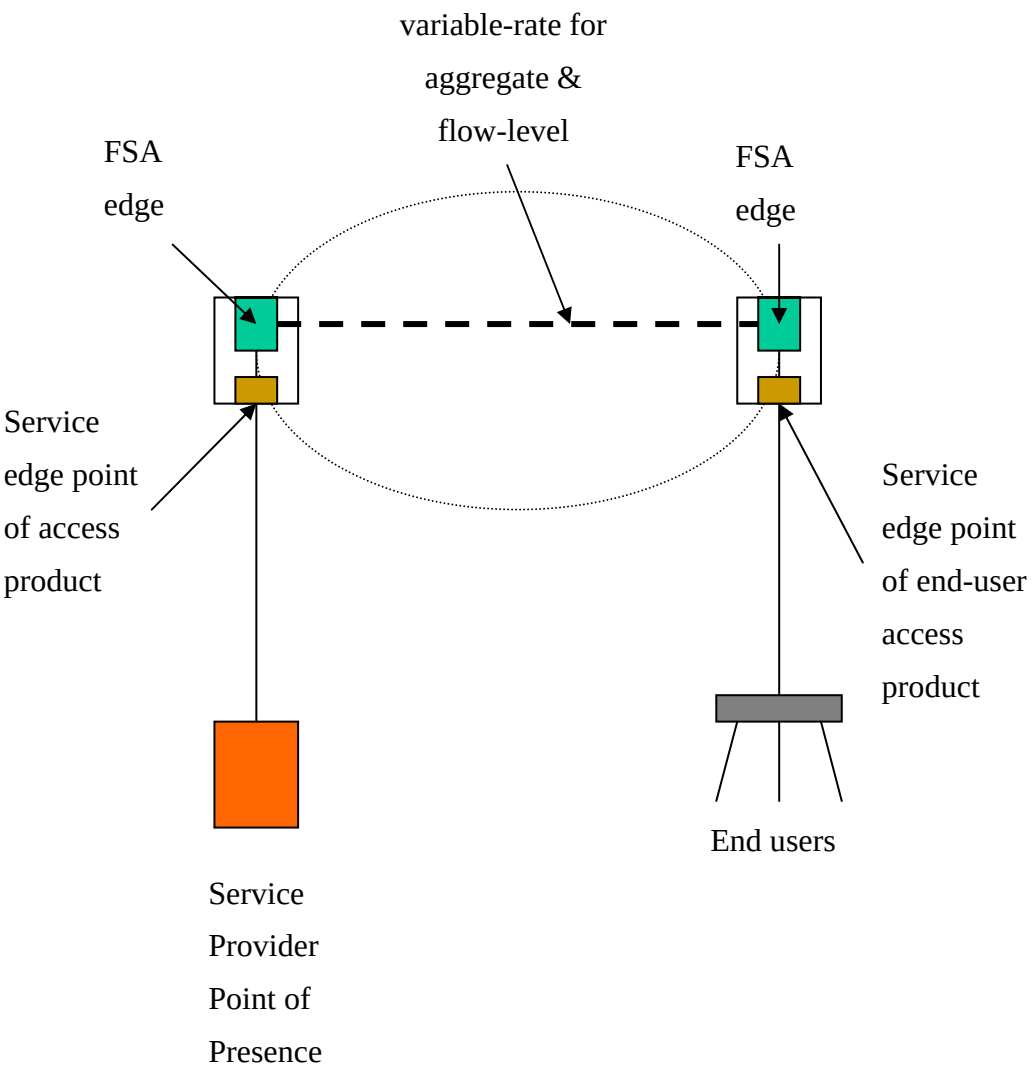


Figure IV.6 FSA variable capacity controls supporting the connection between a Service Provider and specified end users

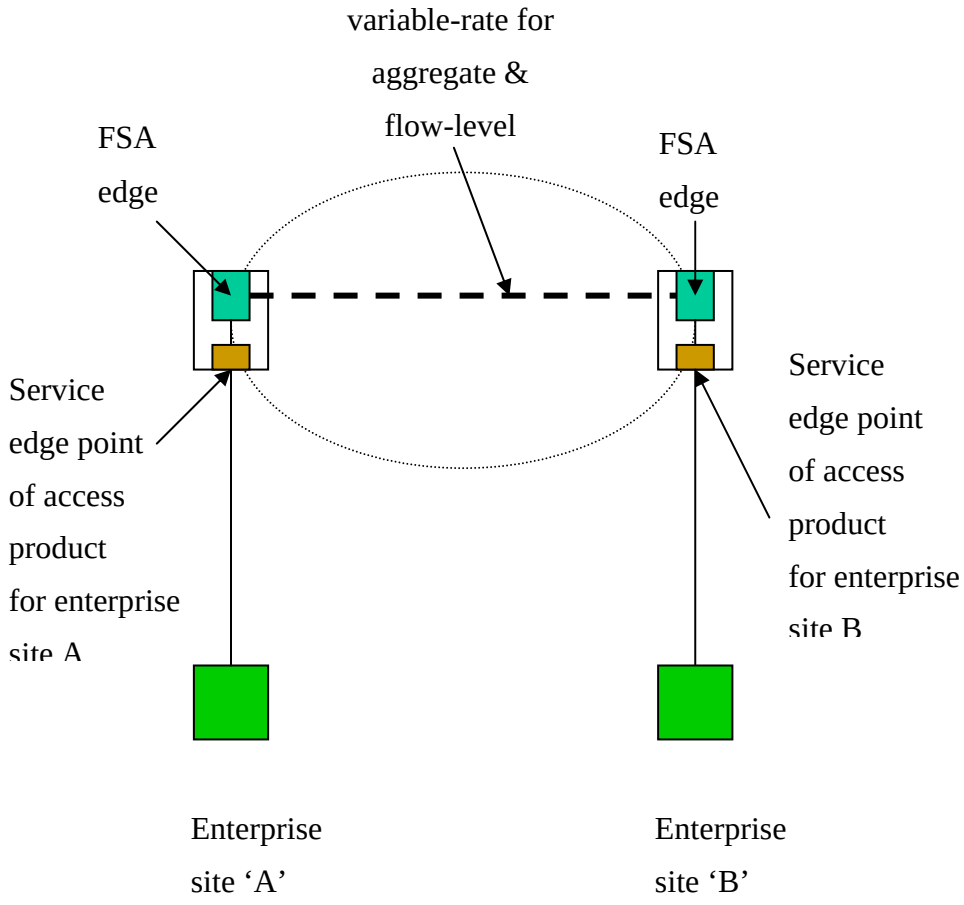


Figure IV.7 Enterprise site-to-site FSA controlled variable capacity

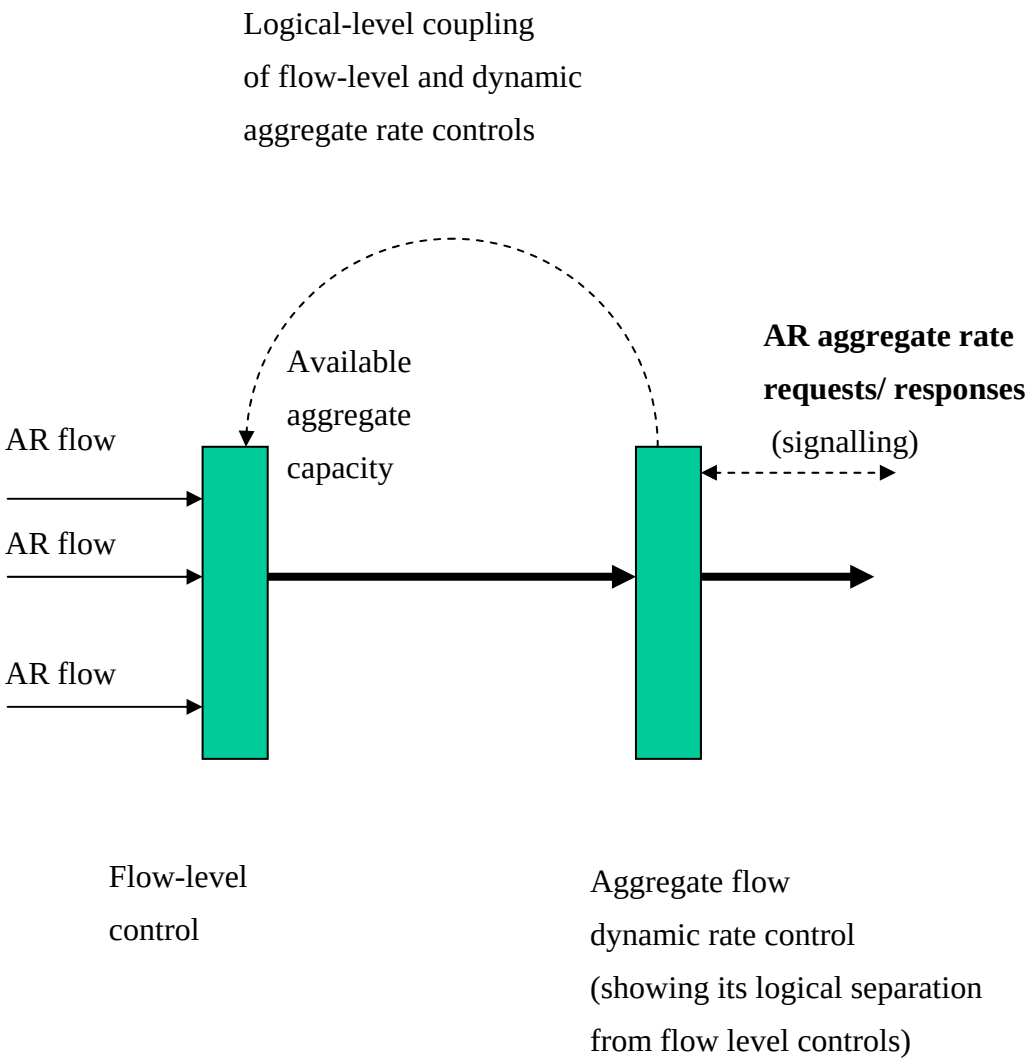


Figure IV.8 AR flows within an AR aggregate access

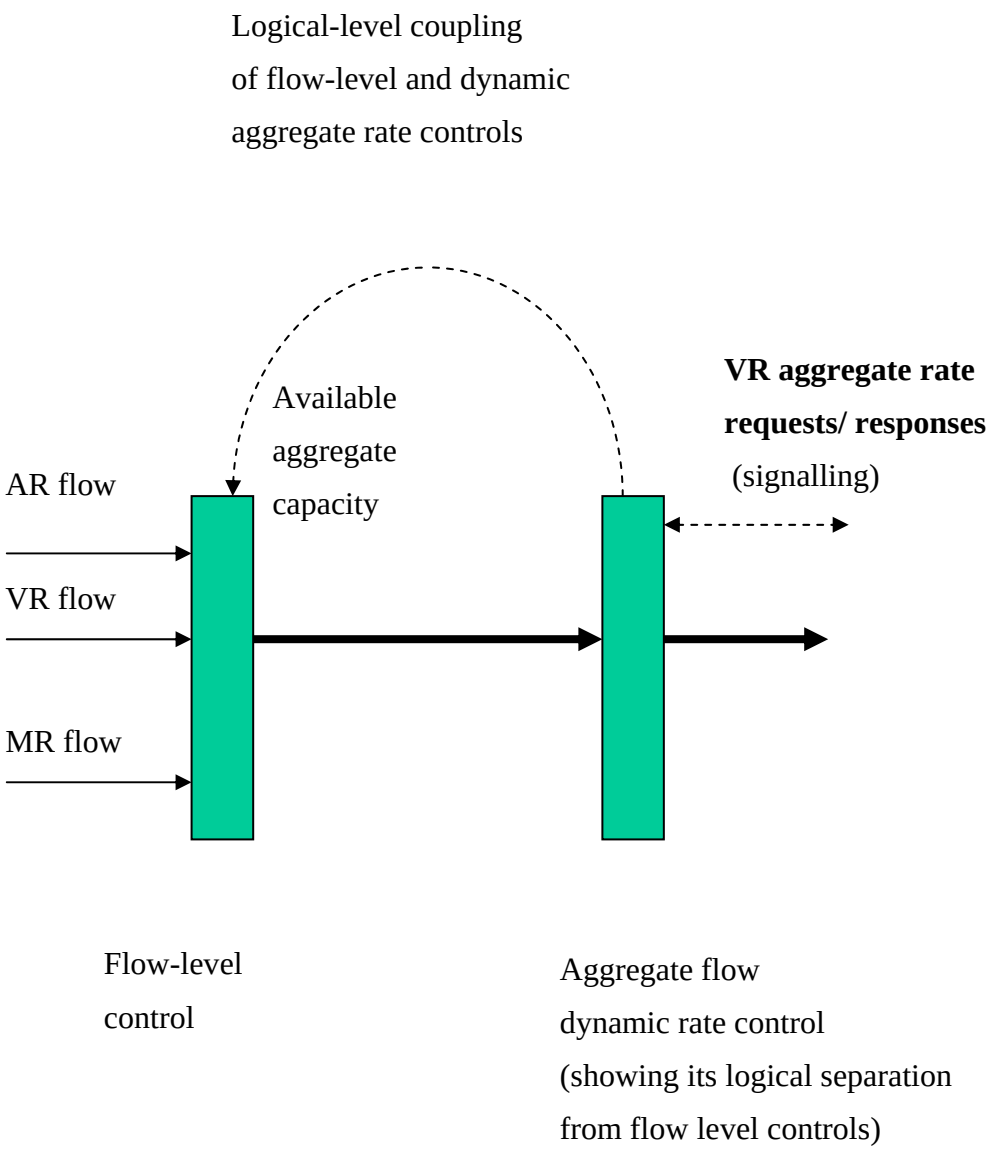


Figure IV.9 AR, VR and MR flows within a VR aggregate access

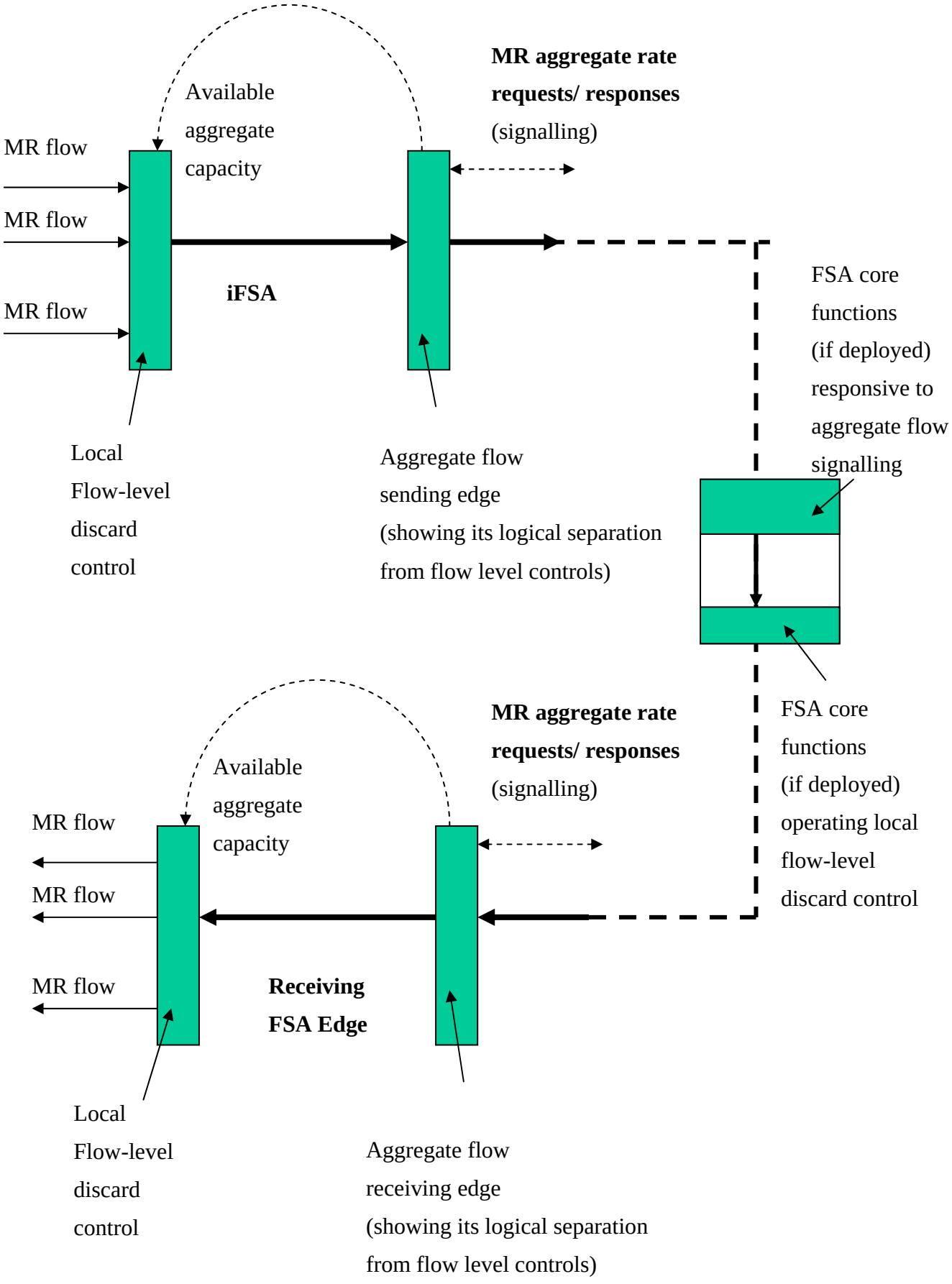

Logical-level coupling
of flow-level and dynamic
aggregate rate controls
Available
aggregate
capacity
MR aggregate rate
requests/ responses
(signalling)
MR flow
MR flow
MR flow
iFSA
FSA core
functions
(if deployed)
responsive to
aggregate flow
signalling
Local
Flow-level
discard
control
Aggregate flow
sending edge
(showing its logical separation
from flow level controls)
MR aggregate rate
requests/ responses
(signalling)
Available
aggregate
capacity
FSA core
functions
(if deployed)
operating local
flow-level
discard control
MR flow
MR flow
MR flow
Receiving
FSA Edge
Local
Flow-level
discard
control
Aggregate flow
receiving edge
(showing its logical separation
from flow level controls)

Figure IV.10 MR flows within an MR aggregate

It may be noted in Figure IV.10 that each FSA node is making a local decision about the rate that could be supported on reception of a MR aggregate rate increase request. Unlike AR, the sending end may immediately send at the new requested aggregate rate. However, network providers may choose to prevent any sending node from increasing a rate beyond some pre-determined maximum rate or beyond some pre-determined maximum allowable percentage increase above a current rate. Other policy options may be applied.

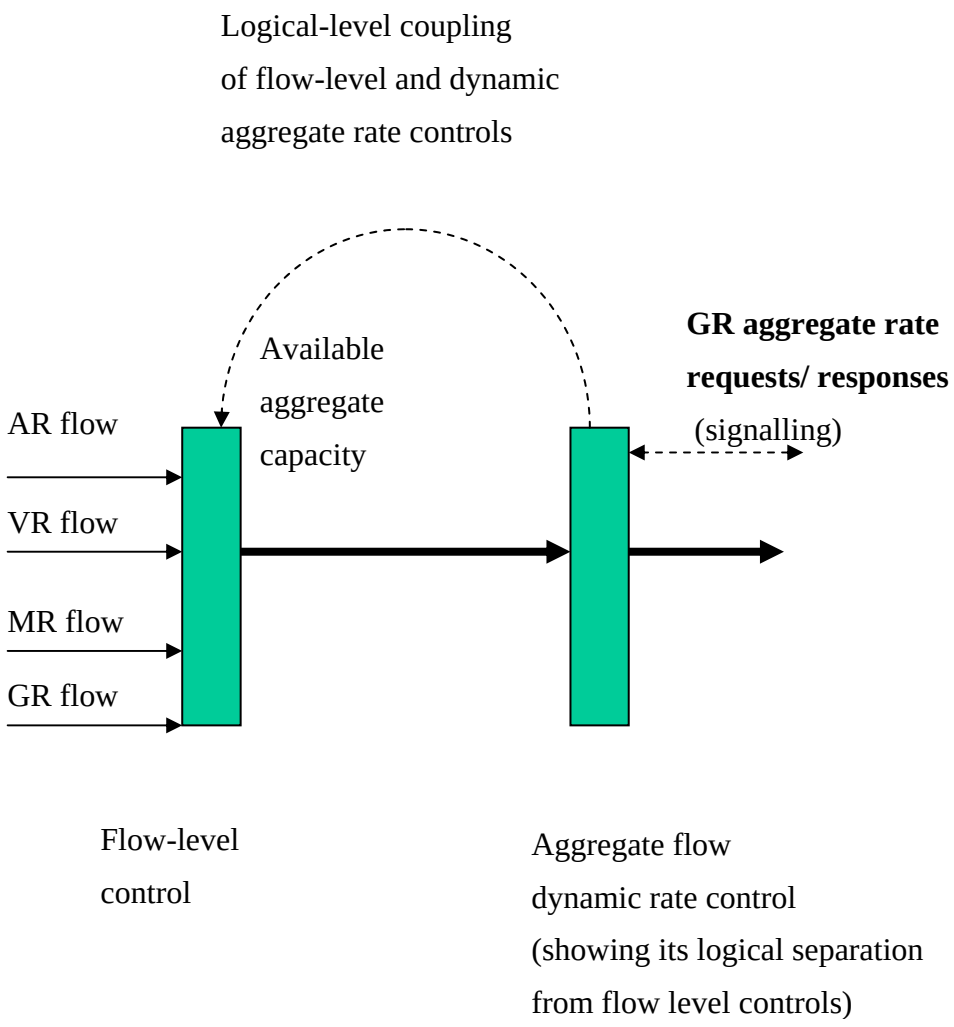


Figure IV.11 GR negotiated aggregate rate changes

Appendix V: Example implementation principles associated with FSA nodes

This appendix does not form an integral part of this Recommendation.

The principles below describe an example implementation, not the only possible implementation of a FSA node:

1. The FSA node determines the identifier of each QoS signalled flow.
2. Having identified a flow, some memory of the flow is required. This memory may be as little as a saved hash entry and a bit to indicate if the flow is “discard last” or discard first”. It may also contain the QoS information for the flow that was in the initial in-band signalling packet.
3. The FSA node also needs to be monitoring the loading of the port on which the flow is exiting the FSA node. From this load information it must determine what capacity is available for a new flow. There will be many techniques for doing this and deciding what rate an AR flow may have or if a MR flow should be “discard first”, “discard last”, or must be denied altogether.
4. The FSA node only needs to route this first packet of the flow to determine what route it should take, what QoS it should have, and if it is subject to Denial of Service (DOS).
5. Once a flow has been accepted at a rate, if the rate is saved in the state information, then the flow may be policed to that rate. If no rate information is kept, the policing could be based on the total port load, and for AR this plus the number of flows may be sufficient. For GR the rate will be required since the total commitment must be controlled. For MR and the MR portion of VR the rate would need to be saved if per flow policing is required, but if only flow acceptance is needed then the “discard first” bit would be sufficient.
6. The flow entries need to be cleared out if no packets arrive for a period thus there also needs to be a time stamp for the last packet arrival.

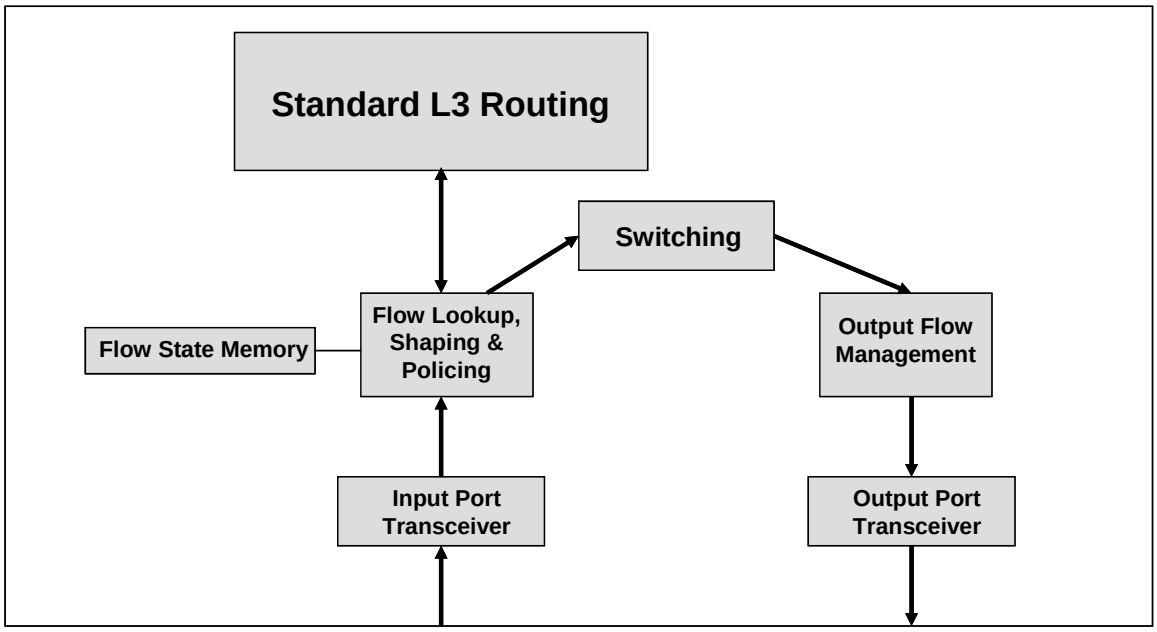


Figure V.1 Example implementation illustrating a FSA node

Figure V.1 shows an example implementation of a FSA node where the data flow follows the path through the switching logic and the first packet is routed with normal L3 FSA node logic.

Once the packet is identified with a flow record which includes the agreed rate, the flow can be policed and thus the total output load on a port controlled to any desired loading. New flows can be accepted by examining the remaining capacity and assigning a rate to the flow. MR flows may be controlled not to exceed some load limit but if too many arrive, they could be accepted as “discard first” or rejected if necessary. Thus the ability to support the envisioned QoS signalling protocol only requires a modest memory and logic capability which need not add any significant cost to the FSA node, and in fact save cost by reducing the routing logic.

Appendix VI: Out-of-band Signalling with a Central Admission Entity

This appendix does not form an integral part of this Recommendation.

Under assumption that a central admission entity such as RACF performs the admission function accurately, the messaging overhead for the response and the confirmation is not required. This we call the simplified FSA signalling.

The simplified FSA signalling can further benefit by combining with the proxy signalling. In this case any registered CPE for FSA treatment, without having the FSA signalling capability, can simply initiate a call. Then through the call-level authorization and admission process, the IP-level traffic parameters (e.g. RACF Traffic descriptor) are notified to the RACF. The RACF then distributes the Traffic descriptor to the appropriate ingress edge node, which will be the signalling proxy. The proxy maps the IP-level traffic descriptor distributed by the RACF to the FSA parameters, prior to sending the Start packet. The CPEs then do not have to register the individual FSA parameters to the proxy. The simplified signalling procedure is depicted in Figure VI.1.

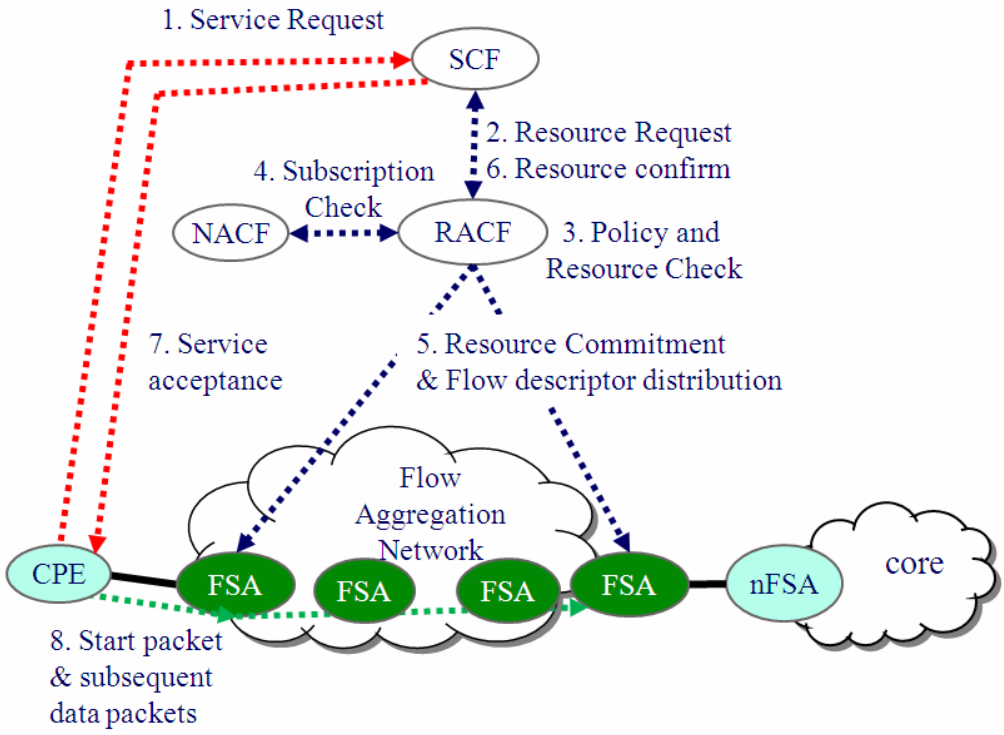


Figure VI. 1 Simplified signalling procedure

1. A CPE requests a service for a call.
2. The SCF requests for available resource to the RACF.
3. The RACF checks the applicable policy to the call, thus to the flow, then check for the available resources in the network.
4. The RACF also checks for the subscription status.
5. Upon passing the checks, the RACF decides to admit and commit resource to the call. The RACF also distributes the IP level flow descriptor. If any of ingress edge nodes work as a signalling proxy, then such an edge node maps the IP level flow descriptor into the FSA

QoS parameters such as Service Context, Burst Tolerance, and Delay Priority. Those FSA parameters are stored in the ingress edge node.

6. The RACF then confirms the available resource.
7. The SCF notifies the CPE of the service acceptance.

The CPE generates the in-band Start packet that corresponds to the FSA Request signal. Every FSA in the path recognizes it, and stores the FSA parameters for the flow. If the ingress edge node performs as a proxy, the CPE just transmits the ordinary data packet. The ingress edge node realizes the first data packet sent from the CPE and generates the Start packet using the FSA parameter mapped from the IP level flow descriptor.

It is likely that the central admission entity examines only the requested data rate of a flow, which corresponds to the Requested Rate in FSA. Other parameters such as Preference priority, Service context, Burst tolerance, and Delay priority can be inferred from the IP level flow descriptor received from the central admission entity.

In summary, the negotiation procedure can be through the complete in-band signalling (including the Response and the Confirmation) or the in-band signalling plus the authorization signalling to the central admission entity. In the latter case the Start packet can be generated at the ingress edge node by mapping the central admission entity IP-level descriptor (such as RACF Traffic Descriptor) to FSA parameters.

BIBLIOGRAPHY

[b-ITU-T I.371] ITU-T Recommendation I.371 (2004), *Traffic control and congestion control in B-ISDN*

[b-RFC 2003] IETF RFC 2003 (1996), *IP encapsulation within IP*

[b-RFC 2748] IETF RFC 2748 (2000), *The COPS (Common Open Policy Service) Protocol*

[b-RFC 2868] IETF RFC 2868 (2000), *RADIUS Attributes for Tunnel Protocol Support*

[b-RFC 3270] IETF RFC 3270 (2002), *Multi-Protocol Label Switching (MPLS) Support of Differentiated Services*