Internet Engineering Task Force
INTERNET DRAFT
Expires May 1999

Juha Heinanen
Telia Finland
Fred Baker
Cisco Systems
Walter Weiss
Lucent Technologies
John Wroclawski
MIT LCS
November, 1998

Assured Forwarding PHB Group

<draft-ietf-diffserv-af-03.txt>

Status of this Memo

Copyright Notice

Abstract

This document defines a general use Differentiated Services (DS) [Blake] Per-Hop-Behavior (PHB) Group called Assured Forwarding (AF). The AF PHB group provides delivery of IP packets in four independently forwarded AF classes. Within each AF class, an IP packet can be assigned one of three different levels of drop precedence. A DS node does not reorder IP packets of the same microflow if they belong to the same AF class.

1. Purpose and Overview

There is a demand to provide assured forwarding of IP packets over the Internet. In a typical application, a company uses the Internet to interconnect its geographically distributed sites and wants an assurance that IP packets within this intranet are forwarded with high probability as long as the aggregate traffic from each site does not exceed the subscribed information rate (profile). It is desirable that a site may exceed the subscribed profile with the understanding that the excess traffic is not delivered with as high probability as the traffic that is within the profile. It is also important that the network does not reorder packets that belong to the same microflow no matter if they are in or out of the profile.

Assured Forwarding (AF) PHB group is a means for a provider DS domain to offer different levels of forwarding assurances for IP packets received from a customer DS domain. Four AF classes are defined, where each AF class is in each DS node allocated a certain amount of forwarding resources (buffer space and bandwidth). IP packets that wish to use the services provided by the AF PHB group are assigned by the customer or the provider DS domain into one or more of these AF classes according to the services that the customer has subscribed to.

Within each AF class IP packets are marked (again by the customer or the provider DS domain) with one of three possible drop precedence values. In case of congestion, the drop precedence of a packet determines the relative importance of the packet within the AF class. A congested DS node tries to protect packets with a lower drop precedence value from being lost by preferably discarding packets with a higher drop precedence value.

In a DS node, the level of forwarding assurance of an IP packet thus depends on (1) how much forwarding resources has been allocated to the AF class that the packet belongs to, (2) what is the current load of the AF class, and, in case of congestion, (3) what is the drop precedence of the packet.

For example, if traffic conditioning actions at the ingress of the provider DS domain make sure that an AF class in the DS nodes is only moderately loaded by packets with the lowest drop precedence value and is not overloaded by packets with the two lowest drop precedence values, then the AF class can offer a high level of forwarding assurance for packets that are within the subscribed profile and offer up to two lower levels of forwarding assurance for the excess traffic.

This document describes the AF PHB group. An otherwise DS-compliant

node is not required to implement this PHB group in order to be considered DS-compliant, but when a DS-compliant node is said to implement an AF PHB group, it must conform to the specification in this document.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [Bradner].

2. The AF PHB Group

Assured Forwarding (AF) PHB group provides forwarding of IP packets in N independent AF classes. Within each AF class, an IP packet is assigned one of M different levels of drop precedence. An IP packet that belongs to an AF class i and has drop precedence j is marked with the AF codepoint AFij, where 1 <= i <= N and 1 <= j <= M. Currently, four classes (N=4) with three levels of drop precedence in each class (M=3) are defined for general use. More AF classes or levels of drop precedence MAY be defined for local use.

A DS node MUST allocate forwarding resources (buffer space and bandwidth) to AF classes so that, under reasonable operating conditions and traffic loads, packets of an AF class x do not have higher probability of timely forwarding than packets of an AF class y if x < y. Similarly, within an AF class, an IP packet with drop precedence p MUST NOT be forwarded with smaller probability than an IP packet with drop precedence q if p < q.

A DS node MUST NOT reorder AF packets of the same microflow when they belong to the same AF class regardless of their drop precedence. There are no quantifiable timing requirements (delay or delay variation) associated with the forwarding of AF packets.

The AF PHB group MAY be used to implement both end-to-end and domain edge-to-domain edge services.

3. Traffic Conditioning Actions

A DS domain MAY at the edge of a domain control the amount of AF traffic that enters or exists the domain at various levels of drop precedence. Such traffic conditioning actions MAY include traffic shaping, discarding of packets, increasing or decreasing the drop precedence of packets, and reassigning of packets to other AF classes. The traffic conditioning actions MUST NOT cause reordering of packets of the same microflow.

4. Queueing and Discard Behavior

A DS node SHOULD implement all four general use AF classes. Packets in one AF class MUST be forwarded independently from packets in another AF class, i.e., a DS node MUST NOT aggregate two or more AF classes together.

Within each AF class, a DS node MUST accept all three drop precedence codepoints and they MUST yield at least two different levels of loss probability. In some networks, particularly in enterprise networks, where transient congestion is a rare and brief occurrence, it may be reasonable for a DS node to implement only two different levels of loss probability. While this may suffice for some networks, three different levels of loss probability SHOULD be supported in DS domains where congestion is a common occurrence.

If a DS node only implements two different levels of loss probability for an AF class x, the codepoint AFx1 MUST yield the lower loss probability and the codepoints AFx2 and AFx3 MUST yield the higher loss probability.

Inconsistent discard behaviors lead to inconsistent end-to-end service semantics. It is RECOMMENDED that the discard mechanism is based on a RED-like [Floyd] algorithm. In any case, the discard control parameters for each precedence within an AF class MUST be separately configurable. In the case of the RED algorithm, this means that the start-drop and hard-drop thresholds for each precedence within a class must be separately configurable. Future versions of this document may say more about specific aspects of the desirable behavior.

5. Tunneling

When AF packets are tunneled, the PHB of the tunneling packet MUST NOT reduce the forwarding assurance of the tunneled AF packet nor cause reordering of AF packets belonging to the same microflow.

6. Recommended Codepoints

It is RECOMMENDED that the AF codepoints AF11, AF21, AF31, and AF41, i.e., the codepoints that denote the lowest drop precedence in each AF class, are mapped to the Class Selector [Nichols] codepoints '010000', '011000', '100000', '101000'. This is done in order to save DS code space, because the forwarding rules associated with these AF codepoints are consistent and compatible with the forwarding rules of the corresponding Class Selector codepoints.

The RECOMMENDED values of the remaining AF codepoints are as follows:

AF12 = '010010', AF13 = '010100', AF22 = '011010', AF23 = '011100', AF32 = '100010', AF33 = '100100', AF42 = '101010', and AF43 = '101100'. The table below summarizes the recommended AF codepoint values.

	Class 1	Class 2	Class 3	Class 4
Low Drop Prec	010000	011000	100000	101000
Medium Drop Prec	010010	011010	100010	101010
High Drop Prec	010100	011100	100100	101100

7. Interactions with Other PHB Groups

The AF codepoint mappings recommended above do not interfere with the local use spaces nor use the Class Selector codepoints '00x000' and '11x000'. The PHBs selected by those Class Selector codepoints may thus coexist with the AF PHB group, and retain the forwarding behavior and relationships that was defined for them in [Nichols]. In particular, the Default PHB codepoint of '000000' may remain to be used for conventional best effort traffic. Similarly, the codepoints '11x000' may remain to be used for network control traffic.

In addition to the Class Selector PHBs, any other PHB groups may co-exist with the AF group within the same DS domain provided that the other PHB groups don't preempt the resources allocated to the AF classes.

8. Security Implications

In order to protect itself against denial of service attacks, a provider DS domain SHOULD limit the traffic entering the domain to the subscribed profiles. Also, in order to protect a link to a customer DS domain from denial of service attacks, the provider DS domain SHOULD allow the customer DS domain to specify how the resources of the link are allocated to AF packets. If a service offering requires that traffic marked with an AF codepoint be limited by such attributes as source or destination address, it is the responsibility of the ingress node in a network to verify validity of such attributes.

Other security considerations are covered in [Blake] and [Nichols].

Appendix: Example Services

The AF PHB group could be used to implement, for example, the so-called Olympic service, which consists of three service classes: bronze, silver, and gold. Packets are assigned to these three

classes so that packets in the gold class experience lighter load (and thus have greater probability for timely forwarding) than packets assigned to the silver class. Same kind of relationship exists between the silver class and the bronze class. If desired, packets within each class may be further separated by giving them either low, medium, or high drop precedence.

The bronze, silver, and gold service classes could in the network be mapped to the AF classes 1, 2, and 3. Similarly, low, medium, and high drop precedence may be mapped to AF drop precedence levels 1, 2, or 3.

The drop precedence level of a packet could be assigned, for example, by using a leaky bucket traffic policer, which has as its parameters a rate and two burst sizes: a committed burst and an excess burst. If a packet falls within the committed burst, it is assigned low drop precedence. If a packet falls between the committed burst and the excess burst, it is assigned medium drop precedence. And finally, if the packet falls out of the excess burst, it is assigned high drop precedence. It may also be necessary to set an upper limit to the amount of high drop precedence traffic from a customer DS domain in order to avoid the situation where an avalanche of undeliverable high drop precedence packets from one customer DS domain can deny service to possibly deliverable high drop precedence packets from other domains.

Another way to assign the drop precedence level of a packet could be to limit the user traffic of an Olympic service class to a given peak rate and distribute it evenly across each level of drop precedence. This would yield a proportional bandwidth service, which equally apportions available capacity during times of congestion under the assumption that customers with high bandwidth microflows have subscribed to higher peak rates than customers with low bandwidth microflows.

The AF PHB group could also be used to implement a low loss, low delay, and low jitter service using an over provisioned AF class, if the maximum arrival rate to that class is known a priori in each DS node. Specification of the required admission control services, however, is beyond the scope of this document.

References

[Blake] Blake, Steve, et al., An Architecture for Differentiated Services. Internet draft draft-ietf-diffserv-arch-01.txt, August 1998.

[Bradner] Bradner, S., Key words for use in RFCs to Indicate

Requirement Levels. Internet RFC 2119, March 1997.

[Floyd] Floyd, S., and Jacobson, V., Random Early Detection gateways for Congestion Avoidance. IEEE/ACM Transactions on Networking, Volume 1, Number 4, August 1993, pp. 397-413.

[Nichols] Nichols, Kathleen, et al., Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. Internet draft draft-ietf-diffserv-header-02.txt, August 1998.

Author Information

Juha Heinanen
Telia Finland
Myyrmaentie 2
01600 Vantaa, Finland
Email: jh@telia.fi

Fred Baker
Cisco Systems
519 Lado Drive
Santa Barbara, California 93111
E-mail: fred@cisco.com

Walter Weiss
Lucent Technologies
300 Baker Avenue, Suite 100,
Concord, MA 01742-2168
E-mail: wweiss@lucent.com

John Wroclawski
MIT Laboratory for Computer Science
545 Technology Sq.
Cambridge, MA 02139
Email: jtw@lcs.mit.edu

Full Copyright Statement