TSVWG Working Group
Internet-Draft
Intended status: Informational
Expires: May 7, 2020

L. Han
Y. Qu
R. Li
Futurewei Technologies
November 4, 2019

Enhanced DiffServ by In-band Signaling
draft-han-tsvwg-enhanced-diffserv-00

Abstract

There are real-time applications requiring deterministic services in terms of bandwidth and latency in various industries, such as network based AR/VR (Augmented Reality and Virtual Reality), industrial internet. This document proposes a solution which enhances DiffServ in IPv6 to provide end to end bandwidth guaranteed service and latency guaranteed service.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

[RFC8557][RFC8578] identifies some use cases from different industries which have a common need for "deterministic flows". Such flows require guaranteed bandwidth and bounded latency.

This document proposes to use in-band signaling together with DiffServ to provide Latency Guaranteed Service and Bandwidth Guaranteed Service. Network traffic is classified according to application requirements and resources such as bandwidth and buffer are dedicated to provide a reliable and scalable service, while unused reserved resources could be used by best effort traffic.

The solution is compatible with existing Quality-of-Service (QoS) mechanism as long as there is enough network resources, and is not restricted to network topologies. Resource Reservation is done by in-band signaling [I-D.han-tsvwg-ip-transport-qos], and IP path can be achieved using existing routing protocols (IGP or BPG), or the explicit path routing such as segment routing [RFC8402].

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] RFC 8174 [RFC8174] when, and only when, they appear in all capitals, as shown here.

Abbreviations used in this documents:

E2E
: End-to-end.

EH
: IPv6 Extension Header or Extension Option.

QoS
: Quality of Service.

OAM
: Operation and Management.

In-band Signaling
: In telecommunications, in-band signaling is sending control information within the same band or channel used for voice or video.

Out-of-band Signaling
: out-of-band signaling is that the control information sent over a different channel, or even over a separate network.

IP flow
: For non-IPSec, an IP flow is identified by the source, destination IP address, the protocol number, the source and destination port number.

IP path
: An IP path is the route that IP flow will traverse. It could be the shortest path determined by routing protocols (IGP or BPG), or the explicit path such as segment routing [RFC8402].

QoS channel
: A forwarding channel that is QoS guaranteed. It provides additional QoS service to IP forwarding. A QoS channel can be

used for one or multiple IP flows depending on the granularity of in-band signaling.

srTCM
: Single Rate Three Color Marker [RFC2697]

trTCM
: Two Rate Three Color Marker [RFC4115]

LGS
: Latency Guaranteed Service

BGS
: Bandwidth Guaranteed Service

BES
: Best Effort Service

CIR
: Committed Information Rate.

PIR
: Peak Information Rate.

HbH-EH
: IPv6 Hop-by-Hop Extension Header.

Dst-EH
: IPv6 Destination Extension Header.

HbH-EH-aware node
: Network nodes that are configured to process the IPv6 Hop-by-Hop Extension Header.

3. DiffServ and IntServ

3.1. Differentiated Services

Differentiated services or DiffServ [RFC2475] provides a simple, scalable and coarse-grained mechanism to support quality of services on IP networks by classifying and managing network traffic.

DiffServ uses DiffServ Code Point (DSCP) which is the first 6 bits of the TOS field [RFC2474] to classify packets when they enter the network and then are separated into different queues. With DiffServ each router handles each packet differently. Per-Hop Forwarding Behavior (PHB) is a way of forwarding a particular flow or group of flows marked with a particular DSCP value using a particular method on a DiffServ node.

3.2. Integrated Services

IntServ [RFC3175] or integrated services specifies more fine-grained QoS, which is often contrasted with DiffServ’s coarse-grained control system. IntServ definitely can support the applications requiring special QoS guarantee if it is deployed in a network, supported by Host OS and integrated with application. However, IntServ works on a small-scale only. When you scale up the network, it is difficult to keep track of all of the reservations and session states. Thus, IntServ is not scalable. Another problem of IntServ is it is not application driven, tedious provisioning cross different network must be done earlier. The provisioning is slow and hard to maintain.

4. In-band Signaling

In-band signaling messages are carried along with the payload, and it is guaranteed that the signaling follows the same path as the data flow.

In-band QoS signaling for IPv6 was discussed by Lawrence Roberts in 2005 [I-D.roberts-inband-qos-ipv6]. The requirements of IP in-band signaling was proposed by Jon Haper in 2007 [I-D.harper-inband-signalling-requirements]. Telecommunications Industry Association (TIA) published a standard for "QoS Signaling for IP QoS Support and Sender Authentication" in 2006 [TIA].

This draft uses the in-band signaling method proposed by [I-D.han-tsvwg-ip-transport-qos], which offers a solution in IPv6 for QoS service using in-band signaling to guarantee certain level of service quality. Advantages of using in-band signaling and what parameters are carried etc. are discussed in great details. Please refer to [I-D.han-tsvwg-ip-transport-qos] for signaling details, which is not in the scope of this document.

5. Design Targets

The design targets of the proposal are outlined as follows:

1. Service Guarantee:

    * Provide guaranteed and minimized (queuing) latency for LGS (Latency Guaranteed Service) flows. The maximum latency is guaranteed, minimized and predictable at each hop, if LGS flow rate is confirmed with their pre-claimed parameters (CIR,PIR,CBS,EBS).

   * Provide guaranteed bandwidth for BGS (Bandwidth Guaranteed Service) flows. The bandwidth of CIR is guaranteed at each hop, if BGS flow rate is confirmed with their pre-claimed parameters (CIR,PIR,CBS,EBS).

2. No Starvation: LGS flows will never starve other types of lower priority flows (BGS and BES). BGS flows will never starve BE flows.

3. No sacrifice of link utilization: When the total rate of LGS flow is less than the committed rate (sum of CIR of all LGS flows), other class flows (BGS and BES) can use the remained bandwidth. When the total rate of LGS flow is less than the committed rate (sum of CIR of all LGS flows), and, the total rate of BGS flow is less than the committed rate (sum of CIR of all BGS flows), other class flows (BES) can use the remained bandwidth.

4. Fairness within the same class: All LGS flows will share the bandwidth within its class. All BGS flows will share the bandwidth within its class. All BE flows will share the bandwidth within BE class.

5. Backward compatible with current DiffServ, and can coexist and can be deployed incrementally.

6. Processing Procedure

6.1. Introduction

The section will describe more details about the in-band signaling integrated with DiffServ to achieve the guaranteed service. It is composed of following contents:

1. The purpose of in-band signaling for enhanced DiffServ.

2. The info carried in the signaling.

3. The admission process with in-band signaling.

4. User traffic classification.

5. User traffic classification and New DSCP class.

6. Queuing and Scheduling Consideration.

7. Class Based Traffic Shaper.

The document [Enhanced_DSCP] has more details for the experiments using the proposal to achieve the latency guaranteed service. The theoretical formula to estimate the maximum latency is given. Different tests with different combination of traffic are done. Results are analyzed and further researches are discussed.

6.2. The purpose of in-band signaling

The in-band signaling procedures described in [I-D.han-tsvwg-ip-transport-qos] is to make network support new services. There are two types of new services provided:

1. Latency Guaranteed Service (LGS): Provide guaranteed and minimized queuing latency. The maximum latency of a LGS flow is guaranteed, minimized and predictable at each hop if the LGS flow rate is confirmed with its pre-claimed parameters (CIR, PIR, CBS, EBS).

2. Bandwidth Guaranteed Service (BGS): provide guaranteed bandwidth. The CIR bandwidth of a BGS flow is guaranteed at each hop if the BGS flow rate is confirmed with its pre-claimed parameters.

In addition to above, the use of in-band signaling with DiffServ will make the flow level QoS expectation available at each network device on the IP path. Then network device could collect all flows signaling info to form an accurate instruction for QoS programming automatically. Without such automation process, DiffServ will only be able to rely on the network management procedures to configure the QoS for each class at each hop one by one, this is called Per Hop Behavior.

6.3. The info carried in in-band signaling

The detailed info carried in the signaling was described in the draft [I-D.han-tsvwg-ip-transport-qos].

In order to achieve the service guarantees, admission control is needed for user's service expectation. Also to get the best service guarantee in network, some network device need to do traffic shaping to measure and control traffic for difference service.

To satisfy above two requirements, each flow's service expectation and traffic pattern should be carried in the signaling.

To describe one or multiple flows traffic character, Single Rate Three Color Marker (srTCM from [RFC2697]) or Two Rate Three Color Marker (trTCM from RFC4115) could be used. The minimum requirement from user's signaling is the Cir (Committed information rate) when srTCM is used, CIR plus PIR (Peak information rate) when trTCM is used. Optionally, the signaling can carry more info about the traffic, such as CBS, EBS. The exact requirements for signaling will be decided by IETF.

6.4. Admission process with in-band signaling

The admission control should be enforced at any ingress interface connecting to user's device or accepting user's traffic, and also enforced at the egress interfaces that will send the user's traffic to next network device. In addition to other irrelevant admission checking, the admission process for the E-DiffServ minimally has three major steps:

1. If a user is allowed to request the desired service described in its in-band signaling. This may include the security, authentication and billing checks.

2. After adding the new flow for a user's request, if the total rate for all flows in the same class at the same ingress interface exceeds the total bandwidth allocated for the class for the ingress interface, this allocated bandwidth is from management procedures and is less than the ingress interface link rate.

3. After adding the new flow for a user's request, if the total rate for all flows in the same class at the same egress interface exceeds the total bandwidth allocated for the class for the egress interface, this allocated bandwidth is from management procedures and is less than the egress interface link rate.

In above admission control process, if a flow is not allowed, the in-band signaling of the flow is updated accordingly to indicate the desired service is rejected, and the in-band signaling will finally notify the user's host. If a flow is allowed, the in-band signaling of the flow will proceed to continue. For the details of how the closed loop control is formed by in-band signaling, please refer to [I-D.han-tsvwg-ip-transport-qos].

6.5. User traffic classification and New DSCP class

After passed the admission control process, user's traffic is classified to different class based on its requested service. This is also done at the ingress interface like the admission control. There are two options for the classification:

1. For a network without DiffServ enabled, the LGS flows can be classified as EF class to achieve the shortest latency; The BGS

flows can be classified as any type of AFxy class; The BE flows can be classified as BE class.

2. For a network that DiffServ is enabled, the LGS and BGS flows can be classified as two new class; The BE flows can be classified as BE class. To be backward compatible and not interrupting existing DiffServ services, new DSCP classes are proposed for LGS and BGS. The values are TBD.

6.6. User traffic conformity checking

After a user’s request service passes the admission control and is accepted, user’s traffic is still needed to be checked for its conformity. i.e, if the traffic is conforming to the traffic pattern claimed in the service request.

This is done by a traffic meter (shaper) associated to ingress interface. The algorithm in the meter is using either srTCM or trTCM. The action of the meter will determine what exact class the traffic will be classified as. Only the Green traffic will be marked to the designated class described in 5.5. Other colored traffic (Yellow or Red) will be classified as lower priority class or just discarded. The exact action is up to the decision of IETF.

6.7. Queuing and Scheduling Consideration

The most important factor to provide the guaranteed service is the queuing and scheduling selection at egress interface on each router. There are many researches about this topic. The current widely accepted methods are Class based Strict Priority Queuing (PQ) and Weighed Fair Queuing (WFQ). It is recommended that the PQ+WFQ are used for different service flows. PQ is the best candidate to provide the shortest latency for LGS flows. The WFQ can be used for BGS flows. There are two options for the Queuing and Scheduling selection:

1. For a network without DiffServ enabled, the LGS flows can be queued into the Highest Priority Queue to achieve the shortest latency; The BGS and BE flows can be queued into different Weighted Fair Queues to share bandwidth. The weight for BGS can be obtained from the sum of all BGS flows’ CIR. The weight for BE can be calculated from the link rate deducting the bandwidth allocated for LGS and BGS flows.

2. For a network with DiffServ enabled, the LGS flows can be queued into the Secondary Highest Priority Queue just below EF queue; The BGS and BE flows can be queued into different Weighted Fair Queues to share bandwidth with other existing classes. The

weight for BGS can be obtained from the sum of all BGS flows' CIR. The weight for BE can be calculated from the link rate deducting the bandwidth allocated for all AFxy classes, LGS and BGS flows.

6.8. Class Based Traffic Shaper

To guarantee each class's traffic are confirming to the pre-allocated bandwidth or resource for the class, class-based traffic shaper can be used for traffics before they are queued into different queues. Only Green traffic will be queued into the queue described above. Other colored traffic (Yellow or Red) will be queued into a lower priority queue or discard. The exact actions will be decided by IETF. The purpose of using traffic shaper for each class is to protect the service for each class when other class traffic is exceeding its allocation. This is critical to satisfy the design target outlined in section 5.

The algorithm for the shaper is similar to the ingress traffic shaper described in section 6.6. The algorithm in the meter is using either srTCM or trTCM.

The Traffic Shaper's Parameter such as PIR, CIR, CBS, EBS can be obtained from all flow's traffic parameters embedded into the in-band signaling. i.e, PIR or CIR can be calculated by the sum of all allowed (same class) flow's PIR or CIR values. CBS and EBS can also be obtained from flow's signaling message if it carries, or from management process. The detailed standard will be decided by IETF.

7. IANA Considerations

The new DSCP values are required to be allocated by IANA.

8. Security Considerations

TBD.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2581] Allman, M., Paxson, V., and W. Stevens, "TCP Congestion Control", RFC 2581, DOI 10.17487/RFC2581, April 1999, <https://www.rfc-editor.org/info/rfc2581>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

9.2. Informative References

[Enhanced_DSCP] "Enhanced DSCP by In-band Signaling", 2018, <https://drive.google.com/file/d/1aBq4z70iJgbKZIBn7TTieQUhZ7MPYOsn/view?usp=sharing>.

[EPC] 3GPP, "The Evolved Packet Core", 2018, <http://www.3gpp.org/technologies/keywords-acronyms/ 100-the-evolved-packet-core>.

[HU5G] Huawei, "5G Vision: 100 Billion connections, 1 ms Latency, and 10 Gbps Throughput", 2015, <http://www.huawei.com/minisite/5g/en/defining-5g.html>.

[I-D.falk-xcp-spec] Falk, A., "Specification for the Explicit Control Protocol (XCP)", draft-falk-xcp-spec-03 (work in progress), July 2007.

[I-D.han-iccrg-arvr-transport-problem] Han, L. and K. Smith, "Problem Statement: Transport Support for Augmented and Virtual Reality Applications", draft-han-iccrg-arvr-transport-problem-01 (work in progress), March 2017.

[I-D.han-tsvwg-cc] Han, L., Qu, Y., and T. Nadeau, "A New Congestion Control in Bandwidth Guaranteed Network", draft-han-tsvwg-cc-00 (work in progress), March 2018.

[I-D.han-tsvwg-ip-transport-qos]
Han, L., Qu, Y., Dong, L., Li, R., Nadeau, T., Smith, K., and J. Tantsura, "Resource Reservation Protocol for IP Transport QoS", draft-han-tsvwg-ip-transport-qos-03 (work in progress), October 2019.

[I-D.harper-inband-signalling-requirements]
Harper, J., "Requirements for In-Band QoS Signalling", draft-harper-inband-signalling-requirements-00 (work in progress), January 2007.

[I-D.ietf-tcpm-dctcp]
Bensley, S., Eggert, L., Thaler, D., Balasubramanian, P., and G. Judd, "Datacenter TCP (DCTCP): TCP Congestion Control for Datacenters", draft-ietf-tcpm-dctcp-03 (work in progress), November 2016.

[I-D.roberts-inband-qos-ipv6]
Roberts, L. and J. Harford, "In-Band QoS Signaling for IPv6", draft-roberts-inband-qos-ipv6-00 (work in progress), July 2005.

[I-D.sridharan-tcpm-ctcp]
Sridharan, M., Tan, K., Bansal, D., and D. Thaler, "Compound TCP: A New TCP Congestion Control for High-Speed and Long Distance Networks", draft-sridharan-tcpm-ctcp-02 (work in progress), November 2008.

[IPv6_Parameters]
IANA, "Internet Protocol Version 6 (IPv6) Parameters", 2015, <https://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml#ipv6-parameters-2>.

[QU2016] Qualcomm, "Leading the world to 5G", 2016, <https://www.qualcomm.com/media/documents/files/qualcomm-5g-vision-presentation.pdf>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <https://www.rfc-editor.org/info/rfc2475>.

[RFC2697] Heinanen, J. and R. Guerin, "A Single Rate Three Color Marker", RFC 2697, DOI 10.17487/RFC2697, September 1999, <https://www.rfc-editor.org/info/rfc2697>.

[RFC3175] Baker, F., Iturralde, C., Le Faucheur, F., and B. Davie, "Aggregation of RSVP for IPv4 and IPv6 Reservations", RFC 3175, DOI 10.17487/RFC3175, September 2001, <https://www.rfc-editor.org/info/rfc3175>.

[RFC4115] Aboul-Magd, O. and S. Rabie, "A Differentiated Service Two-Rate, Three-Color Marker with Efficient Handling of in-Profile Traffic", RFC 4115, DOI 10.17487/RFC4115, July 2005, <https://www.rfc-editor.org/info/rfc4115>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8557] Finn, N. and P. Thubert, "Deterministic Networking Problem Statement", RFC 8557, DOI 10.17487/RFC8557, May 2019, <https://www.rfc-editor.org/info/rfc8557>.

[RFC8578] Grossman, E., Ed., "Deterministic Networking Use Cases", RFC 8578, DOI 10.17487/RFC8578, May 2019, <https://www.rfc-editor.org/info/rfc8578>.

[TIA] TIA 1039 Revision A, "QoS Signaling for IP QoS Support and Sender Authentication", 2015, <https://global.ihs.com/doc_detail.cfm?&csf=TIA&item_s_key=00480715&item_key_date=880431>.

Appendix A. Acknowledgements

TBD.

Appendix B. Testing Results

The analysis of the queuing delay and testing results can be found at [Enhanced_DSCP] .

Authors' Addresses

Lin Han
Futurewei Technologies
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: lin.han@futurewei.com

Yingzhen Qu
Futurewei Technologies
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: yingzhen.qu@futurewei.com

Richard Li
Futurewei Technologies
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: richard.li@futurewei.com