

tsvwg  
Internet-Draft  
Intended status: Informational  
Expires: 26 December 2025

Q. Xiong  
X. Zhu  
ZTE Corporation  
24 June 2025

Admission Control with Gateway  
draft-xiong-tsvwg-admission-control-00

## Abstract

This document proposes to enhance the congestion control with the admission control mechanism at the gateway to achieve fast feedback and per-flow control.

## Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 26 December 2025.

## Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	2
2. Conventions Used in This Document . . . . .	3
2.1. Abbreviations . . . . .	3
2.2. Requirements Language . . . . .	3
3. Admission Control with Gateway . . . . .	4
4. Admission Control Message Formats . . . . .	5
4.1. ICMP Message Format . . . . .	5
4.2. UDP Message Format . . . . .	5
5. Consideration for the Congestion Notification . . . . .	6
6. Security Considerations . . . . .	7
7. IANA Considerations . . . . .	7
8. Informative References . . . . .	7
Authors' Addresses . . . . .	7

## 1. Introduction

Driven by rapid development of big data and AI technologies, the performance of data transmission is critical across various industries. The data is transferred via transport layer protocols such as Transfer Control Protocol (TCP), Quick UDP Internet Connections (QUIC), Remote Direct Memory Access (RDMA) and so on. The congestion control algorithms of these protocols are implemented by controlling the size of the congestion window and adjusting the sending rates upon the network status feedback. The traditional end-to-end congestion control mechanisms will rely on the following signals for rate adjustment.

- \* Explicit Congestion Notification (ECN): notify the sender to slow down by marking packets in the congestion queue.
- \* Delay and bandwidth measurement (RTP/OWD): estimate the network status and bottleneck bandwidth by measuring packet round-trip time or one-way delay.
- \* Network telemetry (INT): inserting queuing status in packets to provide fine-grained congestion information.

When the congestion occurs, the congestion notification will be transmitted through the congested link. The congestion signal path is coupled with the data path. Moreover, it will also delay the network feedback due to the long-distance transmission loop, resulting in the inability to adjust the transmission rate in a timely manner. It needs to achieve fast feedback nearing the senders. The gateway may be used to get congestion status of the network and performs fast feedback and admission control for new traffic entering the network from the same gateway. It will significantly mitigate the congested link when the traffic transfers to the same destination.

This document proposes to enhance the congestion control with the admission control mechanism at the gateway to achieve fast feedback and per-flow control.

## 2. Conventions Used in This Document

### 2.1. Abbreviations

RTT: Round-Trip Time

TCP: Transfer Control Protocol

RDMA: Remote Direct Memory Access Round-Trip Time

QUIC: Quick UDP Internet Connections

ECN: Explicit Congestion Notification

PFC: Priority-based Flow Control

RoCEv2: RDMA over Converged Ethernet version 2

### 2.2. Requirements Language

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] [RFC8174] when, and only when, they appear in all capitals, as shown here.

### 3. Admission Control with Gateway

Multiple source clients may share the same gateway, potentially directing significant traffic volume through it. If these traffic targets the same destination server, it may converge onto a single physical link. Upon detecting congestion, the gateway can record the states and apply immediate congestion control to any subsequent traffic for that server. Furthermore, the gateway can perform admission control, when new traffic arrives, it may reject or pause new flows to prevent worsening existing congestion.

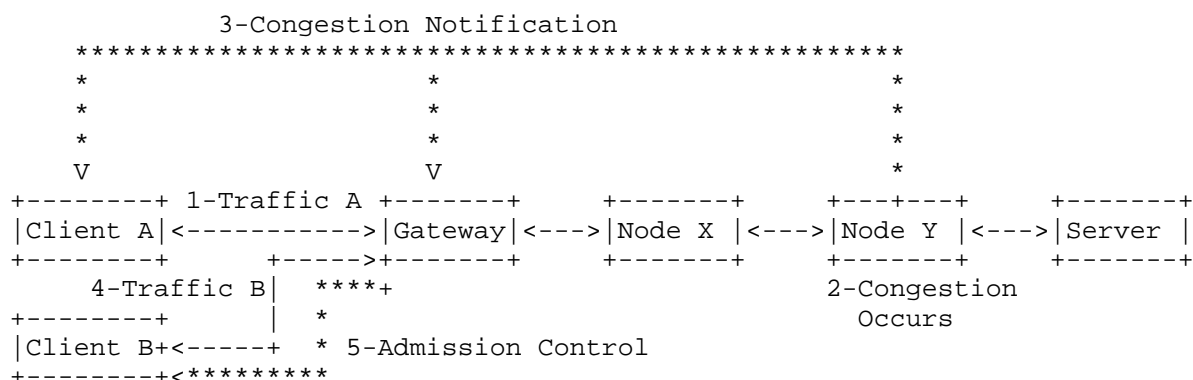


Figure 1: Admission Control at Gateway

The example of admission control at gateway is shown in Figure 1 and the steps are as follows.

Step 1: the client A sends the traffic to the server along the path of network gateway, node X and node Y.

Step 2: a congestion occurs at the node Y.

Step 3: the node Y triggers a congestion notification to the gateway or the client A, the gateway saves the congestion status.

Step 4: the client B sends the traffic to the server entering from gateway.

Step 5: the gateway performs the admission control based on the congestion information to mitigate the congestion when the traffic is sending to the same server.

## 4. Admission Control Message Formats

### 4.1. ICMP Message Format

The admission control message sent from gateway may be a ICMP message which is formatted as Figure 2.

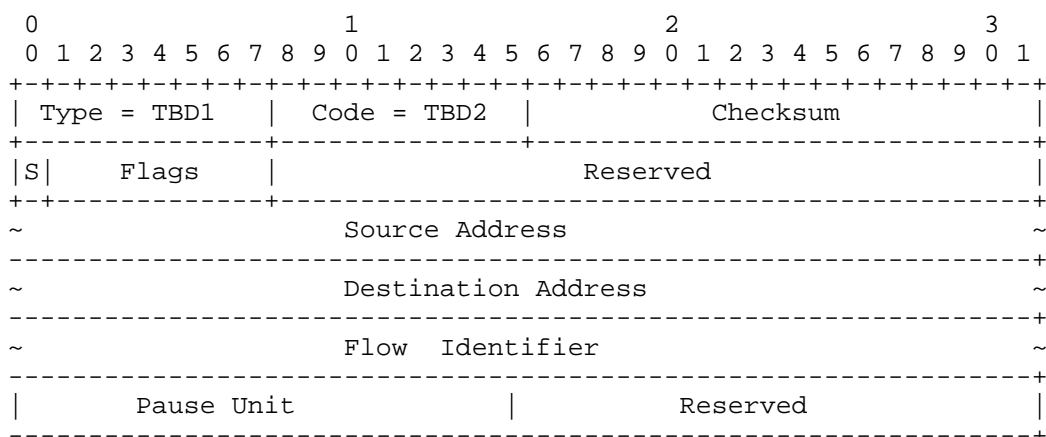


Figure 2: Admission Control Message Format with ICMP

Type and Code: TBD 1 and TBD2, the fields indicate the admission control type and code.

Source Address (variable): indicates the IPv4 or IPv6 address of the sender.

Destination Address (variable): indicates the IPv4 or IPv6 address of the receiver.

Flow Identifier (variable): it indicates the IP 5 tuples when transmitting TCP or QUIC data and it also indicates source QP (Queue Pair) and destination QP when transmitting RDMA data.

Pause Unit (16bits): indicates the quanta time when the sender must stop sending the specified traffic after receiving a pause frame.

### 4.2. UDP Message Format

The admission control message sent from gateway may be a UDP message which is formatted as Figure 3.

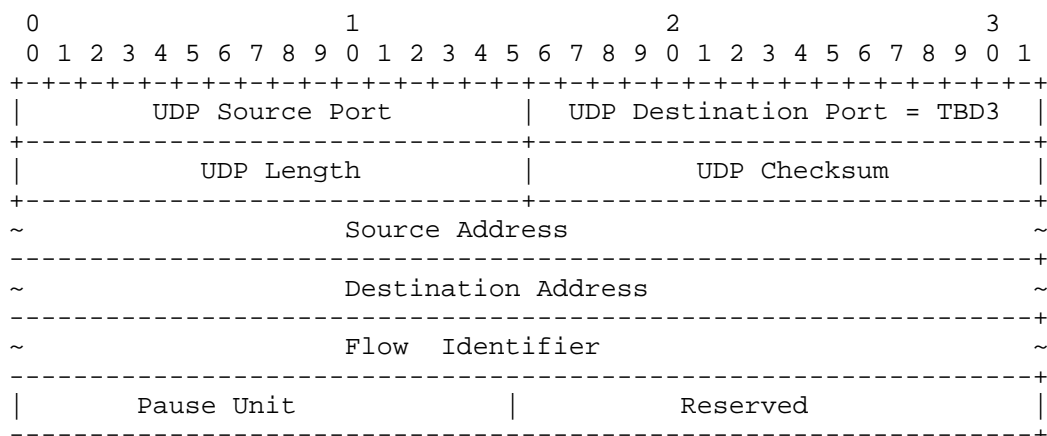


Figure 3: Admission Control Message Format with UDP

UDP Header: The UDP header as specified in [RFC768] includes the UDP source port, UDP destination port, UDP length, and UDP checksum.

UDP Destination port: TBD3, indicates a new well-known UDP destination port needs to be allocated for this admission control message.

The other fields are the same with section 4.1.

## 5. Consideration for the Congestion Notification

The gateway may collect the congestion information as following:

\*The network may signal congestion by ECN markings, and the receiver will pass this information back to the sender in transport-layer acknowledgements. The gateway needs to enable the function of notification intercepting to get the information of congestion.

\*The network may implement the classical stepwise back pressure with dedicated Ethernet pause frame such as Priority-based Flow Control (PFC) frame. And the congestion status may be stepwise notified from congestion point to the gateway.

\*The network may notify the sender with the congestion status directly by the messages defined in section 4, or by a dedicated congestion notification packet as per [I-D.xiao-rtgwg-rocev2-fast-cnp] when RoCv2 implements . The gateway needs to enable the function of notification intercepting to get the information of congestion.

\*The network may notify the gateway (as a proxy) with the congestion status directly by a dedicated congestion notification packet as per [I-D.xiao-rtgwg-proxy-congestion-notification].

## 6. Security Considerations

To be discussed in future versions of this document.

## 7. IANA Considerations

TBD.

## 8. Informative References

- [I-D.xiao-rtgwg-proxy-congestion-notification]  
Min, X., "Proxy for Congestion Notification", Work in Progress, Internet-Draft, draft-xiao-rtgwg-proxy-congestion-notification-00, 16 June 2025, <<https://datatracker.ietf.org/doc/html/draft-xiao-rtgwg-proxy-congestion-notification-00>>.
- [I-D.xiao-rtgwg-rocev2-fast-cnp]  
Min, X. and lihesong, "Fast Congestion Notification Packet (CNP) in RoCEv2 Networks", Work in Progress, Internet-Draft, draft-xiao-rtgwg-rocev2-fast-cnp-03, 9 June 2025, <<https://datatracker.ietf.org/doc/html/draft-xiao-rtgwg-rocev2-fast-cnp-03>>.
- [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/rfc/rfc2119>>.
- [RFC768] "\*\*\*\* BROKEN REFERENCE \*\*\*\*".
- [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/rfc/rfc8174>>.

## Authors' Addresses

Quan Xiong  
ZTE Corporation  
Email: [xiong.quan@zte.com.cn](mailto:xiong.quan@zte.com.cn)

Xiangyang Zhu  
ZTE Corporation

Email: [zhu.xiangyang@zte.com.cn](mailto:zhu.xiangyang@zte.com.cn)