

rtgwg  
Internet-Draft  
Intended status: Informational  
Expires: 23 October 2025

T. He, Ed.  
China Unicom  
H. Shi, Ed.  
Huawei  
Z. Han  
X. Gao  
China Unicom  
T. Zhou  
Huawei  
21 April 2025

Framework for Implementing Lossless Techniques in Wide Area Networks  
draft-hs-rtgwg-wan-lossless-framework-01

Abstract

This document proposes a comprehensive framework to address the challenges of efficient, reliable, and cost-effective large volume data transmission over Wide Area Networks (WANs). The framework focuses on planning and managing traffic paths, network slicing, and utilizing multi-level network buffers. It introduces dynamic path scheduling and advanced resource allocation techniques to optimize network resource and minimize congestion. By leveraging cross-device buffer coordination and real-time adjustments, the framework ensures high throughput and low latency, meeting the demands of modern, data-intensive applications while providing a robust solution for large-scale data transmission.

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 23 October 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. Network Challenges Posed by Large Volume Data Transmission .	3
2.1. Limited Network Capacity . . . . .	3
2.2. Congestion Hotspots . . . . .	4
2.3. Inefficient Buffer Utilization . . . . .	4
3. Framework . . . . .	4
3.1. Adaptive Planning and Management of Network Resource . . .	4
3.1.1. Specific Requirements: . . . . .	5
3.2. Use and Management of Multi-Level Network Buffers . . . .	5
3.2.1. Specific Requirements: . . . . .	6
3.3. Requesting Source Rate Control . . . . .	6
4. Conclusion . . . . .	7
5. Security Considerations . . . . .	7
6. IANA Considerations . . . . .	7
7. Informative References . . . . .	7
Acknowledgements . . . . .	7
Contributors . . . . .	7
Authors' Addresses . . . . .	7

## 1. Introduction

In recent years, the demand for reliable and efficient transmission of large volumes of data across Wide Area Networks (WANs) has surged. [I-D.huang-rtgwg-wan-lossless-uc] highlighted several critical use cases that emphasize the necessity of low packet loss and high throughput in WANs. These requirements are driven by applications that handle massive datasets, such as scientific research, financial transactions, and multimedia content delivery, while the locations of data production and consumption differ, requiring efficient and timely transmission across WANs. The characteristics and requirements of large data transmission are listed as follows:

- \* **Large Volume.** The datasets involved in these transmissions often reach terabyte levels. Traditional fixed bandwidth dedicated lines, while reliable, can be prohibitively expensive. Enterprises must balance the need for high-capacity data transmission with cost considerations. This necessitates exploring more flexible and economical solutions that can handle large-volume data without incurring excessive costs.
- \* **Timeliness.** Timeliness is a critical factor for data transmission over WANs. For instance, in the field of genetic research, the timely transmission of genetic data can significantly influence diagnostic and treatment outcomes. Delays in data transmission can render the data obsolete, e.g., leading to incorrect results and conclusions. Therefore, ensuring that data is transmitted within a specific time window is essential for maintaining its utility and accuracy.
- \* **Predictability.** Large-volume data transmission tasks typically have predictable patterns, allowing for better planning and resource allocation. This predictability helps in designing network solutions that can efficiently manage the anticipated data load. By leveraging predictable traffic patterns, network administrators can optimize resource allocation, minimize congestion, and enhance overall network performance.

This document proposes a comprehensive framework aimed at addressing the challenges associated with large volume data transmission over WANs. The framework focuses on enhancing traffic management and resource allocation strategies to ensure efficient, reliable, and cost-effective data transmission. By implementing these strategies, the framework aims to meet the demands of modern, data-intensive applications, providing a robust solution for large volume data transmission in WAN environments.

## 2. Network Challenges Posed by Large Volume Data Transmission

### 2.1. Limited Network Capacity

WANs have finite carrying capacities. When a significant amount of traffic enters the network simultaneously, it can lead to traffic conflicts, resulting in queuing and jitter. These issues are exacerbated by the continuous nature of large data transfers, which can strain network resources over extended periods. Addressing these challenges requires advanced traffic management techniques that can efficiently utilize available network capacity.

## 2.2. Congestion Hotspots

Packet loss often occurs due to probabilistic simultaneous influxes of large volumes of traffic. This congestion is exacerbated by mechanisms such as Equal-Cost Multi-Path (ECMP) routing, where multiple flows compete for certain bottleneck links, leading to congestion and packet loss. Packet loss in WANs does not lead to permanent data loss since lost packets can be retransmitted. However, retransmissions increase transmission latency, causing delays in data delivery. Moreover, packet loss can trigger congestion control mechanisms, which reduce the network's throughput to prevent further congestion. This reduction in throughput can significantly affect the performance of data-intensive applications, making it critical to minimize packet loss.

## 2.3. Inefficient Buffer Utilization

The network itself has a certain buffer capacity to partially mitigate short-term processing deficiencies. However, current mechanisms only utilize the local device's buffer and do not fully exploit the overall buffer capacity across multiple devices. This fragmented buffer utilization leads to inefficiencies in handling bursty traffic. Advanced congestion management strategies are necessary to coordinate buffer usage across the network, maintaining high throughput and low latency to ensure efficient and reliable data transmission.

## 3. Framework

This document proposes a comprehensive framework to address the challenges of efficient, reliable, and cost-effective large volume data transmission over Wide Area Networks (WANs). The framework focuses on the planning and management of traffic paths, network slicing, and the use and management of multi-level network buffers.

### 3.1. Adaptive Planning and Management of Network Resource

When users seek efficient transmission of large datasets, they can rent temporary network bandwidth in addition to their fixed leased lines (a.k.a. guaranteed bandwidth). This temporary bandwidth is cheaper by sharing but offers weaker Service Level Agreements (SLAs). Due to the predictable nature of the traffic, users can pre-request resource scheduling from the network, including traffic paths and even network slices. The network can allocate resources based on availability, avoiding prolonged congestion through effective planning. If serious congestion occurs, the network scheduler can recalculate paths and slice resources. Network devices can flexibly choose the best available path from multiple pre-allocated paths,

particularly when head-end devices detect local or remote congestion. By adjusting the current and incoming traffic path selection, network devices can optimize traffic distribution and alleviate congestion dynamically.

#### 3.1.1. Specific Requirements:

- \* **\*Network Resource Reporting and User Request\***: Network devices report attributes such as bandwidth, latency through control plane protocols like IGP and BGP-LS. Users provide the overall needs of bandwidth and latency for large volume data transmission, including guaranteed dedicated resources and flexible resources with weaker guarantees. In addition to know network parameters such as bandwidth and latency, the system also needs to know whether network forwarding nodes have the ability to share the buffer with other devices, and confirm the scope of the wide-area lossless network domain. In the centralized mode, the central controller needs to know the network device's capability of buffer and buffer size to do path planning. The information transferring can be done through the BGP-LS protocol extension. In distributed mode, the network forwarding nodes can realize multi-level network buffering and path switching by knowing the neighbour's capability of buffer and buffer size. The information transferring can be done through the IGP protocol extension.
- \* **\*Network Resource Allocation and Policy Distribution\***: Controllers calculate out IP-based dedicated lines (IP tunnels with segment routing) within the WAN domain based on available flexible bandwidth and buffers. Using SR-policy, data traffic is steering into IP tunnels at ingress nodes and directed to dedicated network slicing. Configuration of buffer allocations are distributed via protocols like BGP and PCEP from the controller to the network devices who are executing and enforcing these configurations.
- \* **\*Network State Measurement and Telemetry\***: Real-time bandwidth measurement based on measurement packets helps in sensing utilized and available bandwidth on network links. This information is reported to the controller via telemetry mechanisms and used to adjust paths and slice resources. For example, when a link nears its bandwidth limit, traffic can be rerouted to idle path resources to improve overall network bandwidth utilization.

#### 3.2. Use and Management of Multi-Level Network Buffers

Since temporary bandwidth is shared and not dedicated, it exhibits weaker SLA guarantees. If traffic experiences jitter during transmission, network device buffers can absorb packets to reduce packet loss.

### 3.2.1. Specific Requirements:

- \* **\*Single Device Buffer Sharing and Management\***: Single devices should implement fine-grained buffer divisions based on traffic priority and slice. These buffers should be isolated to avoid mutual interference. Initial buffer resource allocation is determined by the controller and configured across all devices in the domain via control plane protocols.
- \* **\*Cross-Device Buffer Coordination\***: Given the nature of large data transmissions, a single device's buffer might be insufficient for absorbing bursty traffic. Therefore, multiple devices' buffers of the same fine-grained type (e.g., same priority and slice) should be used collectively. For example, if device C in the path A->B->C is congested and its buffer is insufficient, it should notify upstream devices B or A to utilize their similar buffers to absorb some traffic. This involves:
  - **Control Signaling**: Using control signaling packets to notify upstream devices to buffer packets, reducing the burden on the congested device. If upstream device buffers also reach a threshold, further notifications should be triggered upstream. Control signaling should include buffer index (e.g., slice ID), control instructions, and parameters. Controller configuration or segment routing can help determine upstream device addresses. Upon congestion relief, upstream devices should be notified to release buffered traffic. This notification mechanism can be inspired by IEEE PFC mechanisms but requires more granular backpressure.
  - **Trigger Conditions for Buffer Coordination**: The local device-triggering cross-device buffer coordination requires pre-set conditions. Controllers can configure device-specific thresholds to customize trigger conditions for each device, slice, and priority.

### 3.3. Requesting Source Rate Control

Network devices can send rate control requests to the source via data packet marking or separate control packets. This method is useful during widespread network congestion, leveraging source rate reduction to manage traffic. Although this feedback mechanism involves a larger control loop and slower adjustments, efficiency can be improved through fast reverse notifications.

#### 4. Conclusion

The proposed framework addresses the challenges of large volume data transmission over WANs by enhancing traffic management and resource allocation strategies. By implementing dynamic path scheduling, advanced resource allocation, and efficient buffer management, the framework ensures efficient, reliable, and cost-effective data transmission. This approach meets the demands of data-intensive applications, providing a robust solution for large volume data transmission in WAN environments.

#### 5. Security Considerations

TBD.

#### 6. IANA Considerations

TBD.

#### 7. Informative References

[I-D.huang-rtgwg-wan-lossless-uc]

Zhengxin, H., He, T., Huang, H., and T. Zhou, "Use Cases and Requirements for Implementing Lossless Techniques in Wide Area Networks", Work in Progress, Internet-Draft, draft-huang-rtgwg-wan-lossless-uc-01, 8 July 2024, <<https://datatracker.ietf.org/doc/html/draft-huang-rtgwg-wan-lossless-uc-01>>.

#### Acknowledgements

TBD.

#### Contributors

TBD.

#### Authors' Addresses

Tao He (editor)  
China Unicom  
Beijing  
China  
Email: [het21@chinaunicom.cn](mailto:het21@chinaunicom.cn)

Hang Shi (editor)  
Huawei  
Beijing  
China  
Email: shihang9@huawei.com

Zhengxin Han  
China Unicom  
Email: hanzx21@chinaunicom.cn

Xing Gao  
China Unicom  
Email: gaox60@chinaunicom.cn

Tianran Zhou  
Huawei  
Email: zhoutianran@huawei.com