Networking Z. Han, Ed.
Internet-Draft T. He
Intended status: Informational China Unicom
Expires: 3 January 2026 H. Shi
T. Zhou
Huawei
2 July 2025

# Use Cases and Requirements for Implementing Lossless Techniques in Wide Area Networks
draft-hs-rtgwg-wan-lossless-uc-01

## Abstract

This document outlines the use cases and requirements for implementing lossless data transmission techniques in Wide Area Networks (WANs), motivated by the increasing demand for high-bandwidth and reliable data transport in applications such as high-performance computing (HPC), genetic sequencing, multimedia content production and distributed training. The challenges associated with existing data transport protocols in WAN environments are discussed, along with the proposal of requirements for enhancing lossless transmission capabilities to support emerging data-intensive applications.

## Status of This Memo



## Copyright Notice

Table of Contents



## 1. Introduction

With the rapid development of big data and intelligent computing, it is getting more clear that numerous fields need wide area networks (WANs) to provide high-throughput and high-performance transmission services to meet the needs for massive application data transmission over long distance. These typical scenarios include HPC high-performance computing, genetic sequencing, multimedia content production and distributed training etc. Traditional network protocols, designed in an era before these immense data demands, struggle to keep up, particularly when it comes to ensuring extremely low or zero data packet loss over long distance.





This document aims to illustrate on the necessity for advanced lossless transmission technologies in WANs. By identifying the limitations of current network protocols and outlining the requirements for new developments, we hope to pave the way for a new generation of WANs. These networks will not only meet the current demands of data-intensive applications but will also support the next wave of digital innovation.

## 2. Use Cases

The necessity for implementing lossless data transmission techniques in Wide Area Networks (WANs) is underscored by several critical application areas. These use cases highlight the imperative for reliable, high-throughput data transmission capabilities to support the demanding requirements of modern data-intensive operations.

## 2.1. High-Performance Computing (HPC) Services for Scientific Research

High-Performance Computing (HPC) services are fundamental to scientific advancements, where collaborative efforts across various geographical regions are commonplace. For instance, the study of PSII proteins, which are crucial for understanding how water molecules split to produce oxygen, generates between 30 to 120 high-resolution images per second during experiments. This results in 60-100 GB of data every five minutes, necessitating rapid and lossless data transfer from the National Renewable Energy Laboratory’s equipment back to analysis labs such as the Lawrence Berkeley National Laboratory. The efficiency and reliability of WANs in this context are not just beneficial but essential for facilitating the seamless collaboration between scientists in





different domains, enabling them to share and analyze large datasets effectively.

## 2.2. Rapid Transmission Services for Genetic Sequencing of Timely Medical Services

The field of genetic sequencing has seen exponential growth, driven by the decreasing costs and widespread application of sequencing technologies. This growth is matched by the burgeoning data volumes generated, which require efficient and lossless transmission to cloud or private data centers for analysis. For example, sequencing a single human genome produces 100GB to 200GB of data. With daily data production rates reaching 6TB to 12TB and annual data management needs surpassing 1.6PB, the demand for high-speed, reliable data transfer is evident. The existing network transfer efficiencies present significant bottlenecks, extending the turnaround times for sequencing services and impacting the timely delivery of precision medicine.

## 2.3. Large-Scale Audio/Video Data Migration for Multimedia Content Production

The competitive landscape of short video industry, the promotion of 4K ultra-high-definition channels,coupled with the independence of acquisition and shooting, cloud-based post production, and terminal presentation. So that a large amount of audio and video data need to be transmitted across WANs. Traditional methods of data transportation, involving physical media and manual transfer, are time-consuming and inefficient. For instance, film crews generating 2TB of data daily resort to physically moving storage media to processing locations, the process that significantly lengthens the production cycle and slows down the market response. The requirement for network infrastructure capability of handling such extensive data transfers efficiently and without loss is critical for maintaining the pace of production and ensuring the quality of the final multimedia content.

## 2.4. Massive Data Transfer to Intelligent Computing Center for Distributed Training

Transferring massive data to intelligent computing centers is the premise for distributed training. For example, the securities company has a batch of financial models that need to be transmitted to the intelligent computing center for training. The amount of data is huge, and the data transmitted each time reaches TB level. There are usually two kinds of data transmission solutions. One is to use the high-speed dedicated line which is very expensive up to one million yuan monthly. The another is manual transportation of hard





copy, the round trip cycle of each data transferring can be as long as several days and the labor consumption is huge. The reason for the high price of the existing high-speed dedicated line service is mainly because that network need to reserve sufficient bandwidth resources, though the actual network utilization rate is low. High-throughput network is important for distributed training.

## 3. Problem Analysis and Goal

### 3.1. Problem Analysis

The primary objective of Wide Area Networks (WANs) is to provide long-term, stable, high-throughput and high-performance network services that can accommodate the sudden surges in data transmission demands, essential for data migration across diverse geographical locations. This goal is predicated on leveraging the inherent statistical multiplexing advantage of IP networks, which allows for cost-effective bandwidth allocation and enhanced overall network throughput.

Despite the advantages of statistical multiplexing in IP networks, such as cost reduction and throughput optimization, this model introduces significant challenges in ensuring absolute resource guarantee and and extremely low packet loss especially when there are micro-bursts and congestion. The practice of over provisioning bandwidth, common among service providers, does not equate to lossless data transmission, which is a critical shortfall when compared to dedicated light networks or resources with hard isolation.

#### 3.1.1. Impact of Packet Loss

In the scenarios outlined for data migration whether for high-performance computing services, genetic sequencing, or audio/video data migration the reliance on traditional transmission protocols like TCP or RDMA [RoCEv2] is common. However, both protocols are adversely affected by packet loss, especially over long distance transmissions.

For TCP, algorithms such as CUBIC, a loss-based congestion control mechanism, see a dramatic throughput decline of up to 89.9% with just a 2% packet loss when the Round-Trip Time (RTT) is 30ms. BBR, another TCP congestion control that based on bandwidth and delay, also suffers significantly when packet loss exceeds 5%, with throughput plummeting in scenarios where packet loss reaches 20%. The cost of retransmissions in these conditions is notably high, with slight packet loss (<1%) scenarios showing a retransmission rate 6-10 times higher than CUBIC, and in severe packet loss scenarios, the rate can increase exponentially.

For RDMA, often used within data centers for inter-node data access over UDP, relies on a go-Back-N retransmission mechanism. Its throughput dramatically decreases with packet loss rates greater than 0.1%, and a 2% packet loss rate effectively reduces throughput to zero. To maintain unaffected throughput, the packet loss rate must be kept below one in a hundred thousand.

These challenges underscore a critical gap in the current capabilities of IP networks to support the demanding requirements of modern, data-intensive applications. The inability to ensure extremely low or zero packet loss across WANs not only impacts application performance but also limits the potential for innovation and collaboration across key sectors reliant on rapid and reliable data transmission.

### 3.2. Goal

The overarching goal in the evolution of Wide Area Networks (WANs) to serve the mentioned use cases is to enable lossless, extremely low or zero packet loss transmission services customized for the seamless migration of data across different geographical areas. In an age where digital data’s volume, velocity, and variety are expanding exponentially, ensuring the lossless transmission of this data during inter-regional migration activities becomes indispensable. This is critically important for applications and operations that rely on the integrity and timeliness of data, such as AI/HPC computing and data backup and recovery.

## 4. Challenges and Requirements

The quest for lossless data transmission in Wide Area Networks (WANs) is confronted with significant challenges, notably the phenomenon of elephant flows-large, bursty data transfers that can cause instantaneous congestion and packet loss within network device queues. This not only increases application latency but also diminishes throughput, adversely affecting application performance. In data centers, certain lossless technologies are deployed to enhance the performance of such applications:

1) Priority-based Flow Control (PFC): Widely adopted for its ability to manage traffic flow, PFC [PFC] works by halting the transmission of specific queues when downstream congestion is detected, thereby achieving zero packet loss. The foundational





flow control mechanism, defined by IEEE 802, involves sending a pause frame from a receiving device to a sending device to temporarily halt traffic, allowing time for congestion to clear before resuming transmission.

2)Explicit Congestion Notification (ECN) with Data Center Quantized Congestion Notification (DCQCN): DCQCN [DCQCN], the most extensively used congestion control algorithm in RDMA networks, requires network devices to support ECN functionality [RFC3168], with other protocol functionalities implemented on the network card of the host machine. DCQCN ensures high throughput in RDMA networks needing zero packet loss by signaling congestion through ECN markers sent from congested nodes to the sender, prompting a reduction in sending rate.

However, the application of these data center-oriented lossless techniques to WANs encounters obstacles due to the larger scale and longer RTTs inherent in WAN environments. Challenges and corresponding requirements arise such as:

1)Backpressure from PFC: The widespread application of PFC in large-scale networks can lead to head-of-line blocking, deadlocks, and congestion spreading, which degrade network throughput. Such challenges make the traditional PFC backpressure mechanisms poorly suited for the high stability demands of WANs, necessitating innovation in protocol design to alleviate issues like deadlocks and PFC storms.

Requirement 1: Innovate and improve upon the PFC backpressure mechanism for WANs, addressing and mitigating the risk of deadlocks and congestion spreading to ensure stable and lossless data transmission.

2)ECN-Based Congestion Control Limitations: While ECN facilitates sender rate control through network collaboration, its effectiveness diminishes over longer distances typical of WANs. The delayed congestion notifications result in prolonged control loops, making it challenging to quickly alleviate congestion.

Requirement 2: Optimize the ECN control loop for WANs, enhancing the network’s ability to manage congestion through improved routing and control strategies, thereby ensuring efficient and lossless transmission across vast geographical distances.

These challenges underscore the need for tailored solutions that address the unique demands and conditions of WANs. By adapting and innovating on existing lossless transmission technologies from data center networks, the goal of achieving extremely low or zero packet loss in WANs becomes attainable, paving the way for enhanced data mobility and application performance.





## 5. Security Considerations

This document does not introduce any new security considerations.

## 6. IANA Considerations

This document has no IANA actions.

## 7. Informative References

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001,

<https://www.rfc-editor.org/rfc/rfc3168>.

[RoCEv2] "Supplement to InfiniBand architecture specification volume 1 release 1.2.2 annex A17 - RoCEv2 (IP routable RoCE).", n.d..

[DCQCN] et.al., Y. Z., "Congestion Control for Large-Scale RDMA Deployments", August 2015, <https://conferences.sigcomm.org/sigcomm/2015/pdf/papers/p523.pdf>.

[PFC] "IEEE Standard for Local and metropolitan area networks--Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks--Amendment 17- Priority-based Flow Control", n.d..

[ESnet] "Energy Sciences Networks", n.d..

## Acknowledgements

TBD.

## Contributors

TBD.

## Authors' Addresses





Zhengxin Han (editor)
China Unicom
Beijing
China
Email: hanzx21@chinaunicom.cn

Tao He
China Unicom
Beijing
China
Email: het21@chinaunicom.cn

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

Tianran Zhou
Huawei
Email: zhoutianran@huawei.com