hpwan
Internet-Draft
Intended status: Standards Track
Expires: 12 December 2025

Q. Xiong
G. Huang
ZTE Corporation
K. Yao
China Mobile
C. Lin
New H3C Technologies
10 June 2025

Framework for High Performance Wide Area Network (HP-WAN)
draft-xhy-hpwan-framework-02

Abstract

This document defines a framework to enable the host-network collaboration for high-speed and high-throughput data transmission, coupled with fast completion time of High Performance Wide Area Networks (HP-WAN). It focuses on key congestion control functions to facilitate host-to-network collaboration and perform rate negotiation, such as QoS policy, admission control, and traffic scheduling.

Table of Contents

1. Introduction

Data-intensive applications always demand high-speed data transmission over WANs such as scientific research, academia, education as discussed in [I-D.kcrh-hpwan-state-of-art] and other applications in public networks as per [I-D.yx-hpwan-uc-requirements-public-operator]. The specific requirements of HP-WANs applications mainly focus on job-based massive data transmission over long-distance WANs, with set completion times. High, reliable and effective data throughput is the fundamental requirement for HP-WAN. It is crucial to achieve high throughput while ensuring the efficient use of capacity as per [I-D.xiong-hpwan-problem-statement]. Current technology does not guarantee these goals, and the issues may impact performance related to existing transport protocols and congestion control mechanisms such as poor convergence speed, long feedback loop, and unscheduled traffic.

High-level requirements for HPWAN can be summarized as:

*Multiple data transfer requests should be scheduled in terms of available capacity and the requested completion time in terms of transmission performance;

*From the routing aspect, the optimal path and resources should be scheduled based on the QoS policy for the high-speed flows to travel through the network with the negotiated data transfer rate;

*From the transport aspect, it ensures the reliable delivery of data with traffic scheduling and admission control to effectively handle the flow of data during transmission, reducing congestion and ensuring timely delivery of data packets;

*The host should consider signalling and collaborating with the network to negotiate the rates of differentiated traffic (especially when the traffic is encrypted) to avoid the congestion and optimize the overall efficiency of data transfer.

This document defines a framework for these requirements, including the signaling goals to enable the host-and-network collaboration for the high-speed and high-throughput data transmission, coupled with fast completion time in High Performance Wide Area Network (HP-WAN). It particularly enhances the congestion control and facilitates the functionalities for the host to collaborate with the network to perform rate negotiation, such as QoS policy, admission control and traffic scheduling.

2. Definition of Terms

This document uses the terms defined in [I-D.kcrh-hpwan-state-of-art] and [I-D.xiong-hpwan-problem-statement]:

3. Framework for HP-WAN

3.1. Overview

The framework is formulated to enable the host-network collaboration upon more active network involvement. The client and server could adjust the rate efficiently and rapidly with the negotiated rate-based congestion control in a fine-grained way. The network could enhance the capability to regulate the traffic and schedule the resources which could provide predictable network behaviour and mitigate incast network congestion preemptively.

The following diagram illustrates the functionalities between Client/ Server and WAN including:

*Host-network collaboration signalling or protocol

*Active network-collaborated traffic enforcement and scheduling

*Negotiated rate-based congestion control algorithms

```
                          +-------------------------------+
                          |             WAN               |
   +--------+             |                               |              +--------+
   |        |        +----+----+   +-------------+   +----+----+          |        |
   | Client |<------>|Edge Node|...|Transit Nodes|...|Edge Node|<-------->| Server |
   |        |        +----+----+   +-------------+   +----+----+          |        |
   +--------+             |                               |              +--------+
       *collaboration     |                               |     *collaboration
      signalling/protocols|                               |     signalling/protocols
                          +-------------------------------+
  \_________/              \______________________________/
*Negotiated rate-based      *Active network-collaborated
congestion control            enforcement and scheduling
algorithms
```

Figure 1 HP-WAN framework

3.2. Workflow and Functions

The following diagram illustrates the workflows among client, server and network nodes (e.g. edge nodes and transit nodes).

*The request of scheduled traffic will be signalled from the client to the network based on the negotiated rate. Furthermore, the traffic pattern and job-based requirements, such as completion time, should be included in the request.

*The edge node will perform admission control and acknowledge the traffic, reserving the resource quota, but it will reject access when the network capacity cannot guarantee the job’s completion time.

*The acknowledgement will be signalled back from the network to the client, including the response with the negotiated rate and QoS policy for the client to send traffic.

*The notification will be signalled from the client to the network to notify the completion of traffic, and the network will release the resource quota and cancel the acknowledgement of this job.

*The update may signal to the client from the network to update the acknowledgement of the negotiated rate when new traffic requests are received.

```
 +--------+              +-----------+       +------------+       +-----------+
+--------+
 | Client |              | Edge Node |       |Transit Node|       | Edge Node |
| Server |
 +----+---+              +-----+-----+       +-----+------+       +-----+-----+
+----+---+
      |                        |                   |                    |
     |
      |Requests(traffic pattern)|                   |                    |
     |
      |------------------------>|*Rate negotiation  |                    |    Requests
     |
      |                        |*Traffic scheduling|                    |-------------
---->|
      |Acknowledgement          |*Admission control |                    | Acknowledgem
ent  |
      |(negotiated rate)        |        *Reserve resource quota        |<------------
-----|
      |<-----------------------|*Negotiated rate-based traffic engineering|
     |
      |                        |<#######################################>|
     |
      |New Request              |                   |                    |
     |
      |----------------------->|                   |                    |
     |
      |Update(negotiated rate)  |                   |                    |
     |
      |<-----------------------|                   |                    |
     |
      |Notification(completion) |                   |                    | Notification
     |
      |----------------------->|        *Release resource quota        |-------------
---->|
      |Acknowledgement(cancel)  |<#######################################>| Acknowledgem
ent  |
      |<-----------------------|                   |                    |<------------
-----|
      |                        |                   |                    |
     |
      V                        V                   V                    V
     V
```

Figure 2 The workflow of signalling between host and network

The client could send traffic according to the negotiated rate policy to achieve a high throughput within the completion time. And the edge node will send fast feedback with the advised rate when the traffic rate does not apply to the network. It could also pause the traffic when congestion occurs (e.g. the traffic is exceeding the threshold of the server, the network performs the flow control).

```
 +--------+                 +-----------+   +------------+  +-----------+            +--
------+
 | Client |                 | Edge Node |   |Transit Node|  | Edge Node |            | S
erver |
 +----+---+                 +-----+-----+   +-----+------+  +-----+-----+            +--
--+---+
      |                           |               |               |
  |
      |                           |               |               |
  |
      | Traffic(negotiated-rate)  |    Traffic(negotiated-rate)   |Traffic(negotiated-rat
e)|
      |-------------------------->|***************************>|---------------------
->|
      |                           |               |               |   Exceeding threshold
  |
      |                           |               |*Flow control  |<---------------------
--|
      |                           |*Flow control  |<--------------|
  |
      | Fast Feedback(pause)      |<--------------|               |
  |
      |<--------------------------|               |               |
  |
      |  Traffic(wrong-rate)      |               |               |
  |
      |-------------------------->|               |               |
  |
      |Fast Feedback(advised-rate)|               |               |
  |
      |<--------------------------|               |               |
  |
      |                           |               |               |
  |
      V                           V               V               V
  V
```

Figure 3 The workflow of traffic between host and network

The functions are described in the sections below including transport-related technologies such as rate negotiation, admission control, traffic scheduling and enforcement and routing-related technologies like traffic engineering, resource scheduling and load balancing.

3.2.1. Rate Negotiation

In HP-WAN, the host could negotiate the sending rate with the network due to the predictability of jobs. The client communicates the traffic patterns of high-speed flows to the network to negotiate rate. The traffic patterns may cover traffic information such as job ID, start time, completion time, data volume, traffic type and so on. The network responds to the negotiated rate and QoS policy for the client to send traffic. There are three kinds of rate policy as follows:

*Optimal rate or optimal rate range negotiation. The network provides resource reservation for high-speed data to guarantee the transmission capacity and achieve optimal rate transmission. The client could transmit flows according to the negotiated optimal rate or optimal rate range.

*Minimum rate negotiation. The network provides the minimum resource guarantee. The client could transmit at a rate not less than the negotiated rate.

*Maximum rate negotiation. The network provides an upper limit for resource guarantee. The client could transmit at a rate not greater than the negotiated rate.

3.2.2. Admission Control

The network node should perform admission and traffic control based on negotiated QoS and rate. By combining the admission control with congestion control, it can provide high throughput associated with completion time while efficiently using the available network capacity. The strategies of admission control are different based on the QoS policy. For example, one strategy is to immediately grant or reject admission to a reservation request on its arrival time, which is called on-demand admission control. If a reservation request can not be granted or rejected at the time of its arrival, it will be put in a queue, which is called queue-based admission control. Furthermore, a time-slot based admission control is used for scheduling the elastic and flows requests.

3.2.3. Traffic Scheduling and Enforcement

The network node (e.g.edge node) performs rate-based traffic scheduling and enforcement. For example, traffic classification may be needed based on the traffic type. If it needs to prioritize critical traffic for acceleration, it should upgrade the priority of QoS. Moreover, if the traffic needs a guaranteed QoS, it should provide guaranteed bandwidth for this flow. It also could perform the aggregation of mouse flows or the fragmentation of an elephant flow if needed. Splitting data across multiple paths for load balancing can increase the throughput and provide redundancy. If one path experiences congestion, alternate paths compensate, ensuring timely delivery. The traffic enforcement at network edges can used to regulate data flow to eliminate congestion and minimize the flow completion time. For example, it could enforce the rate limits based on the negotiated rate to access traffic.

3.2.4. Optimization of Congestion Control Algorithms

The client should perform the improvement of congestion control algorithms based on the negotiated-rate from the network. The negotiated-rate can be viewed as an initial congestion signal to assist the client in selecting a suitable sending rate with the network resource scheduling acknowledgement. And it also needs to turn off and on or adjust the rate reasonably and rapidly when receiving the fast feedback from the node nearing the client.

3.2.5. Negotiated Rate-based Traffic Engineering

The negotiated rate-based traffic engineering should be provided by routing technologies and the signaling from client will assist the network operator’s traffic management and corresponding resource planning and scheduling. The edge node may get information (topology, bottleneck link bandwidth, queue and buffer) from a centralized controller or through IGP advertisement. The network should provide resource scheduling at nodes along the path and it is not bandwidth allocation but quota reservation which can be used for admission control. The client and network can also negotiate rate based on the quota of each job. Quota is expressed as a vector of resource quantities (bandwidth, buffer, queue, etc.) at a given priority, for a time frame. The network can make dynamic bandwidth reservation upon different time frames defined by quota. It will differ based on the different QoS policy. For example, it is required to reserve the minimum bandwidth quota for the minimum rate policy.

3.2.6. Fast Feedback

The fast feedback function is optional for HP-WAN. The edge node will send fast feedback with the advised rate when the traffic rate is not applicable to the network. It could also pause traffic when congestion occurs and resume it when congestion is mitigated.

3.2.7. Flow Control

The specific elements along the path may be optional to provide active and precise flow control to mitigate network congestion to control the packet loss. Flow control refers to a method for ensuring the data is transmitted efficiently and reliably and controlling the rate of data transmission to prevent the fast sender from overwhelming the slow receiver and prevent packet loss in congested situations. For example, the receiver node could signal the sender node to control the traffic on or off to guarantee the packet loss. When the data sent by the client exceeds the threshold, the network should provide fast and accurate quantitative feedback to control the traffic on or off.

4. Applicability of Host-network Collaboration Signalling

There are several existing signalling options for HP-WAN host-network collaboration signalling such as RSVP and GRASP. There will be two deployment scenarios in HP-WAN. The first one will be the central controller deployment which will have a hierarchical planning and resource reservation in the network like CERN deployment and the SENSE architecture. In this case, the host-newtork signalling (between client and edge node) may be peer-to-peer solution and both GRASP and RSVP may be applicable. And the second case will be distributed or hybrid deployment in the network which needs distributed signalling along the path for resource reservation. In this case, the host may signal from the client to the network nodes along the path. RSVP may be applicable but not GRASP.

GRASP is peer-to-peer signalling and is designed for synchronization and negotiation between autonomic service agents, which reduces the need for hierarchy and allows the intelligence to be distributed rather than centralized. However it is not applicable when the signalling should be performed along the end-to-end path.

Although RSVP may not be deployable with complex configuration and management which requires precise configuration across all network devices along the path. It will also add administrative complexity between host and network in HP-WAN with operational issues. But SR, slicing, diffServ QoS and SDN-based approaches may be used to largely improve RSVP in HP-WAN. Moreover, RSVP reservations often allocate fixed resources in the nodes along the path, which can lead to underutilization if the reserved resources are not fully used. The extensions may be required to applied to HP-WAN that the bandwidth and rate vary over time and it requires scalable throughput, dynamic bandwidth reservation and efficient use of capacity.

5. Security Considerations

To be discussed in future versions of this document.

6. IANA Considerations

Currently this document does not make an IANA requests.

7. Informative References

[I-D.kcrh-hpwan-state-of-art]
King, D., Chown, T., Rapier, C., and D. Huang, "Current State of the Art for High Performance Wide Area Networks", Work in Progress, Internet-Draft, draft-kcrh-hpwan-state-of-art-01, 8 January 2025, <https://datatracker.ietf.org/doc/html/draft-kcrh-hpwan-state-of-art-01>.

[I-D.xiong-hpwan-problem-statement]
Xiong, Q., Yao, K., Huang, C., Zhengxin, H., and J. Zhao, "Problem Statement for High Performance Wide Area Networks", Work in Progress, Internet-Draft, draft-xiong-hpwan-problem-statement-02, 25 February 2025, <https://datatracker.ietf.org/doc/html/draft-xiong-hpwan-problem-statement-02>.

[I-D.yx-hpwan-uc-requirements-public-operator]
Yao, K. and Q. Xiong, "High Performance Wide Area Network (HPWAN) Use Cases and Requirements -- From Public Operator’s View", Work in Progress, Internet-Draft, draft-yx-hpwan-uc-requirements-public-operator-00, 20 February 2025, <https://datatracker.ietf.org/doc/html/draft-yx-hpwan-uc-requirements-public-operator-00>.

Authors’ Addresses

Quan Xiong
ZTE Corporation
Email: xiong.quan@zte.com.cn

Guangping Huang
ZTE Corporation
Email: huang.guangping@zte.com.cn

Kehan Yao
China Mobile
Email: yaokehan@chinamobile.com

Changwang Lin
New H3C Technologies
Email: linchangwang.04414@h3c.com