Resource Reservation Protocol for IP Transport QoS

draft-han-tsvwg-ip-transport-qos-00

Lin Han (lin.han@huawei.com)

Yingzhen Qu (yingzhen.qu@huawei.com)

Lijun Dong(lijun.dong@huawei.com)

Thomas Nadeau (tnadeau@lucidvision.com)

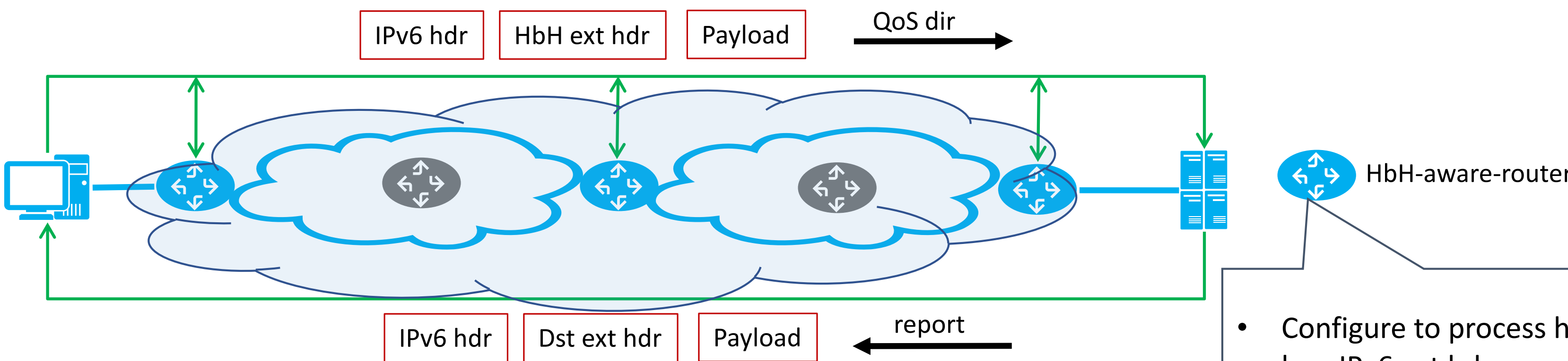
Kevin Smith (kevin.smith@vodafone.com)

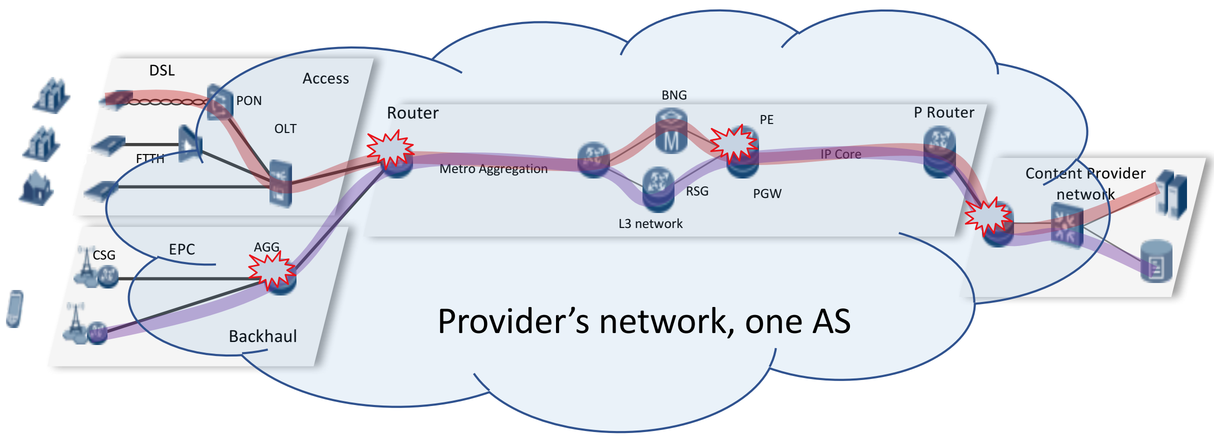
Introduction

- Presented in IETF 100. This is the re-written draft for TSVWG.
- The presentation will answer some comments and give more details.
- Objective

  A simpler/faster/more scalable resource reservation protocol to achieve bandwidth and/or bounded-latency guaranteed QoS for IP flow(s) that need this service.
- Solution: In-band signaling by IPv6 extension header. Not associated with specific QoS implementation.
- Design principles
  - Backward compatible, coexist with current services
  - Agnostic to transport layer protocols
  - Practical performance and scale targets
  - Basic signaling and data security
- Scope and assumptions
  - Targeted for applications that are bandwidth and/or latency sensitive
  - Within one service domain
  - Limited scalability requirement

How it works





- Configure to process hop-by-hop IPv6 ext hdr
- Only needed for throttle devices
- Signal processing distributed to NPU on line card
- Per flow state kept in line card
- Per flow state self-maintained by data flow, deleted if aged
- Basic security for signaling and data forwarding are provided

Flow level QoS and Aggregated flow QoS

- Flow level

  Identified by 5 tuples: source and destination address, protocol number, source and destination port number. or 3 tuples: source and destination address, and flow label

- Transport level

  Packets share the same source and destination address, and protocol number, e.g. TCP or UCP flows that started and terminated at the same IP addresses

- Address Level

  Packets share the same source, destination IP address, but with different protocol number.

- DiffServ Level

  Packets share the same DSCP value

Scalability and Performance Analysis

- **Distributed Processing**: No extra protocol, such as RSVP run by CPU. In-band signal processing is distributed in NPUs on line cards.
- **Modern Hardware Architecture:** More ports or higher throughput for a system, more NPUs are used. This means the system scalability and performance is almost not changing with the growth of the number of transport sessions.

Scalability example

- The scalability is related to the queue supported on NPU. More flows, more queue needed
- Industry fastest NPU – port speed: 400 G.
- If 50% of link capacity (200G) is for TCP that needs resource reservation
- and per TCP flow requires 100M bandwidth
- There are only 200G/100M = 2000 flows need the in-band signalling processing, and associated QoS
- Normally, there is no problem for NPU to support more than 2000 flows (queues) on a NPU

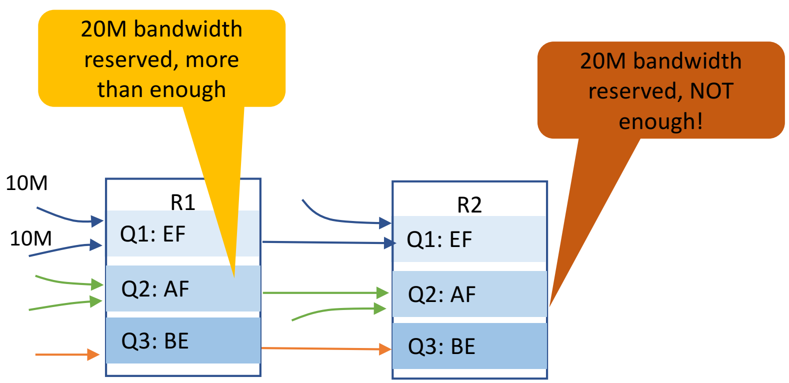
This is a conservative example. In reality, there will be smaller number of larger flows that needs this QoS service. i.e, the AR service needs much higher bandwidth than 100M, and flow number < 2000

Per-Hop Behavior DiffServ

Manual configure needed
No way to accurately know the bandwidth
Configuration example (100M interface):

```
Class-map class1
  match dscp af42
Policy-map policy1
  bandwidth percent 20
```

Bandwidth reservation is configured based on estimation

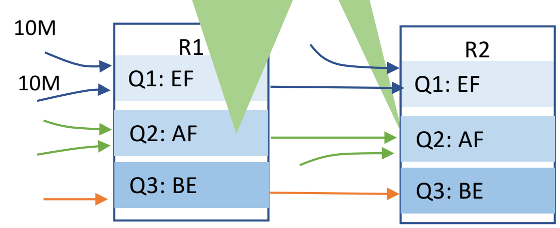


IP QoS with Resource Reservation

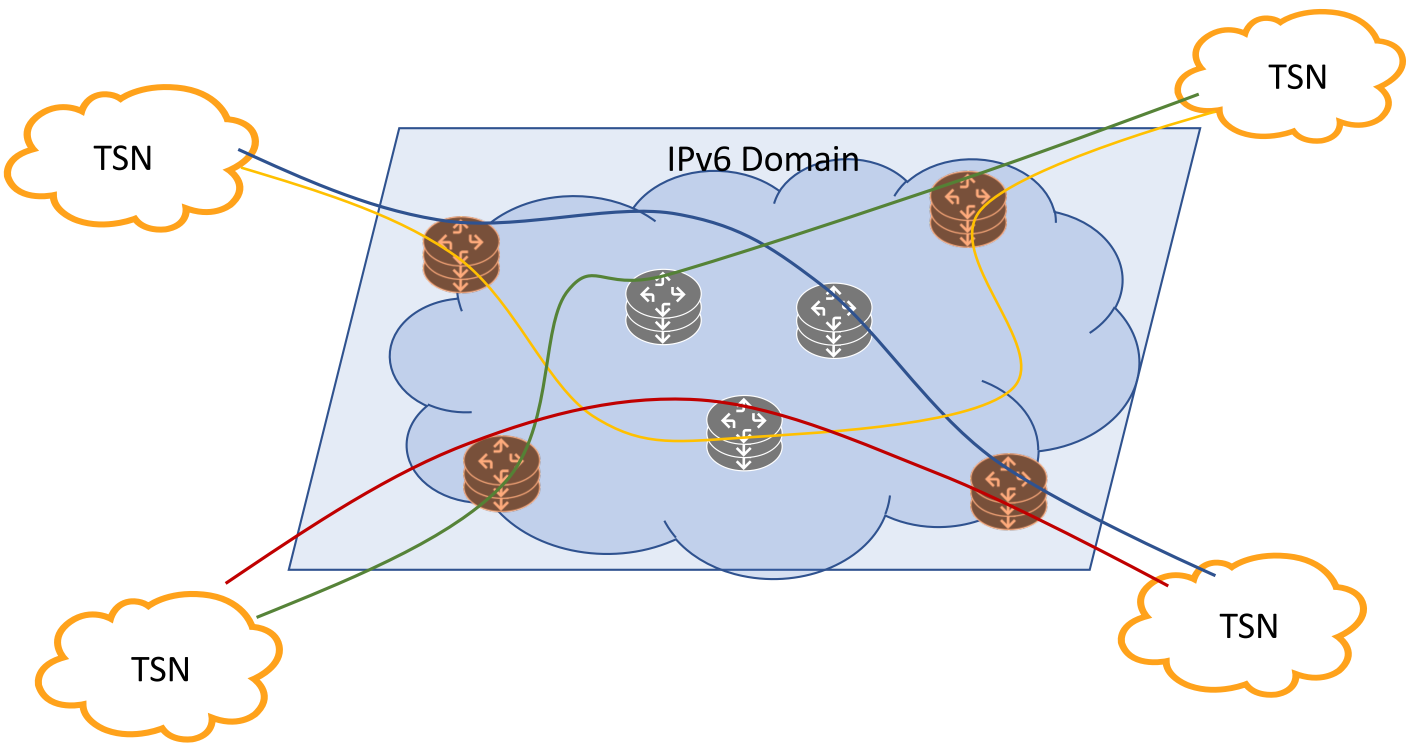
Per flow state makes it possible to have accurate QoS control for DiffServ

$$Wq = \frac{f1+f2+..fn}{B}$$

Queuing scheduling is dynamically adjusted based on flows.



Use case 1 - Detnet



The protocol makes the per-flow state available and easily maintained on device, this is the key to the realization of bounded-latency in Detnet.

TSN interconnect using IPv6:

- Guaranteed bandwidth
- Guaranteed and predictable minimum per-hop-latency.
- No MPLS/LDP needed

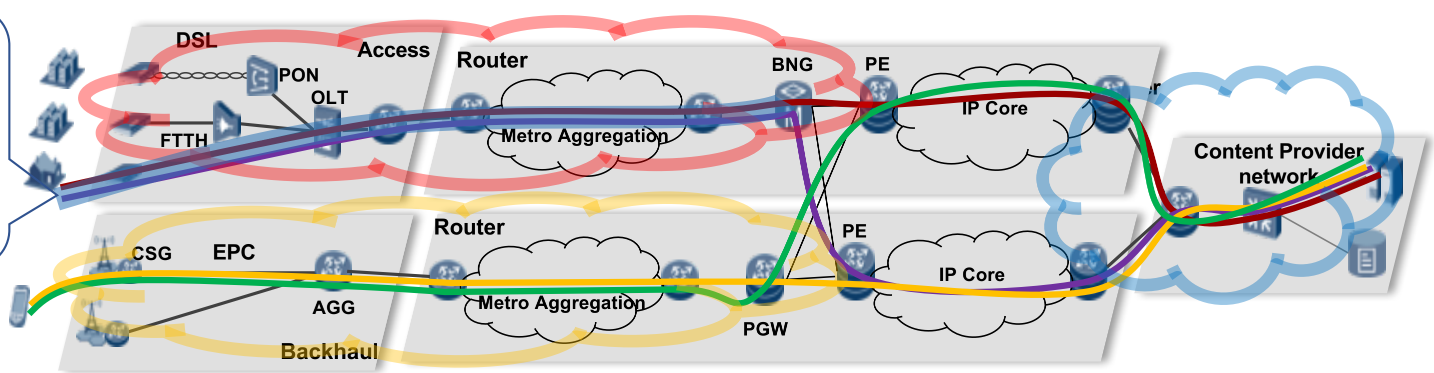
Two possible working modes:

- Aggregated mode: Encap/decap at gateway routers, can be used to connect IPv4 networks or private address spaces
- Native mode: TSN network routes populated to IPV6 domain

Use Case 2 - PANRG

- QoS for each MPTCP sub-flow in a access network through resource reservation protocol.
- Overcome the constraint of MPTCP fairness principal (Multipath TCP should take as much capacity as TCP at a bottleneck link, no matter how many paths it is using*)*
- Integrated with multi-path in Internet to support MPTCP, and Bringing path-aware networking in current Internet that is not path-aware

QoS for segments of sub-flows by the protocol



Q&A

More detailed works in
ETSI NGP (Next Generation Protocol, WI#10: New transport technology):
https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=52932