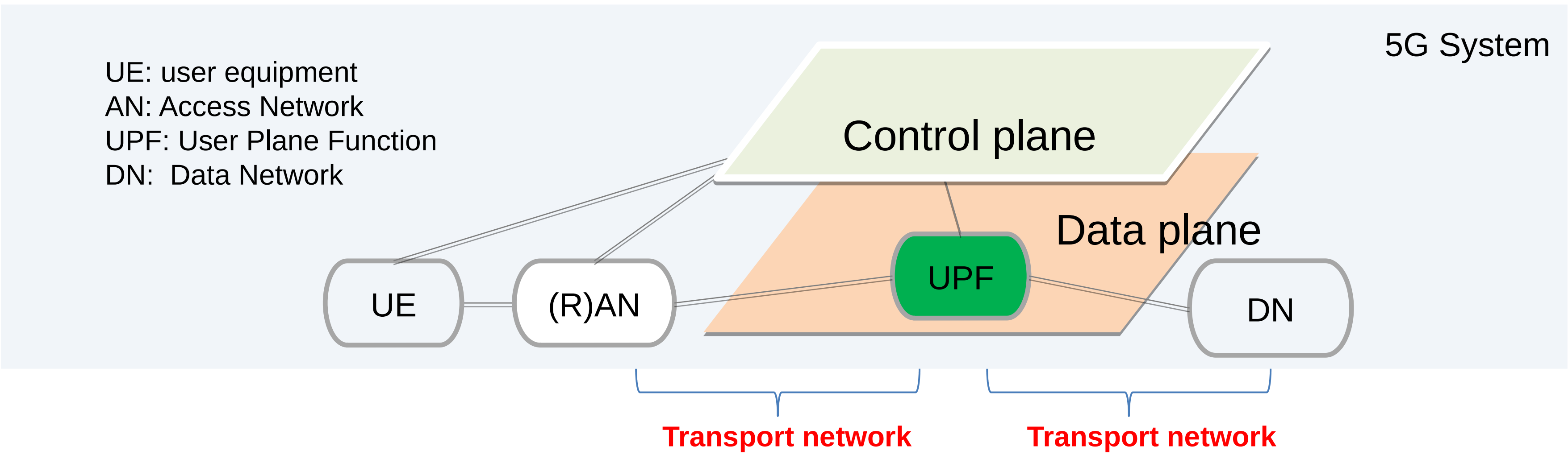
Micro-burst Decreasing in Layer3 Network for Low-Latency Traffic (revised 1206)

draft-du-detnet-layer3-low-latency-01

Zongpeng Du duzongpeng@chinamobile.com
Peng Liu liupengyjy@chinamobile.com
Liang Geng gengliang@chinamobile.com

IETF109

Motivation in the 5G bearing



- The Figure shows a simple architecture including the transport network under the 5G Core
  - Transport networks provide connection service for 5G Data plane and Control plane, and are also called the **5G bearing network**
- Though not in the standardization scope of 3GPP, some of the 5G requirements also motivate the evolution of the 5G bearing network, which is often **IP-based**

Deterministic requirement in operator's 5G bearing network

- URLLC (Ultra-reliable and Low Latency Communications) is one of the key aspects in 5G
  - It is an **E2E** requirement, so that the transport network (the bearing fixed network) also needs to provide Low Latency service
- However, it is **hard to provide deterministic service** in the current 5G bearing network
  - Traditional IP network is based on statistical multiplexing, and can only provide **Best Effort** service, short of SLA guaranteed mechanisms

Related works in IEEE and IETF

- IEEE 802.1 Time-Sensitive Networking (**TSN**) Task Group (TG) has completed some projects for providing deterministic services through IEEE 802 networks
  - i.e., guaranteed packet transport with bounded latency, low packet delay variation, and low packet loss
- IETF **DetNet** Working Group also has done a lot of jobs, to enable deterministic data paths that operate over Layer 2 bridged and Layer 3 routed segments
  - ~~However,~~ DetNet now is supposed to work in campus-wide networks and private WANs, ~~and hasn't covered the large-scale ISP network scenario~~ which inclue the large-scale ISP network scenario, such as the 5G bearing network
  - ~~In summary,~~ However, there are still **some gaps** for 5G bearing network to become deterministic

Gaps for the large-scale layer 3 deterministic network

- Currently, we are short of **common and mature mechanisms** that can provide deterministic transport in the 5G bearing network
  - Draft *qiang-detnet-large-scale-detnet* provides a potential solution based on some enhancements to the CQF in TSN
    - We think it is a good exploration, but more explorations are needed
- TSN mechanisms are designed for L2 network originally, and cannot be directly used in the large-scale L3 network because of various reasons
  - In which the main aspects are the **simplicity** and the **scalability**
  - The former can ensure that the mechanism is easy to deploy, and the second can ensure that the mechanism is able to fulfill the complex requirements

Why TSN mechanism can not be directly used in large-scale L3 network

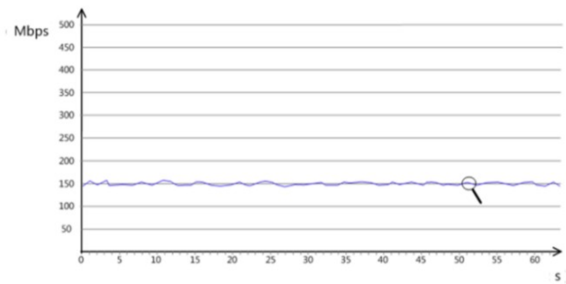
- Some TSN mechanisms need **synchronization** of the network equipment, which is easier in a small network, but hard in a large network
  - i.e., it brings in some complex maintain jobs across a large distance that is not needed before
- Some TSN mechanisms need a **per-flow state** in the forwarding plane
  - which is un-scalable
- Some TSN mechanisms need a **constant** and forecastable **traffic characteristic**
  - But it is more complicated in a large network
  - which includes much more flows joining in or leaving randomly and the traffic characteristics are more dynamic

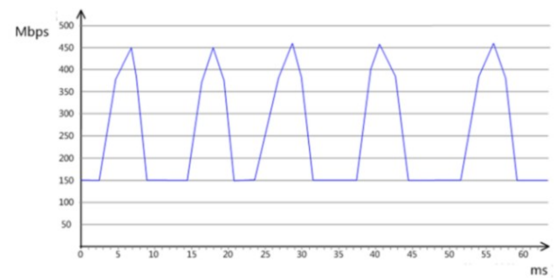
Rethinking the problem in the current IP forwarding mechanism

- The main problem is that
  - In the current IP network, a long delay in queuing, or some **packet losses** due to burst are acceptable
  - But it is **unacceptable** in the deterministic forwarding
- Even be given a high priority, a packet can experience a long congestion delay or be lost in a relatively light-loaded network, because of micro-burst in the network
  - It is a special case of network congestion, which typically lasts a short period, at the granularity of millisecond
  - In a micro-burst, a lot of data are received on the interface suddenly, and the temporary bandwidth requirement would be tens of or hundreds of the average bandwidth requirement, or even exceed the interface bandwidth

Introduction of micro-burst

- In the currently transport network, if we observe the traffic utilized bandwidth on the interface at a granularity of per-second, a smooth curve may be obtained

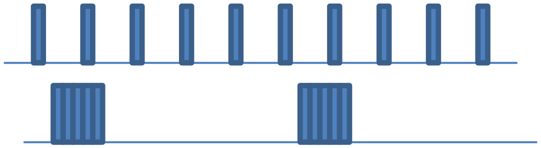




- However, if we observe the traffic utilized bandwidth at a granularity of **per-millisecond**, we can see micro-bursts on the interface
  - **In most cases, the buffer on the equipment can handle the micro-bursts**
  - However, in some corner cases, micro-bursts bring in a long delay (at the granularity of millisecond) or even packet loss

Causes of the micro-burst

- Firstly, IP traffic has a instinct of **burstiness** no matter in the macro or micro aspect
  - i.e., it does not have a constant traffic model even after aggregations
- Secondly, IP network has a flexible topology, where the incoming traffic may exceed the bandwidth of the outgoing interface
  - For example, an interface with a large bandwidth may need to send traffic to an interface with a smaller bandwidth
  - or **multiple flows** from several incoming interfaces may need to occupy the **same outgoing interface**
- Thirdly, the IP node has been designed to send traffic **as quickly as possible**, and it is not aware whether the downstream node's buffer can handle the traffic
  - For example, the figure below shows the problem of the current IP scheduling mechanism. Before the scheduling in an IP network, the packets are well paced, but after the scheduling, the packets will be gathered even the total traffic rate is unchanged

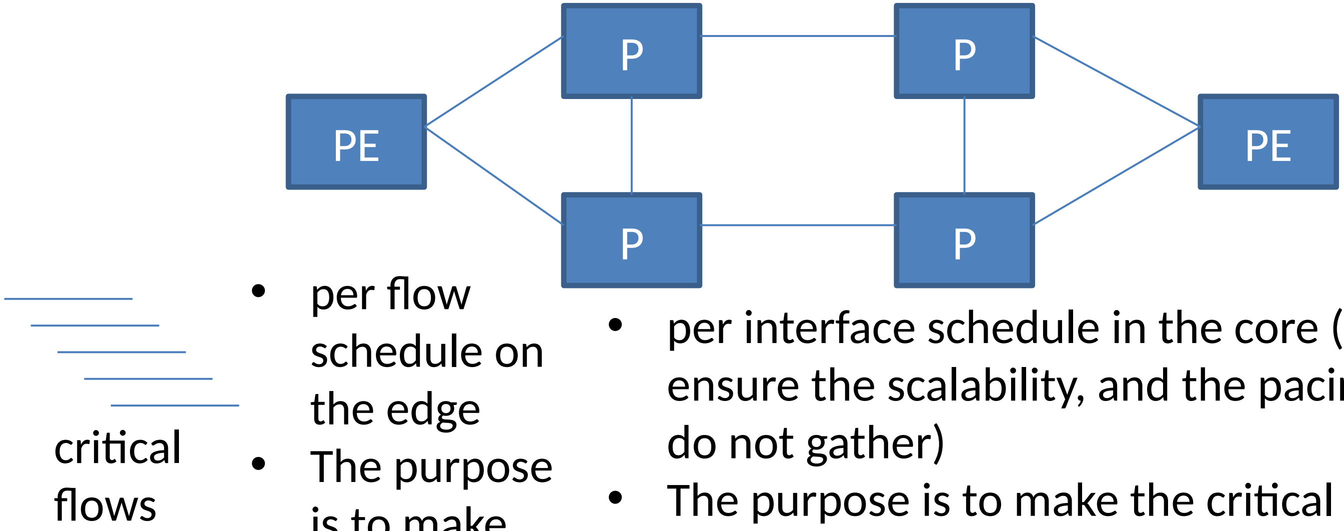


A potential idea to decrease micro-burst

- In fact, mechanisms in TSN can handle the micro-burst for the deterministic traffic, but **at a relatively higher cost** than the current IP forwarding
- Our draft-du-detnet-layer3-low-latency proposes a mechanism:
  - In the PE node, do the per-flow shaping and color all the deterministic traffic
  - In the P node, recognize the deterministic traffic and do a per-interface shaping
    - The purpose is to **maintain a proper buffer depth on the P node**, and forward the deterministic packets orderly, **instead of as quickly as possible**
- Future jobs include:
  - firstly, we need a mechanism in the control plane that guarantees the total bandwidth on the interface is not exceeded
  - secondly, we need a convenience way for the service to inform the network about the needed bandwidth
  - thirdly, the deterministic traffic should be colored so that the P node can recognize them

Thanks for listening and
Welcome for comments

Explanations of the proposed mechanism



- per flow schedule on the edge
- The purpose is to make the critical traffic **have a constant traffic model**

- per interface schedule in the core (traffic are aggregated to ensure the scalability, and the pacing also makes sure they do not gather)
- The purpose is to make the critical traffic are forwarded as **the shape when outgoing the edge**, not as quickly as possible
  - We assume the sending rate of the buffer for the critical traffic is the same as the receiving rate (maybe an algorithm is needed here). If all works good, the buffer will be maintained with a proper depth
- Others requirements:
  - RSVP liked mechanisms with a good scalability should be used to make sure the bandwidth is not exceeded on the interface

Explanations of the potential mechanisms

Mechanisms	Easy to deploy	Scalability	100% low-latency assurance for critical traffic
Traditional IP forwarding	easy	good	No
Traditional IP forwarding with priorities	easy	good	good if only a small amount of critical traffic
The mechanism in our draft (also assumed a high priority for critical traffic)	relatively easy, per flow schedule in the edge, per interface schedule in the core	relatively good	good if the bandwidth is reserved, and the critical traffic are forwarded in order
TSN mechanisms defended by IEEE	normally No	normally not good, designed for L2	Yes