Large-Scale Deterministic Network Update

draft-qiang-detnet-large-scale-detnet-04

Cristian Qiang (qiangli3@huawei.com)
Xuesong Geng (gengxuesong@huawei.com)
Bingyang Liu (liubingyang@huawei.com)
Toerless Eckert (tte+ietf@cs.fau.de)
Liang Geng (gengliang@chinamobile.com)

Requirements for this work

Bounded Jitter requirement from DetNet

Charter: *The Deterministic Networking (DetNet) Working Group focuses on deterministic data paths that operate over Layer 2 bridged and Layer 3 routed segments, where such paths can* ***provide bounds on*** *latency, loss, and packet* ***delay variation (jitter)****, and high reliability*

Architecture: Primary goals defining the DetNet QoS: Minimum and maximum end-to-end latency from source to destination; timely delivery, **and bounded jitter (packet delay variation)** derived from these constraints.

Asynchronous Traffic Shaping only provides upper bound on delay

Jitter ~= propagation delay variation = [0 … max-end-to-end-queuing-delay] - “”no lower bound””!

Network Speed: DetNet needs to support fast networks with 100++ Gbps transit link speeds (common outside building/campus networks).

Need a per-link QoS option that is viable at this speed.

Network link speed not determined by only one traffic class (DetNet)

Other

Easily calculated delay/jitter (centralized or distributed)

Badly behaving links (jitter)

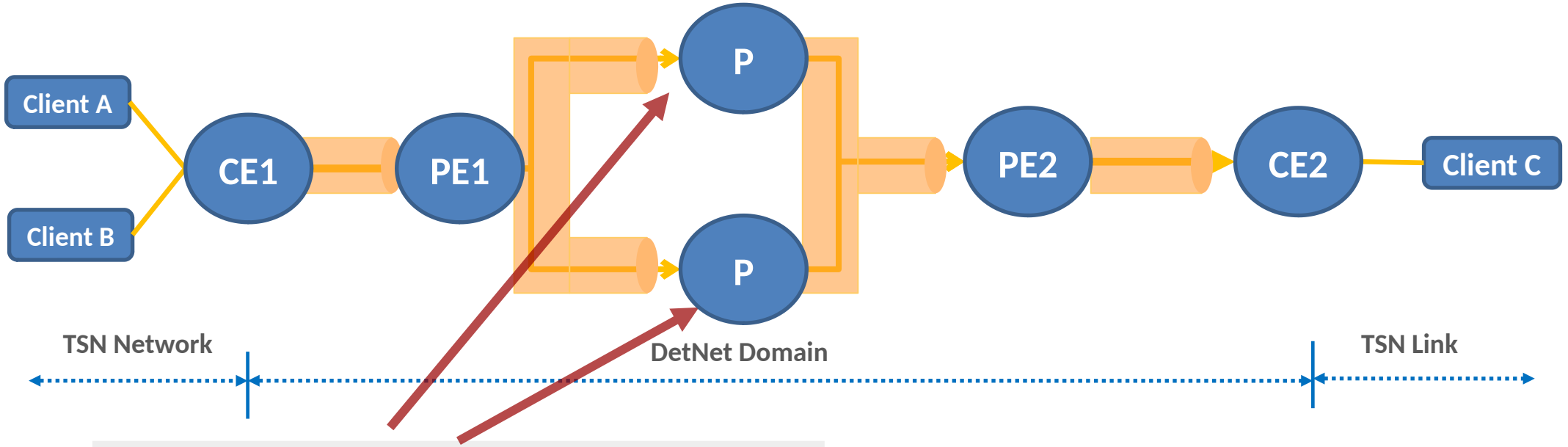
Jitter of mechanism not subject to link jitter but per-link cycle-mapping (shape out link jitter)

Goals of this work

- Informational DetNet WG document
- Introduce Requirements and Framework
  - Independent of specific forwarding plane options
  - Generically applicable to DetNet scenarios

- Use as justification and reference for normative work in other WGs
  - TBD, figure out later, but for example:
  - QoS model in TSVWG (PHB?) or DetNet
  - Forwarding plane encodings, TBD, e.g.:
    - IPv4/IPv6 DSCP or UDP extensions TSVWG
    - IPv6 extension header in 6MAN
    - SR/SIF encoding SPRING
    - SR-MPLS/MPLS specific encodings in MPLS
    - …

Simple Scalable Queuing Solution in the Scenario



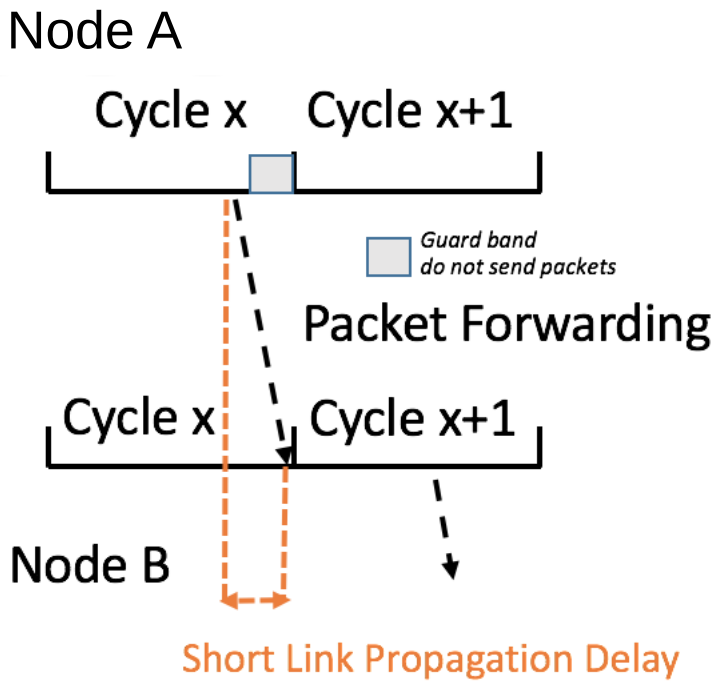
Congestion Protection

Avoiding packet loss and latency because of congestion

- *Resource reservation for DetNet flows*
- ***Queuing Management (Shaping, Scheduling, etc.)***

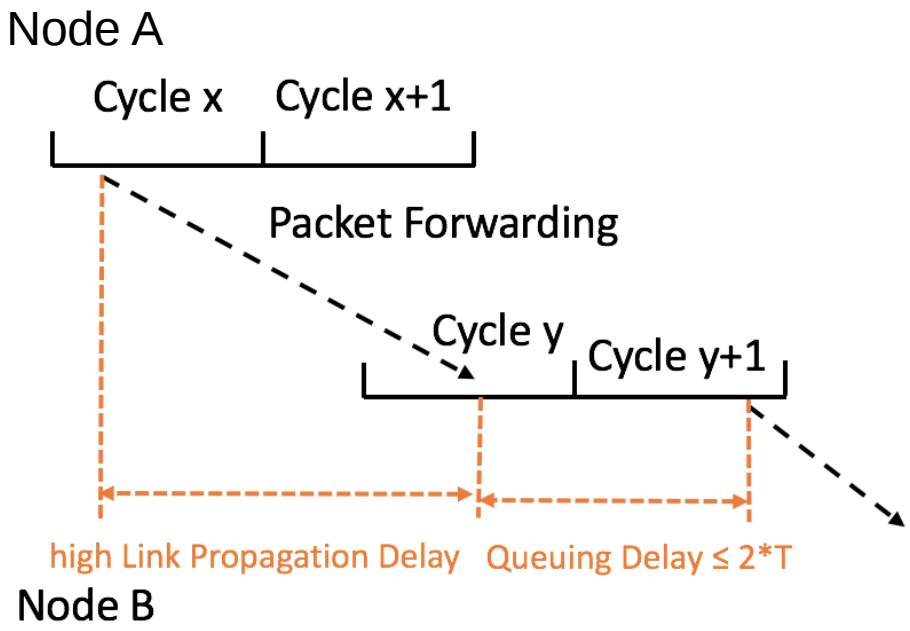
Cyclic forwarding in TSN

- TSN-CQF: synchronous forwarding scheme without per-flow queuing state on every hop
  - Synchronous packets forwarded across all hops within a single synchronized cycle (10..100 usec)
- **Challenges**
  - High accuracy time synchronization requirement (nsec)
  - Limited physical size due to cycle time
  - The larger the network, the smaller the percentage of traffic that can be synchronous.
  - Example (extreme to make point):
    - 10 usec cycle time: max network size: 2 Km, after < 1 Km only < 50% traffic could be synchronous



Proposed Large-Scale Network Cyclic Forwarding

- Carry cycle-identifier in packet
- No synchronous forwarding:
  - Buffer cycles and send after all packets for cycle arrived
- **Results**
  - Keep the key benefits
    - Easy calculated end-to-end-delay (sum(per-hop-cycle-delay))
    - Tight bounded jitter O(cycle-time) [usec]
  - Eliminated physical scale limitations
    - Can support arbitrary link-propagation delay, hop, end-to-end delay
  - Eliminates need for tight time-synchronization
    - Requires only frequency synchronization in order to control drift bet ween adjacent node cycle times (usec instead of nsec)
    - Frequency synchronization is much easier than time synchronization
      - E.g.： no problems with the difficult asymmetric link problem



Major Changes Since Last IETF Meeting

- Add a new figure to illustrate that common IP/MPLS forwarding + Priority Queueing couldn't guarantee bounded latency and delay variance (jitter) d ue to micro-burst and micro-bust iteration

If no control of packet behavior, then the wort cast will be packets from the same priority flows arrive at the same time

Packets wait in the queue, and produce micro-burst

Micro-burst iterating after several hops, no way to guarantee bounded latency and jitter anymore

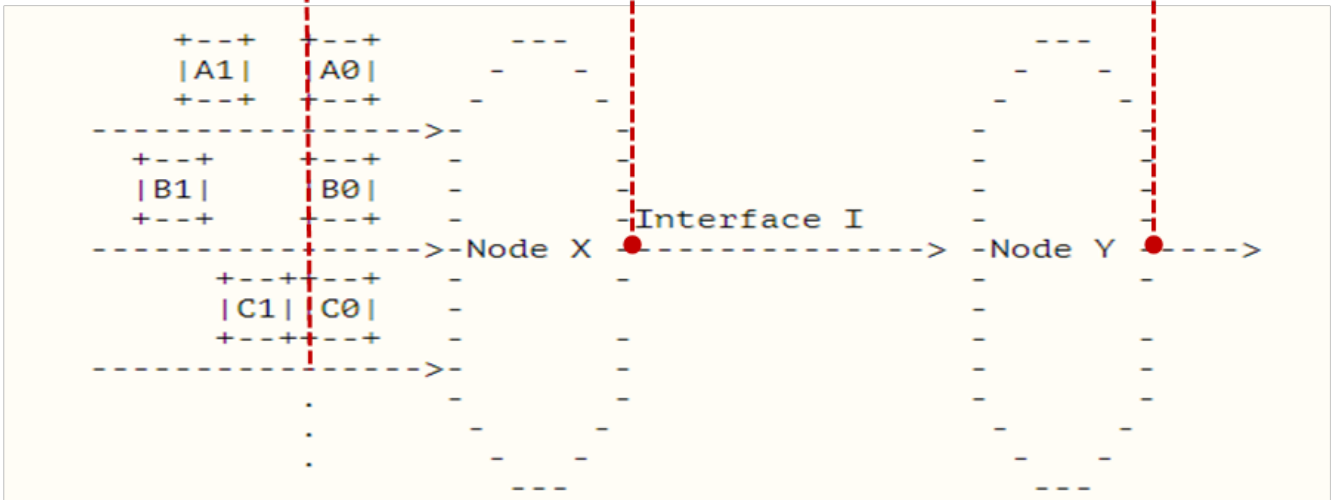


Figure 1: Micro-burst and micro-burst iteration

Major Changes Since Last IETF Meeting

- Use cycle identifier to indicate packet behavior (sending timing) every hop, r efine the potential ways to carry **minimal** 2 bits cycle identifier

List in former version

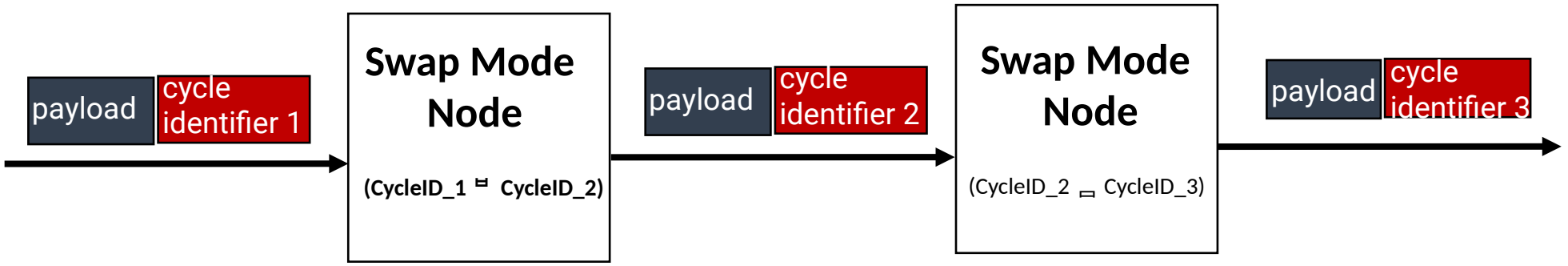
- DSCP of IPv4 Header
- Traffic Class of IPv6 Header
- TC of MPLS Header (used to be EXP)
- EtherType of Ethernet Header — Removed
- IPv6 Extension Header
- TLV of SRv6
- TC of MPLS-SR Header (used to be EXP)
- Three labels/adjacency SIDs for MPLS-SR

List in 04 version

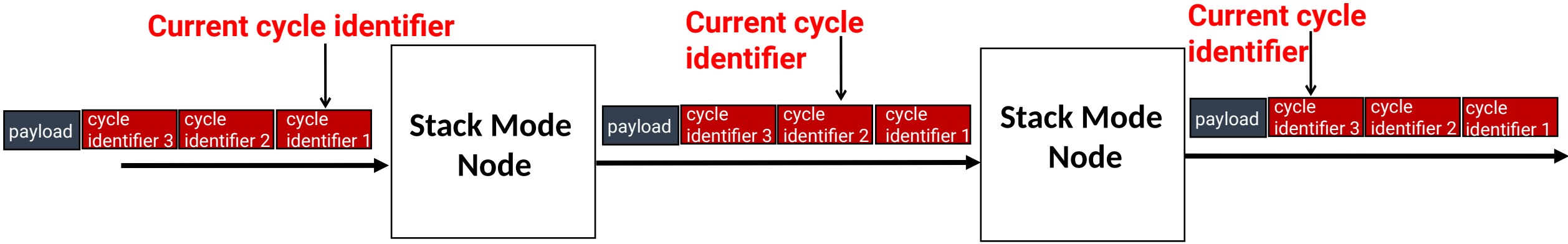
- DSCP of IPv4 Header
- Traffic Class of IPv6 Header
- TC of MPLS Header (used to be EXP)
- IPv6 Extension Header
- UDP Option — Newly added
- SID of SRv6
- Flag fields of SRH
- 2 bits in IPv6 DA
- TLV of SRv6
- TC of SR-MPLS Header (used to be EXP)
- 3 (or 4) labels/adjacency SIDs for SR-MPLS

- Add description for two modes (i.e., swap mode and stack mode) th at implement LDN cyclic forwarding method

Swap Mode: In-packet: Cycle-identifier (applicable to any forwarding plane)
Each node pre-provisioned with cycle-mapping table (e.g.: from PCE-CC)



Stack Mode: In-packet: stack of cycle-identifiers , one for each hop (for SR-MPLS, SRv6)
Each hop maps based on next cycle-identifier



In Summary

- Want to make sure requirements and solution are well un derstood through draft
  - And verify that DetNet WG agrees on requirements being valu able
- Want to ask for working group adoption before next IETF
  - Will do another version resulting from feedback
- Disclaimer:
  - This is NOT the only QoS model useful for DetNet, but is impor tant: least complex per-hop QoS (we think).

Thanks