Timeslot Queueing and Forwarding (TQF)
draft-peng-detnet-packet-timeslot-mechanism-04

Shaofu Peng  ZTE
Peng Liu  China Mobile
Kashinath. Basu  Oxford Brookes University
Aihua Liu  ZTE
Dong Yang  Beijing Jiaotong University
Guoyu Peng  Beijing University of Posts and Telecommunications

IETF-118  Nov 2023
Updates

- Supplement PIFO based timeslot scheduling.
- Simplify the timeslot mapping relationship to reduce maintenance and advertisement costs.
- Clarify the forwarding states related with ideal position maintained on the network entry, and the relevant processing based on actual position.
- Fix the latency equation per-hop and E2E.
- Supplement in-time scheduling mode.
- Update the evaluation table.
- Supplement security considerations.
Motivations

- TSN TAS introduces a synchronous TDM method based on gate control list (GCL) rotation cycle (i.e., gating cycle) in Ethernet LAN.
  - The basic idea is to calculate when the packets of the Scheduled Traffic arrive at a certain node, then the node will turn on the green light for it.
  - A gating cycle contains multiple varialbe length timeslots each for specific flow(s). Here, the timeslot is different with the time frame such as SONET/SDH implemented in L1.
  - However, TAS requires time synchronization and has scalability issues on GCL calculation, update and installation.

- To meet the large scaling requirements, this document enhance TAS to introduce timeslot resources to layer-3 and related timeslot scheduling on DetNet Data Plane.
  - It belongs to TSN mechanism category “PERIODIC SCHEDULING”.
Overview of TQF Mechanism

- Decouple the Orchestration Period from the Scheduling Period.
  - **Orchestration Period** includes timeslot resources, equals to LCM of service burst interval;
  - **Scheduling Period** matches the actual capacity of the device, requiring only a few RR queues or a single sorted-queue.
    - All RR queues have equivalent binary semantic Open/Closed, no other semantics.
- Path calculation based on timeslot resource reservation. Obtain a flexible mapping relationship between the incoming and outgoing timeslots on each node.
Basic Timeslot Mapping (BTM) Between Adjacent Nodes

- For an unidirectional link, we may detect a single timeslot mapping relationship: “the outgoing timeslot of local node’s outgoing port” -> “the ongoing sending timeslot of downstream node’s incoming port”.
- Or, may detect the orchestration period offset relationship between adjacent nodes.

Based on the above detection, we can deduce trivial BTM: “any outgoing timeslot $X$ of local node’s outgoing port” -> “the ongoing sending timeslot $Y$ of downstream node’s outgoing port”, for the purpose of resource reservation later.
States Related with Ideal Position on the Network Entry

- On the network entry, for each regulated sub-burst of specific flow, the reserved outgoing timeslot of NNI port is based on the ideal incoming timeslot of UNI port.
  - e.g., states for the flow:
    <OPL, incoming slot i_1, outgoing slot z_1>;  
    ...
    <OPL, incoming slot i_n, outgoing slot z_n>
  - The network entry should take the closest ideal incoming slot based on the actual position of the regulated packet.

![Diagram of network entry with actual and ideal position](image-url)
PIFO Based Scheduling

• On-time scheduling mode for low jitter

In the case of on-time mode, packets should only be permitted to send at the beginning of the reserved outgoing timeslot Z. So:

packet’s rank = Z.begin

where, Z is the one closest to the arrival time.

Note that there are multiple Z in multiple OP.

![Diagram showing packet scheduling in PIFO based scheduling](image-url)
• In-time scheduling mode for low delay

In the case of in-time mode, packets may be sent before the beginning of the reserved outgoing timeslot $Z$. So:

$$\text{packet's rank} = Z.\text{begin}$$

where, $Z$ is the one closest to the “arrival time plus $E$”. $E$ should be carried in the packets.
E2E Latency Equation

- Best E2E latency (linear with hops)
  \[ \sum_{1 \leq i \leq n} (F_i + T_i + o_i \times TL_i) - TL_n + F_e \]
  It occurs when the packet arrived at the end of the ideal incoming timeslot of UNI port and sent at the head of outgoing timeslot of all NNI port.

- Worst E2E latency (linear with hops)
  \[ \sum_{1 \leq i \leq n} (F_i + T_i + o_i \times TL_i) + TL_uni + F_e \]
  It occurs when the packet arrived at the head of the ideal incoming timeslot of UNI port and sent at the end of outgoing timeslot of all NNI port.

- Jitter is constant TL_uni + TL_n.
Multiple Orchestration Period Instances

• Multiple orchestration periods may be provided by the network. A TQF enabled link can be configured with multiple TQF scheduling instances each corresponding to specific orchestration period length.
• Interworking between different nodes is based on the same orchestration period instance.

\[
\sum_{1 \leq i \leq n} BW_{U_i} \leq \text{link-a.bandwidth} \quad \sum_{1 \leq i \leq n} BW_{V_i} \leq \text{link-b.bandwidth}
\]
### Evaluation

<table>
<thead>
<tr>
<th>Requirement items</th>
<th>Evaluation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Tolerate Time Asynchrony</td>
<td>Yes</td>
<td>No full time synchronization needed, but need frequency sync(3.1.3).</td>
</tr>
<tr>
<td>3.2. Support Large Single-hop Propagation Latency</td>
<td>Yes</td>
<td>The timeslot mapping covers any value of link propagation delay.</td>
</tr>
<tr>
<td>3.3. Accommodate the Higher Link Speed</td>
<td>Partial</td>
<td>The higher service rate, the more buffer needed for the same timeslot length.</td>
</tr>
<tr>
<td>3.4. Be Scalable to The Large Number of Flows and Tolerate High Utilization</td>
<td>Yes</td>
<td>Multiple OPL instance, each for a set of service flows, without overprovision. Utilization may reach 100% link bandwidth. The unused bandwidth of the timeslot can be used by best-effort flows. Calculating paths is NP-hard.</td>
</tr>
<tr>
<td>3.5. Tolerate Failures of Links or Nodes and Topology Changes</td>
<td>N/A</td>
<td>No relationship with queueing mechanism...</td>
</tr>
<tr>
<td>3.6. Prevent Flow Fluctuation</td>
<td>Yes</td>
<td>Flows are permitted based on timeslot reservation, isolated from each other through timeslots.</td>
</tr>
<tr>
<td>3.7. Be scalable to a Large Number of Hops with Complex Topology</td>
<td>Yes</td>
<td>E2E latency is linear with hops, from ultra-low to low latency by multiple OPL. E2E jitter is low by on-time mode. Calculating paths may be NP-hard.</td>
</tr>
<tr>
<td>3.8 Support Multi-Mechanisms in Single Domain and Multi-Domains</td>
<td>N/A</td>
<td>No relationship with queueing mechanism...</td>
</tr>
</tbody>
</table>
Next step

• The content is basically mature and detailed for implementation, and we would like to request WG adoption.
• Any questions/comments?

Thank you!