Deadline Evaluation

Requirement items	Evaluation	Notes
3.1. Tolerate Time Asynchrony	Yes	No full time synchronization needed, only need frequency sync(3.1.3).
3.2. Support Large Single-hop Propagation Latency	Yes	The eligibility arrival of flows is independent with the link propagation delay.
3.3. Accommodate the Higher Link Speed	Partial	The higher service rate, the more burst resource may provided by each delay level, and more buffer space is needed. And, extra instructions ....
3.4. Be Scalable to The Large Number of Flows and Tolerate High Utilization	Partial	Multiple delay levels, each with limited delay resources, can support lots of flows. The unused bandwidth of the high delay level can be used by the low levels or BE flows.
3.5. Tolerate Failures of Links or Nodes and Topology Changes	N/A	No relationship with queueing mechanism...
3.6. Prevent Flow Fluctuation	Yes	Flows are permitted based on the resources reservation of delay levels, and isolated from each other.
3.7. Be scalable to a Large Number of Hops with Complex Topology	Partial	More buffer may be needed when the latency compensation + in-time is used.
3.8 Support Multi-Mechanisms in Single Domain and Multi-Domains	N/A	No relationship with queueing mechanism...

Some Further Explanations of Deadline

3.3. Accommodate the Higher Link Speed: (partial)

- **Buffer cost**: for a given service data amount requirement, the **minimum delay level provided can be correspondingly smaller**; Otherwise, **burst resource** of delay level di (if its duration remained independent of the link speed) is proportionate to the link speed, thus the higher link speed, the more burst resource provided, and more buffer size needed. In the case of limited total buffer of hardware (if this is the case), only a smaller number of round robin queues can be provided, which again means that the maximum delay level is limited. In brief, it is related to hardware capability.
- **Processing cost**: **instructions** of NP microcode programming include: 1) calculate Q = D + E - F per received packet; 2) rewrite E = E + (D-R), where R = tx_timestamp - rx_timpstamp, to the departure packet.

3.4. Be Scalable to The Large Number of Flows and Tolerate High Utilization: (partial)

- **State cost: No state per flow** is maintained on the nodes. Instead, the state, i.e., D & E, is carried in the packet.
- **Overprovision**: The burst resource and bandwidth **resource pool of each delay level is limited** and derived from schedulability condition, and cannot support flow scales exceeding the pool size.

  e.g, let link speed be C, delay levels be d1=10us, d2=20us, d3=30us, ..., d10 = 100us.
  the limited resources of each delay level may be:
  d1: b1 = C*10us, r1 = 0.1*C,   d2: b2 = C*8us, r2 = 0.1*C,   d3: b3 = C*8us, r3 = 0.1*C, ...
  No overprovision, that is , it does not take the calculation result of b_i/d_i to get an overprovision bandwidth.
- **More Levels**: We can flexibly define new delay levels (e.g, d11 = 110 us, d12 = 120 us, ...) on demand to achieve high utilization. That means, although each delay level has only limited burst resource, more bursts resource can be provided by defining more delay levels (however with larger delay value) to admit more services. Note that the minimum delay level is still the most expensive.
- **Path calculation**: it may firstly calculate some deadline path with reserved resource larger than service required, and ready to admit more service, i.e., the calculation may not be per flow.

Some Further Explanations of Deadline (cont...)

3.6. Prevent Flow Fluctuation: (yes)

- All delay levels have their own resource pool, isolated from each other.
- Resource reservation is executed at the granularity of the path, and it is assumed that all flows of the same delay level arrive simultaneously (which is the worst case) on the merged node, without exceeding the deadline. That is, the amout of eligibility arrivals (after latency compensation) is always meet the schedulability condition in despite of flow fluctuation.

3.7. Be scalable to a Large Number of Hops with Complex Topology: (partial)

- **Path calculation**: The calculation is simple, and only independently check if the resource (burst, bandwidth)@level of each link is enough.
- **Burst accumulation:** A flow may traverse a path with a large number of hops. In this case:
  in-time mode may produce more accumulated burst at the remote hop, and need more buffer. For example, d_i is 10 us, burst_interval is 250 us, there may be one b_10 burst in the queue within the 25th hop, and two b_10 bursts simutaneously in the queue during 25~50 hops, etc.
  On-time mode has not this issue of accumulated burst.
- **E2E latency**: The E2E latency formula: N * D is a perfect linear relationship, where N is hops, D is the selected delay level. Different delay levels provide different slopes of E2E latency, and can flexibly meet the latency requirements of service.
- **E2E jitter**: Latency compensation can be used to reduce jitter for each delay level to the value of delay level in theory (i.e., the worst case is that on the egress node all delay levels have the full traffic and discharge floodwater at the same time, however, in reality, the service flow of the output port facing the destination customer may only involve one delay level, then the jitter may be only one AT (e.g, 10us)).

TQF Evaluation

Requirement items	Evaluation	Notes
3.1. Tolerate Time Asynchrony	Yes	No full time synchronization needed, but need frequency sync(3.1.3).
3.2. Support Large Single-hop Propagation Latency	Yes	The detection of timeslot mapping covers link propagation delay.
3.3. Accommodate the Higher Link Speed	Partial	The higher service rate, the more buffer needed for the same timeslot length.
3.4. Be Scalable to The Large Number of Flows and Tolerate High Utilization	Partial	Calculating paths for as many flows as possible is an NP-hard problem. The unused bandwidth of the timeslot can be used by best-effot flows.
3.5. Tolerate Failures of Links or Nodes and Topology Changes	N/A	No relationship with queueing mechanism...
3.6. Prevent Flow Fluctuation	Yes	Flows are permitted based on timeslot reservation, isolated from each other through timeslots.
3.7. Be scalable to a Large Number of Hops with Complex Topology	Partial	Calculating TQF paths for all services is NP-hard problem, related to hops count.
3.8 Support Multi-Mechanisms in Single Domain and Multi-Domains	N/A	No relationship with queueing mechanism...

Some Further Explanations of TQF

3.3. Accommodate the Higher Link Speed: (partial)

- **Buffer cost**: For a given service data amount requirement, the **timeslot length of the orchestration period instance can be correspondingly smaller;** Otherwise, **burst resource** of each timeslot (if its duration remained independent of the link speed) is proportionate to the link speed, thus the higher link speed, the more burst resource provided, and more buffer size needed. In the case of limited total buffer of hardware (if this is the case), only a smaller number of round robin queues can be provided, i.e., scheduling period or PIFO is relatively short and may affect the success of path calculation for flow interleaving.
- **Processing cost: instructions** of NP microcode programming include:

  Round Robin: Scheduling Period outgoing timeslot = Orchestration Period outgoing timeslot MOD M.

  PIFO: rank = ideal outgoing timeslot, which is calculated based on arrival-time, deviation E, orchestration period length, and time slot length.

3.4. Be Scalable to The Large Number of Flows and Tolerate High Utilization: (partial)

- **State cost: No state per flow** is maintained on the nodes. Instead, the state, i.e., orchestration period length, timeslot-id (and deviation E in the case of PIFO), is carried in the packet.
- **Overprovision**: **No overprovision**. It does not take the calculation result of burst / timeslot to get an overprovision bandwidth.
- **More Levels**: **Multiple orchestration period length**, each is LCM of ther served services burst interval and naturally serves the flows in an interleaving flow manner (isolated through timeslots) from headend to intermediate, and to endpoint.

  e.g, service-1 with burst interval 100us may be served by OP-instance-100us;

  service-2 with burst interval 1ms may be served by OP-instance-1ms;

  service-3 with burst interval 1s may be served by OP-instance-1s;
- **Path calculation**: The path calculation may be **NP-hard** to find a path meet the expected E2E latency. For example, some nodes may contribute less latency than budget, and some other nodes may contribute more latency than budget, but the total latency of the path meet the total budget. In brief, the resource reservation results of multiple links will affect each other.

Some Further Explanations of TQF (cont...)

3.7. Be scalable to a Large Number of Hops with Complex Topology: (partial)

- **Path calculation**: The path calculation may be **NP-hard** to find a path meet the expected E2E latency.
- **Burst accumulation:** In-time mode (with PIFO) may have burst accumulation; On-time has not.
- **E2E latency**: The E2E latency of the candidate path is based on reserved timeslot resource (thus the per hop latency may be different due to different o), which must be less than or equals to budget value.
  - the E2E latency formula: , where F_i is the forwarding delay of node i, Remaining_i the remaining time of ongoing sending timeslot of node i, TL_i is the timeslot length of node i, o_i is the timeslot offset (between outgoing timeslot and ongoing sending timeslot) of node i.
- **E2E jitter**: For the ideal arrival position of orchestration period of UNI port, the E2E jitter is only 2 timeslot and independent of hops.

  e.g, best case: arrived at i.end and departure at z.begin;

  worst case: arrived at i.begin and departured at z.end;

  thus 2 timeslot jitter.