Problems with existing DetNet bounded latency queuing mechanisms

draft-eckert-detnet-bounded-latency-problems-00

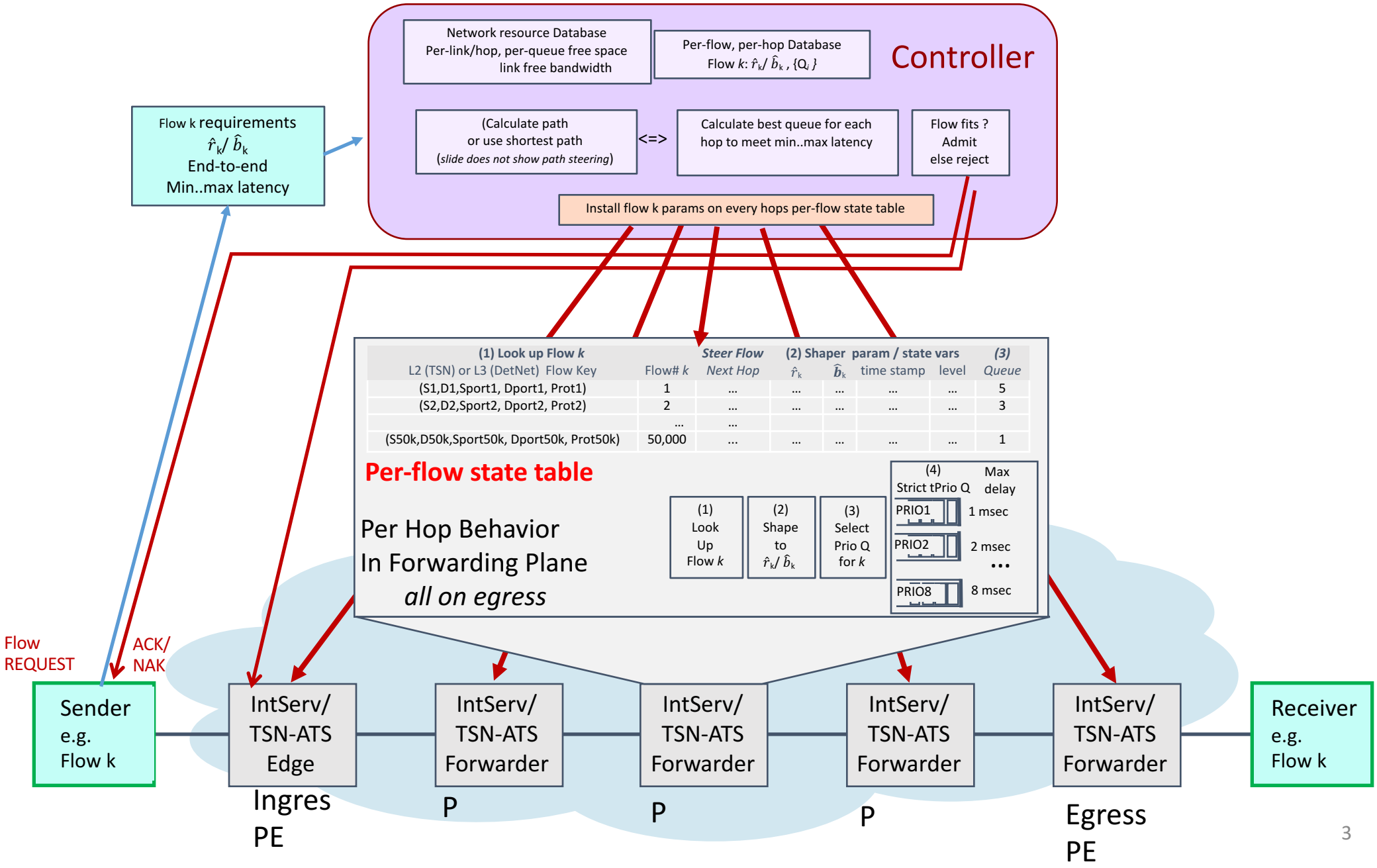
Toerless Eckert, Futurewei USA (tte@cs.fau.de)
Stewart Bryant, Stewart Bryant Ltd (sb@stewartbryant.com)
09/09/2021

Summary

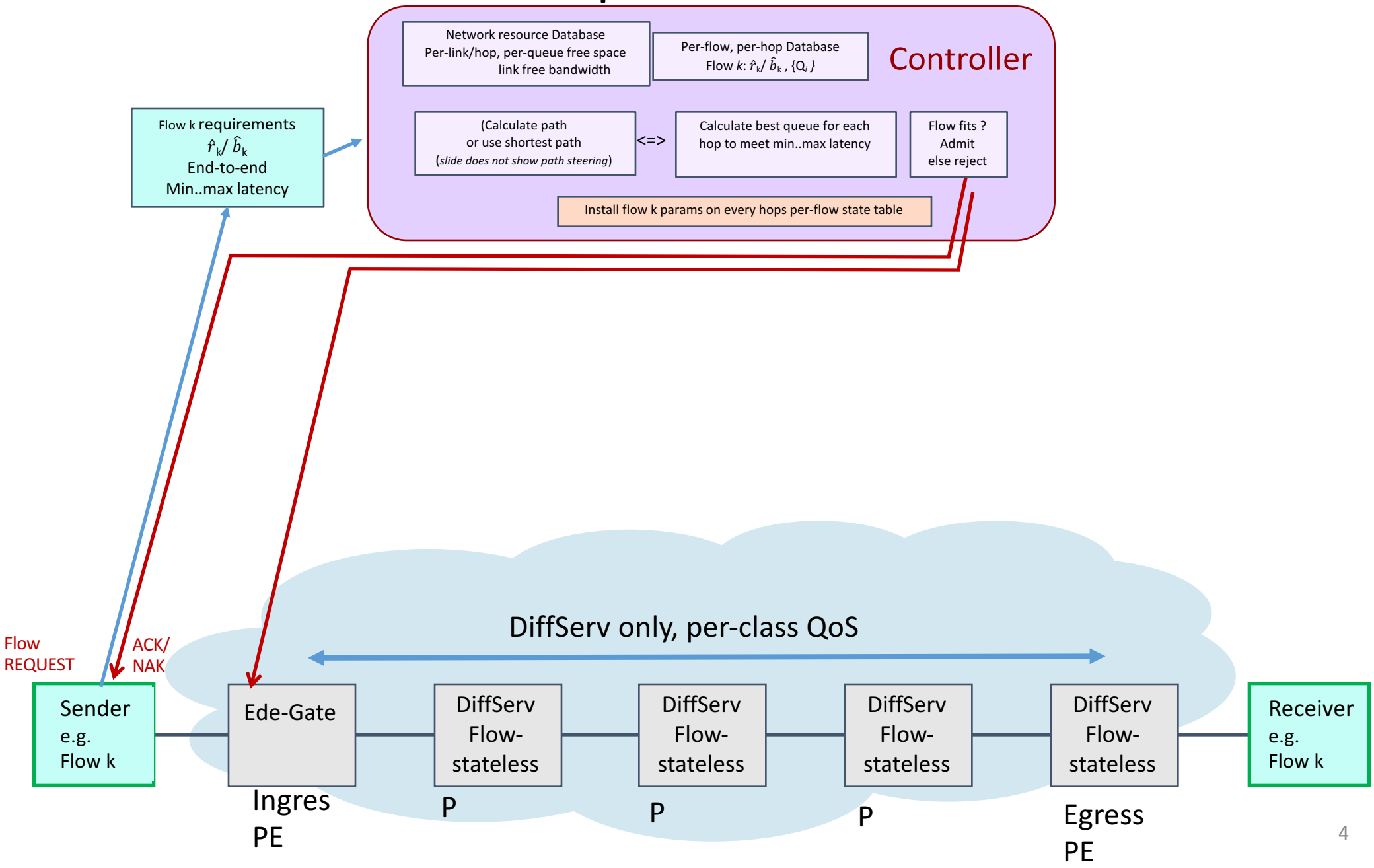
- Various open issues in detnet related to queuing
- High level:
- Network / deployment
  - Can we/how-can-we support popular network designs: SR-MPLS/SRv6 (and futre ones: BIER/BIER-TE)
  - Can we/how-can-we support large-scale/wide-area networks with DetNet (MPLS WAN)
    - Including preference for DiffServ, no controller-plane to P-LSR per-flow signaling
  - How much can-we/should we have lower-cost DetNet options via no or reduced clock synchronization requirements ?
- Application
  - Bounded latency not always enough!
  - Tightly bounded jitter can improve value of service

- *Slides highlighting only subset of issues, draf more comprehensive*

IntServ / TSN-ATS, DetNet model (issues)



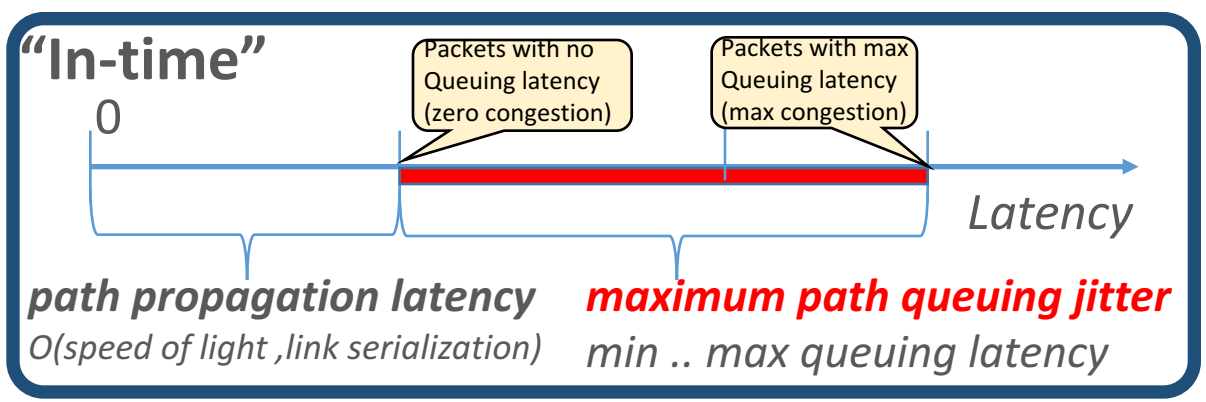
Desirable DetNet QoS option

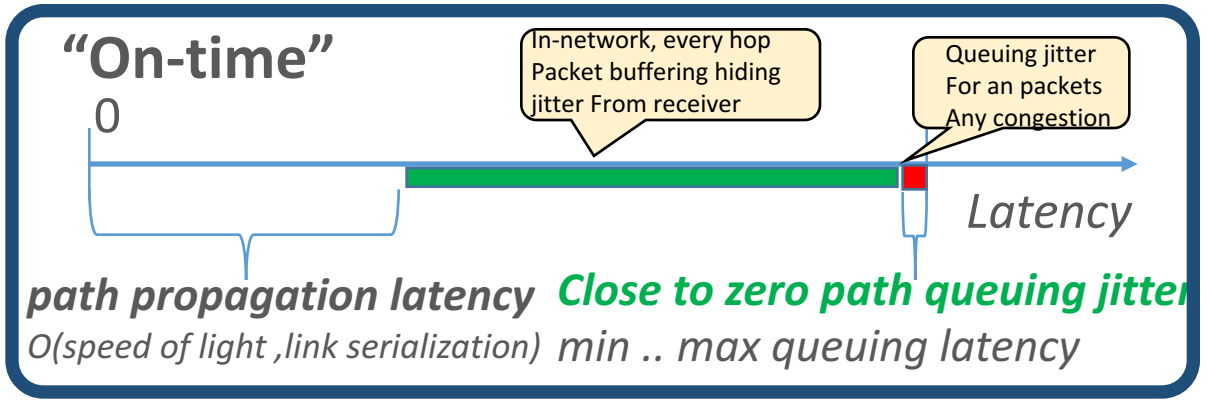


Per-Hop, Per-flow state issues

- Core of IntServ (RFC2212), quickly amended by DiffServ. No DiffServ deployed significantly in networks larger than campus.
  - Reduced IntServ, RSVP-Traffic-steering, NOT-per-hop-queuing was used until better technology was available (Segment Routing).
  - This is more expensive the faster the network is.
- Several issues with per-hop, per-flow state
  - QoS hardware cost limitation: Shaper – IntServ -> Interleaved Regulators (UPS, TSN-ATS), still too expensive for large-scale, high-speed forwarders with many interfaces.
  - Churn through signaling updates. Per-Hop, Per-Flow state upates upon chage: new/dead flows, path changes. Update to hardware.
  - If state is driven by application, state on P nodes even more problematic (unplannable). Biggest experiences from IP multicast and evolution of IETF standards for that.
- Current standard or proposed standard for large-scale network models: no per-hop, per-flow statee: Segment Routing (source routing), BIER(-TE) or multicast , simple DiffServ QoS
  - Need DetNet QoS option supporting SR, BIER…

Tightly bounded jitter – “In-time” vs. “On-time” (1)





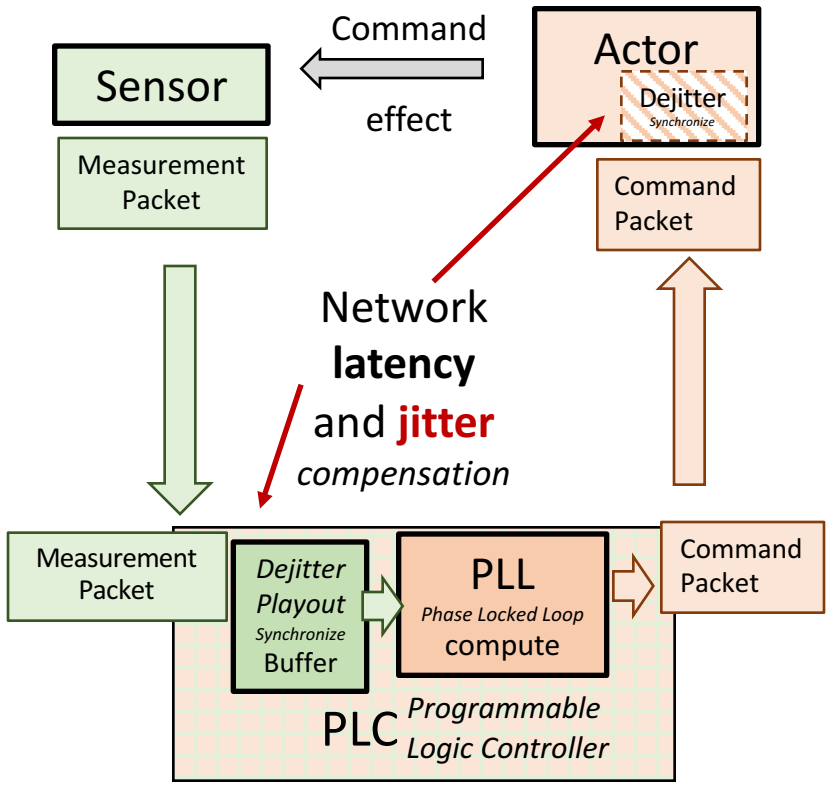
- “In-time” delivers packets as soon as possible.
  - Under no congestion, no in-network queuing latency incurred
  - IETF IntServ/Guaranteed Service, TSN-ATS use this model
  - Queuing latency only because of
    1. Temporary congestion / burst accumulation of packets from multiple interfaces arriving into same output queue
    2. Per-flow interleaved regulator / shaper – to remove bursts introduced by 1.
- “In-time” creates worst-case jitter between no..max traffic load
- On-time eliminates this jitter
  - By networking buffering independ of traffic load

Tightly bounded jitter – “In-time” vs. “On-time” (2)

- Why/when in-time vs. on-time ?
- Non-application reasons:
  - Current on-time (e.g.: CQF) requires clock synchronization. In-time can work without any clock synchronization. If applications do not benefit from on-time (tight jitter), then in-time can be a simpler solution
  - Curent in-time solutions (e.g.: TSN-AT IETF-GS) require per-hop, per-flow controller-plane instantiatd state – this can be a challenge (see prior slides). If applications do not benefit from in-time (lower latency under lower network load), then on-time can be the better solution.
- The application view:
  - All deterministic application MUST be prepared for any packet to arrive as late as (guaranteed) bounded latency
  - Only SOME deterministic applications can be built to operate opportunistically better with in-time
    - Network load low ➔ lower latency/RTT ➔ deterministic application may operate faster / more-accurate / “somewhat better”
    - But there is no GUARANTEE for any of this – it is OPPORTUNISTIC! Network load can always change unexpectedly, burst collisions can happen stochastically !
  - In-time often also shifts work from network into application/device !

Tightly bounded jitter – “In-time” vs. “On-time” (3)

- Application traffic profiles, e.g.: Industrial Internet Consortium (IIC) _Time Sensitive Networks for Flexible Manufacturing Testbed https://www.iiconsortium.org/white-papers.htm
  - Isochronous, Cyclic, Audio-Video, (on-time), Alarm and Events (in-time), …
- Media playout and most control loops want on-time
  - Media: Synchronous playout
  - If network is in-time, packets must be buffered in application and consumed at maximum guaranteed latency time
- Playout buffer size requirement depend on network jitter == network size
  - Expensive war stories in industry when equipment was used in networks with higher jitter
- In-time also raises clock synchronization needs on devices
  - On-time delivered packets carry implicit timing information !
  - Dumb devices (actors) may not be able to support dejittering and/or accurate clock
  - Classical example: accurate PLC with periodic “polling” of dumb actors/sensor without accurate clock: on-time allows for isochronous operation.
    - In-time would require much more complex sensor/actor/control-loops



Comparison of TSN/DetNet options	**IntServ/GS** RFC2212	**TSN-ATS** Latest (2020) TSN standard, also prime target for Detnet	**TSN-CQF (Qcr)** Original/simplified TSN option (over Qbv)	**Packet taged CQF** draft-dang-queuing-with-multiple-cyclic-buffers	**Packet tagged per-hop deadlines** draft-stein-srtsn
DiffServ / SR-MPLS,v6 / BIER design goal compatible	**NO**	**NO**	**YES**	**YES**	**YES**
Per-hop-per-flow state *Hardware-cost/scale* *Signalling-complexit/churn*	**YES**	**YES** Interleaved regulators (simplified over GS)	**NO**	**NO**	**NO**
Clock synchronization required *Additional PTP hardware and network operational requirements*	**NO**	**NO**	**YES** High accuracy (nsec)	**YES** Low accuracy (usec)	**TBD ?** *Not considered to be a deployment cost by author ?!*
Tighly bounded jitter	**No**	**No**	**Yes** Usec jitter (cycle size)	**Yes** Usec jiter (cycle size)	**No ?**
Target deployment scale	**1990th "Internet"**	**Building/ Campus?**	**Building/ Campus?**	**Campus/Metro/ Country (tested)**	**Metro ?**
Arbitrary physical distance network / network links	**YES**	**YES**	**NO** Throughput deteriorating to 0 at ca. 2 km	**YES**	**Yes ?**
Latency calculus (for PCE)	Complex, deterministic	Simple, deterministic	Trivial, deterministic	*Trivial, deterministic*	*TBD – ongoing work to be published*

Q & A