Strict Priority RA Class

LNI 4.0 contribution to IEEE P802.1Qdd
for future RAP Strict Priority RA Class

November 2021

Deterministic QoS (dQoS)

Deterministic QoS provides bounds on latency, jitter, packet (congestion) loss.

In this presentation, we present the LNI's motivation, view and expectations regarding the upcoming "Strict Priority" RAP RA Class Template. Our assumptions are based on "Bride local guaranteed latency with Strict Priority Scheduling" (see dd-grigorjew-strict-priority-latency-0320-v02.pdf (ieee802.org)).

According to the LNI working model, we will implement, test and provide feedback to the IEEE 802.1 TSN WG.

Focus of LNI 4.0 TSN Testbed

Currently, most approaches to provide dQoS are based on the following methods:

- Centralized Schedule
- **Resource Reservation**
- Traffic Shaping
- **End Station Models**

LNI testbed focuses on the resource reservation method.
As a first step, Traffic Shaping and End Station Model based on TAS was evaluated.

Recap: Resource Reservation

Basic Principle: Dedicated Ressource Allocation for Streams

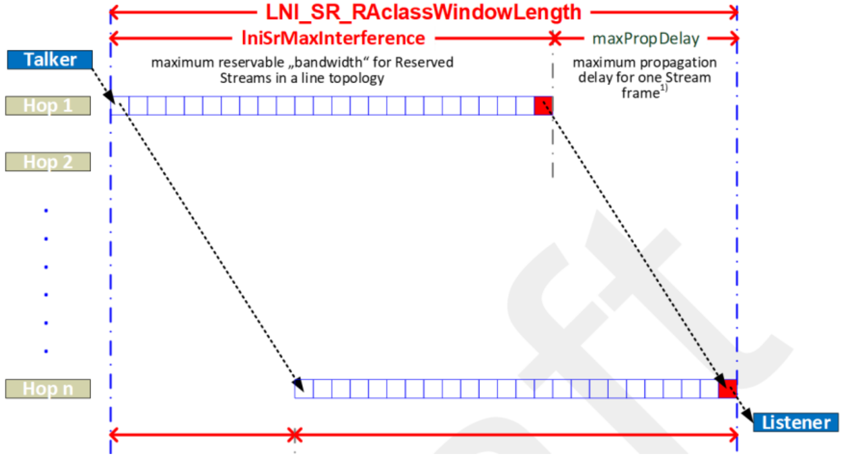
- All talkers announce their streams to resource reservation entities assigned to bridges and potential listeners. The active network topology is given and used to propagate the information hop-by-hop.
- Listeners can join a stream by issuing an attach request to the reservation entities assigned to bridges and talkers.
  The attach request is propagated hop-by-hop following the reverse path back to the talker.
- On each hop, the reservation entity allocates dedicated resources (e. g. queuing resources), bandwidth and latency per stream to guarantee zero congestion loss and bounded latency. In contrast to best effort, resource reservation completely excludes overbooking.
- The distributed user network interface (dUNI) provides application agnostic access to network resources.
- An application independent network service interface is provided, allowing clear separation of application from network.

Resource reservation can achieve dQoS without synchronization and scheduling

Recap: LNI “TAS” Approach

Features: Resource Reservation + TAS + Synchronization

https://www.ieee802.org/1/files/public/docs2020/liaison-LNI40-LRP-RAP-WP-0520-v01.pdf



Technologies selected:

- dQoS by Distributed Resource Reservation:
  Resource reservation entities hosted on end stations and bridges (IEEE 802.1Qdd RAP draft)
- Traffic shaping by synchronized time based shaping:
  Time-Aware Shaping (TAS, IEEE 802.1Qbv) based on network cycle (gating cycle).
  Protected windows for streams open and close at same time.
- Two end station models are supported:
  - Synchronous talkers using transmission time model,
    all frames ready for transmission at network cycle start and sent „bursty“.
  - Periodical free running talkers using data rate model.
- Network topology is constrained to line/ring meeting industrial network requirements.
- Homogeneous link speed between bridges is required.

Evaluation of the “TAS” Approach

What are the Pros?

Pros:

- The distributed resource reservation model works without the need of a CNC component.
- The distributed user network interface (dUNI) provides application agnostic access to network resources without CUCs.
- No additional efforts for schedule calculation.
- “Bus-like” deployment:
  The latency bound is independent from the connection point of an end station in the line/ring.

Evaluation of the “TAS” Approach

What are the Cons?

- High accurate synchronization requirements on bridges and end stations have to be met to enter operational state, currently further complicated by diverse sync profiles in place.
- Data update cycles of all streams must be aligned to the network cycle, causing overprovisioning in many scenarios.
- The amount of reservable streams depends on the available queuing resources, and depends on the window size.
- Increasing the network cycle reduces the reservation efficiency of given queuing resources.
- „Inefficiency penalty“ for free running talkers.
- The window size defines the maximum accumulated latency for all streams of the stream class, no maximum latency variation possible.

Obstacles to start the TSN journey are currently high

SP Model Technology Selection

LNI 4.0 TSN Testbed is going to evaluate the “Bride local guaranteed latency with Strict Priority Scheduling” selecting the following technologies:

- dQoS by distributed Resource Reservation.
- End station Traffic shaping for streams only
- End Stations apply the dQoS data rate model where the maximum data rate must not be exceeded. (see dd-grigorjew-strict-priority-latency-0320-v02.pdf (ieee802.org))
- “Strict Priority” operation in the network - No shaping in the bridges
- All traffic classes above SP class(es) have to be assigned to deterministic traffic

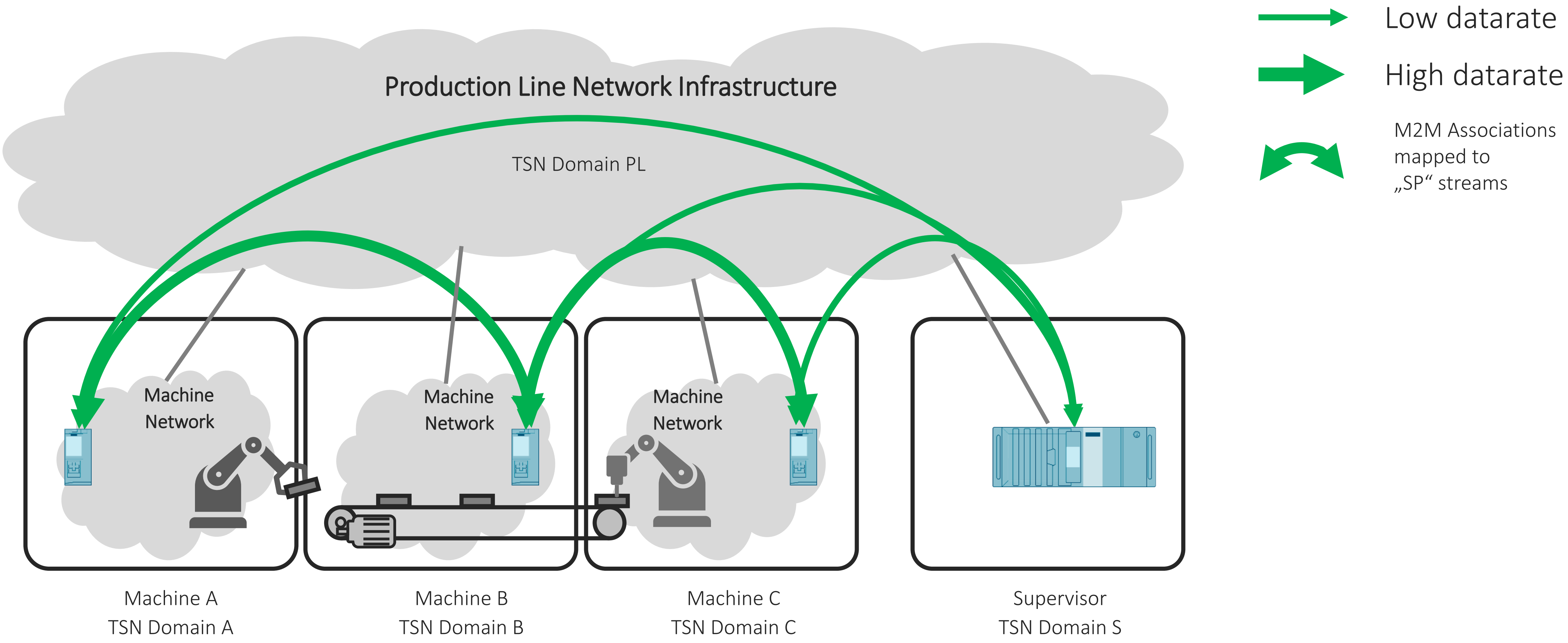
Note: Synchronization is currently not selected

We assume to match the RAP RA Class Template for SP and would appreciate your feedback.

Example

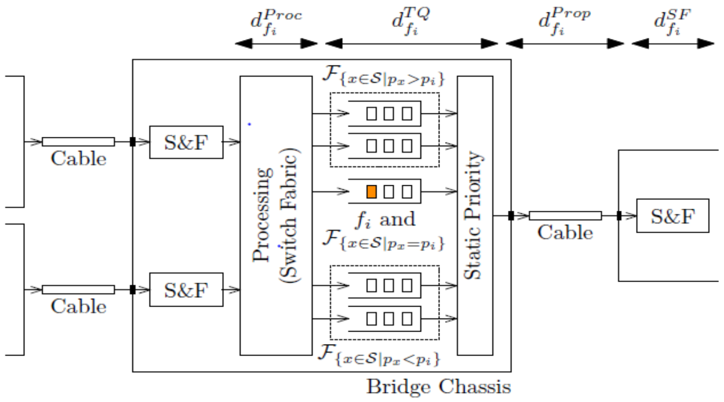
Production Line Use Case

https://www.ieee802.org/1/files/public/docs2021/60802-Weber-LNI40-ApproachProductionLine-0721-v01.pdf



Formula for Provisioning & RAP

https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-198310



Note: Picture from University Whitepaper please see there for details

Knowledge about topology (physical links , paths …) and streams (Talker, Listener, Tspec …) is required.

Provisioning of network is based on following per hop latency formula:

Per Hop Latency Guarantee: $d^{Proc} + d^{Prop} + d^{TQ} + d^{SF}$

$$d_i^{TQ,SF} \leq \sum_{\{x\in\mathcal{S}|p_x>p_i\}} y_{i,x} b_x/r + \sum_{\{x\in\mathcal{S}|p_x=p_i\}} z_x b_x/r + \max_{\{x\in\mathcal{S}|p_x<p_i\}} \hat{\ell}_x/r$$

Simplification: Highest traffic class for streams and max. interference by lower traffic classes (max frame size) by 1GBit/s

$$d_i^{TQ,SF} \leq \frac{0 + Num_{SR} \cdot (mFS_{SR} + IFG + SFD + Prea) + (mFS_{LTC} + IFG + SFD + Prea)}{0.125\ Bytes/ns}$$

NumSR := Number of Streams, mFSSR := Maximum Framesize Streams, IFG:= Inter Frame Gap (Bytes),
SFD := Start Frame Delimiter (Bytes), Prea := Preamble (Bytes), mFSLTC := Maximum Framesize Lower Class Traffic

Results of a Simulation

Assumptions:

- Only one deterministic traffic class based on SP (highest traffic class)
- Per stream measurement interval is larger than latency bound
- All streams have same max. frame size
- All streams share same path with max. interference on each hop
- d^{Prop}=350ns; d^{Proc}=1988ns; mFSLTC=1522 Bytes; Prea=1 Byte; SFD=1 Byte; IFG=12 Bytes

Framesize (Bytes)	Maximum Hops	Maximum Streams per Hop	Latency Bound
100	2	128	< 276 us
500	2	48	< 430 us
1522	2	32	< 1 ms
100	5	128	< 690 us
200	5	96	< 1 ms
500	5	48	< 2 ms
200	20	96	< 4 ms
500	20	48	< 5 ms
200	50	96	< 10 ms

LNI Approaches easy to combine

The LNI “TAS” and LNI “SP” approaches can be combined easily in a single network:

- Both approaches are providing dQoS using the same method:
  Resource reservation IEEE 802.1Qdd Resource Allocation Protocol (RAP).
- The RAP “per hop max latency model” supports both LNI models of traffic shaping:
  - A LNI “TAS” stream class occupies a fixed time interval which is well known in all synchronized bridges supporting this class.
  - A LNI „SP“ stream class considers the class interference with LNI „TAS“ in it‘s per hop latency calculation when traversing a synchronized bridge supporting LNI „TAS“.

Further traffic shaping models for dQoS will be supported by RAP to cover a wide range of applications.

Evaluation of the “SP” Approach (1)

What are the Cons?

- The latency bounds for time-sensitive streams (accumulation of per hop max latency) depend on the paths between talkers and their listeners.
- The amount of reservable streams depends on the available queuing resources.
- Large frames transmitted within large stream measurement intervals cause inefficient queuing resource usage because of dedicated queuing resource allocation.
- To ensure zero congestion loss, additional dedicated queuing resource for multiple intervals need to be allocated in case of the stream measurement interval is smaller than max. accumulated latency*).

*) Note: Traffic shaping in bridges would improve efficiency of queuing resource usage and would decrease interference for lower latency.

Evaluation of the “SP” Approach (2)

LNI4.0 LABS NETWORK INDUSTRIE 4.0

What are the Pros?

Pros as before with LNI “TAS”:

- The distributed resource reservation model works without the need of a (C)NC component and the distributed user network interface (dUNI) provides application agnostic access to network resources without CUCs.
  No additional efforts for schedule calculation.

Pros in addition:

- Local clocks for traffic shaping in end stations is sufficient, no need for synchronization.
- Almost any existing HW for end stations or bridges can be used, SW extensions only.
- Independent data update cycles for streams are supported by max data rate per stream.
- All talker are „free running“, no „Inefficiency penalty“ as in LNI “TAS”.
- All network topologies including star / mesh are supported, the network topology can be designed to e.g. minimize interference.
- Various link speeds are supported.
- Allowing applications to vary their maximum required latency per stream can be used to increase the number of reservable streams.

Obstacles to start the TSN journey are low

Thank You