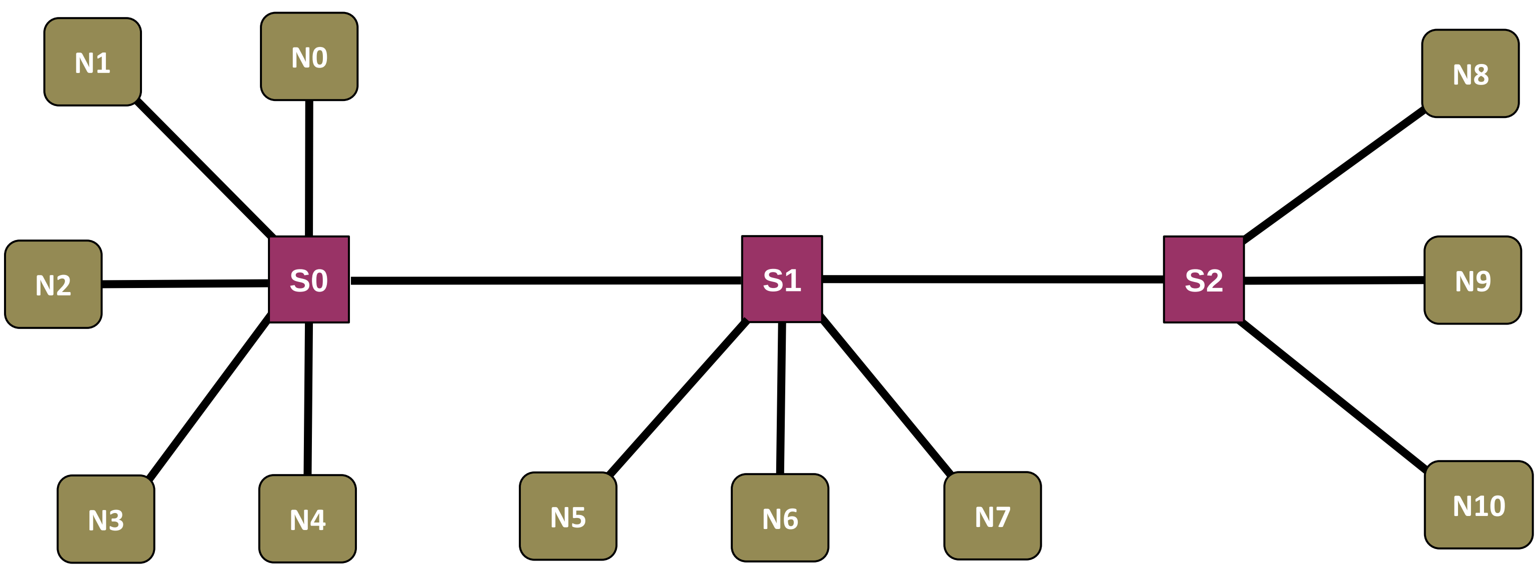
Switched Ethernet in Time Sensitive Networking

Based on Internet sources, work of Jingyue Cao and Siva Thangamuthu

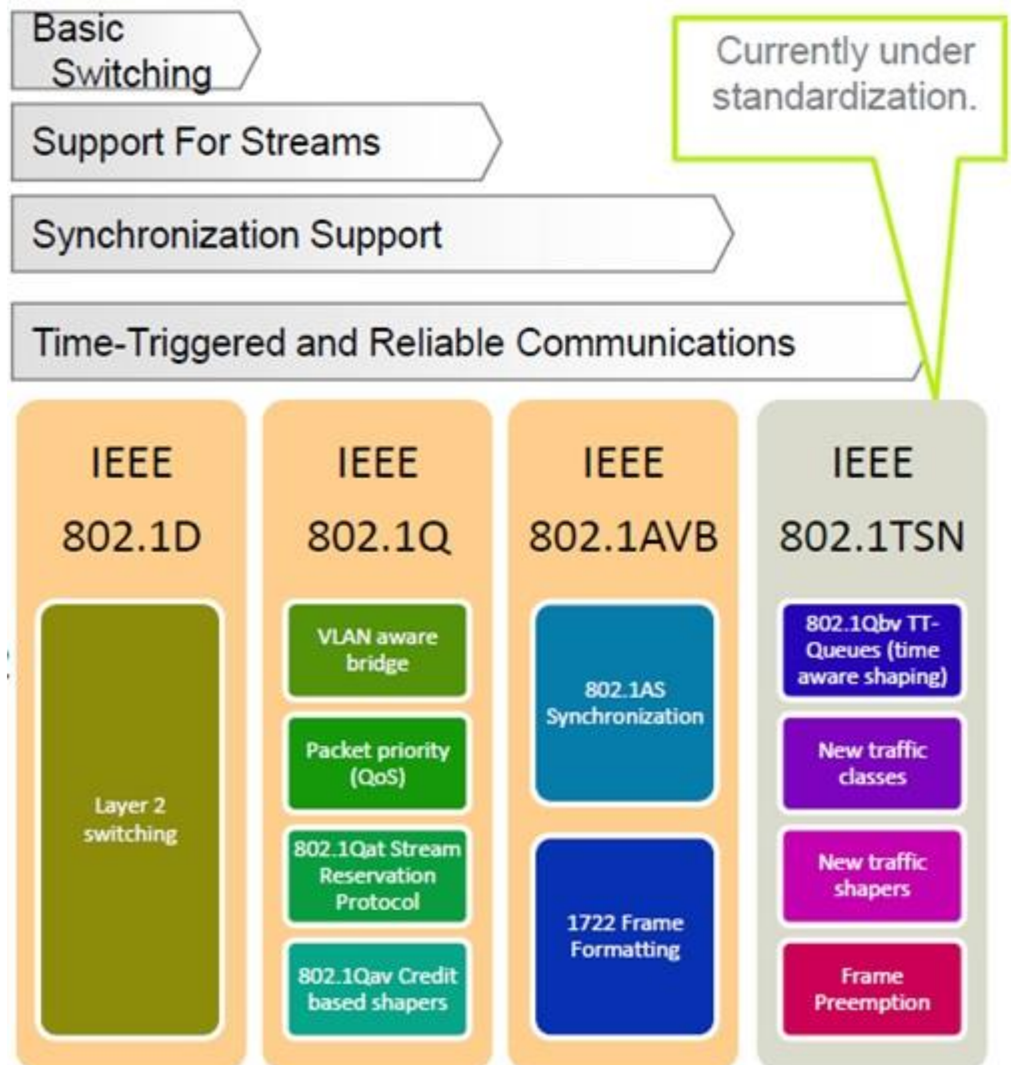
edited by Johan Lukkien

(Switched) Ethernet for real time

- Advantages: High-Bandwidth, Commercialization, Low cost
- Disadvantage: does not provide real-time guarantees
- Solutions:
  - Traffic shaping (plus a lot more) in IEEE 802.1AVB/TSN
  - TTP (time triggered protocol)



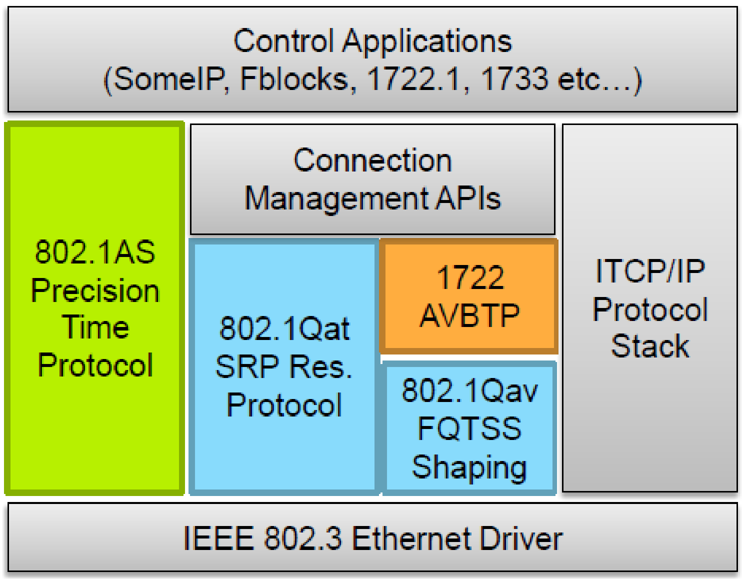
Evolution of Ethernet Switching



From: Sivakumar Thangamuthu's master thesis presentation

IEEE 802.1 AVB Overview

- IEEE grouped a set of 802.1 standards created by the Audio Video Bridging (AVB) Task Group
  - AVB Task Group was first formed in 2006, its target applications were Audio & Video
  - aims to extend the functionalities of standard Ethernet with synchronization and guaranteed latency services
  - Upon the completion of the AVB standards in 2013, it became clear that streaming data can also be Control, so the Task Group changed its name to Time Sensitive Networking (TSN)

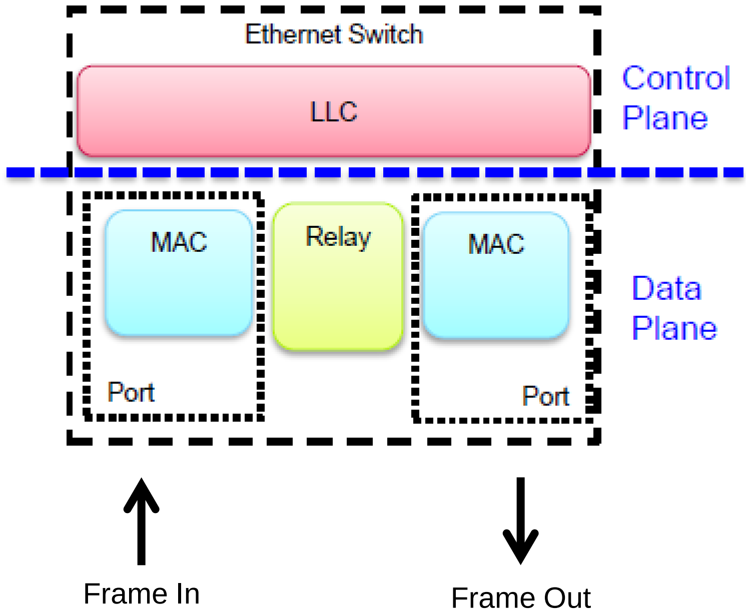


Name	Description
802.1AS-2011	Timing and Synchronization for Time Sensitive Applications (generalized Precision Time Protocol)
802.1Qat-2010	Stream Reservation Protocol
802.1Qav-2009	Forwarding and Queuing Enhancements for Time-Sensitive Streams (Credit based shaping)
802.1BA-2011	AVB Systems
1722-2011	Audio Video Transport Protocol
1722.1-2013	Device Discovery, Enumeration, Connection Management and Control

(mostly) From: Sivakumar Thangamuthu’s master thesis presentation

Ethernet Switch - Components

- Switches are separated in two main parts
  - Control Plane: Usually implemented in SW and provides automatic configuration features and services
  - Data Plane: Usually implemented in HW and is composed of the MAC and Relay Modules
- Traffic Shaping is a functional part of *Queuing and Transmission Selection* which is implemented in the MAC-output module



From: Sivakumar Thangamuthu's master thesis presentation

Essentially,

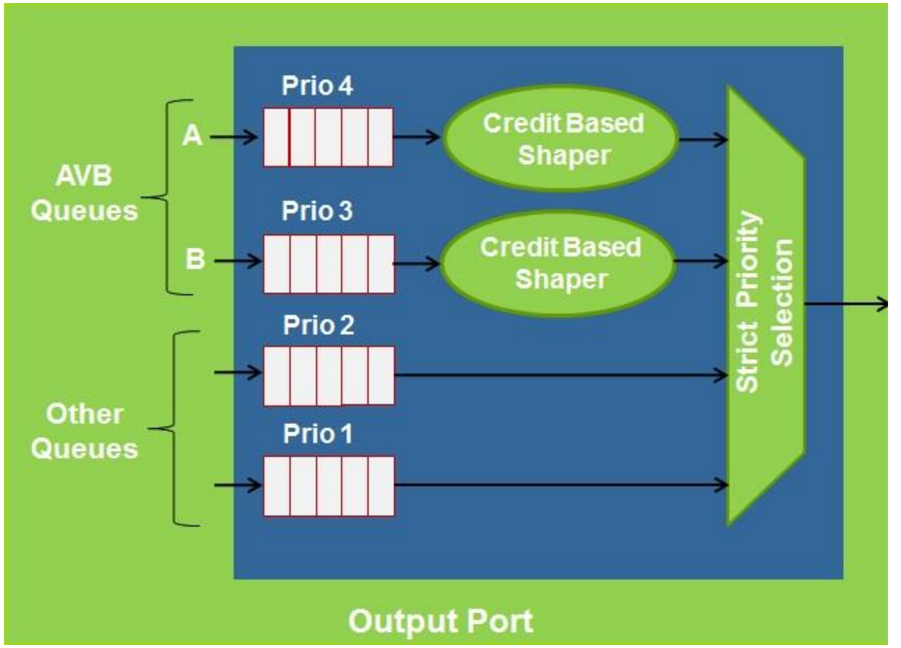
- Flows (streams, frames) are divided into classes
- The classes have real-time requirements and are associated with an output queue of frames
  - locally generated traffic
  - forwarded
- Queues are admitted to the medium using *shaper rules*
  - shapers are functions of time, queue history, other queues, and include further priority rules
- Actual guarantees also depends on *admission control*
  - overloading particular queues will lead to deadline misses
- Discussed and analyzed shapers are
  - Credit Based
  - Time aware
  - Peristaltic (name changed?)
  - Burst Limiting
- Analysis yields response time expressions, hence total delays

AVB Traffic Classes for Audio and Video

Traffic Class	Transmission Period	Expected Latency (7 hops)
Class A	125 us	2 ms
Class B	250 us	50 ms

Application Dependent
Max Frame Sizes

- Best Effort (BE) or Legacy Ethernet Traffic in addition
- Class A stream traffic has the highest priority
- Max 75% of the bandwidth can be allocated for AVB streams
- One Queue per port per Traffic Class
- No offline scheduling, flexibility using online stream reservation protocol (SRP)
- AVB traffic shaped by “Credit Based Shaper“ (CBS) in order to prevent bursts



Traffic Classes in IEEE802.1TSN

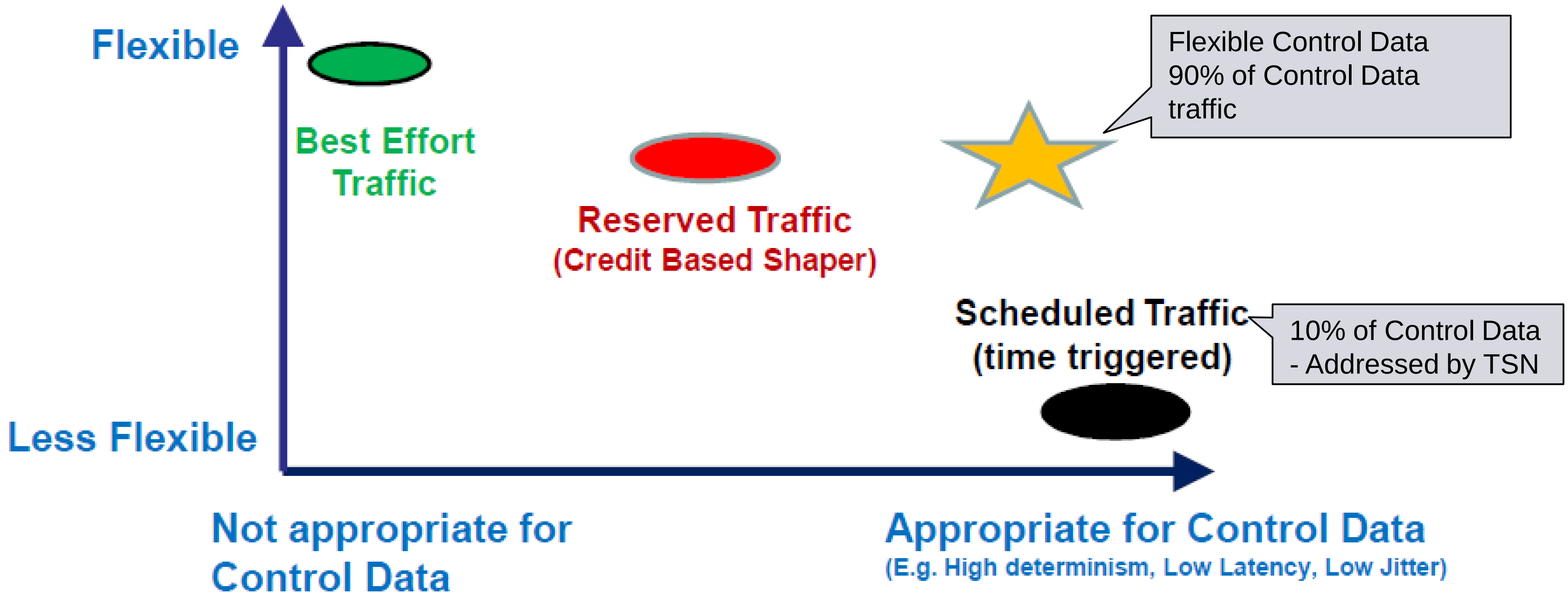
- IEEE802.1TSN (Time Sensitive Networking) is a set of standards to enhance switched Ethernet for time critical industrial and automotive networks

- In-addition to AVB Class, traffic class for time-critical control data

Traffic Class	Max. Frame Size (MTU)	Transmission Period	Expected Latency
Class CDT	128 bytes	500 µs	100 µs (over 5 hops)
Class A	256 bytes	125 us	2 ms (over 7 hops)
Class B	256 bytes	250 us	50 ms (over 7 hops)
Class BE	256 bytes	-	-

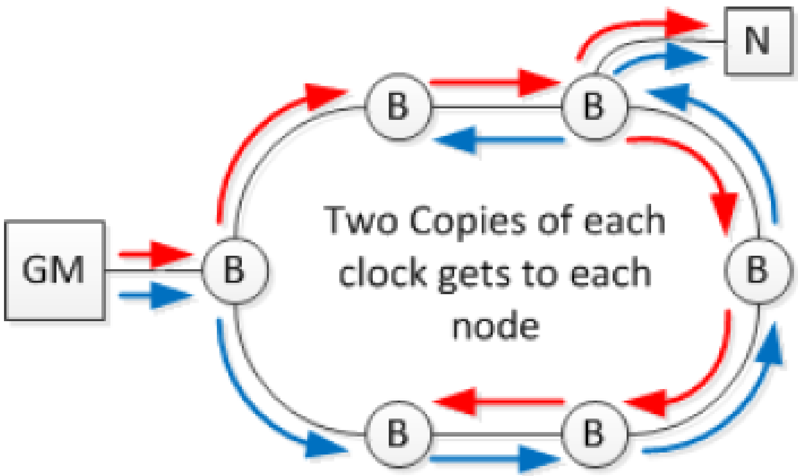
- These requirements are not yet standardized (2014) and might change

Automotive Control Data Traffic Requirement

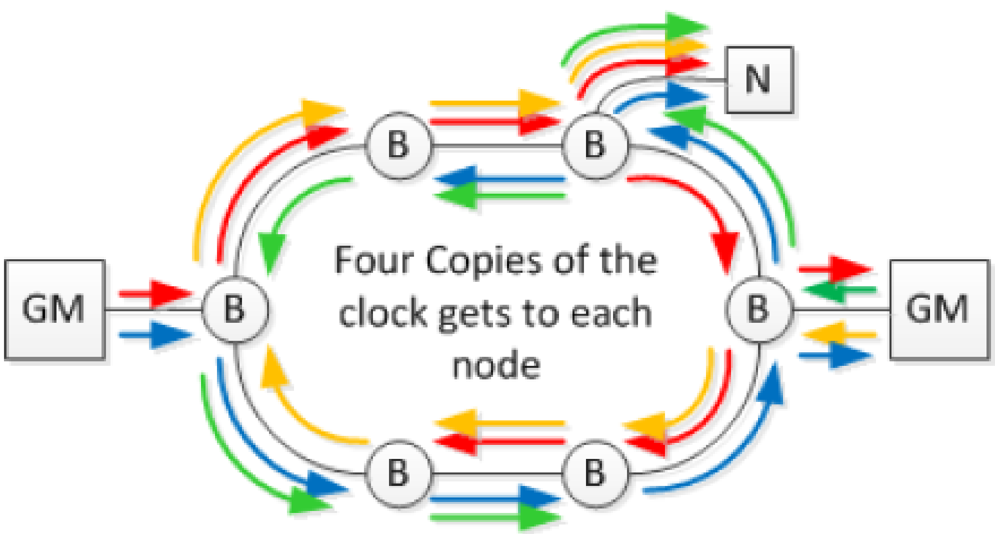


- At least 90% of automotive control applications will not have periods shorter than 5ms.
  - Min: 5ms Max: 1000ms Typical: 8ms
- Shorter periods (e.g. 1ms, 2.5ms) are desirable
- Need for flexibility to configure the required periods for periodic messages

Enhance Generic Precise Timing Protocol



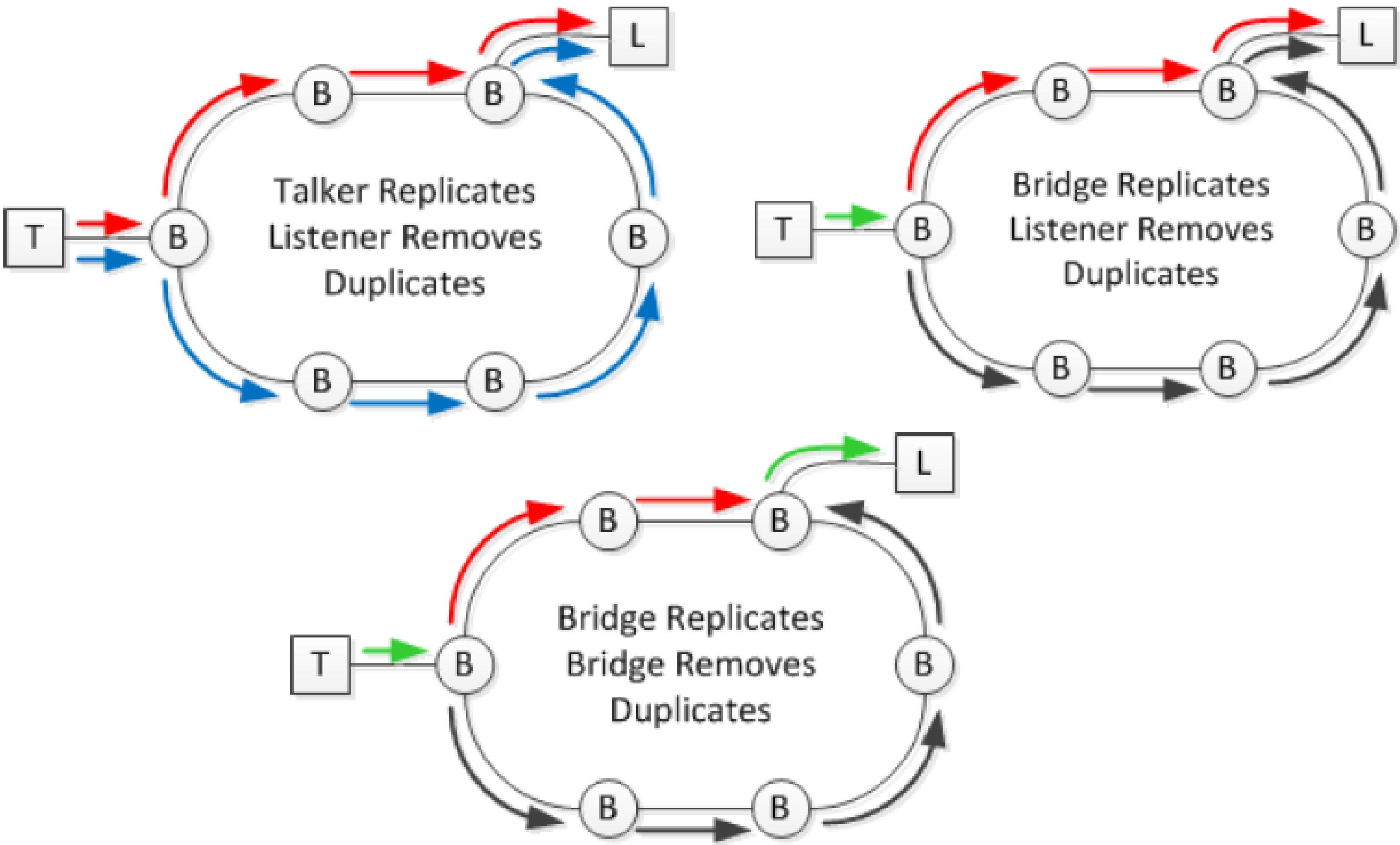
- Single Grand Master transmitting two copies of its clock using separate paths



- Dual Active Grand Masters each transmitting two copies of their clock using separate paths

From: https://standards.ieee.org/events/automotive/2015/03_IEEE_TSN_Standards_Overview_and_Update_v4.pdf

Add Frame Replication & Elimination

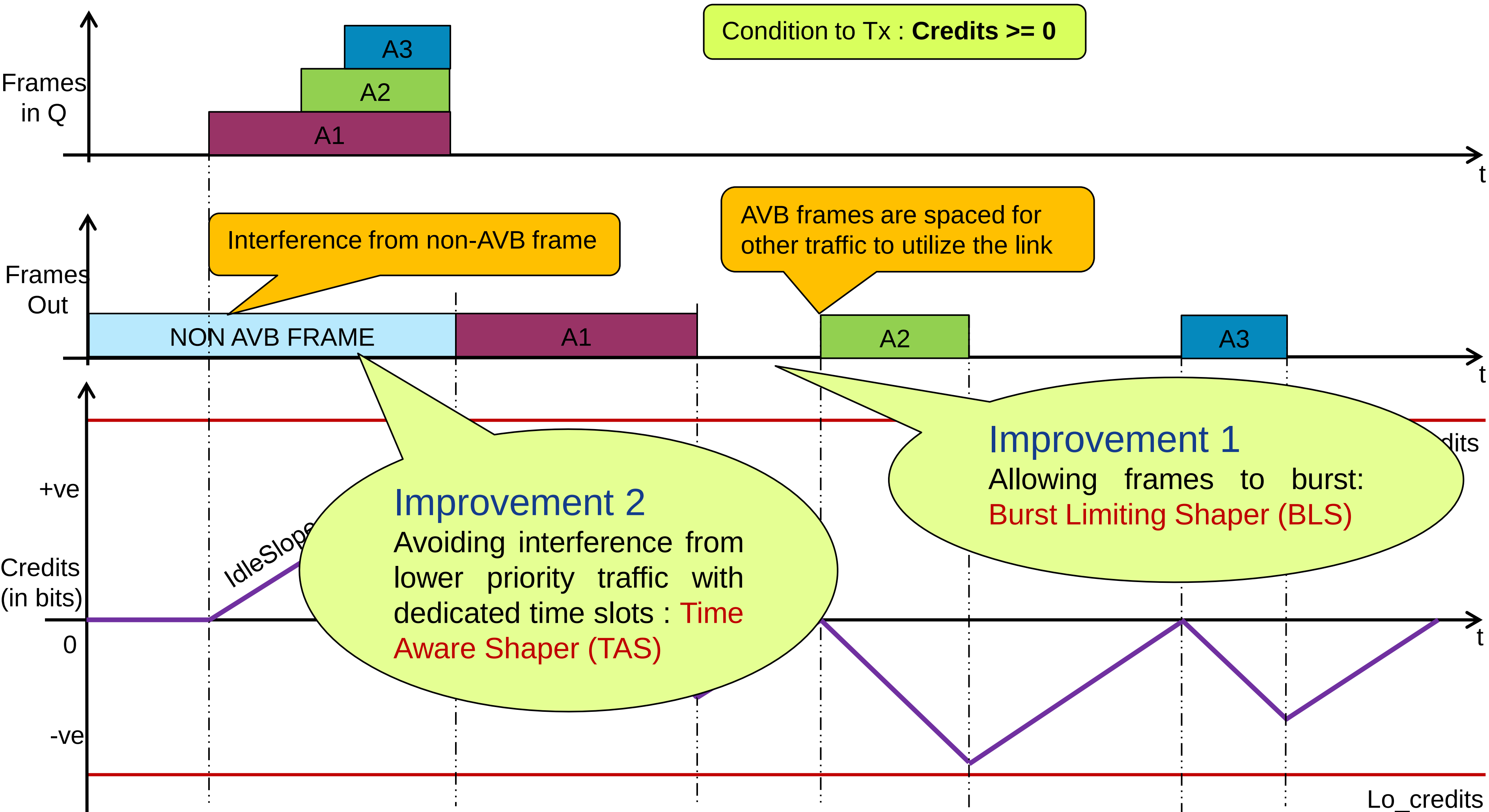


From: https://standards.ieee.org/events/automotive/2015/03_IEEE_TSN_Standards_Overview_and_Update_v4.pdf

Credit Based Shaper

– A frame is only allowed to transmit when the corresponding credit is not negative
– Credit decreases at the rate of *sendslope* (${\alpha_X}^-$) when a frame of corresponding class is in transmission
– Credit increases at the rate of *idleslope* (${\alpha_X}^+$) when frames of corresponding class are waiting or when there is no frame of the corresponding class but the credit is negative
– Positive credit is set to zero if there is no frame of the corresponding class

Credit Based Shaper – From IEEE802.1AVB

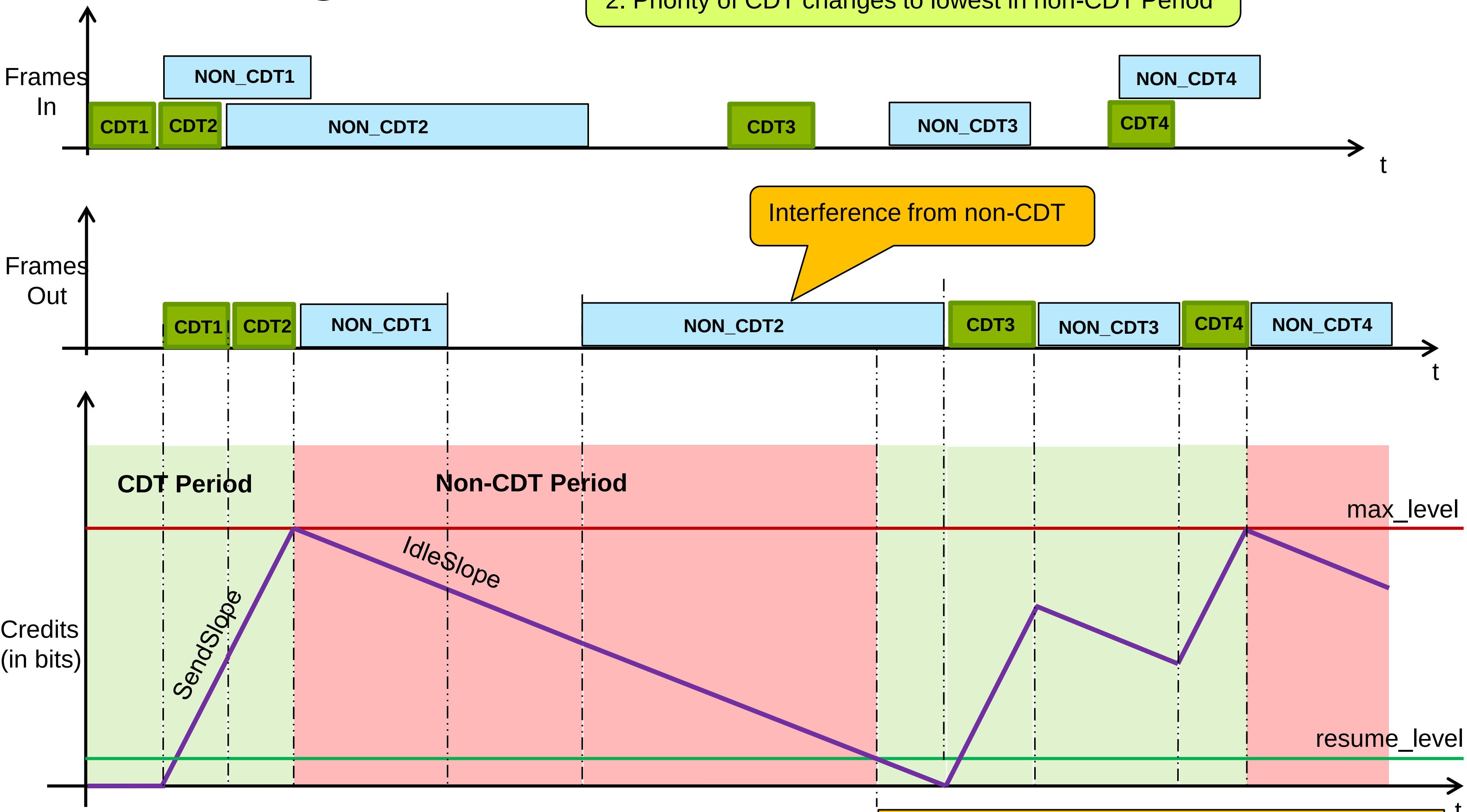


Hi_credits is fixed based on the maximum size of the interfering frame

Lo_credits is fixed based on the maximum size of the AVB frame

BLS - Algorithm

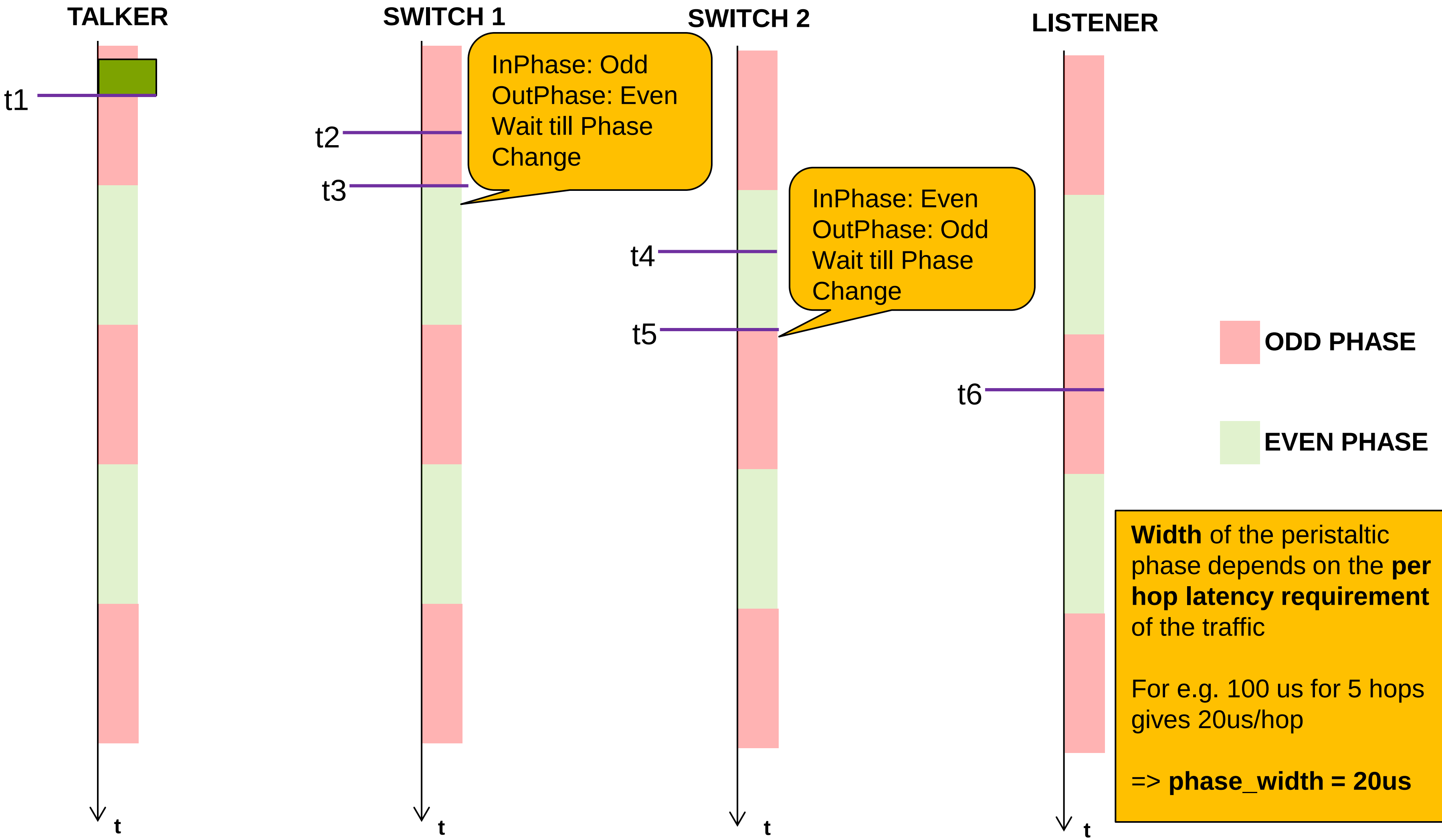
1. Condition to Tx : **Credits <= max_level**
2. Priority of CDT changes to lowest in non-CDT Period



max_level depends on the *transmission period* and *bandwidth fraction*
resume_level is fixed percentage of max_level

Peristaltic Shaper

Condition for Tx:
PeristalticOutPhase != PeristalticInPhase



Width of the peristaltic phase depends on the **per hop latency requirement** of the traffic

For e.g. 100 us for 5 hops gives 20us/hop

=> **phase_width = 20us**

Peristaltic Shaper - Implications

- Since all devices are time aware and in the same timing domain sharing the same cycle duration phase

  Worst Case Delay/hop = Cycle duration

- Multiple traffic classes corresponds to multiple delay classes
  - E.g: Class A AVB is 250us/hop, Control Data Traffic 100us/5 hops i.e, 20us/hop

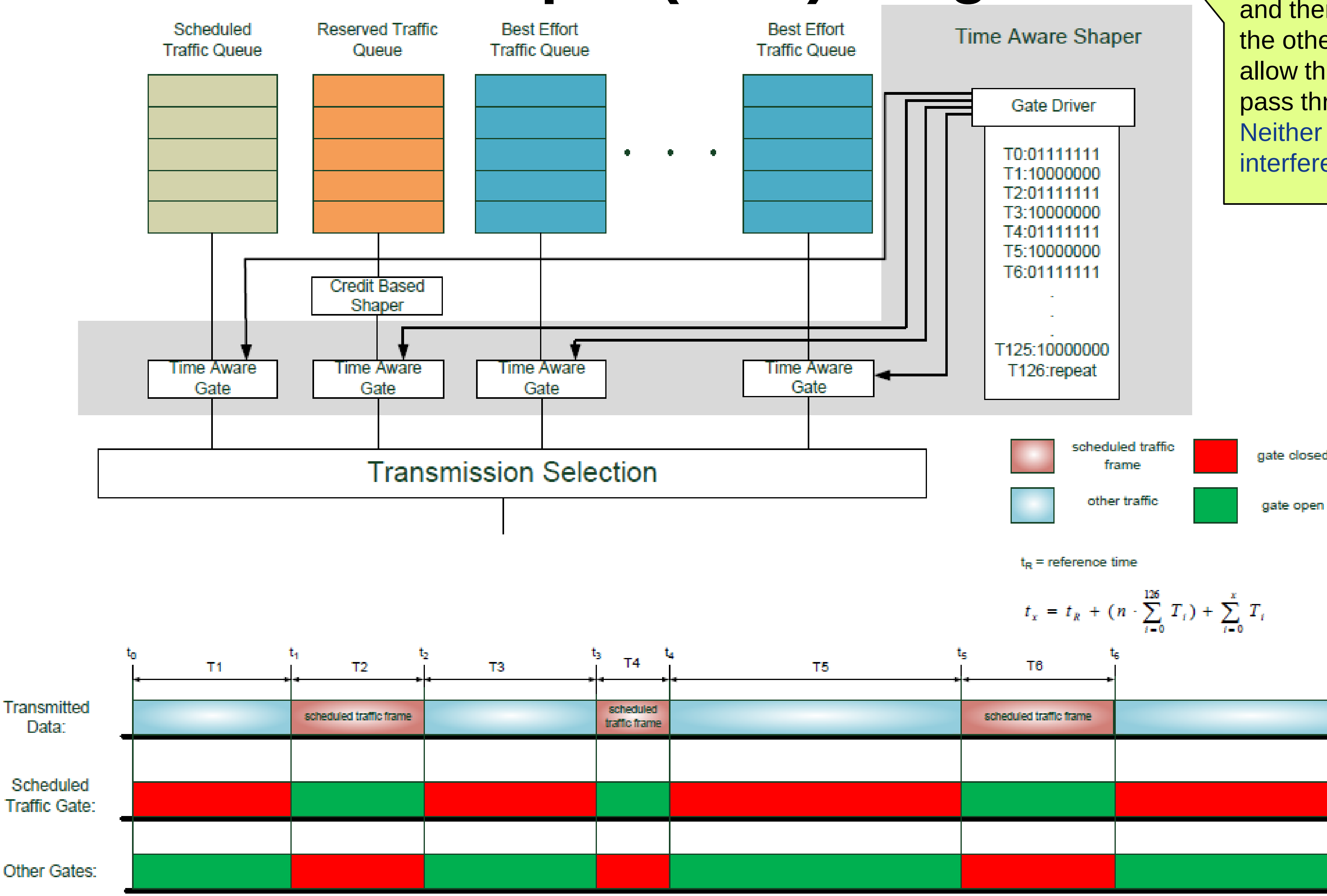
- One cycle time for one traffic/delay class
  - So need to track all the cycles

- Also possible to run a single cycle duration for all classes

- Simple assumption – Cycle durations are integer multiples of each other
  - Offers flexibility of less reconfiguration in case of addition of new streams
- Deterministic worst case latency in multiples of cycle times

Time Aware Shaper (TAS) - Algorithm

Precisely knowing when the scheduled traffic (CDT) arrives and then blocking all the other traffic to allow the CDT to pass through. Neither bursting nor interference.



From: http://www.ieee802.org/1/files/public/docs2013/bv-boiger-gates-vs-windows-scheduled-traffic-0113.pdf

Time Aware Shaper – Transmission

- A frame on a traffic class queue is not available for transmission
  - if the transmission gate is in the closed state or
  - if there is *insufficient time available to transmit the entirety of that frame* before the next gate-close event associated with that queue

- One GateEventList per port
  - Changes the gate state for each traffic class queue into open or closed

- The time-interval associated with a gate event is measured relative to the time at which the previous gate event in the list completed its execution

- The start time of the first gate event is calculated as

  tFirstGateEvent = n x tCycle

  n: an integer

  tCycle: length of the gating cycle for the queue, calculated as the sum of the time-interval parameters of the events in the list up to and including the Repeat event

From: http://www.ieee802.org/1/files/private/bv-drafts/d1/802-1Qbv-d1-1.pdf (802.1Qbv Draft 1.1)

Time Aware Shaper – Implications

- The entire schedule of Control Data traffic is to be known for one repeat period

- To allow multiple classes the repeat period could be fixed as the Least Common Multiple of all the transmission periods
  - For ex. CDT Class A 25us, CDT Class B 125us the repeat period can be 125us
  - Longer the period of CDT, longer the repeat period
  - But it is to be noted that the Gate event List has an upper bound

- The width of each gating event (gate open for Control Data) can be adjusted according to the Bandwidth Fraction allowed for CDT traffic class
  - A margin can be included for future use to avoid re-configuration
  - But this will be wasted if not used

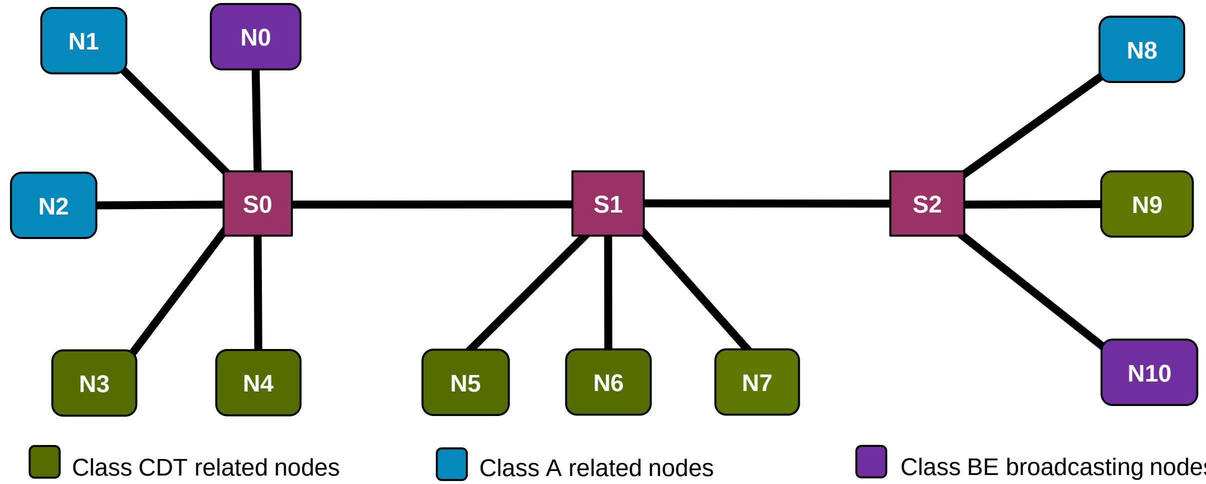
- The size of the frame in each queue needs to be known in-order to ascertain whether frame can be transmitted before gate close event
  - To minimize overhead: assume all are equal MaxSized frames and MaxSize is known

Questions

- How much delay can a frame experience when traveling multiple hops?
- For 1, or for multiple switches, what is the delay for a particular set of shaped streams?

- Methods:
  - simulation
  - analysis, analytical expressions

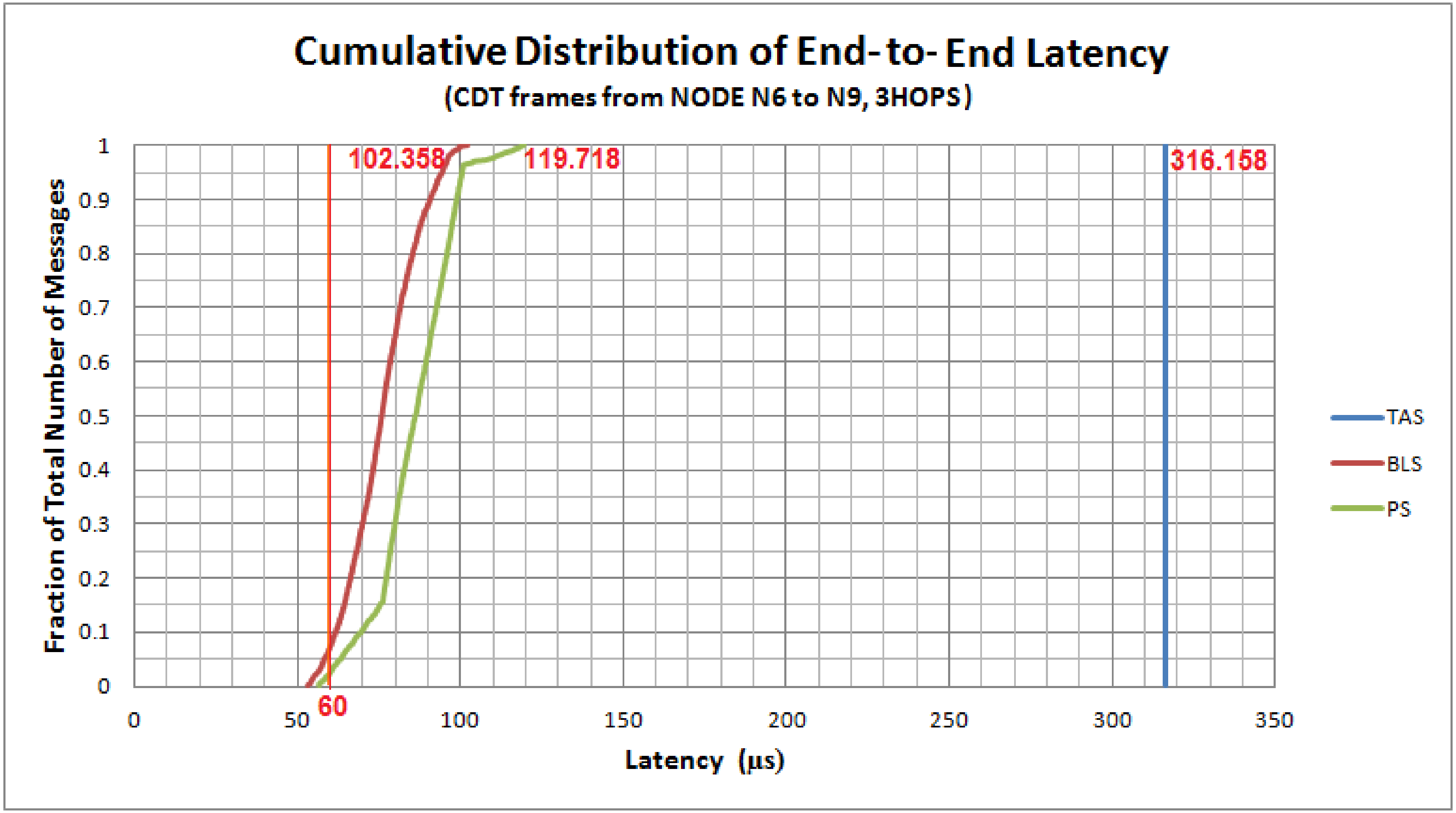
Traffic Information



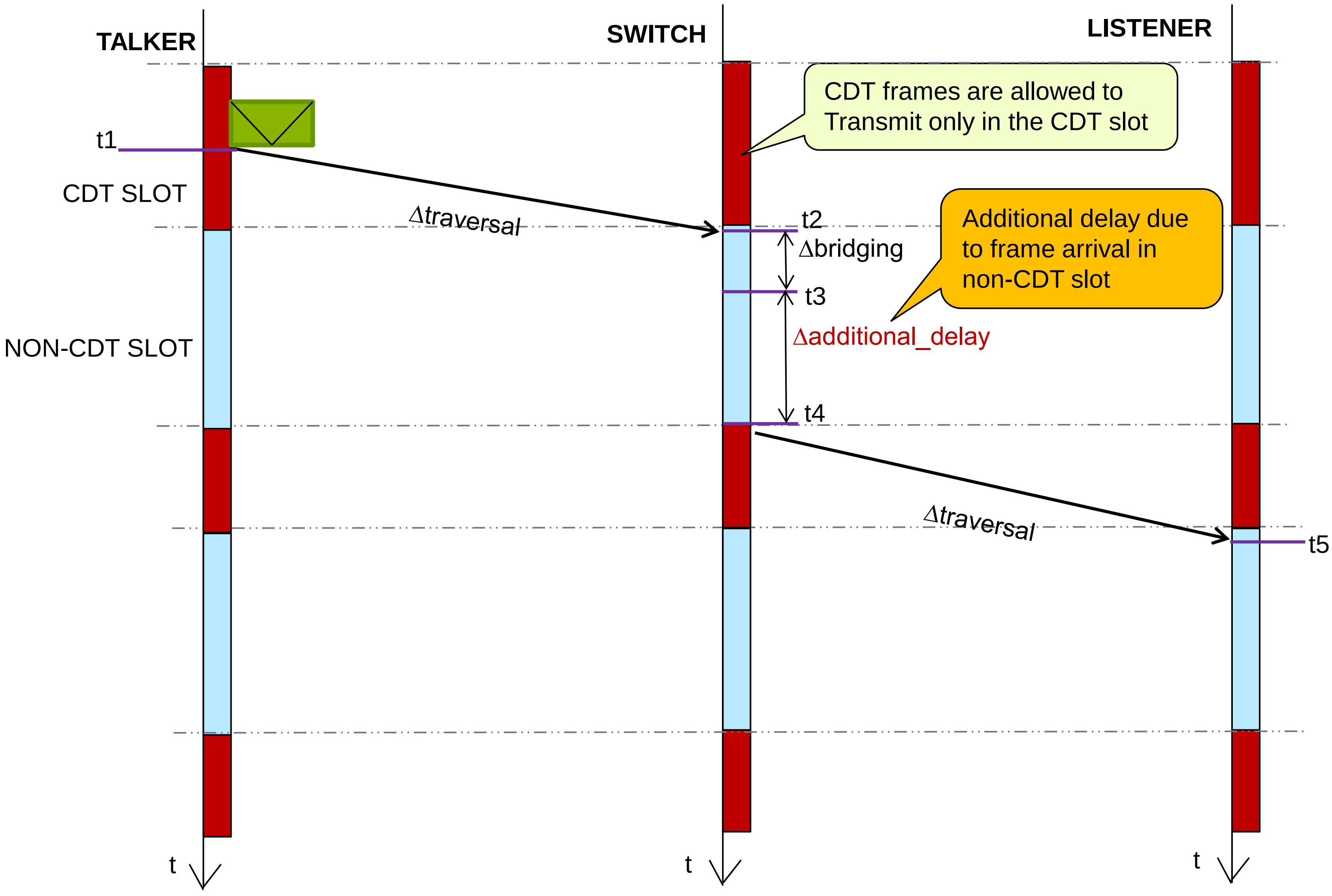
Source	Destination	Messaging Type	Class	Interval (µs)	Frame Size (Bytes)		Bandwidth with Header (% of total bandwidth)	End-to-End Latency Expectation
					Payload	With Header		
N0	All nodes except N0	Broadcast	BE	Random (Poisson)	256	298	#	-
N1	N8	Unicast	AVB_A	125	256	322	20.608	2ms/7 hops
N2	N8	Unicast	AVB_A	125	256	322	20.608	2ms/7 hops
N3	N7	Unicast	CDT	500	128	170	2.72	100 µs/ 5 hops
N4	N7	Unicast	CDT	500	128	170	2.72	100 µs/ 5 hops
N5	N9	Unicast	CDT	500	128	170	2.72	100 µs/ 5 hops
N6	N9	Unicast	CDT	500	128	170	2.72	100 µs/ 5 hops
N10	All nodes except N10	Broadcast	BE	Random (Poisson)	256	298	#	-

- Share of rest of the total bandwidth

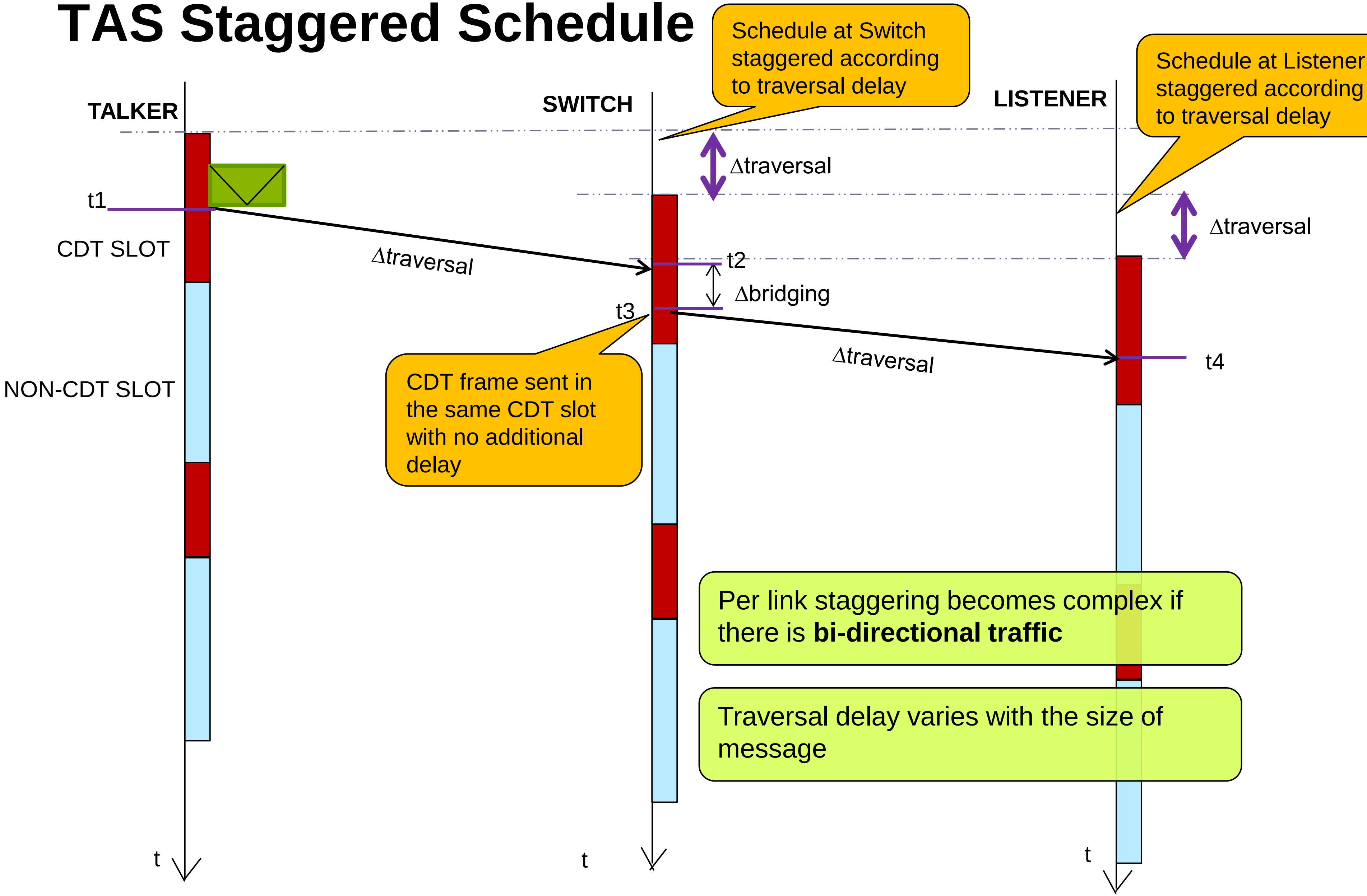
Maximum Latency for Control Data Traffic



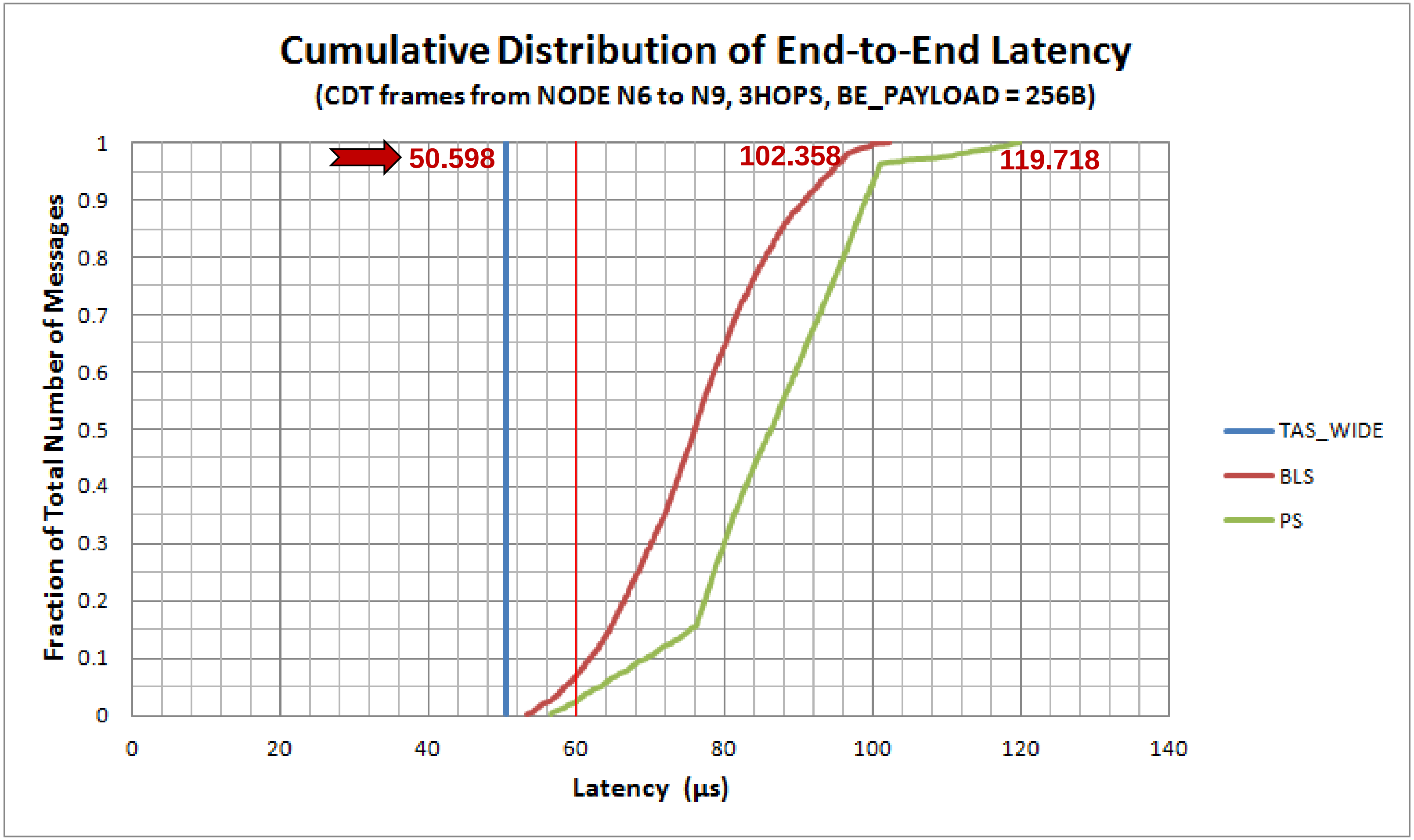
TAS Worst Case



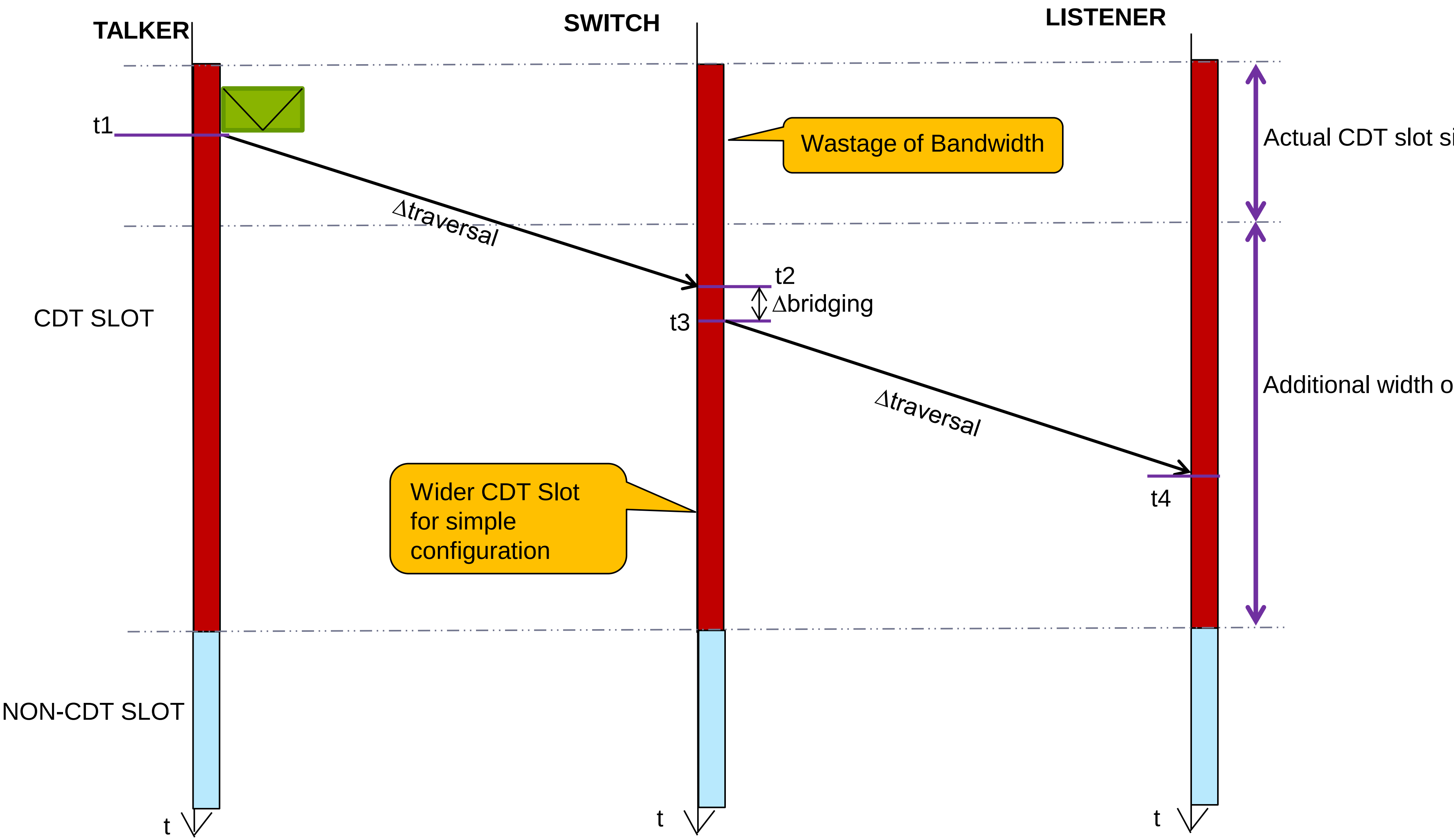
TAS Staggered Schedule



Maximum Latency for Control Data Traffic – TAS Wider Slot



TAS with Wider CDT Slots



Analysis

- **How to derive a tight WCRT bound?**