Analysing TSN within network calculus framework

Marc Boyer

École d'Été Temps Réel 2021
23 septembre 2021

Outline

What is TSN ?

TSN is not a technology

- TSN is the name of a IEEE task group of the IEEE 802.1 Working Group
  - TSN : Time-Sensitive Networking
  - http://ieee802.org/1/pages/tsn.html
  - https://1.ieee802.org/
- Documents : Naming : 802.1Q, 802.1ad, and 802.1Qat...
  From one up to even four letters after 802.1
  - Uppercase : standards
  - Lower-case : amendments
  - -REV : revision (more extensive changes to the existing text than can be undertaken in an amendment)
- Document Access :
  - Working documents : need to be member (≈)
  - Published standard :
    - ≈ free after 6 months : "IEEE Standards runs a Get IEEE802 program that allows anyone to download the standards for free, 6 months after publication."
    - Or buy it

TSN promises

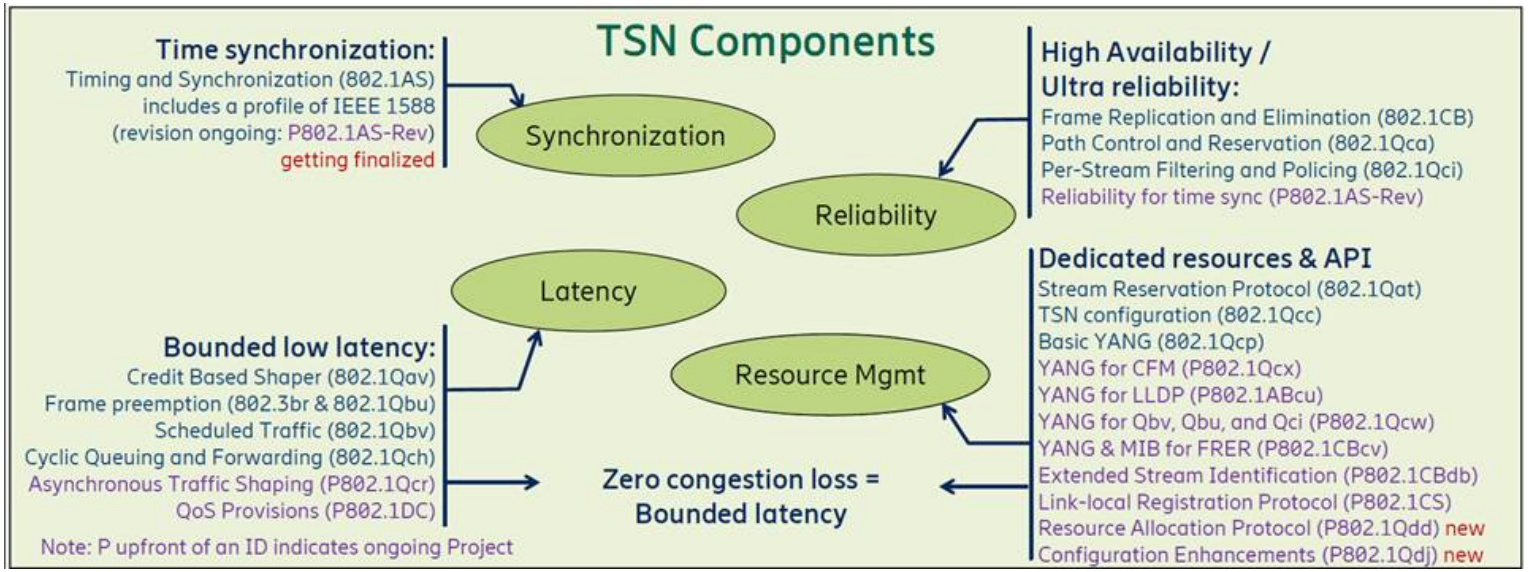


Figure – TSN Overview, J. Farkas [?]

Outline

An Ethernet network

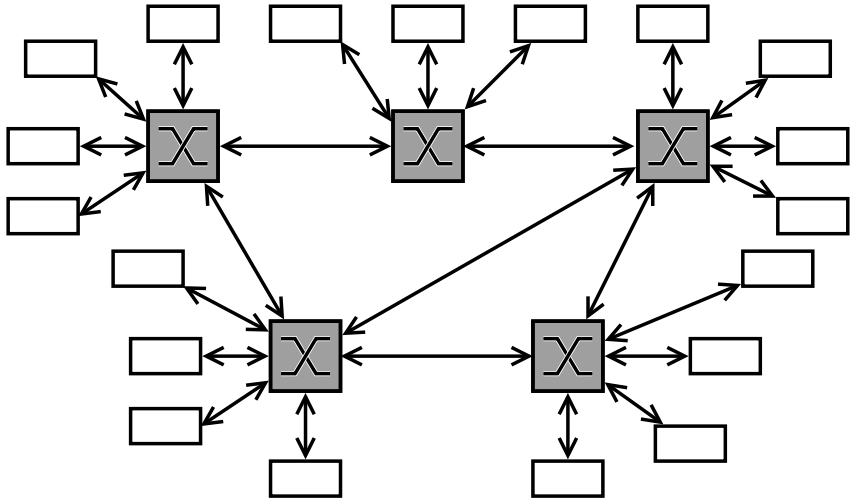


Figure – Principle of Ethernet network (switch-based)

- full duplex links
- propagation delay : signal transmission ($\approx$ 60% light speed)
- main delay : in switches
- routing, frame format : lack of time

An Ethernet switch

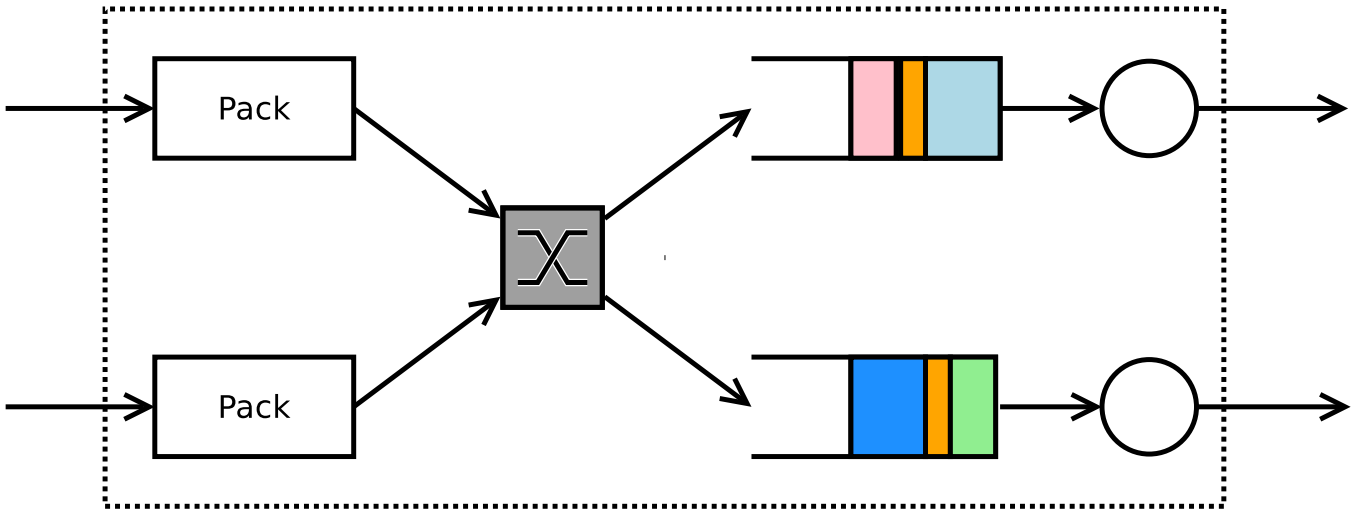


Figure – Common architecture of Ethernet switch

- input ports : frame arrivals
- switching : copy in destination port(s)
- output port : queuing + transmission

An 8 priority level Ethernet switch

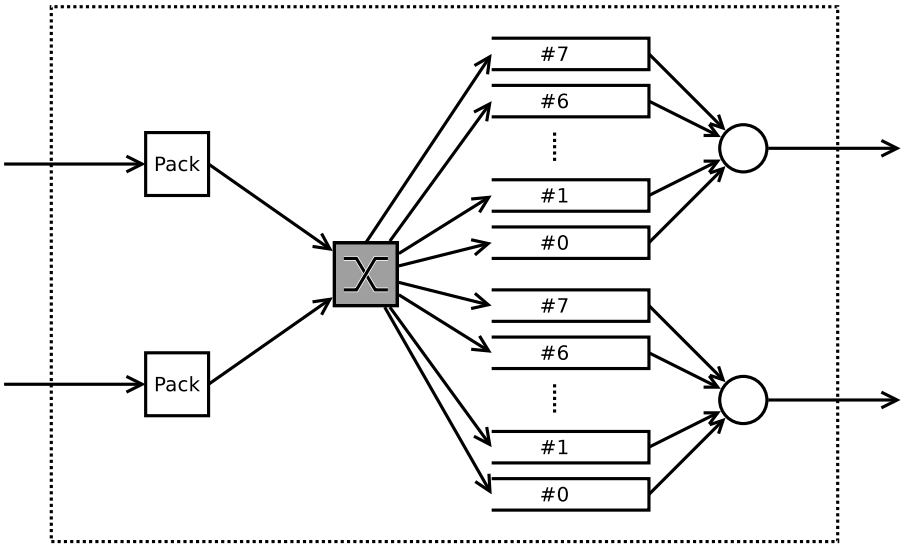


Figure – Ethernet switch with priority levels

- non-preemption : up to 1542B blocking
- preemption (802.3br, 802.1Qbu) :
  - partial blocking (up to 148 B) + overhead
  - single-level preemption

Outline

Outline

Main TSN addenda

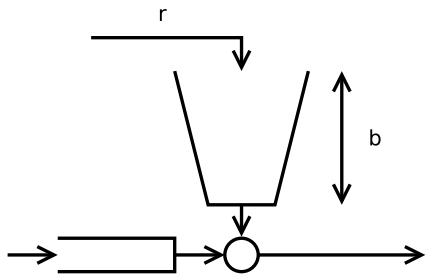
- Frame preemption (802.3br, 802.1Qbu)
- Synchronisation mechanisms (algorithms, architecture, protocols) 802.1AS-Rev
- Resource reservation, access control, configuration, signalisation, stream identification (802.1Qat, 802.1Qcc, 802.1CBdb, 802.1Qca, 802.1Qdd...)
- Safety and reliability :
  - Input port policing : 802.1Qci
  - Redondancy : 802.1CB
- Output port scheduling :
  - Credit Based Shaper, CBS (802.1Qav)
  - Scheduled Traffic (802.1Qbv)
  - Cyclic Queuing and Forwarding (802.1Qch)
  - Asynchronous Traffic Shaping, ATS (802.1Qcr)
  - ETS for bandwidth sharing (802.1Qaz, pre-TSN)

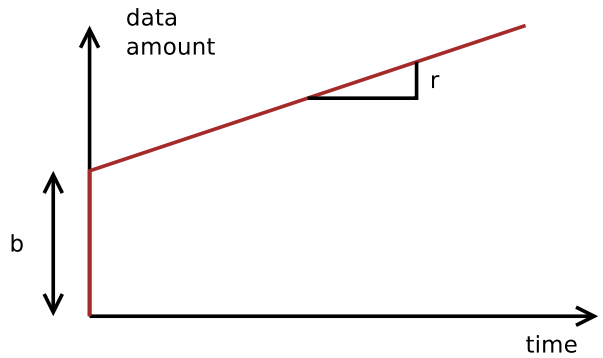
Outline

The token-bucket model

- two parameters :
  - throughput r,
  - burst b (aka capacity, depth)
- the bucket rules
  - the bucket is initially full of b tokens
  - sending a frame of size s consumes s tokens
  - the bucket fills with rate r tokens per time unit
  - can never be negative nor exceed b
- in case of insufficient tokens
  - drop the frame : policing
  - queue until enough : shaping
- property : on *any* observation interval of duration d, the data amount is less than

$$b + d \cdot r \qquad (1)$$

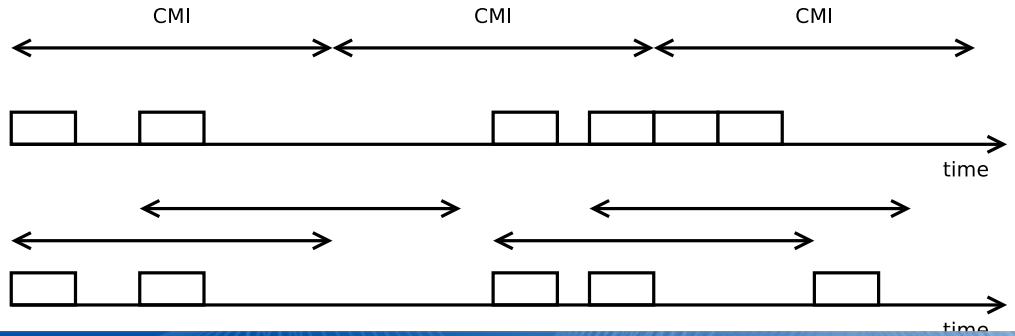




- a periodic flow with frames of size S and period T respects token-bucket $b = S$, $r = S/T$

Flows contract

- notion of *stream*
- several "traffic specification"
- the AVB stream traffic specification
  - Traffic Specification associated with a Stream [**?**, § 35.2.2.8.4 TSpec]
    - MaxFrameSize : the maximum frame size
    - MaxIntervalFrames : the maximum number of frames that the Talker may transmit in one "class measurement interval" (34.4).
    - Class Measurement Interval (CMI) : static, per class (in 0-7)
  - Semantics : tumbling window vs. sliding window
    TSpec as token-bucket
    - sliding $\implies r^s = r, b^s = b$
    - tumbling $\implies r^t = r, b^t = 2b$



$$b = \text{MFS} \cdot \text{MIF}$$

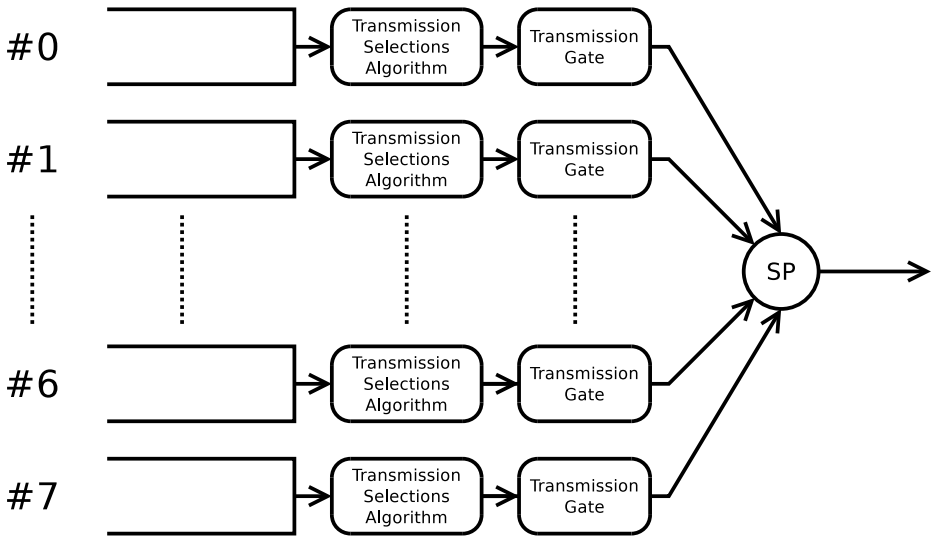
$$r = \frac{b}{\text{CMI}}$$

Input port policing : 802.1Qci

- 802.1Qci : Per-Stream Filtering and Policing – PSFP
- done at input port
- associates a token-bucket to a (configurable) set of streams
- drop “out of contract” frames

Outline

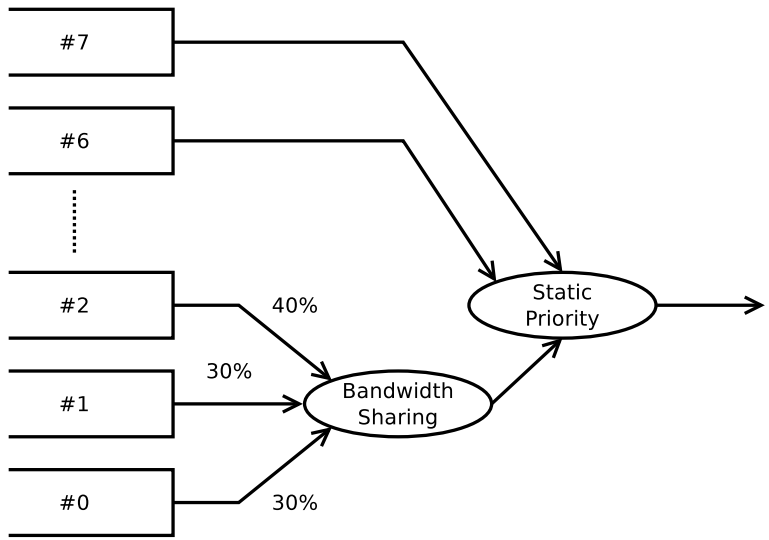
TSN output port



- Transmission Selection Algorithm :
  - per queue choice
  - one in "none, CBS, ATS, ETS"
- Transmission gate :
  - a gate is either open or closed
  - based on a static cyclic schedule

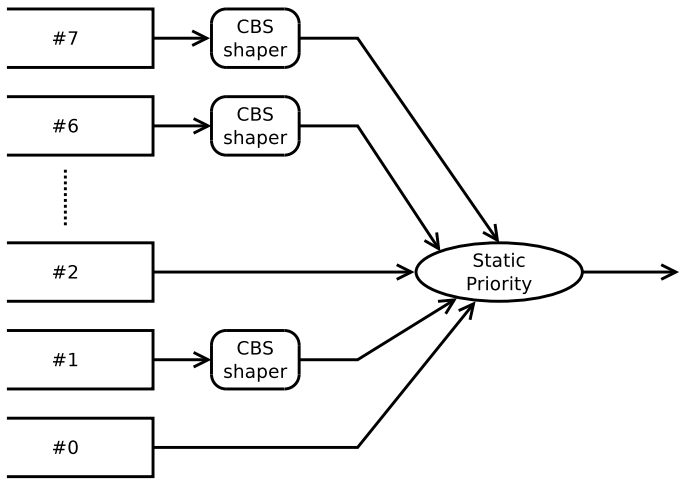
802.1Qaz : Bandwidth Sharing (SP/WRR, SP/DRR... – pre-TSN)

- Enhanced Transmission Selection for Bandwidth Sharing Between Traffic Classes (aka ETS)
- 802.1Qaz, 2011 (pre-TSN)
- Simple hierarchical scheduling : Static priority + Round-Robin-like
- Introduced for data centers
- Sharing the leftover bandwidth
- Bandwidth Sharing is implementation-defined
  - WRR cited in the standard
  - DRR used in Linux
  - not able to find choice of Cisco, Juniper...

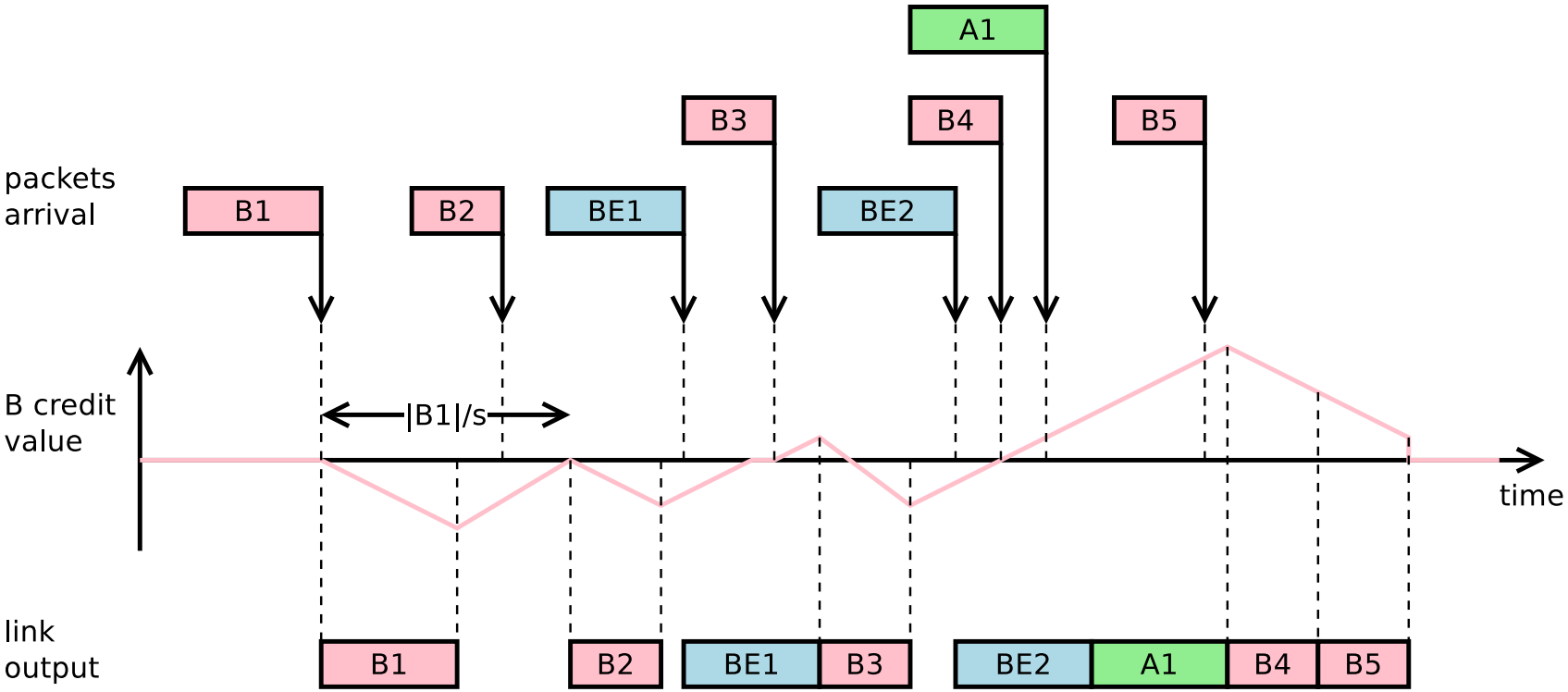


802.1Qav : Credit-Based Shaped (CBS)

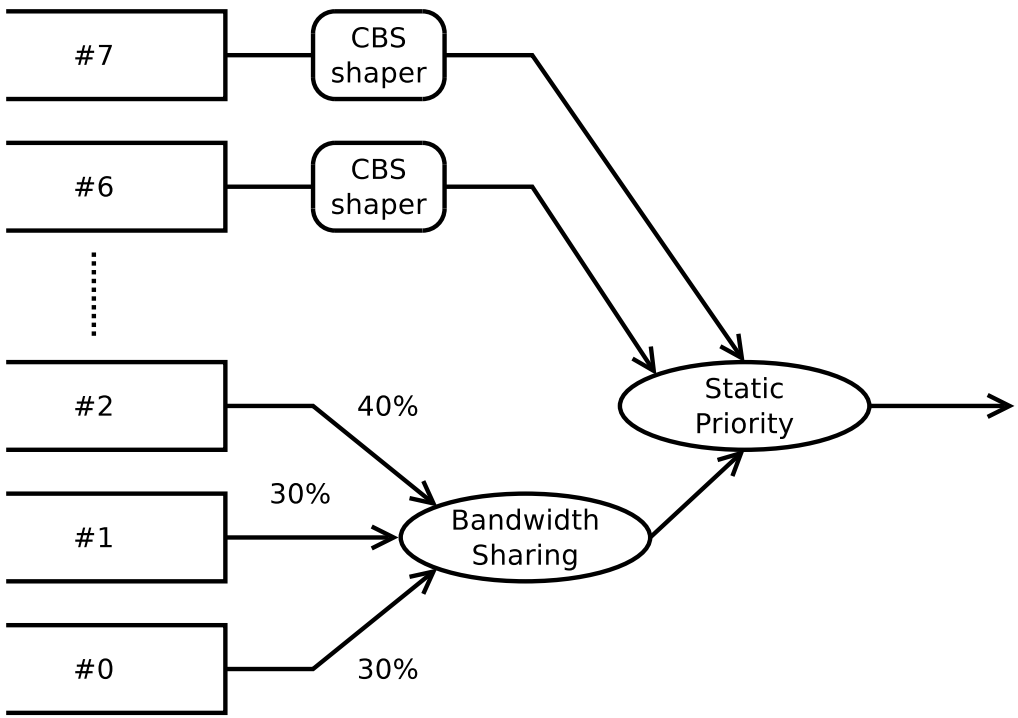
- "Forwarding and Queuing Enhancements for Time-Sensitive Stream – FQTSS"
- 802.1Qaz, 2011 (AVB, pre-TSN)
- CBS shaper is optionnal
- Each CBS shaper has a "slope" s parameter (in bit per second)
- A credit increases when the queue waits, and decreases when the queue transmits
- It limits the associated queue to throuput s
- Its shapes/spreads/smoothes the output
- Designed to
  - avoid starvation
  - limit jitter



Example of CBS credit evolution rule

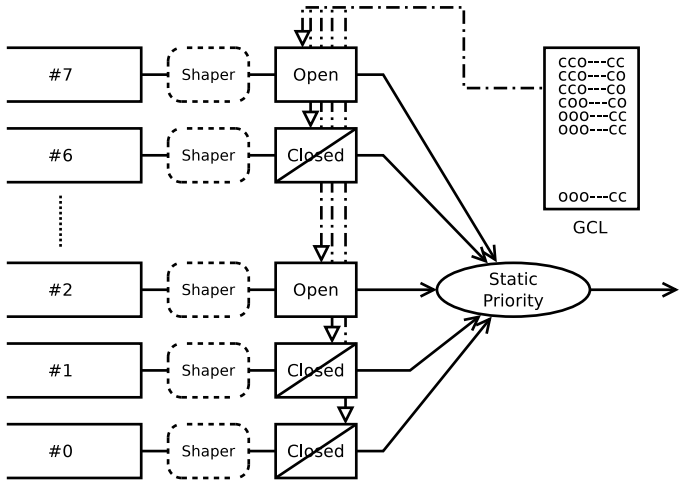


802.1Qaz + 802.1Qav : ETS+CBS

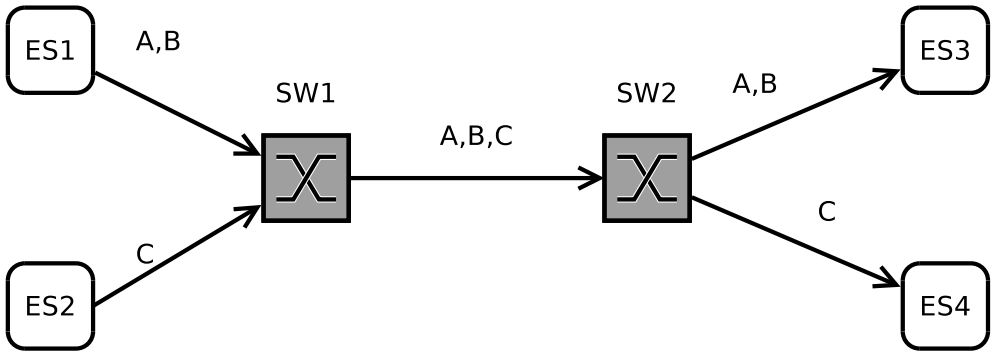


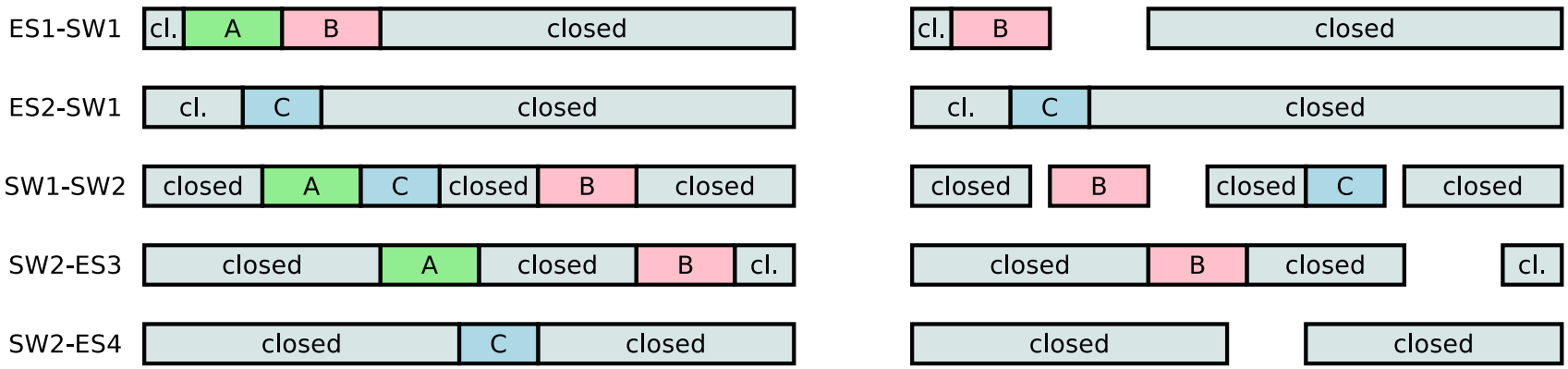
802.1Qbv : Time Aware Shaper – TAS

- "Enhancements for Scheduled Traffic"
- A *gate* is associated to each queue
- The gate is either open or closed
- A global cyclic schedule (Gate Control List – GCL), w.r.t local clock
- Building schedule is out of standard
- "Exclusive gating" ≈ one gate opened at a time
- Integration with GCL : update of credit evolution rules
- End-to-end TT schedule requires
  - global build of local schedules
  - synchronisation of local clocks (eg. 802.1AS)



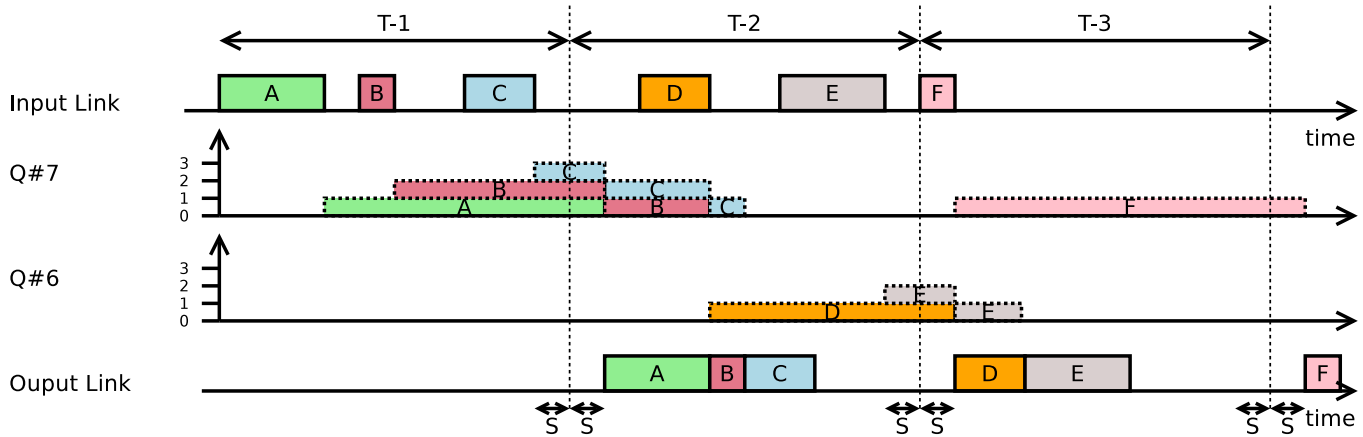
TAS : a Time-Trigerred implementation ?





802.1Qch : Cyclic Queuing and Forwarding – CQF

- Not a new "mechanism" : based on 802.1Qci (Filtering) and 802.1Qbv (Time Aware Shaper)
- Divide time into time intervals of common length T
- Frames received in one interval are forwarded in the next one

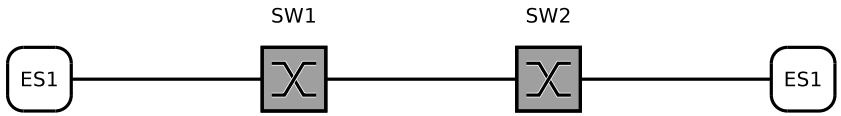


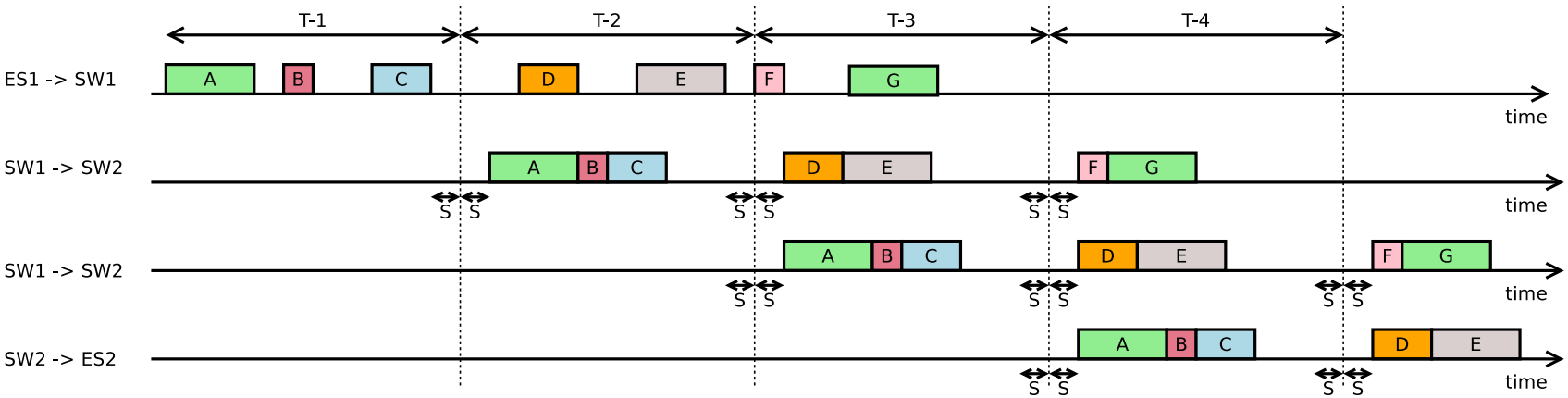
CQF performances

- Global synchronisation

  $\Longrightarrow$ Low jitter (2T)

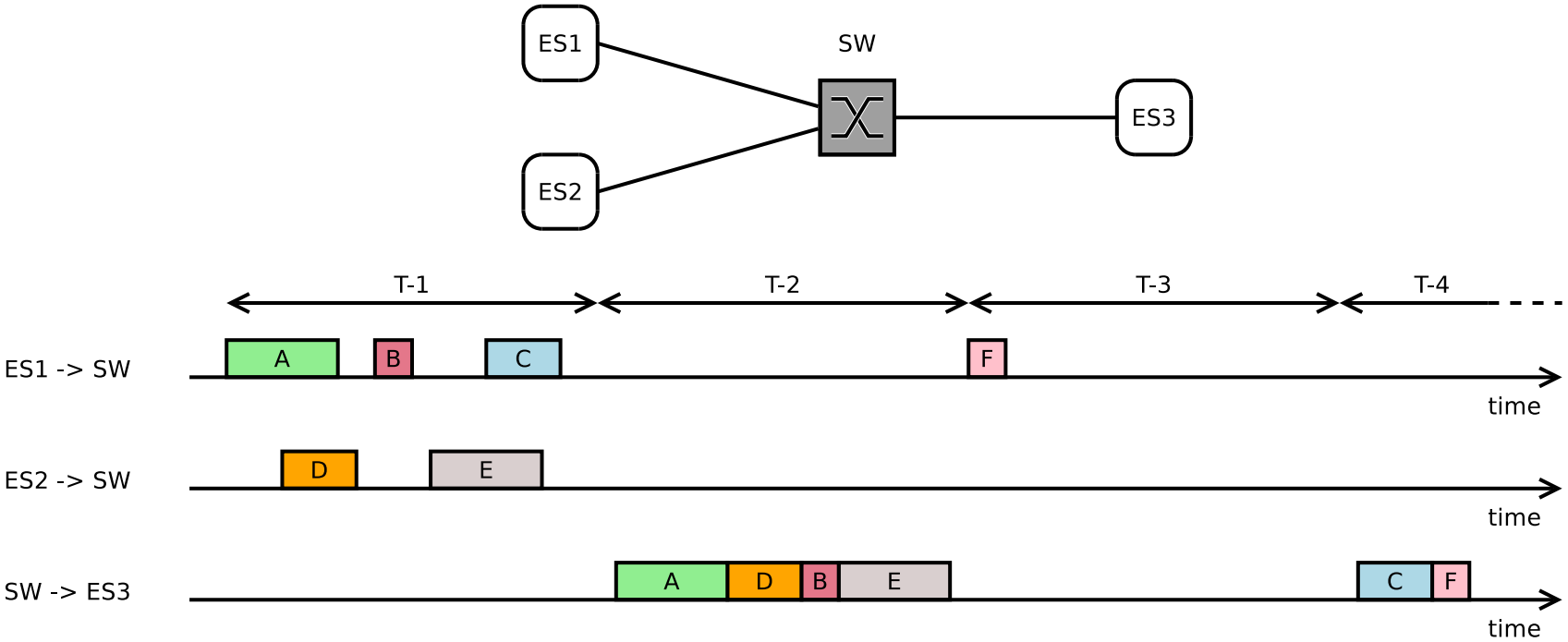
  $\Longrightarrow$ simple delay computation (T $\times$ nb of hops)





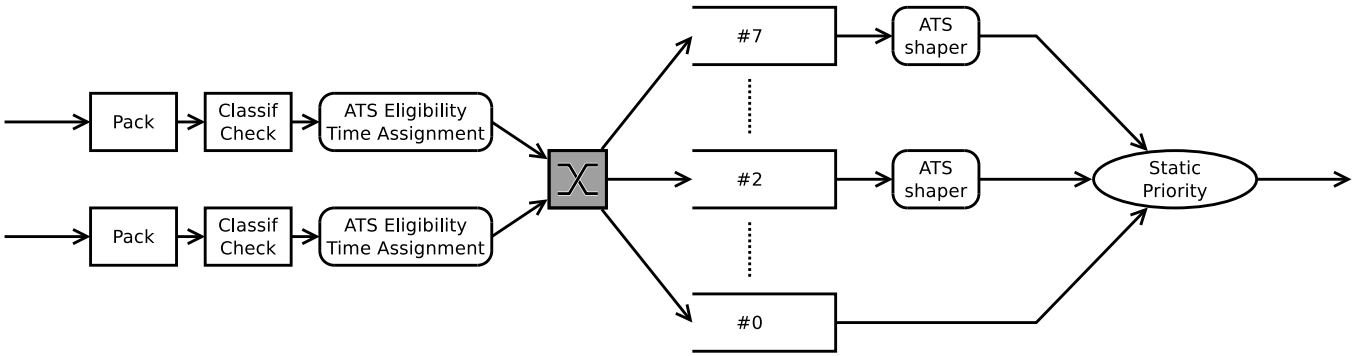
CQF configuration

- Cycle time must be “large enough” w.r.t. bursts



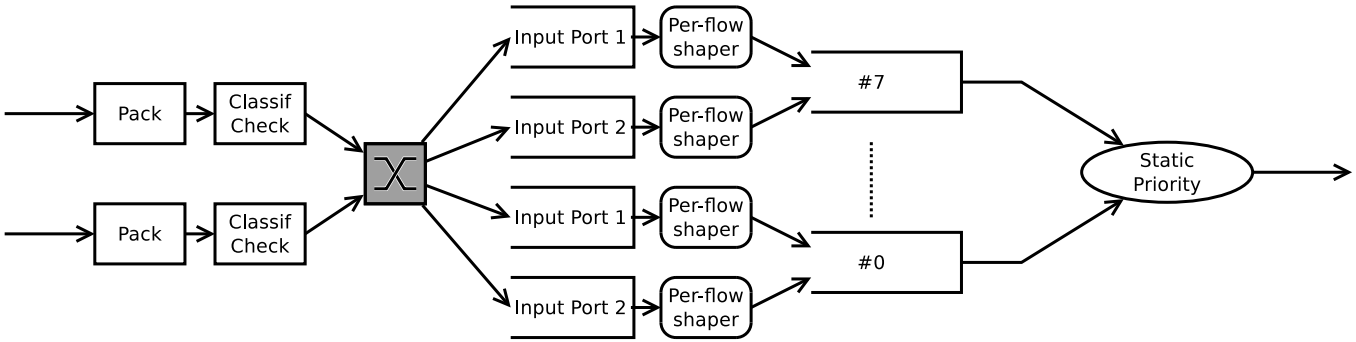
802.1Qcr : Asynchronous Traffic Shaping – ATS

- Queue waiting create bursts / jitter
- ATS introduces delay to absorb the jitter
  - computes a “Eligibility Time” per frame
    - a local value (no global synchronisation)
    - token-bucket parameters
    - use some share variables between ATS schedulers
  - head of queue can not be selected before this Eligibility Time



ATS : implementation and equivalent model

- Complexity relies in computattion of “Eligibility Time”
- Computed in order to be equivalent to group reshaping (token bucket) with aggregate queuing
- A major theoretical breakthrough
  - reshaping comes for free
  - avoid cyclic dependency problem



Outline

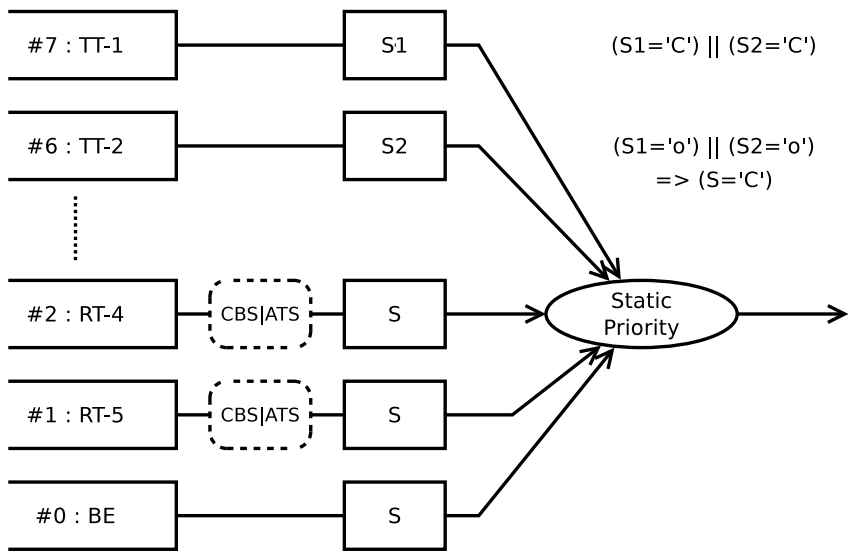
The most obvious one : TT/Shaper/BE

- TT queues : for very low latency and jitter flows
- CBS|TAS queues : for real time
- Best Effort

Principles :

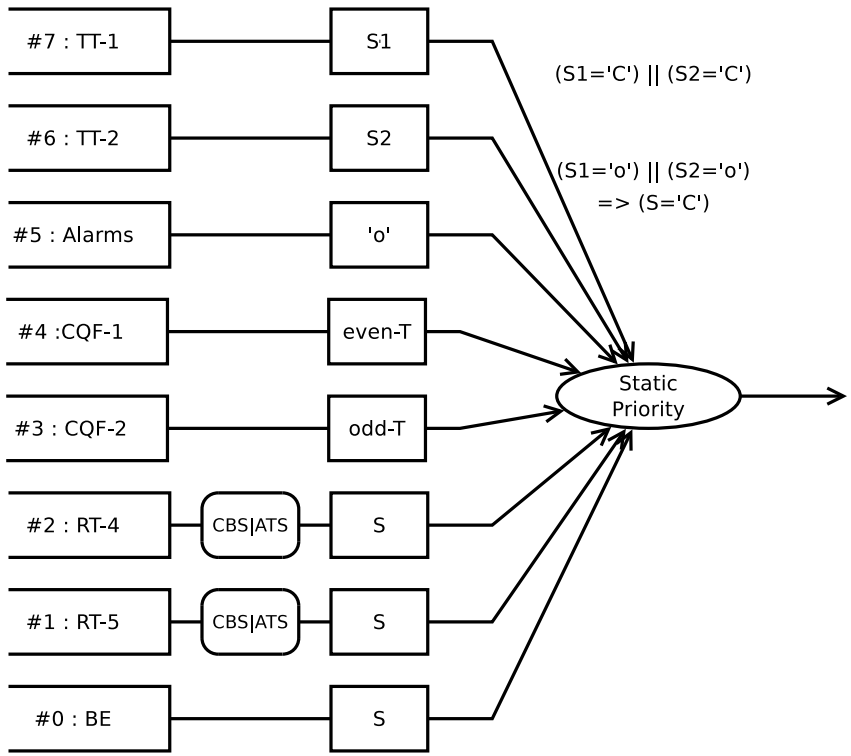
- build TT queue GCL wrt TT behaviour, no shaper for TT queues
- set other GCL queues as the opposite (exclusive gating)
- set BE at lower priority
- configure CBS or ATS wrt expected workload

Rq : exclusive gating allows TT files to use any priority level.



With alarms and CQF

- TT queues : for very low latency and jitter flows
- Static priority : for asynchronous alarms
- CQF
- CBS|TAS queues : for real time
- Best Effort



Outline

TSN conclusion

- the next real-time network
- a lot on industry involved
- able to host several kinds of flows
- offering several scheduling policies
- how to configure it ?
- how to bound buffer usage and delay ?

Outline

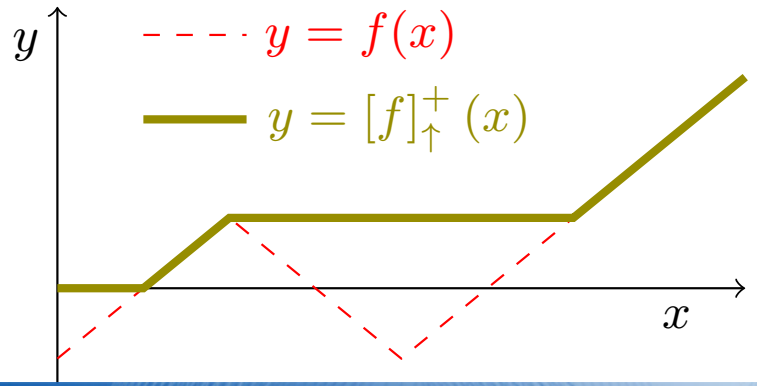
What is network calculus ?

- another theory for real-time (computes response-time bound)
- based on (min,plus) dioid theory
- used to certify AFDX network in A380, A400M, etc.
- several tools (e.g. RTaW-PEGASE)
- share common aspects with Event Stream theory [**?**]

Notations

- $\mathbb{R}$: the set of real numbers,
- $\mathbb{R}^+$ the subset of non-negative real numbers,
- $\mathbb{Z}$ the set of integers,
- $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{Z}$ the ceiling function ($\lceil 1.2 \rceil = 2, \lceil 4 \rceil = 4, \lceil -1.2 \rceil = -1$)
- $\forall x \in \mathbb{R}, [x]^+ = \max(x, 0)$
- $\forall f : \mathbb{R}^+ \to \mathbb{R}$, its non-decreasing non-negative closure is defined by

$$[f]^+_\uparrow (t) = \max_{0 \leq s \leq t} [f(s)]^+ . \tag{2}$$



Outline

Modeling data flows

Definition : Cumulative curve

$\mathcal{C}$ denotes the set of functions

- from $\mathbb{R}^+$ to $\mathbb{R}^+$
- non-decreasing
- piece-wise continuous
- left-continuous

- An element $A \in \mathcal{C}$ is used to model a data flow in the network
- A is called "cumulative curve"
- $A(t)$ represent the amount of data the amount of data from a flow observed at some point up to time t
- A lot of information lost