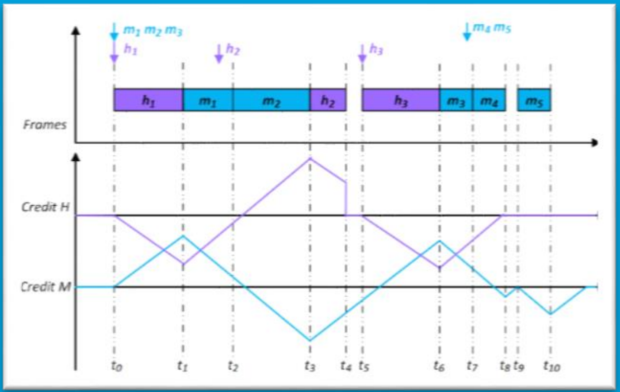
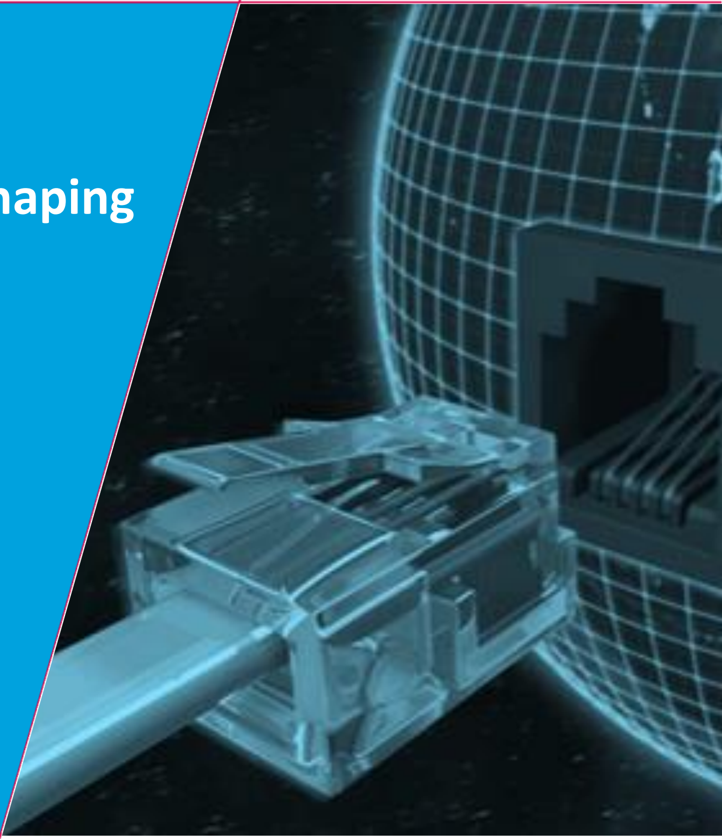
Ethernet TSN

and the worst-case response time of credit-based shaping



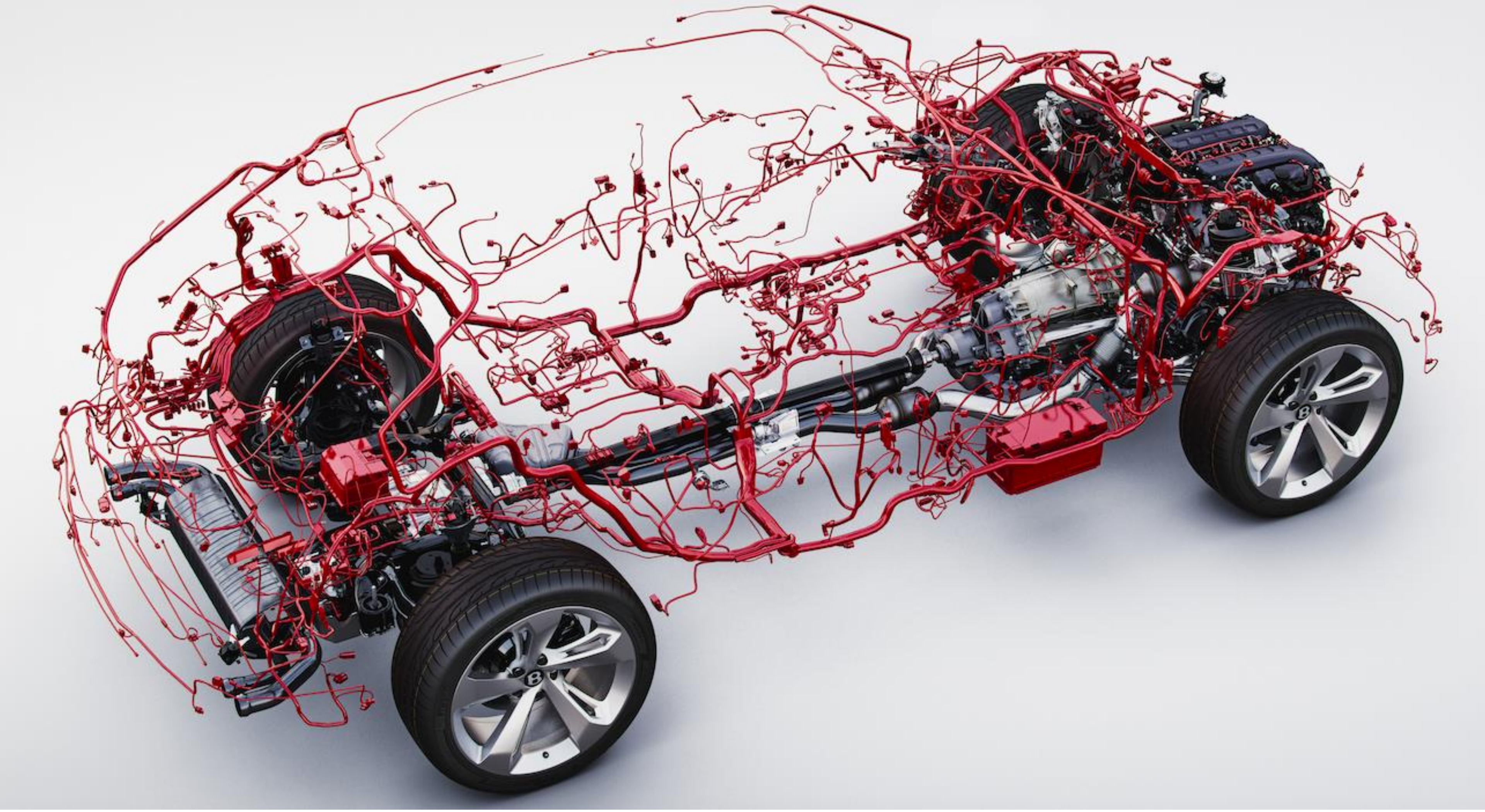
by Pieter Cuijpers



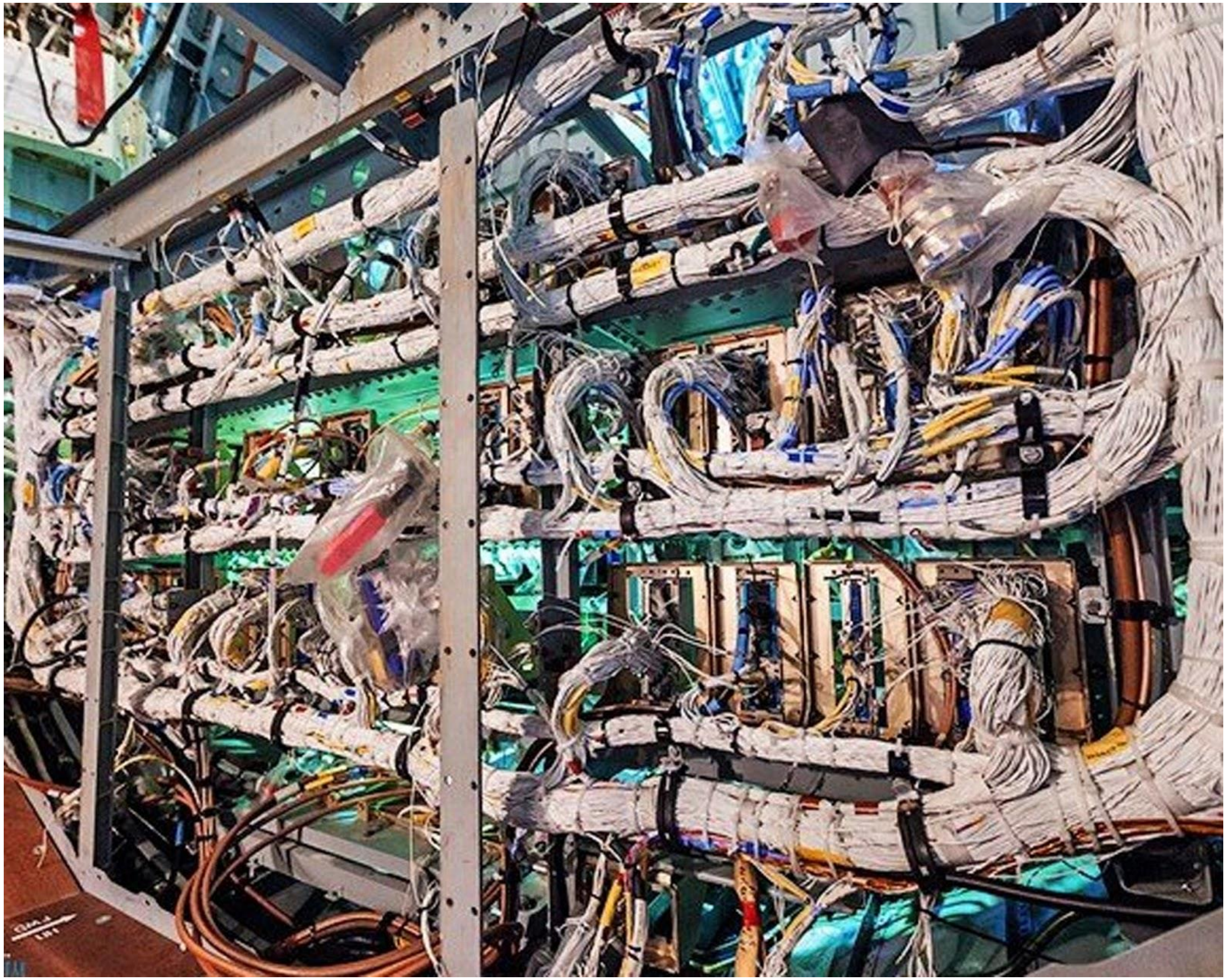
Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

Why are real-life networks so complex?



Why are real-life networks so complex?



source: electrical-engineering-portal.com

Why are real-life networks so complex?

source: electrical-engineering-portal.com

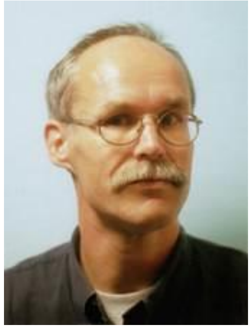
The system architecture and networking group

Internet of Things
Automotive

Platforms with predictable performance

Quantitative Formal Modeling

SAN

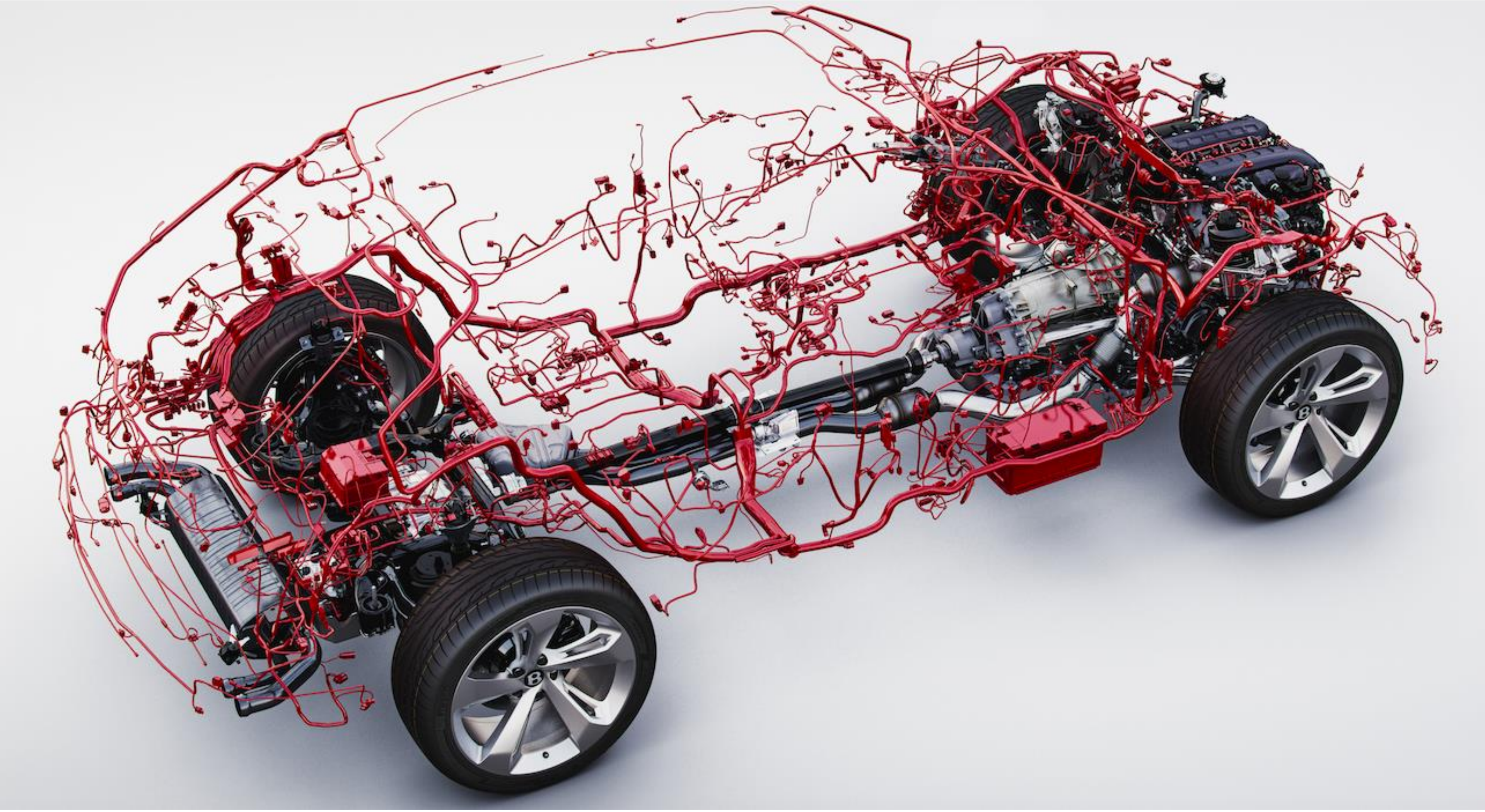
Jingyue Cao | me ☺ | Reinder Bril | Johan Lukkien

Jingyue Cao, Pieter J.L. Cuijpers, Reinder J. Bril, and Johan J. Lukkien. 2016. Independent yet Tight WCRT Analysis for Individual Priority Classes in Ethernet AVB. In Proceedings of the 24th International Conference on Real-Time Networks and Systems (RTNS '16). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/2997465.2997493

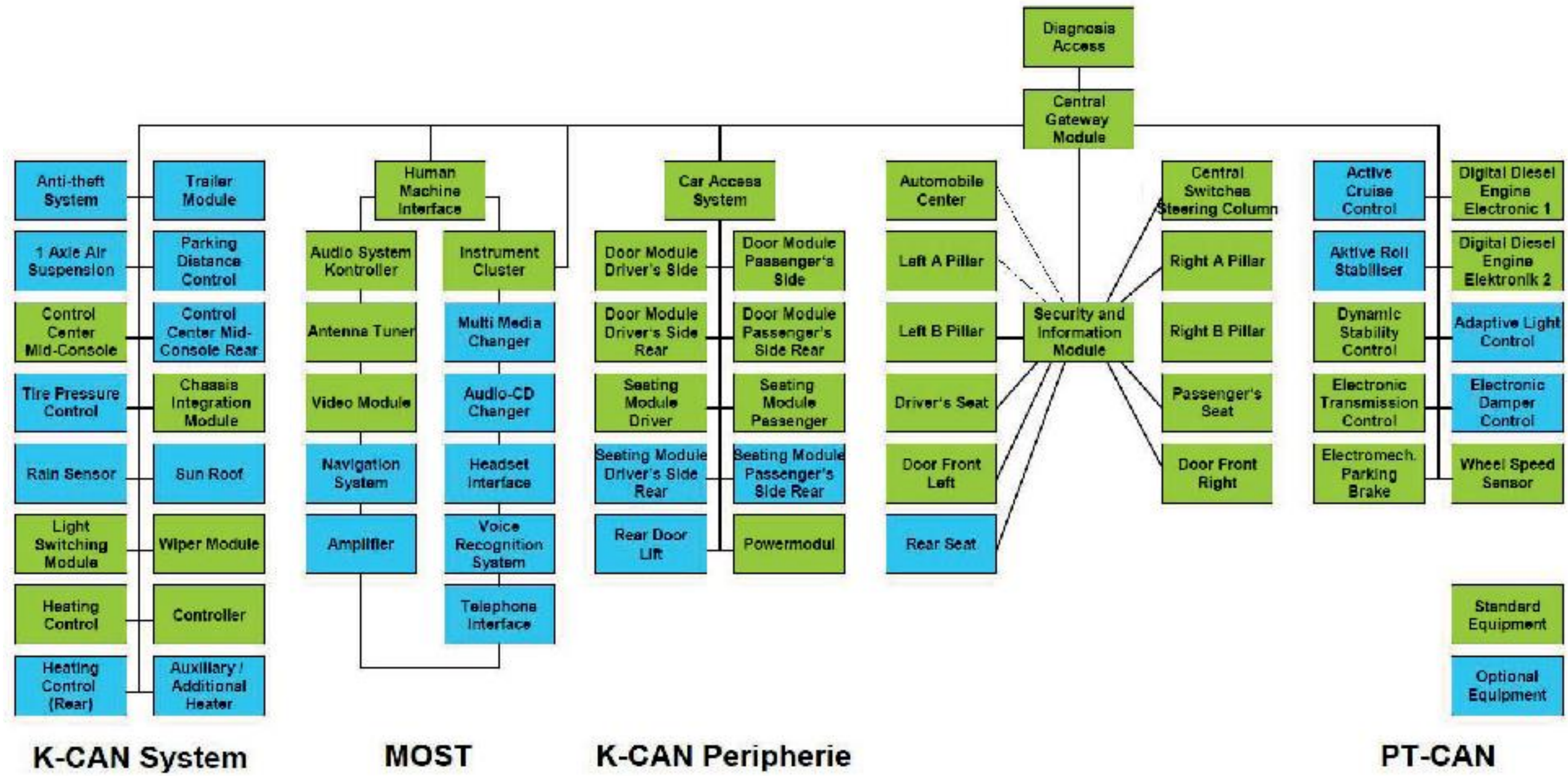
H. J. Rivera Verduzco, P. J. L. Cuijpers and J. Cao, "Work-in-Progress: Best-Case Response Time Analysis for Ethernet AVB," *2017 IEEE Real-Time Systems Symposium (RTSS)*, Paris, 2017, pp. 378-380. doi: 10.1109/RTSS.2017.00043

J. Cao, M. Ashjaei, P. J. L. Cuijpers, R. J. Bril and J. J. Lukkien, "An independent yet efficient analysis of bandwidth reservation for credit-based shaping," *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*, Imperia, 2018, pp. 1-10. doi: 10.1109/WFCS.2018.8402345

Why are real-life networks so complex?

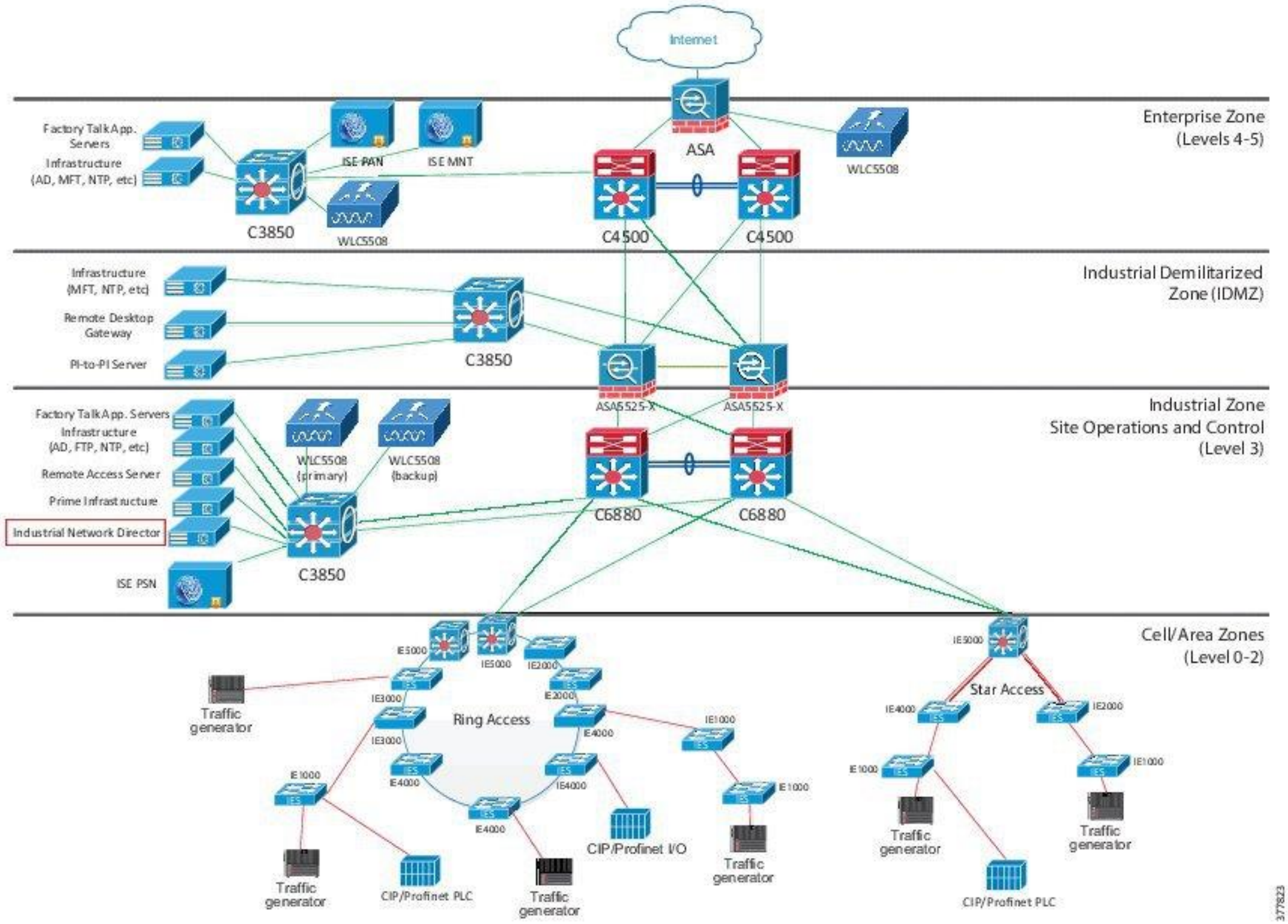


Why are real-life networks so complex?



Source:
Performance Study of an In-Car Switched Ethernet Network without Prioritization
Hyung-Taek Lim, Kay Weckemann, Daniel HerrscherPublished 2011 in Nets4Cars/Nets4Trains
DOI:10.1007/978-3-642-19786-4_15

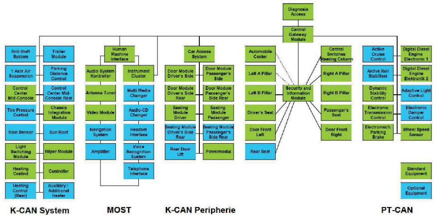
Why are real-life networks so complex?

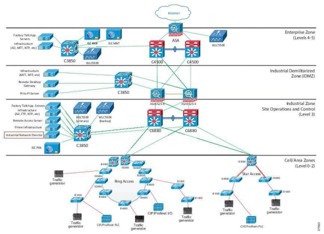


source: www.cisco.com

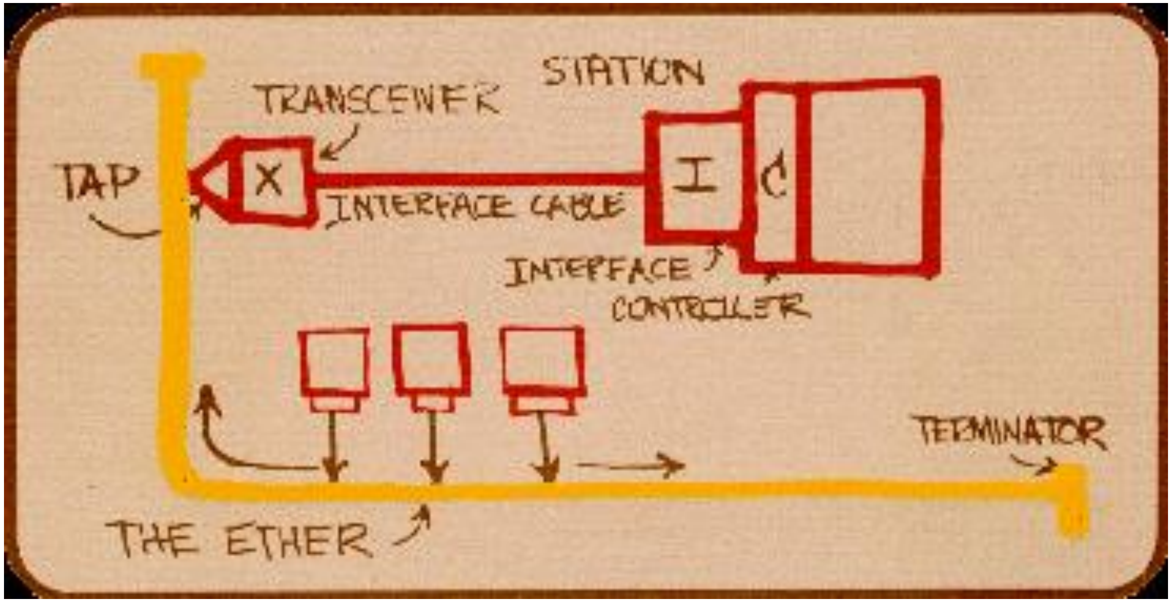
Why are real-life networks so complex?

Why are real-life networks so complex?





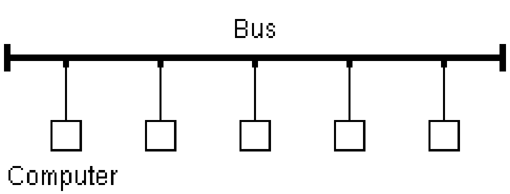
A little history of Ethernet

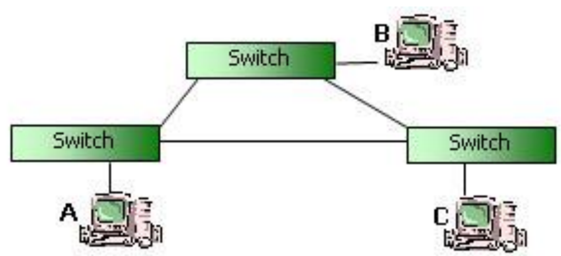


The original 1976 slide to explain Ethernet

Robert Metcalfe

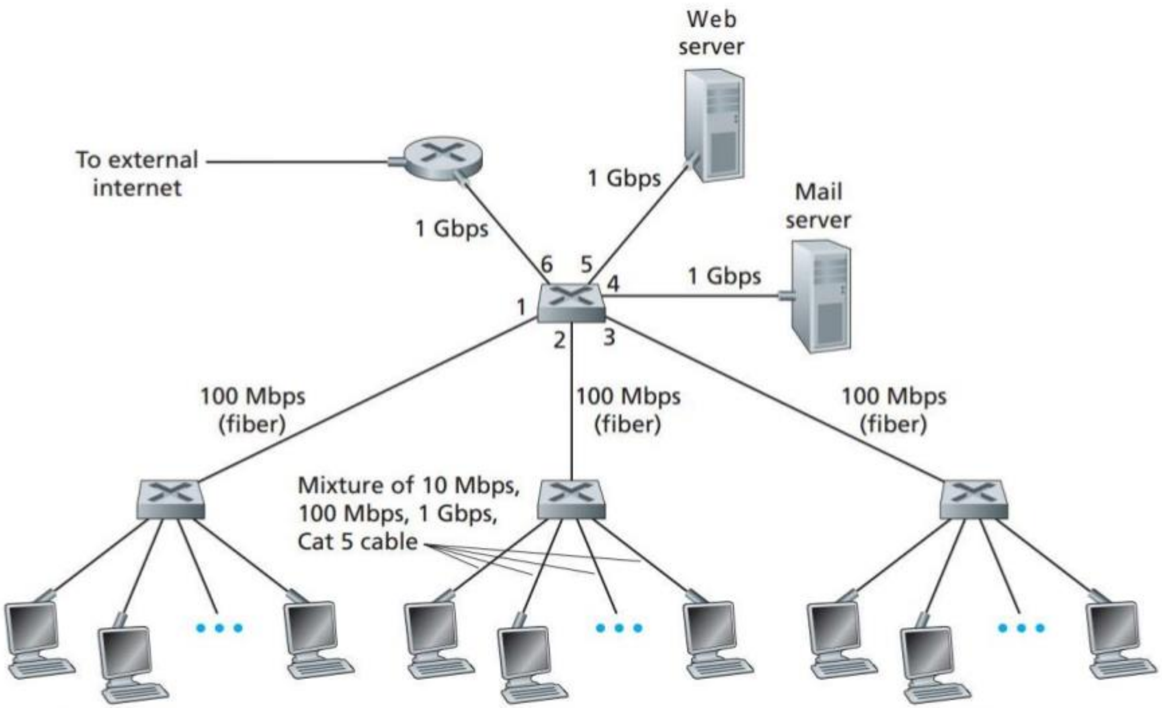
David Boggs

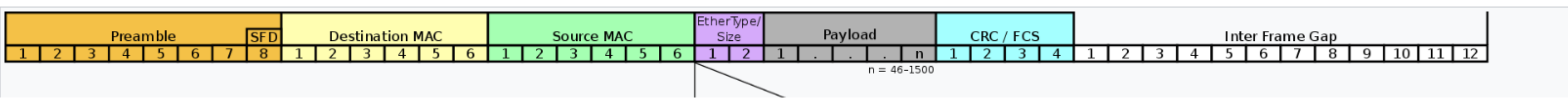




source: https://en.wikipedia.org/wiki/Ethernet

Switched Ethernet and Virtual LAN (IEEE 802.1Q)





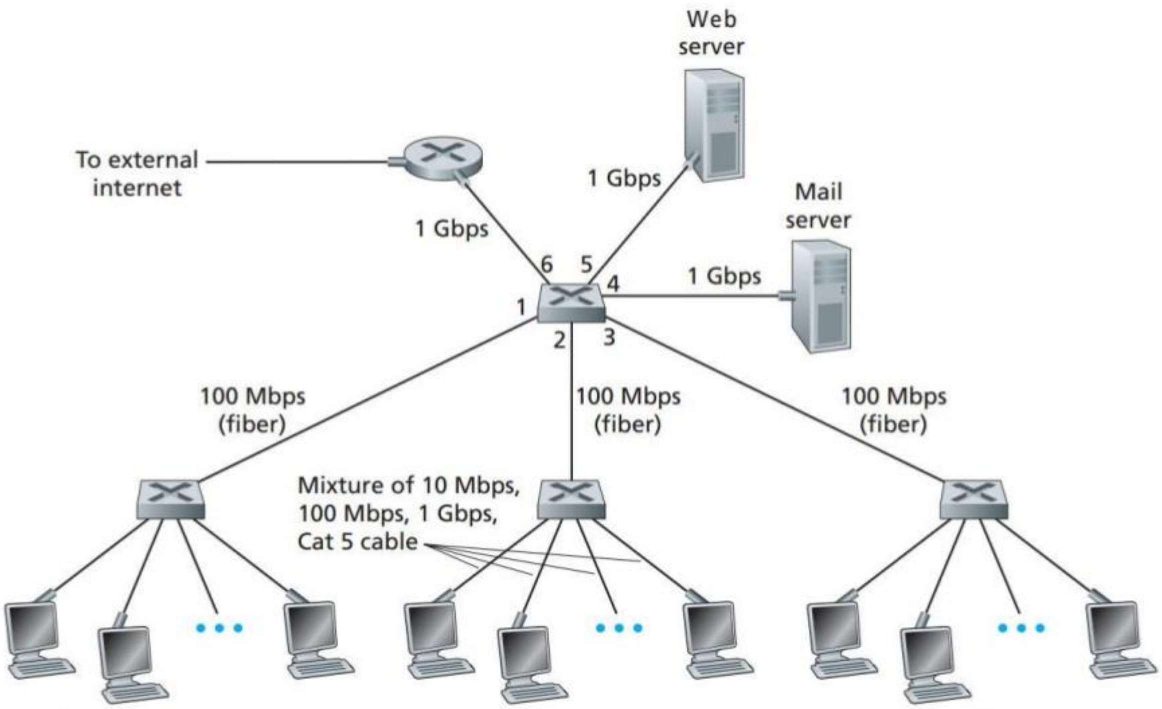
source: https://en.wikipedia.org/wiki/Ethernet

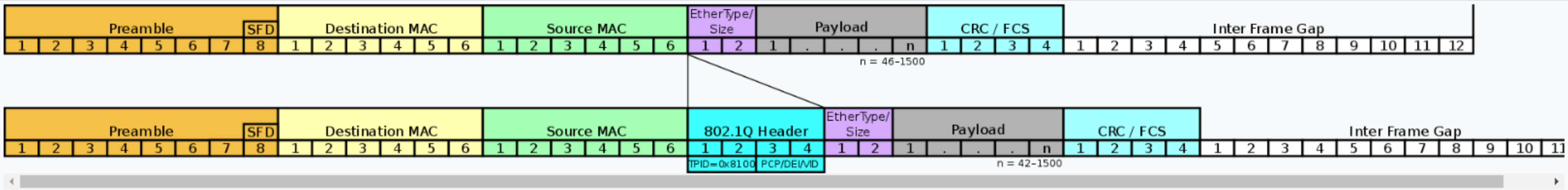
What's the difference: switch vs router?

OSI Model			
Layer		Protocol data unit (PDU)	Function[3]
Host layers	7. Application	Data	High-level APIs, including resource sharing, remote file access
	6. Presentation		Translation of data between a networking service and an application; including character encoding, data compression and encryption/decryption
	5. Session		Managing communication sessions, i.e. continuous exchange of information in the form of multiple back-and-forth transmissions between two nodes
	4. Transport	Segment, Datagram	Reliable transmission of data segments between points on a network, including segmentation, acknowledgement and multiplexing
Media layers	3. Network	Packet	Structuring and managing a multi-node network, including addressing, routing and traffic control
	2. Data link	Frame	Reliable transmission of data frames between two nodes connected by a physical layer
	1. Physical	Symbol	Transmission and reception of raw bit streams over a physical medium

source: https://en.wikipedia.org/wiki/Ethernet

Switched Ethernet and Virtual LAN (IEEE 802.1Q)





Insertion of 802.1Q tag in an Ethernet frame

source: https://en.wikipedia.org/wiki/Ethernet

Ethernet TSN *(not a standard, but a suite of standards)*

Time Synchronization

IEEE 802.1AS: Time Synchronization
P802.1AS-Rev: Time Synchronization Redundancy

Latency

IEEE 802.1Qav: Credit Based Shaper
IEEE 802.1Qbu: Frame Preemption
IEEE 802.1Qbv: Scheduled Traffic – Time Aware Shaper
IEEE 802.1Qch: Cyclic Queuing and Forwarding
P802.1Qcr: Asynchronous Traffic Shaper

Reliability

IEEE 802.1Qca: Path Control
IEEE 802.1Qci: Per-Stream Filtering and Policing
IEEE 802.1CB: Frame Replication and Elimination
+ P802.1CBdb: Extended Stream Identification
P802.1AS-Rev: Time Synchronization Redundancy

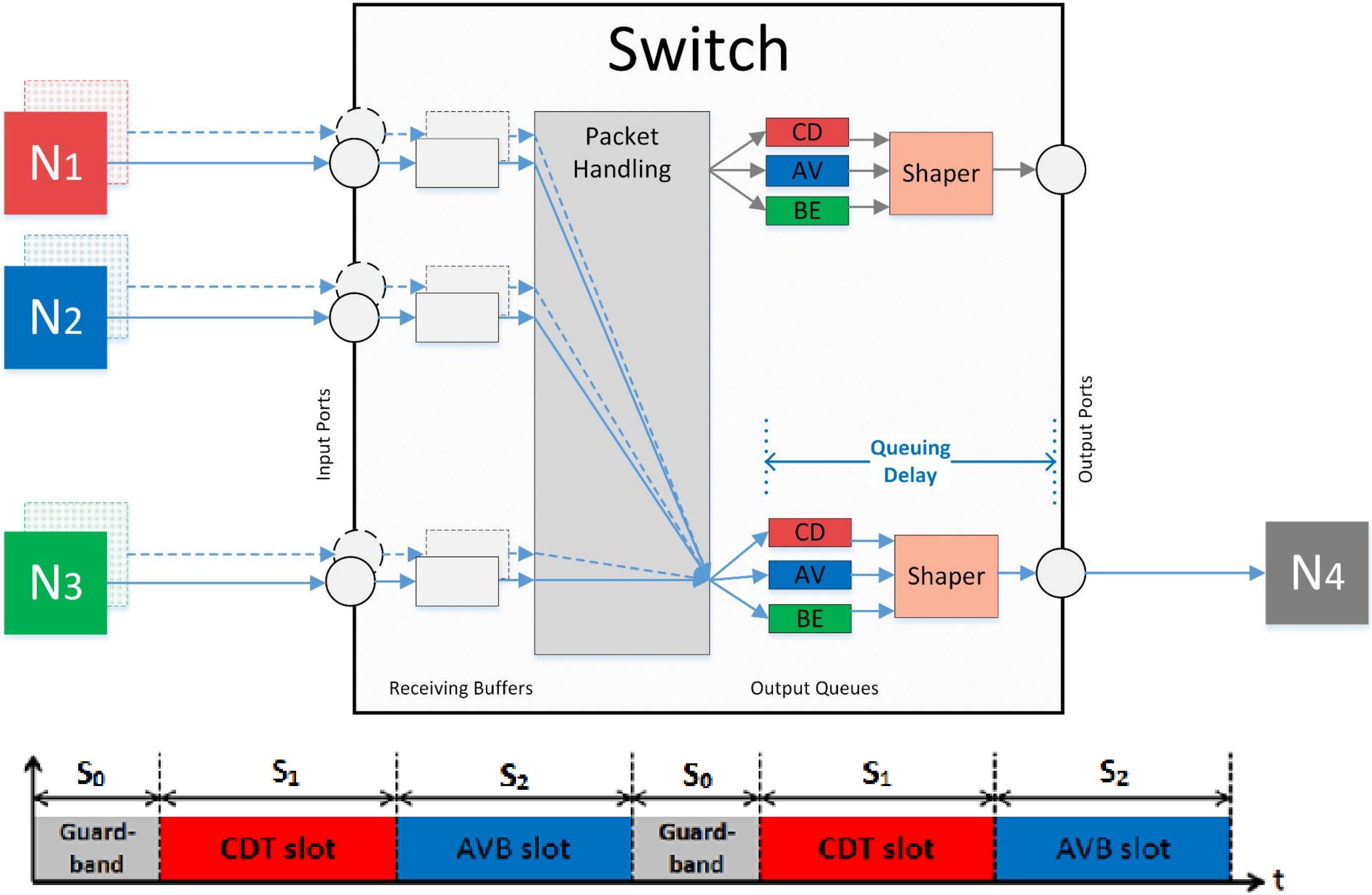
Management

IEEE 802.1Qat: Stream Reservation Protocol
P802.1Qcc: TSN Configuration
P802.1Qcp: YANG Data Model
+ P802.1Qcw: YANG Data Models for Qbv, Qbu, Qci
+ P802.1ABcu: LLDP YANG Data Model
+ P802.1CBcv: YANG Data Model for CB
P802.1CS: Link-local Reservation Protocol

source: https://www.tttech.com/

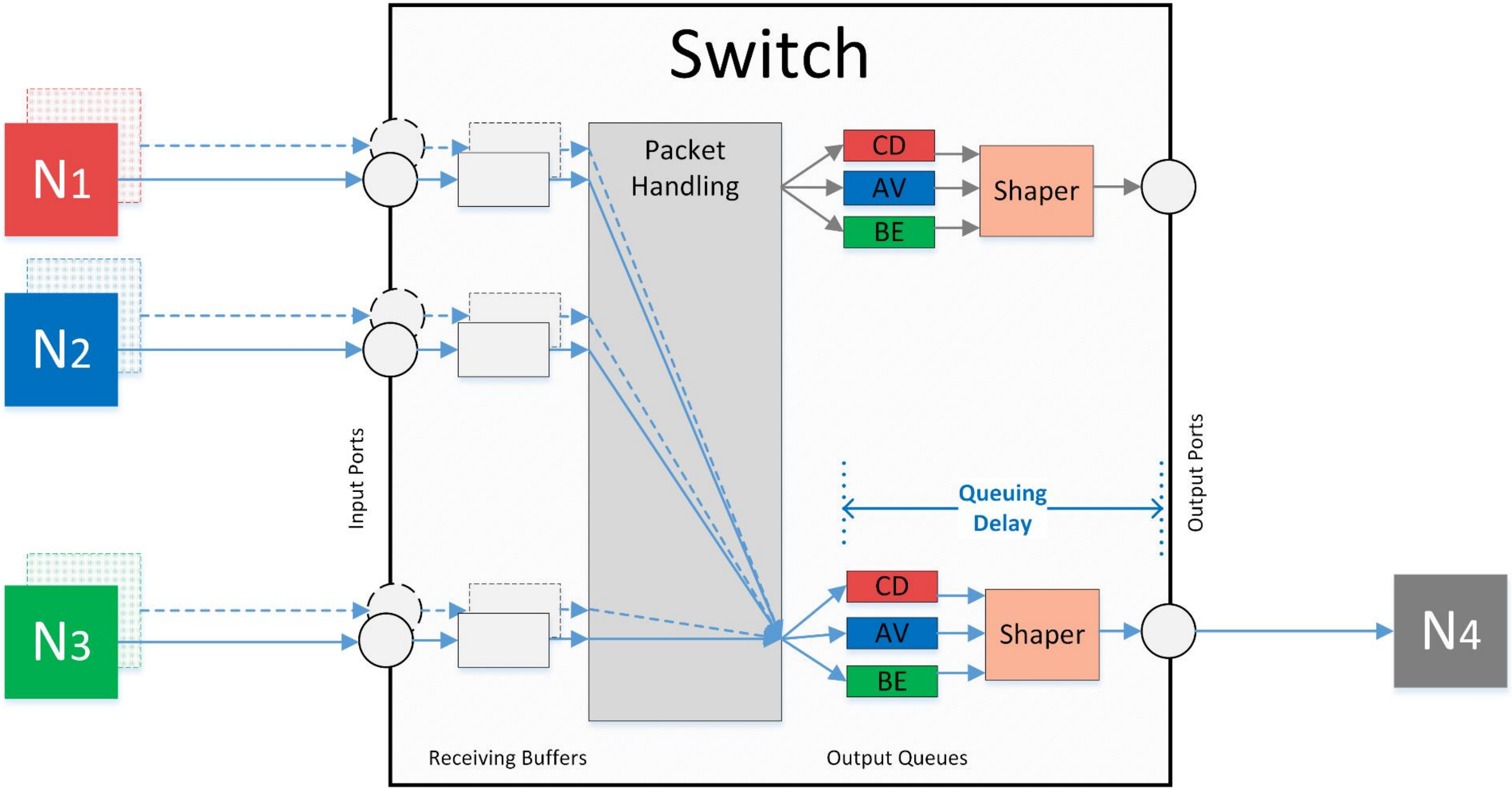
source: http://www.ieee802.org/1/pages/tsn.html

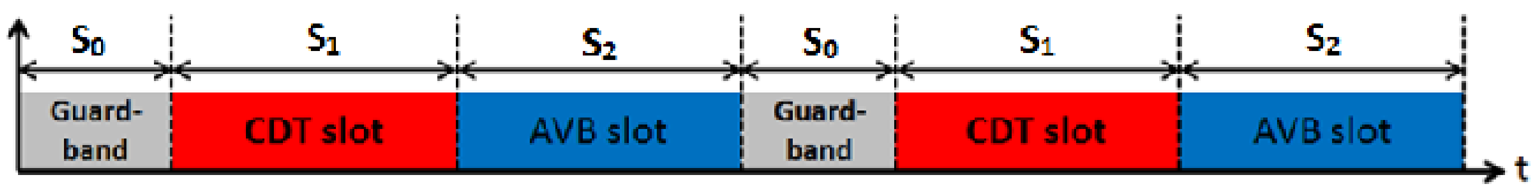
Timing and traffic shaping



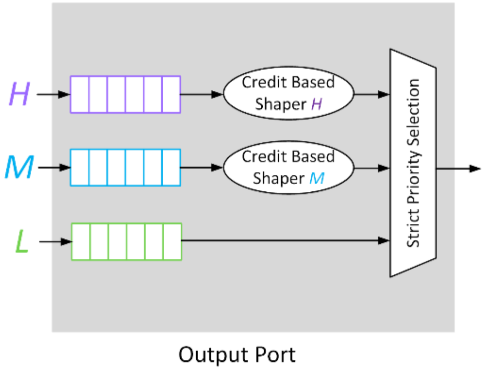
BREAK

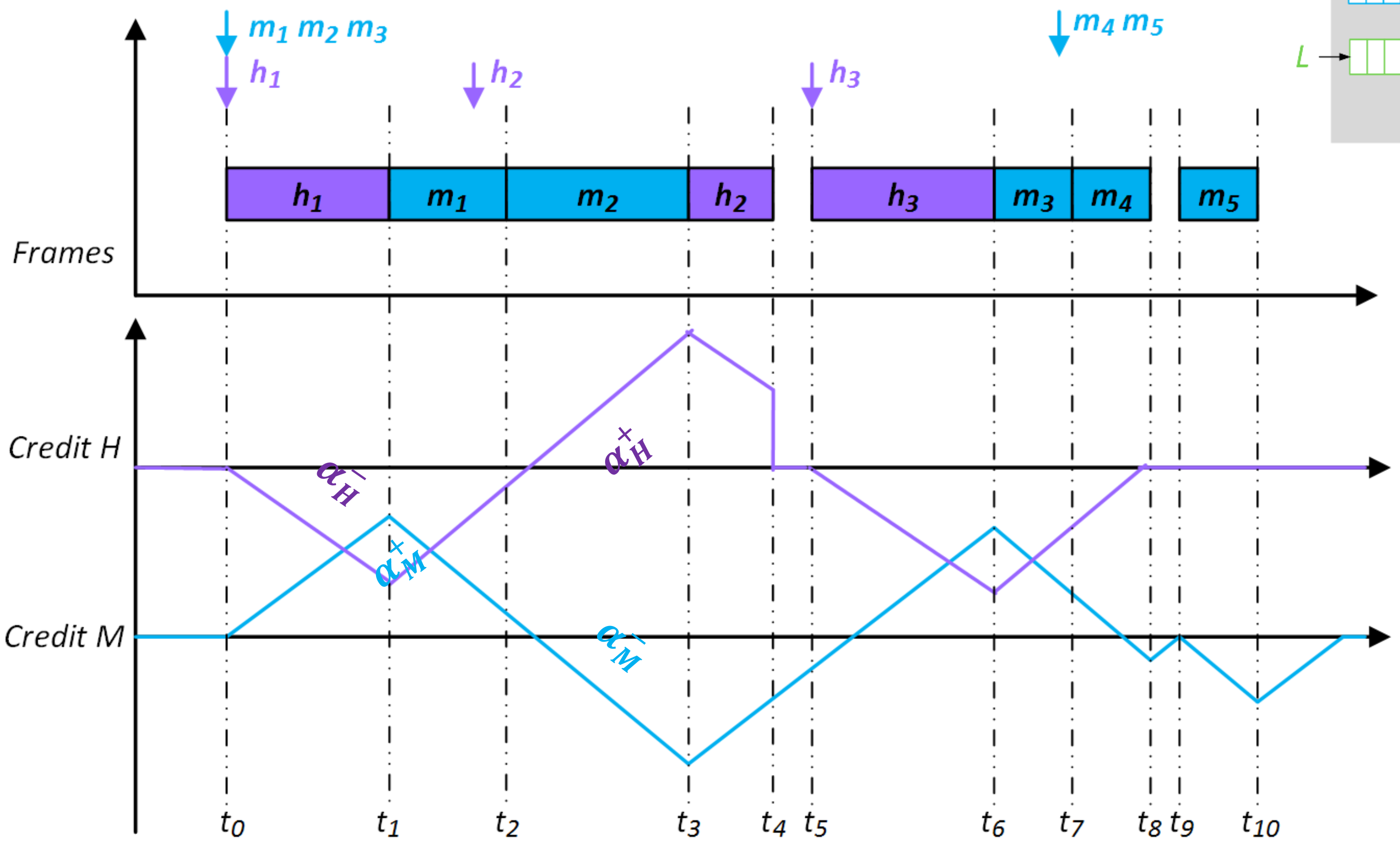
Credit-based shaping in Ethernet TSN



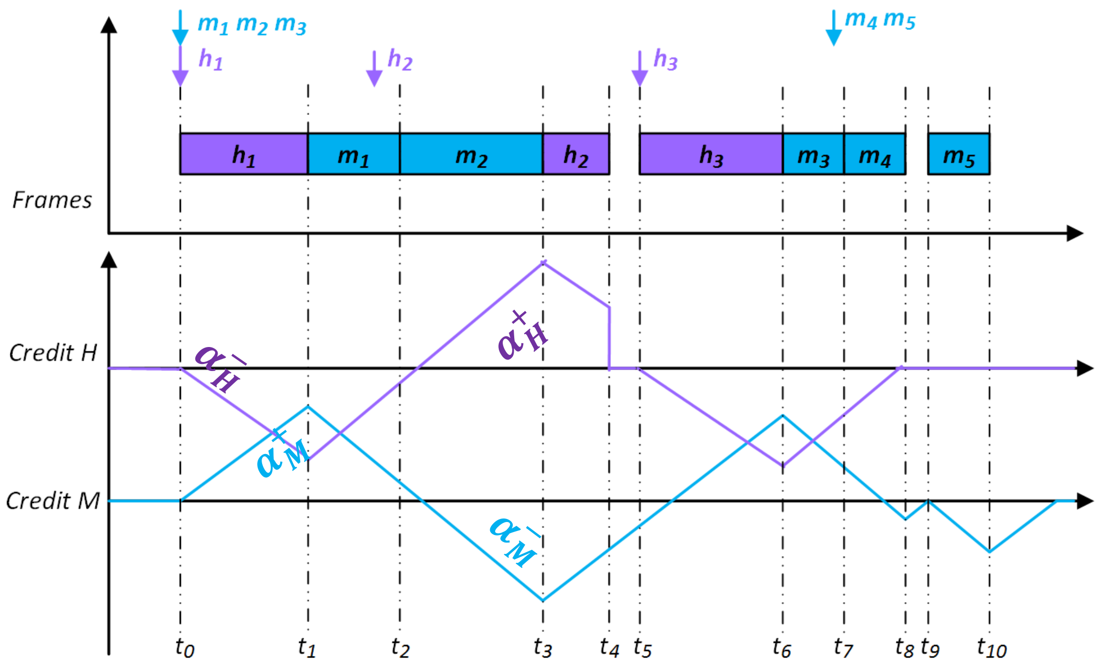


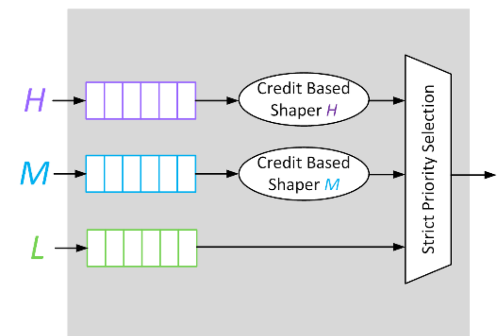
Credit-based shaping in Ethernet TSN





Exercise : what is the wcrt of a single class M frame?

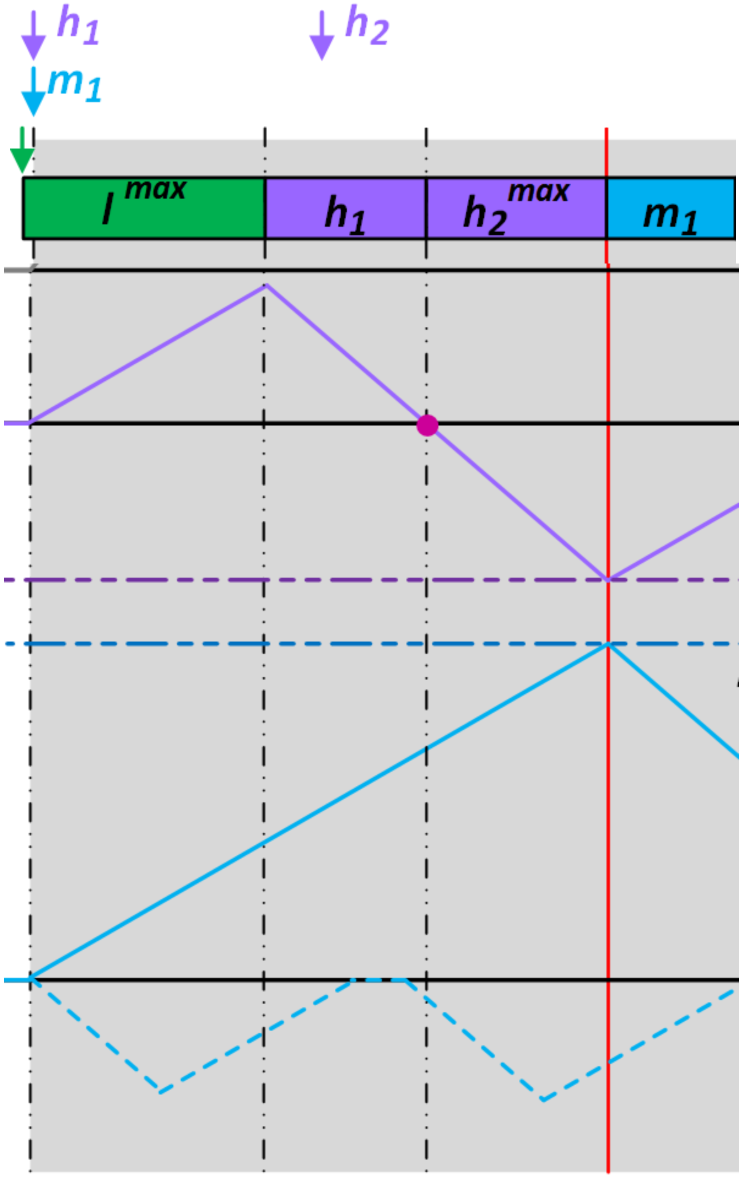




Output Port

- Transmit when credit is non-negative and higher-priority transmissions are not possible
- All transmissions are non-preemptive
- Loose credit at rate α_M^- while transmitting
- Increase credit at rate α_M^+ when credit is negative
- Increase credit at rate α_M^+ when there are pending frames
- Reset credit to 0 when positive and no pending frames

Solution… but how do we extend and prove this?



$$C_L^{max} + C_L^{max} \cdot \frac{\alpha_H^{+}}{\alpha_H^{-}} + C_H^{max}$$

A formal model of network traffic

Execution:

$$u : T \rightarrow \wp(F) \times \wp(F)$$

The cumulative set of frames that arrived at the switch at time t:

$$u_a(t) \subseteq F$$

The cumulative set of frames that finished transmission at t:

$$u_f(t) \subseteq F$$

A formal model of network traffic

The *arrival time of a frame x in execution u:*

$$a^u(x) = \min\{t \in T \mid x \in u_1(t)\}$$

The *finish time of a frame x in execution u:*

$$f^u(x) = \min\{t \in T \mid x \in u_2(t)\}$$

A formal model of network traffic

The *worst-case response-time*

$$WCRT(X) = \sup_{u \in U} \sup_{x \in X} |f^u(x), a^u(x)|$$

Giving formal semantics to behavior

Definition:

The *behavior* of a credit-based shaper is characterized by

a set U of executions $u : T \rightarrow \wp(F) \times \wp(F)$ such that

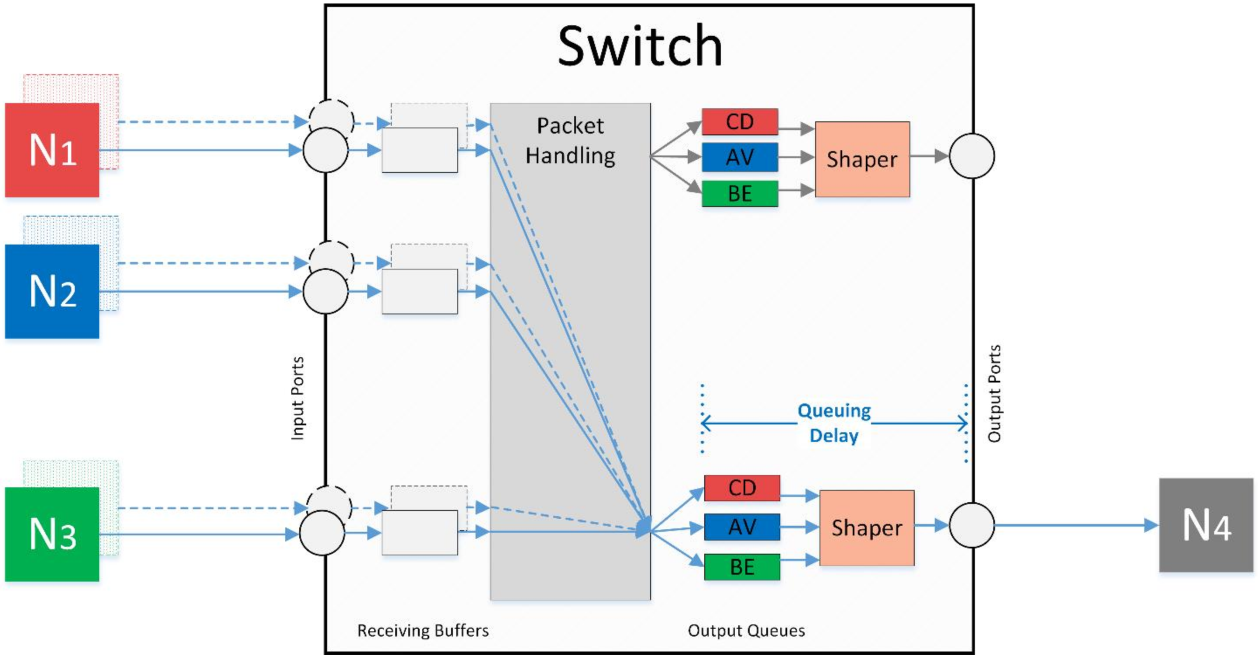
for every $u \in U$ we find … *(fill in a set of axioms here) .*

Theorem:

Given a behavior U of a credit-based shaper we find

$WCRT(X) = \sup_{u \in U} \sup_{x \in X} |f^u(x), a^u(x)| \leq \cdots (fill\ in\ result\ of\ analysis\ here)$.

A formal model of credit-based shaping



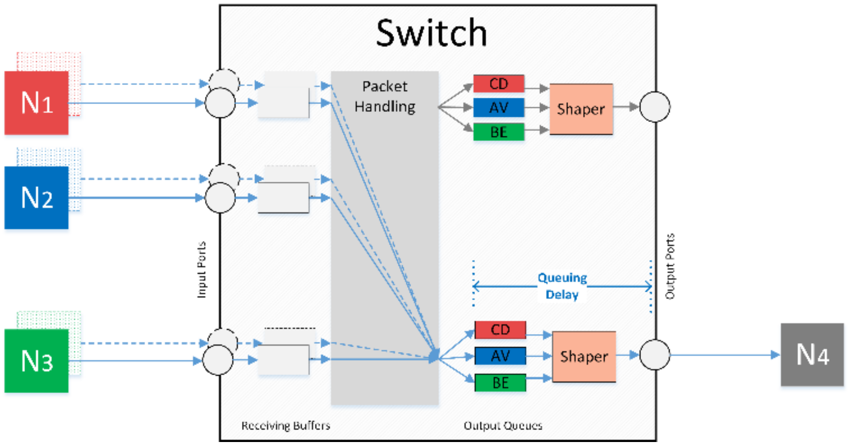
The class, ~~source~~, ~~destination~~, size, and arrival time of frames and interfering frames together fully determine their transmission time.

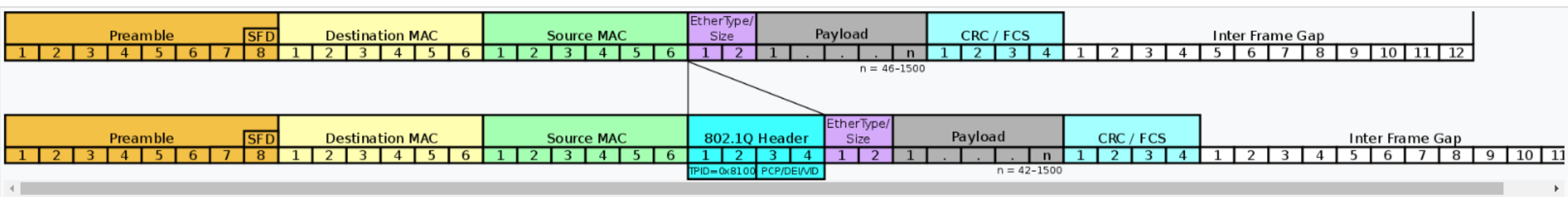
A formal model of credit-based shaping

According to the 802.1 standard, *all* relevant information on frames can be found in the header structure:

$$size : F \rightarrow \mathbb{R}$$

$$class : F \rightarrow \{CD, A, B, BE\} \text{ or } \{H, M, L\}$$





Insertion of 802.1Q tag in an Ethernet frame

A formal model of credit-based shaping

Credit of class H at time t during execution u:

$$CR_H^u : T \rightarrow \mathbb{R}$$

Credit of class M at time t during execution u:

$$CR_M^u : T \rightarrow \mathbb{R}$$

Start-time of the transmission of frame x during execution u:

$$s^u : F \rightarrow T$$

Maximum frame size of a class:

$$C_H^{max}, C_M^{max}, C_L^{max}$$

Credit increase and decrease rate (idleslope and sendslope):

$$\alpha_H^+, \alpha_M^+, \alpha_H^-, \alpha_M^-, BW$$

Axioms of credit based shaping

- NON-PREEMPTIVE TRANSMISSION:

$$f^u(x) = s^u(x) + size(x)$$

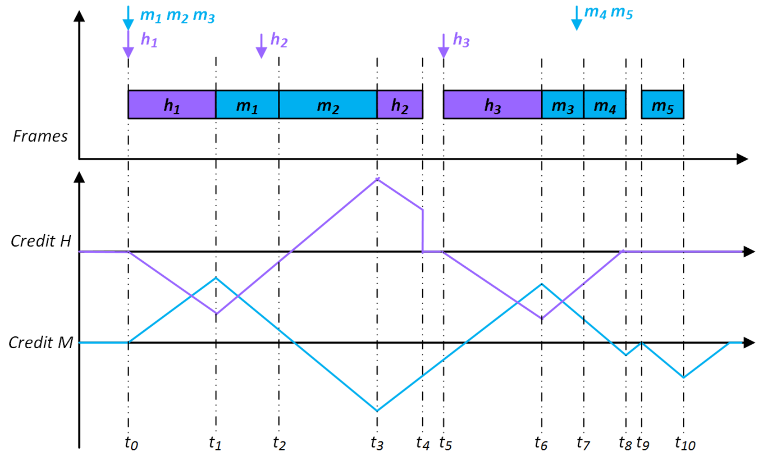
- SINGLE TRANSMISSION

$$x \neq x' \Rightarrow s^u(x) \neq s^u(x')$$

and

$$s^u(x) < s^u(x') \Rightarrow f^u(x) \leq s^u(x')$$

- FIFO (?)



Axioms of credit based shaping

- NON-PREEMPTIVE TRANSMISSION:

$$f^u(x) = s^u(x) + size(x)$$

- SINGLE TRANSMISSION

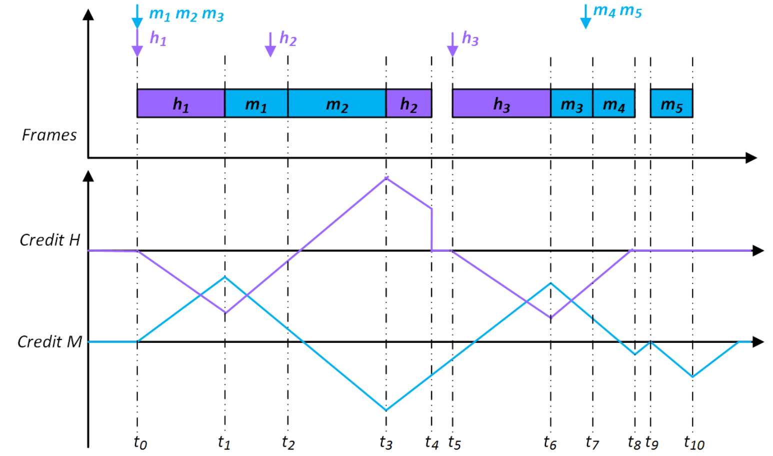
$$x \neq x' \Rightarrow s^u(x) \neq s^u(x')$$

and

$$s^u(x) < s^u(x') \Rightarrow f^u(x) \leq s^u(x')$$

- FIFO

$$a^u(x) < a^u(x') \Rightarrow s^u(x) < s^u(x')$$



Axioms of credit based shaping

- NON-PREEMPTIVE TRANSMISSION:

$$f^u(x) = s^u(x) + size(x)$$

- SINGLE TRANSMISSION

$$x \neq x' \Rightarrow s^u(x) \neq s^u(x')$$

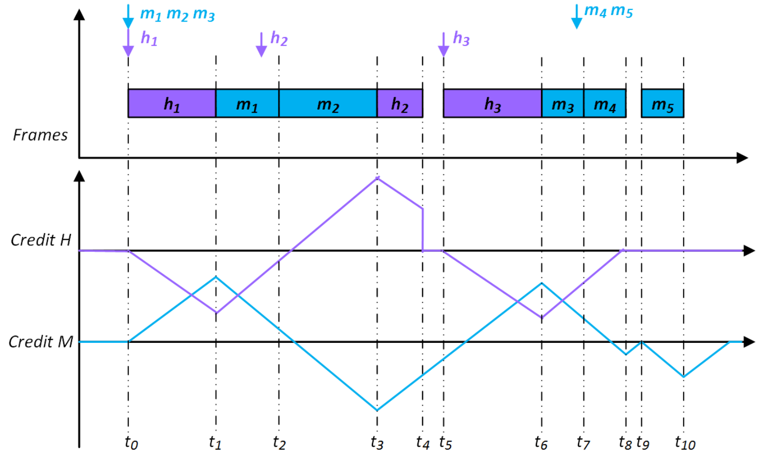
and

$$s^u(x) < s^u(x') \Rightarrow f^u(x) \leq s^u(x')$$

- FIFO

$$a^u(x) < a^u(x') \Rightarrow s^u(x) < s^u(x')$$

- CREDIT DROP



Axioms of credit based shaping

- NON-PREEMPTIVE TRANSMISSION:

$$f^u(x) = s^u(x) + size(x)$$

- SINGLE TRANSMISSION

$$x \neq x' \Rightarrow s^u(x) \neq s^u(x')$$

and

$$s^u(x) < s^u(x') \Rightarrow f^u(x) \leq s^u(x')$$

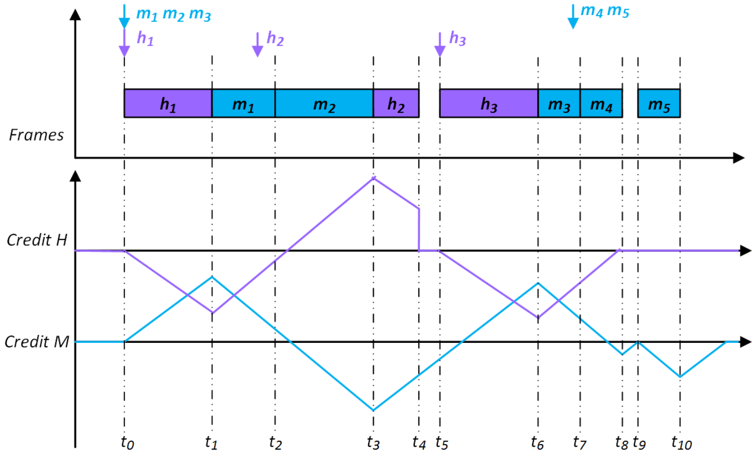
- FIFO

$$a^u(x) < a^u(x') \Rightarrow s^u(x) < s^u(x')$$

- CREDIT DROP

if $s^u(x) \leq t \leq t' \leq f^u(x)$ and $class(x) = H$

then $CR_H^u(t') = CR_H^u(t) - \alpha_H^- \cdot |t', t|$



Axioms of credit based shaping

- A class H has *pending frames* whenever a frame has arrived that has not yet been transmitted

$$Pending_H(t) \stackrel{\text{def}}{=} \exists_{x:class(x)=H}\ a^u(x) \leq t\ \wedge\ f^u(x) > t$$

- Credit recovery:
  if $t \leq t'$ and $|t, t'| \leq -CR_H^u(t)$
  then $CR_H^u(t') = CR_H^u(t) + \alpha_H^+ \cdot |t', t|$

- Gaining credit:
  if $class(x) = M$ and $class(y) \neq M$ and
  if $a^u(x) \leq t\ \wedge\ s^u(x) > t'$ and $s^u(y) \leq t\ \wedge\ f^u(y) > t'$
  then $CR_M^u(t') = CR_M^u(t) + \alpha_H^+ \cdot |t', t|$

- Credit reset:

$$\neg Pending_H(t)\ \Rightarrow\ CR_M^u(t) \leq 0$$

(this one is true, but actually a bit more complicated)

Axioms of credit based shaping

- A frame may only start transmission if there is no higher priority frame pending for which credit is available…

Axioms of credit based shaping

- A frame may only start transmission if there is no higher priority frame pending for which credit is available…

  if $class(x) < class(y)$ *and* $a^u(y) \leq s^u(x)$ *and* $CR^u_{class(y)}\left(s^u(x)\right)$
  then $s^u(y) \leq s^u(x)$

Axioms of credit based shaping

- A frame may only start transmission if there is no higher priority frame pending for which credit is available…

  if $class(x) < class(y)$ *and* $a^u(y) \leq s^u(x)$ *and* $CR^u_{class(y)}\left(s^u(x)\right)$
  then $s^u(y) \leq s^u(x)$

- A frame will start as soon as it is allowed to, given all previous rules…

Axioms of credit based shaping

- A frame may only start transmission if there is no higher priority frame pending for which credit is available…

  if $class(x) < class(y)$ *and* $a^u(y) \leq s^u(x)$ *and* $CR^u_{class(y)}\big(s^u(x)\big)$
  then $s^u(y) \leq s^u(x)$

- A frame will start as soon as it is allowed to, given all previous rules…

$$s^u(x) = \inf\left\{ t \;\middle|\; \begin{array}{c} \forall_z\, s^u(z) < t \Rightarrow f^u(z) \leq t \\ CR^u_{class(x)}(t) \geq 0 \\ \forall_{x':class(x)=class(x')}\, a^u(x') < a^u(x) \Rightarrow s^u(x') < t \\ \forall_{y:class(y)>class(x)}\big(CR^u_{class(y)}(t) \geq 0 \wedge a^u(y) \leq t\big) \Rightarrow \big(f^u(y) \leq t\big) \end{array} \right\}$$

Giving formal semantics to behavior

Definition:

The *behavior* of a credit-based shaper is characterized by

a set U of executions $u : T \rightarrow \wp(F) \times \wp(F)$ such that

for every $u \in U$ we find … *(fill in a set of axioms here) .*

Theorem:

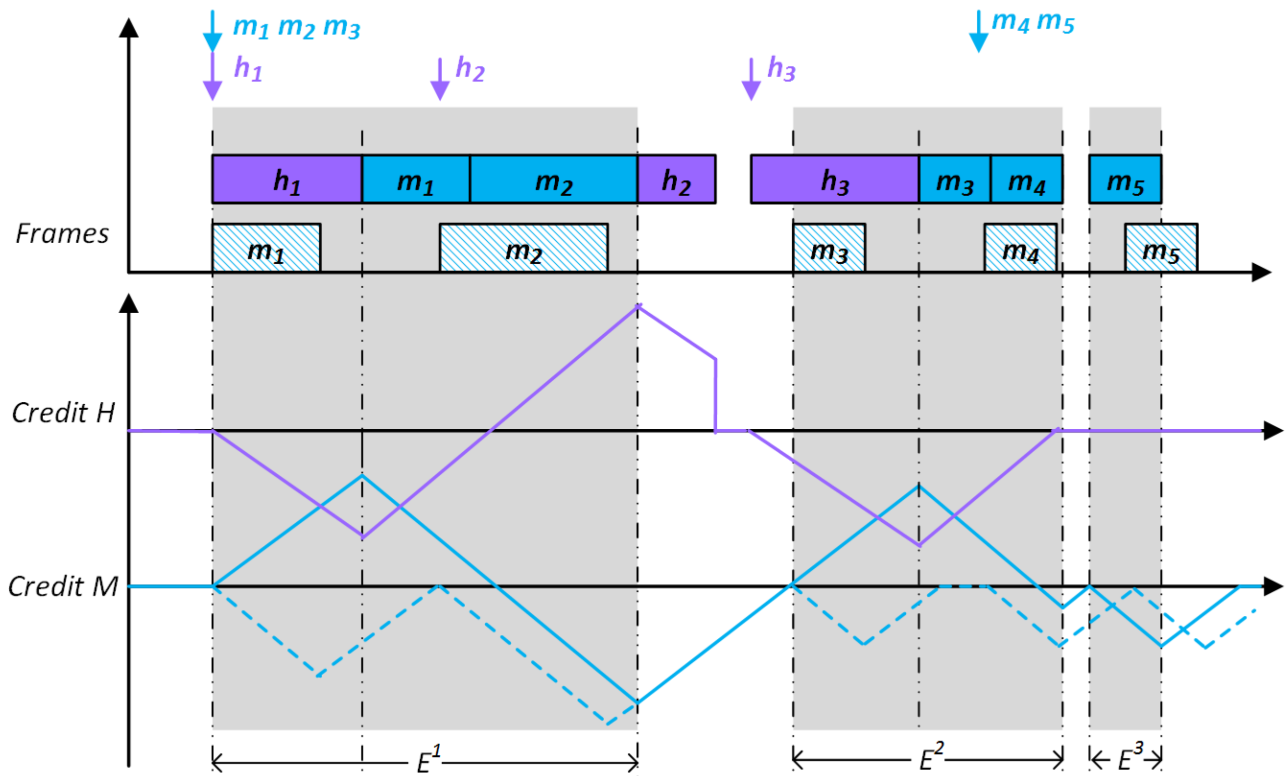
Given a behavior U of a credit-based shaper we find

$$WCRT(X) = \sup_{u \in U} \sup_{x \in X} |f^u(x), a^u(x)| \leq \cdots (fill\ in\ result\ of\ analysis\ here).$$

Eligible interval analysis

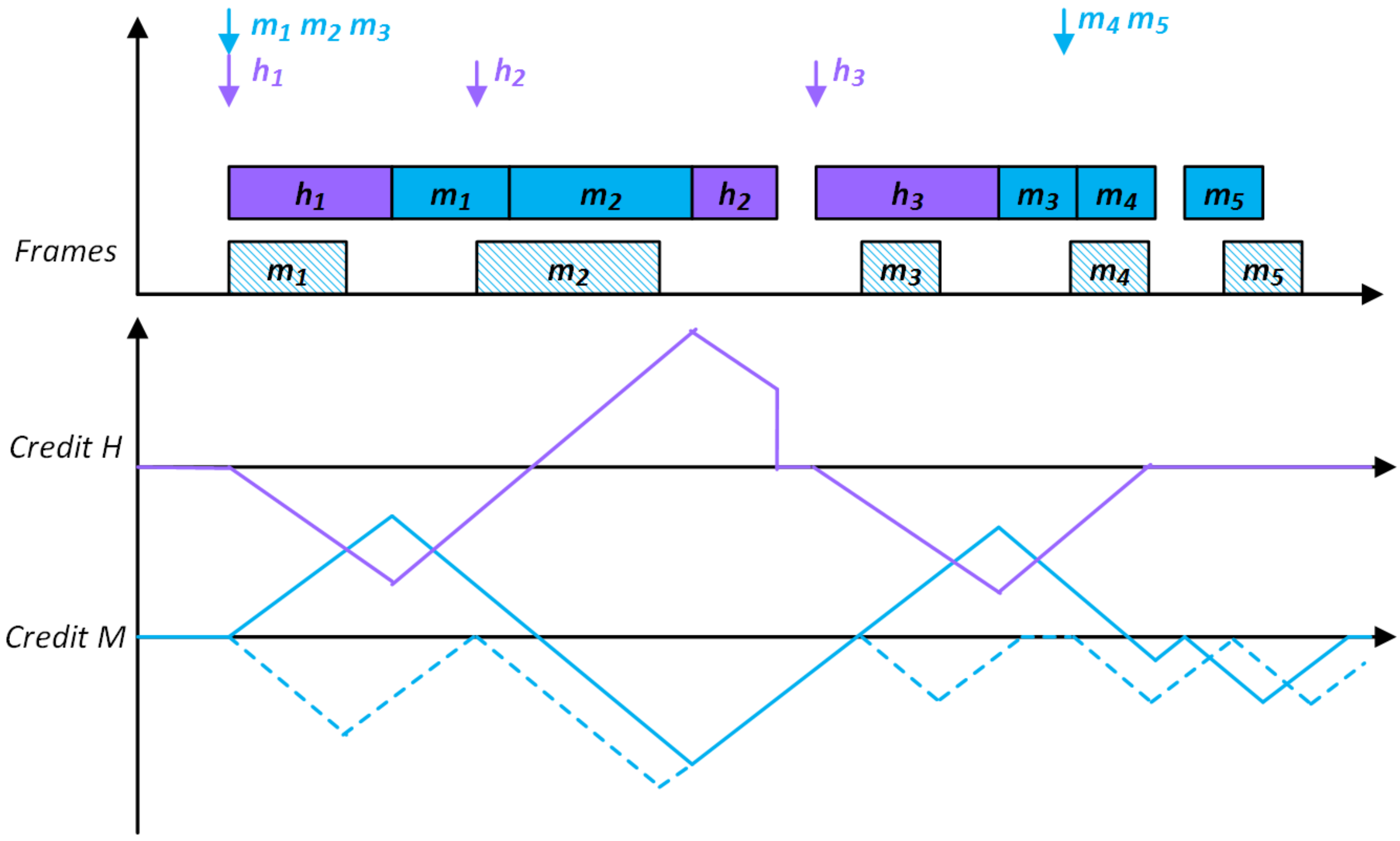
- An eligible interval $E \subseteq T$ for class X is an interval of time during which frames of X are eligible for transmission; i.e. both pending load and credit available, or an actual transmission is in progress.

$$\forall_{t \in E}(Pending_X(t) \wedge CR_X^u(t) \geq 0) \vee Transmitting_X(t)$$

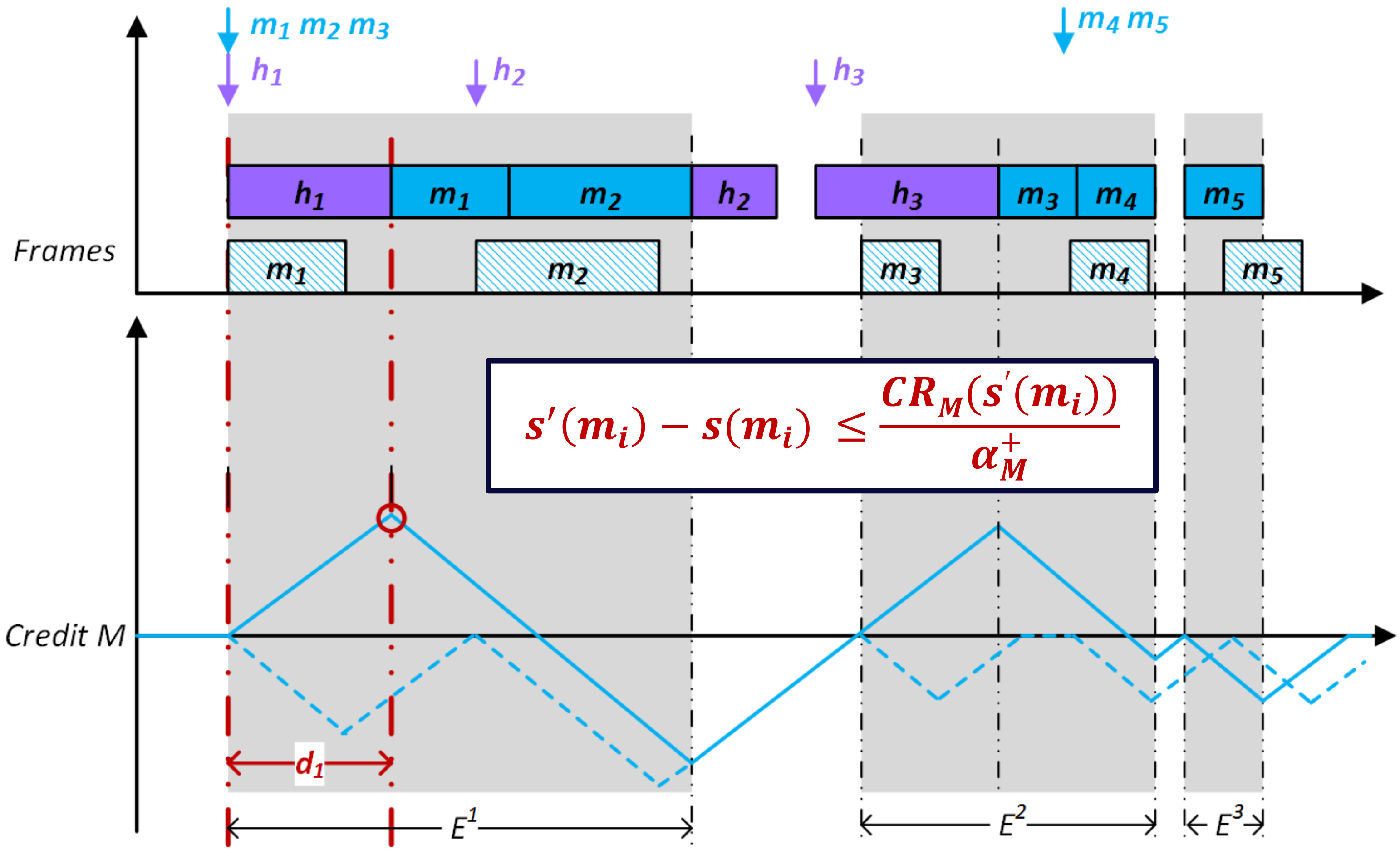


TU/e Technische Universiteit Eindhoven University of Technology

Eligible intervals & relative wcrt-analysis



Eligible intervals & relative wcrt-analysis



$$s'(m_i) - s(m_i) \leq \frac{CR_M(s'(m_i))}{\alpha_M^+}$$

TU/e Technische Universiteit Eindhoven University of Technology

A bound on the minimum and maximum credit

- **What is the minimum credit that can be achieved?**

A bound on the minimum and maximum credit

- **What is the minimum credit that can be achieved?**

$$CR_M^u(t) \geq -\alpha_M^- \cdot C_M^{max}$$

A bound on the minimum and maximum credit

- **What is the minimum credit that can be achieved?**

$$CR_M^u(t) \geq -\alpha_M^- \cdot C_M^{max}$$

- **What is the maximum credit for M without pending load?**

A bound on the minimum and maximum credit

- **What is the minimum credit that can be achieved?**

$$CR_M^u(t) \geq -\alpha_M^- \cdot C_M^{max}$$

- **What is the maximum credit for M without pending load?**

$$CR_M^u(t) \leq 0$$

A bound on the minimum and maximum credit

- **What is the minimum credit that can be achieved?**

$$CR_M^u(t) \geq -\alpha_M^- \cdot C_M^{max}$$

- **What is the maximum credit for M without pending load?**

$$CR_M^u(t) \leq 0$$

- **What is the maximum credit for M when L is transmitting?**

A bound on the minimum and maximum credit

- What is the minimum credit that can be achieved?

$$CR_M^u(t) \geq -\alpha_M^- \cdot C_M^{max}$$

- What is the maximum credit for M without pending load?

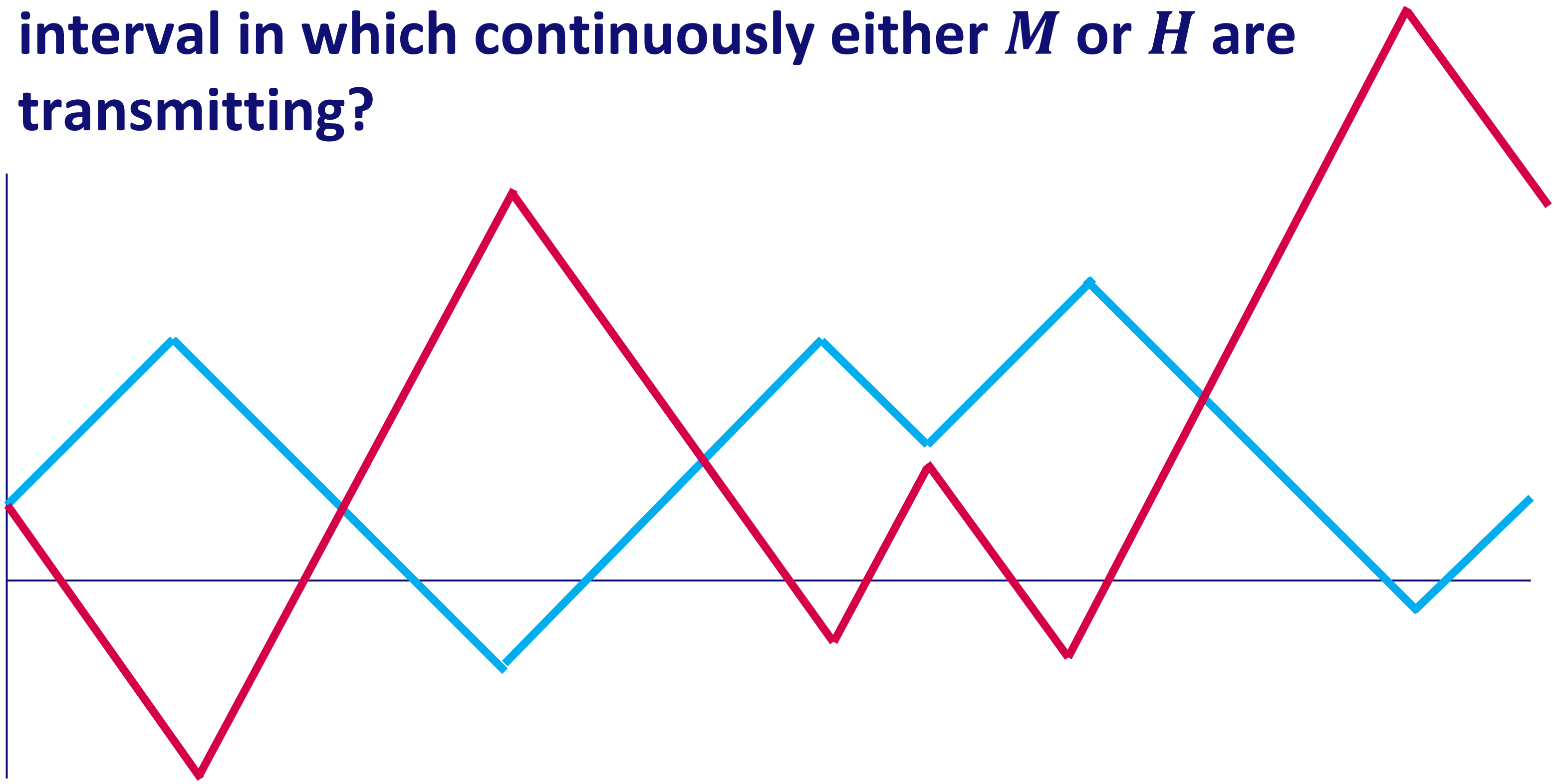
$$CR_M^u(t) \leq 0$$

- What is the maximum credit for M when L is transmitting?

$$CR_M^u(t) \leq \alpha_M^+ \cdot C_L^{max}$$

A bound on the minimum and maximum credit

- **What is the maximum increase in credit of M during an interval in which continuously either M or H are transmitting?**



A bound on the minimum and maximum credit

- **What is the maximum increasing in credit of M during an interval in which continuously either M or H are transmitting?**

  Let $I = (t, t')$ be a length of time during which either M or H are continuously transmitting (alternatingly).

  Let T_M, T_H denote the total time that M and H are transmitting, respectively. So $|t, t'| = T_M + T_H$.

  Then:

$$CR_H^u(t') = CR_H^u(t) + \alpha_H^+ \cdot T_M - \alpha_H^- \cdot T_H$$
$$CR_M^u(t') = CR_M^u(t) + \alpha_M^+ \cdot T_H - \alpha_M^- \cdot T_M$$

A bound on the minimum and maximum credit

- **What is the maximum increasing in credit of M during an interval in which continuously either M or H are transmitting?**

So:

$$\Delta CR_H^u + \Delta CR_M^u = (\alpha_M^+ - \alpha_H^-) \cdot T_H + (\alpha_H^+ - \alpha_M^-) \cdot T_M$$

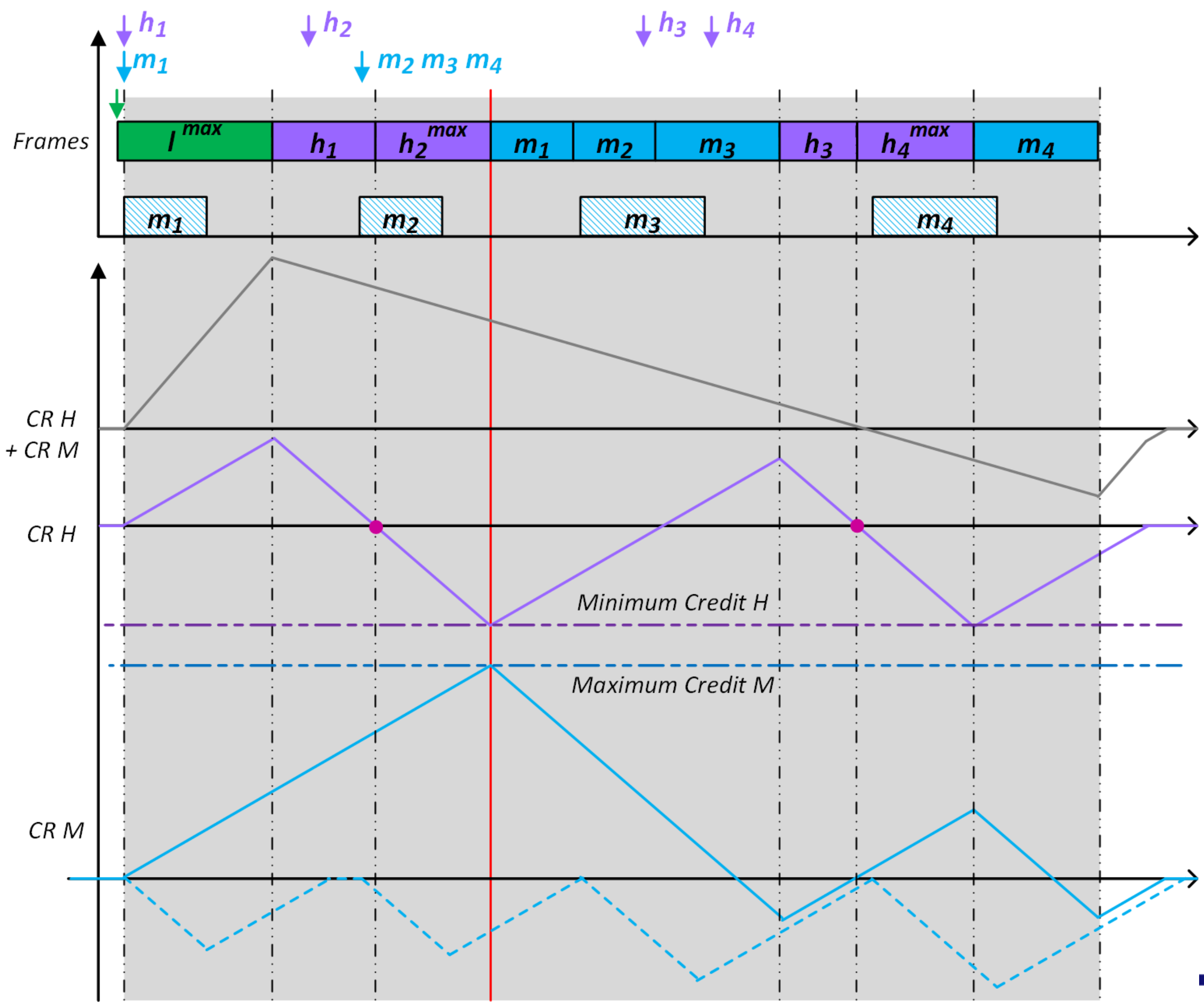
And using $\alpha_M^- = BW - \alpha_M^+$ and $\alpha_H^- = BW - \alpha_H^+$

$$\Delta CR_H^u + \Delta CR_M^u = (\alpha_M^+ + \alpha_H^+ - BW) \cdot (T_H + T_M)$$

The *total* credit drops as long as the total reservation is less than the bandwidth. So the rise in one credit is at most proportional to the drop in the other.

Eligible interval analysis

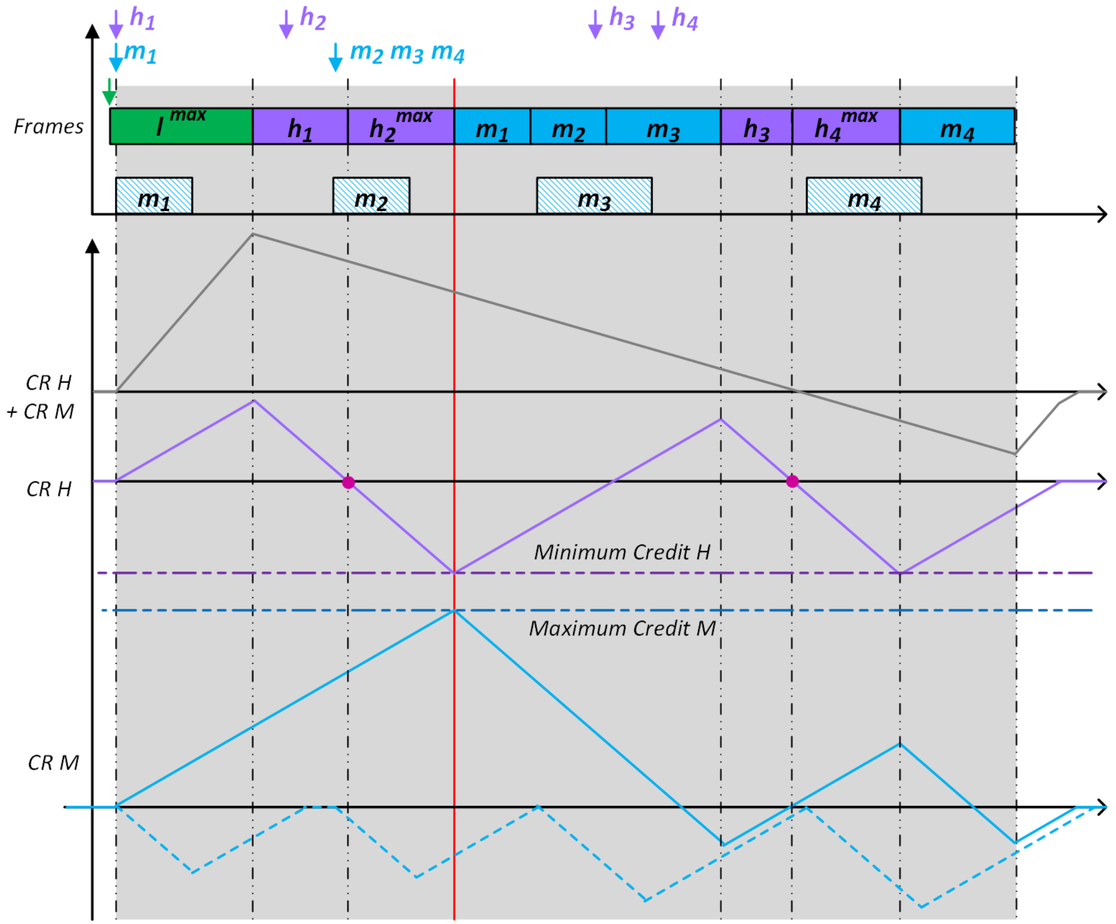
- **Maximum relative delay:** $C_L^{max} + C_L^{max} \cdot \frac{\alpha_H^+}{\alpha_H^-} + C_H^{max}$



Eligible interval analysis

- **Worst-case response time:**

$$WR_{int} = WRino_{int} + C_L^{max} + C_L^{max} \cdot \frac{\alpha_H^+}{\alpha_H^-} + C_H^{max}$$

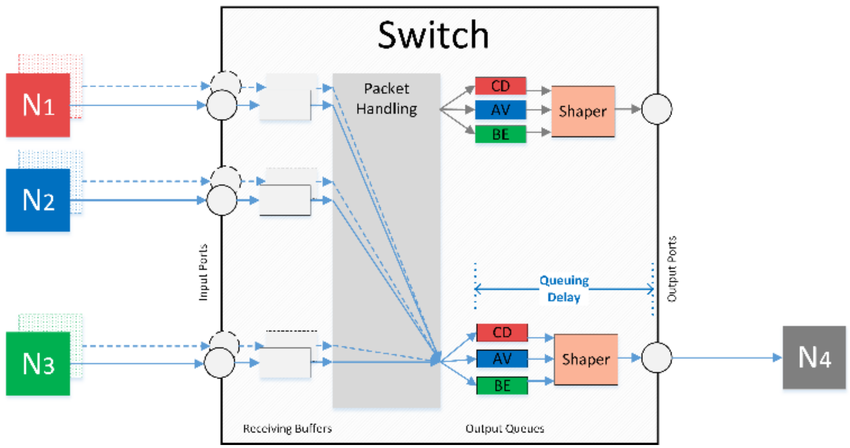


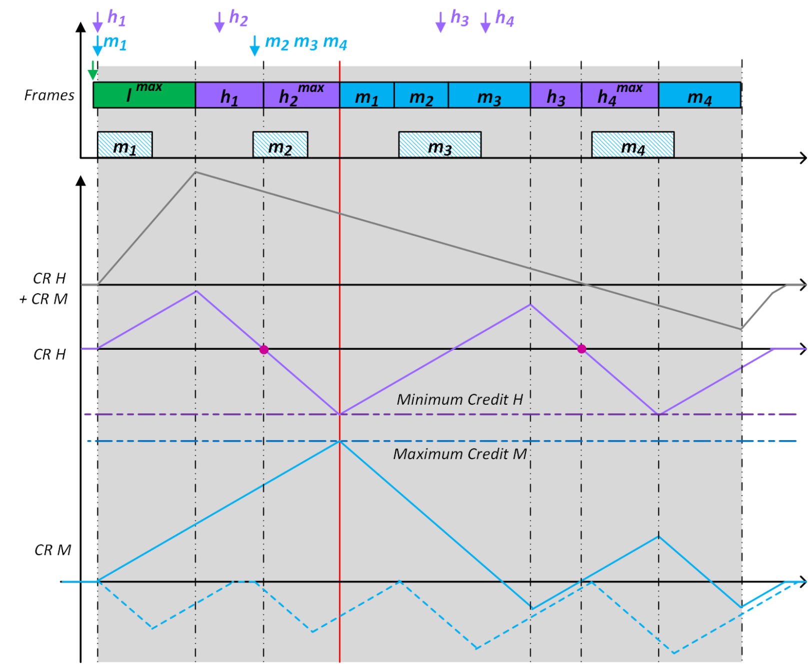
TU/e Technische Universiteit Eindhoven University of Technology

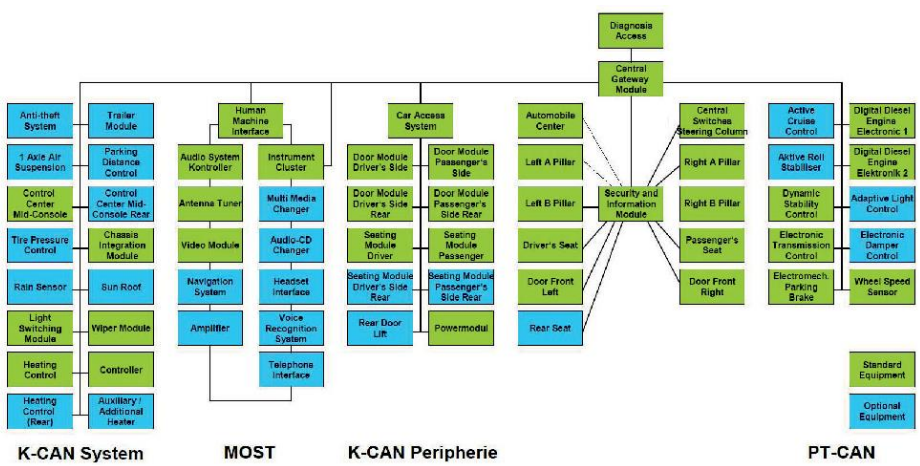
BREAK

Ethernet TSN
Worst-case response time analysis of credit-based shaping

$$WR_{int} = WRino_{\,int} + C_L^{max} + C_L^{max} \cdot \frac{\alpha_H^{+}}{\alpha_H^{-}} + C_H^{max}$$







Analysis based on `just the standard'

Time Synchronization

IEEE 802.1AS: Time Synchronization
P802.1AS-Rev: Time Synchronization Redundancy

Latency

IEEE 802.1Qav: Credit Based Shaper
IEEE 802.1Qbu: Frame Preemption
IEEE 802.1Qbv: Scheduled Traffic – Time Aware Shaper
IEEE 802.1Qch: Cyclic Queuing and Forwarding
P802.1Qcr: Asynchronous Traffic Shaper

Reliability

IEEE 802.1Qca: Path Control
IEEE 802.1Qci: Per-Stream Filtering and Policing
IEEE 802.1CB: Frame Replication and Elimination
+ P802.1CBdb: Extended Stream Identification
P802.1AS-Rev: Time Synchronization Redundancy

Management

IEEE 802.1Qat: Stream Reservation Protocol
P802.1Qcc: TSN Configuration
P802.1Qcp: YANG Data Model
+ P802.1Qcw: YANG Data Models for Qbv, Qbu, Qci
+ P802.1ABcu: LLDP YANG Data Model
+ P802.1CBcv: YANG Data Model for CB
P802.1CS: Link-local Reservation Protocol

Axioms that capture engineering principles

$$f^u(x) = s^u(x) + size(x)$$

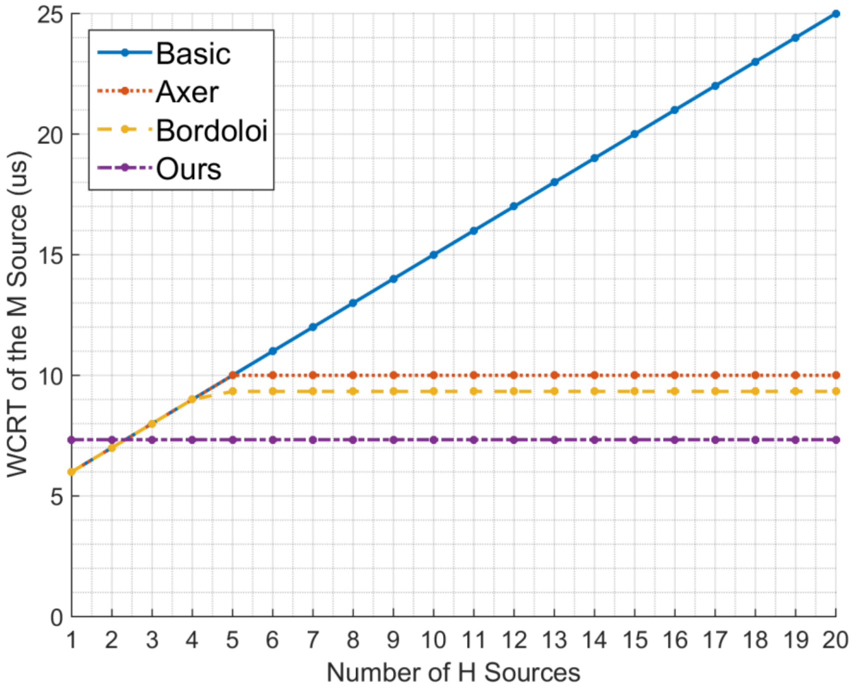
$$x \neq x' \Rightarrow s^u(x) \neq s^u(x')$$

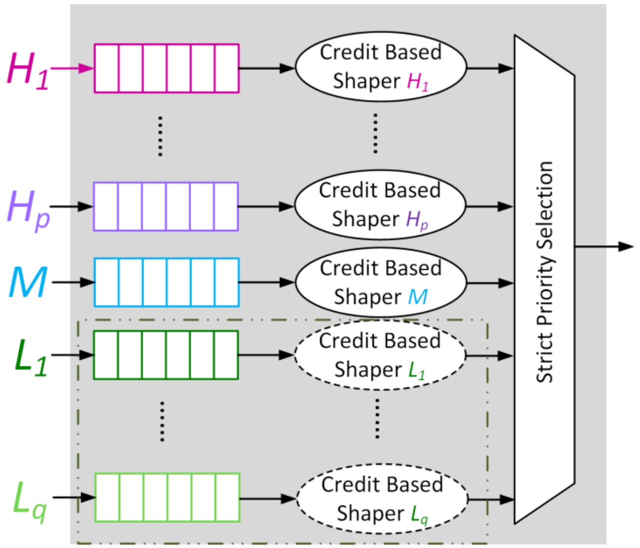
$$s^u(x) < s^u(x') \Rightarrow f^u(x) \leq s^u(x')$$

$$a^u(x) < a^u(x') \Rightarrow s^u(x) < s^u(x')$$

$$s^u(x) = \inf \left\{ t \;\middle|\; \begin{array}{c} \forall_z\, s^u(z) < t \Rightarrow f^u(z) \leq t \\ CR^u_{class(x)}(t) \geq 0 \\ \forall_{x':class(x)=class(x')}\, a^u(x') < a^u(x) \Rightarrow s^u(x') < t \\ \forall_{y:class(y)>class(x)} \left(CR^u_{class(y)}(t) \geq 0 \wedge a^u(y) \leq t\right) \Rightarrow \left(f^u(y) \leq t\right) \end{array} \right\}$$

Improvement over traditional busy period analysis





Eligible interval analysis is independent and tight, and improves busy period analysis for idling servers

We analyze the platform, not the final product!

Technische Universiteit
Eindhoven
University of Technology

Validation

Prof. Edward Lee, Berkeley

Validation

Prof. Edward Lee, Berkeley

Physicists build their models
to fit reality as closely as possible…

Validation

Prof. Edward Lee, Berkeley

Physicists build their models
to fit reality as closely as possible…

Engineers build reality
to fit their model as closely as possible…