Modeling Packet Delay in Ethernet and IP Networks

André Vallat

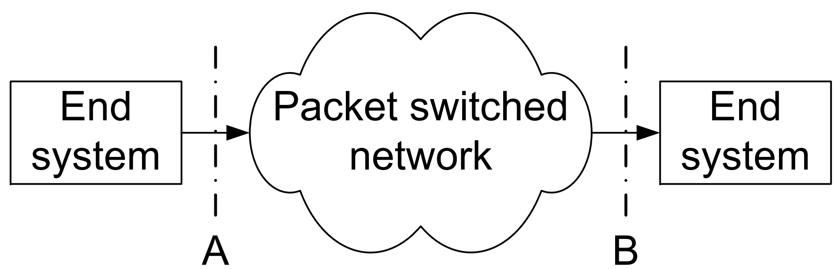
Dominik Schneuwly

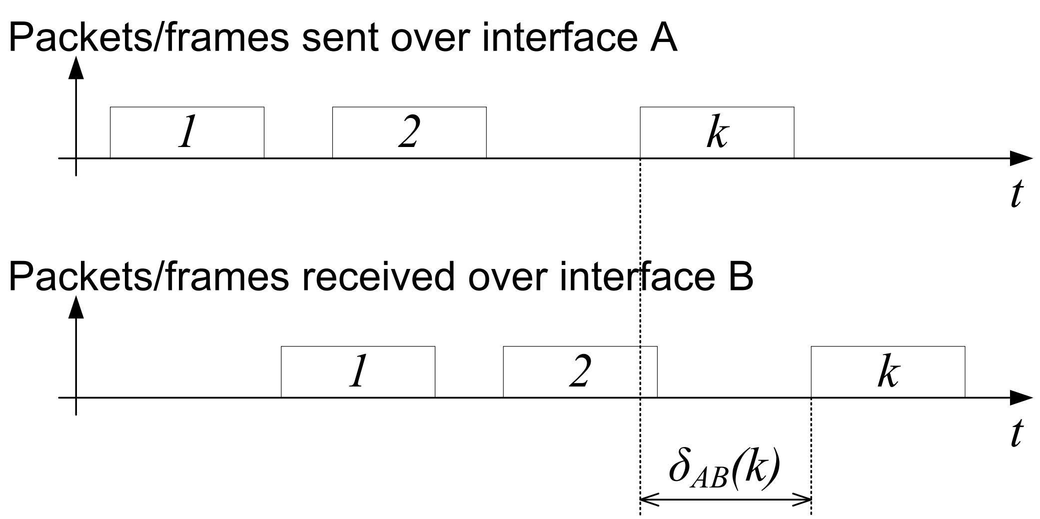
Introduction

- Is it possible to transfer time and frequency over packet switched networks with accuracies commonly required by telecom applications and equipment?
- If yes, under what conditions?
- Take as an example the ‘Precision Time Protocol’ (IEEE 1588 v2)
- Use simulations to study packet propagation properties and predict performance
- Study the influence of traffic load
- Study the influence of protocol support in switching/routing nodes (e.g. Transparent Clocks)

Definition: Packet Delay $\delta_{AB}(k)$

Consider two interfaces A and B traversed by a given packet flow.



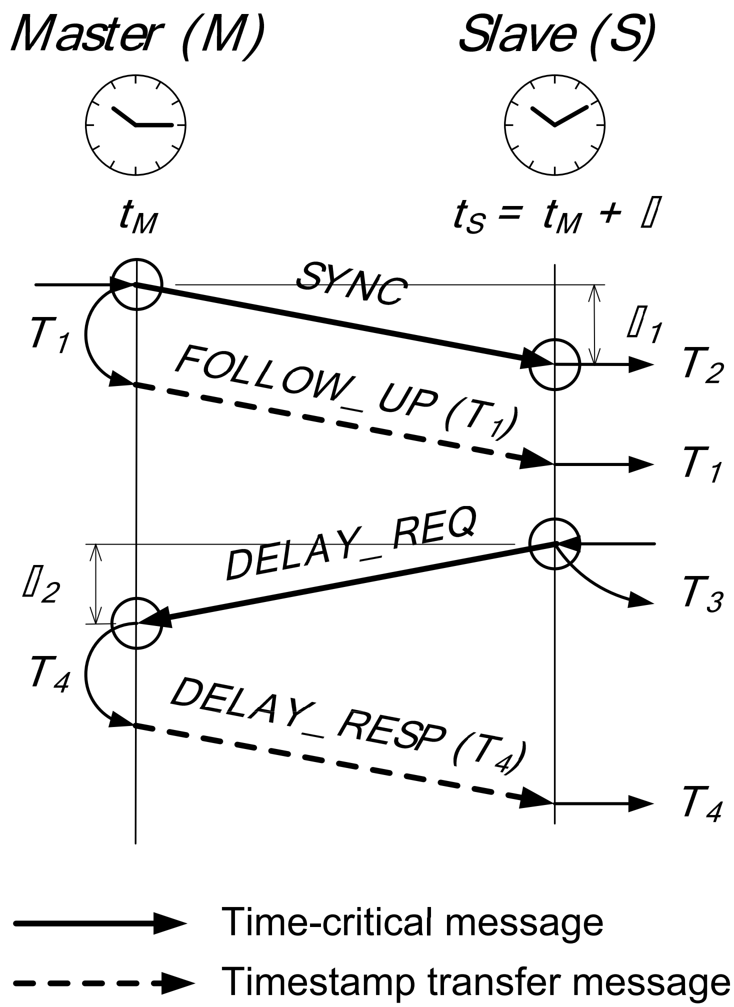


Definition: Packet Delay Asymmetry A(k)

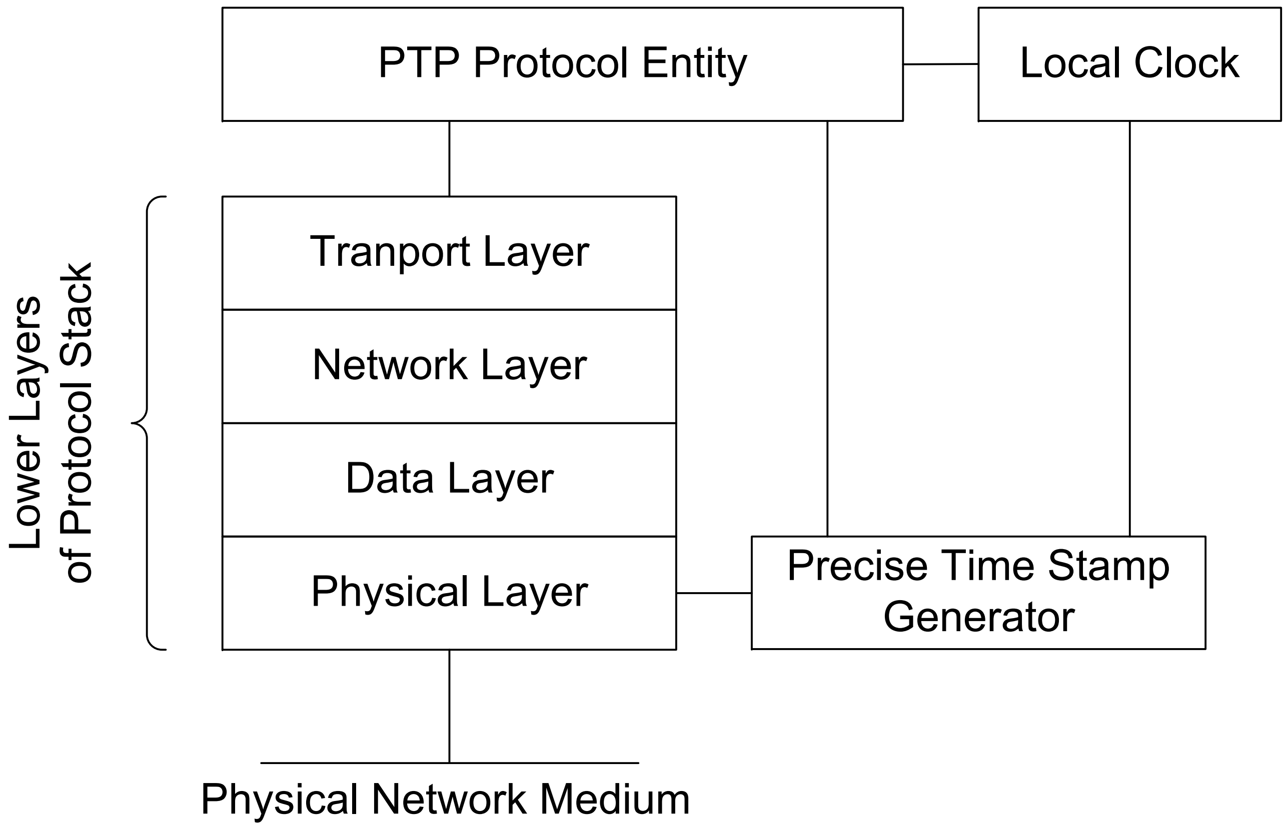
Consider two interfaces A and B traversed by bi-directional paired packet flows, where the k-th packet pair experiences the delays $\delta_{AB}(k)$ and $\delta_{BA}(k)$:

$$A(k) = \delta_{AB}(k) - \delta_{BA}(k)$$

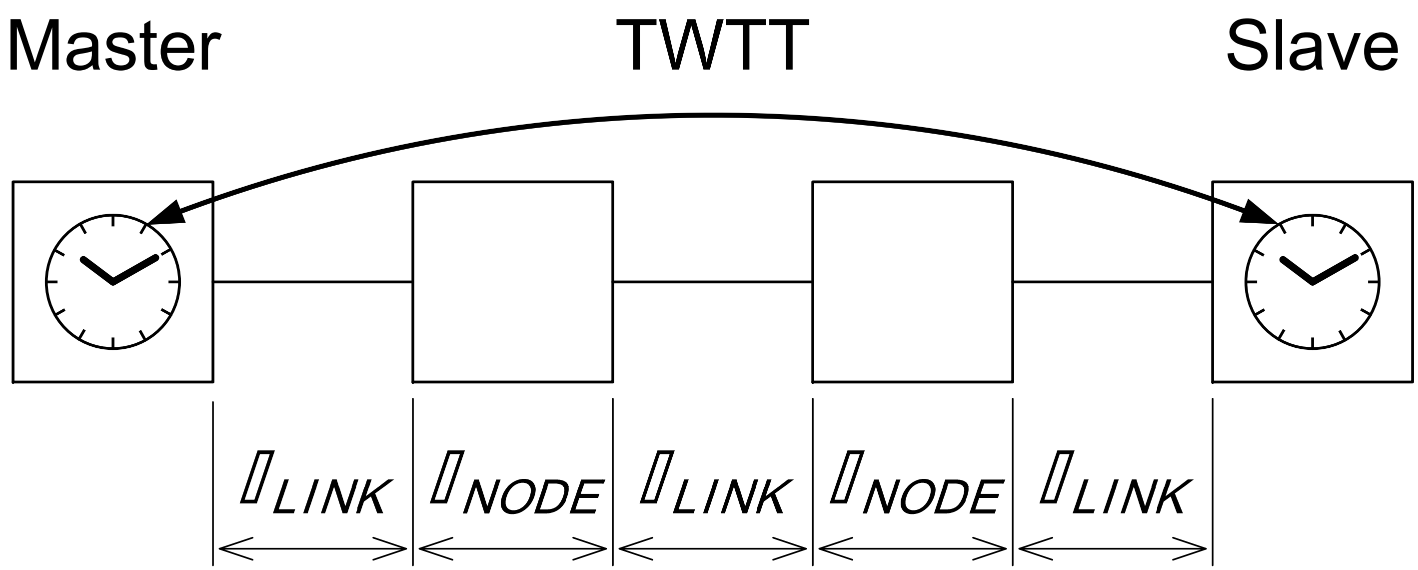
PTP IEEE 1588: TWTT



PTP IEEE 1588: Time-stamping



PTP IEEE 1588: Transparent Clocks



End-to-end Transparent Clock

Ordinary TWTT:

$$\bar{\theta} = \frac{(T_2 - T_1) - (T_4 - T_3)}{2} = \theta + \frac{\delta_{M \to S} - \delta_{S \to M}}{2} = \theta - \frac{A}{2}$$

PTP with End-to-end Transparent Clocks:

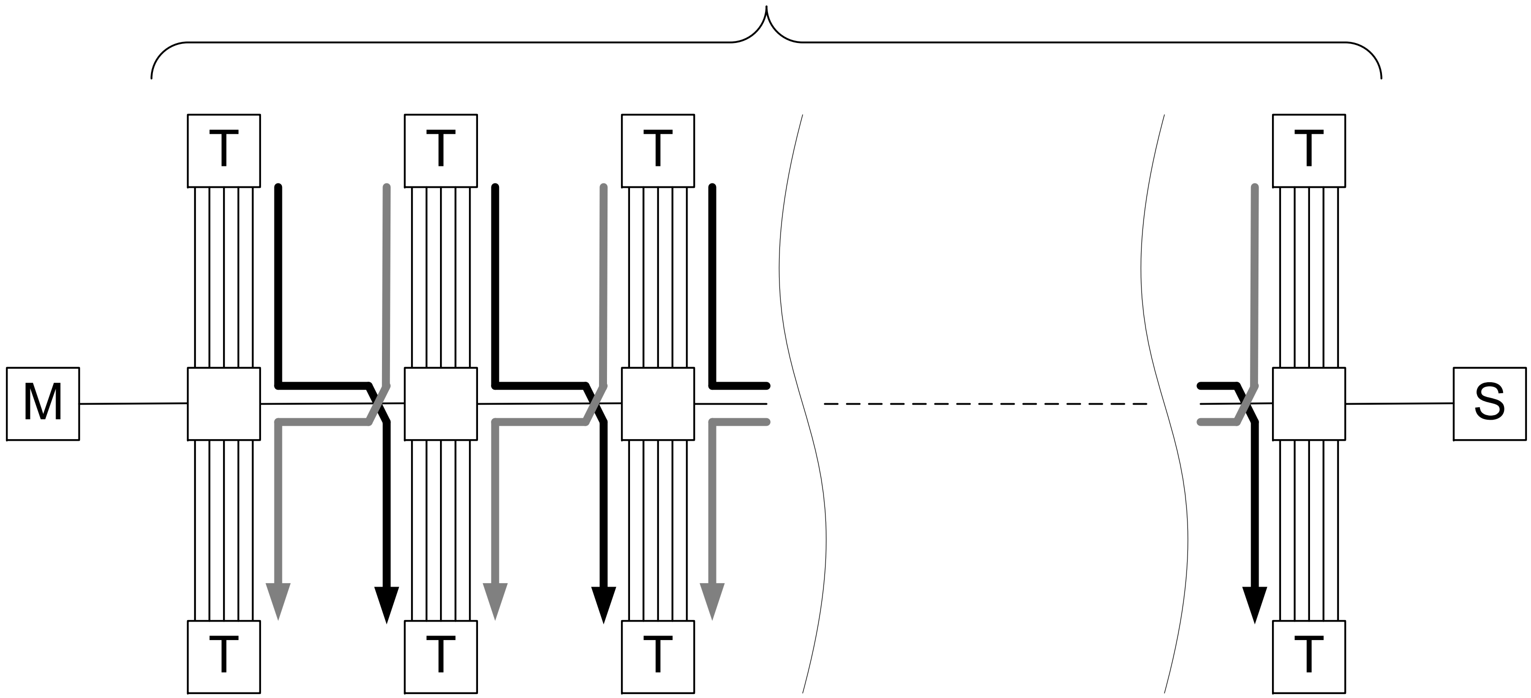
$$\bar{\delta}_{M \to S} = \sum \bar{\delta}_{NODE, M \to S}$$

$$\bar{A} = \sum \bar{\delta}_{NODE, M \to S} - \sum \bar{\delta}_{NODE, S \to M}$$

$$\bar{\theta} = \frac{\left(T_4 - T_3\right) - \left(T_2 - T_1\right)}{2} - \frac{\bar{A}}{2}$$

Simulation: network structure

10 nodes, traffic sources and traffic sinks



M	= PTP Master		= Switch or Router
S	= PTP Slave	T	= Traffic Source & Sink

Traffic sources

- Each source bloc shown in the diagram represents 5 independent smaller sources with the same stochastic properties
- Traffic generation process: see next slide
- Packet sizes:

Packet size [octet]	1500	500	125
Probability	0.6	0.1	0.3

Traffic generation process

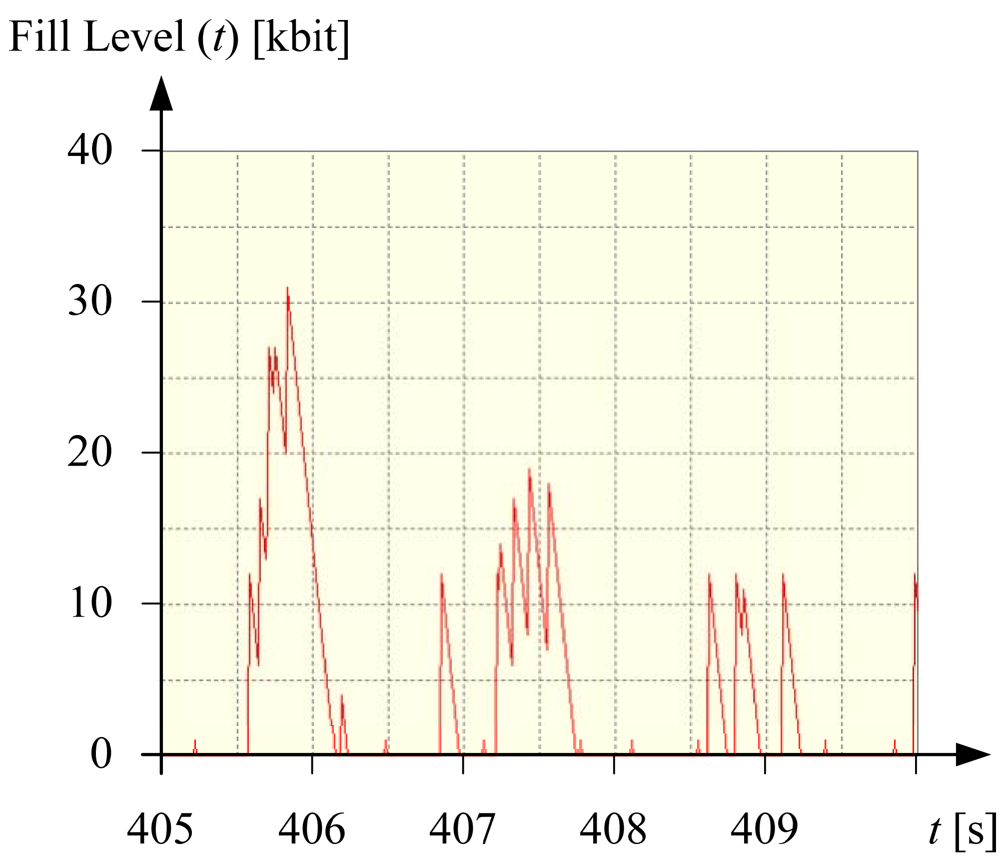
At each sampling instant where a traffic source has the opportunity to start sending a new packet, the source does so with a probability p. The probability p stays the same for periods of $1000 \cdot T_S$. At the beginning of each such period, p changes according to the algorithm described by the following pseudo-code (the parameter a is related to the average traffic load):

R is drawn randomly in $[-a/2, +a/2]$, where a is a constant in $[0, 1]$
if $(R < 0)$ then $\Delta p := R$
if $(R >= 0)$ then $\Delta p := R \cdot (1 - p(t))$
if $(p(t) + \Delta p < 0)$ then $p(t + T_S) := -(p(t) + \Delta p)$
if $(p(t) + \Delta p >= 0)$ then $p(t + T_S) := p(t) + \Delta p$

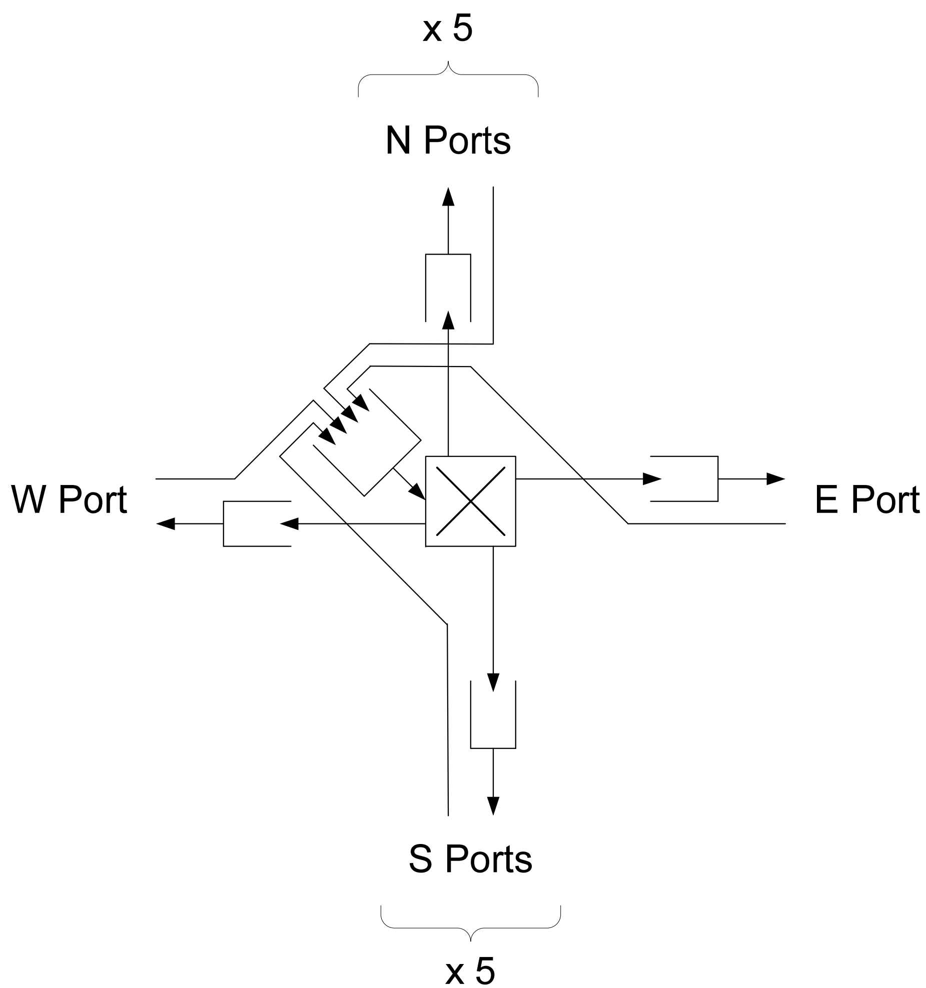
The algorithm for changing p resembles the random walk process. The ordinary random walker is modified in the sense that its excursion is limited to the interval [0,1] by a sort of hard saturation at the lower, and soft saturation at the upper end of the interval.

Traffic burstiness

- Fill level of an output buffer on the PTP path reflects traffic evolution:



Simulation: node structure



Simulation scenarios

- Traffic load cases:
  - $\lambda = 0.2$
  - $\lambda = 0.4$
  - $\lambda = 0.8$

  where λ = (**average** data rate / link capacity) on a longitudinal link
- Number of non-PTP-capable nodes:
  - $N_N = 2$: (- O - X - X - O - O - O - O - O - O - O -)
  - $N_N = 4$: (- O - X - X - X - X - O - O - O - O - O -)
  - $N_N = 10$: (- X - X - X - X - X - X - X - X - X - X -)

  where X = non-PTP-capable node, O = PTP-capable node
- A total of 9 simuation scenarios

Simulation parameters and output

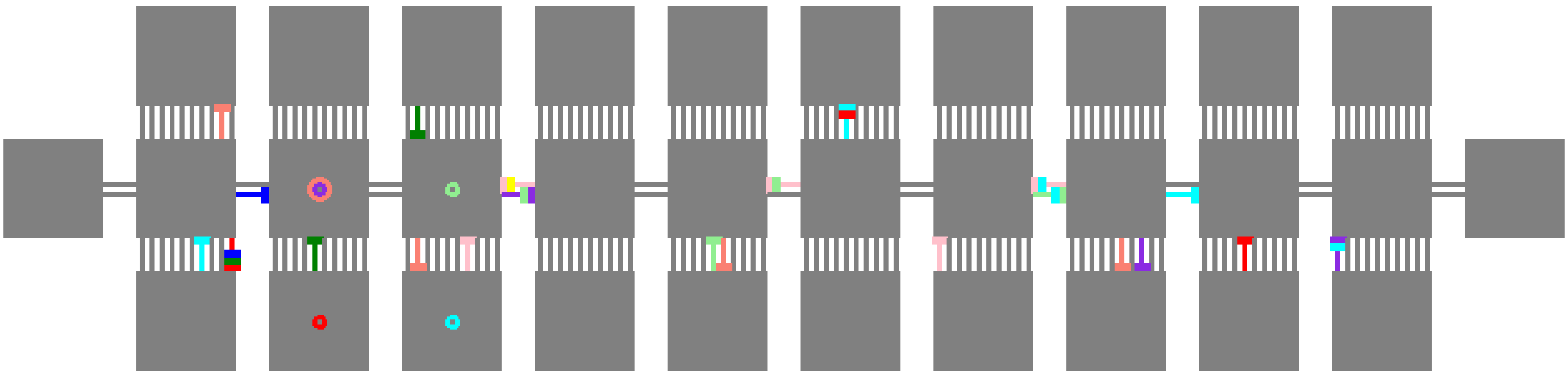
Simulation parameters:

- Input queues = 100,000 octets
- Output queues = 100,000 octets
- Service time of switching matrix: 10 µs per packet
- Link capacity C = 100 Mbit/s
- TWTT interogation rate = 100 s^{-1} (SYNC & DELAY_REQ)
- Sampling period T_S = 10 µs
- Simulation length = 30,000 s
- => 400 mio. samples on 112 elements (nodes, sources, etc.)

Main simulation output:

- Residual Packet Delay $\delta_{R,M\to S} = \delta_{M\to S} - \overline{\delta}_{M\to S}$
- Residual Packet Delay Asymmetry $A_R = A - \overline{A}$

Simulator State Display



minTDEV and 'min'TDEV

minTDEV of residual delay δ_R:

$$\text{minTDEV}(\tau) = \sqrt{\tfrac{1}{6}\left\langle\left[\delta_{R,\min}(i+2n) - 2\delta_{R,\min}(i+n) + \delta_{R,\min}(i)\right]^2\right\rangle}$$

$\tau = n\tau_0$, $\tau_0 =$ sampling period

$\delta_{R,\min}(i) = \min\left[\delta_R(j)\right]$ for $\forall j : i <= j <= i+n$

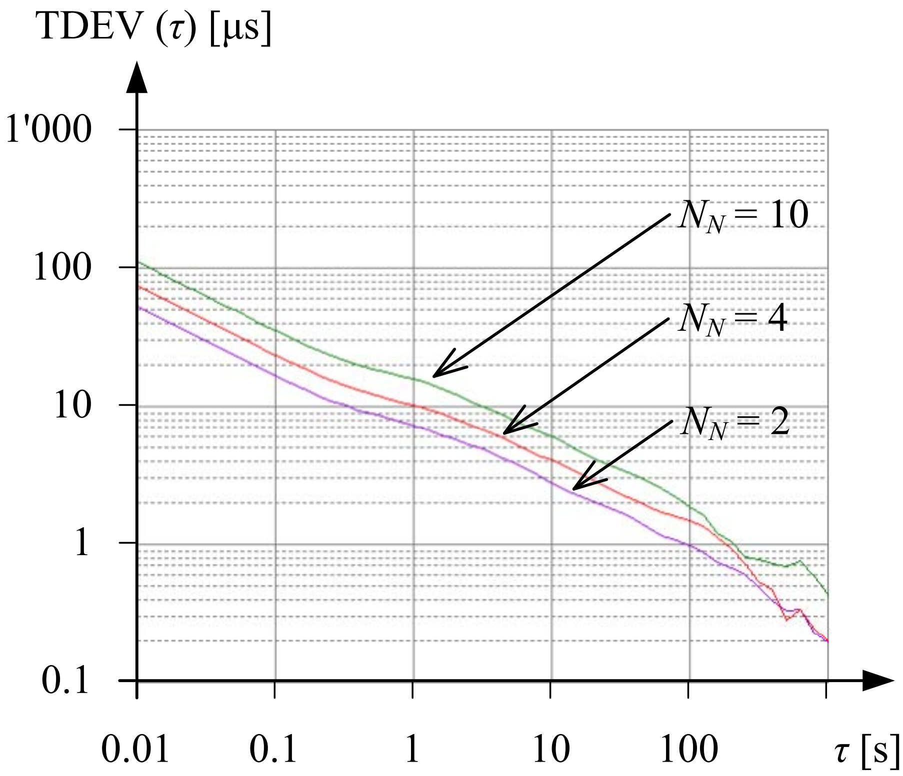
'min'TDEV of residual asymmetry A_R:

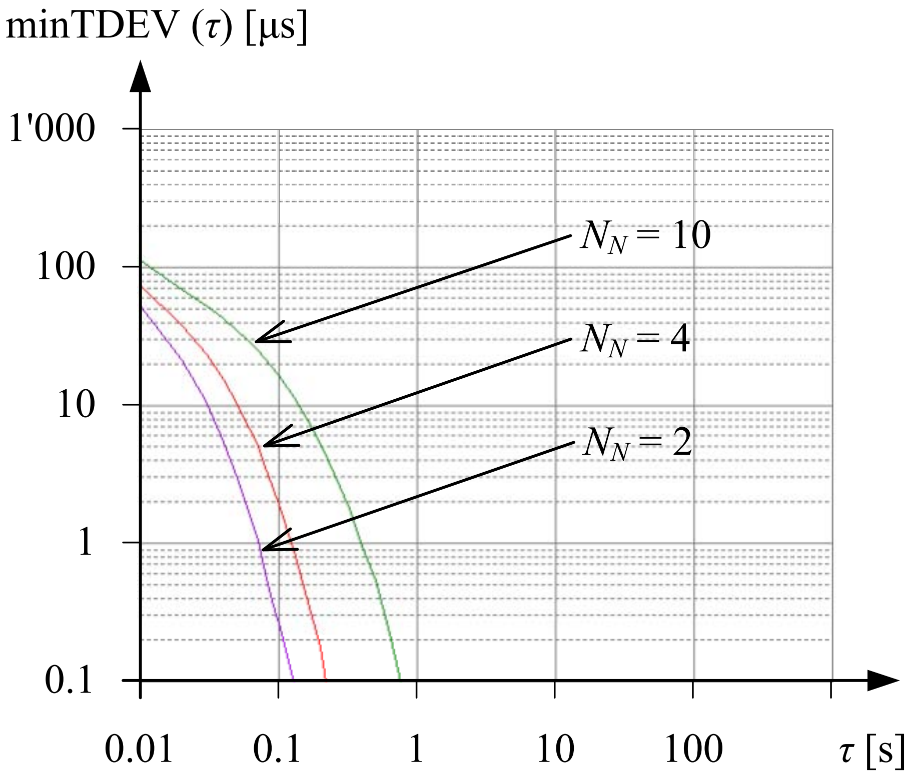
$$\text{'min'TDEV}(\tau) = \sqrt{\tfrac{1}{6}\left\langle\left[A_{R,\min}(i+2n) - 2A_{R,\min}(i+n) + A_{R,\min}(i)\right]^2\right\rangle}$$

$\tau = n\tau_0$, $\tau_0 =$ sampling period

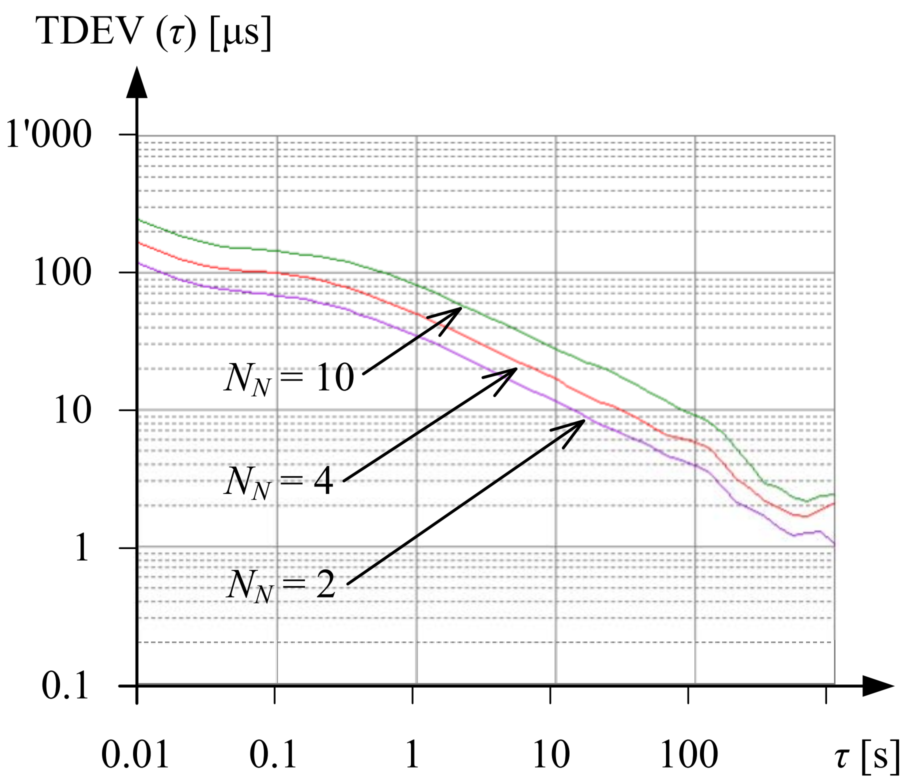
$A_{R,\min}(i) = \min\left[\delta_{R,M\to S}(j)\right] - \min\left[\delta_{R,S\to M}(j)\right]$ for $\forall j : i <= j <= i+n$

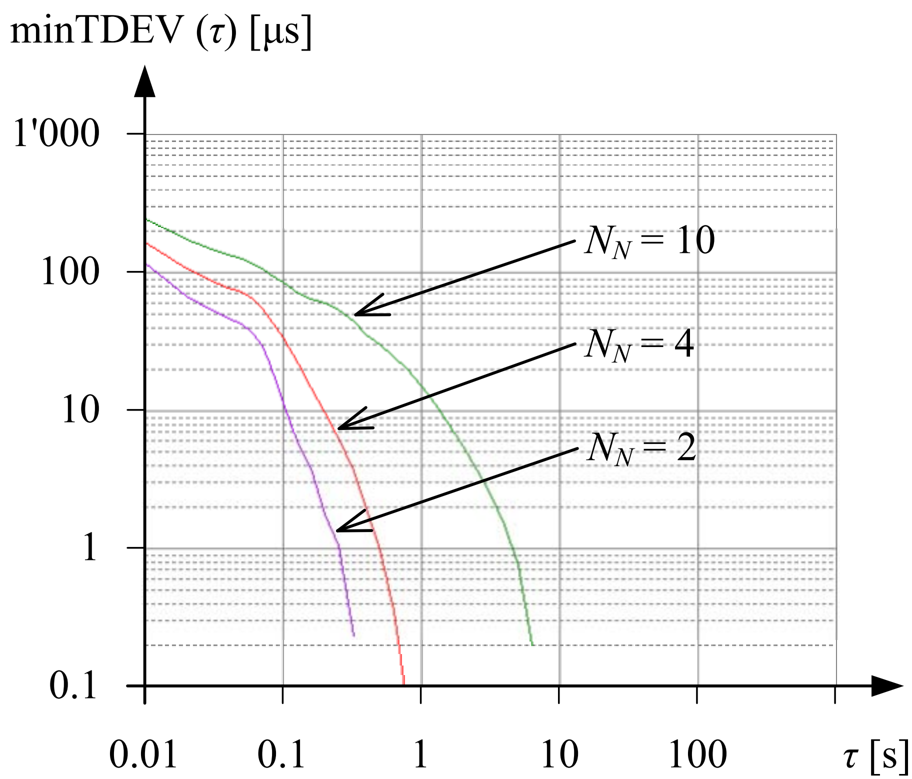
Residual Delay at λ = 0.2



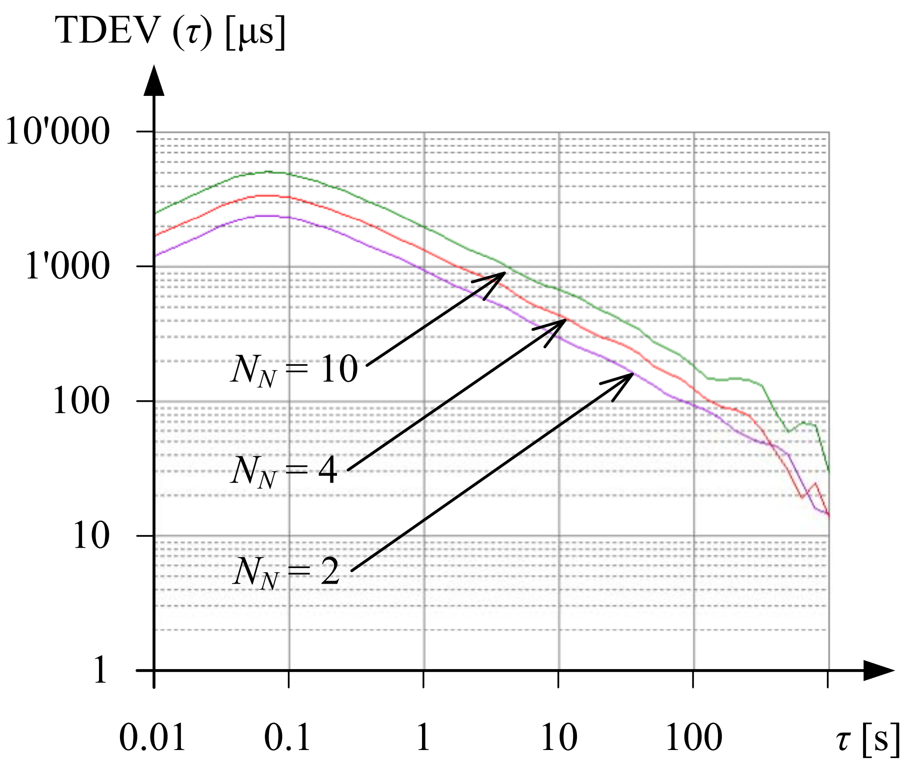


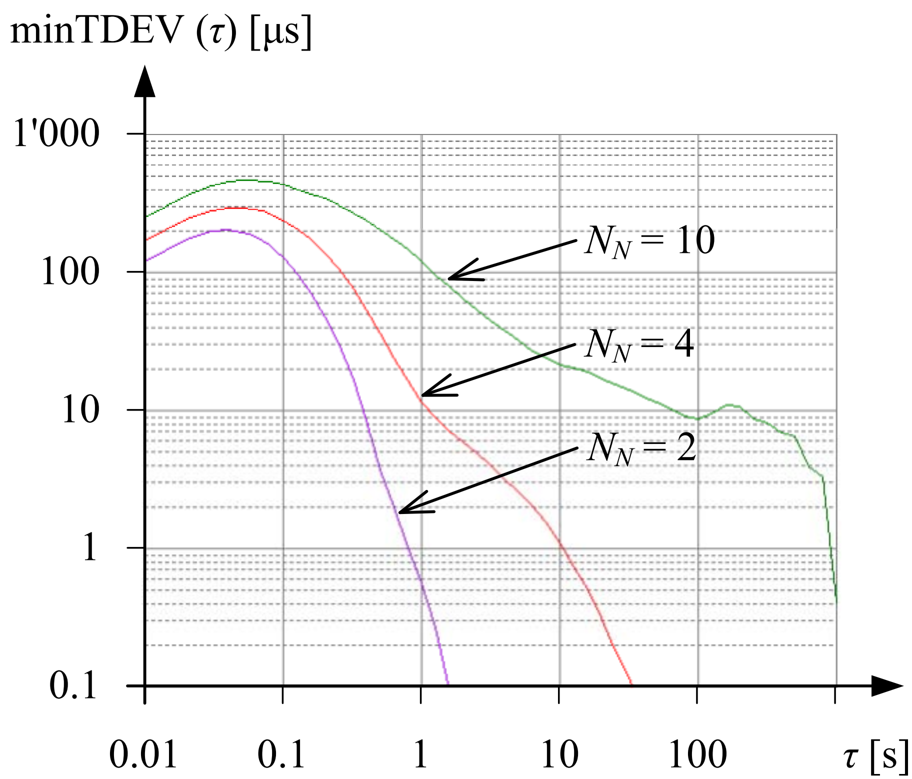
Residual Delay at λ = 0.4



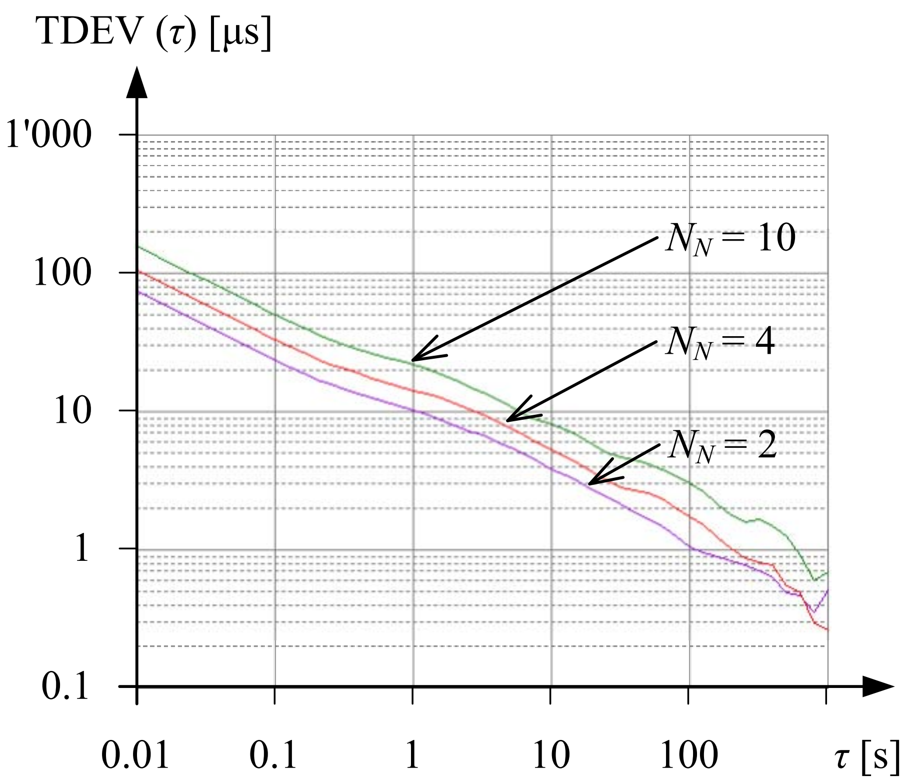


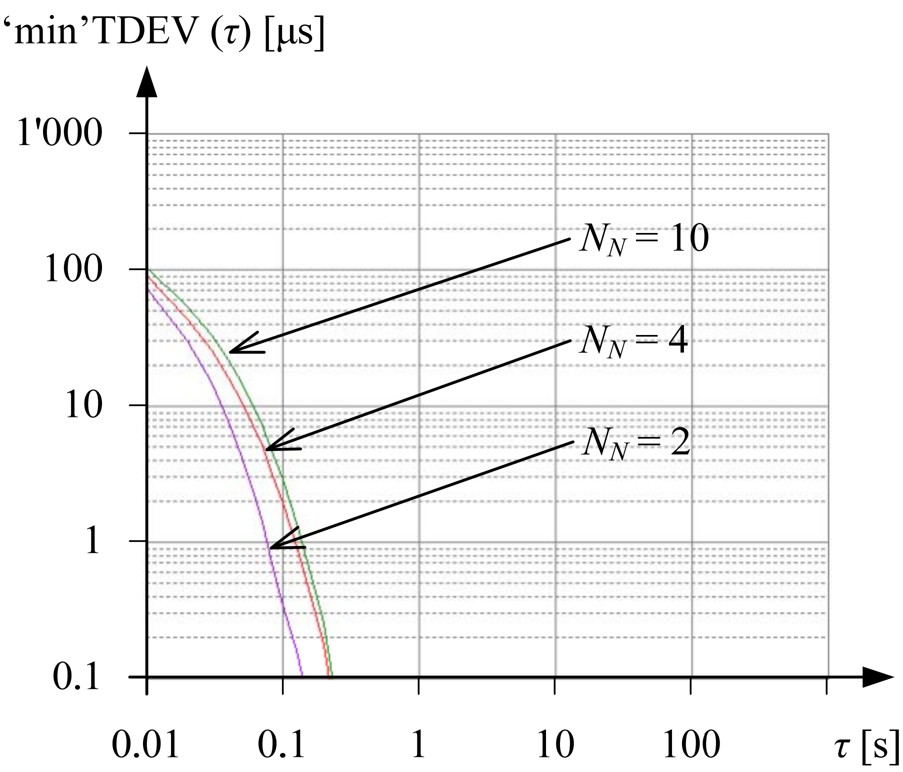
Residual Delay at λ = 0.8



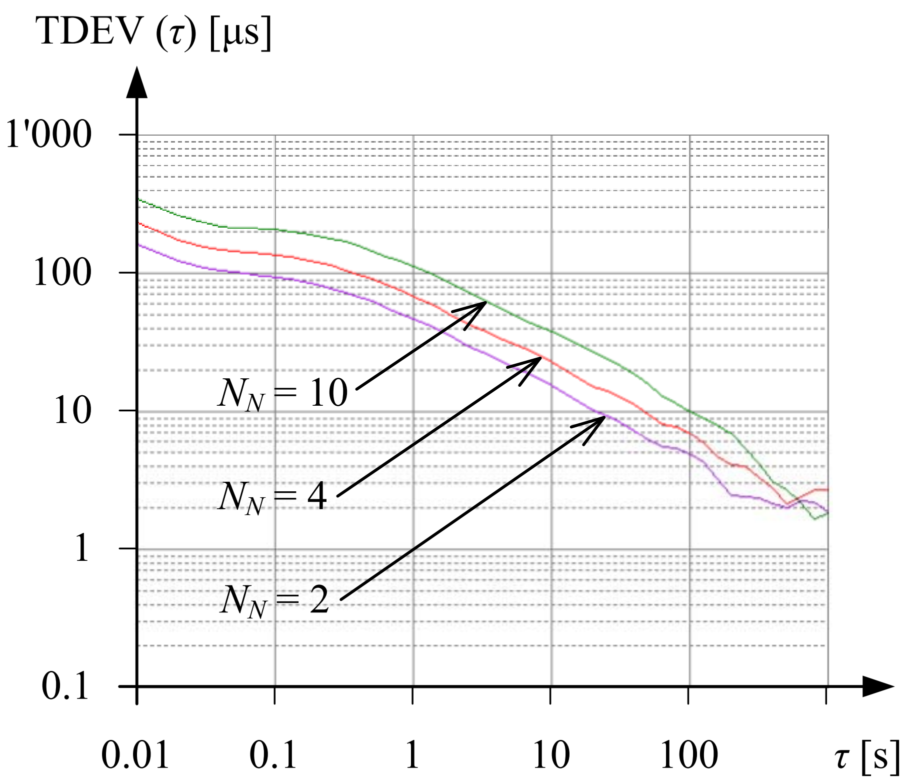


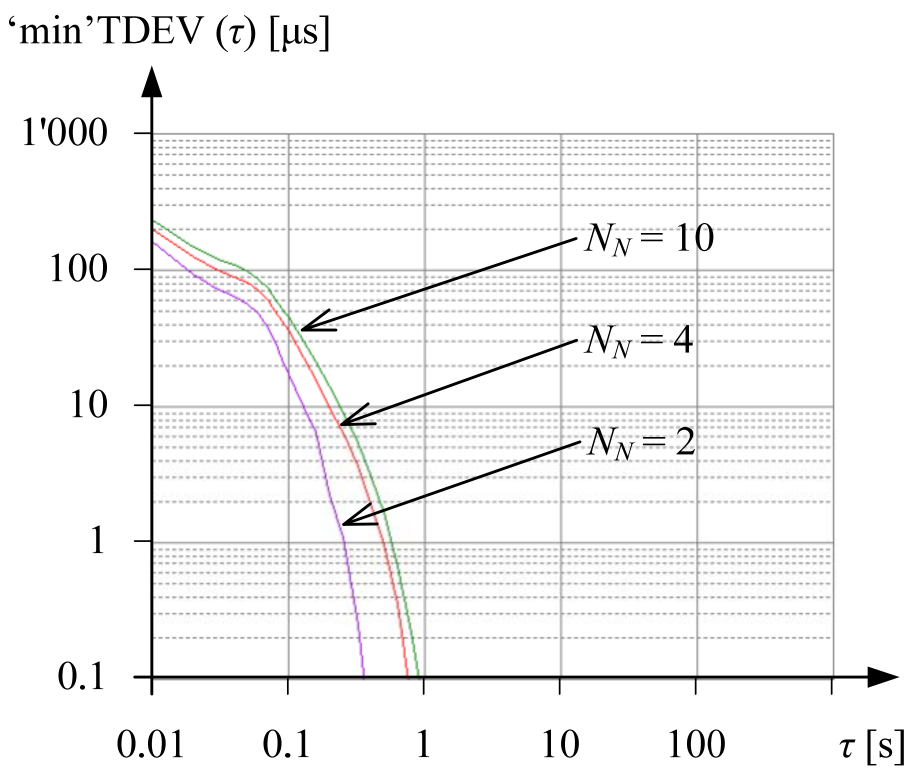
½(Residual Asymmetry) at λ = 0.2



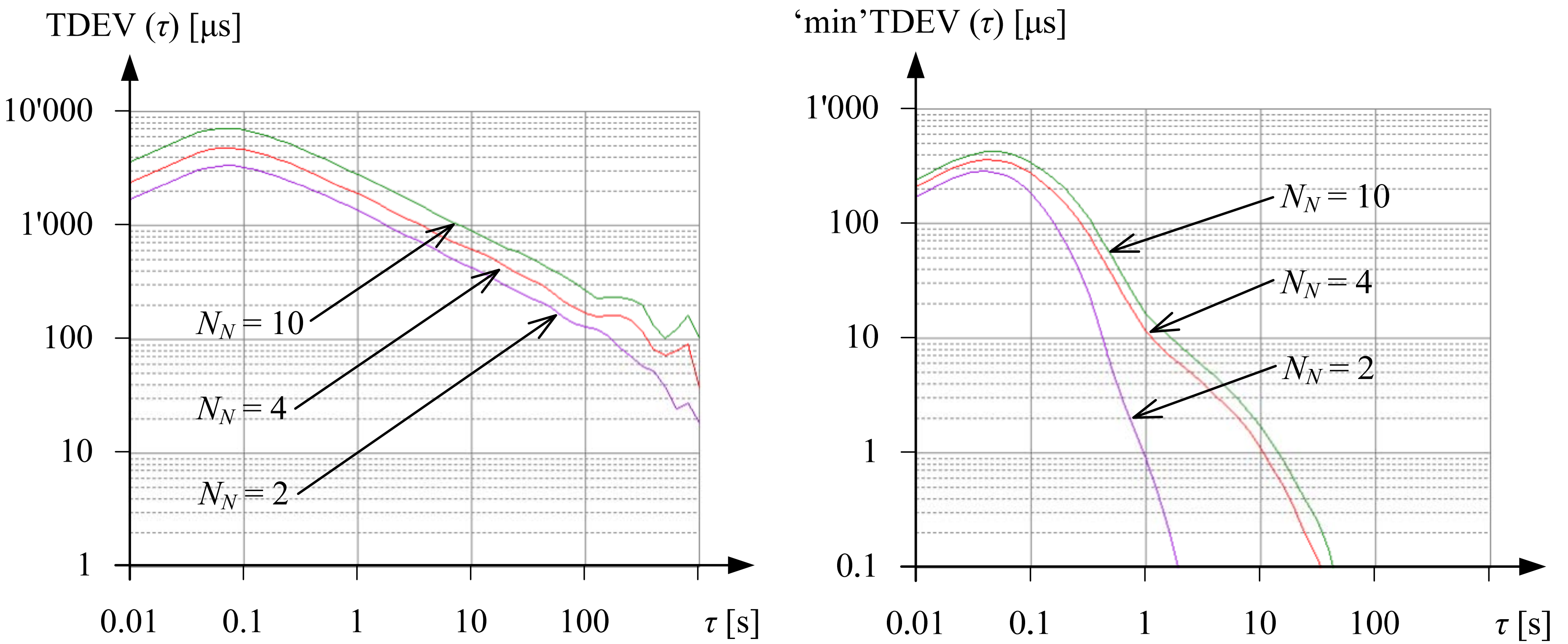


½(Residual Asymmetry) at $\lambda = 0.4$





½(Residual Asymmetry) at λ = 0.8



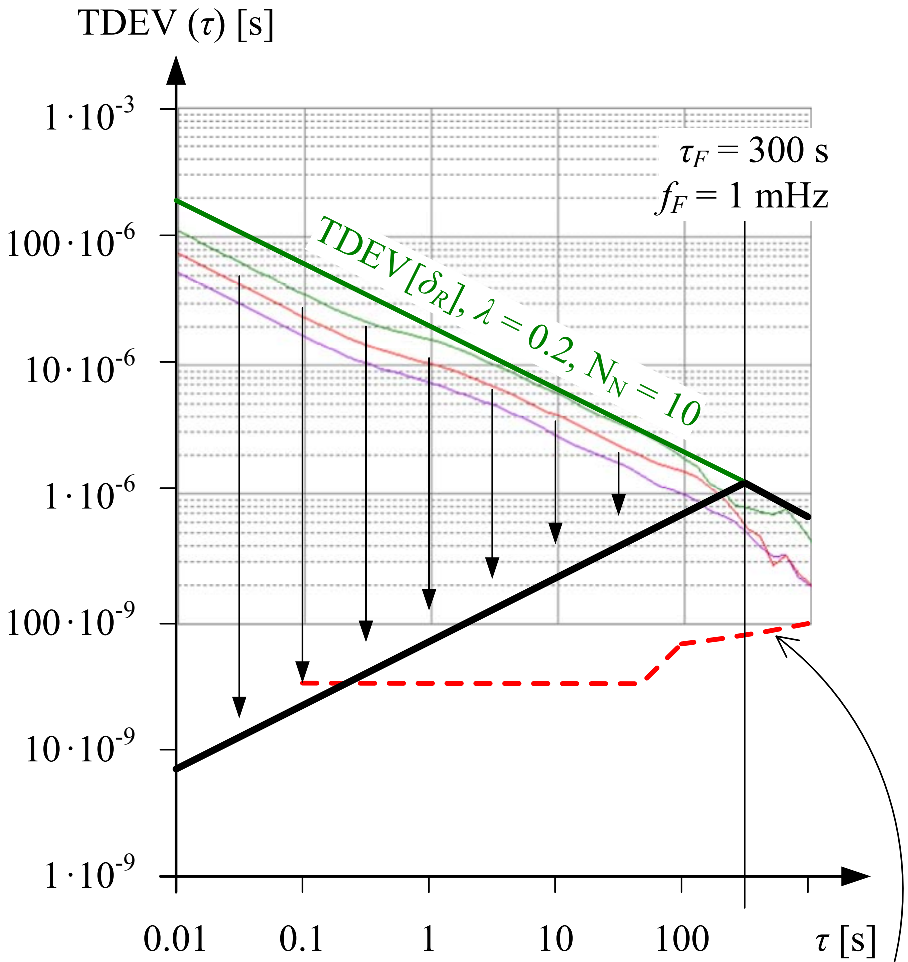
Imagine a PTP slave with …

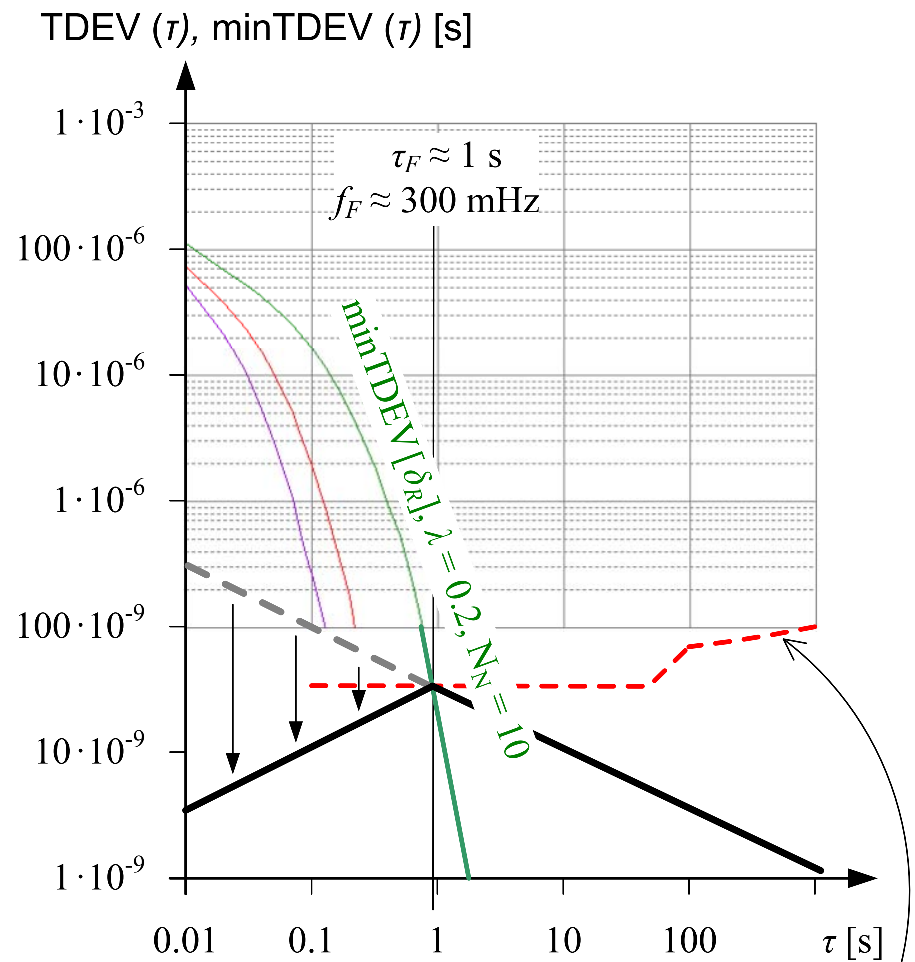
$$x_1(k) \longrightarrow \boxed{\text{F1}} \xrightarrow{x_2(k)} \boxed{\text{F2}} \xrightarrow{x_3(k)}$$

- F1 picks the fastest packets within a time window τ_{F1}
- F2 is a linear low-pass filter with bandwidth τ_{F1}
- F1 conserves minTEV ($\tau \geq \tau_{F1}$)
- For x_2: TDEV ($\tau = \tau_{F1}$) = minTDEV($\tau = \tau_{F1}$)
- F2 attenuates TDEV ($\tau < \tau_{F2}$)

- Assume $x_2(k)$ is white noise and take $\tau_{F1} = \tau_{F2}$: see following slides …

Filtering Residual Delay at $\lambda = 0.2$

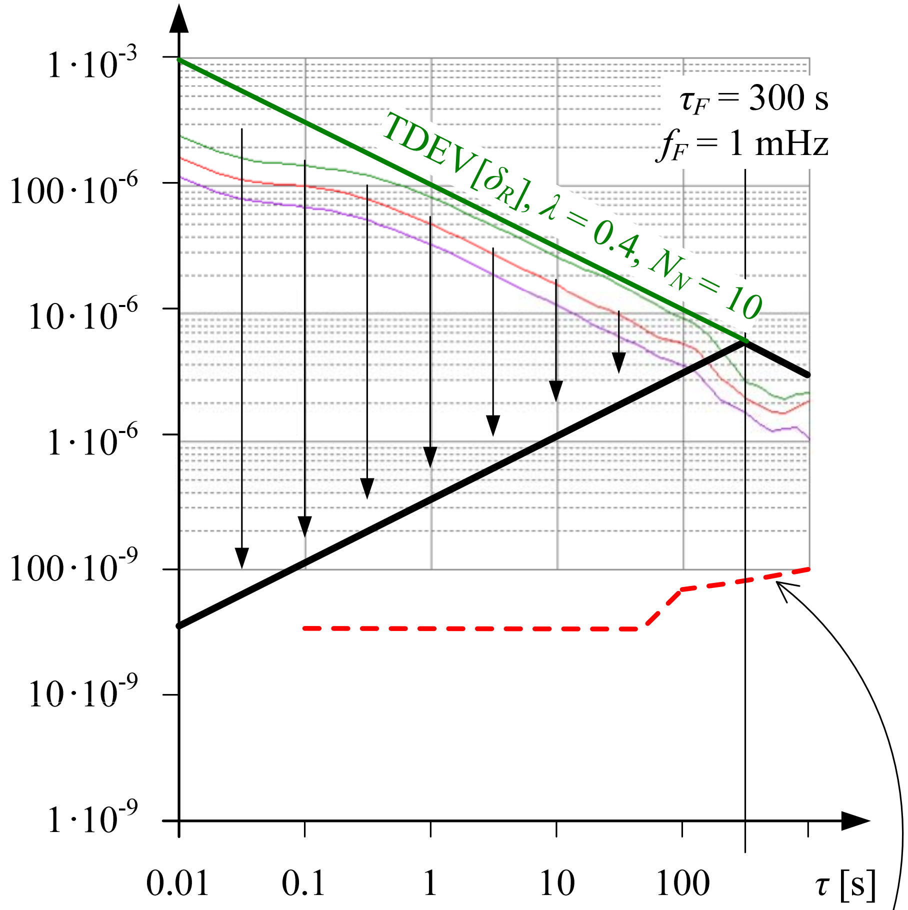


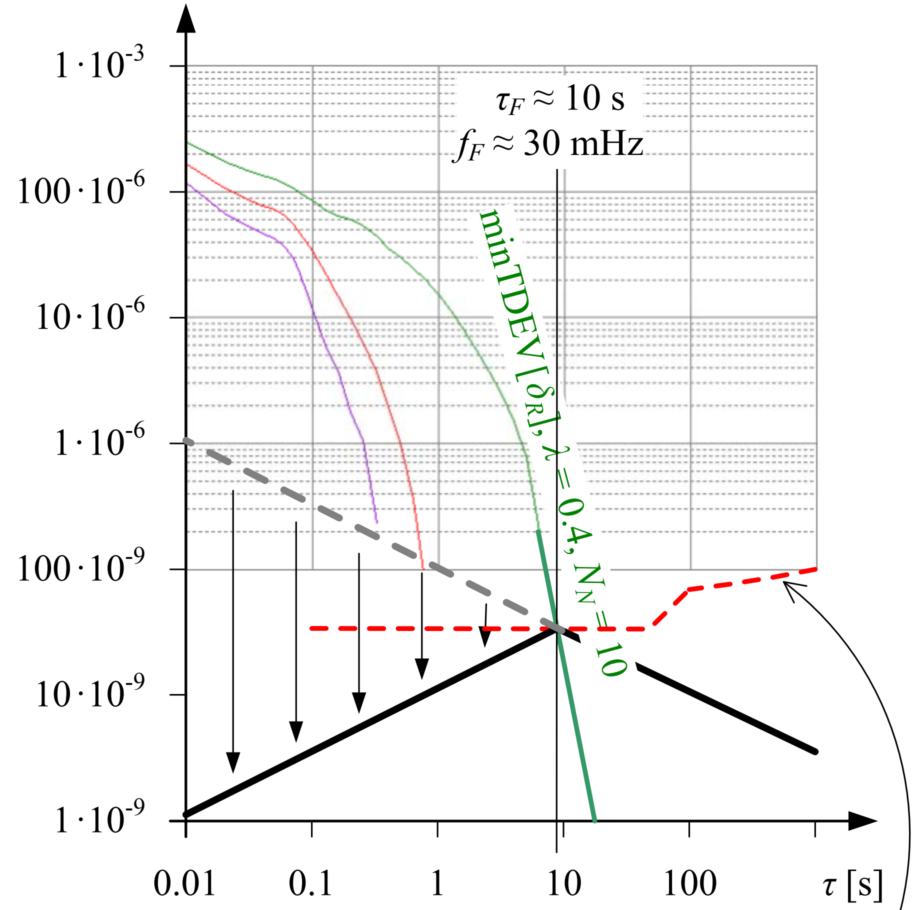


Network Limit for PDH synch. interfaces, ITU-T Rec. G.823

Network Limit for PDH synch. interfaces, ITU-T Rec. G.823

Filtering Residual Delay at $\lambda = 0.4$

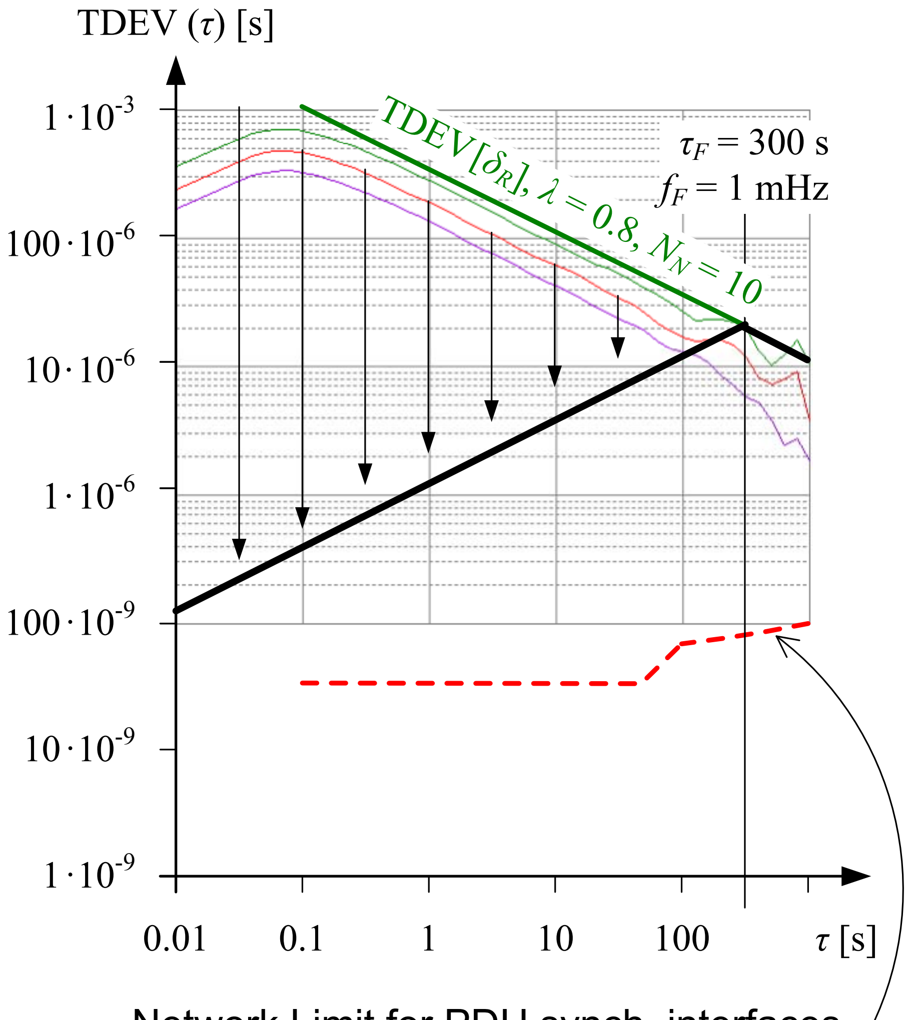


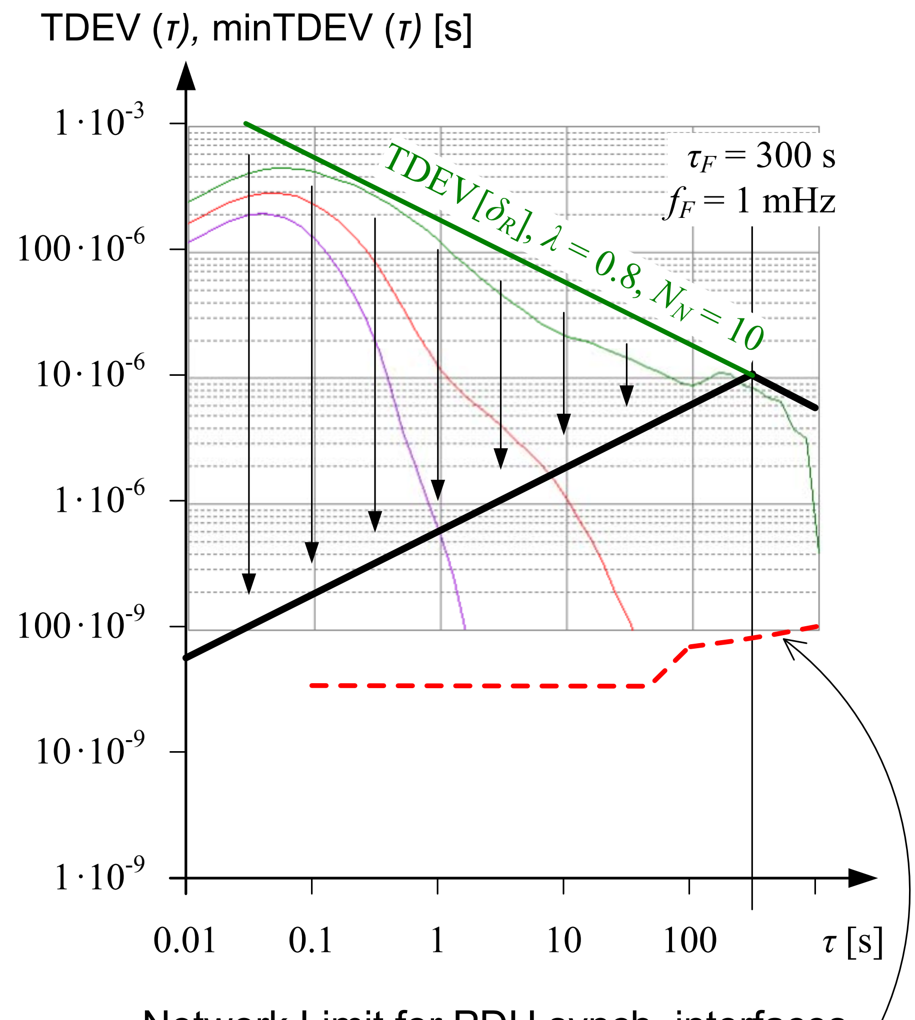


Network Limit for PDH synch. interfaces, ITU-T Rec. G.823

Network Limit for PDH synch. interfaces, ITU-T Rec. G.823

Filtering Residual Delay at $\lambda = 0.8$





Network Limit for PDH synch. interfaces,
ITU-T Rec. G.823

Network Limit for PDH synch. interfaces,
ITU-T Rec. G.823

Filtering Residual Delay: Summary

Filter bandwidth required for compliance with Network Limit		λ: Average traffic load		
		0.2	0.4	0.8
N_N: Number of non-PTP-capable nodes	2	2 Hz	700 mHz	140 mHz
	4	1 Hz	300 mHz	5 mHz
	10	300 mHz	30 mHz	**Not feasible**

Performance Suitable for Telecoms?

- The filter bandwidth required for compliance with the Network Limit is a good indication about how difficult (expensive) it is to achieve the required performance.
- A 1 mHz filter bandwidth is a practical feasibility limit (below 1 mHz, oscillator cost gets prohibitive).
- Without the use of Transparent Clocks, performance is suitable with traffic loads up to 0.4.
- WIth Transparent Clocks, performance is suitable with traffic loads up to 0.8.

Conclusions

- With PTP IEEE 1588 v2, telecom performance level is achievable over networks with up to 10 switching/routing nodes under certain conditions.
- In networks with PTP-capable nodes (Transparent Clocks), telecom performance is achievable with traffic loads up to 0.8.
- If traffic load is limited to about half the capacity, telecom performance is achieveable in PTP networks without PTP-capable nodes.
- PTP v2 suitable for the distribution of time and frequency in network types such as Metro Area Networks, base-station backhaul networks, access aggregation networks, etc.

Thank you

Objectives, questions