TIMING IN PACKET NETWORKS

STEFANO RUFFINI – WSTS 2017

CONTENTS

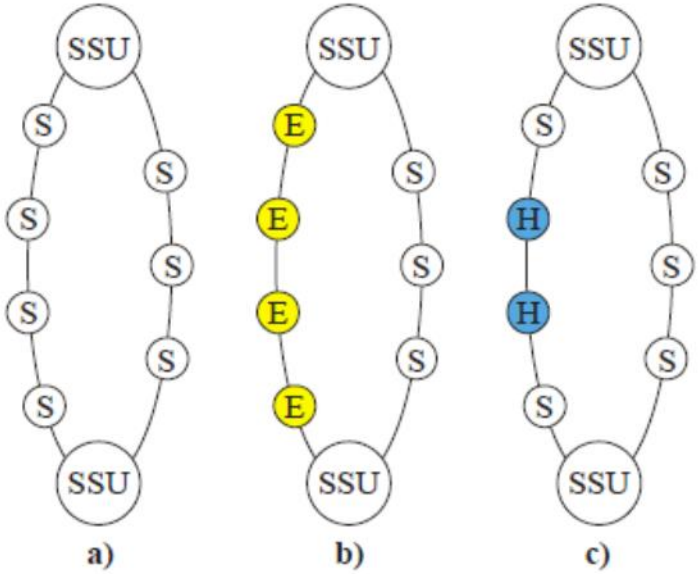
  - Note: some of this material is based on earlier presentations from Christian Farrow and Kishan Shenoi

HISTORICAL BACKGROUND

- Packet switching network does not require sync itself (at least traditional packet networks ..)
- CBR (Constant Bit Rate) services over ATM, one of the first packet–sync related examples
  - Methodologies to recover the CBR sync rate (e.g. 2 Mbit/s services, including 50 ppm frequency deviation):
    - **Network Synchronous**
    - **Adaptive**
    - Differential
- Generalization due to **migration to packet networks** (Ethernet-IP; Ethernet Physical layer traditionally defined as «asynchronous»):
  - Support timing requirements of the connected networks (e.g. **Mobile applications**)
  - **Circuit Emulation** detailed performance analysis
  - Frequency sync distribution via dedicated protocols (NTP, PTP)
  - Standardized performance objectives over reference networks (e.g. ITU-T Recc. G.8261)
  - Definition of a synchronous Ethernet (syncE) physical layer (G.8261, G.8262, G.8264)
- Current main focus is to deliver **time/phase sync** reference
  - **Packet-based sync technologies** required (may be combined with synchronous physical layer)
- «Deterministic» packet networks (e.g. **TSN**-IEEE, **Detnet**-IETF) as a related topic

SYNCE: INTRODUCTION

- Several applications requiring **accurate frequency reached by Ethernet**
  - Since the very start of timing over packet network activities, it was proposed to define possible use of synchronous Ethernet physical layer
  - Not in contradiction with IEEE (10^{-11} within the +/-100 ppm - 20 ppm)
  - Only in full duplex mode (continuous signal required)
- **Based on SDH** specification (for interoperability and simplifying the standardization task)
  - Synchronous Ethernet equipment equipped with a **synchronous Ethernet Equipment Clock – EEC** (G.8262). Synchronous Ethernet interfaces extract the received clock and pass it to the system clock.
  - Synchronization Status Message as per G.8264
  - Ongoing work to define an **enhanced SyncE** (G.8262.1)
- It does not transport Time (but it has been proposed)
- All nodes must support SyncE: sync chain as per G.803
  - Cannot be transported transparently across network boundaries



G.8261-Y.1361(13)_FD.1

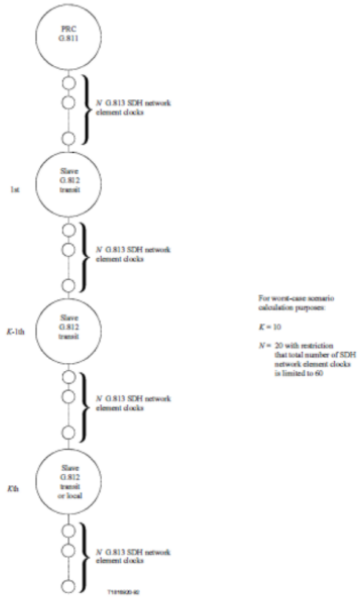
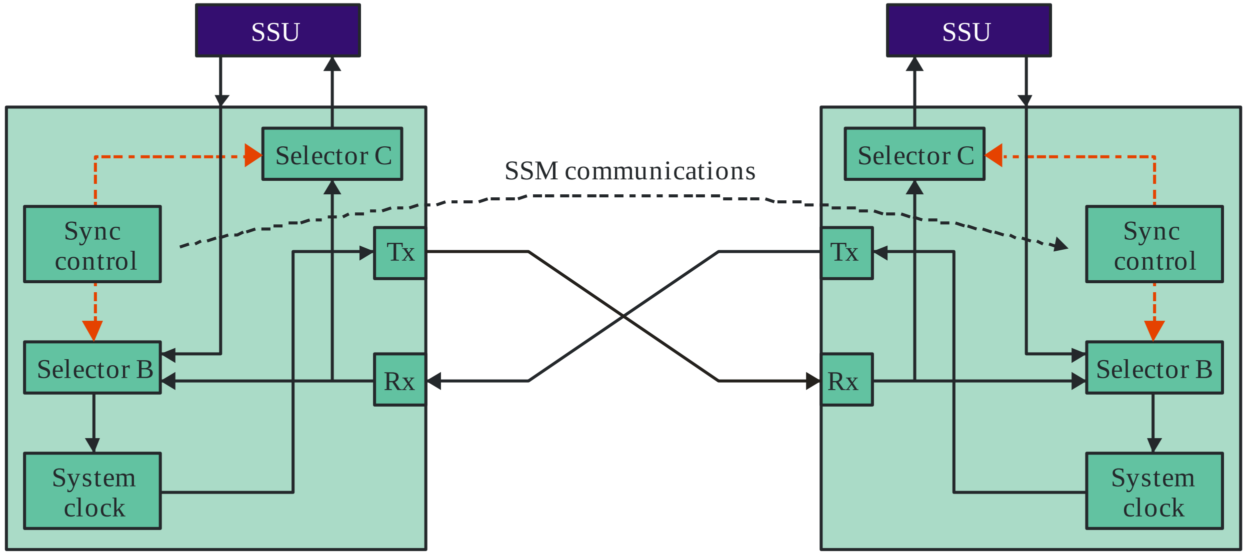


Figure 8-5/G.803 – Synchronization network reference chain

SSM (SYNCHRONIZATION STATUS MESSAGE) IN SYNCE

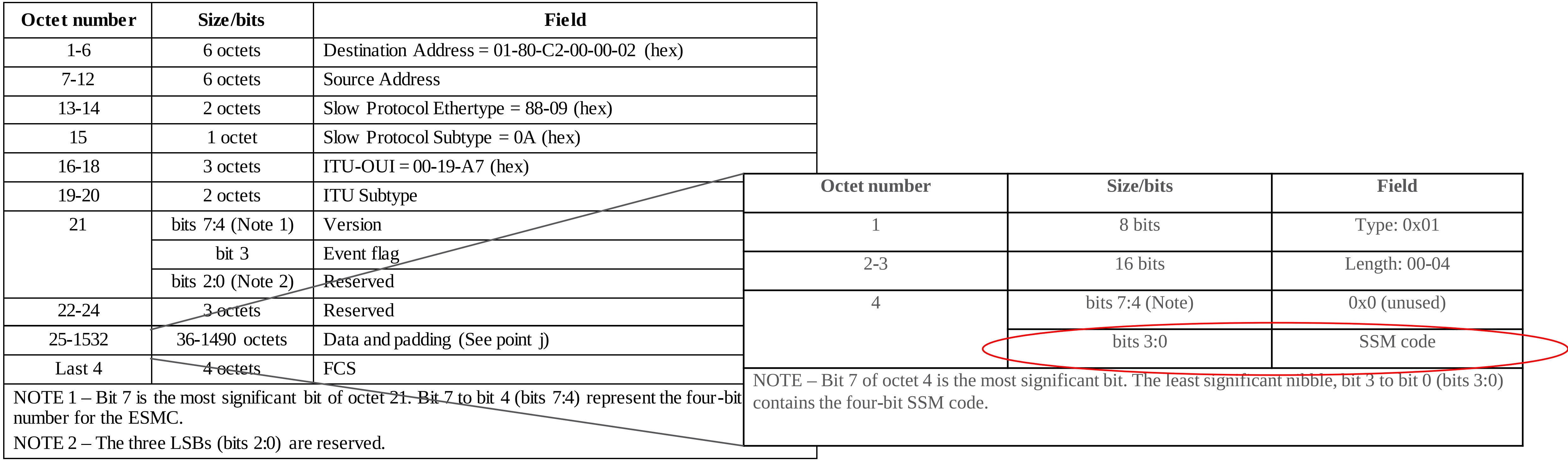
- SSM required to prevent timing loops and to support reference selection (as per SDH);
  - details according to G.781 and G.8264
- In SDH SSM delivered in fixed locations of the SDH frame
  - Packet based mechanism required in case of SyncE
- OUI (organizationally unique identifier) from IEEE reused to specify exchange of QLs over the OAM specific slow protocol (OSSP)
- EEC option 1 clock treated as G.813 option 1 (QL-SEC), EEC option 2 as an G.812 type IV clock (QL-ST3).
- Two types of protocol message types are defined
  - "heart-beat" message (once per second)
  - Event message generated immediately
- SSM QL value is considered failed if no SSM messages are received after a five second period.



G.8264-Y.1364(14)_F11-1

ETHERNET SYNCHRONIZATION MESSAGING CHANNEL (ESMC) FORMAT

- ESMC PDU with QL TLV always sent as the first TLV in the Data and padding field



Octet number	Size/bits	Field
1-6	6 octets	Destination Address = 01-80-C2-00-00-02 (hex)
7-12	6 octets	Source Address
13-14	2 octets	Slow Protocol Ethertype = 88-09 (hex)
15	1 octet	Slow Protocol Subtype = 0A (hex)
16-18	3 octets	ITU-OUI = 00-19-A7 (hex)
19-20	2 octets	ITU Subtype
21	bits 7:4 (Note 1)	Version
	bit 3	Event flag
	bits 2:0 (Note 2)	Reserved
22-24	3 octets	Reserved
25-1532	36-1490 octets	Data and padding (See point j)
Last 4	4 octets	FCS

NOTE 1 – Bit 7 is the most significant bit of octet 21. Bit 7 to bit 4 (bits 7:4) represent the four-bit number for the ESMC.

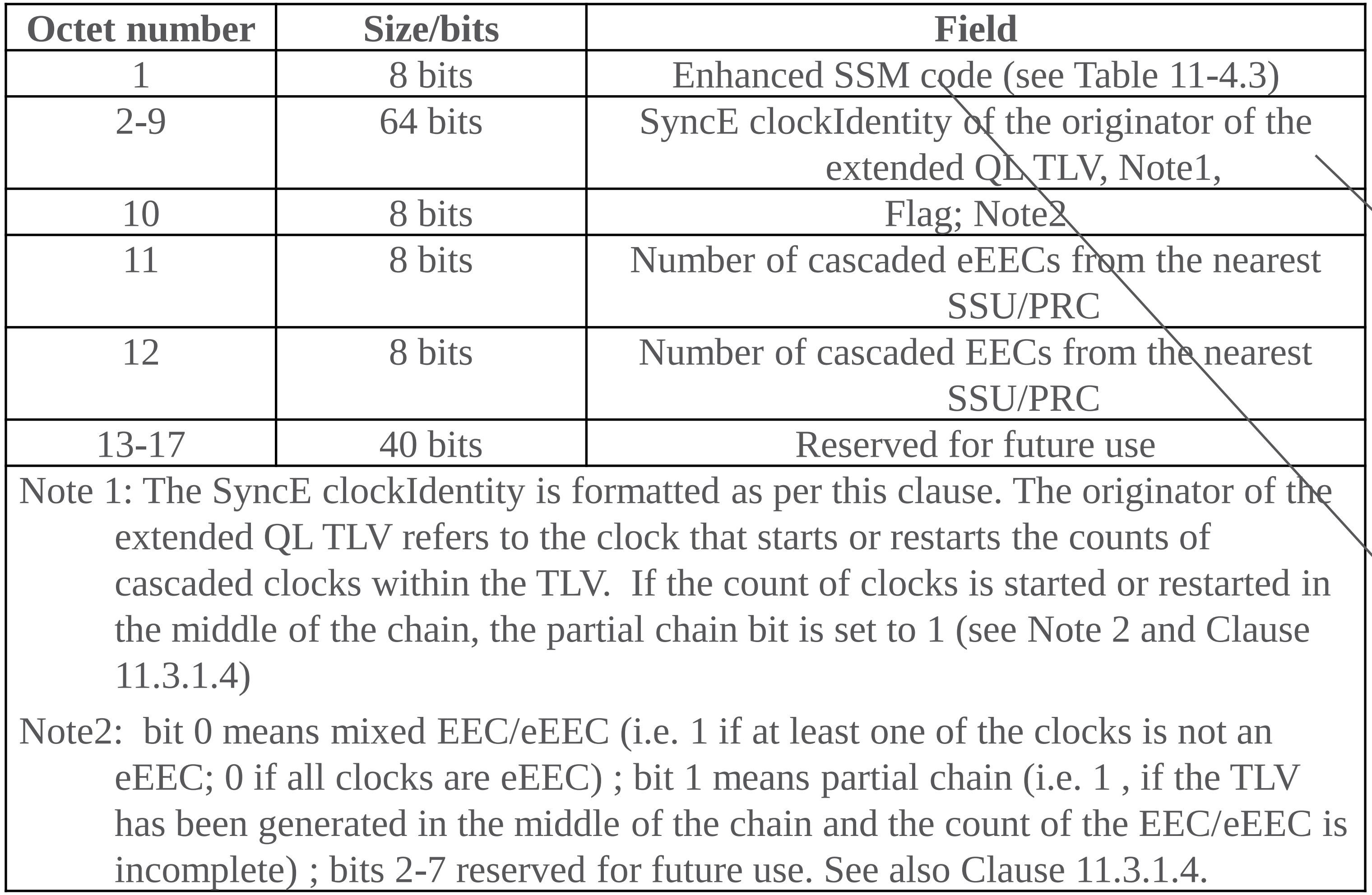
NOTE 2 – The three LSBs (bits 2:0) are reserved.

Octet number	Size/bits	Field
1	8 bits	Type: 0x01
2-3	16 bits	Length: 00-04
4	bits 7:4 (Note)	0x0 (unused)
	bits 3:0	SSM code

NOTE – Bit 7 of octet 4 is the most significant bit. The least significant nibble, bit 3 to bit 0 (bits 3:0) contains the four-bit SSM code.

- Recently extended to carry new clock types (and inform on PRTC traceability)
  - Extended QL TLV

EXTENDED QL

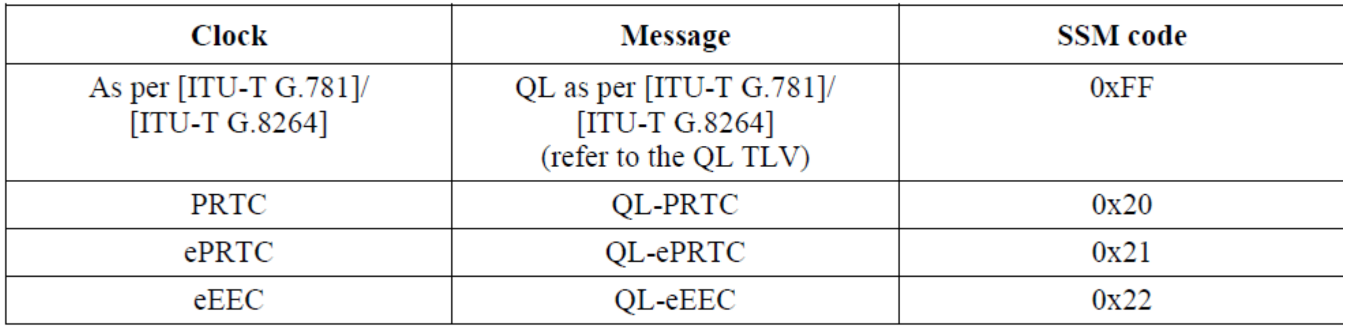


Octet number	Size/bits	Field
1	8 bits	Enhanced SSM code (see Table 11-4.3)
2-9	64 bits	SyncE clockIdentity of the originator of the extended QL TLV, Note1,
10	8 bits	Flag; Note2
11	8 bits	Number of cascaded eEECs from the nearest SSU/PRC
12	8 bits	Number of cascaded EECs from the nearest SSU/PRC
13-17	40 bits	Reserved for future use

Note 1: The SyncE clockIdentity is formatted as per this clause. The originator of the extended QL TLV refers to the clock that starts or restarts the counts of cascaded clocks within the TLV. If the count of clocks is started or restarted in the middle of the chain, the partial chain bit is set to 1 (see Note 2 and Clause 11.3.1.4)

Note2: bit 0 means mixed EEC/eEEC (i.e. 1 if at least one of the clocks is not an eEEC; 0 if all clocks are eEEC) ; bit 1 means partial chain (i.e. 1 , if the TLV has been generated in the middle of the chain and the count of the EEC/eEEC is incomplete) ; bits 2-7 reserved for future use. See also Clause 11.3.1.4.

SyncE clockIdentity follows the IEEE 1588 rules



Clock	Message	SSM code
As per [ITU-T G.781]/ [ITU-T G.8264]	QL as per [ITU-T G.781]/ [ITU-T G.8264] (refer to the QL TLV)	0xFF
PRTC	QL-PRTC	0x20
ePRTC	QL-ePRTC	0x21
eEEC	QL-eEEC	0x22

SSM CODES FOR SYNCE

Table 11-5 (G.8264 Draft): Option I

Clock	Quality level	SSM code	Enhanced SSM code
PRC	QL-PRC	0010	0xFF
SSU-A	QL-SSU-A	0100	0xFF
SSU-B	QL-SSU-B	1000	0xFF
EEC1	QL-EEC1	1011	0xFF
Note	QL-DNU	1111	0xFF
PRTC	QL-PRTC	0010	0x20
ePRTC	QL-ePRTC	0010	0x21
eEEC	QL-eEEC	1011	0x22
Note: There is no clock corresponding to this quality level.			

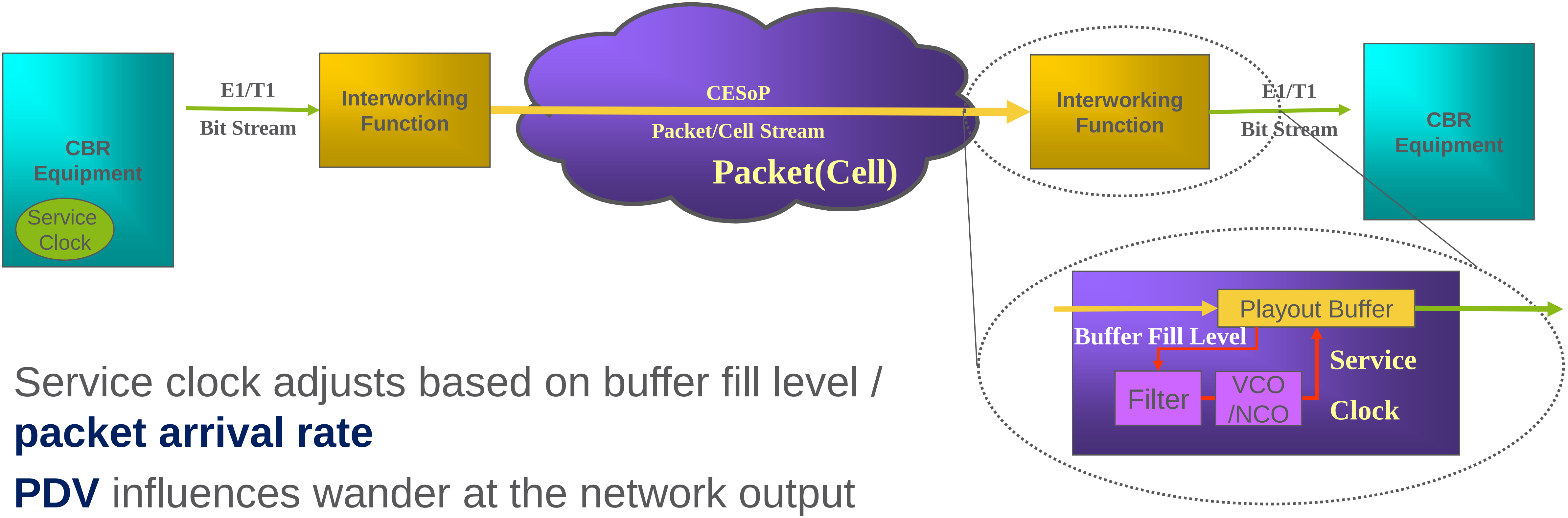
If a clock supports both the QL TLV and the extended QL TLV, it should set the SSM code and the enhanced SSM code according to table 11-5/11-6, and send both the QL TLV and the extended QL TLV.

If a clock supports only the QL TLV, it should set the SSM code according to table 11-5/11-6, and send the QL TLV.

Table 11-6 (G.8264 Draft): Option II

Clock	Quality level	SSM code	Enhanced SSM code
PRS	QL-PRC	0001	0xFF
Note	QL-STU	0000	0xFF
ST2	QL-ST2	0111	0xFF
TNC	QL-TNC	0100	0xFF
ST3E	QL-ST3E	1101	0xFF
ST3	QL-ST3	1010	0xFF
EEC2	QL-EEC2	0010	0xFF
Note	QL-PROV	1110	0xFF
Note	QL-DUS	1111	0xFF
PRTC	QL-PRTC	0010	0x20
ePRTC	QL-ePRTC	0010	0x21
eEEC	QL-eEEC	1011	0x22
Note: There is no clock that corresponds to this quality level.			

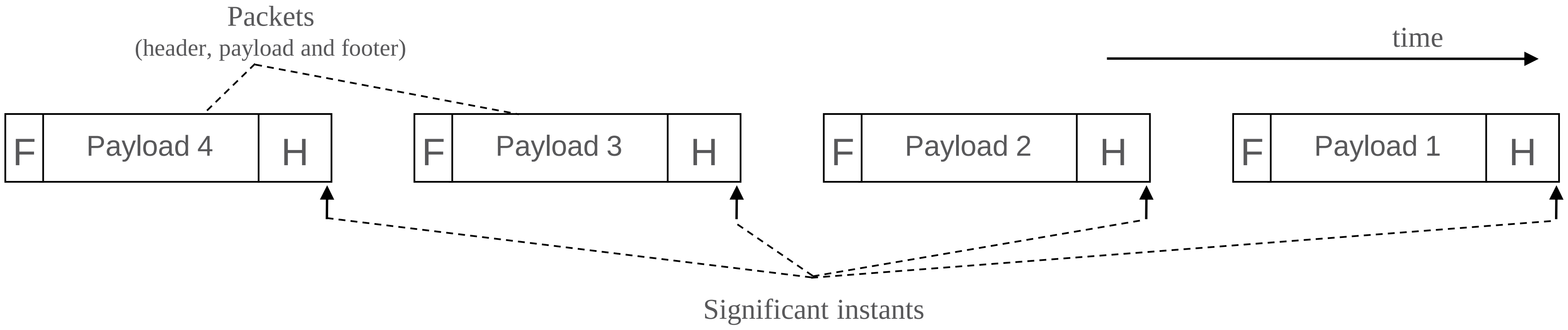
PACKET-BASED TIMING: ADAPTIVE CLOCK OPERATION



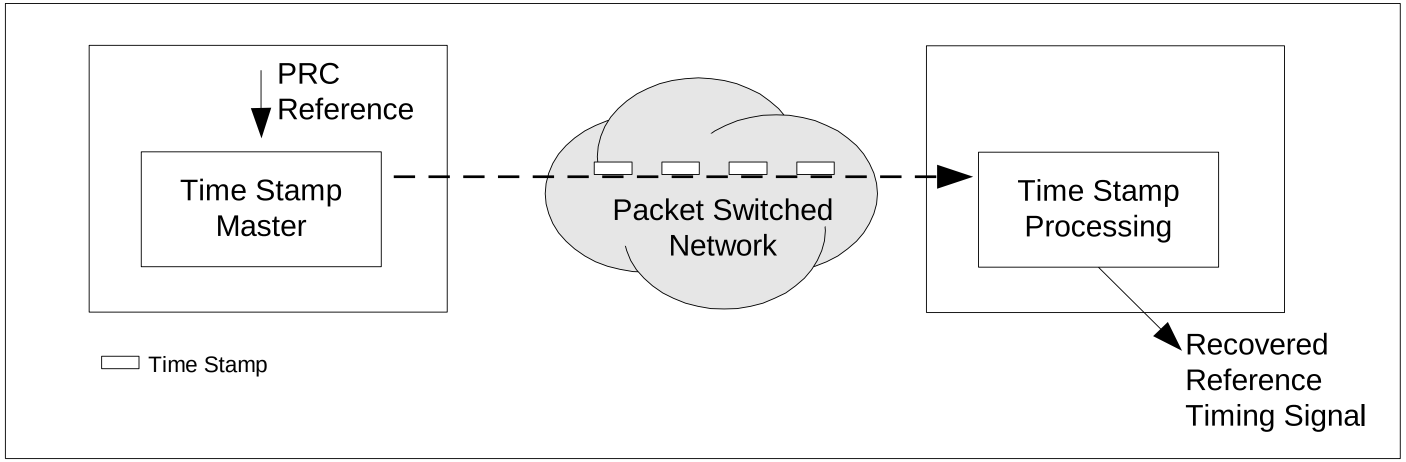
- Service clock adjusts based on buffer fill level / **packet arrival rate**
- **PDV** influences wander at the network output

FROM CLOCKS TO “PACKET CLOCKS”

- “**Packet clocks**” can be described in a similar way …
- CES Packets do have a regular rhythm
- Extension to using dedicated protocols: **NTP, PTP**
  - Packets may not arrive regularly, but **timestamps** mean time information can be extracted
  - Timing information contained in the arrival/departure time of the packets
  - **Two-way or one-way** protocols
  - Timing recovery process requires **PDV filtering**
- Time and frequency can be distributed from point A to point B



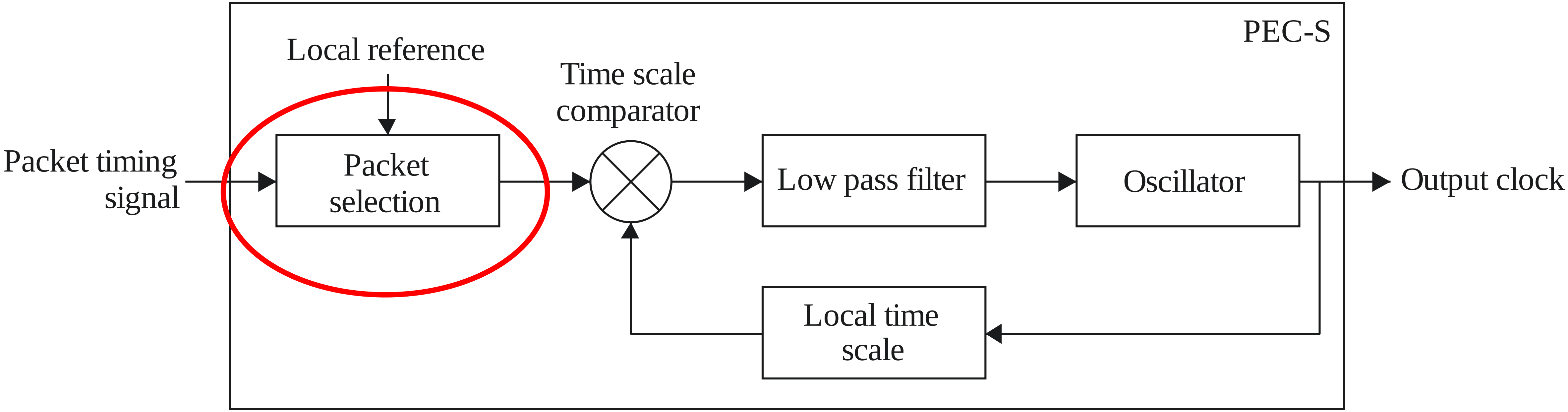
PACKET-BASED METHODS



From ITU-T Recc. G.8261

- Timing information carried by **dedicated timing packets**:
  - Network Time Protocol (NTP) – IETF RFC 5905
  - Precision Time Protocol (PTP) – IEEE1588-2008

PACKET-BASED EQUIPMENT CLOCK

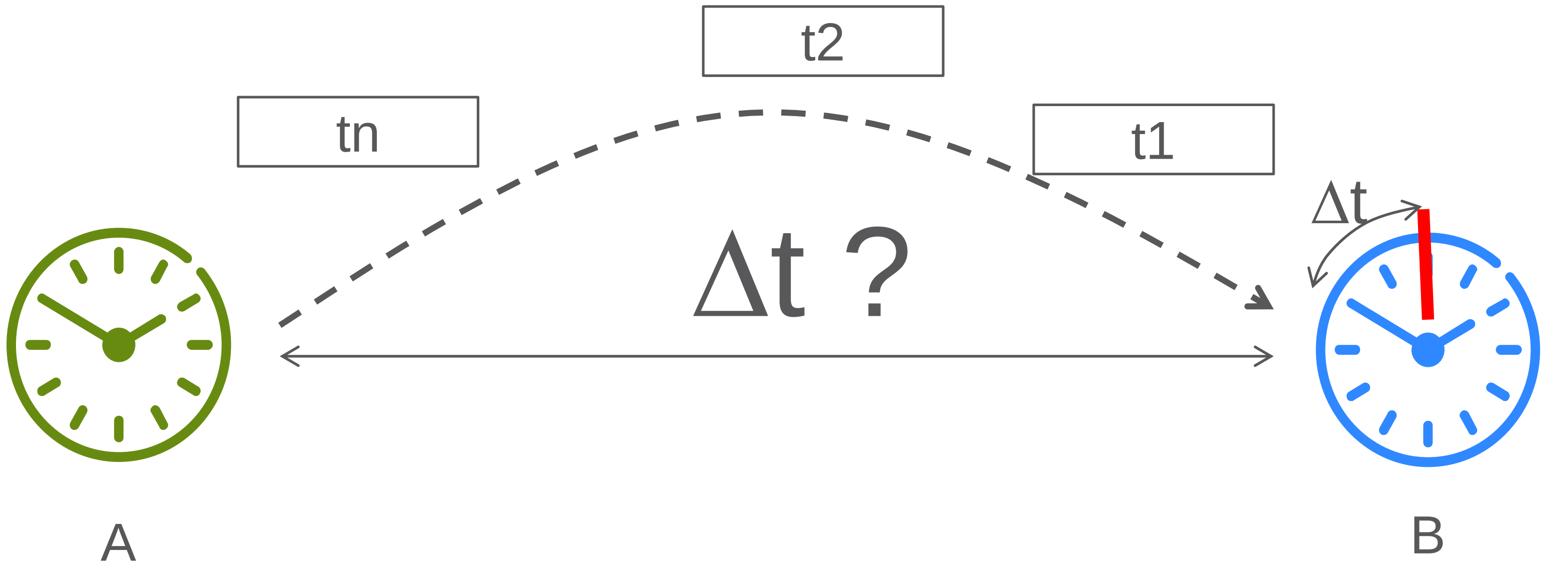


G.8263-Y.1363(12)_FA.1

- Concept of «**Packet Selection**»:
  - Pre-processing of packets before use in a traditional clock to handle PDV

TWO-WAYS TIME TRANSFER

- Delivery of Time synchronization requires also the knowledge of «transit delay» from A to B

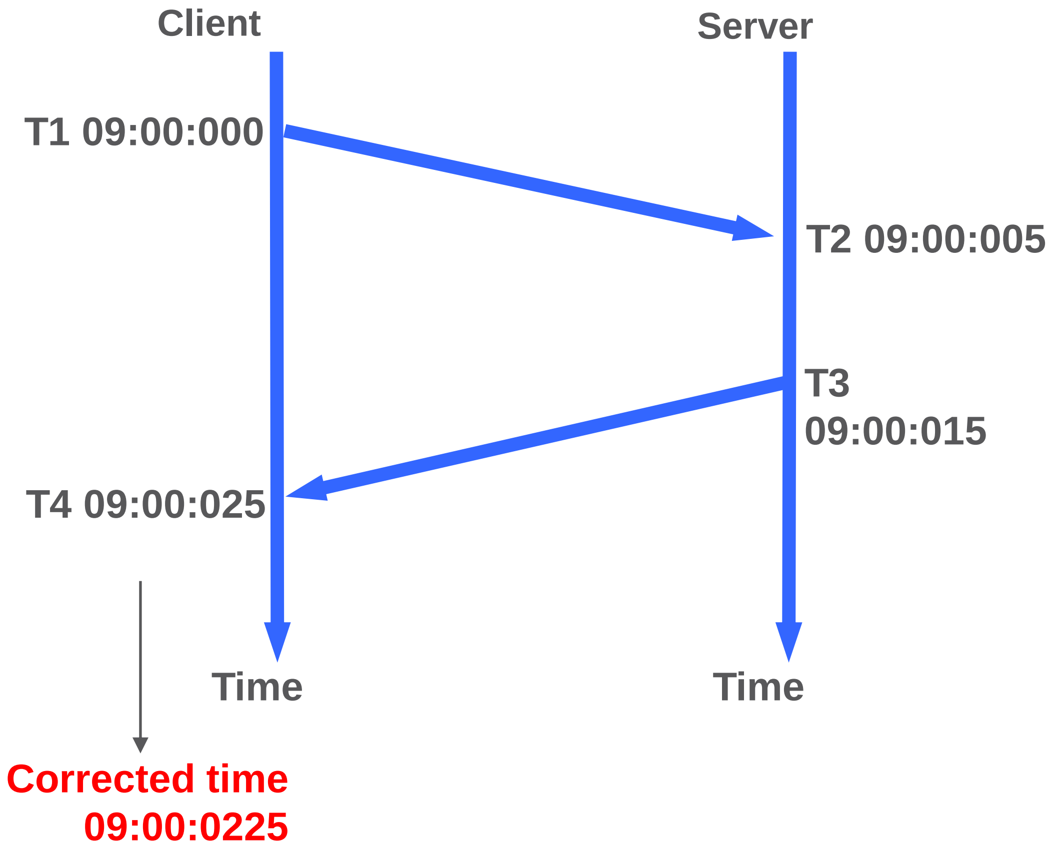


- **Two-ways transfer protocols** (round trip delay)
  - Assumption for **symmetric channel**

HOW NTP WORKS

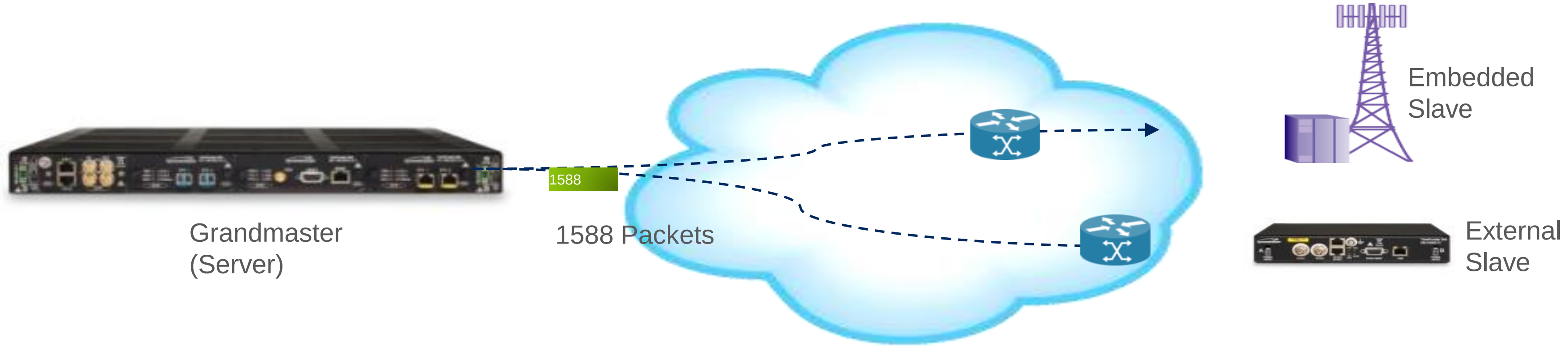
- T1 Originate Timestamp
  - Time request sent by client
- T2 Receive Timestamp
  - Time request received by server
- T3 Transmit Timestamp
  - Time reply sent by server
- T4 Destination Timestamp
  - Time reply received by client
- Round Trip Delay=(T4-T1)-(T3-T2)
  - Round Trip Delay =25-10=15
- Clock Offset= [(T2-T1)-(T4-T3)]/2
  - Clock Offset =[5-10]/2= **-2.5** (Clients actual time when reply received was therefore **09:00:0225**)

- **Key Assumptions:**
  - **One way delay is half Round Trip (symmetry!)**
  - Drift of client and server clocks are small and close to same value
  - Time is traceable



IEEE 1588-2008 PTPV2 OVERVIEW

- The **Grandmaster** “reference clock” sends a series of time-stamped messages to slaves.
- Slaves process the round-trip delay & synchronize to the Grandmaster.
- Frequency can be recovered from an accurate time of day reference (but **L1 can also be used** …)
- **Best Master Clock Algorithm** to define the hierarchy
- Accuracy is possible by means of:
  - Proper packet rate (up to 128 per second)
  - **Hardware time-stamping** (eliminate software processing delays)
  - **Timing support in the network** (e.g. transparent clocks, boundary clocks)

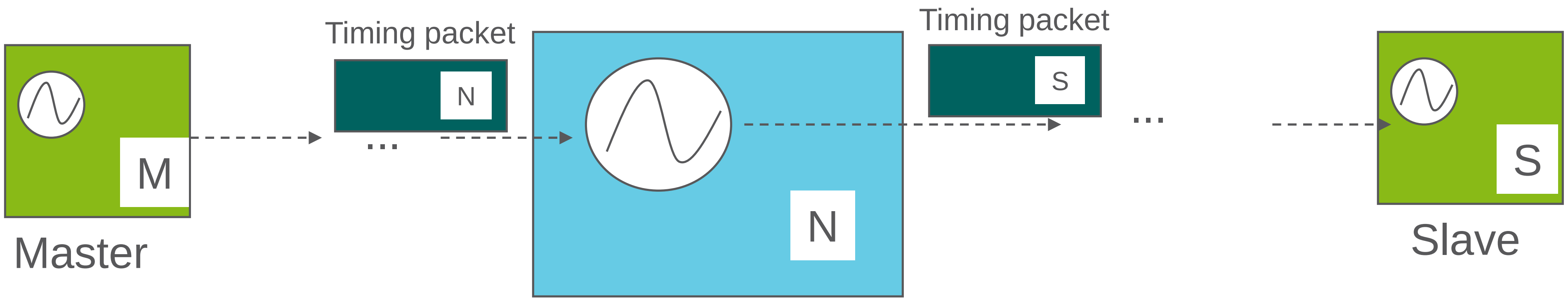


Packet Slave clocks can be either stand-alone or embedded in network equipment

- Note: IEEE 1588 under revision (planned 2017)

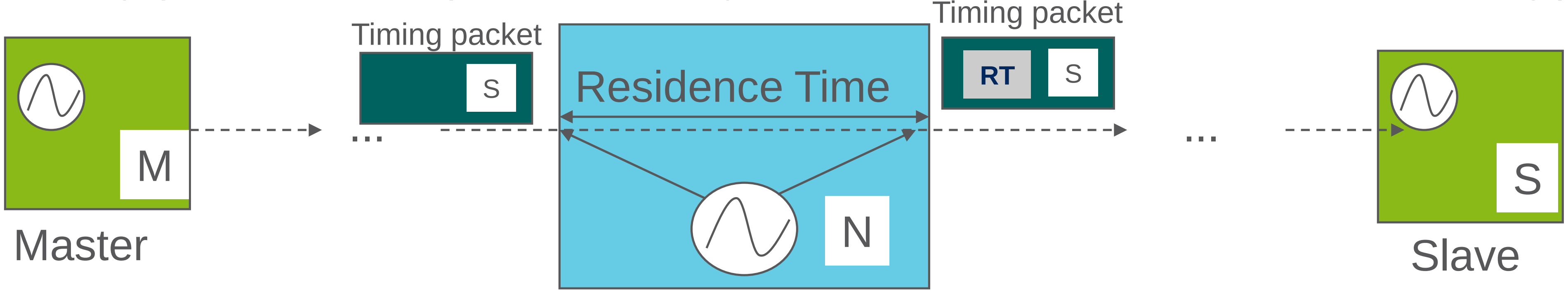
TIMING SUPPORT

Timing packets are terminated and regenerated by N



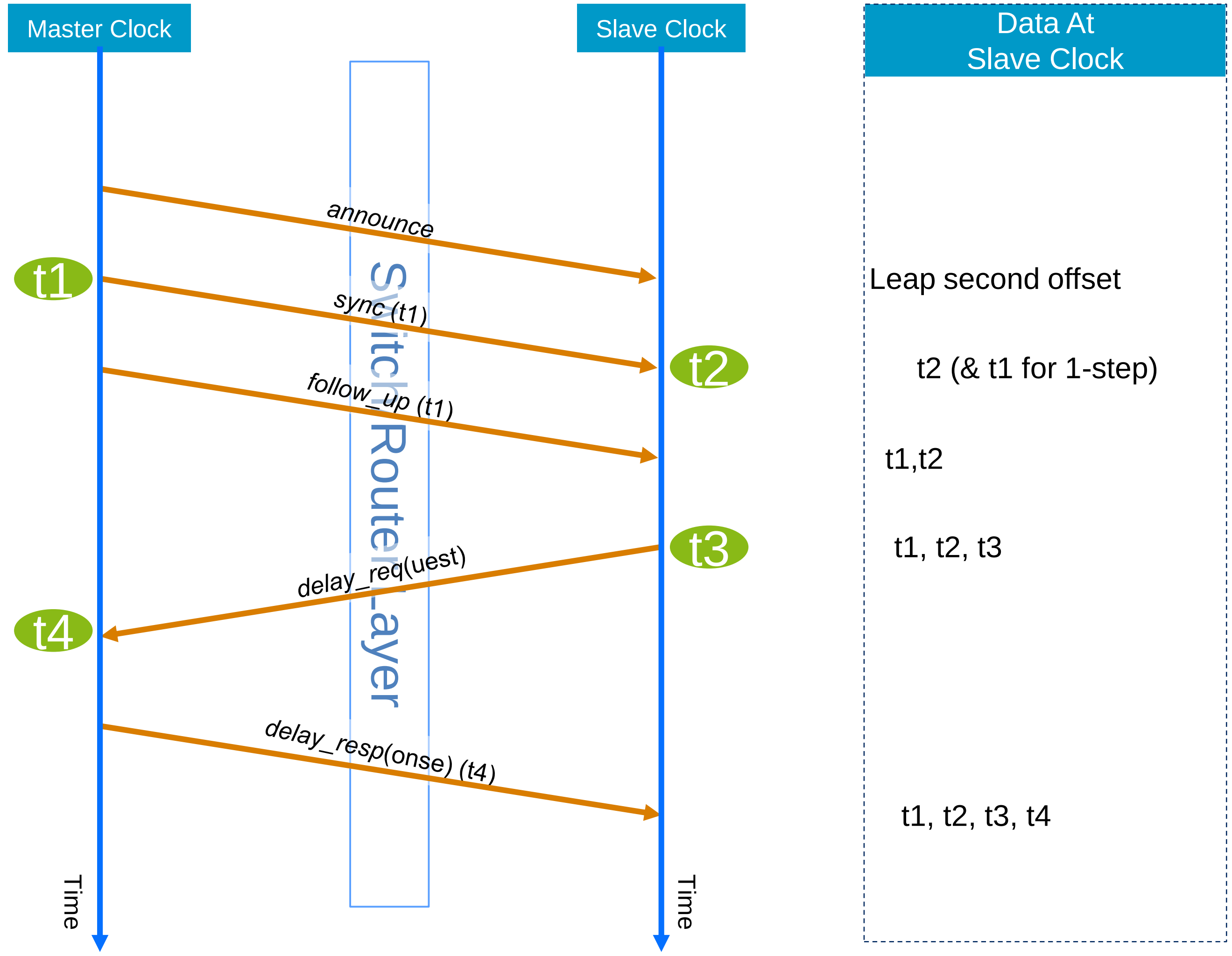
e.g. **IEEE1588 Boundary Clock**, NTP Stratum Clock

Latency (Residence Time) is calculated by NE and the information is added in the timing packet



e.g. **IEEE1588 Transparent Clock**

PTP TIME TRANSFER TECHNIQUE



The process is repeated up to 128 times per second.
(Announce rate is lower than Sync rate)

Round Trip Delay

$RTD = (t2 - t1) + (t4 - t3)$

Offset:

(slave clock error and one-way path delay)

$Offset_{SYNC} = t2 – t1$

$Offset_{DELAY\_REQ} = t4 – t3$

We assume path symmetry, therefore

One-Way Path Delay = RTD ÷ 2

Slave Clock Error = (t2 - t1) - (RTD ÷ 2)

Notes:

1. One-way delay cannot be calculated exactly, but there is a bounded error.
2. The protocol transfers TAI (Atomic Time).
   UTC time is TAI + leap second offset from the *announce* message.

“THE TELECOM PROFILE” (G.8265.N/G.8275.N)

- A **profile** is a subset of required **options**, prohibited options, and the ranges and defaults of configurable attributes
  - e.g. for Telecom: Update rate, unicast/multicast, etc.

- PTP profiles are created to allow organizations to specify selections of attribute values and optional features of PTP that, when using the same transport protocol, **inter-works** and achieve a **performance** that meets the requirements of a particular application

- Other (non-Telecom) profiles:
  - IEEE C37.238 (Standard Profile for Use of IEEE 1588 Precision Time Protocol in Power System Applications,)
  - IEEE 802.1AS (Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks); Under revision (targeting a full compliance with the next IEEE 1588 revision)

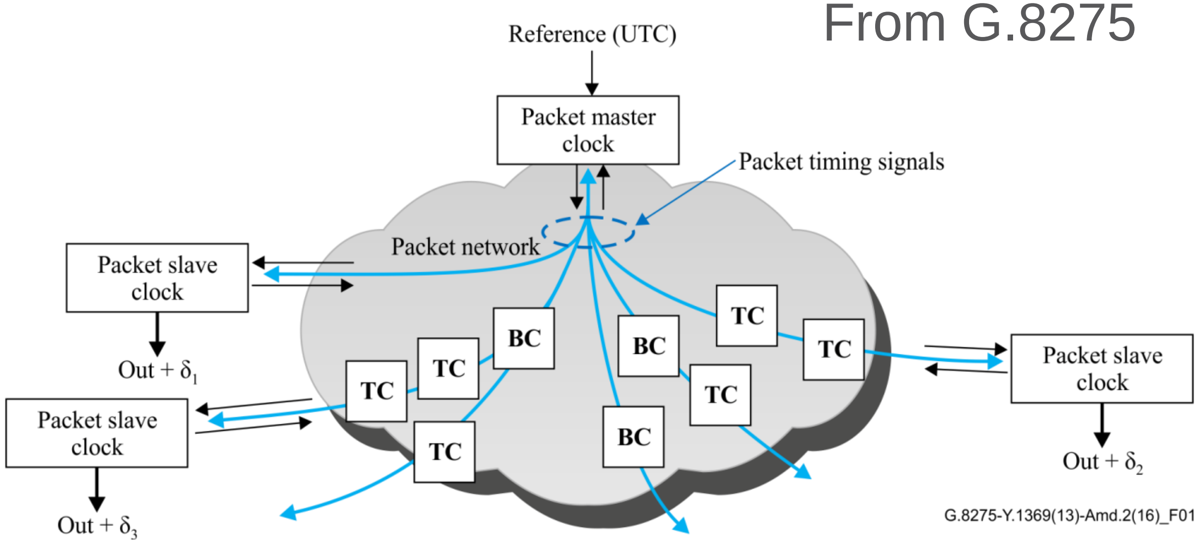
TIME SYNCHRONIZATION ARCHITECTURE

- General network topology for time/phase distribution from a packet master clock PRTC to a telecom time slave clock (T-TSC)

- The synchronization flow is from the master to slave, although the timing messages will flow in both directions.

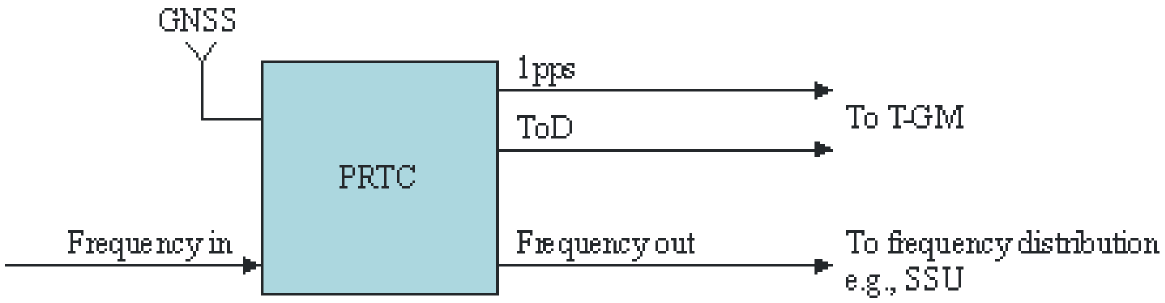
- Individual nodes are T-BCs or T-TCs in the case of full support from the network

From G.8275



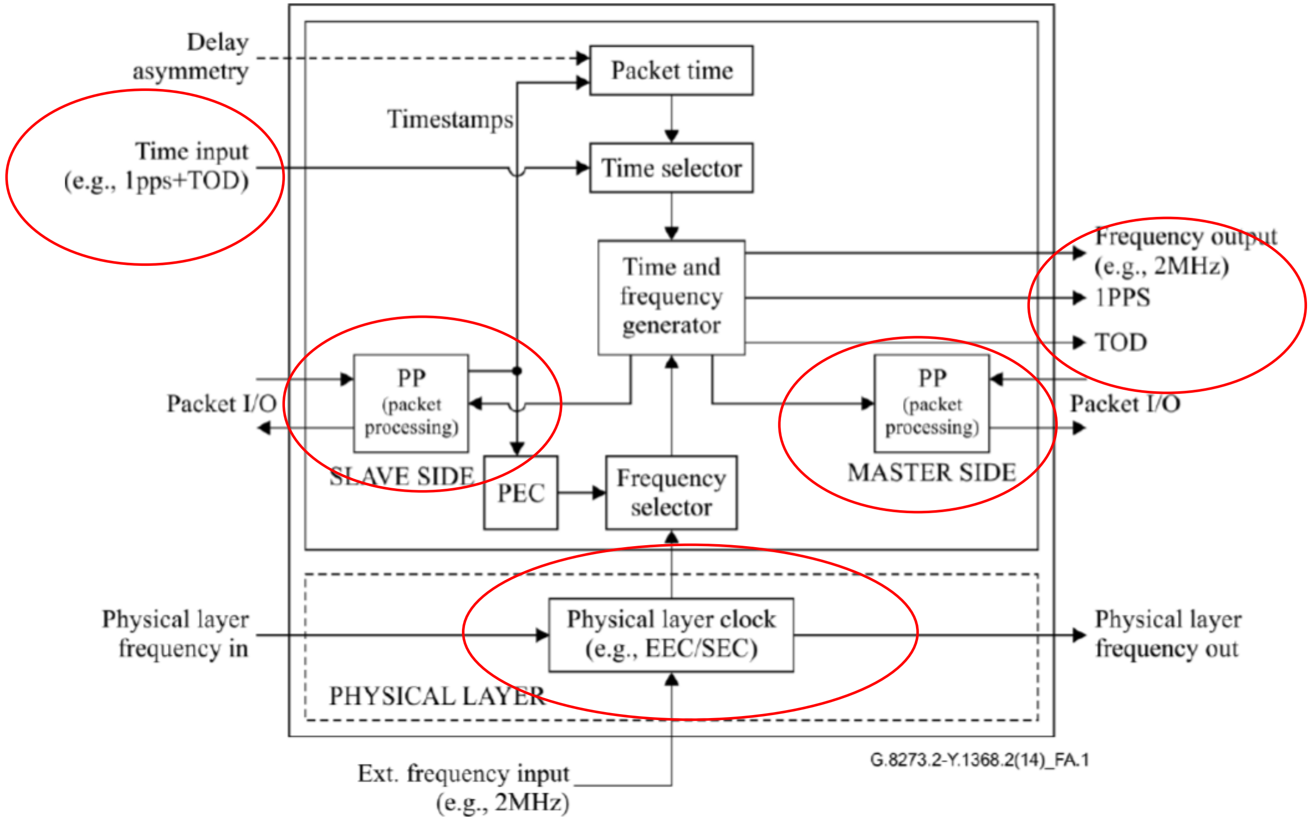
G.8275-Y.1369(13)-Amd.2(16)_F01

- Primary Reference Time Clock (PRTC) is the master of the time synchronization network

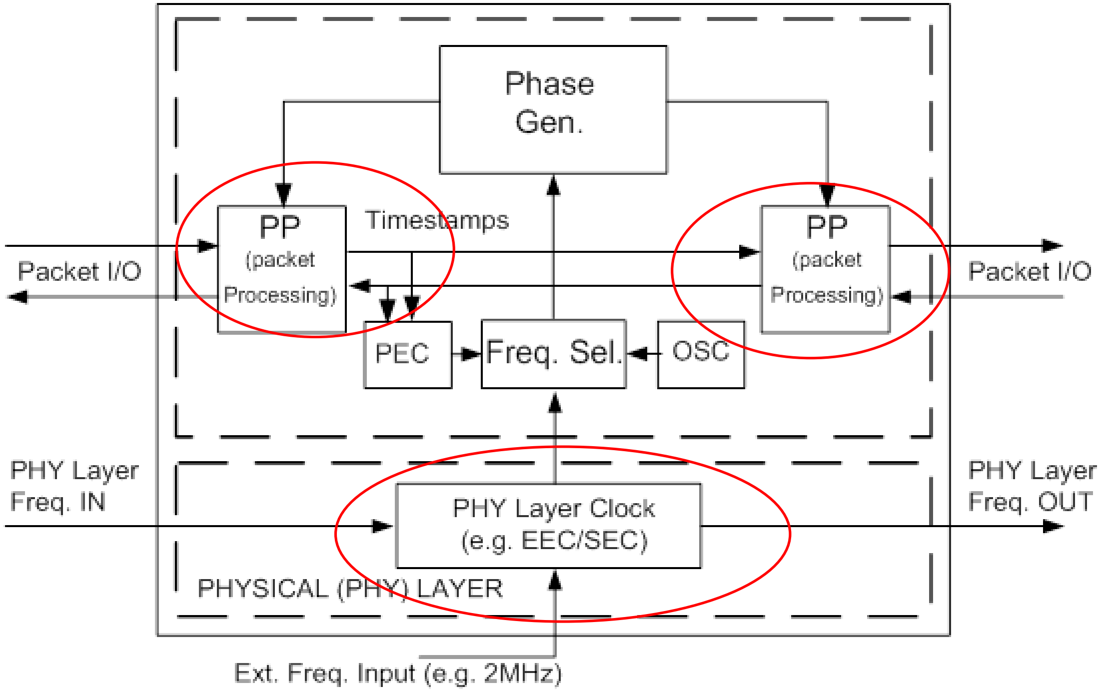


G.8275-Y.1369(13)_F07

T-BC AND T-TC CLOCK MODELS



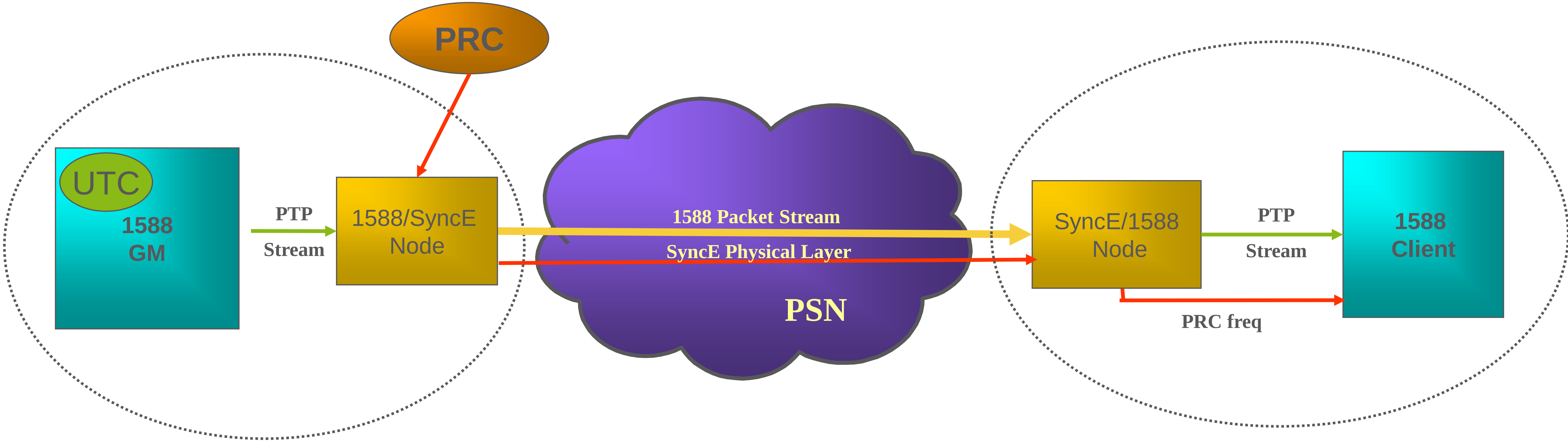
G.8273.2-Y.1368.2(14)_FA.1



- **G.8273.2** and **G.8273.3** provide models for the Telecom Boundary and Transparent Clocks
  - **Frequency sync via physical layer** initially considered

COMBINED PTP-SYNCE

- SyncE as “frequency assistance” to 1588



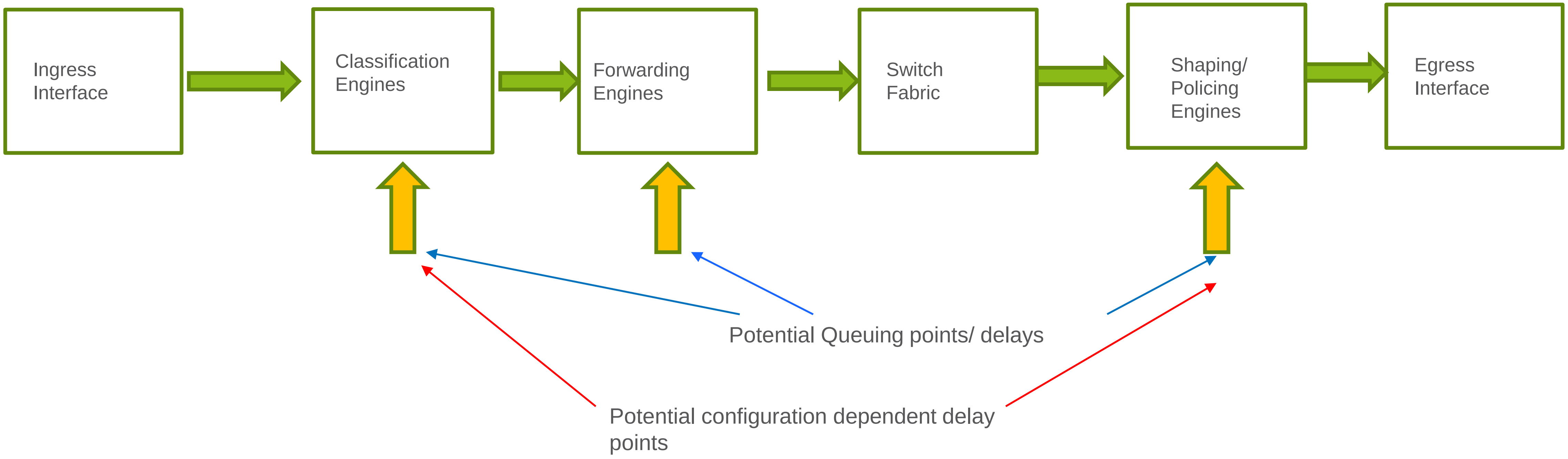
- Gives immediate “frequency lock” to 1588 client
- SyncE & 1588 functionality may be in the same node/element
- SyncE might be used for “**Time sync holdover**”

IMPAIRMENTS IN PACKET NETWORKS

- **Packet delay variations** [PDV], depending on
  - Network dimension
  - Traffic load
  - QoS
- Path dependent aspects
  - Physical path **asymmetry** (*particularly relevant for time synchronization*)
  - **Path rerouting**
- Interactions between the packet streams

PACKET DELAY VARIATION (PDV)

- Queuing
- Equipment Configuration
- Priority/ QoS



Ingress Interface

Classification Engines

Forwarding Engines

Switch Fabric

Shaping/ Policing Engines

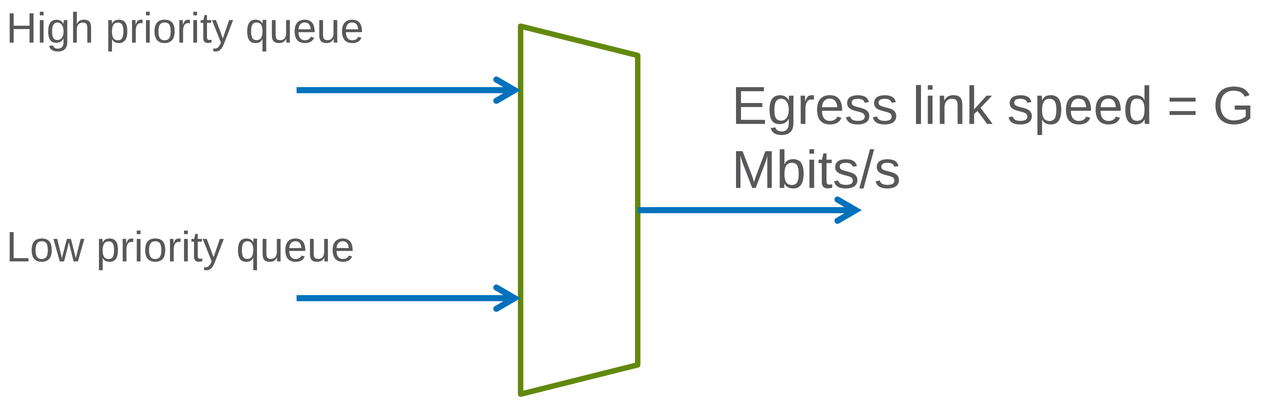
Egress Interface

Potential Queuing points/ delays

Potential configuration dependent delay points

PACKET DELAY VARIATION (PDV), CONT.

- Head of line blocking



High priority queue

Low priority queue

Egress link speed = G Mbits/s

MTU size M byte
Strict priority queue

$$\left(\Delta_{pp}\right)_{\max} \geq \left(\frac{M}{G}\right) \mu s$$

- A packet arrives in the HPQ, just when a packet from the LPQ has begun transmission
- The packet from HPQ is blocked till the LPQ packet is transmitted
- With more complex prioritization scheme the delay due to head of line blocking could vary significantly
- Tools being specified by IEEE 802.1 to address this issue (e.g. frame preemption, scheduled traffic)

Ex. : at 1Gbit/s,
1000 byte packet = 8 x 1000 / 1000 x 10^6 = 8 µs

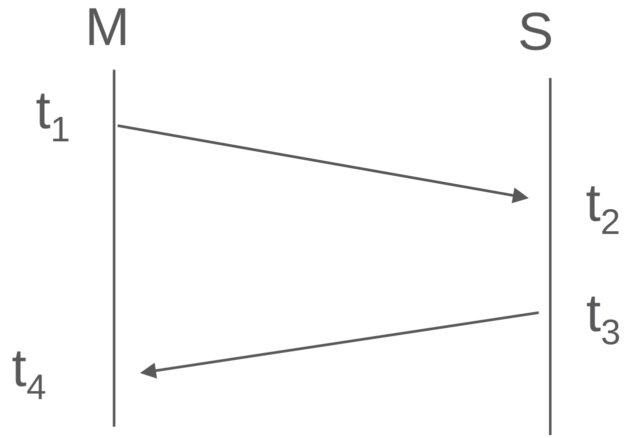
PATH DEPENDENT IMPAIRMENTS

- Asymmetry
  - Static difference in paths between the forward and reverse paths. E.g difference in lengths of fiber
  - Forward and reverse paths pass through different node
- Rerouting
  - Leads change in path delays and can “confuse” the algorithms.

TIME SYNCHRONIZATION VIA PTP: ASYMMETRY RELATED IMPAIRMENTS

- The basic principle is to distribute Time sync reference by means of two-way time stamps exchange

Time Offset= $t_2 - t_1$ – Mean path delay
Mean path delay = $((t_2 - t_1) + (t_4 - t_3))/2$



- As for NTP, also in case of PTP, symmetric paths are required:
  - Basic assumption: $t_2 - t_1 = t_4 - t_3$
  - Any asymmetry will contribute with half of that to the error in the time offset calculation (e.g. 3 μs asymmetry would exceed the target requirement of 1.5 μs)

ASYMMETRY DUE TO THE TRANSPORT TECHNOLOGIES

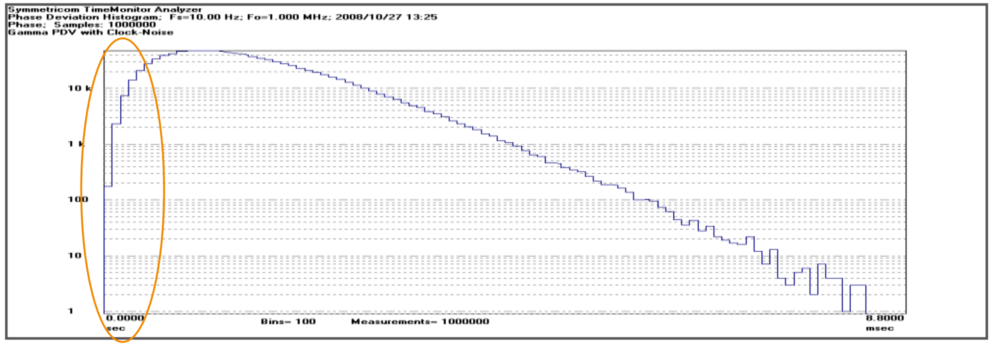
- Different paths in Packet networks
  - Traffic Engineering rules in order to define always the same path for the forward and reverse directions
- Different Fiber Lengths in the forward and reverse direction
  - Additional problem: DCF (Dispersion Compensated Fiber)
- Different Wavelengths used on the forward and reverse direction
- Asymmetries added by specific access and transport technologies
  - GPON
  - VDSL2
  - Microwave
  - OTN

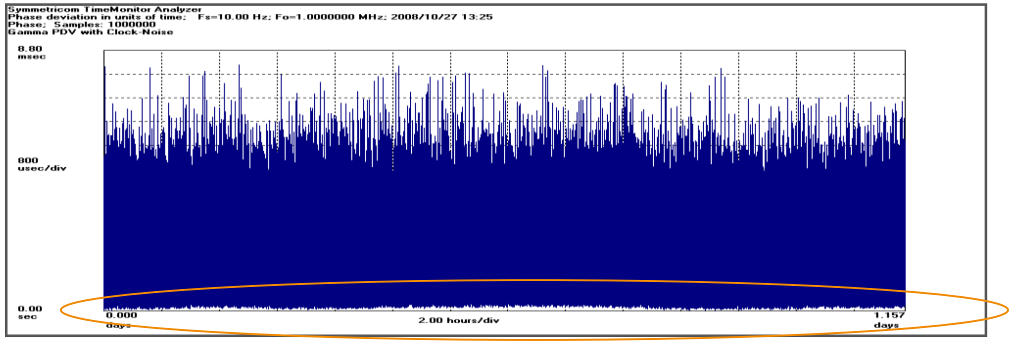
KEY ASPECTS OF PERFORMANCE

- **Packet Delay Variation** (PDV) is a major contributor to “clock noise”
  - Related to number of hops, congestion, line-bit-rate, queuing priority, etc. Time-stamp-error can be viewed as part of PDV
- Clock recovery involves low-pass-filter action on PDV
  - **Oscillator** characteristics determine degree of filtering capability (i.e. tolerance to PDV)
    - Higher performance oscillators allow for longer time-constants (i.e. stronger filtering)
    - Lower performance (less expensive) oscillators may be used (may require algorithmic performance improvements)
- Performance improvements can be achieved by
  - Higher packet rate
  - Controlling PDV in network (e.g. network engineering, QoS)
  - Timing support from network (e.g. *boundary clocks* in PTP)
  - Packet selection and/or nonlinear processing

NOTION OF “BEST PACKETS”

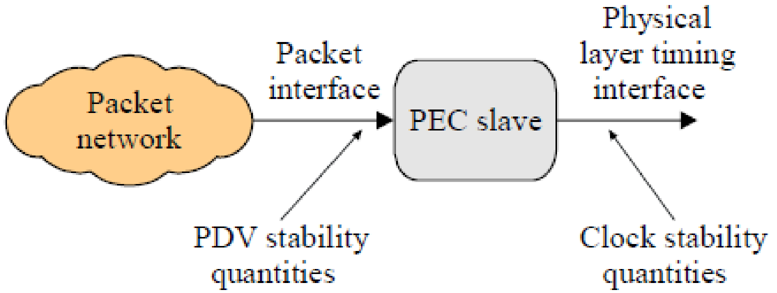
- Impact of PDV can be mitigated by means of a suitable classification and **selection** of packets
- The “**minimum delay**” approach is presented as an example. Depending on the network characteristics other approaches may be more suitable
- The assumption that the path is constant over the interval of observation implies a PDV with a distribution function with a slowly changing floor (i.e. minimum delay that a packet can experience)
- In many cases it has been observed that a reasonable **fraction (e.g. x%) of the total number of packets** will traverse the network at or **near this floor**
- Using only these packets in the timing recovery mechanism would allow to significantly reduce the impact of the PDV on the quality of the recovered reference timing signal





SYNC METRICS IN PACKET NETWORKS

- The Network Element clock output metrics as per TDM networks (e.g. MTIE/MRTIE/TDEV)
  - Some distinctions are required in case of packet clock integrated in the Base Station (no standardized output MTIE/TDEV by 3GPP)
- Specific Metrics have been defined to better characterize the behavior of packet networks (PDV) delivering the timing reference
  - Metrics that associate PDV with Frequency Offset or phase variation
  - Tolerance masks/Network limits are used by network operators and clock manufacturers
  - Packet selection methods can be justified



[*Clock stability quantities estimation*] = *function* (*PDV stability quantities*)

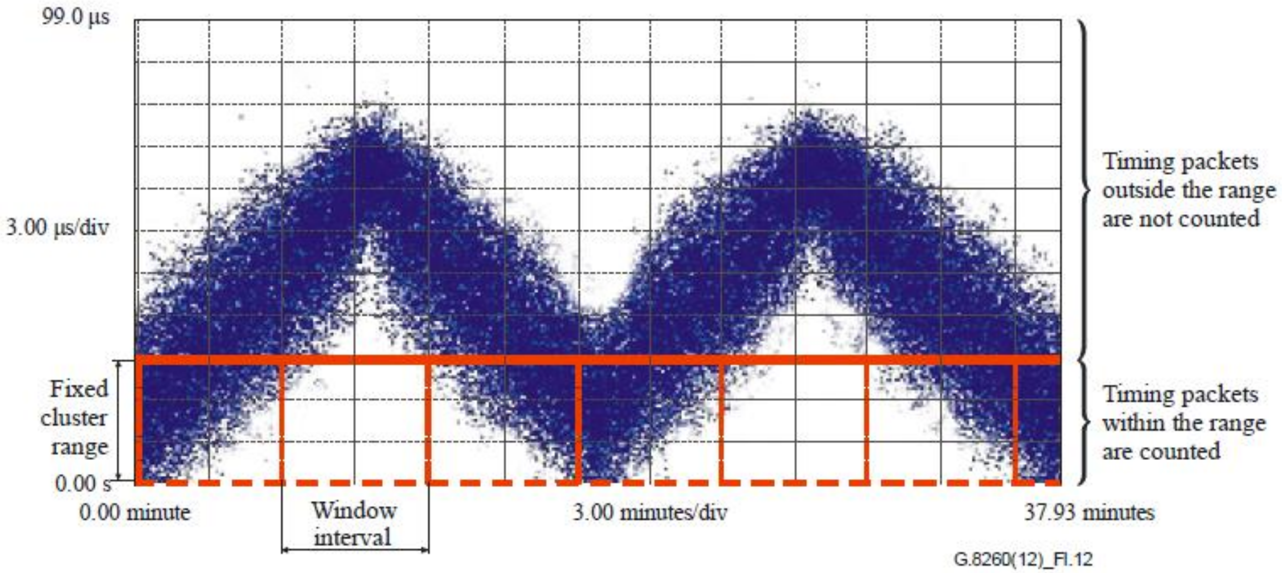
G.8260(10)_FI.1

PDV METRICS

- Several metrics have been defined by ITU-T:
  - These include minTDEV, MATIE, MAFE, percentileTDEV, bandTDEV, clusterTDEV, FPP, etc.
- minTDEV is analogous to TDEV
  - TDEV utilizes the average over windows
  - minTDEV utilizes the minimum over windows
- MATIE is related to MTIE
  - MTIE computed directly on the time error sequences $\{x_k\}$ or $\{y_k\}$ is not that meaningful because of large “jitter” (PDV)
  - MATIE is computed on the sequence following the pre-filtering (packet-selection) and emulates the low-pass nature of the traditional clock model (bandwidth / time-constant)
- Metrics Studying floor delay packet population
  - **FPP, Floor Packet Percent** (selected for defining network performance objectives for frequency sync)

FLOOR PACKET PERCENTAGE

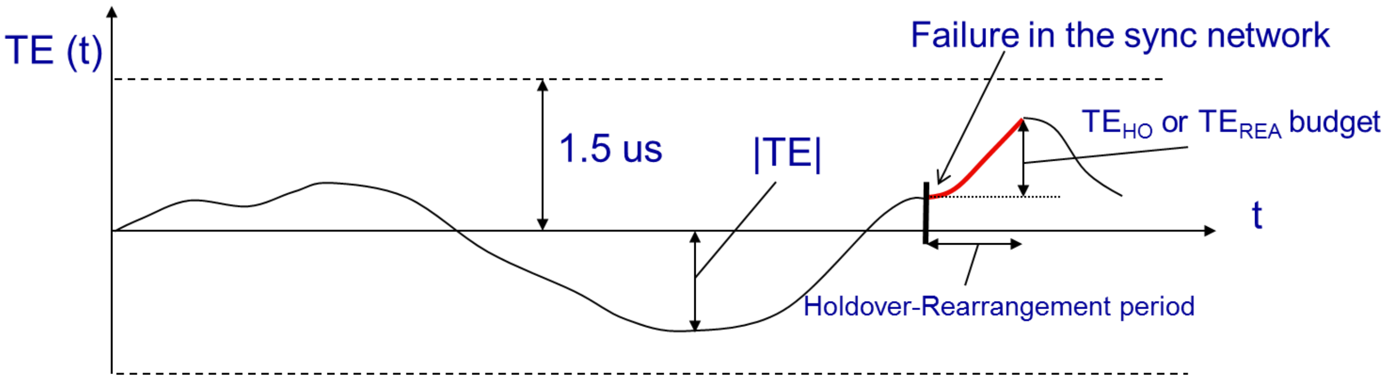
- Family of metrics based on counting amount of packets, observed for any window interval of t seconds within a fixed cluster range starting at the observed floor delay and having a size δ



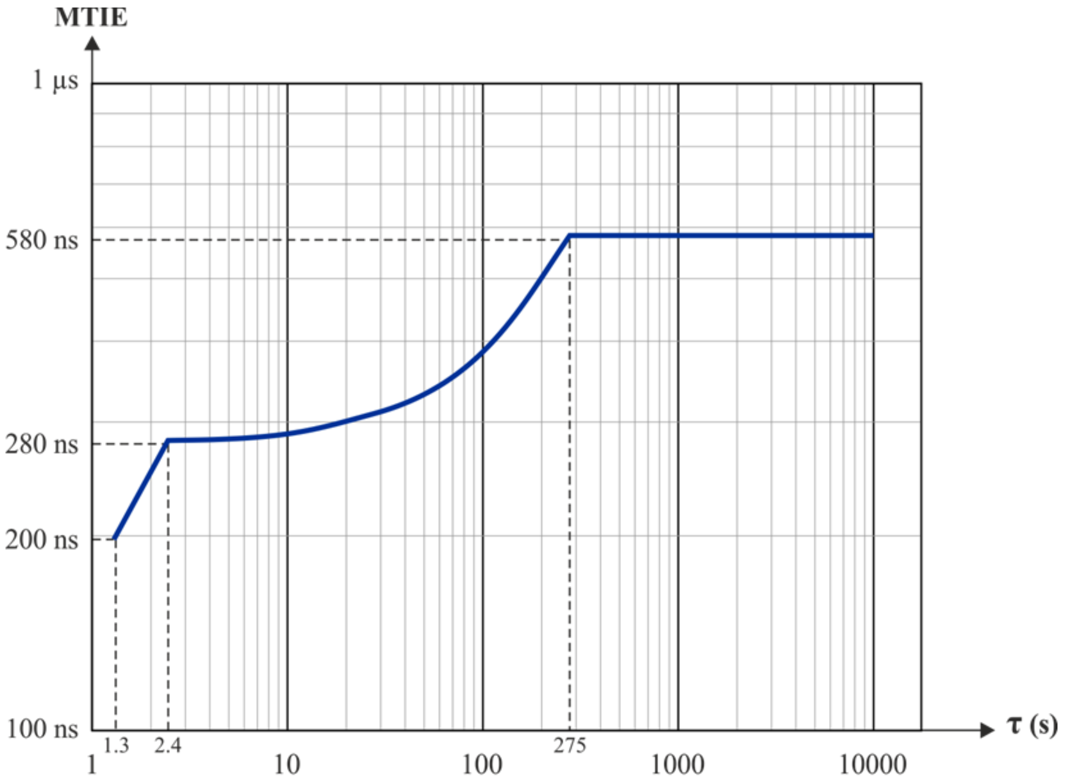
- Floor Packet Percent (FPP) defined in terms of percentage of packets meeting these criteria
- Basis for the G.8261.1 network limits (150 / 75 us)

TIME SYNC PERFORMANCE METRIC: FULL TIMING SUPPORT

- Max abs(TE) for combined **dynamic** and **constant time error**
- MTIE (low frequency) and «peak-to-peak TE amplitude» (high frequency) for dynamic time error



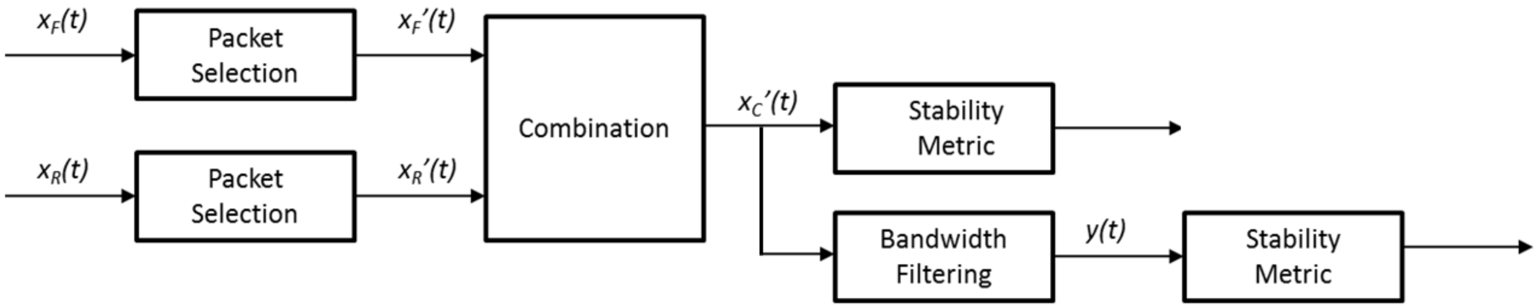
TE_{HO} applicable to the network (End Application continues to be locked to the external reference)
TE_{REA} applicable to the End Application (End Application handles short rearrangement periods)

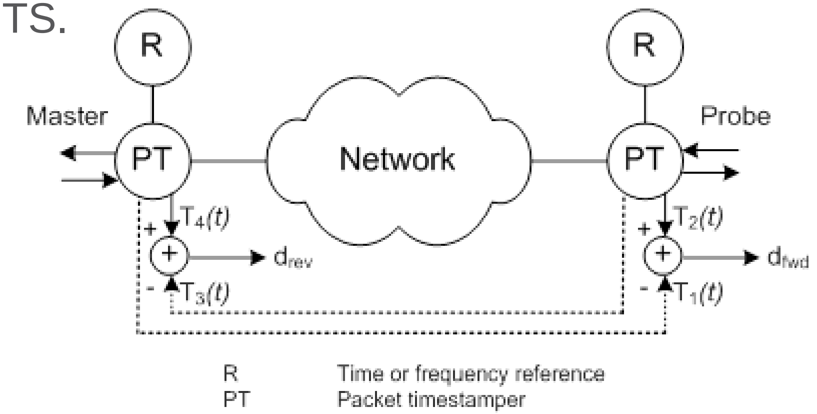


G.8271.1-Y.1366.1(13)-Amd.2(14)_F02

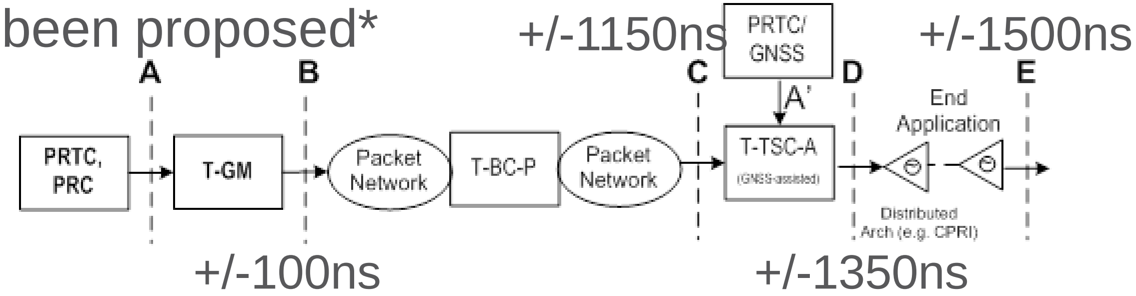
TIME SYNC PERFORMANCE METRIC: PARTIAL TIMING SUPPORT

- Metric agreed : «**Packet selected 2WayTE**»
  - Packet Selection criteria ;
    window interval of 200s/100s - percentage of 0.25%/0.5% as initial assumptions for high/normal stability clocks
  - Applicable to both «APTS» (Assisted Partial Timing Support) and «PTS»:
    - Peak-to-peak pktSelected2WayTE for APTS, max |pktSelected2WayTE| for PTS.





- 2 classes of network limits addressing different end applications cases.
- APTS Network Limit: 1.35µs in terms of maximum absolute time error (at the output of the clock).
- PTS Network Limits: pktSelected2WayTE; values have been proposed*
- MTIE mask also needed ?



*G.8273.4 tolerance should be the same for PTS and APTS

REFERENCES

- Packet Timing in ITU-T: ITU-T G.826x series, G.827x series,
- ITU-T general definitions: G.810, G.8260
- NTP: IETF RFC 5905/6/7/8
- PTP: IEEE 1588-2008
- CES: RFC 5087, RFC 5086, RFC4533, ITU-T Y.1413, ITU-T Y.1453, MEF3, MEF 8

ERICSSON