TIME APPLIANCES (TAP)



Bounding Time Uncertainty in Distributed Systems

Kishan Shenoi *Sr. Principal System Architect, SiTime*

Gary Giust *Sr. Director System Architecture, SiTime*

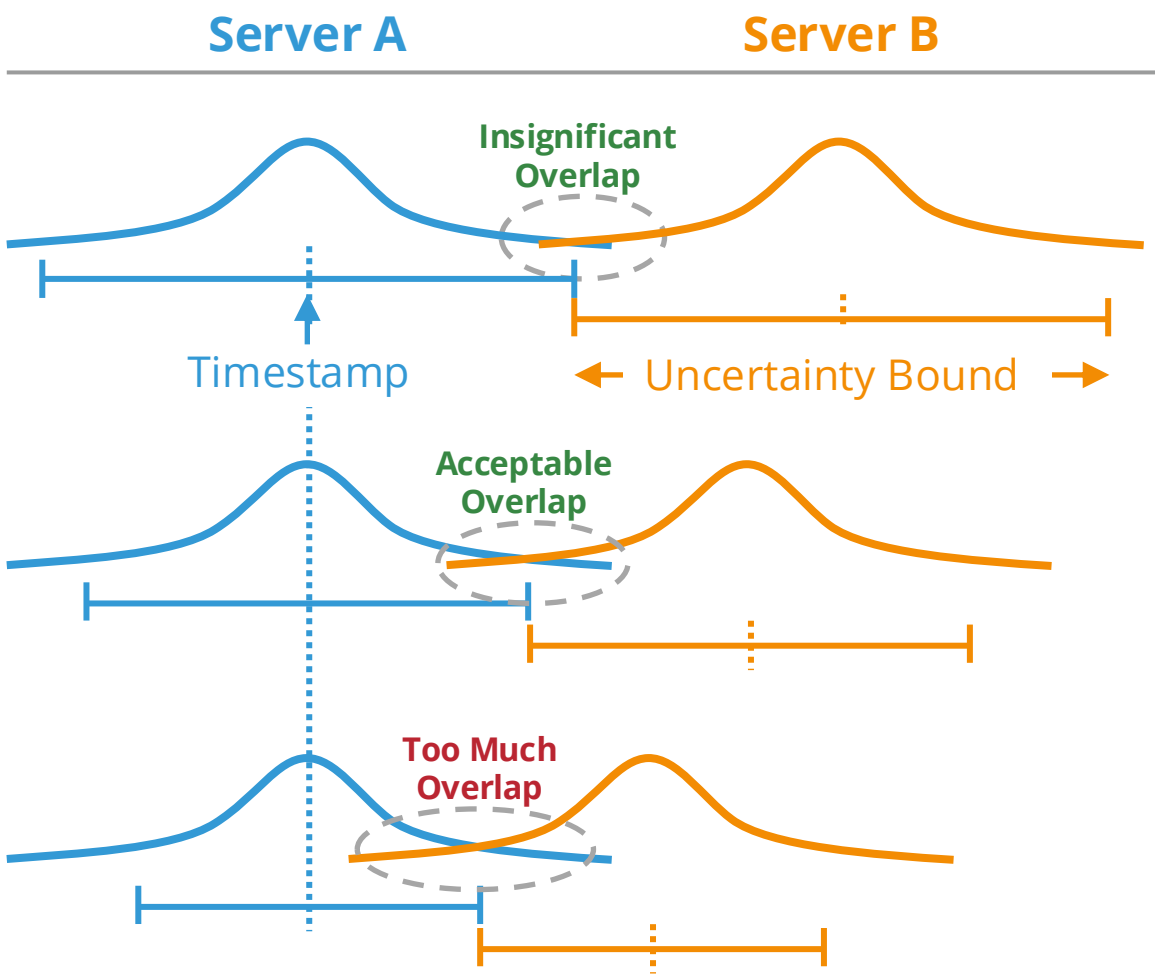


Distributed Systems Bound Time Uncertainty to Prevent Silent Errors

Too Conservative
Bound is wider than needed

Right Sized
Bound matches the true uncertainty

Too Optimistic
Bound is tighter than the true uncertainty



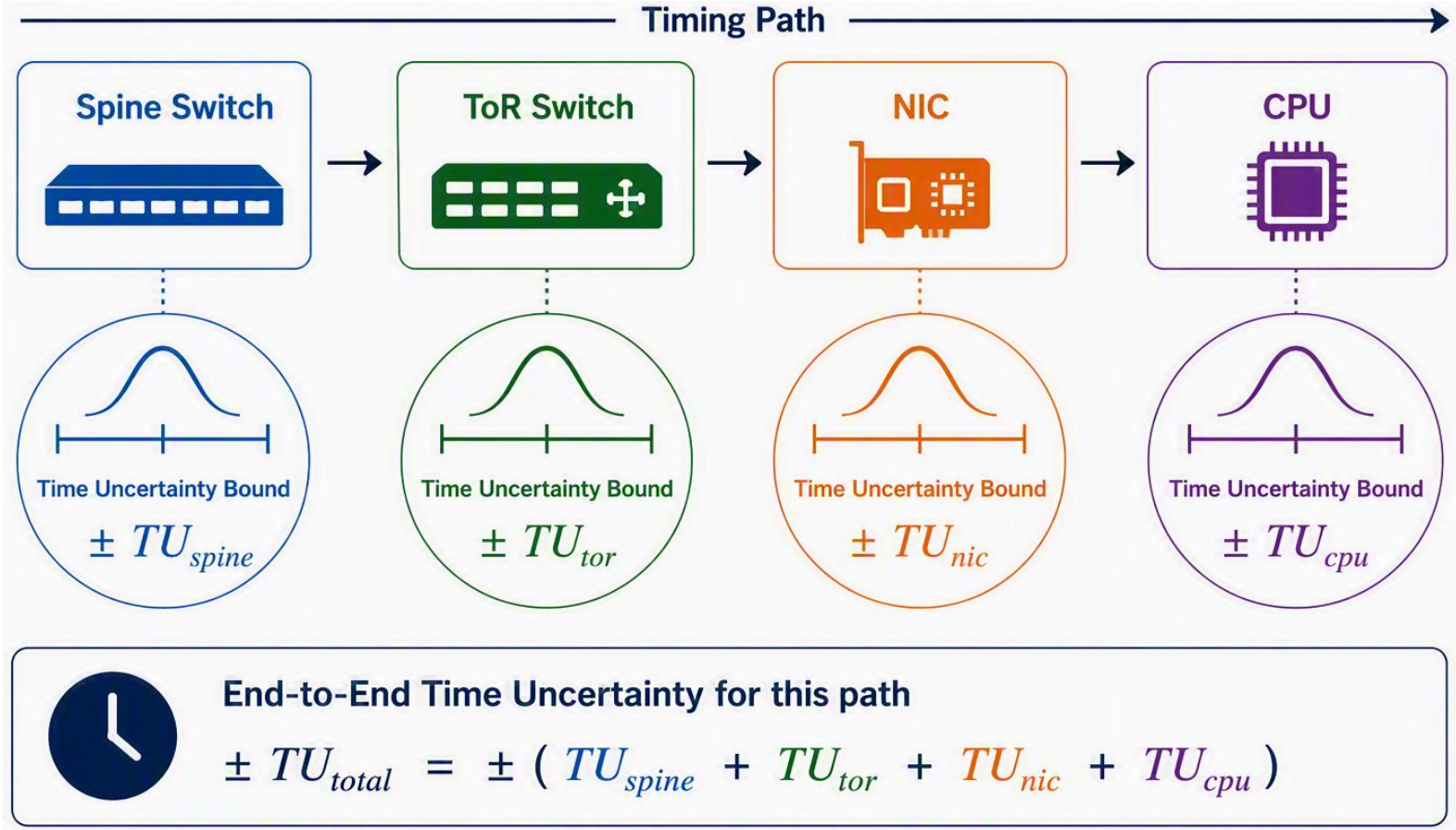
Wasted Time
Waiting out uncertainty reduces efficiency

Efficient & Safe
Best efficiency without silent data errors

Silent Data Errors
Probability of error causes wrong ordering with no error reported

Goal: Quantify & Optimize Time Uncertainty at Network Level

End-to-end time uncertainty is a combination of per-node bounds along the path



Time is Critical Metric, not Frequency

- PTP synchronizes.
  The data center bounds time uncertainty.
  - To ensure no silent data errors, the operator must establish a defensible end-to-end bound across oscillator behavior, residual asymmetry, timestamping, servo response, environmental effects, and other worst-case implementation errors.

- Start by bounding uncertainty in oscillator datasheet
  - **Today**: Difficult to bound uncertainty from datasheet specs
  - **Future**: Total Time Uncertainty (TTU) – a single spec for synchronization applications

Frequency Stability					
Frequency Stability over Temperature	F_stab	–	–	±0.1	ppm
Initial Tolerance	F_init	–	–	±0.3	ppm
Supply Voltage Sensitivity	F_Vdd	–	±0.5	±2.5	ppb
Output Load Sensitivity	F_load	–	±0.05	±0.4	ppb
Frequency vs. Temperature Slope	ΔF/ΔT	–	±0.9	±2	ppb/°C
		–	±1	±3.5	ppb/°C
		–	±0.9	±3.3	ppb/°C
Dynamic Frequency Change during Temperature Ramp	F_dynamic	–	±0.008	±0.02	ppb/s
		–	±0.01	±0.03	ppb/s
		–	±0.008	±0.028	ppb/s
24-hour holdover stability	F_24_Hold	–	–	±0.15	ppm
Hysteresis Over Temperature	F_hys	–	±25	±42	ppb
		–	±15	±27	ppb
		–	±10	±20	ppb
One-Day Aging	F_1d	–	±0.5	±2.0	ppb
One-Year Aging	F_1y	–	±57	±230	ppb
5-Year Aging	F_5y	–	±73	±320	ppb
10-Year Aging	F_10y	–	±80	±360	ppb
20-Year Aging	F_20y	–	±87	±400	ppb
Allan deviation	ADEV	–	1.5e-11	–	–

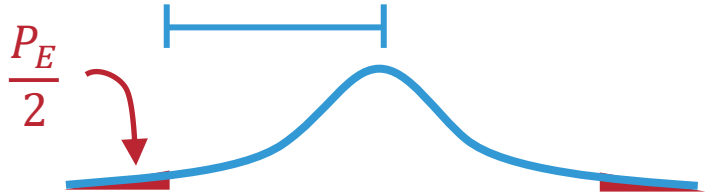
Total Time Uncertainty	TTU	–	±1e-9	±3e-9	ns

Total Time Uncertainty (TTU) Must Account for Deployed Fleet of Oscillators

Deterministic Component

Random Component

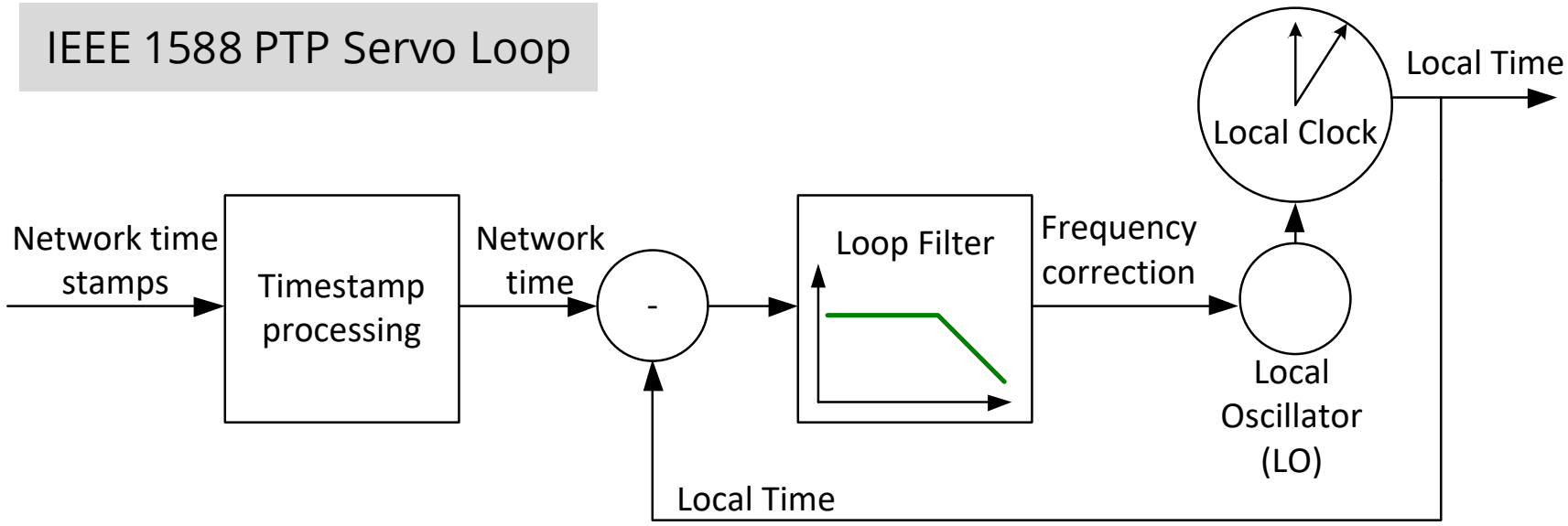
$$\pm TTU(P_E) = TU_{Peak} + Q(P_E) \times TU_{RMS}$$

Distribution		
Noise Source	• Variation in temperature, voltage, load, process • Aging	• Allan Deviation
Deployment	• Thermal profile • Airflow profile	• Constant temperature • Acceptable annual probability of error, P_E • # of oscillators deployed, N

$P_E = \mathrm{N} \times \mathrm{erfc}(Q/\sqrt{2})$ Solve for Q iteratively

TTU Calculation – Account for Location of Oscillator in Network

- Transparent clock (switch) – Oscillator TTU contribution adds linearly to network uncertainty
- Ordinary clock (NIC) – Oscillator TTU contribution is high-pass filtered by servo loop



Oscillator Datasheet Specifies TTU for Data Center Statistics

Example Data

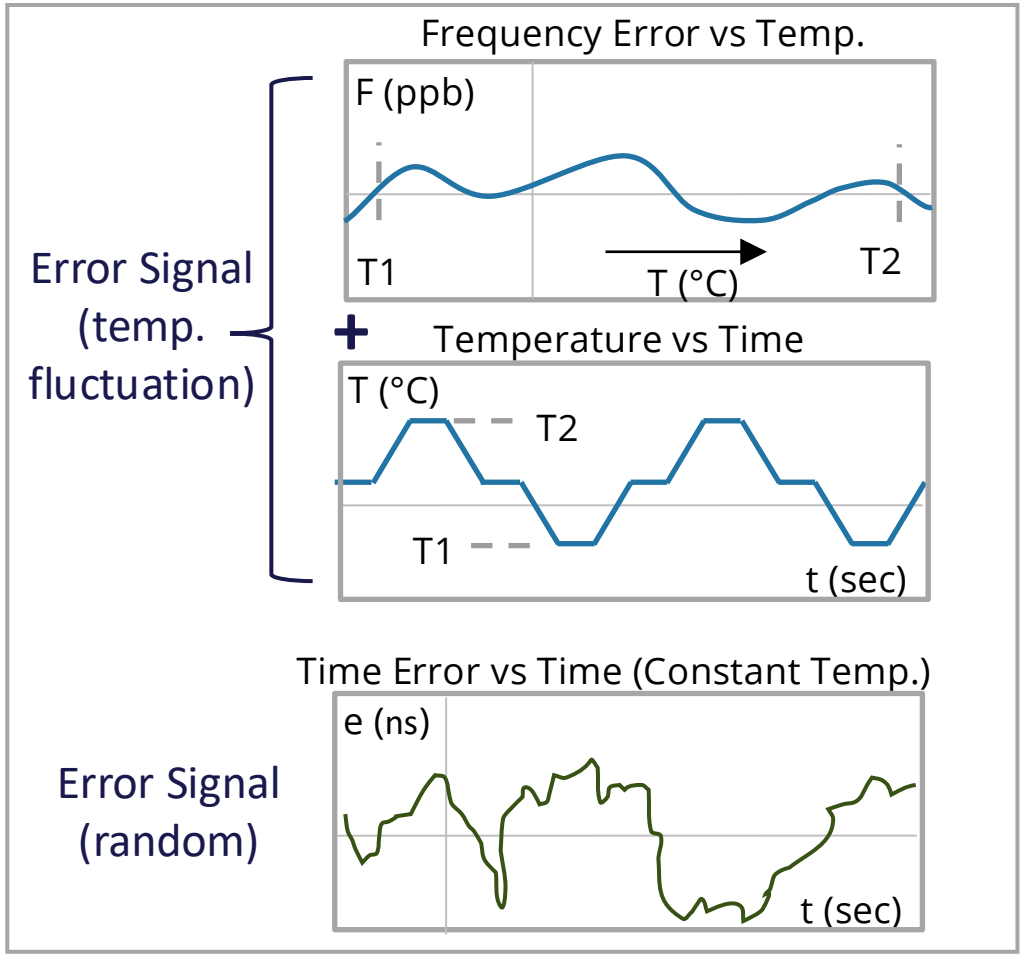
Time Uncertainty (TU)	Symbol	TU for a given servo bandwidth			Condition
		0.1 Hz	0.01 Hz	0.001 Hz	
Random component	TU_{RMS}	0.13 ns	2.6 ns	82 ns	Ambient temperature within ±1 C
Deterministic component	TU_{Peak}	17 ns	181 ns	1856 ns	Any thermal profile up to 6 C/min over -40 C to 105 C
Total TU[1]	TTU_{OSC}	$TU_{Peak} + TU_{RMS} \times Q$			Q is the "crest factor" that converts a rms to a peak value for probability P_E

Notes

1. Solve Q iteratively for an acceptable annual probability of error, P_E, and deployed fleet of N oscillators: $P_E = N \times erfc(Q/\sqrt{2})$
2. TU_{RMS} and TU_{Peak} values derived from common oscillator measurements

Time Uncertainty (TU) from *Measurement Analysis*

MEASUREMENT (Sample size = M devices)



SIMULATION (for each of k = 1,2,…M devices)

- Simulate Servo for temperature error and for random error
- BW choices: 0.1Hz, 0.01Hz, 0.001Hz
- For each choice of BW establish:
  - Temperature effect: $P(k) = |TE|_{MAX}$
  - Random behavior: $S(k) = |TE|_{RMS}$

STATISTICS

- Statistically extend results from sample size of M to whole population
- For each choice of BW:
  - Temperature: $TU_{Peak} \leftarrow \{P(k)\}$
  - Random component: $TU_{RMS} \leftarrow \{S(k)\}$

Implementations to Reduce TTU

- Physical layer synchronization
- PTP or 1 PPS time transfer
- XO → TCXO → OCXO (lower noise oscillators)
- EPPS for in-rack and in-server phase distribution

Standardize TTU Framework – Components & Equipment

- Challenge – Reduce guardband. Increase efficiency. Preserve correctness.
- TTU framework is new, which relies on existing test methods
- Contribute
  - Questions: Standardize TTU in OCP? Typical or maximum values? Thermal profiles? etc.
  - Offer feedback: ggiust@sitime.com, kshenoi@sitime.com
- Get Involved
  - Project Wiki: https://www.opencompute.org/w/index.php?title=Time_Appliances_Project
  - Workstream Wiki: https://www.opencompute.org/w/index.php?title=TAP_Oscillators
  - Mailing list: https://ocp-all.groups.io/g/OCP-TAP

Thank You!

