

EffiTest: Efficient Delay Test and Statistical Prediction for Configuring Post-silicon Tunable Buffers

Grace Li Zhang, Bing Li, Ulf Schlichtmann
Institute for Electronic Design Automation
Technical University of Munich (TUM)

Overview

Motivation

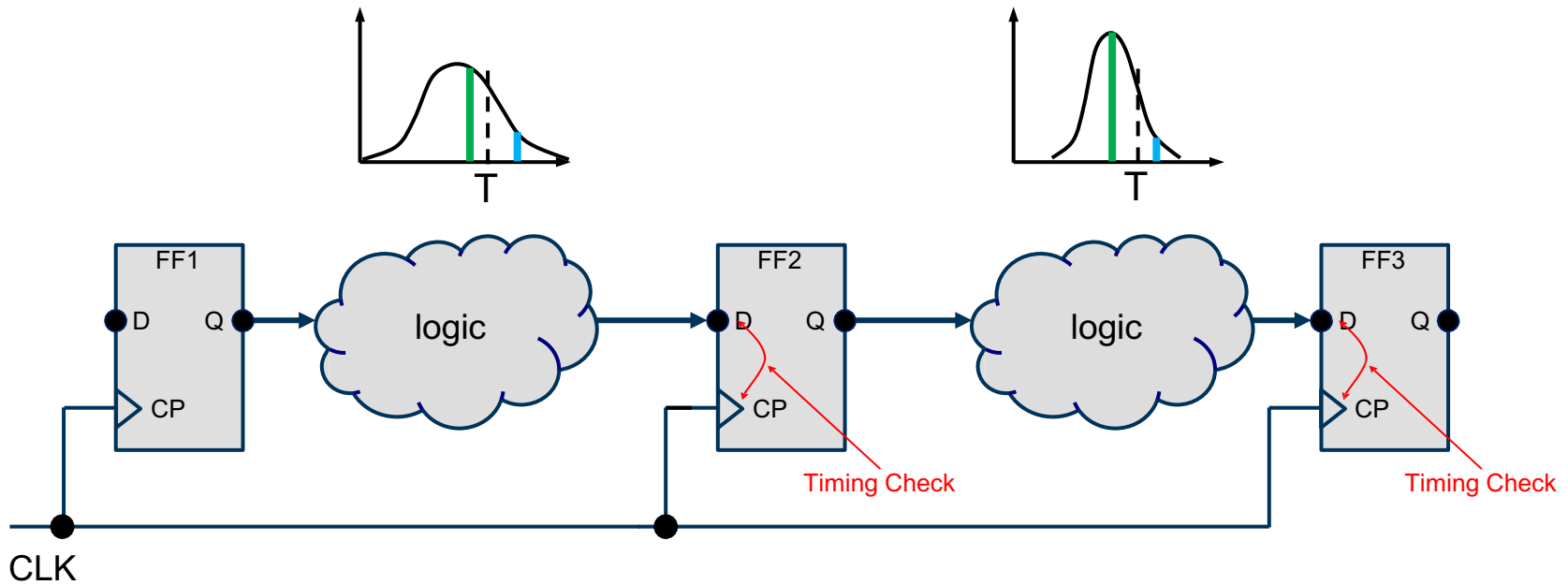
Post-silicon delay test and buffer configuration

- Delay alignment and statistical prediction
- Post-silicon configuration

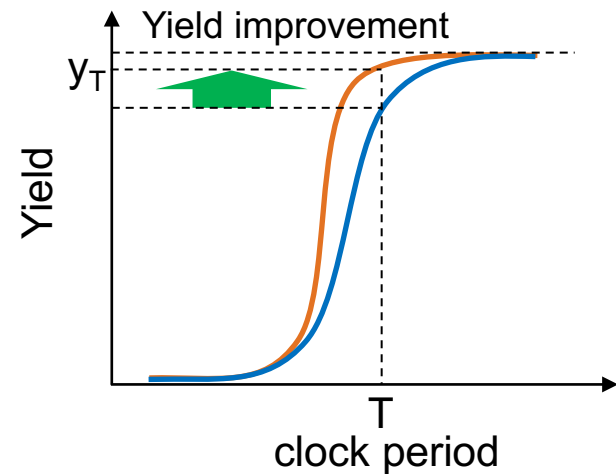
Experimental results

Summary

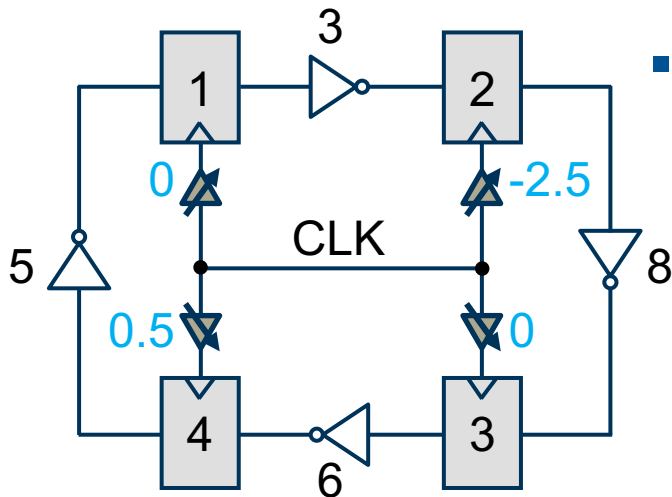
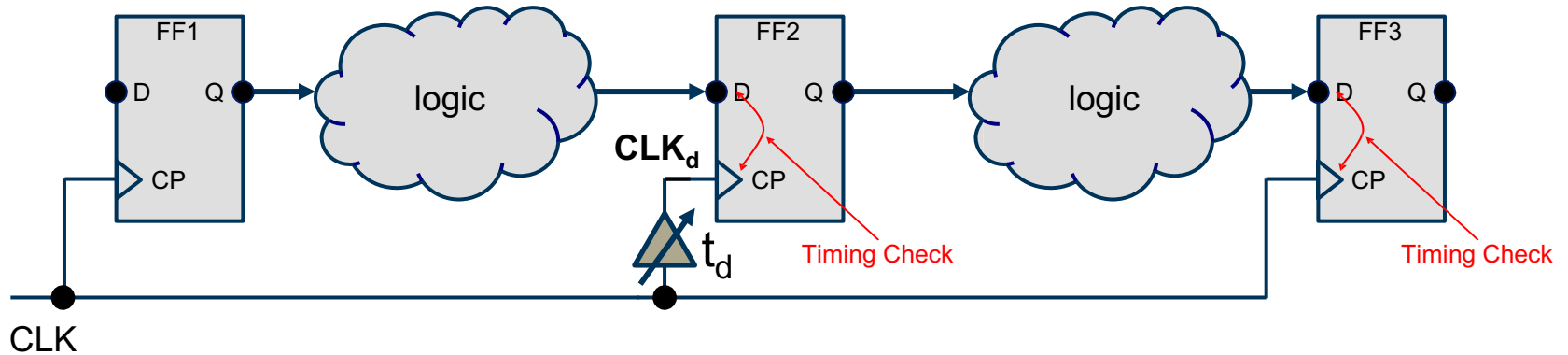
Chips with post-silicon tuning



Logic delays		Timing
Fast	Fast	✓
Fast	Slow	✗ ✓
Slow	Fast	✗ ✓
Slow	Slow	✗



Post-silicon clock tuning

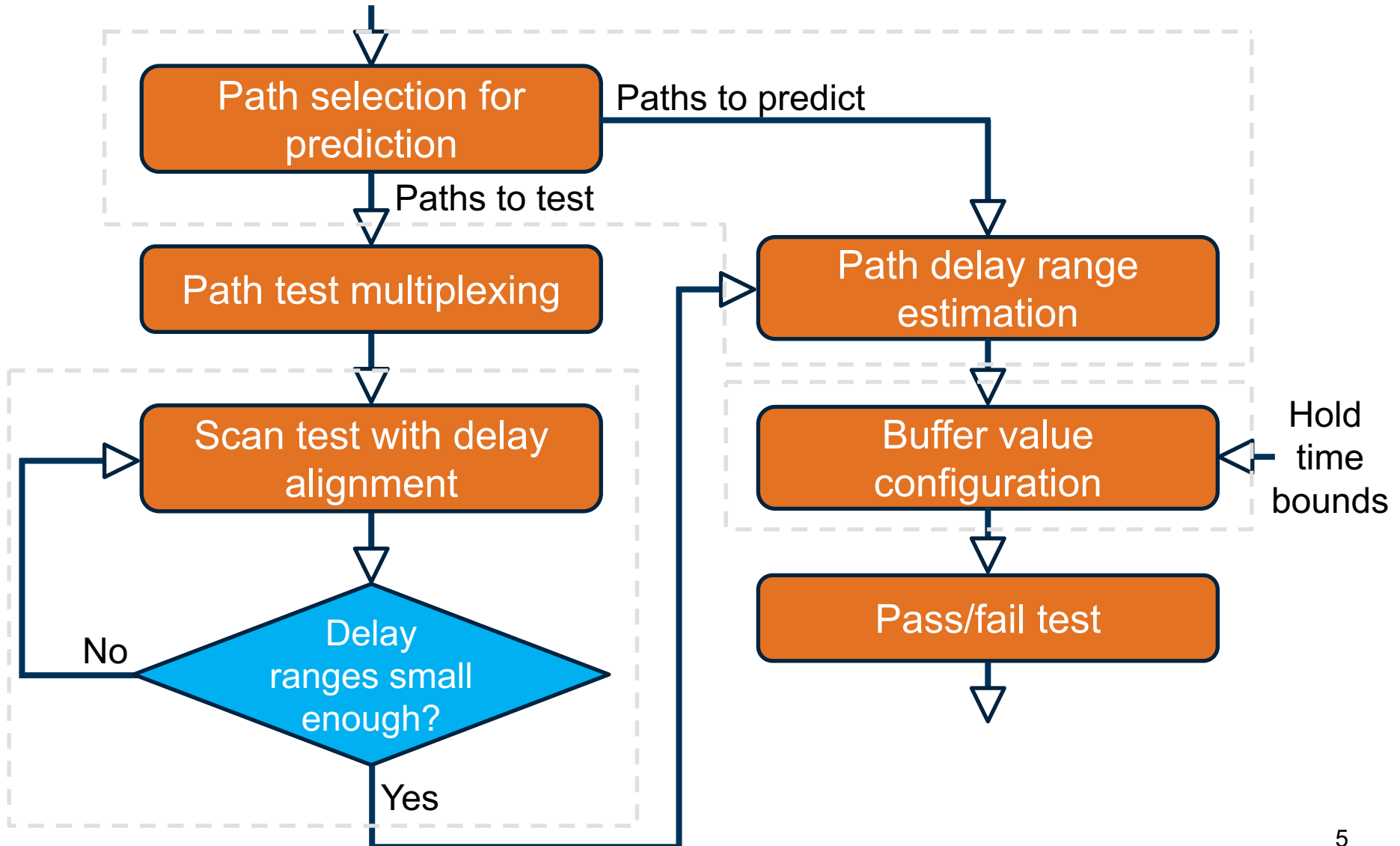


T: 8 → 5.5

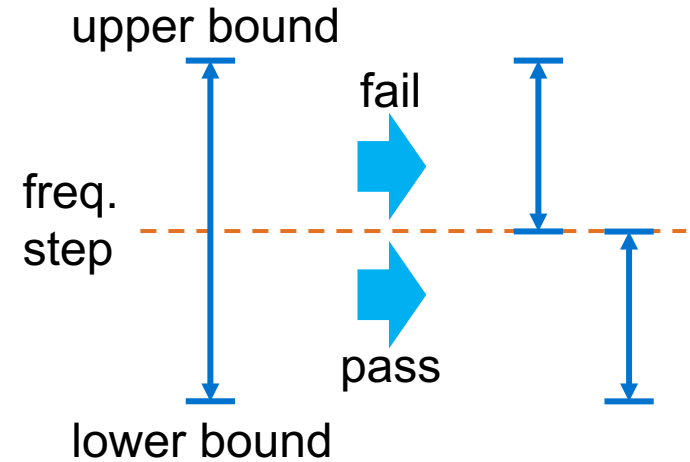
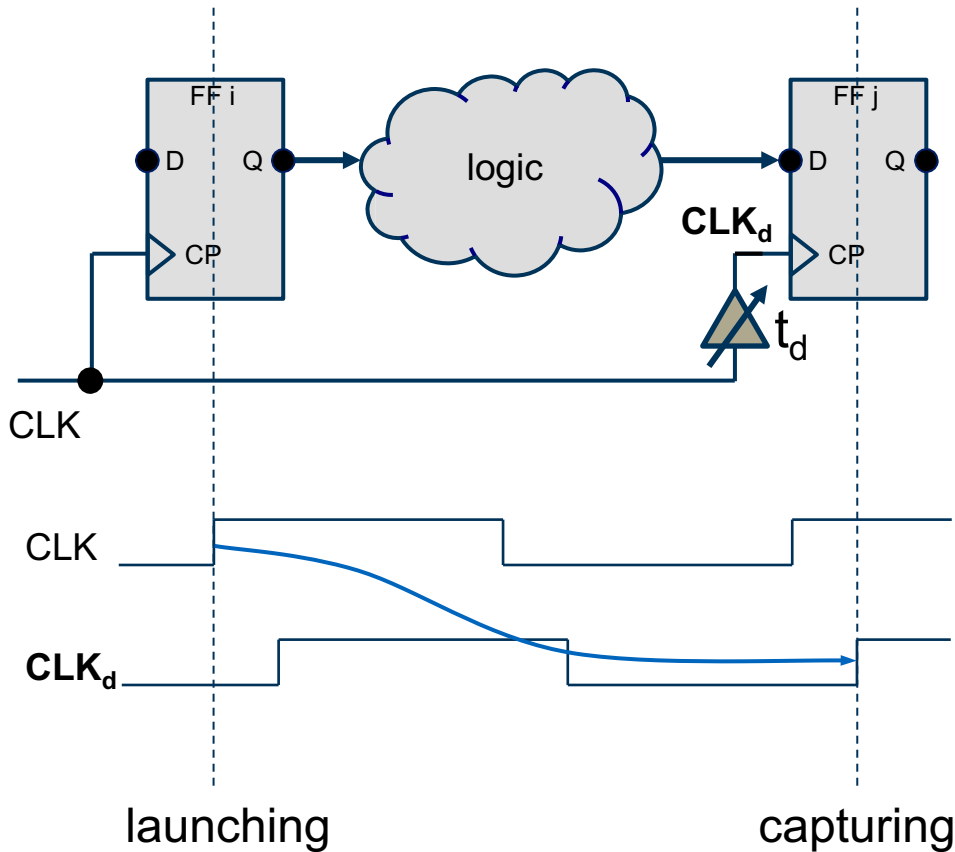
- Challenges in post-silicon clock tuning
 - Buffer insertion during design phase: area vs. yield*
 - ***Post-silicon test after manufacturing: test cost vs. yield***

*Design, Automation and Test in Europe 2016, Dresden

Test and configuration flow

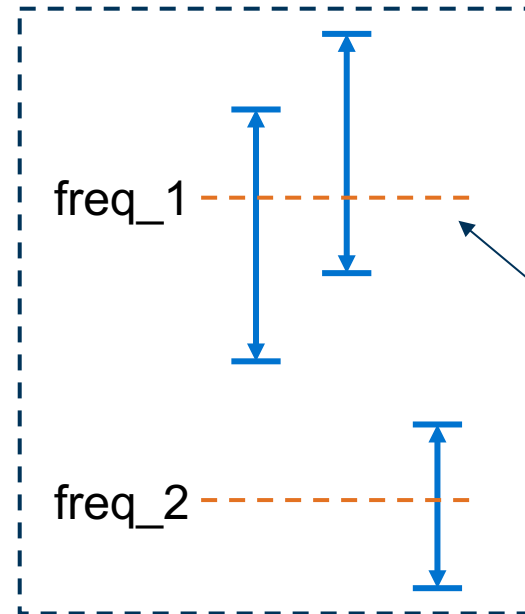
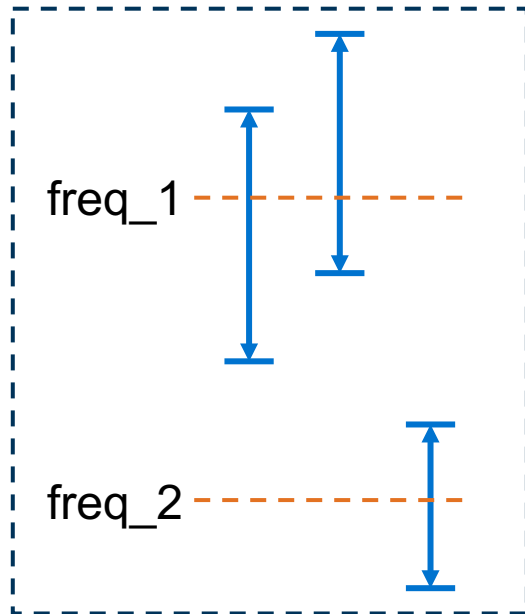


Delay range test with frequency stepping



- Combinational delays can be approximated by partitioning delay ranges iteratively, i.e., *frequency stepping*.

Delay range alignment with tuning buffers

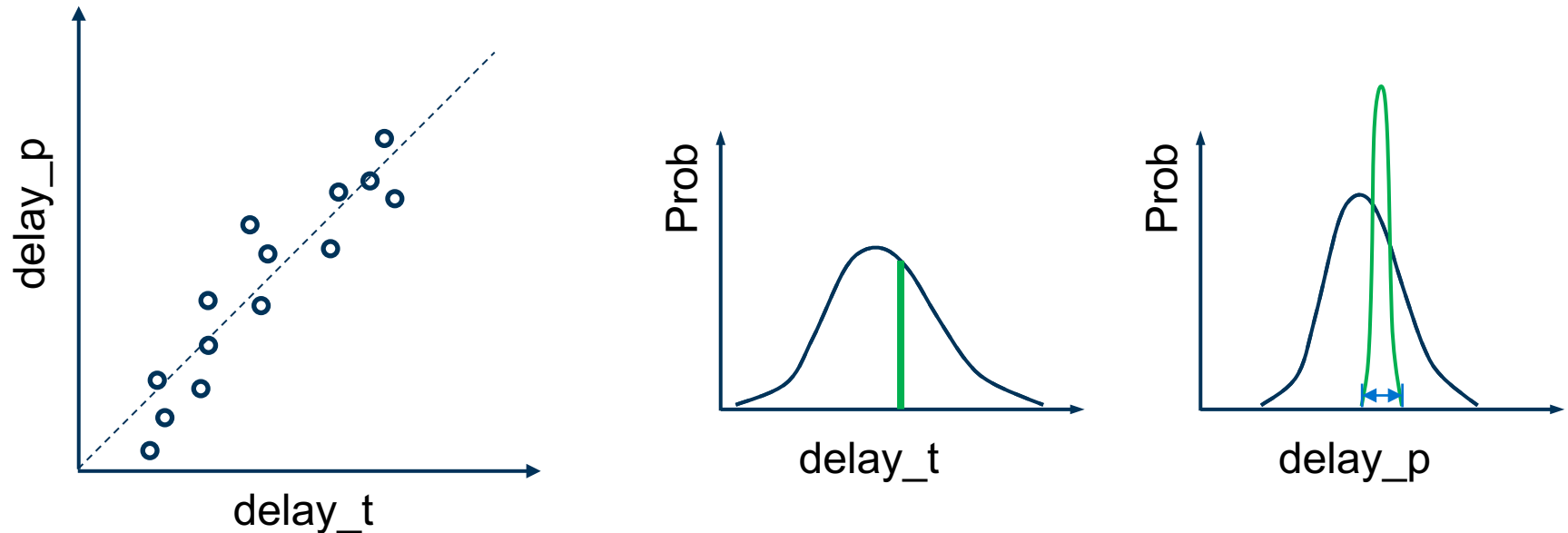


Timing budget for the path is reduced by moving clock edges

Delay test batch: maximum two ranges can be tested by a frequency.

Aligned delay ranges by tuning buffers: three ranges can be tested at the same time.

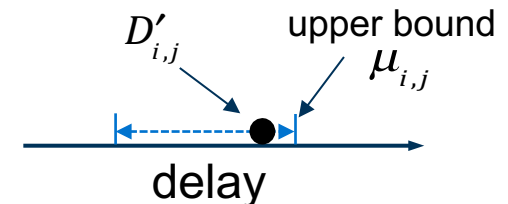
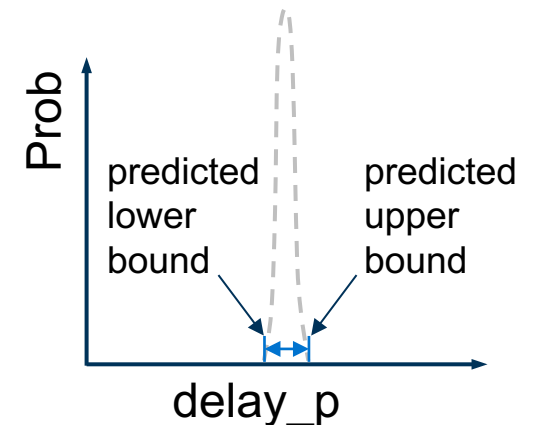
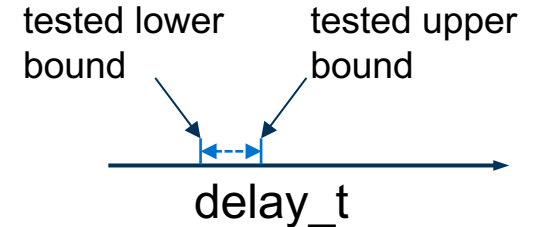
Statistical delay prediction



- Highly-correlated delays resemble each other in manufactured chips.
- Test information about one delay narrows the ranges of other correlated delays.
- Delay prediction is performed in path clusters individually.

Buffer configuration after test

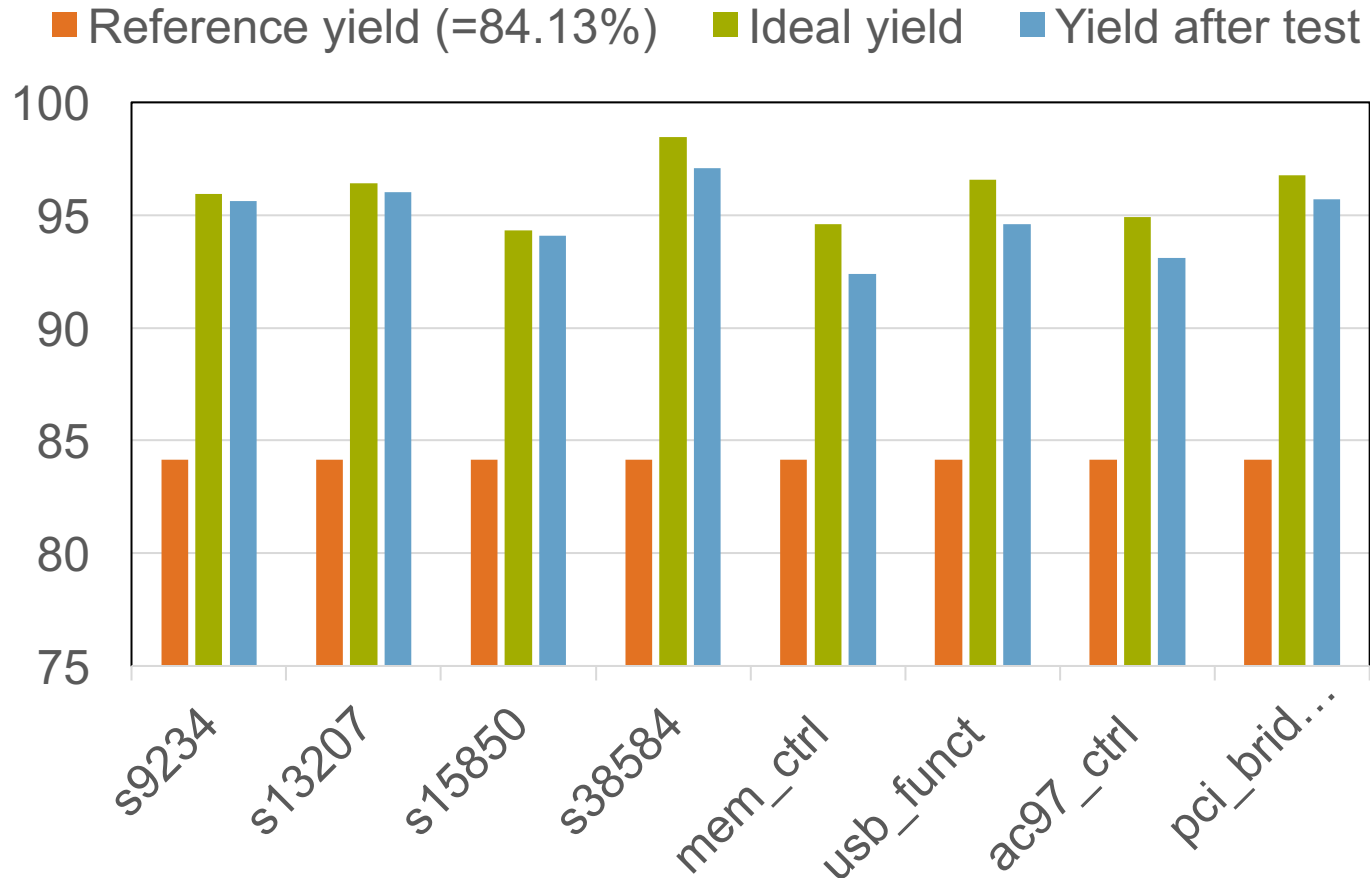
- Delays after test are small ranges.
- Real delays may take any value in the ranges.
- Buffer configuration:
 - Upper bounds of all ranges are used to configure buffers, **if** a valid configuration can be found to meet the given clock period;
 - **Else**: assume delays as large as possible
➔ risk of yield loss vs. feasibility



Results of test iterations

Circuit	Path-wise test	Aligned test with prediction	Test reduction
s9234	700	37	94.71%
s13207	4001	39	99.03%
s15850	3684	76	97.94%
s38584	3093	62	98.00%
mem_ctrl	27415	195	99.29%
usb_funct	4569	114	97.51%
ac97_ctrl	7340	288	96.08%
pci_bridge32	29061	298	98.97%

Results of yield improvement



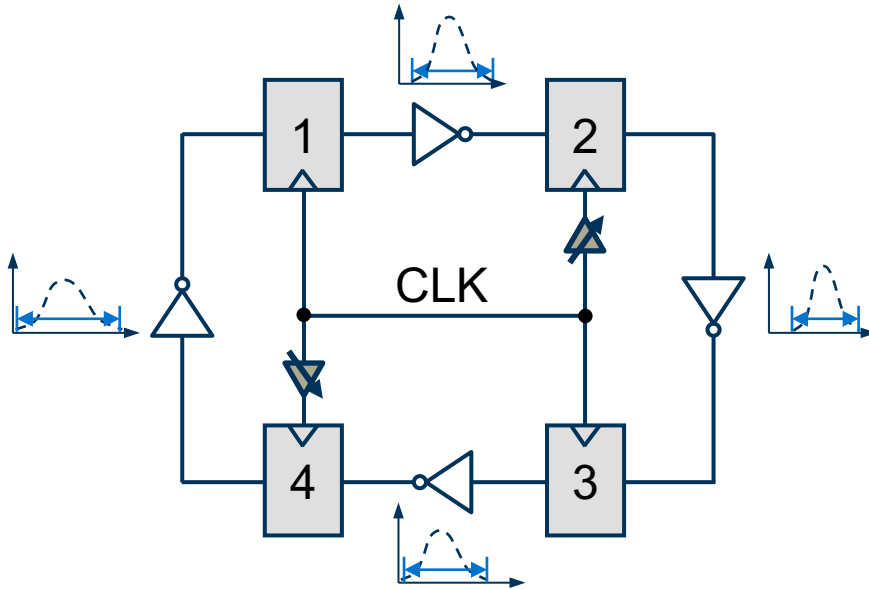
- No more than 1% of flip-flops have buffers

Summary

- Post-silicon tuning deals with the effect of process variations in each chip individually.
- Delay test can be performed efficiently, with
 - delay alignment using existing tuning buffers during tests
 - statistical delay prediction
 - post-silicon configuration with delay tolerance in ranges
- Our ongoing work includes post-silicon tuning considering reliability and noise.

Thank you for your attention!

Post-silicon delay test problem



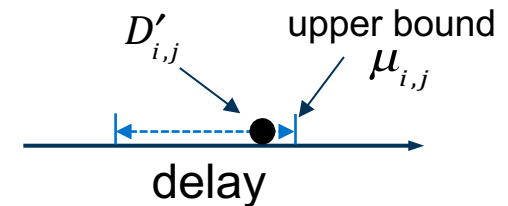
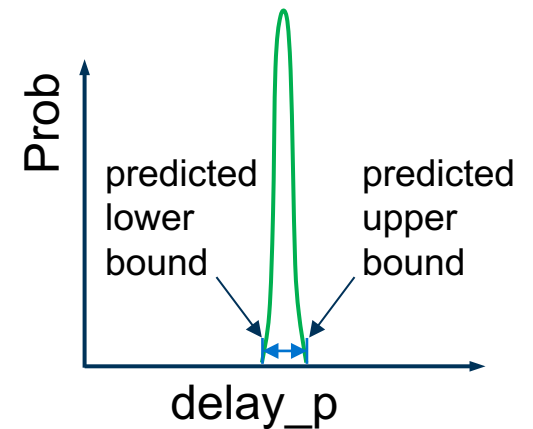
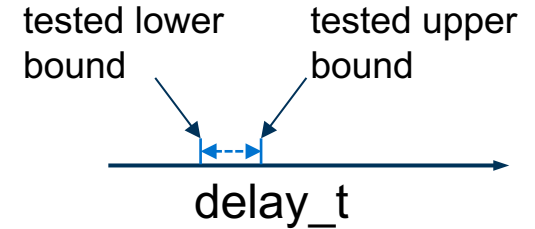
Sample chip after manufacturing

- Logic delays need to be measured for buffer configuration.
- Narrower delay ranges
 - ➡ better clock tuning
 - ➡ higher yield

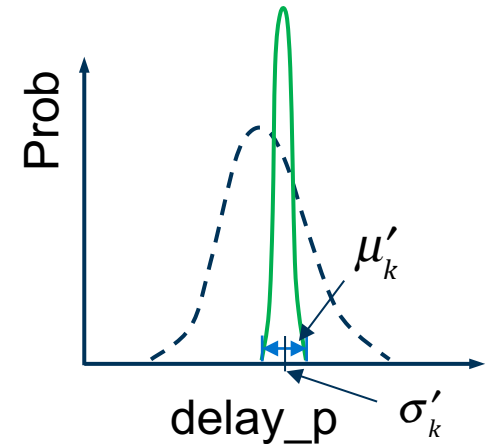
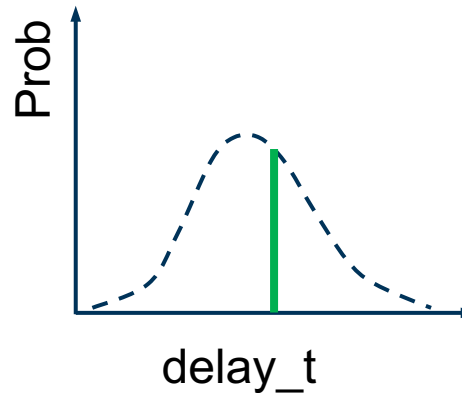
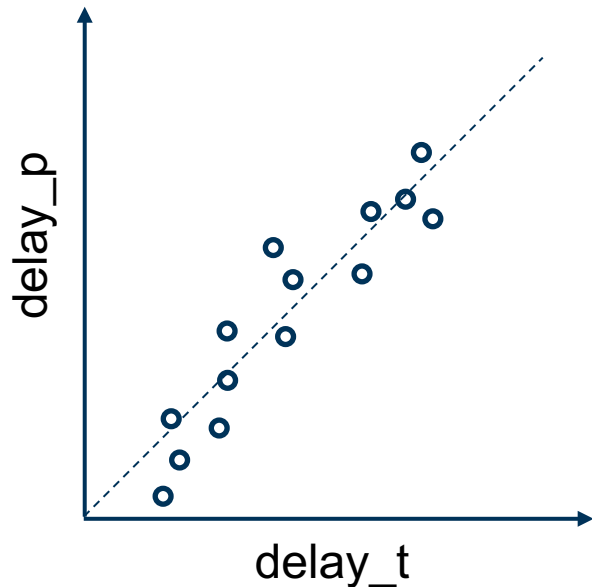
Buffer configuration after test

- Delays after test are small ranges.
- Real delays may take any value in the ranges.
- Buffer configuration:
 - Upper bounds of all ranges are used to configure buffers, **if** a valid configuration can be found to meet the given clock period;
 - **Else**: assume delays as large as possible
 → risk of yield loss vs. feasibility

Minimize: ξ
 Subject to: $T \geq D'_{i,j} + x_i - x_j$
 $\xi \geq \mu_{i,j} - D'_{i,j}$
 x_i, x_j : buffer values



Statistical delay prediction



- Mean and variance of the predicted delay

$$\mu'_k = \mu_k + \Sigma_{k,t} \Sigma_t^{-1} (d_t - \mu_t)$$

μ_k, σ_k : original delay distribution

$$\sigma'^2_k = \sigma_k^2 - \underbrace{\Sigma_{k,t} \Sigma_t^{-1} \Sigma_{t,k}}_{\text{variance decrease}}$$

$\Sigma_{k,t}, \Sigma_{t,k}, \Sigma_t$: (co)variance matrix

Hold time constraints

- Circuit tuning for performance may affect hold time constraints.

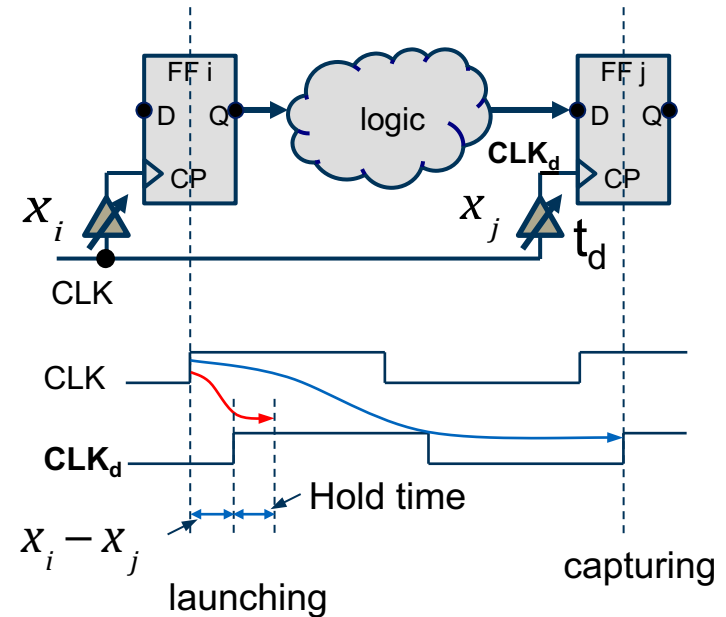
$$T \geq D'_{i,j} + x_i - x_j$$

$$x_i - x_j \geq d'_{i,j}$$

- Hold time are constraints by lower bounds as

$$x_i - x_j \geq \lambda'_{i,j}$$

- $\lambda'_{i,j}$ are set to guarantee a small yield loss



Runtime

Circuit	T_p (s)	T_t (s)	T_s (s)
s9234	6.58	0.09	0.00
s13207	16.75	0.06	0.00
s15850	50.51	0.17	0.01
s38584	90.45	0.15	0.01
mem_ctrl	622.63	0.36	0.02
usb_funct	118.48	0.17	0.02
ac97_ctrl	81.63	0.30	0.01
pci_bridge32	749.31	1.19	1.59

T_p : runtime for path grouping and selection, test multiplexing and hold time bound computation;

T_t : average test time of a chip for post-silicon tuning;

T_s : runtime to determine the final buffer values.