**Development of Complex Curricula for Molecular Bionics and Infobionics Programs within a consortial* framework****

**Consortium leader**

# PETER PAZMANY CATHOLIC UNIVERSITY

**Consortium members**

# SEMMELWEIS UNIVERSITY, DIALOG CAMPUS PUBLISHER

The Project has been realised with the support of the European Union and has been co-financed by the European Social Fund ***

**Molekuláris bionika és Infobionika Szakok tananyagának komplex fejlesztése konzorciumi keretben

***A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

# VLSI Design Methodologies

(VLSI tervezési módszerek)

# Digital design flow: timing

(A digitális tervezési folyamat: időzítések)

PÉTER FÖLDESY

# The topics are covered in this chapter:

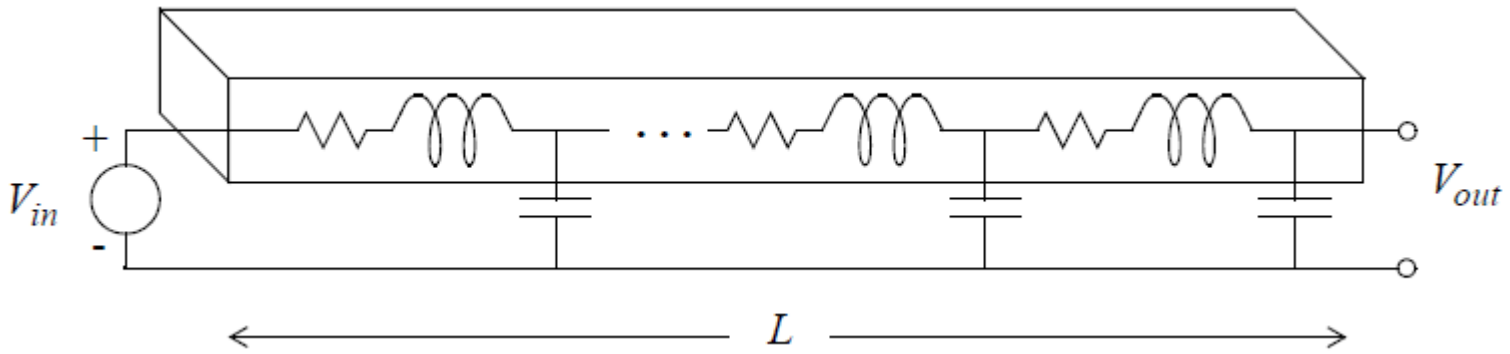
- Interconnection delay sources are summarized
- Classic wire model and its simplification
- Estimating the wire delay
- Methods to reduce the wire delay
- A simple design task to introduce the timing related tasks
- Introduction of the static timing delay

# Section I

# Interconnection delay sources, wire models

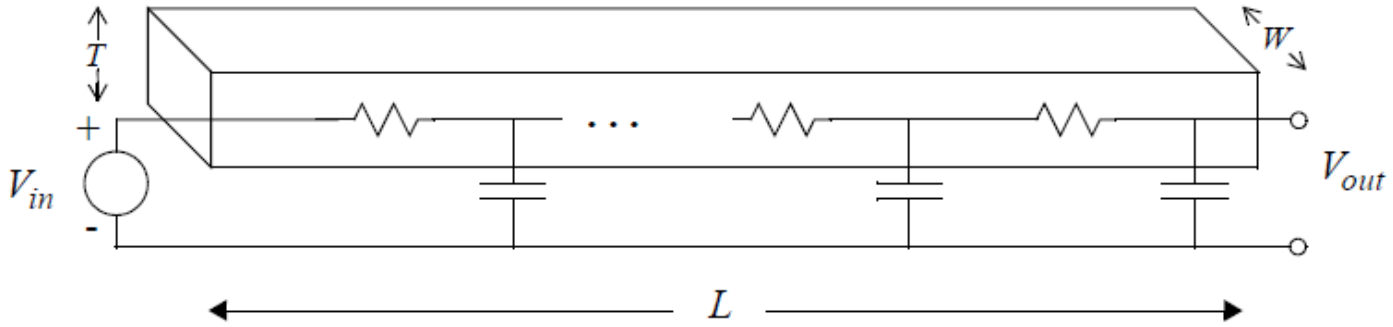
- Delays occur and cause:
  - Signal does not arrive at once to the target node, why?
  - Voltage slew degrading for a long wire
    - Signal propagation limited to light speed
    - Driver needs to charge or discharge wire's capacitance through its self resistance and wire resistance
  - Going through pass transistors (passive elements) or buffers (active elements)
  - Thresholding it to binary value results in a delayed acknowledge of the signal change

- Classic wire model
  - the conducting wires have parasitic resistance, inductance and capacitance.
  - These parasitic determine how fast signals propagate through the wires.
  - For maximum accuracy, the wires can be modeled as an RLC transmission line.

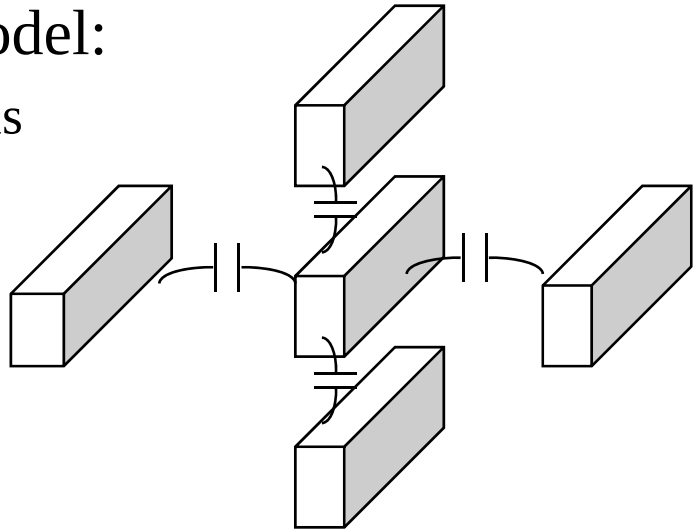


- Classic wire model
  - In digital simulation, simpler models are used. First, the inductance is neglected.
  - To handle the delay in a simple form, we can approximate it by the so-called Elmore delay, that is a simple approximation to the delay through an RC network in an electronic system.
  - Its definition is a time, when a step function is delayed so that the area under the step is the same as the area under the actual wave form at point

- Simplified wire model




- Generic wire capacitance model:
  - Area to upper and lower metals and substrate
  - Fringe to adjacent metal lines

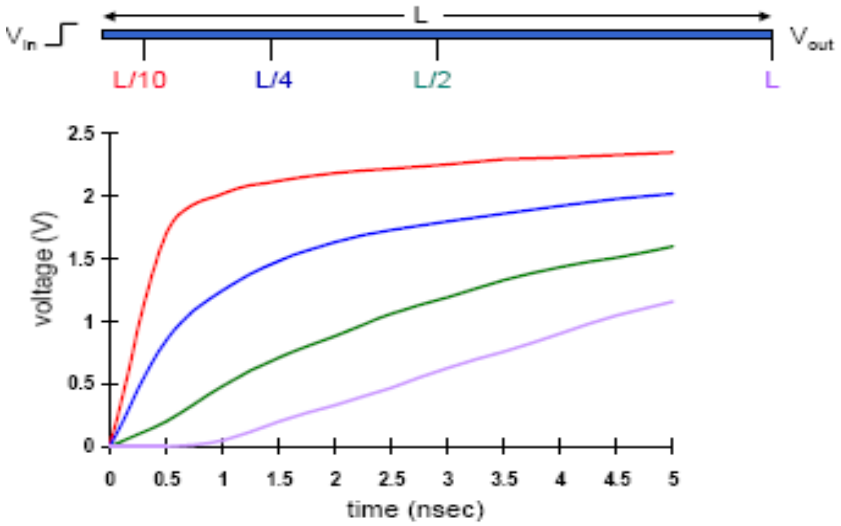


- Estimating the delay
  - The delay is estimated to a total length L of the wire as

$$t_E = rc\frac{L^2}{2} = R_s\left[\left(\frac{C}{A}\right) + \frac{2}{W}\left(\frac{C}{P}\right)\right]\frac{L^2}{2}$$

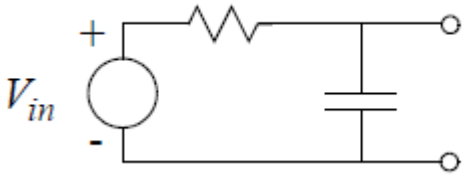
  - Where the *r, c* are the resistance and capacitance per unit width, *A* is the wire area, *P* is the perimeter
  - Note that the capacitance is composed of the area and the perimeter related fringe capacitances.

- The consequences are:
  - The wire delay is proportional to the square of the length



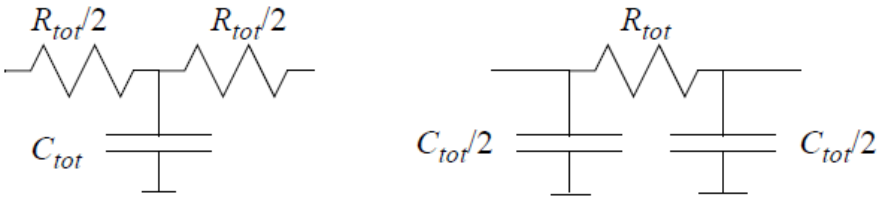

  - In modern processes with large (C/P), wire delay is smaller for wider lines
  - In modern processes, wire delay is smaller for thicker lines

- Modeling the delay
  - We can use simple RC transmission line for simplicity with the same Elmore delay (the form of the wavefront will be different!)




$$RC = rc\frac{L^2}{2} = \frac{1}{2} \cdot rL \cdot cL = \frac{1}{2} R_{tot} C_{tot}$$

  - Or with one of the followings:



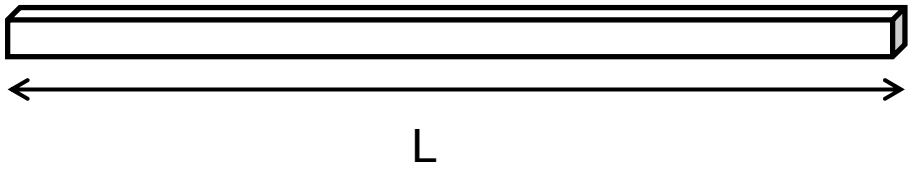
# Section II

# Reducing the delay of interconnections

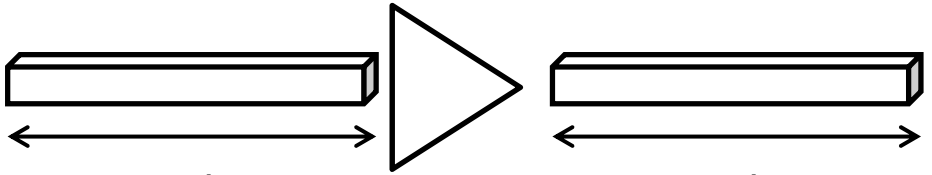
- Reducing the delay
  - Usage of low resistance line (copper wires!)
  - Usage of thicker wires (the higher level metals are usually thicker, good for both power distribution and high speed lines, like the clock tree)
  - Insert buffers in it (as long as the buffer delay is smaller than the wire delay, but in modern technologies, this is simple requirement)

- Reducing the delay

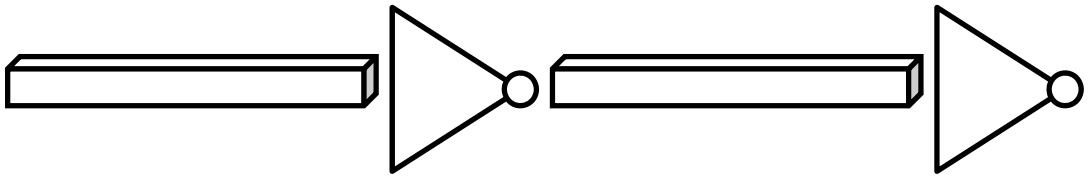
delay ~ $rcL^2$




delay ~ $t_{buff} + rc/4L^2$

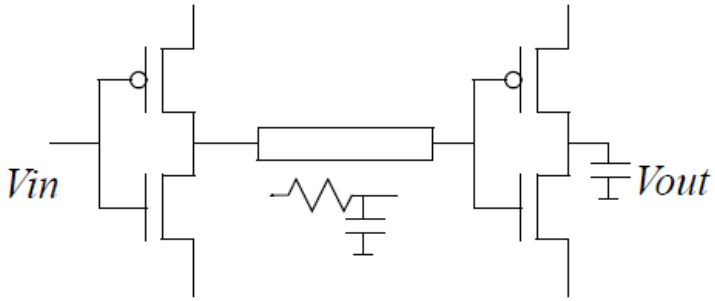



or, two inverters

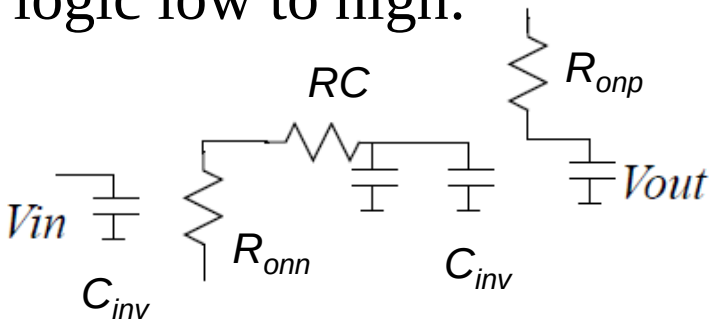


(a buffer is usually two inverters in series)

- Reducing the delay
  - In a circuit example:




  - And its rough model, supposing that the *Vin* is just switched from logic low to high:



- If the $C_{inv}$ capacitance is smaller than the wire capacitance and the transistor on resistance is neglible, than we get the previous estimates

$$RC >> R_{inv} C_{inv}$$

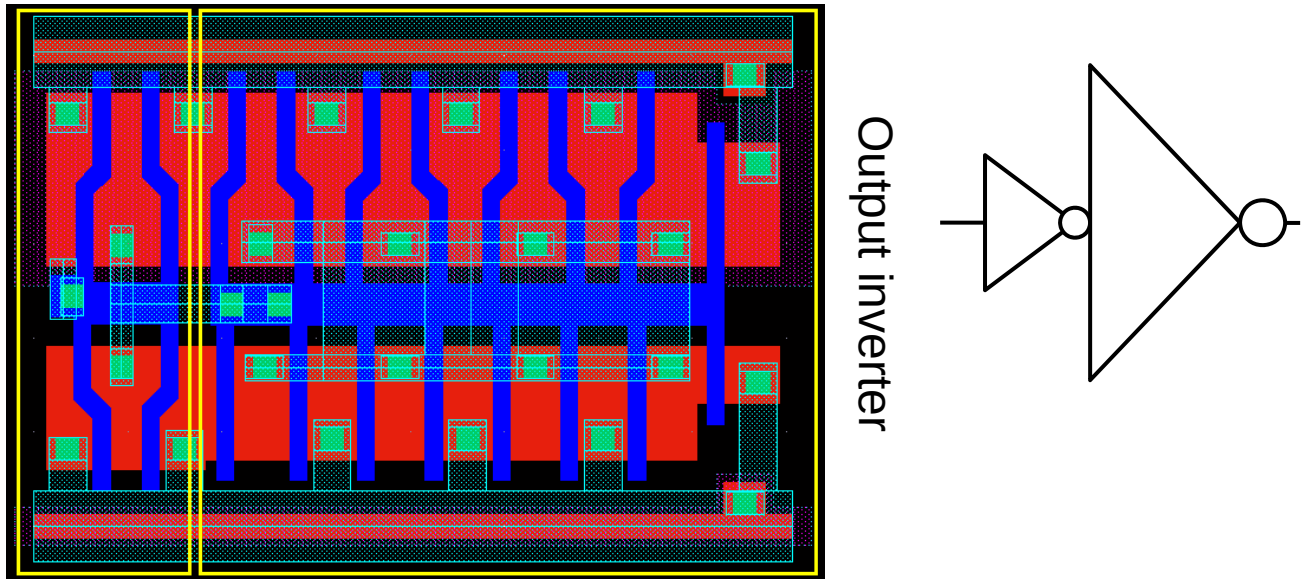
- If the wire delay is smaller

$$RC << R_{inv} C_{inv}$$

than we get a delay that is governed by the transistor sizes.

- Rule of thumbs
  - Remember, that the transistor capacitance is proportional to the area of the gate (~W*L), and the
  - on resistance (with rough estimate) is proportional to the width (W) and inversely proportional to the length (L) of the transistor's channel (~L/W)
  - By increasing only the width of both devices, the $R_{inv} C_{inv}$ value remains the same ($R_{inv} C_{inv} \sim L^2$)!
  - This raises the idea to increase the buffer transistor width to suppress the wire delay.

- Well, the wise step would be to increase the driving capability of the buffer and decrease its input capacitance. And the *buffers* in *standard cell libraries* are designed so.




A typical high drive buffer

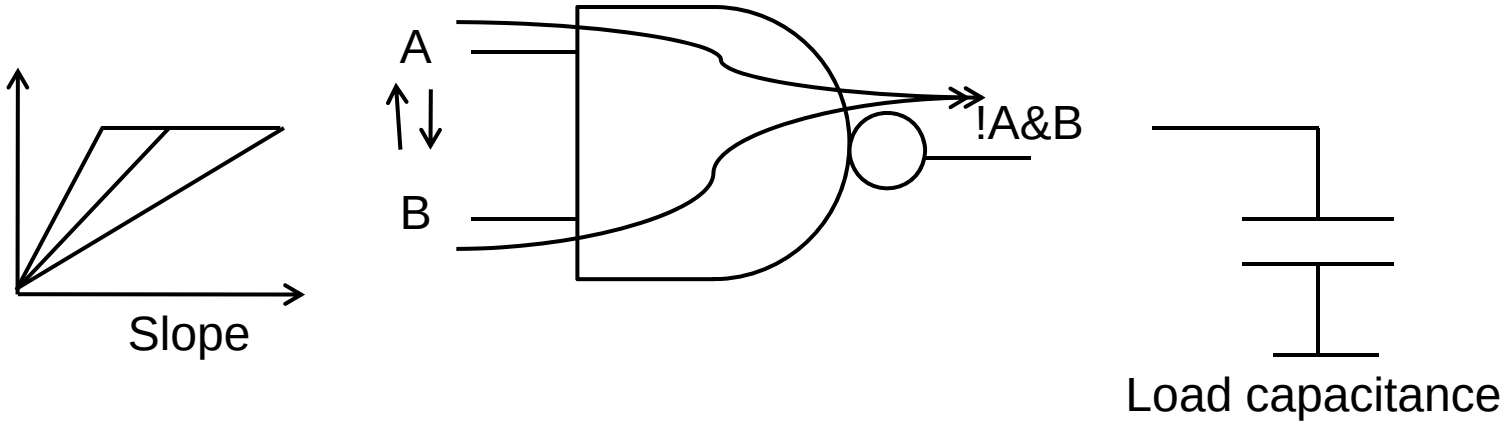
# Section III

# Modeling delay in digital design environment

- Modeling the delay in digital design environment
  - Circuit simulators such as SPICE
  - *K-factor model*, which is a very simple model for delay estimation
  - *Degrading models* are used for interpolating non measured situations.
  - Two dimensional look-up tables of many paths describing the connection between the delay and the skew and load capacitance.

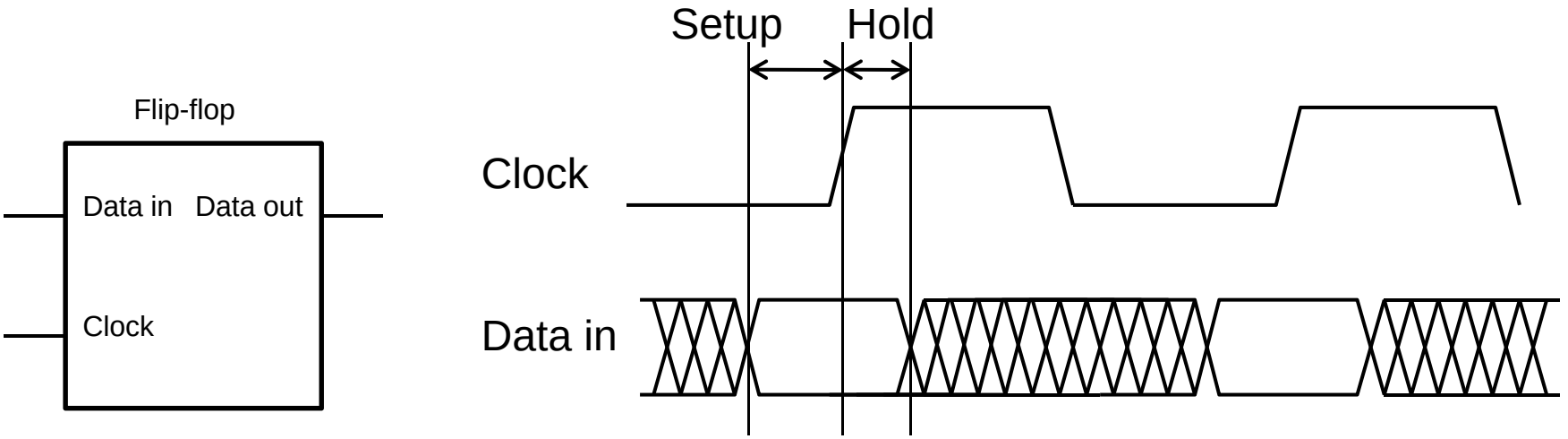
- Modeling in digital design environment
  - Circuit simulators such as SPICE can be used with RLC models. This is the most accurate, but slowest, method.
    - This technique is used for simulation of the extracted clock trees.
  - Simple modeling using the K-factor model. This approximates the delay as a constant coming from the driver and k times the load capacitance.
    - Used in synthesis, when exact length are not available.
  - Degrading models when the given model parameter is not characterized well (such as temperature, power supply drop, statistically significant delay spread is considered as an accumulated value).

- Modeling in digital design environment
  - Two dimensional tables. The most useful and used way in digital design (also power consumption). These tables take an output load and input slope, and generate a circuit delay and output slope.
    - Table size is about 5x5 to 10x10 for each arch in the modeled block (all input to all output, both rise/fall conditions)!



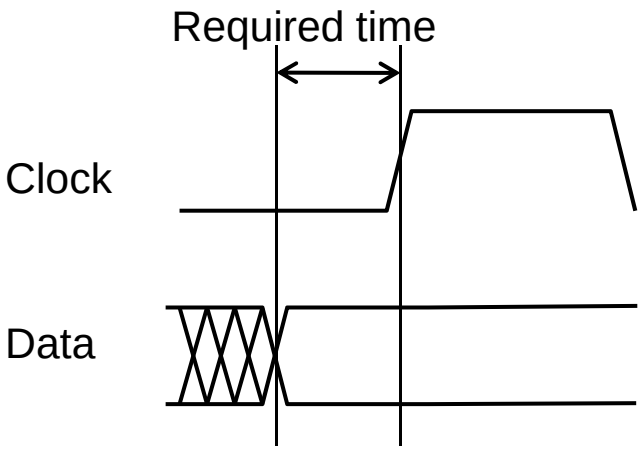
# Timing analysis: Important notations and checked values

- The flip-flops and other synchronous elements are sensitive mostly to two time period around the clock edges:
  - *Hold time*, that needed for the FFs to store the information
  - *Setup time*, that needed to be prepared for storage

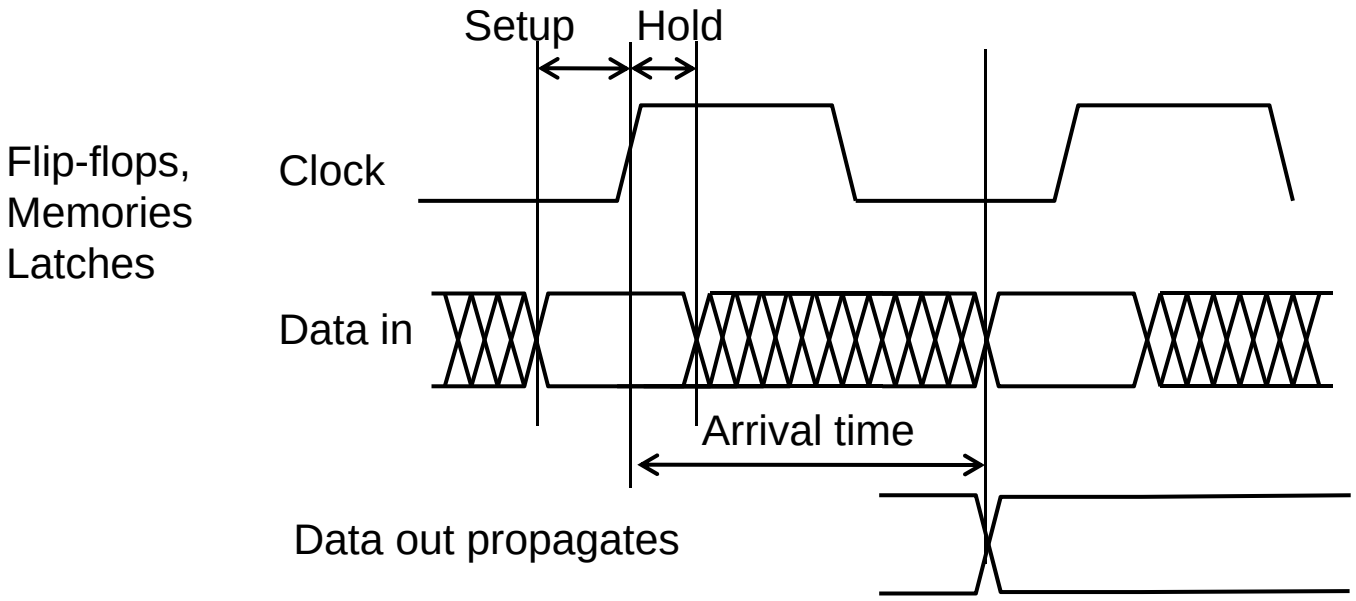


- Between these time markers, the inputs should be stable. The failures are the following:
  - A hold time violation, when an input signal changes too quickly
  - A setup time violation, when a signal arrives too late
- The *critical path* is defined as the path between an input and an output with the maximum delay of all possible paths.

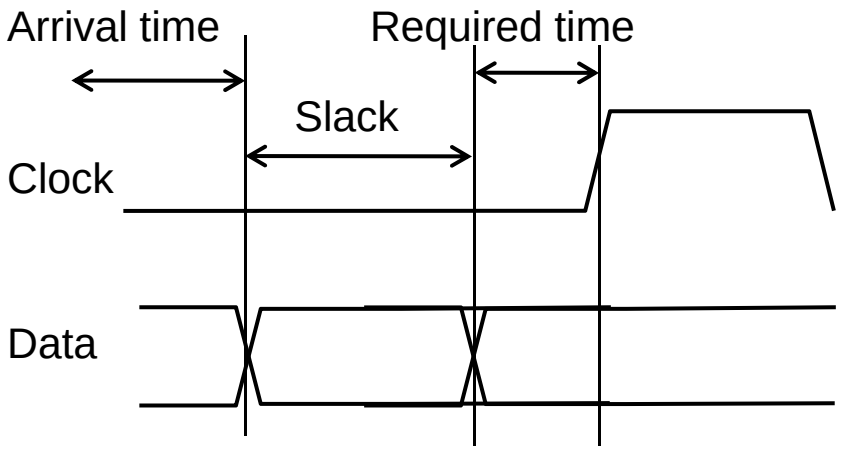
The required time is the latest time at which a signal can arrive without causing any violation.



- The arrival time is the time elapsed for a specific signal to arrive at a certain point.
- The *timing slack* is the difference between the required and the arrival time. If they are more paths, then the worst of them.



- The slack is the difference between the required and the arrival time. The most important timing result, showing the robustness and timing margin.



# Corner analysis

- In the circuit operation the operating conditions affect the switching speed, wire delay, etc.
- The most important conditions are the following of which we can define SLOW, FAST, TYPICAL cases:
  - Temperature, as increases, the speed drops
  - Power supply, as increases, the speed increases
  - Manufacturing differences in active element doping level, it happens to be a transistor to be more or less conductive
- The corner analysis means any simulation, timing checks under permuted collection of conditions

# Static timing analysis (STA)

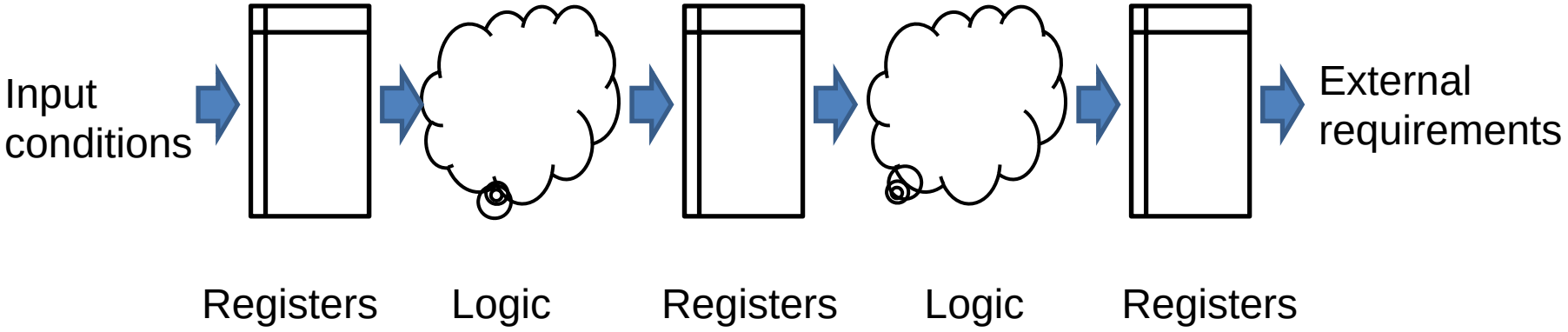
- This is a method to systematically calculate timing
- The STA is much faster than functional simulation and plays a key role in digital design
- This process takes the calculated loads and driving capabilities (from prepared libraries of tables) as input.
- The user defined environmental requirement is also needed (external loads, required arrival time, etc.)

## *Static timing analysis (STA)*

- Go though every possible arches and calculate the signal propagation through them.
- The STA checks the elements signal arrival and stability conditions by propagating the interleaved logic paths and interconnect delays.
- Its result is the slack and arrival timing information.
- This information is fed back to the synthesis, to the placement and the routing, finally the sign-off verifications

# Static timing analysis

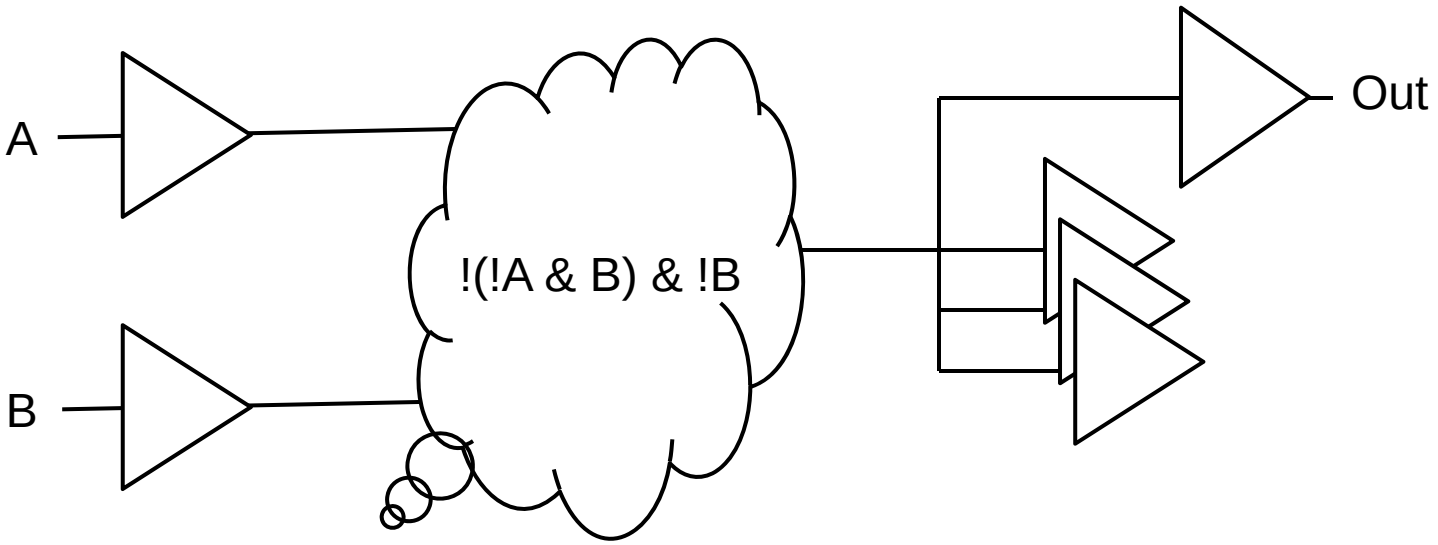
- In general, it calculates with synchronizing elements, such as flip-flops or latches (actually, this is key for synchronous systems)
- Its more advanced version takes into account the interconnection mismatch in a statistical way (SSTA)



# Section IV

# Explanatory task for manual synthesis with timing and area awareness

- Homework: synthesis with timing and area optimization
  - Get the minimal delay and area versions of the function by hand STA!
  - Logic function to create: Out = !(!A & B) & !B
  - Given a set of cells, and the environment with output load



- The available gates:
  - Given gates: BUF1, INV1, INV2, NAND1, NAND2
  - Interconnect delay: for simplicity, nothing
  - Inputs are driven by BUF1
  - Output load is 4 * BUF1
  - Note: two or more inverters or buffers can be connected in parallel to increase its driving force

Cell area table

| Cell | Area |
|---|---|
| INV1 | 1 |
| BUF1, NAND1, INV2 | 2 |
| NAND2 | 3 |

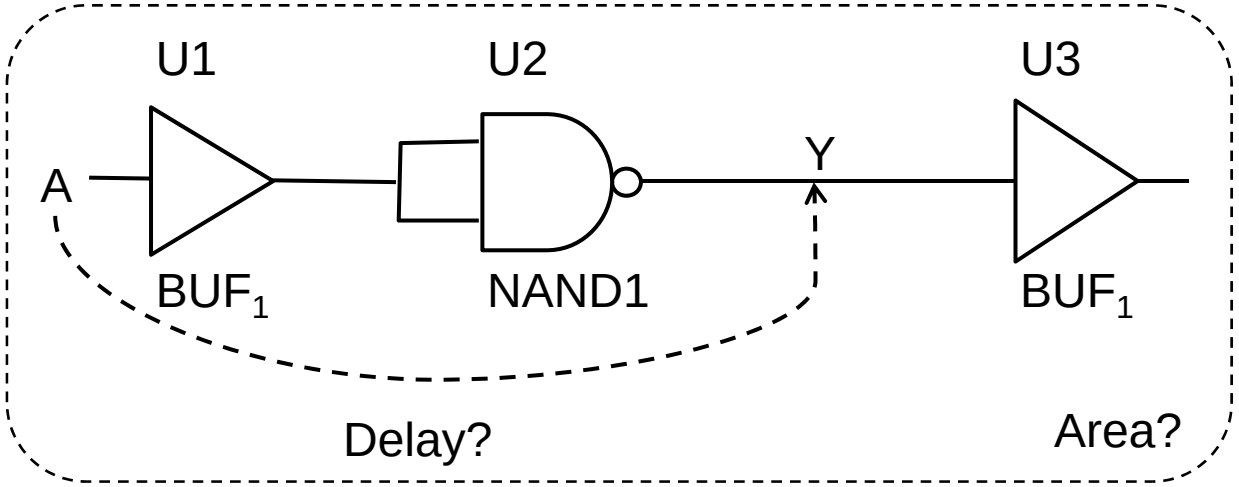
Cell input capacitance table

| Cell | Input C |
|---|---|
| INV1, BUF1, NAND1 | 1 |
| NAND2, INV2 | 2 |

Cell intra delay as a function of the output load

| Load in C | INV1 | INV2 | BUF1 | NAND1 | NAND2 |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 2 | 2 |
| 2 | 2 | 1 | 3 | 3 | 2 |
| 3 | 3 | 2 | 4 | 4 | 3 |
| 4 | 4 | 2 | 5 | 5 | 3 |

- Example for delay and area calculation




- Output loads: U1 has 2 units, U2 has 1 unit
- Delay per cells: U1 has 3 units, U2 has 2 units
- Delay calculation from A to Y, $T_{delay}$ = 3+2 = **5 units**
- Area is BUF1 and NAND1 and BUF1 = **6 units**

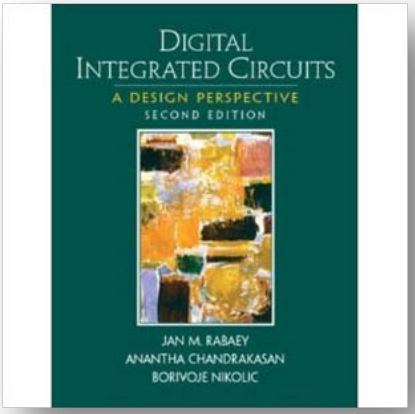
- Algorithm to go systematically gets the results:
  - Subtask 1: Synthesize the circuit in a few minimal forms
    - Subtask 2: Calculate the delay in each input to output arch
    - Subtask 3: Find the worst path. That is the critical path.
    - Subtask 4: Change the driving capability of elements
    - Subtask 5: Still all paths and cell version are listed go to step 2
  - Subtask 6: Report
    - Subtask 7: Select the quickest with its critical timing.
    - Subtask 8: Select the smallest with its critical timing.
  - Subtask 9: If there are other form go to step 1

- Question: Are the quickest and the fastest the same?

# Conclusions

- We saw that the delay in digital circuits comes from block internal delay and the interconnection signal propagation.
- The general timing requirements and their concept are introduced.
- As an automatic timing checking method, we saw the static timing analysis.

# Recommended literature




**Digital Integrated Circuits (2nd Edition)**
Jan M. Rabaey, Anantha Chandrakasan
Publisher: Prentice Hall; 2 edition (January 3, 2003)

# Comprehension questions:

I. What are the sources of the parasitic?

II. How the circuit parasitic are described?

III. Describe the modeling delay in digital design environment.

IV. What is the static timing analysis?