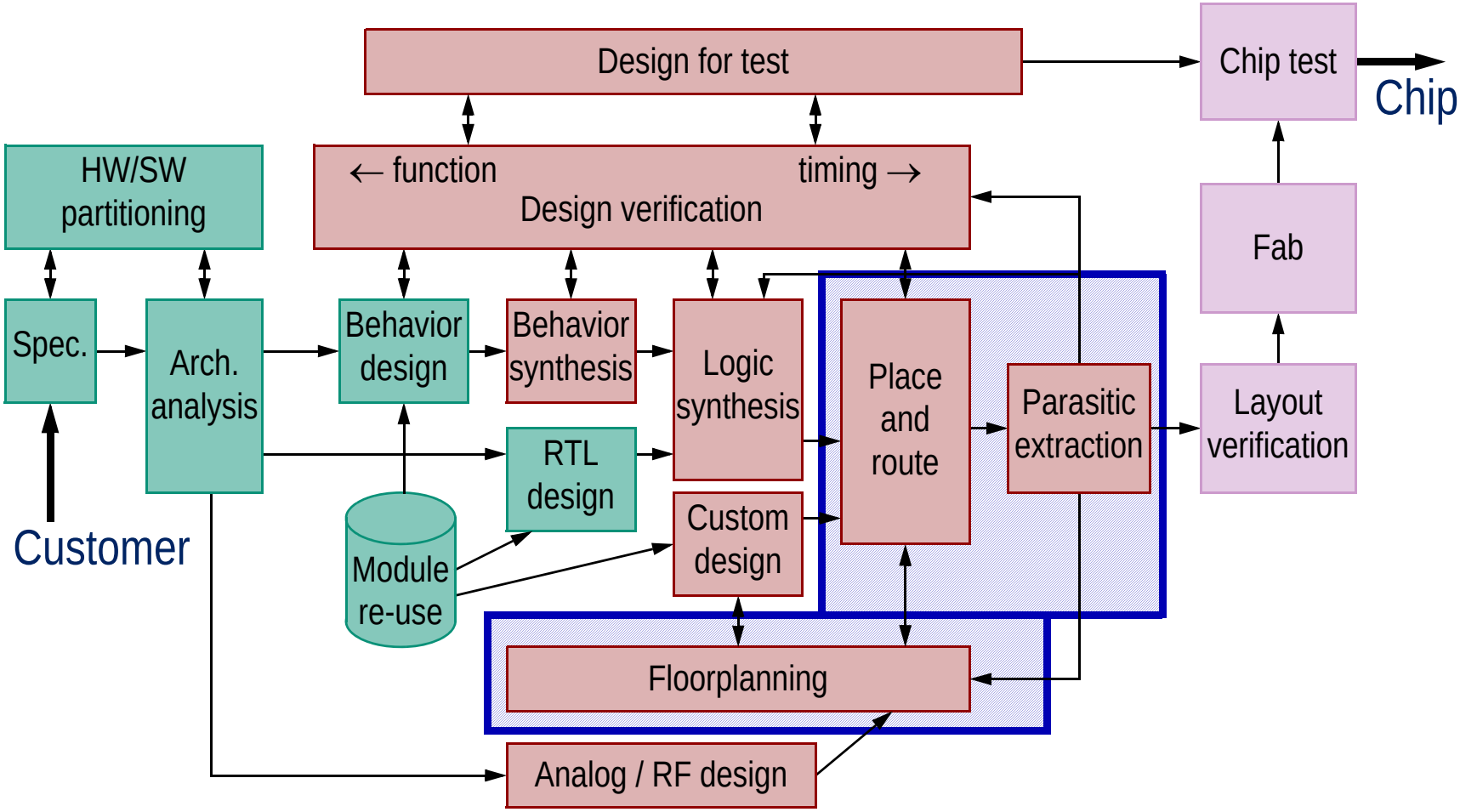
DAT110
Methods for Electronic System Design and Verification

Per Larsson-Edefors
VLSI Research Group

Lecture 7: Physical Design.

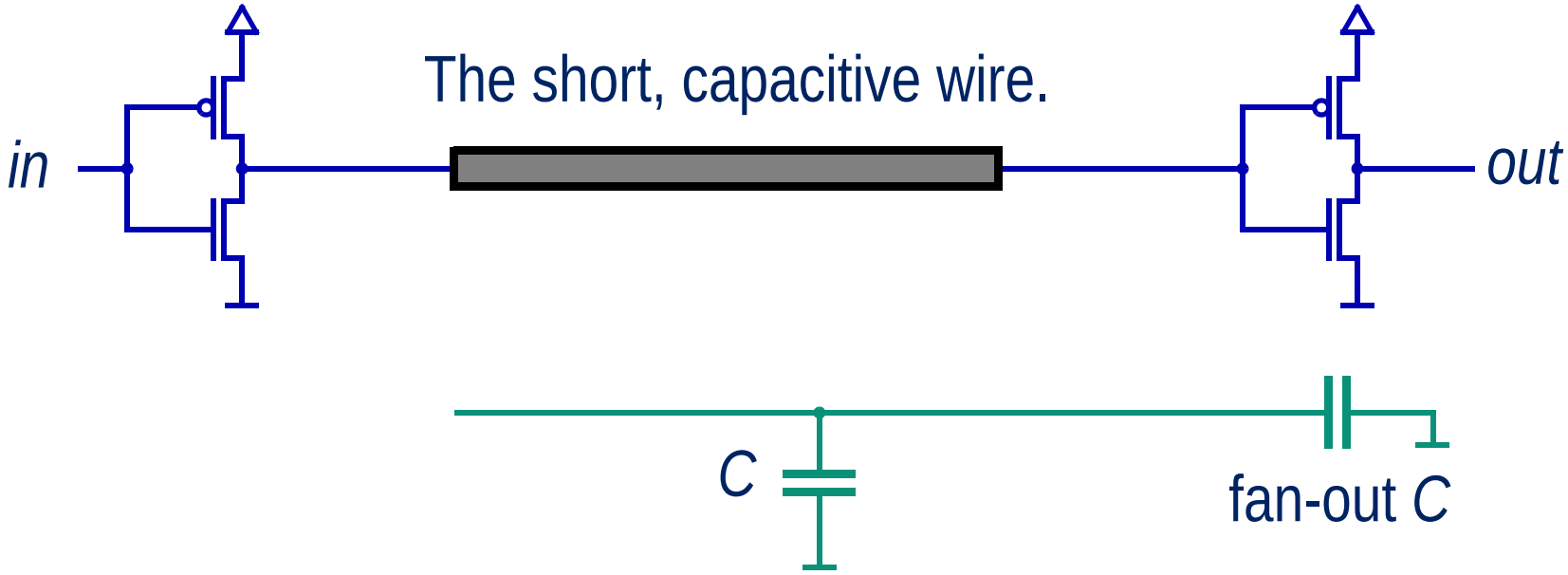
Present Scenario: Physical Design



Interconnect basics

Capacitive Model - Short, Local Wires

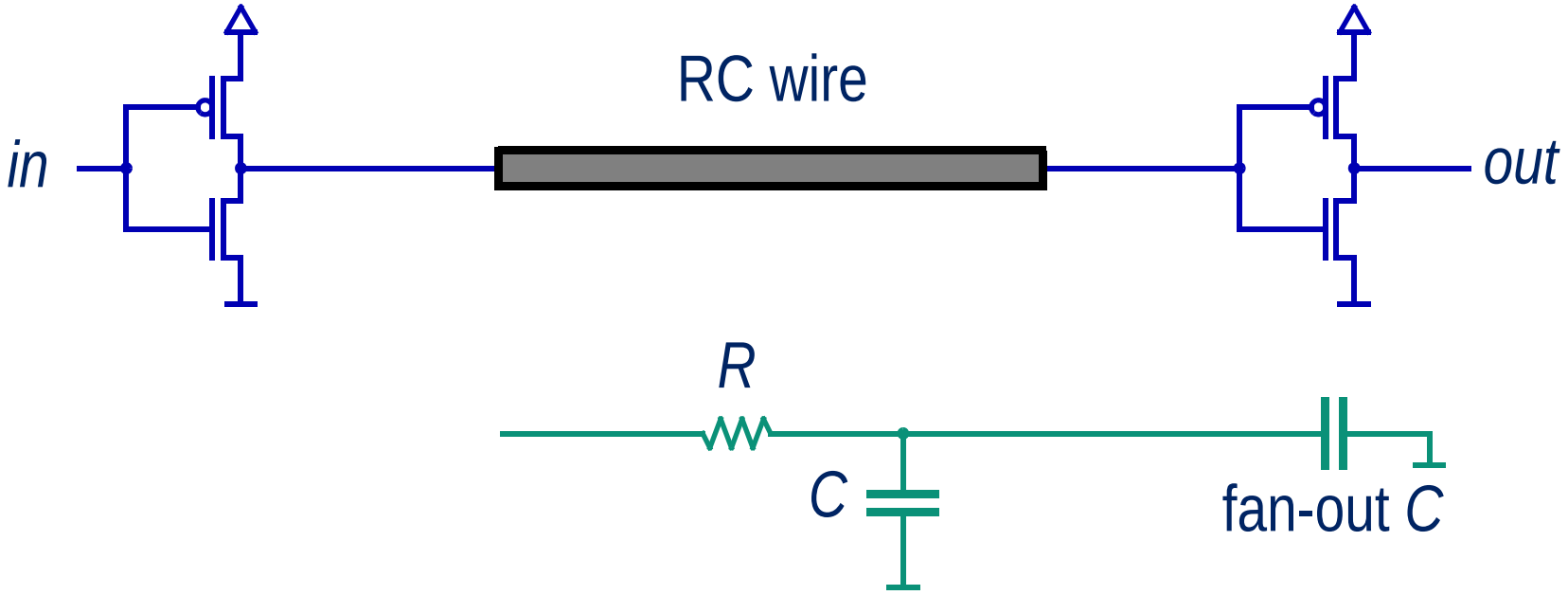
- Wire capacitance adds to the fan-out *C* load of gates.



- The wire introduces a delay penalty of $t_d = \frac{C}{k \cdot V_{DD}}$.

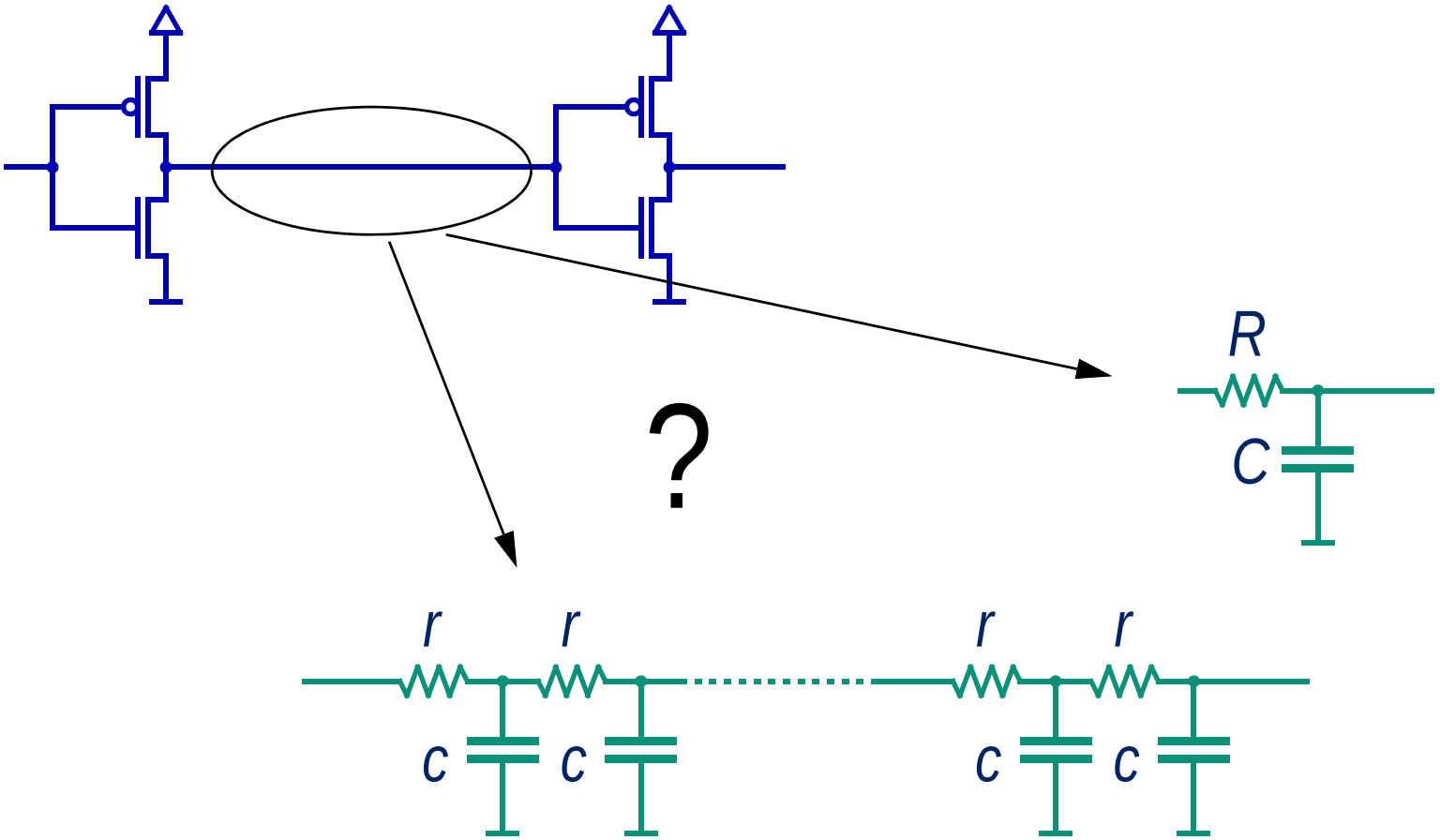
Lossy Model - Longer Wires

- Device scaling forces lower metal layers to scale (wire width), to make transistor-wire connections possible.

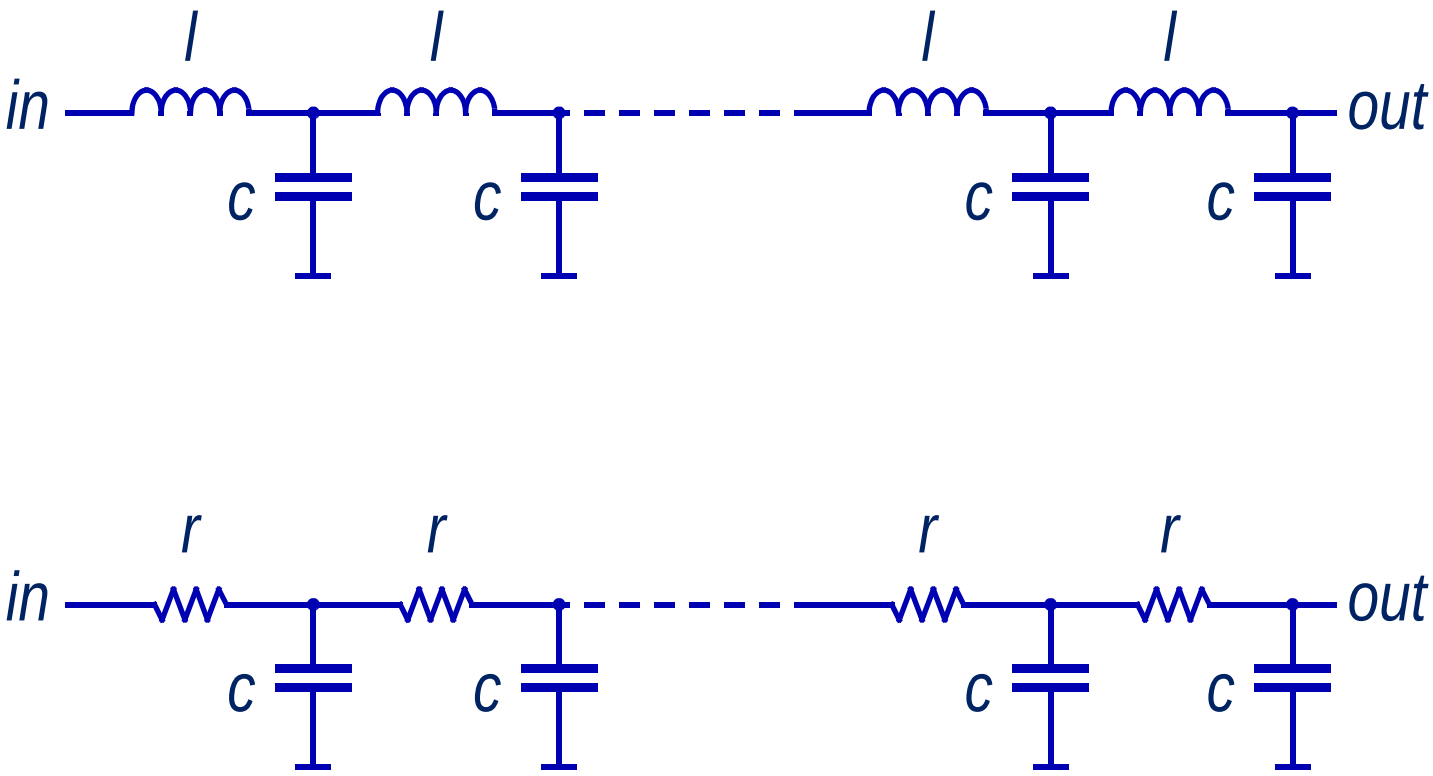


- Resistance increases delay significantly!
- Resistance decouples wire input and wire output transition.

Lumped or Distributed Models?



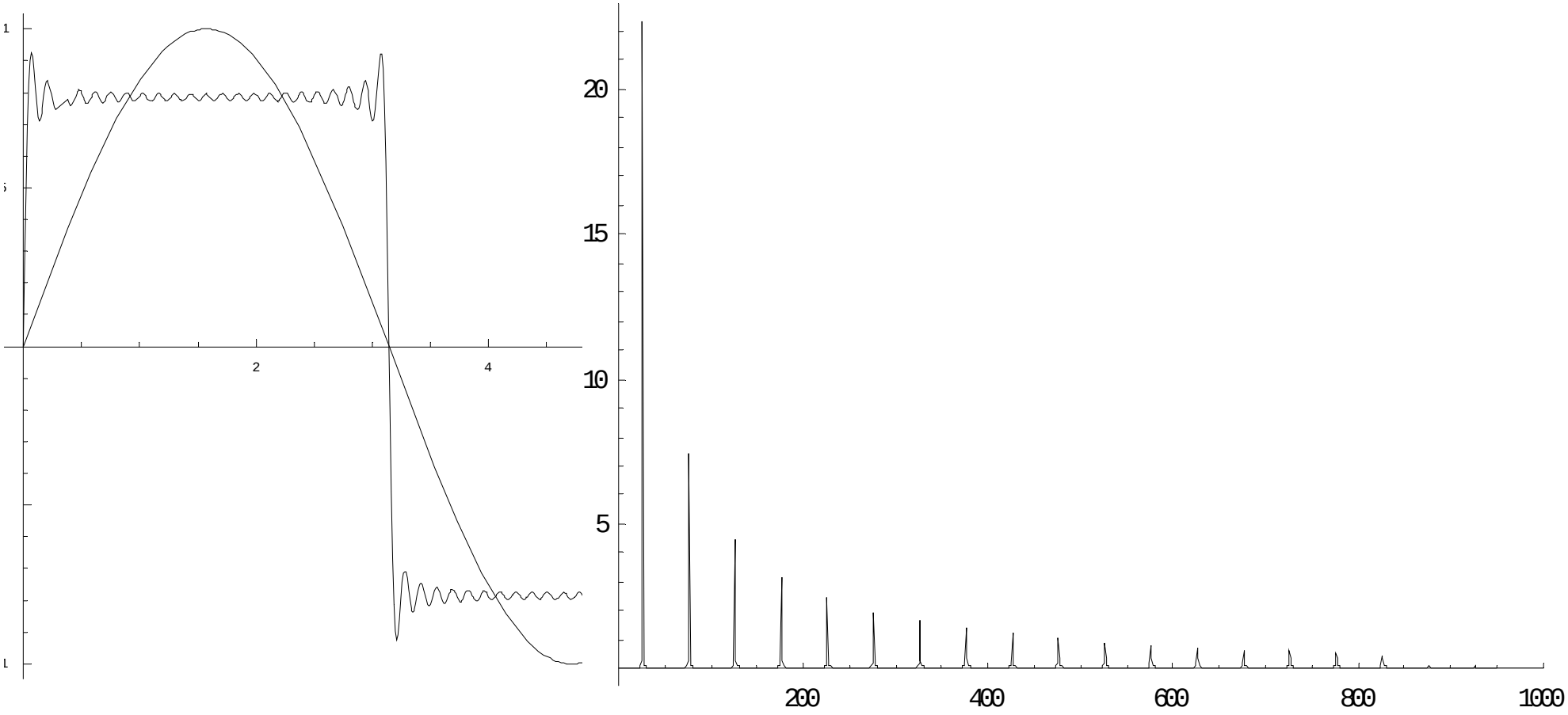
Lossless or Lossy Models?



Selecting Models

- *Lumped or distributed models?*
  This depends on wire length and signal bandwidth.
- *C, RC, LRC or LC models?*
  This depends on wire length,
  cross section, and signal bandwidth.
- Signal bandwidth refers to frequency content of edge
  (what we capture in rise/fall time = slope = slew rate = edge rate).

Frequency Content



Short edges imply high frequency content (harmonics).

Electromagnetics (Electrostatics)

- D field = electric flux density (*swe*: elektrisk flödestäthet).
- E field = electric field.
- The dielectric constant ε establishes a relation between D- and E-fields: $D = \varepsilon E$.
- In air $D = \varepsilon_0 E$, where $\varepsilon_0 = 8.854\ 10^{-12}$ As/Vm.
- In a dielectric material, an electric field gives rise to dipoles, (a polarization of the electron clouds of the atoms) which displaces the relation between the fields $\Rightarrow D = \varepsilon_r\ \varepsilon_0 E$
- ε_r is also called κ (kappa) or *k*.

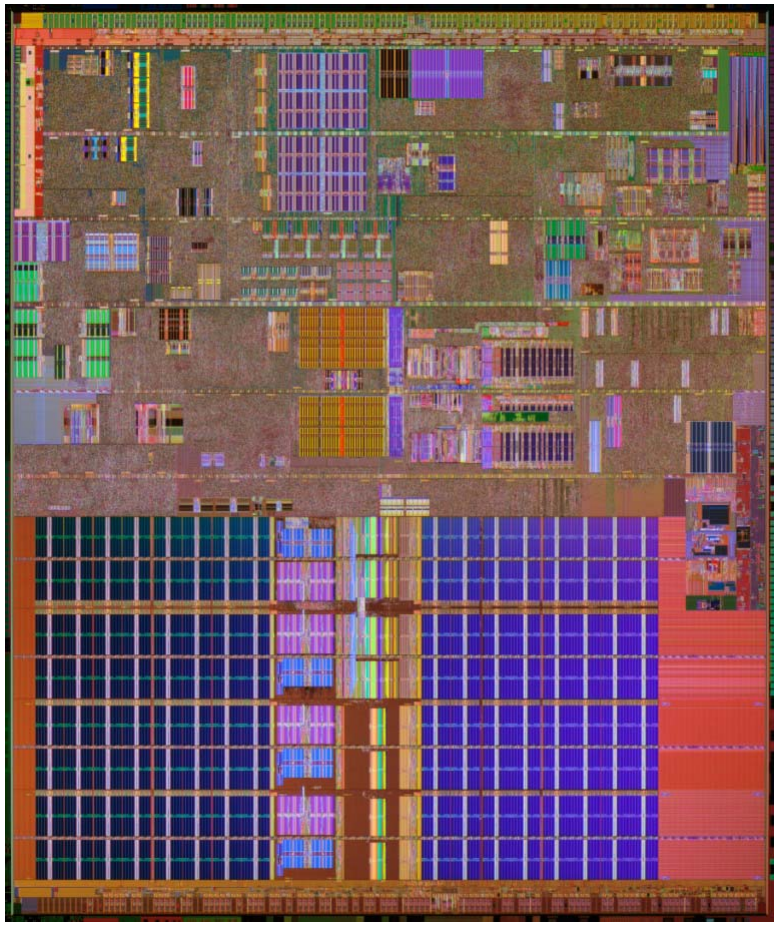
Signal Propagation Speed 1(2)

- At 3.73 GHz (overclocking $\Rightarrow$ 8.3 GHz)
  the core of a P4 runs at twice the speed = 7.46 GHz $\Rightarrow$
  clock period of 134 ps. How far does a signal reach in this time?
- Assume lossless wire, for which the dielectric is speed limiting,

  $v = \frac{c}{\sqrt{\varepsilon_r}}$, c = 300 000 000 m/s (vacuum). But what is ε_r ?
- Assume ILD is SiO_2, for which $\varepsilon_r = 3.9$ $\Rightarrow$
  v_{max} is 150 000 000 m/s $\Rightarrow$
  in 134 ps, a signal reaches around 20 mm.

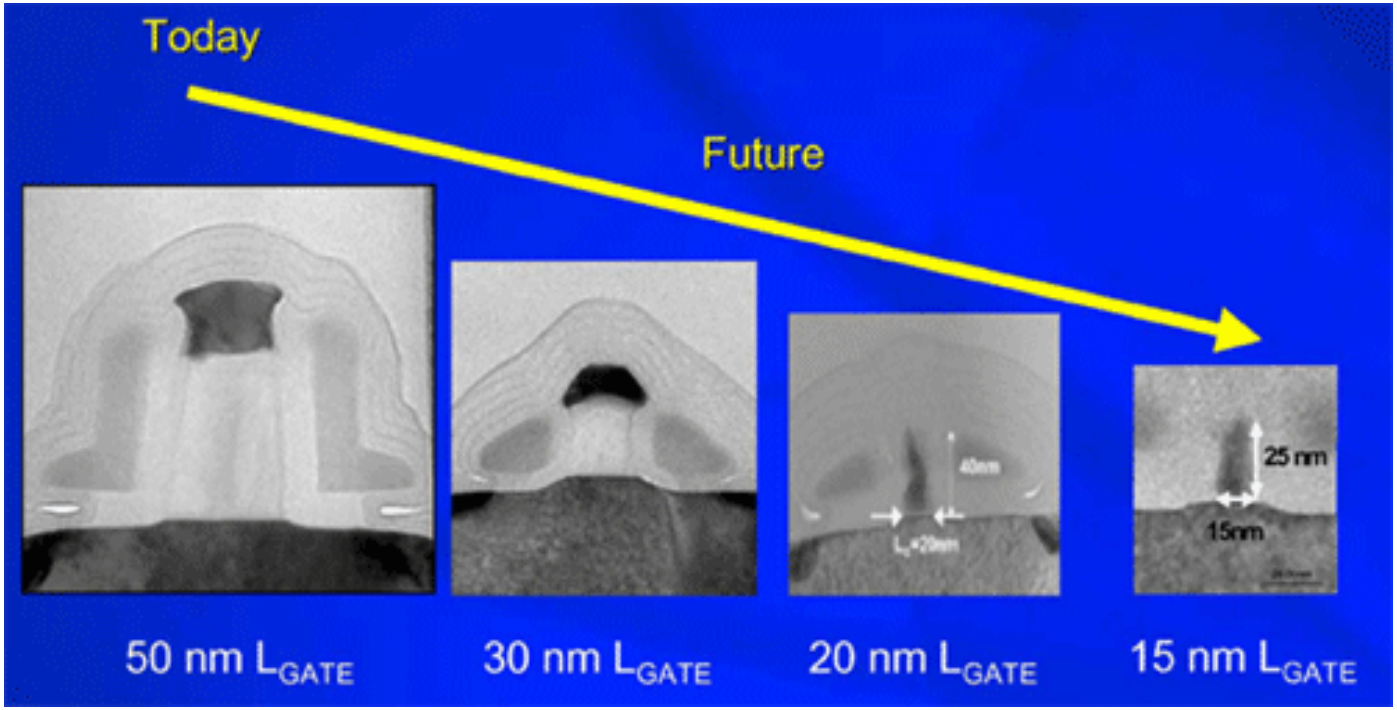
Signal Propagation Speed 2(2)

- P4 (6xx), 170Mtrans, 90 nm.
- Area = 10.7 mm x 12.6 mm.
- Signal propagation distance is of the same order as die side.
- However, we assumed lossless wires; but this is seldom correct!



Source: Intel

Scaling - Technology Front-end

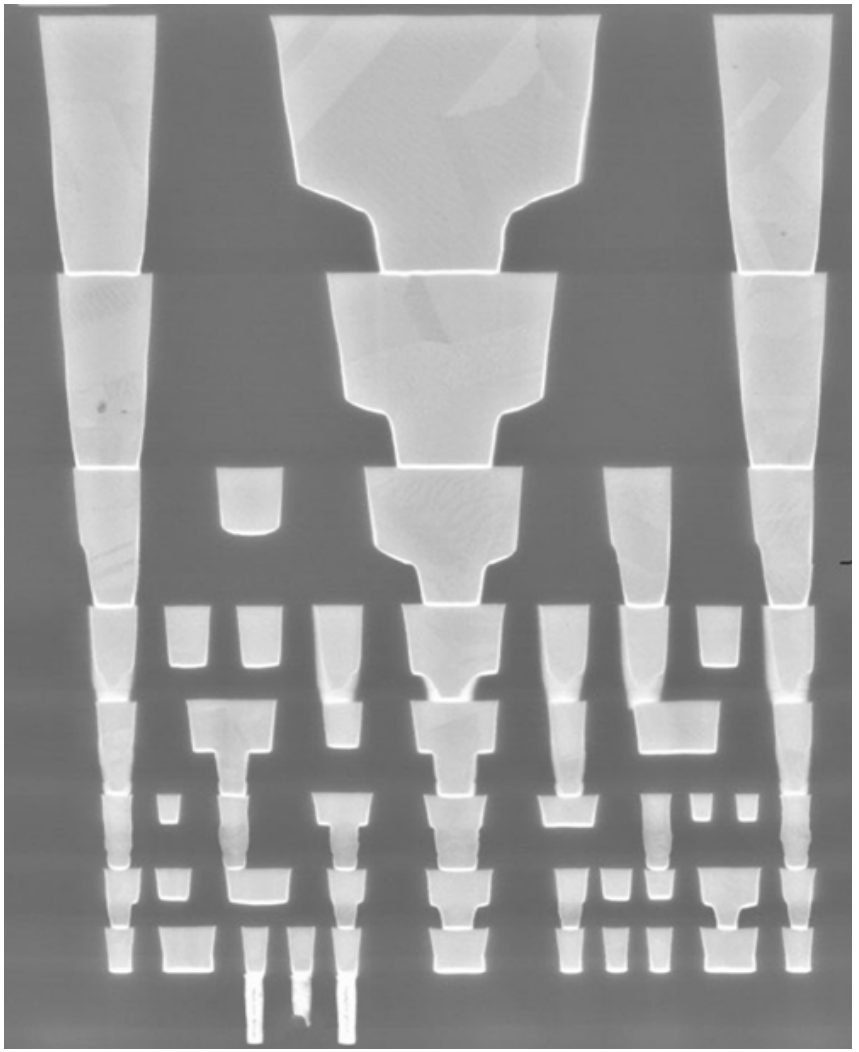


Source: Intel

Scaling - Technology Back-end (Intel 65nm)

- 1 km routing/cm^2 (2000)
  10 km routing/cm^2 (2016)
- Different metal layers have diverse characteristics.
  - Top level - lossless, bottom-level - lossy.
- Inter-Layer Dielectric (ILD) needs to have low *k*:

$$C = \varepsilon \cdot \frac{w}{t_{\text{oxide}}}$$

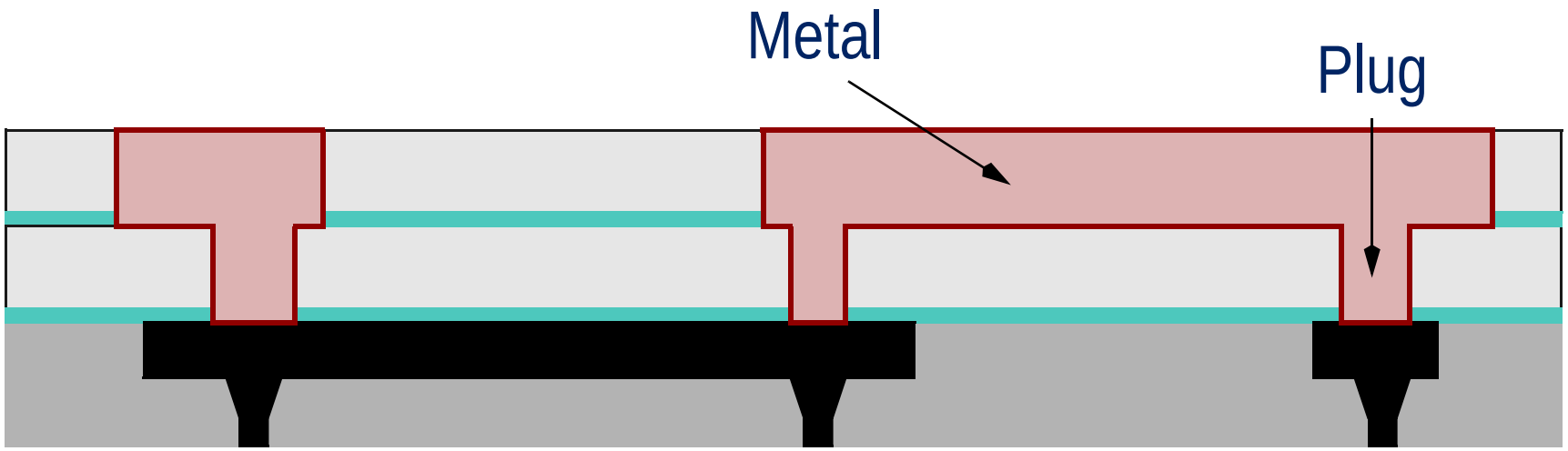


Source: Intel

Back-end Technology Development

- In the past, tungsten/aluminum (W/Al) and silicon dioxide (SiO_2).
- Nowadays, copper and low-*k* dielectrics.
- Low-*k* dielectrics.
  - Non-organic materials, e.g. SiOF with $\varepsilon_r = 3.5$.
  - Organic materials, e.g. polymers with $\varepsilon_r \sim 2$. Adhesion problems!
  - Intel's 90-nm process (2002):
    Carbon doped oxide (CDO) with $\varepsilon_r = 2.9$
    (yields a 4-mm signal propagation extension in our P4 example).
  - Current trademarks: MSQ, SiLK, Coral, Black Diamond, and Flowfill.

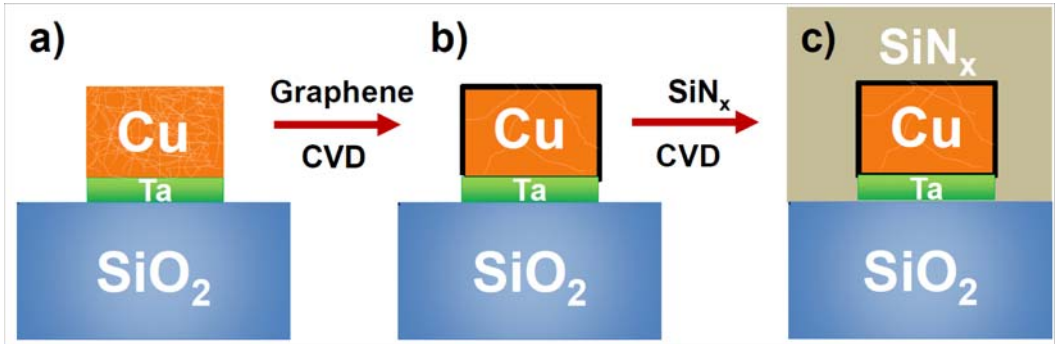
Formation of Wires and Contacts



- Oxide growth in two steps simplifies metal deposition (dual damascene).
- Tantalum nitride (TaN) for liner.
- Vias connect metal layers.
- Contacts connect metal-1 with transistors.

Next Barrier?

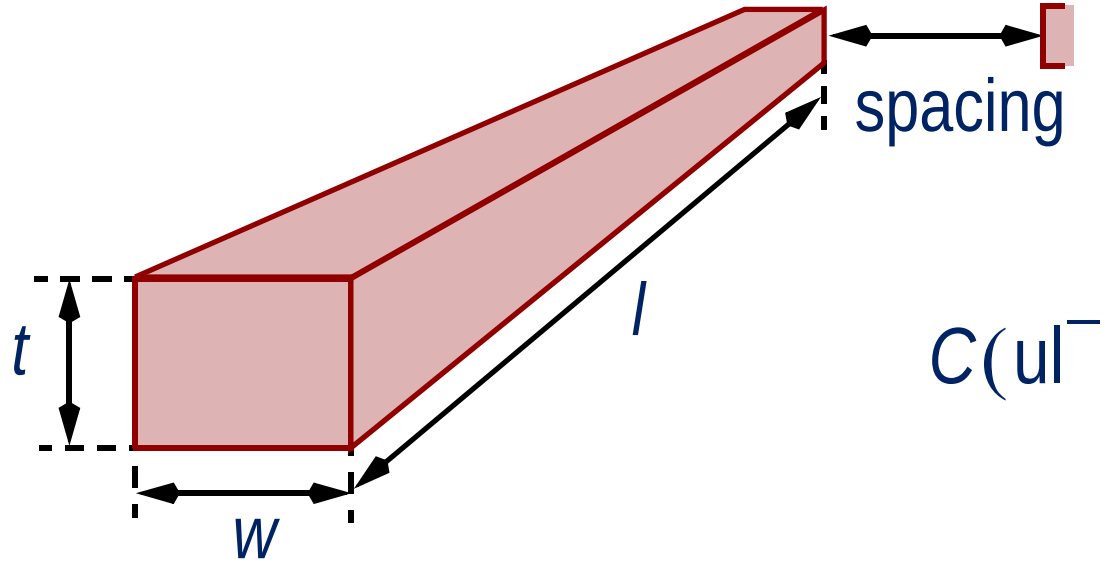
- Ongoing scaling and transistor improvements can make wires the next barrier.
- Graphene/Cu composite has 2X lower resistivity and 40X longer electromigration lifetime than as-deposited Cu
  [Li et al., “BEOL Compatible Graphene/Cu with Improved Electromigration Lifetime for Future Interconnects”, IEDM’16]



Source: Li et al. / IEDM’16

Wire Scaling

- Small wire cross sections at bottom layers ⇒ high resistance.



$$R(\text{ul}^{-1}) = \rho \cdot \frac{1}{wt}$$

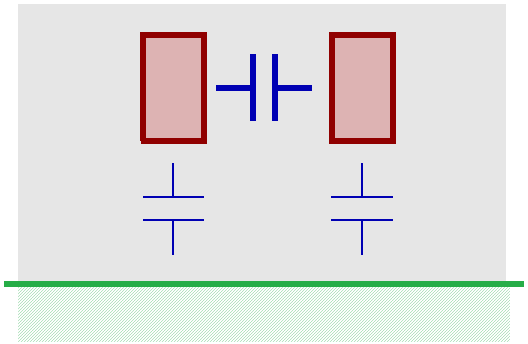
$$C(\text{ul}^{-1}) = \varepsilon \cdot \left(\frac{w}{t_{\text{oxide}}} + \frac{t}{\text{spacing}} \right)$$

ul = unit length

- w and t_{oxide} downscale ~0.7 per technology node generation; but t scales less to mitigate resistance effects.
  C remains constant with scaling, but R increases.

Wire Aspect Ratio Promotes Crosstalk

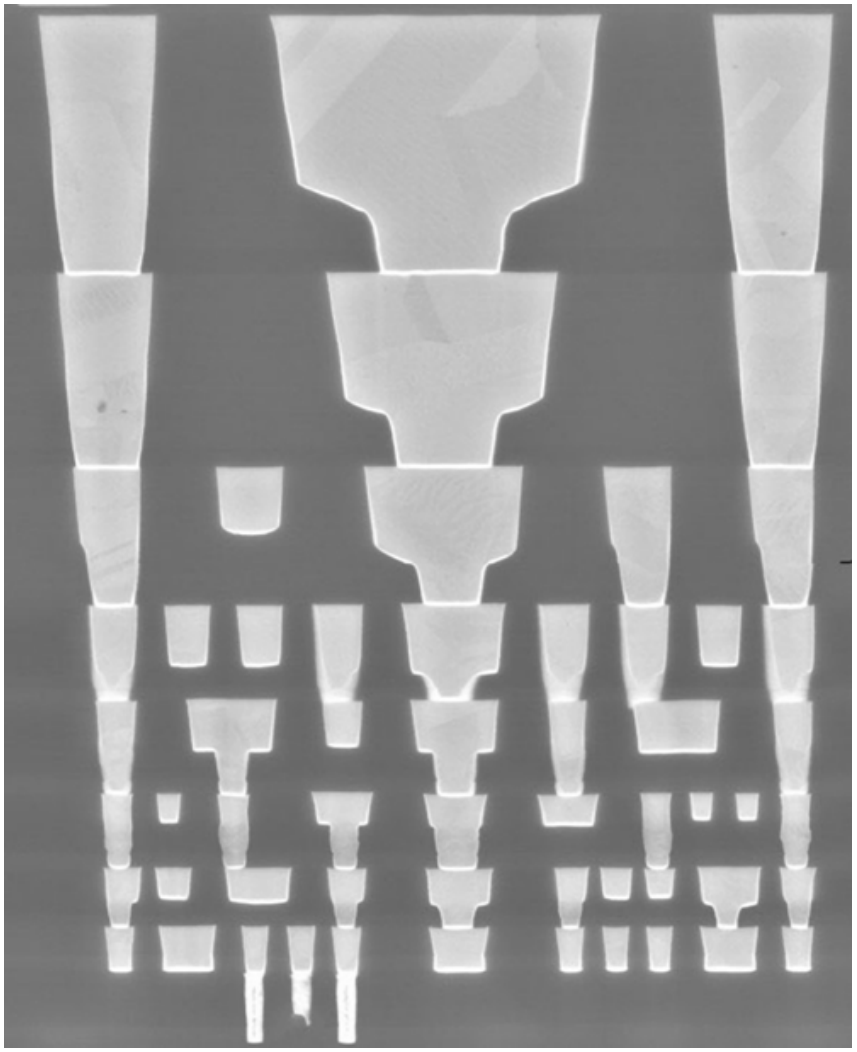
1. With technology scaling,
   wires grow taller and narrower.
2. As more metal layers are added,
   many layers are getting far away
   from ground reference (substrate).



Result:

Wire sidewall capacitance
dominates.

Intel 65 nm



Notice the wire aspect ratio.

After lab exercise 4, consider the impact of all the wires that are added in the routing phase.

Gains from Scaling

- Scaling improves integration density;
  historically each new technology node offered
  X scaling x Y scaling = 0.7 x 0.7 = 0.5 in area.
- The concept of technology nodes is no longer valid.
- Emerging issues:
  - Compensation circuits (fault and variation tolerance)
    offset scaling gains.
  - Post 10 nm wire issues: The liner needed for
    Cu interconnects (to prohibit diffusion) occupies precious area.

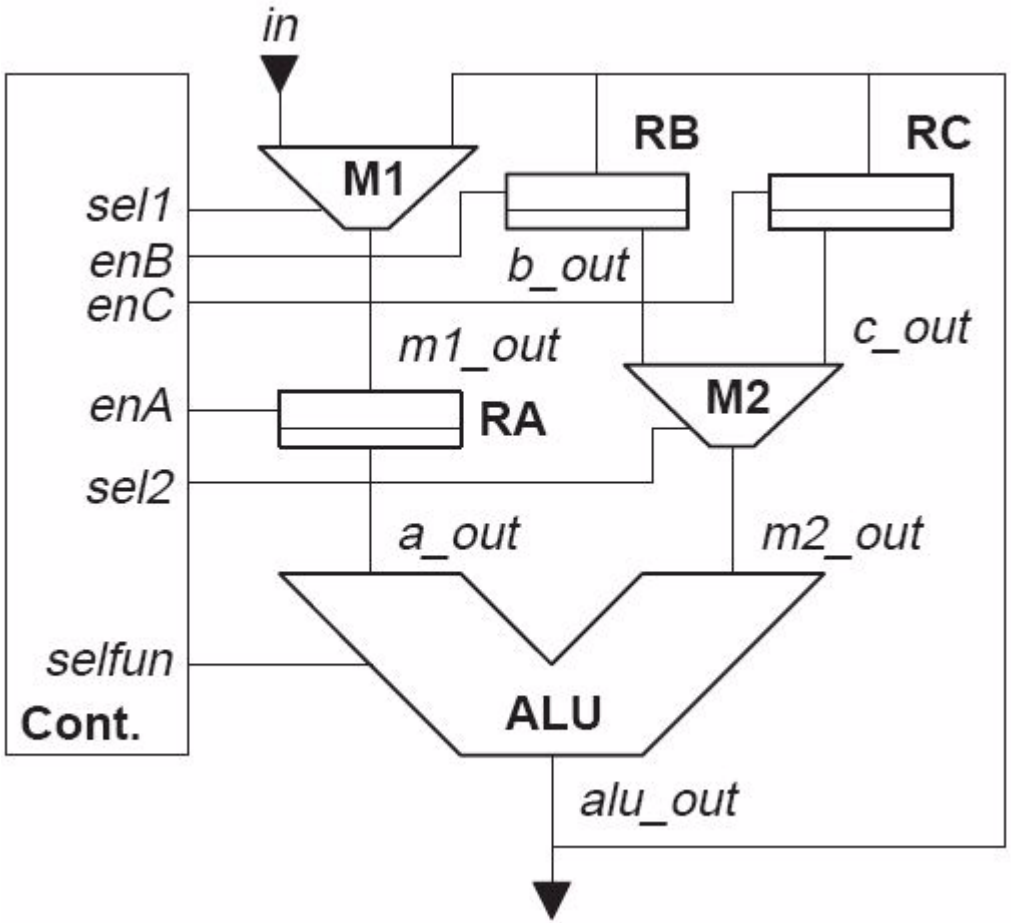
EDA for physical design
[*flow partly according to Gerez*]

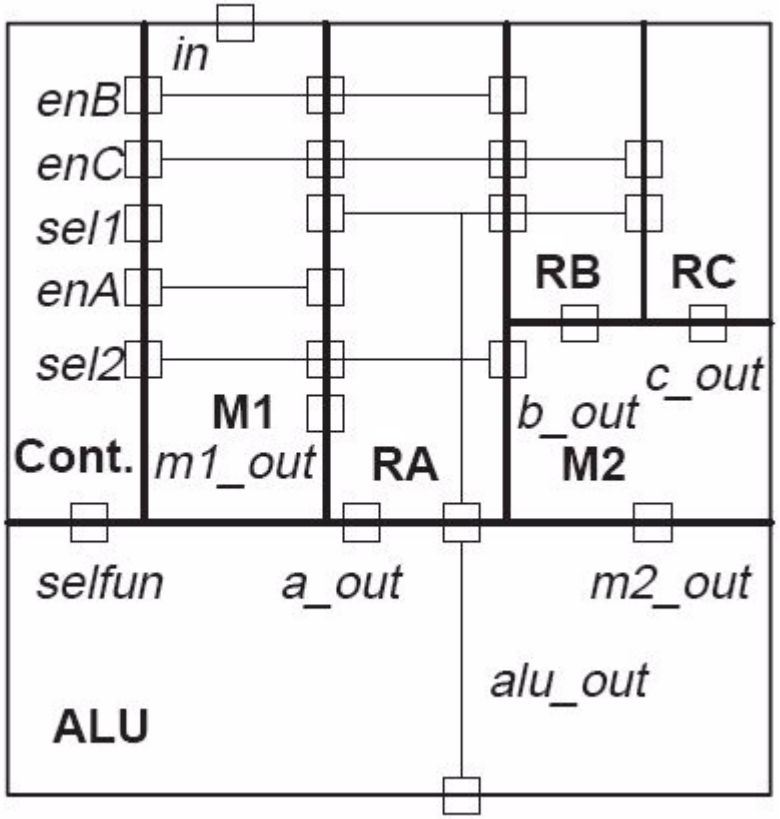
Physical Design

- Floorplanning.
  - Organization of big blocks.
  - User experience is critical. Manual floorplan is often the best.
- Placement.
  - Usually infers standard-cell placement.
  - The problem overwhelms human designers; EDA support needed.
- Routing.
  - Involves cell, power and clock routing.
  - The problem overwhelms human designers; EDA support needed.

Floorplanning

FLOORPLANNING EXAMPLE



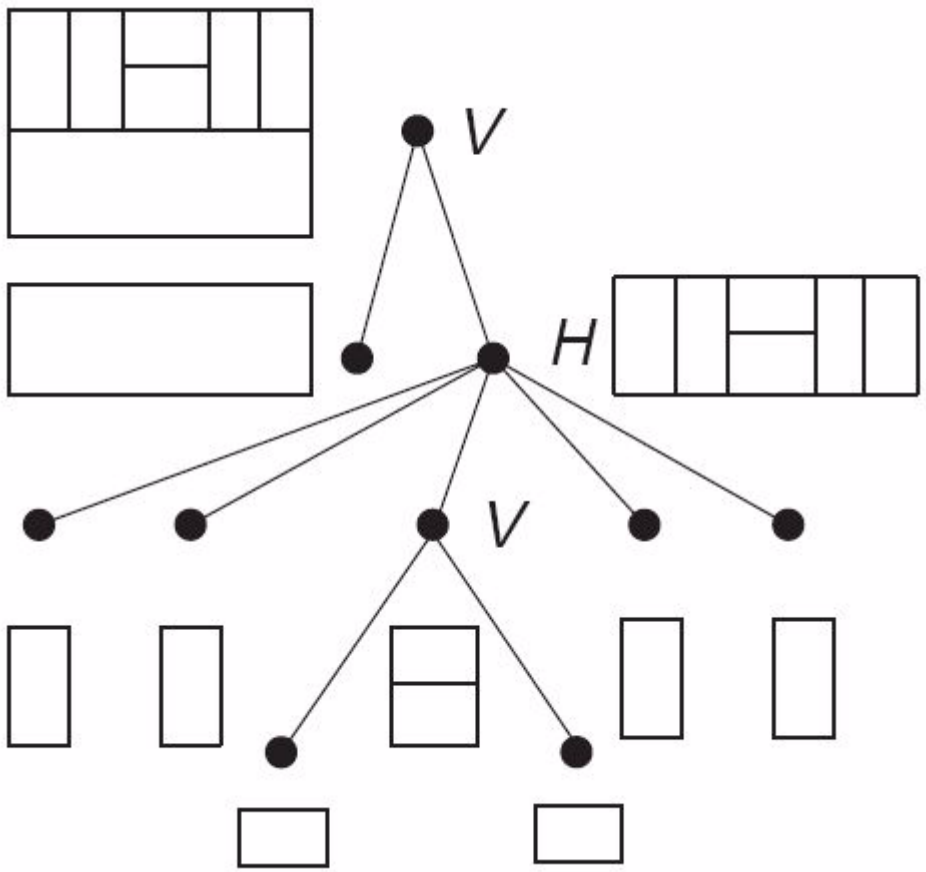


Floorplanning Concepts

- Abutment: Establishing connections between cells by putting them directly next to each other, without the necessity of routing.
- Leaf cell: A cell at the lowest level of the hierarchy; a leaf cell does not contain any other cell.
- Composite cell: A cell that is composed of either leaf cells or other composite cells. The entire IC is the highest-level composite cell.
- A restriction in floorplanning is that all leaf cells and composite cells are supposed to be rectangular.

Floorplan Representation

- A composite cell's subcells are obtained by a horizontal or vertical bisection of the composite cell.
- Slicing floorplans can be represented by a slicing tree.



Tasks of Automated Floorplanning

- Floorplan generation: How to construct the slicing tree, given a structural description (netlist).
  - Min-cut partitioning from placement can be used. (*Slide later.*)
  - Cost function example: *Cost* = *Area* + λ *Wire length*, where λ gives the relative importance of *Area* and *Wire length*.
- Floorplan sizing: Optimization of the floorplan area taking advantage of the flexibility in the cells.
- Generation of flexible cells: Generation of cell layouts once the cell shapes have been fixed by the sizing procedure.
- Other issues: Buffer allocation, wire congestion avoidance ...

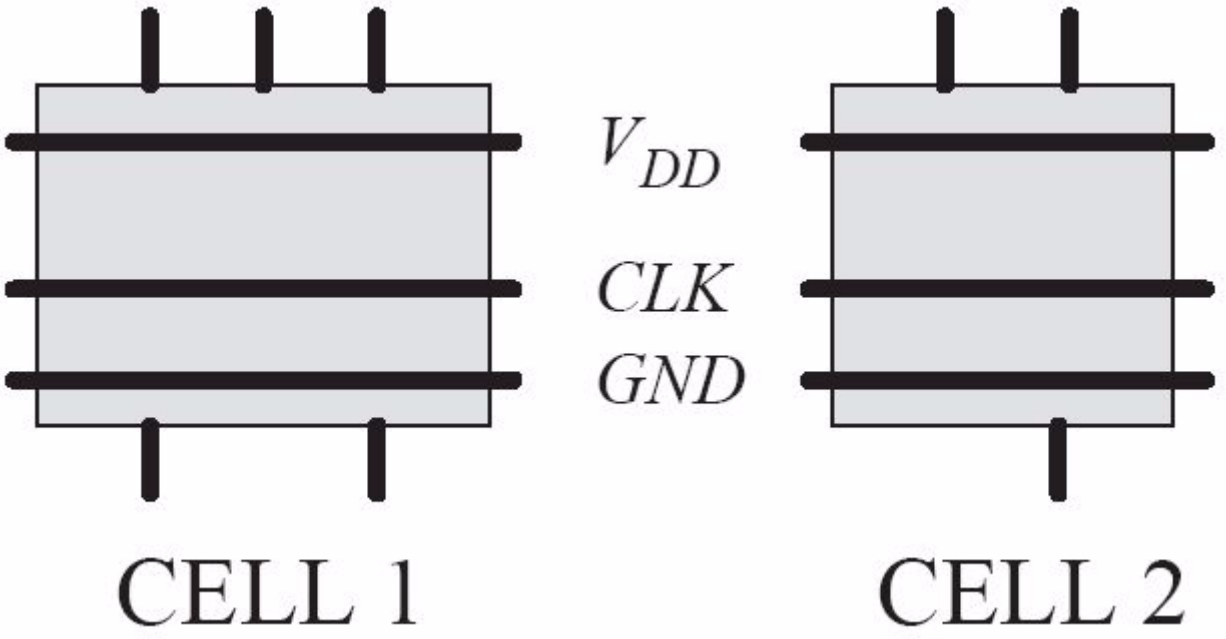
Placement

Placement

- Placement is the problem of automatically assigning correct positions to predesigned cells, such that some cost function is optimized.
- Different placement problems.
  - Standard-cell placement.
  - Building-block placement.
  - A combination of the above.

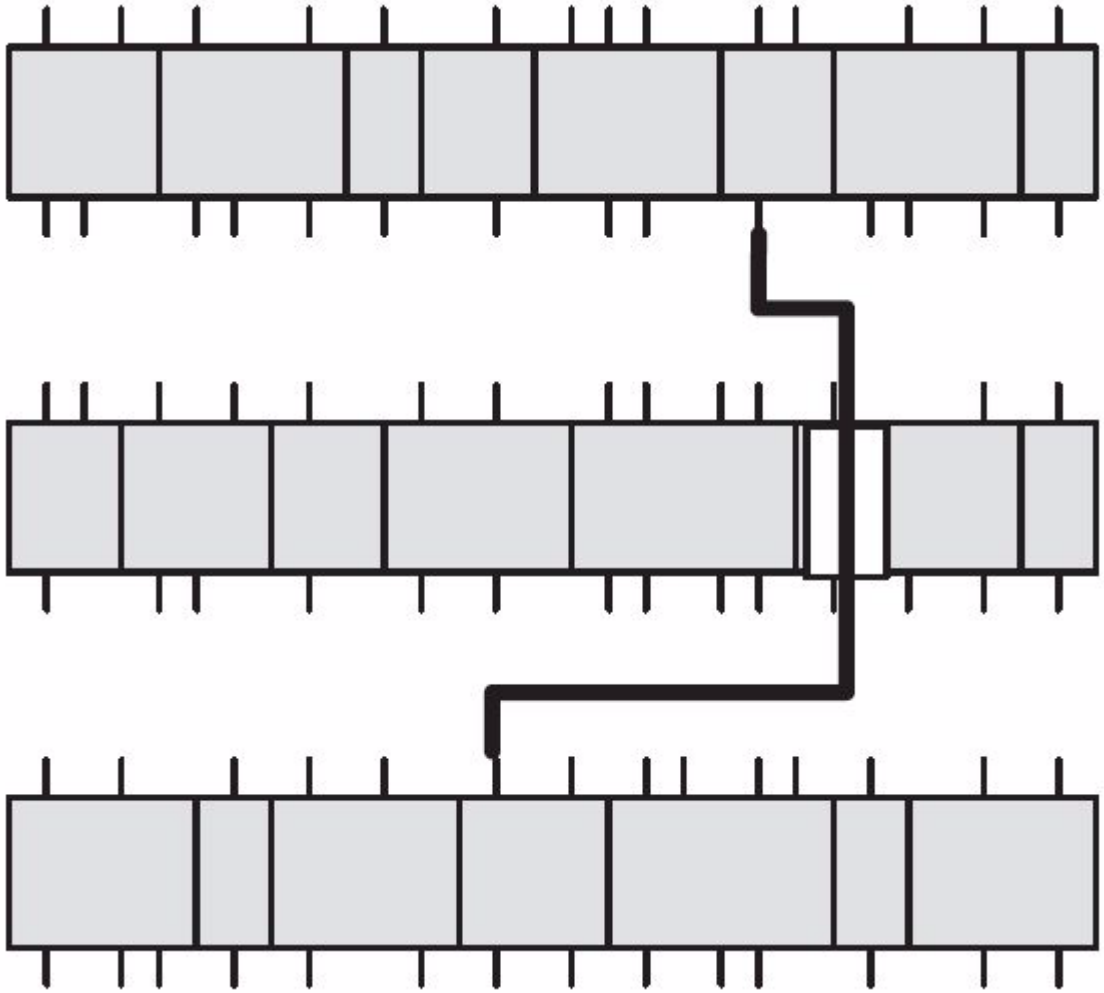
Standard-Cell Organization

- Standard cells are organized so that ...
  - power (VDD and ground) and clock wires run horizontally.
  - logic signals run vertically.



Standard Cell Rows

- One cell abuts to neighbor cells, left and right.
- Cell pitch is constant.
- Over-the-cell routing is often possible. (If not, use feedthrough cells.)
- Filler cells are used in empty space. These provide metal fill and decaps.



Cell Trends

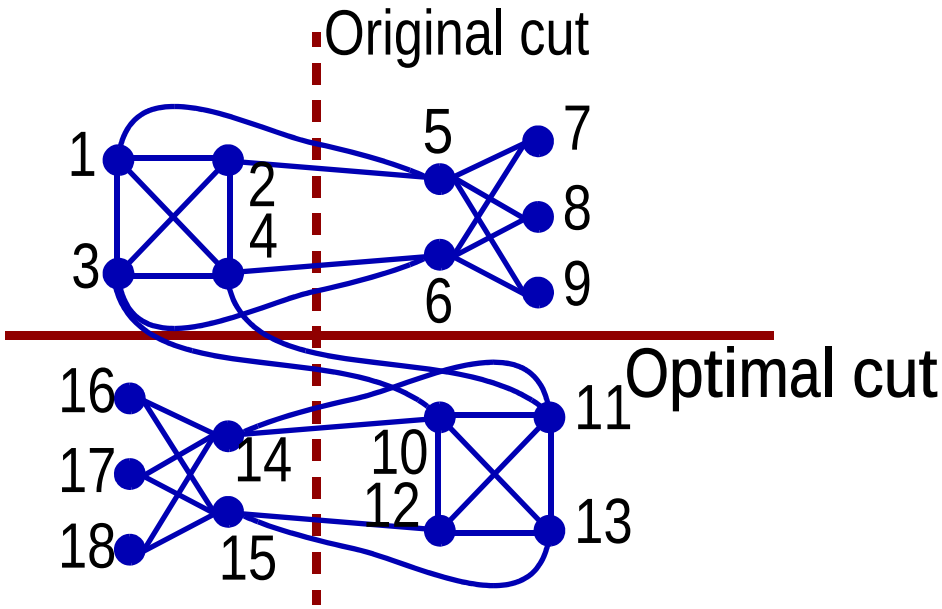
- Relevant metrics for cells of today:
  - contacted poly pitch (CPP) - related to width.
  - minimum metal pitch (MMP) - related to height/cell pitch.
- Examples:
  - Intel 14 nm: MMP = 52 nm and #tracks = 7.67 $\Rightarrow$ cell pitch 399 nm.
  - Intel 10 nm: MMP = 36 nm and #tracks = 7.56 $\Rightarrow$ cell pitch 272 nm.
  - TSMC 7 nm: cell pitch 240 nm.
  - Alliance (Samsung, IBM, and GlobalFoundries) 7 nm: MMP = 36 nm.
  - Compare to 22 nm's cell pitch $\approx$ 550-650 nm, or ST 65 nm with cell pitch of 2.6 μm.

Placement Algorithm

- What does a placement algorithm try to optimize?
  - Total area, or
  - total wire length, or
  - number of horizontal/vertical wire segments crossing a line.
- Constraints.
  - The placement should be routable.
  - Timing constraints (some wires should always be shorter than a given length).
- There exist both top-down and bottom-up strategies.

Top-Down Placement - Partitioning

Reduce a cost function (cut size) based on interconnection count and priority



Step no	Vertex pair	Resulting cut size
0	-	10
1	{4,10}	12
2	{2,12}	12
3	{1,13}	8
4	{3,11}	2
5	{7,18}	6
6	{8,17}	10
7	{5,15}	12
8	{9,16}	12
9	{6,14}	10

Kernighan-Lin bipartitioning (min-cut)

Bottom-Up Placement

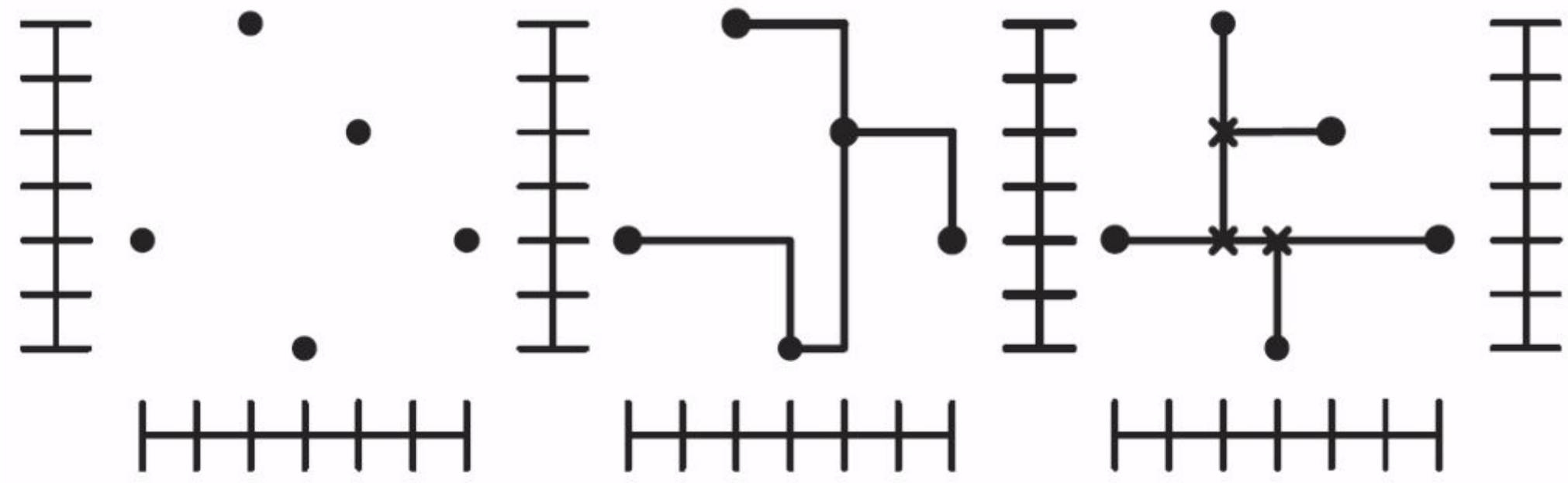
- Constructive algorithm followed by iterative algorithm:
  1. Initial so-called clustering:
     Tie together gates according to local interconnectivity.
     Starting from a seed cell, choose neighbor cells
     based on interconnection weight
     (those can e.g. include timing priorities).
  2. Improvement through iteration, for example
     simulated annealing. Perturb the clustered
     solution into something better.

Wire-Length Estimation

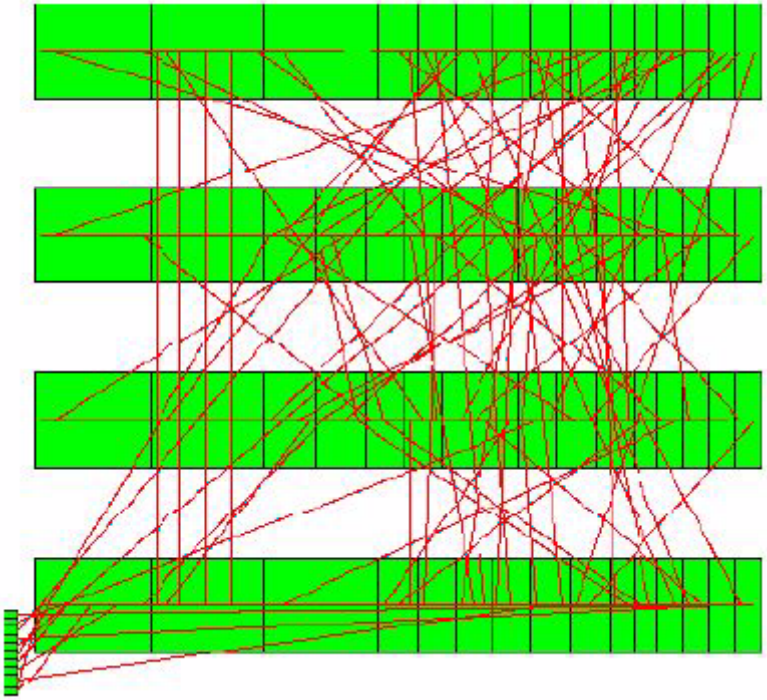
- Ideally, placement and routing (P&R) should be performed simultaneously as they depend on each other's results. This is, however, complicated ...
- In practice, most of the placement is done prior to routing. The placement algorithm estimates the wire length of a wire net using some metric (the wire load model).
  - Half of the perimeter of the rectangle enclosing all terminals in a net.
  - Minimum rectilinear spanning/Steiner tree.
  - Squares of all pairwise terminal distances in a net using a quadratic cost function.

Example of Wire Length Approximation

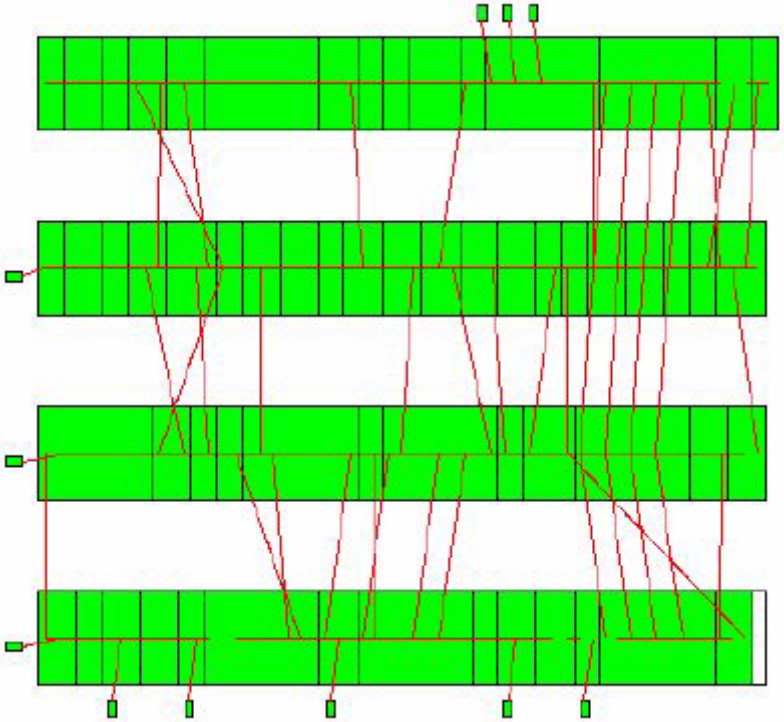
- The minimum spanning tree (MST) algorithm (middle) can give an approximation of shortest possible wire net length (right).



Impact of Placement Results



Poor quality ⇒
high power, long delay,
large area.



Source Keutzer/Kahng

Good quality ⇒
low power, short delay,
small area.

Routing

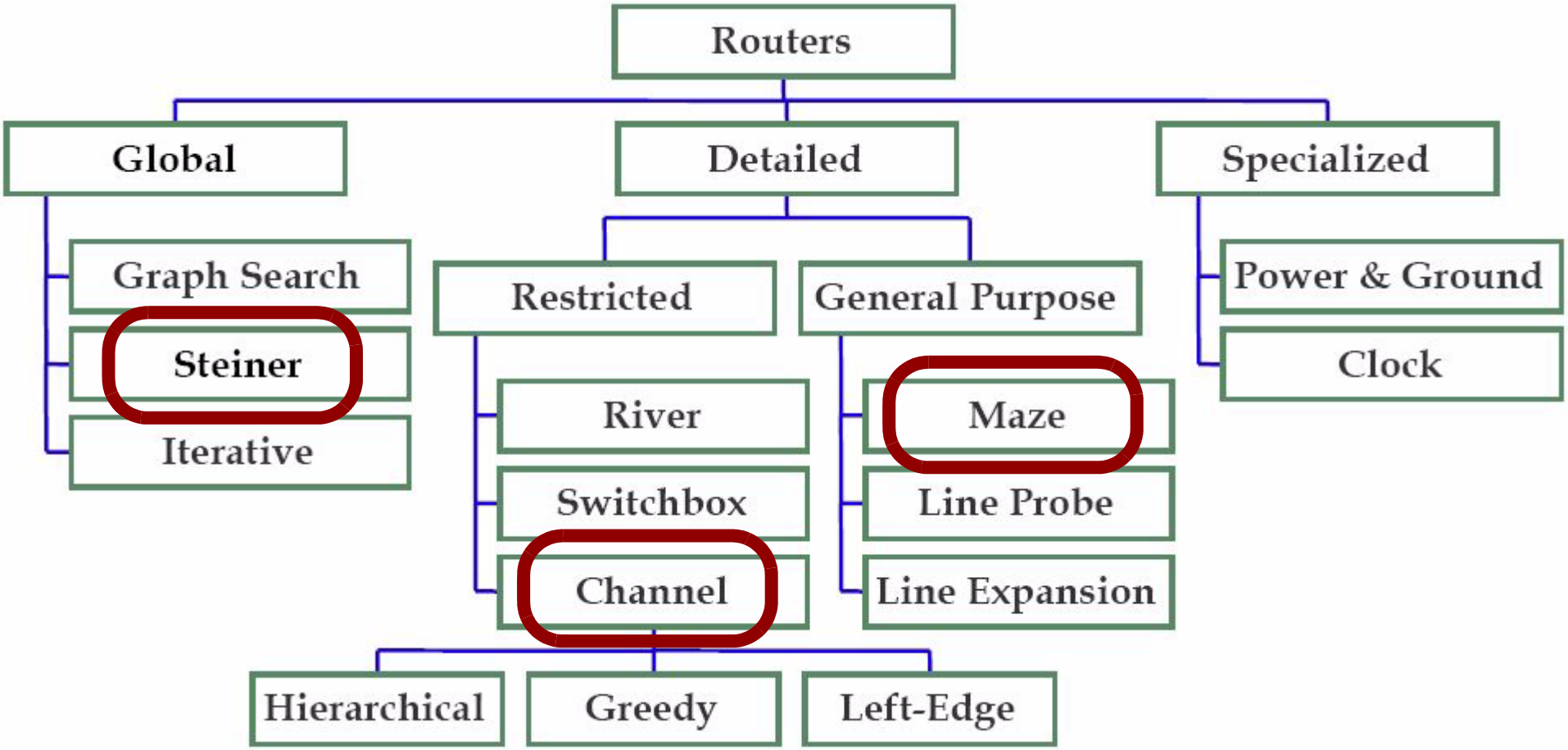
Overview on Routing

- Global routing: To establish through which routing areas a connection will run.
- Local or detailed routing: To find the exact interconnections for a set of terminals in a given routing area.
- Special routing.
  - Power routing for VDD and ground.
  - Clock routing. This involves Clock Tree Synthesis (CTS), in which routing and buffer insertion and sizing are integrated.
- General routing criteria:
  Total wire length, congestion, crosstalk, timing, etc.

Logic Signal Routing

- Placement determines position of cells = position of terminals.
- Global routing (is also used in floorplanning and placement).
  - Identify routing resources to be used.
  - Identify layers (and tracks) to be used.
  - Assign particular nets to these resources.
- Detailed routing.
  - Defines actual pin-to-pin connections.
  - Must understand most or all design rules.
  - May use a compactor to optimize result.

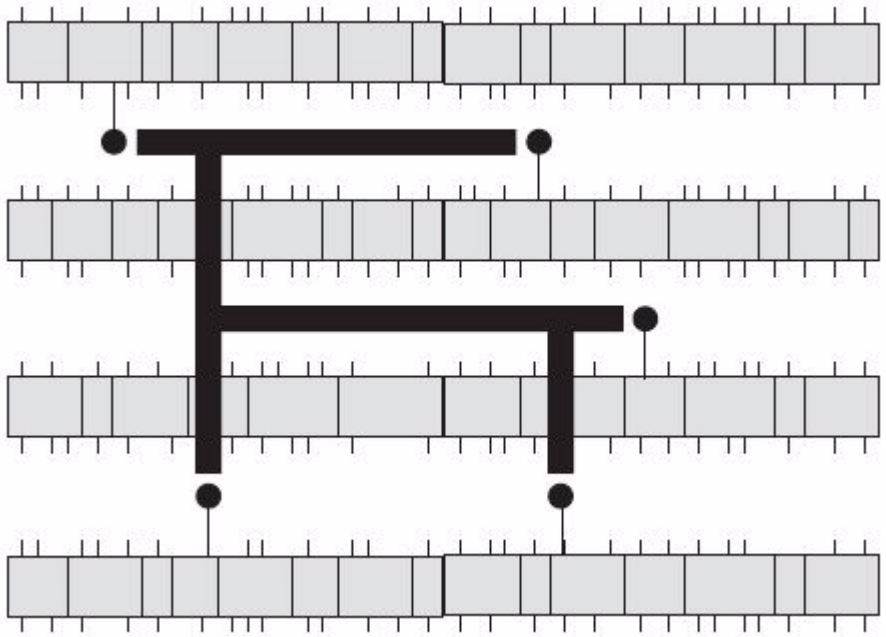
Map of Router Principles

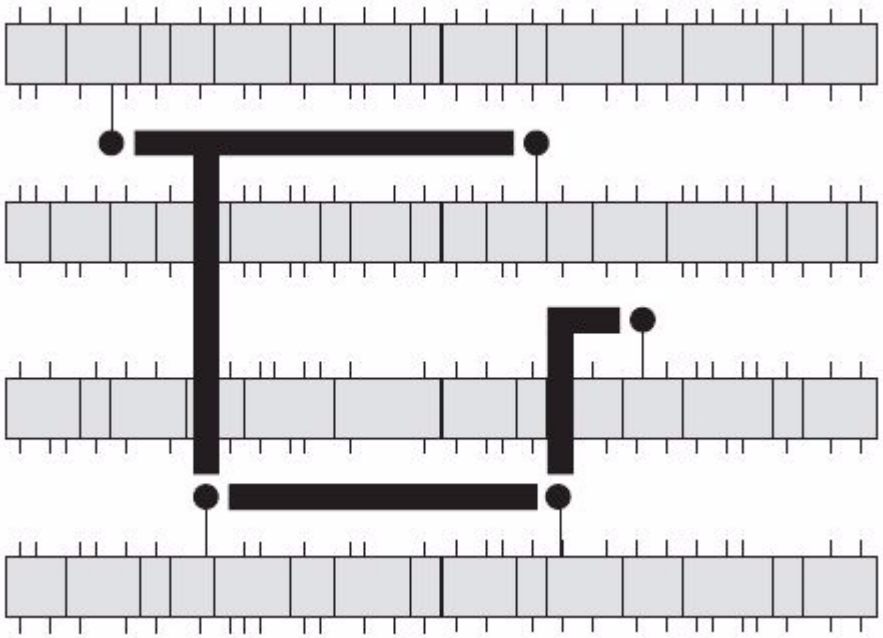


source Keutzer

Global Routing

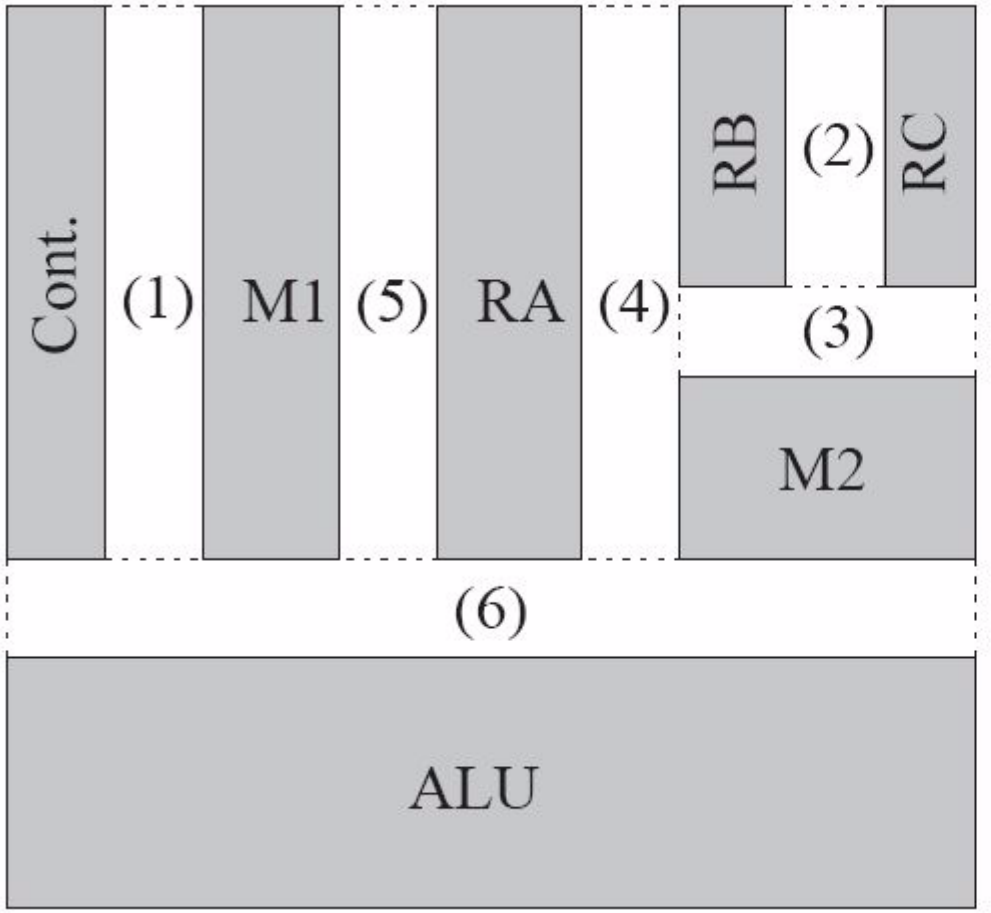
- Establishes the shapes of the connections for each net by distributing the wiring segments among the available channels.
- Each shape is a rectilinear Steiner tree.

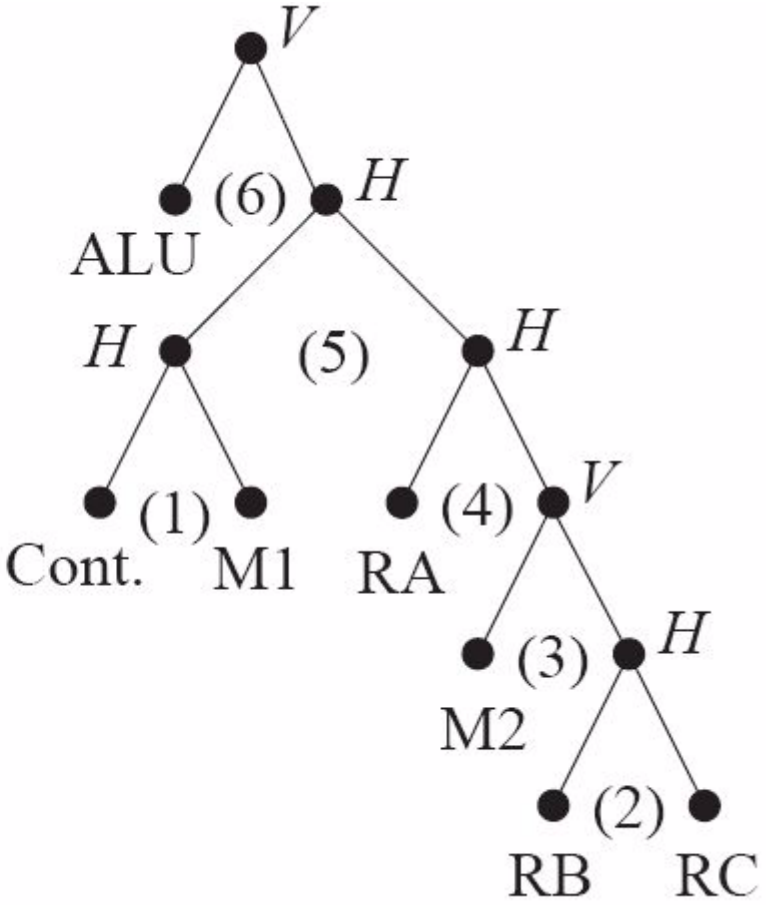




FINDING THE CHANNELS

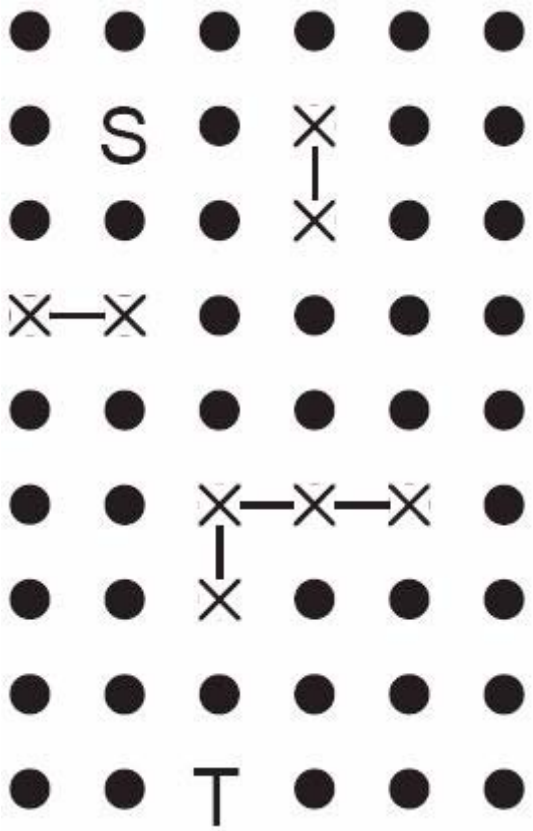
- Slicing resulting from floorplan can guide channel allocation.





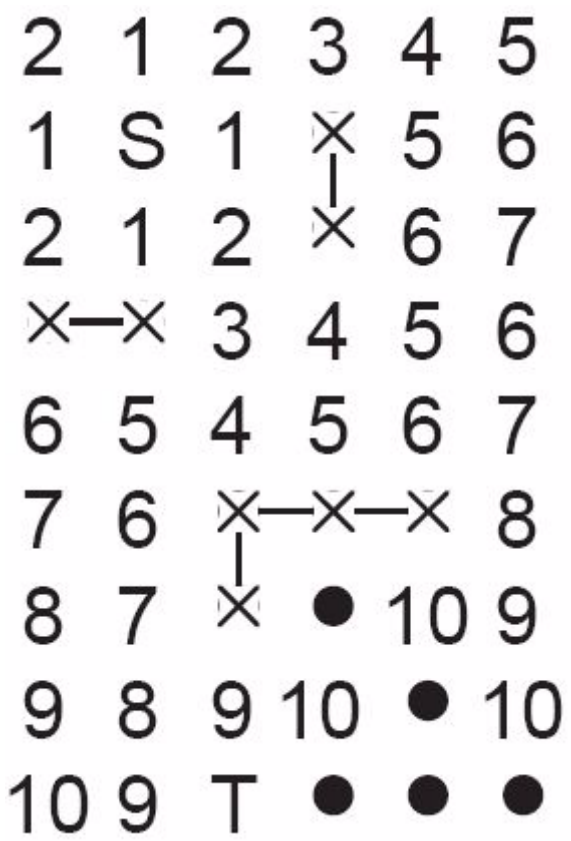
Maze Routing (Lee's Algorithm) 1(4)

- We wish to route a wire from source (S) to sink (T).
- But there exist obstacles (X = previously routed wires).



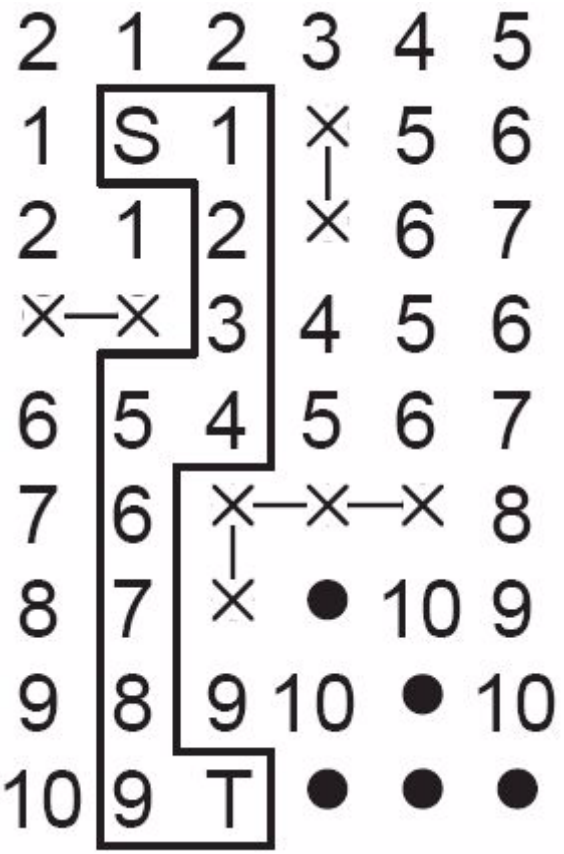
Maze Routing (Lee's Algorithm) 2(4)

- A wave of incrementing numbers emanates from the source, until the sink is reached.



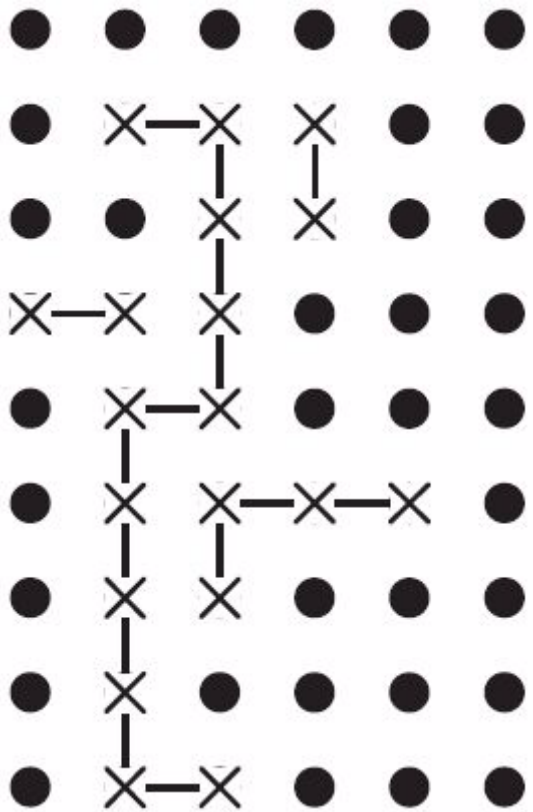
Maze Routing (Lee's Algorithm) 3(4)

- Backtracking the shortest distance gives the routing pattern.



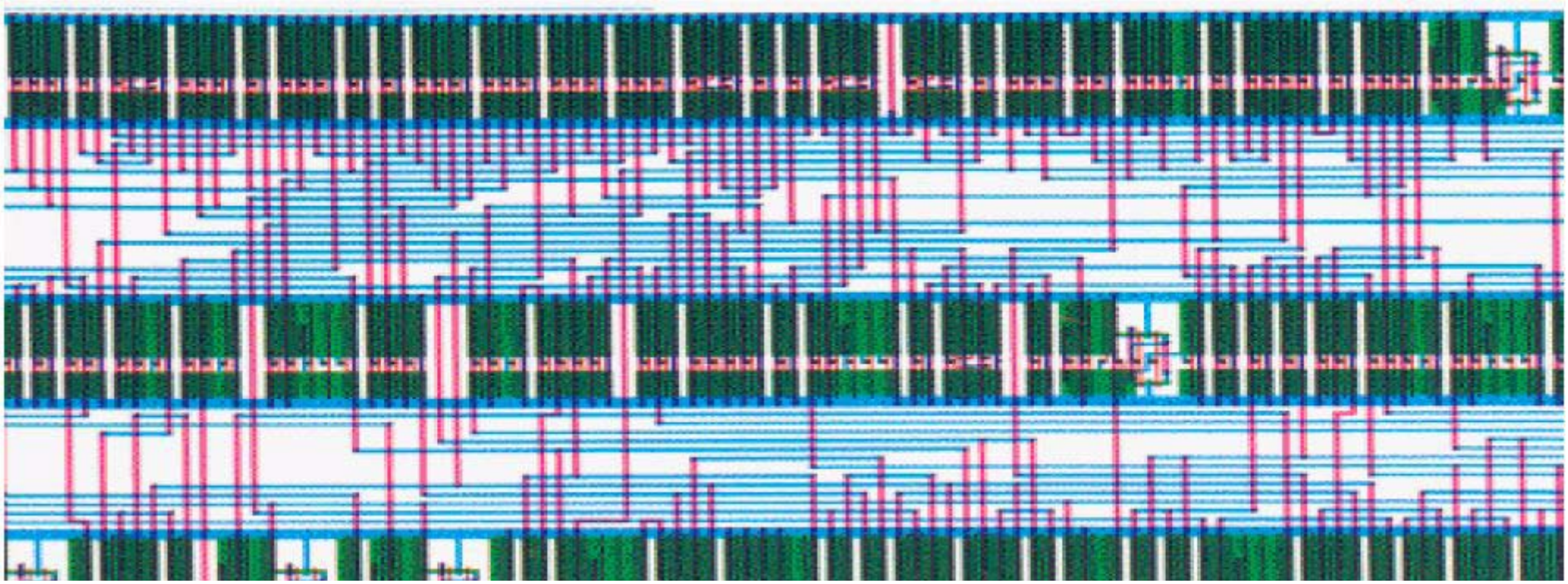
Maze Routing (Lee's Algorithm) 4(4)

- The chosen path is selected and becomes an obstacle for future routing.



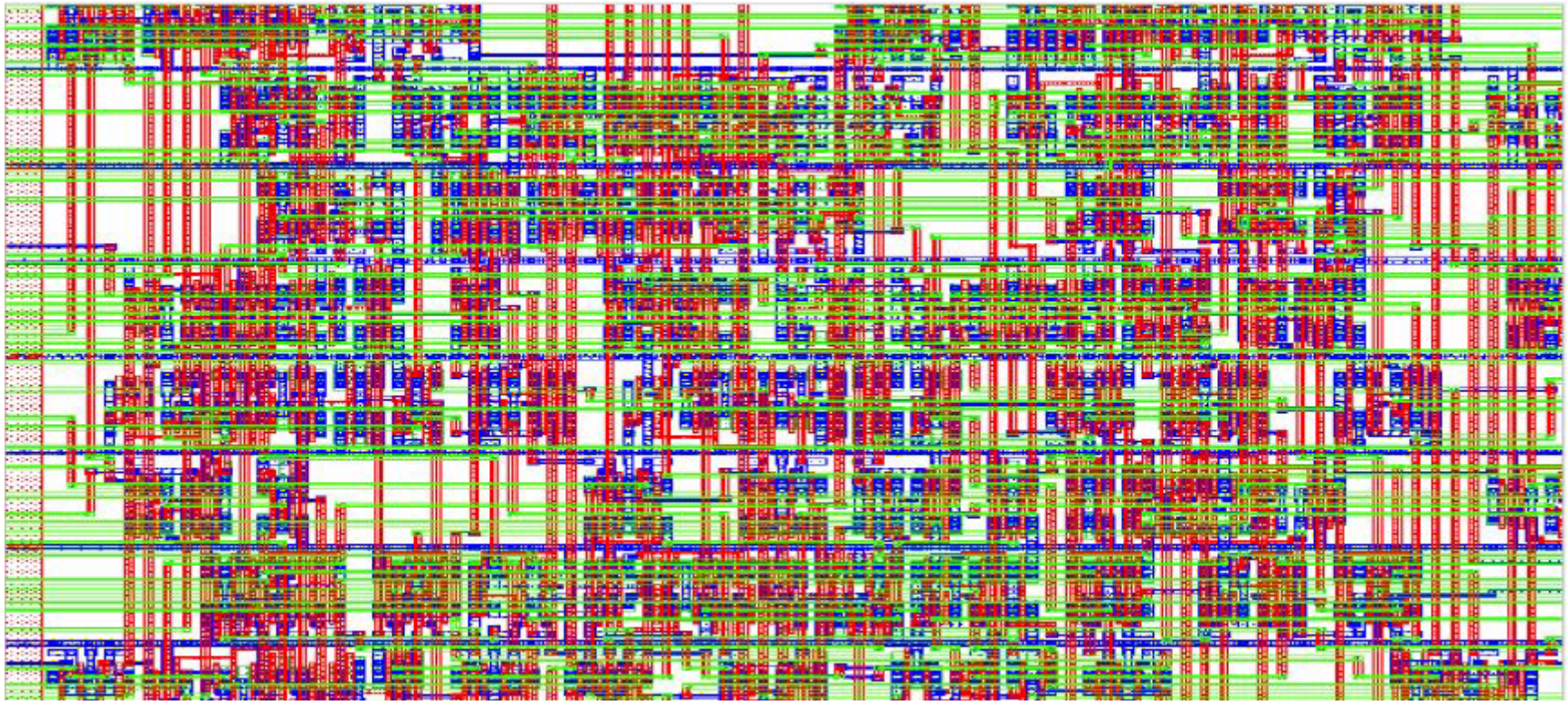
Fixed Grid - Detailed Channel Routing

1. Feedthrough cells are needed to go through rows.
2. Design may use metal-1 or metal-2.
3. Cells must bring primary logic signals out to the channel.



Area Routing

1. Wiring can go over cells.
2. Careful layout of cells: minimize routing obstacles, i.e. use of m1/m2.
3. Cells must not bring signals to channel; the route will come to them.



Physical Design: Conclusion

- Consider the wires!
- Early floorplan.
  - Use designer's understanding.
- Placement & Routing.
  - Make use of EDA software support.