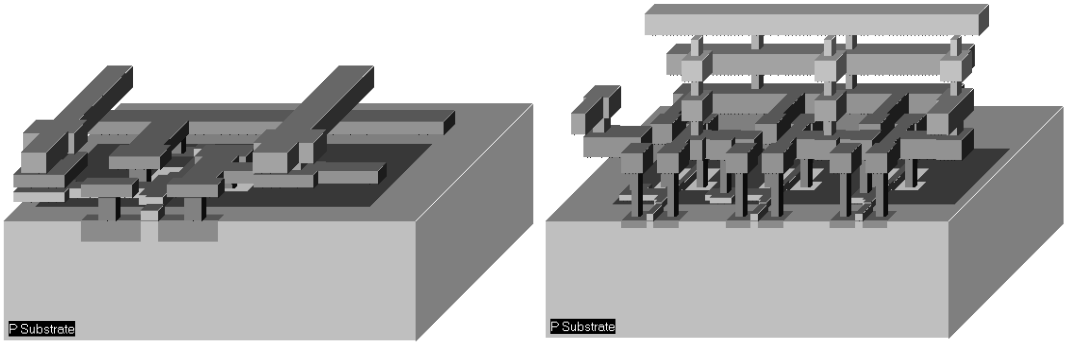
5 Interconnects

1. Introduction

The role of interconnects in integrated circuit performances has considerably increased with the technology scale down. Figure 5-1 shows the evolution of the aspect of the integrated circuit. In 0.12µm, 6 to 8 metal layers are available.



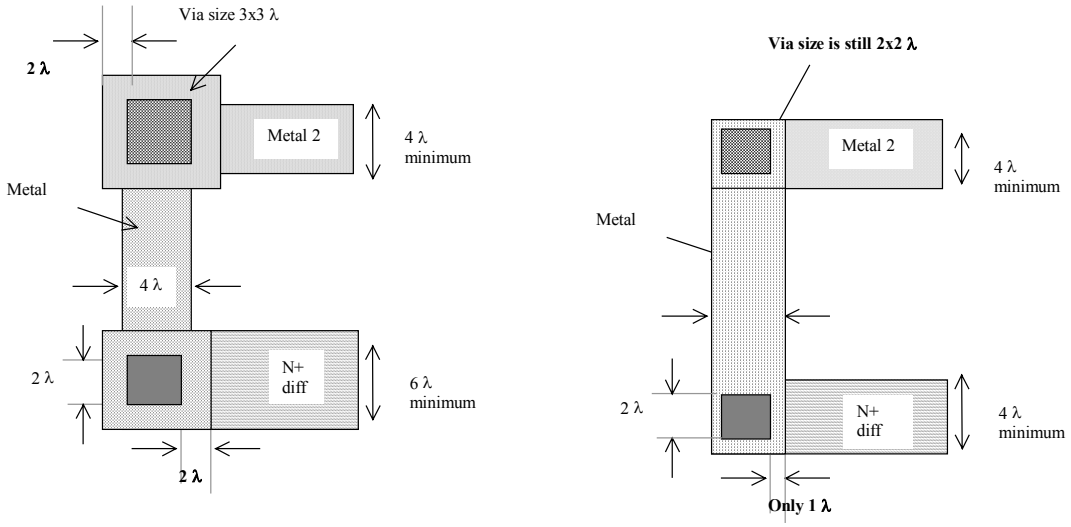
(a) 0.7µm *(b) 0.12µm technology*

Figure 5-1: Evolution of interconnect between 0.7µm technology and 0.12µm technology (Inv3.MSK)

2. Metal Layers

In the previous chapter, we designed the CMOS inverter using two layers of metal. However, up to 6 metal layers are available for signal connection and supply purpose. A significant gap exists between the 0.7µm 2-metal layer technology and the 0.12µm technology in terms of interconnect efficiency.

Firstly, the contact size is 6 lambda in 0.7µm technology, and only 4 lambda in 0.12µm. This features a significant reduction of device connection to metal and metal2, as shown in figure 5-2. Notice that a MOS device generated using 0.7µm design rules is still compatible with 0.12µm technology. But a MOS device generated using 0.12µm design rules would violate several rules if checked in 0.7µm technology.

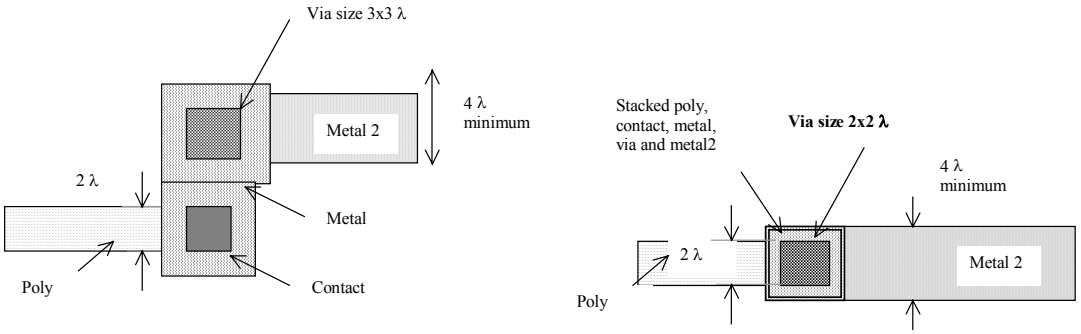


(a) Contacts in 0.7µm technology

(b) Contacts in 0.12µm technology

Figure 5-2: Contacts in 0.7µm technology require more area than in 0.12µm technology

Secondly, the stacking of contacts is not allowed in micro technologies. This means that a contact from poly to metal2 requires a significant silicon area (Figure 5-3a) as contacts must be drawn in a separate location. In deep-submicron technology (Starting 0.35µm and below), stacked contacts are allowed (Figure 5-3b).



(a) Poly to metal2 contact in 0.7µm technology

(b) Poly to metal2 contact in 0.12µm technology

Figure 5-3: Stacked vias are allowed in 0.12µm technology, which saves a significant amount of silicon area compared to 0.7µm design style.

Metal layers are labeled according to the order in which they are fabricated, from the lower level 1 (metal 1) to the upper level (metal 6 in 0.12µm). Each layer is embedded into a silicon oxide (SiO2) which isolates layers from each other. A cross-section of a 0.12µm CMOS technology is shown in figure 5-4.

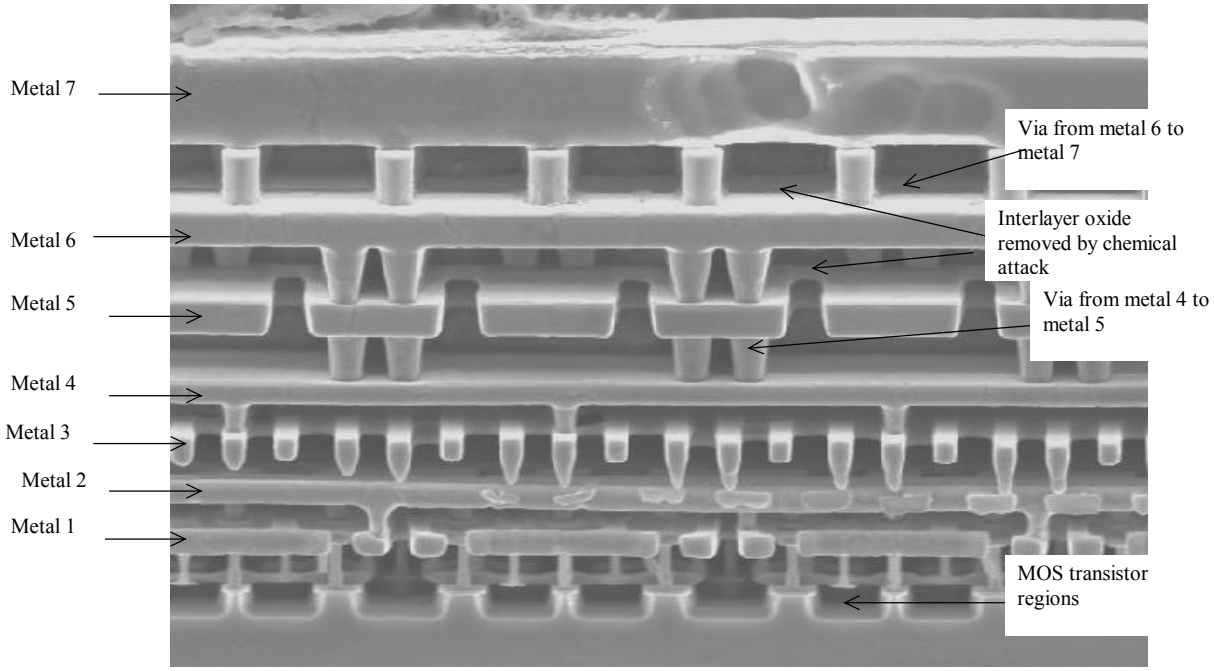


Figure 5-4: Cross-section of a 0.12µm technology (Courtesy Fujitsu)

3. Contact & Vias

The connection material between diffusion and metal is called "contact". The same layer is also used to connect poly to metal, or poly2 to metal. The connection material between metal and metal2 is called "via". By extension, the material that connects metal2 to metal3 is "via2", metal3 to metal4 "via3", etc..

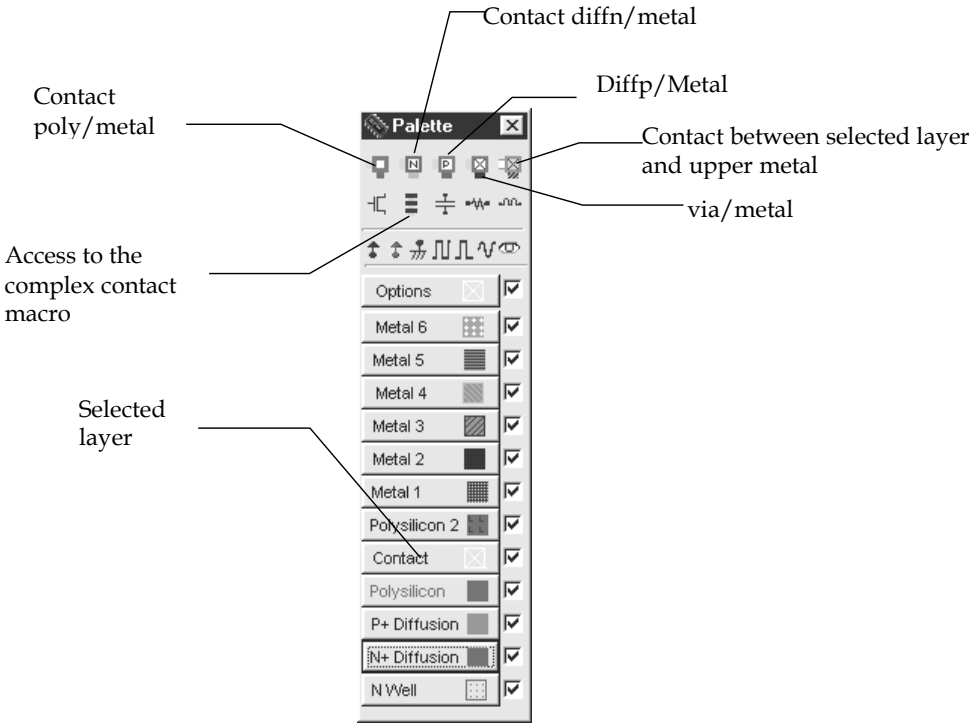


Figure 5-5: Access to basic contact macros

In Microwind, specific macros are accessible to ease the addition of contacts in the layout. These macros may be found in the palette, as shown in figure 5-6. As an example, you may instanciate a design-error free poly/metal contact by a click on the upper left corner icon in the palette. You may obtain the same result by drawing one box of poly (4x4 lambda), one box of metal (4x4 lambda) and one box of contact (2x2 lambda), according to design rules.

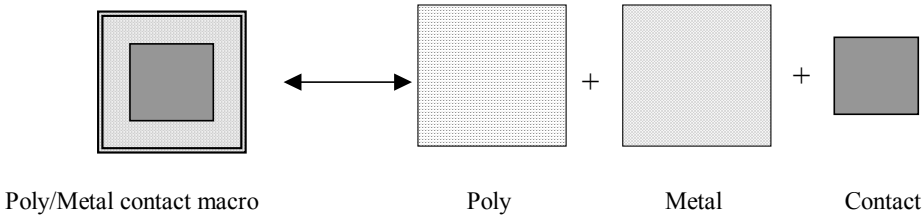


Figure 5-6: Access to basic contact macros

Additionally, an access to complex stacked contacts is proposed thanks to the icon "complex contacts" situated in the palette, second row, second column. The screen reported in figure 5-7 appears. By default you create a contact from poly to metal1, and from metal1 to metal2. Change the tick to build more complex stacked contacts.

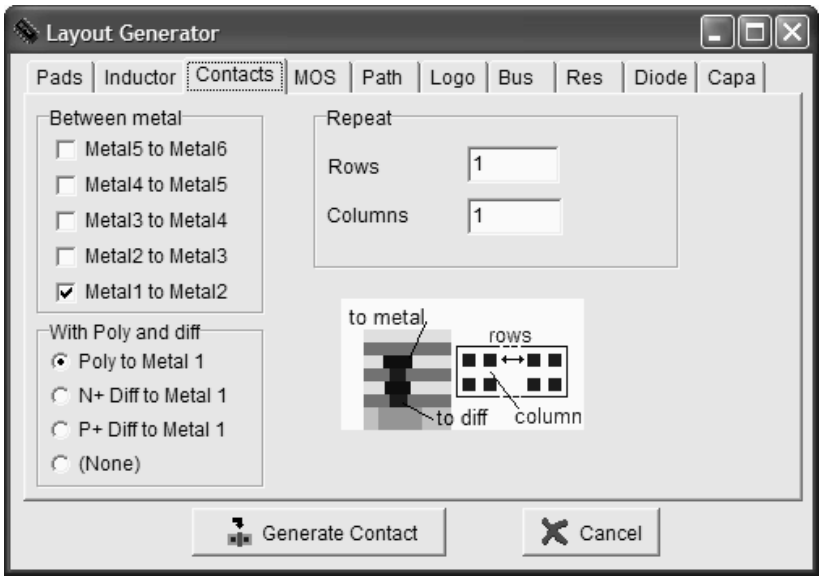


Figure 5-7: Access to complex stacked contact generator

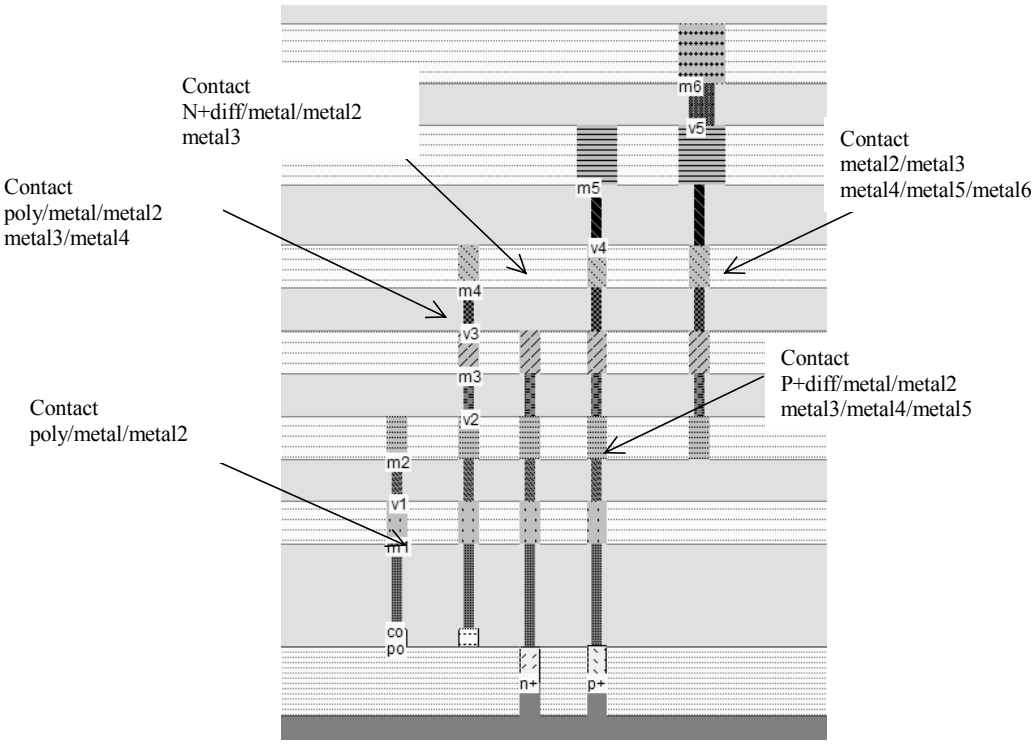


Figure 5-8: Examples of layer connection using the complex contact command from Microwind (Contacts.MSK)

DIRECT LAYER CONNECTION

A convenient command exists in Microwind to add the appropriate contact between two layers. Let us imagine that we need to connect two signals, one routed in polysilicon and an other in metal3. Rather than invoking the complex macro

command, we may just select the icon "connect layers". As a result a stack of contacts is inserted at the desired location to connect the lower layer to the upper layer. An illustration of this command is shown in figure 5-9.

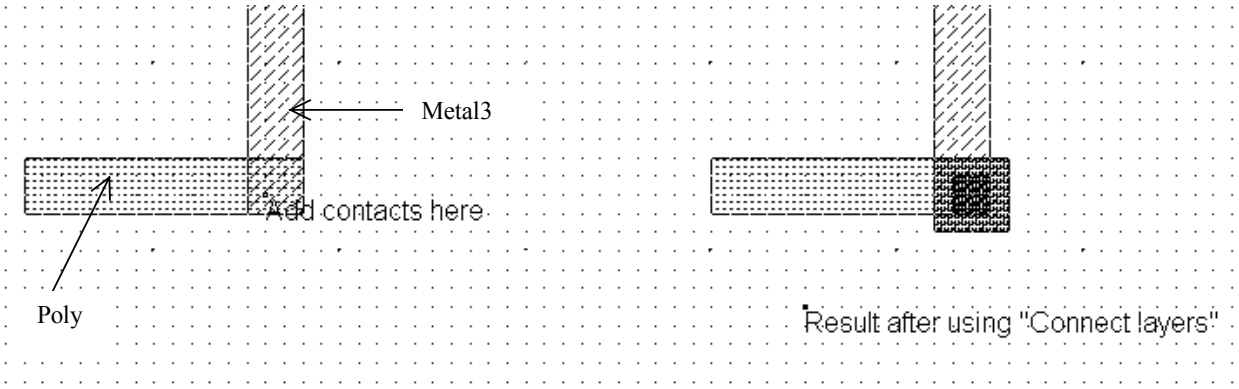


Figure 5-9: The command "Connect layers" inserts the appropriate stacked contacts to build the connection between the desired layers (ConnectLayers.MSK)

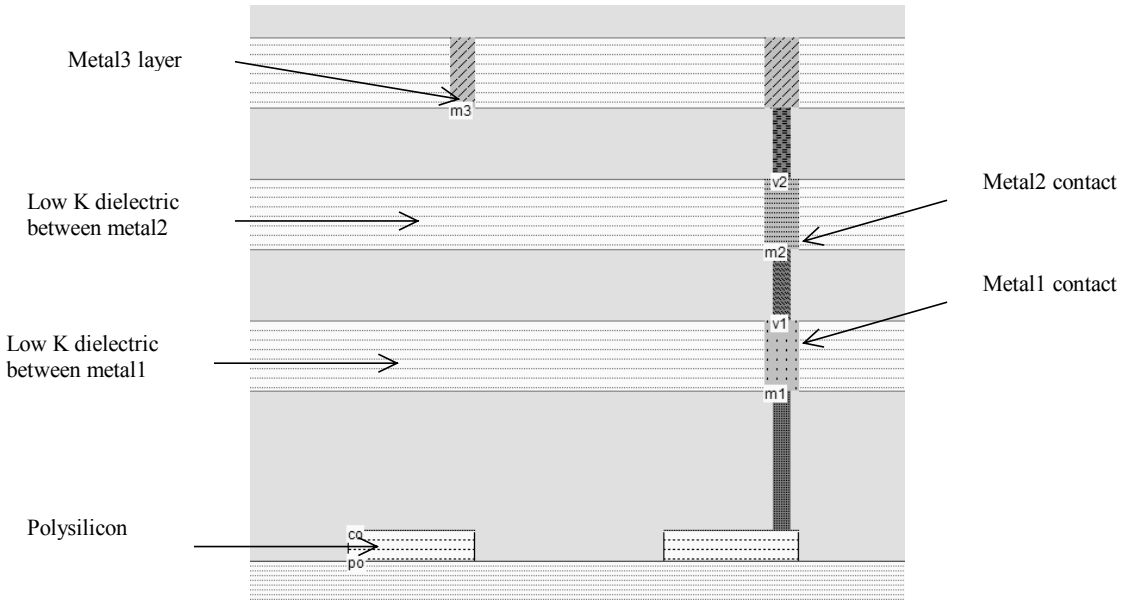


Figure 5-10: 2-D cross-section of the layout before and after connecting poly and metal3 layers(ConnectLayers.MSK)

4. Design rules

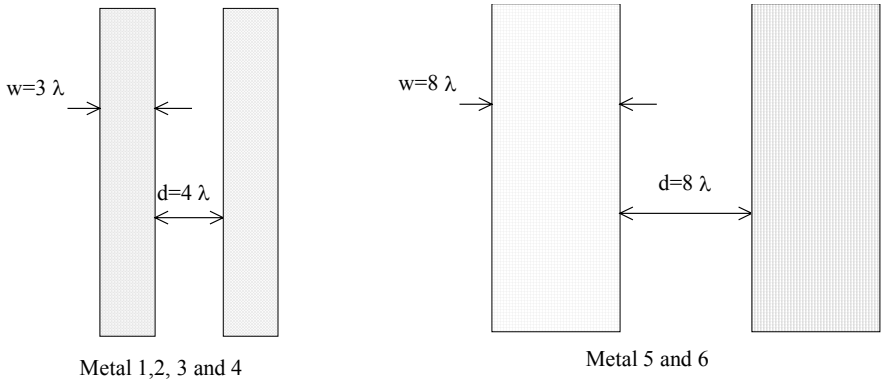


Figure 5-11: Minimum width w and distance d between metal layers

In 0.12µm technology, the metal layers 1, 2 , 3 and 4 have almost identical characteristics. Concerning the design rules, the minimum size w of the interconnect is 3 lambda. The minimum spacing is 4 lambda (Figure 5-11). In Microwind, each interconnect layer is drawn with a different color and pattern. Examples of minimum width and distance interconnects are reported in figure 5-12.

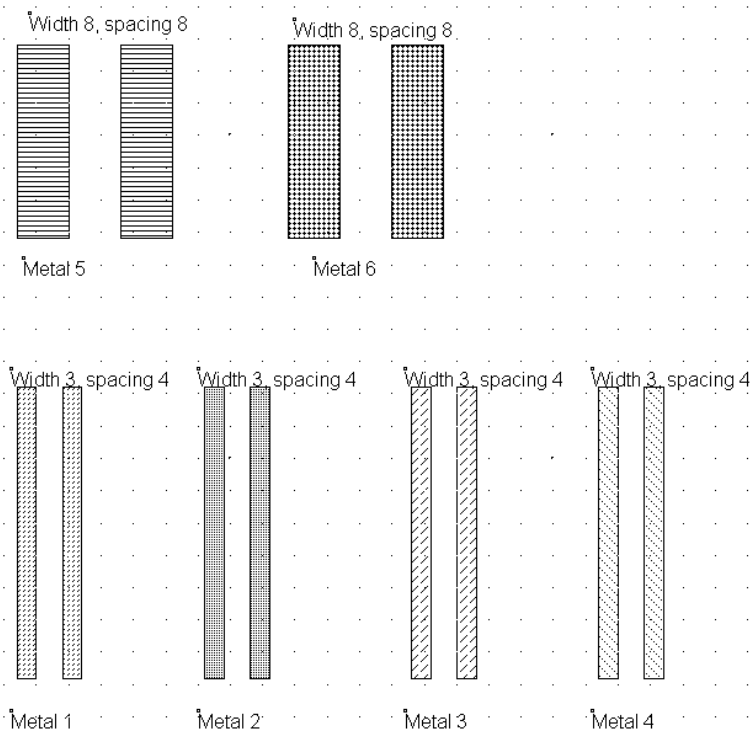


Figure 5-12: Metal layers, associated rules and patterns as appearing in the Microwind editor (DesignRulesMetal.MSK)

These minimum width and spacing are critical dimensions. They define the limit below which the probability of manufacturing error rises to an unacceptable level. If we draw metal 1 lines with 2 lambda width and 2 lambda spacing, interconnect interruptions or short-cuts may appear, as illustrated in figure 5-13. Prior to fabrication, the design rules must be checked to ensure that the whole circuit complies with the width and spacing rules, to avoid unwanted interruptions or bridges in the final integrated circuit. The Microwind command to get access to the design rule checker is **Analysis → Design Rule Checker**. Still, there exists a significant probability of manufacturing error even if the circuit complies to all design rules. A wafer of 500 integrated circuits has a typical yield of 70% in mature technologies, which means that 30% of the circuits have a fabrication error and must be rejected. The yield may drop to a percentage as low as 20%, for example in the case of state-of-the art technologies with all process constraints pushed to their limits, and very large silicon dies.

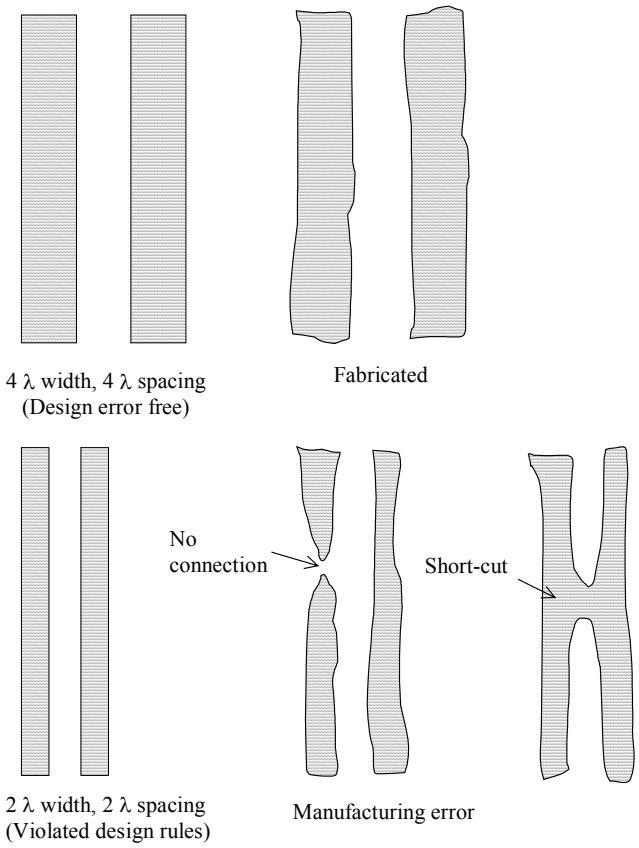


Figure 5-13: The manufacturing of interconnects which violate the minimum width and distance may result in interruptions or short-cuts, which may have catastrophic consequences on the behavior of the integrated circuit.

The practical design width for metal interconnects is usually a little higher than the minimum value. In Microwind, the routing interconnects are drawn in 4 lambda width. The pitch is the usual distance that separates two different interconnects. In 0.7µm, due to severe constraints in the contact size, the pitch has been fixed to 10 lambda. In deep-submicron technology, improvements in contact sizing may reduce that pitch to 8 lambda. In 0.12µm technology, this routing pitch is equivalent to 0.48µm.

Technology	Filename	Upper metal			Lower metal		
		Width	Thickness	Distance	Width	Thickness	Distance
0.8µm	cmos08. RUL	1.2	0.70	1.20	3.20	0.70	6.00
0.6µm	cmos06. RUL	0.9	0.70	0.90	2.40	0.70	4.50
0.35	cmos035.RUL	0.6	0.70	0.60	1.60	0.70	3.00
0.25	cmos025.RUL	0.38	0.60	0.50	1.00	0.70	1.88
0.18	cmos018.RUL	0.30	0.50	0.40	0.80	1.00	0.80
0.12	cmos012.RUL	0.18	0.40	0.24	0.48	0.80	0.48
90nm	cmos90n.RUL	0.15	0.35	0.20	0.40	1.60	0.40
70nm	cmos70n.RUL	0.10	0.30	0.14	0.28	1.00	0.52
50nm	cmos50n.RUL	0.08	0.25	0.10	0.20	0.50	0.38

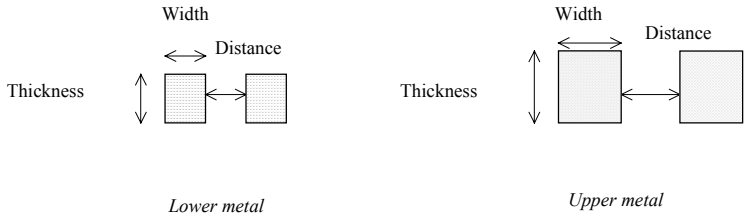


Table 5-2: Conductor parameters vs. technology

Integrated circuit manufacturers usually specify this routing pitch in their non-confidential technological descriptions, as one of the commercial arguments for designing compact (and behind this, low cost) integrated circuits. Common industrial pitch in 0.12µm CMOS process is around 0.4µm. The design styles in 0.7µm and 0.12µm are illustrated in figure 5-14.

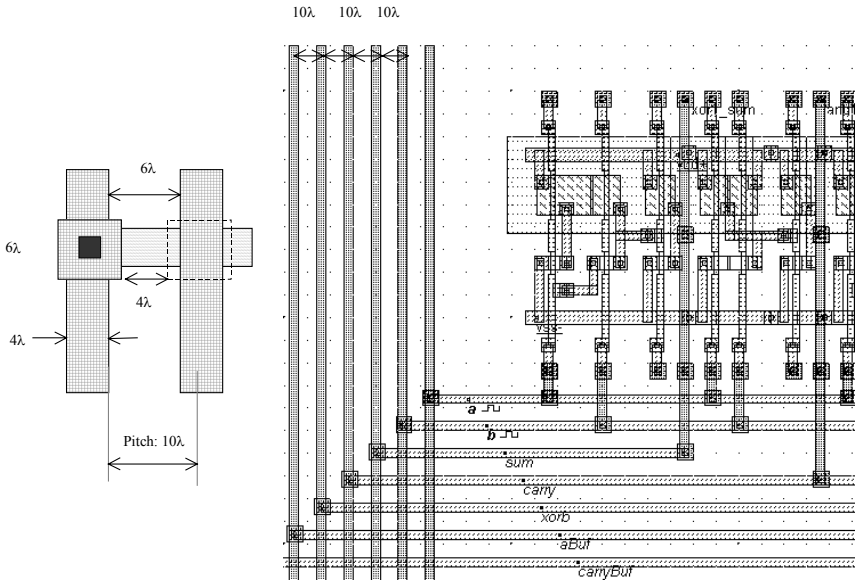


Figure 5-14: Illustration of the routing pitch in 0.7µm, set to 10 lambda due to the large size of the contacts

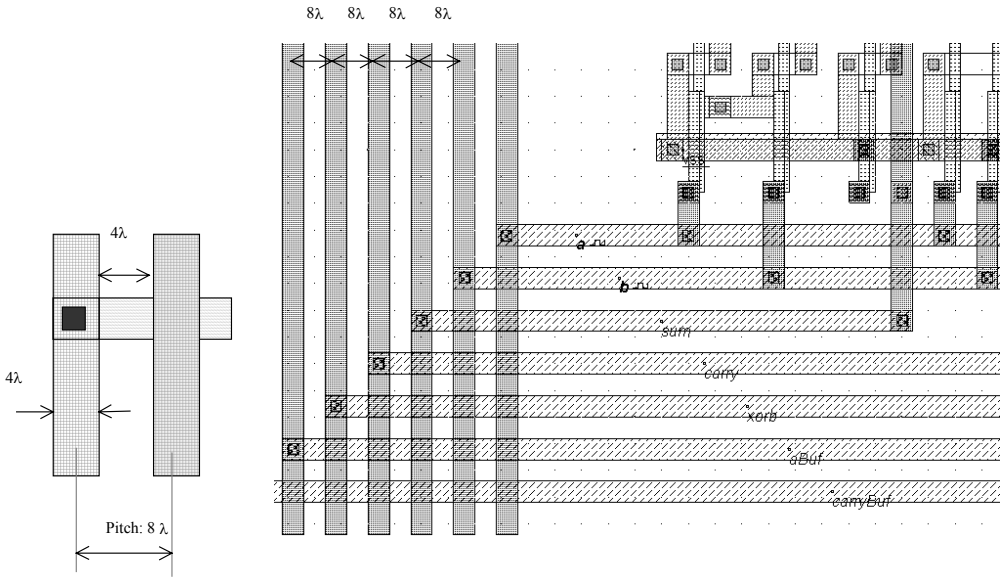


Figure 5-15: Illustration of the routing pitch in 0.12µm, set to 8 lambda thanks to the reduced size of the contact

5. Capacitance associated with interconnects

Interconnect lines exhibit the property of capacitance, as they are able to store charges in the metal interface with oxide. The capacitance effect is not simple to describe and to modelize. This is due to the fact that interconnects are routed very close to each other, as shown in the example in figure 5-16. The capacitance effects are represented by a set of capacitors which link interconnects electrically.

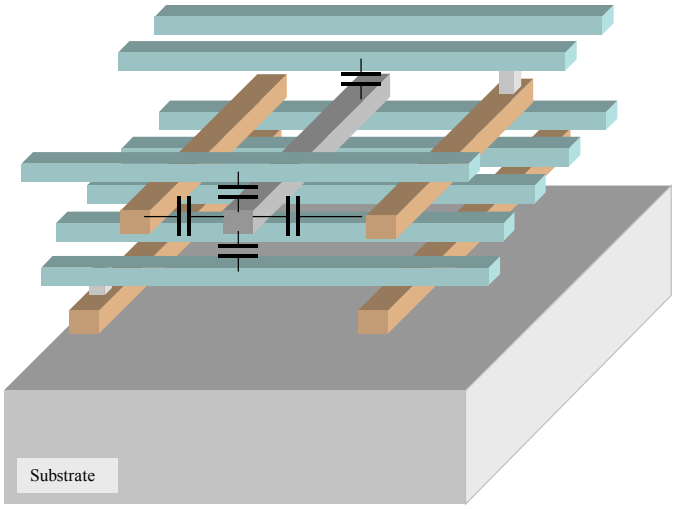


Figure 5-16: One interconnect is coupled to other conductors in several ways, both lateral and vertical

LARGE PLATES

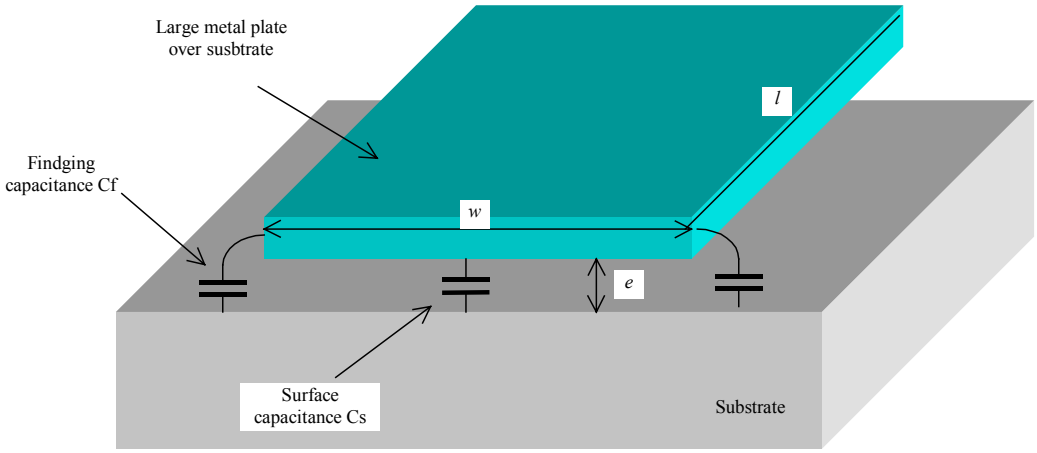


Figure 5-17: Large plate of metal above the substrate

In the case of a large metal area (width w, length l) separated from the substrate or other metal areas by an oxide with a thickness e, the formulation 5-1 is quite accurate. We neglect the fringing capacitance *Cf* in that case. Large plates of metal are used in pads (See chapter 15 for input/output pad description) and supply lines.

$$Cs = \varepsilon_0 \varepsilon_r \frac{w.l}{e} \qquad (5\text{-}1)$$

$\varepsilon_0 = 8.85\ e^{-12}$ Farad/m

$\varepsilon_r = 3.9$ for SiO_2

w= conductor width (m)

l = conductor length (m)

e = dielectric thickness (m)

CONDUCTOR ABOVE A PLANE

Several formulations have been proposed [Sakurai][Delorme] to compute the capacitance of a conductor when the width is comparable to the oxide thickness, which is the case with the large majority of conductors used to transport signals. We give the formulation of the total capacitance which consists in the sum of Cs and twice the fringing capacitance Cf, from [Delorme].

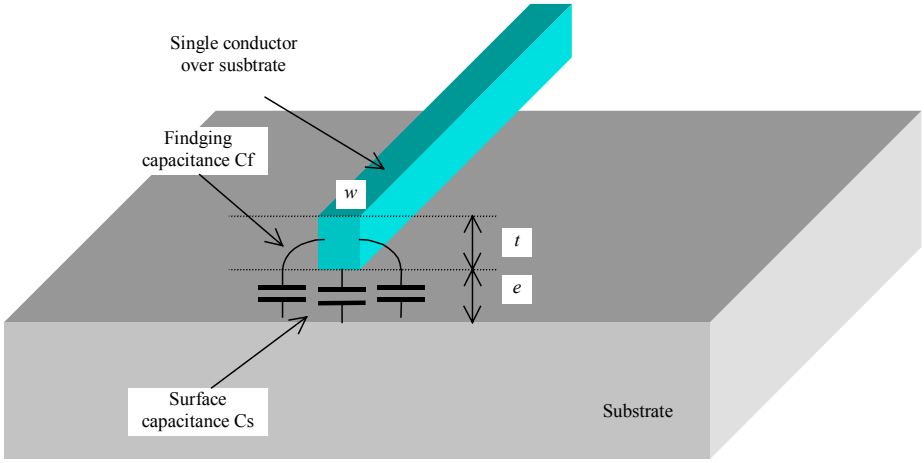


Figure 5-18: One conductor above a ground plane

$$C = C_s + 2.C_f = \varepsilon_0 \varepsilon_r (1{,}13.\frac{w}{e} + 1{,}44.(\frac{w}{e})^{0{,}11} + 1{,}46.(\frac{t}{e})^{0{,}42}) \qquad (5\text{-}2)$$

C = total capacitance per meter (Farad/m)

Cs= surface capacitance (Farad/meter)

Cf = fringing capacitance (Farad/m)

$\varepsilon_0 = 8.85\ e^{-12}$ Farad/m

$\varepsilon_r = 3.9$ for SiO_2

w= conductor width (m)

t= conductor thickness (m)

e = dielectric thickness (m)

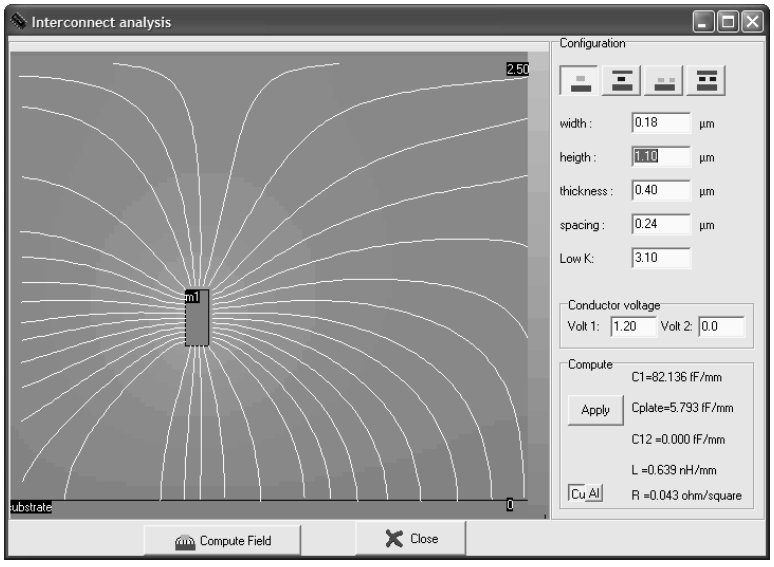


Figure 5-19: 2D simulation of the field lines between the conductor and the ground

The command **Analysis → Interconnect Analysis** with FEM gives some interesting information about the electric coupling between the conductor and the ground (Figure 5-19). The palette of colors indicate the potential in the oxide region, in volt. The field lines are the white wires which link the conductor to the ground. Some field lines are directly coupled to the substrate ground, some other field lines go to the free space. In 0.12µm technology, the metal conductor is 1.2µm above the ground plane, which corresponds to the simulation reported in figure 5-19. Its minimum width is 0.18µm (Equivalent to 3 lambda), its thickness is 0.4µm. These physical dimensions are listed in the parameter menu situated on the right side of the window. The formulations given in (5-1) produce an underestimated value for *Cplate*, equal to 6fF/mm. The formulations proposed in (5-2) give a more reliable result for *C1*, equal to 82fF/mm.

Two Conductors above a ground plane

When a conductor is routed close to another conductor, a crosstalk capacitance, defined as C12 is created between the two conductors (Figure 5-20). In 0.12µm technology, a specific dielectric with a low permittivity (This parameter, called LowK, is approximately 3 instead of 4) is used to fill the gaps between interconnects. This is an efficient technique to reduce the crosstalk capacitance while keeping the upper and lower capacitance almost unchanged. Consequently, the oxide stack alternates between high K and low K materials, as shown in the cross-section. Low K dielectrics were introduced with 0.18µm technology. Air gaps are the ultimate low K materials, with a lowest possible K =1. Intensive research is being conducted on the enclosure of air gaps in between coupled connectors, and could become a standard in future technologies.

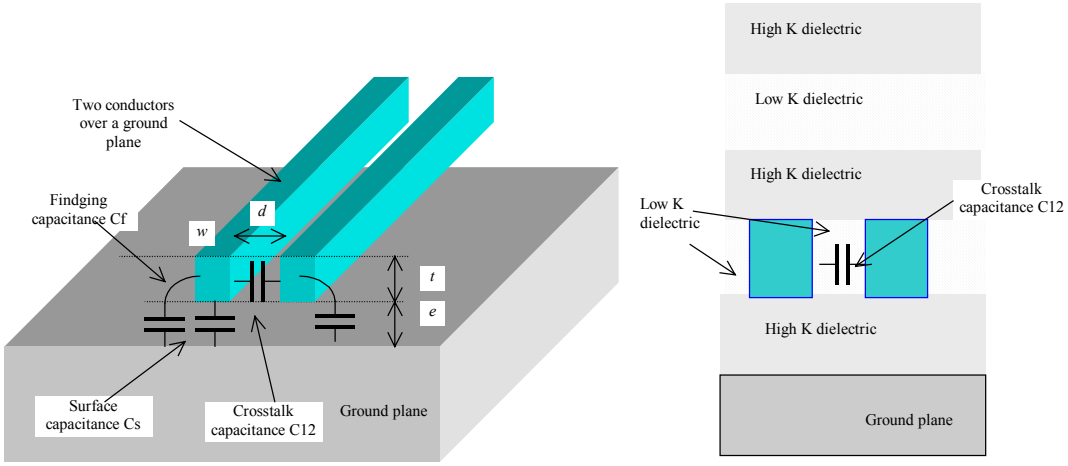


Figure 5-20: Two conductors above a ground plane

$$C = C_s + C_f = \varepsilon_0 \varepsilon_r (1.10\frac{w}{e} + 0.79(\frac{w}{e})^{0.1} + 0.46.(\frac{t}{e})^{0.17}(1 - 0.87e^{(\frac{-d}{e})})) \qquad (5\text{-}3)$$

$$C_{12} = \varepsilon_0 \varepsilon_{rlowK} (\frac{t}{d} + 1.2(\frac{d}{e})^{0.1}.(\frac{d}{e} + 1.15)^{-2.22} + 0.253ln(1 + 7.17\frac{w}{d}).(\frac{d}{e} + 0.54)^{-0.64}) \qquad (5\text{-}4)$$

C = conductor capacitance to ground per meter (Farad/m)

Cs= surface capacitance (Farad/meter)

Cf = fringing capacitance (Farad/m)

C12 = crosstalk capacitance (Farad/m)

$\varepsilon_0 = 8.85\ e^{-12}$ Farad/m

$\varepsilon_r = 3.9$ for SiO_2

ε_{rlowK} = permittivity of low dielectric material (around 3.0 in 0.12µm)

w= conductor width (m)

t= conductor thickness (m)

e = dielectric thickness (m)

d = conductor distance (m)

In the default 0.12µm technology, two coupled interconnects with the minimum width and distance routed with the first level of metallization have a cross-section shown in figure 5-21. The benefits of the low permittivity dielectric appear clearly in the value of the coupling capacitance C12 (75fF/mm), which is comparable to the ground capacitance C1 (51fF/mm). If we change the parameter *Low K* to the silicon dioxide permittivity equal to 4, the coupling capacitance becomes much higher than the ground capacitance. More about the crosstalk effect, including simulations and measurements, is given at the end of this chapter.

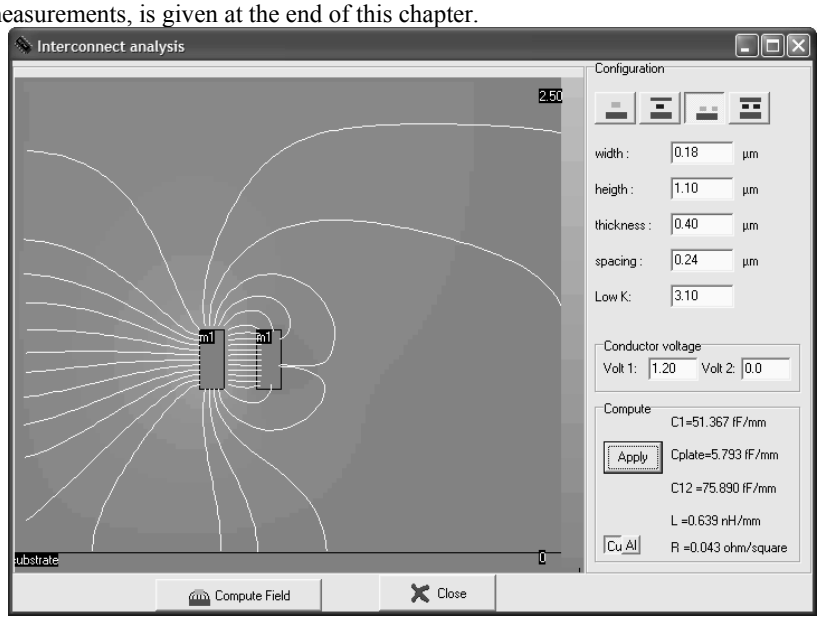


Figure 5-20: 2D simulation of the field lines between two conductors and the ground

Real case conductors

In practice, the accurate extraction of the capacitance of each node is based on a 2 dimensional partitioning of the layout, and the 3D computing of capacitance for each elementary configuration. Even a simple interconnect configuration, such as the one shown in figure 5-21, is equivalent to a large set of 3D configurations, each of them requiring the use of a 3D static field solver.

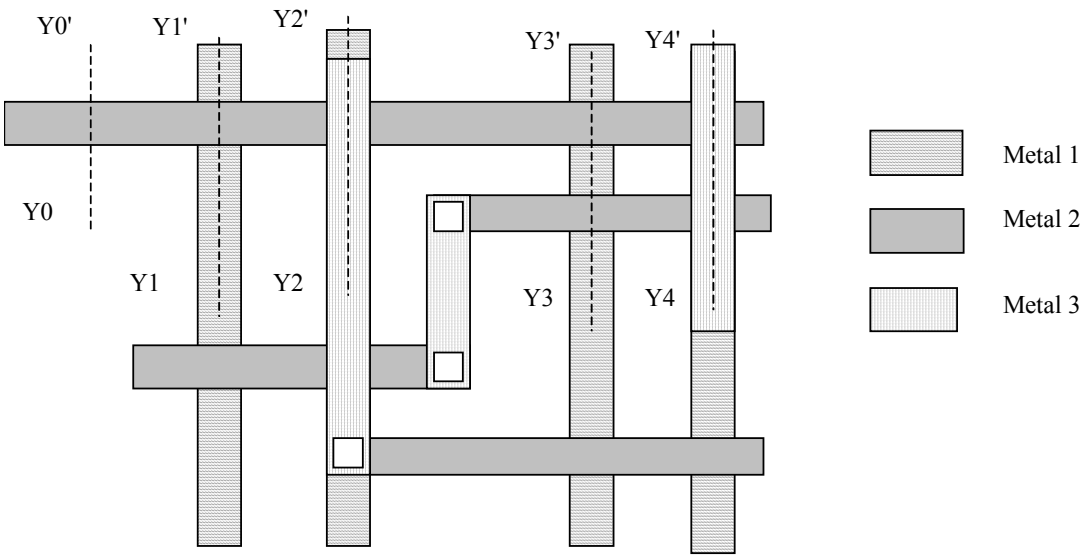


Figure 5-21: Example of interconnects routed in metal 1, 2 and 3

Examples of standard configurations that may be found in the layout of figure 5-21 are shown in figure 5-22, with numerical values corresponding to 0.12µm technology, metal wire width 4 lambda (0.24µm), and wire spacing 4 lambda (0.24µm).

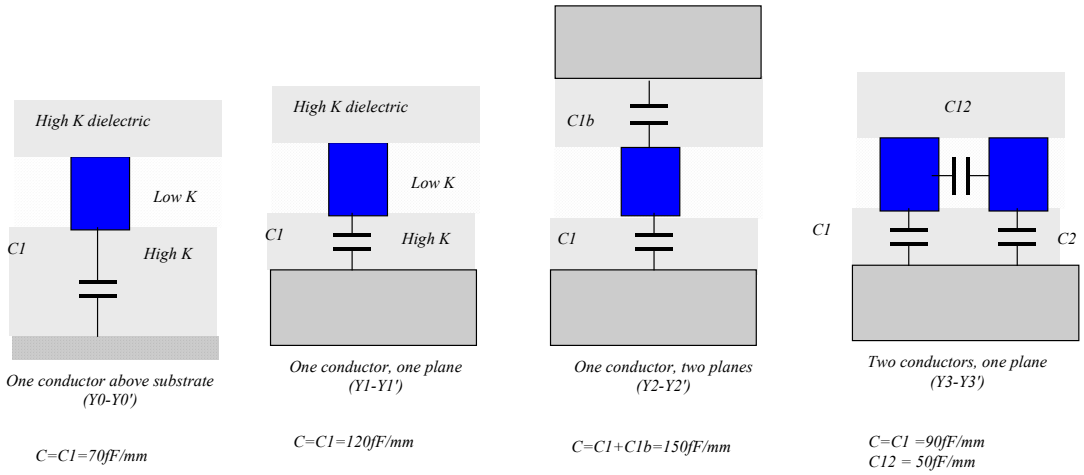


Figure 5-22: Basic configurations illustrating the cross-sections of the interconnect network (0.12µm)

The extraction of the layout is conducted according to the flow described in figure 5-23. The MOS devices and interconnects are extracted separately. To compute the crosstalk capacitance, the interconnect network is parsed into elementary structures, each one having a fixed stack of layers. An iterative procedure permits to locate vertical and horizontal elementary crosstalk contributions and to compute the sum which appears in the SPICE netlist and the electrical node properties.

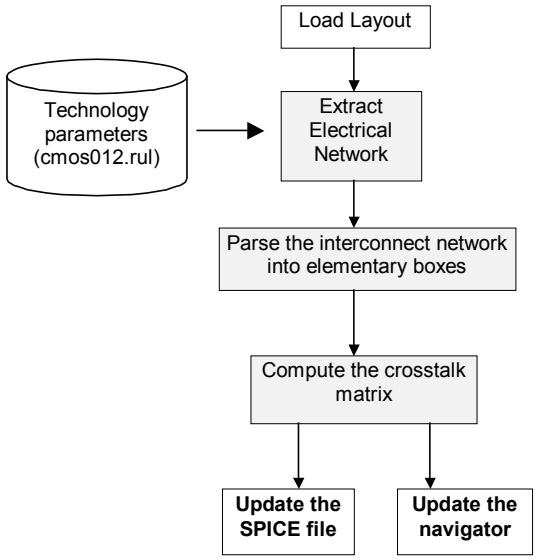


Figure 5-23: The crosstalk capacitance extraction steps in Microwind

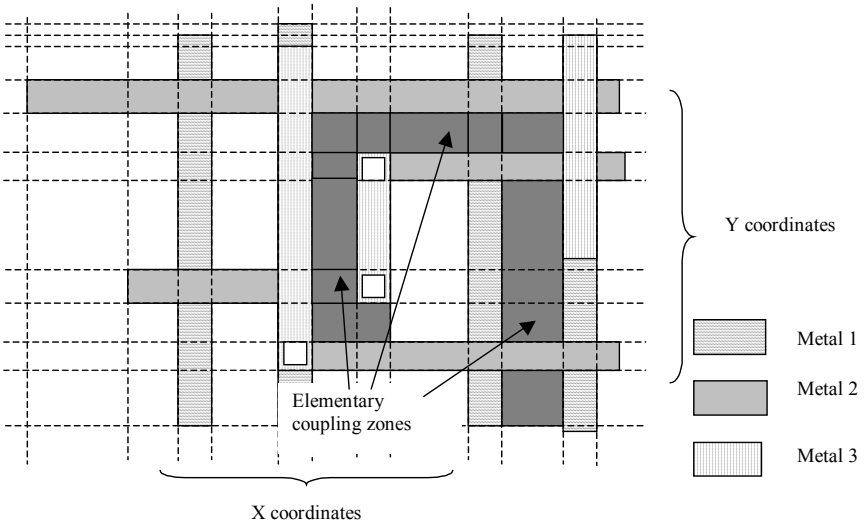


Figure 5-24: Identification of coupling zones in the test case presented in figure 5-21

Figure 5-24 shows where the elementary coupling zones have been detected. Only lateral coupling between identical layers, close enough to generate a significant crosstalk contribution are considered. The other couplings are neglected in this extraction phase. In the test case proposed in figure 5-21, couplings occur in metal1, metal2 and metal3, as seen in figure 5-24. Each contribution is very small (less than 1femto-Farad). However, the sum of elementary coupling contribution may result in hundreds of femto farad, which may be comparable to the ground capacitance, and thus create significant crosstalk coupling effects.

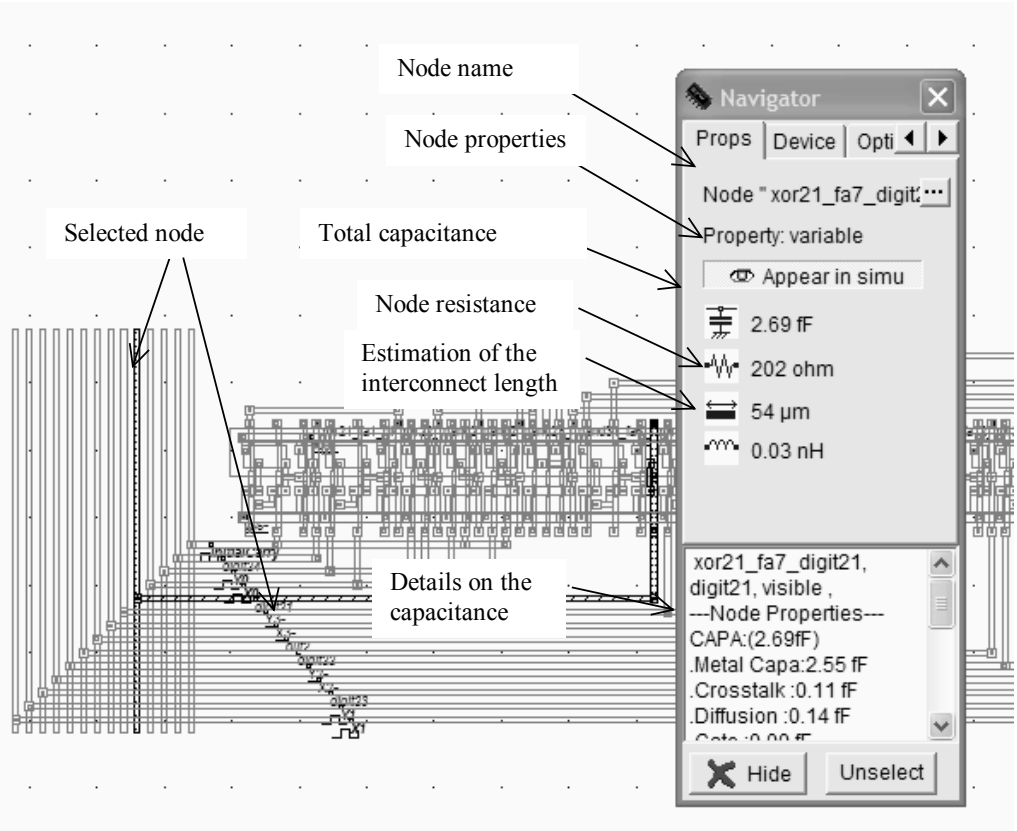


Figure 5-25: Extraction of a real-case interconnect capacitance using the command "View Electrical Node"

The command **View →View Electrical Node** or the above icon above launches the extraction of ground and crosstalk capacitance for each electrical net of the layout. The ground capacitance of the node and the crosstalk capacitance are detailed in the navigator window, as illustrated in figure 5-25. Notice that the global capacitance is split into metal, crosstalk, gate and diffusion capacitance. The selected net has weak crosstalk coupling as compared to ground coupling.

6. Resistance associated with interconnects

The resistivity of interconnect materials used in CMOS integrated circuits is listed in table 5-3. Conductors have very low resistivity, while semiconductor materials such as highly doped silicon have a moderate resistivity. In contrast, the intrinsic silicon resistivity is very high.

Symbol	Description	Used for	Resistivity at 25°C
ρ_{cu}	Copper resistivity	Signal transport	$1.72\ 10^{-6}\Omega.cm$
ρ_{al}	Aluminum resistivity	Signal transport	$2.77\ 10^{-6}\Omega.cm$

ρ_{Ag}	Gold resistivity	Bonding between chip and package	$2.20\ 10^{-6}\ \Omega.cm$
$\rho_{tungsten}$	Tungsten resistivity	Contacts	$5.30\ 10^{-6}\ \Omega.cm$
ρ_{Ndiff}	Highly doped silicon resistivity	N+ diffusions	$0.25\ \Omega.cm$
ρ_{Nwell}	Lightly doped silicon resistivity	N well	$50\ \Omega.cm$
ρ_{si}	Intrinsic silicon resistivity	Substrate	$2.5\ 10^{5}\ \Omega.cm$

Table 5-3: Resistivity of several materials used in CMOS circuits

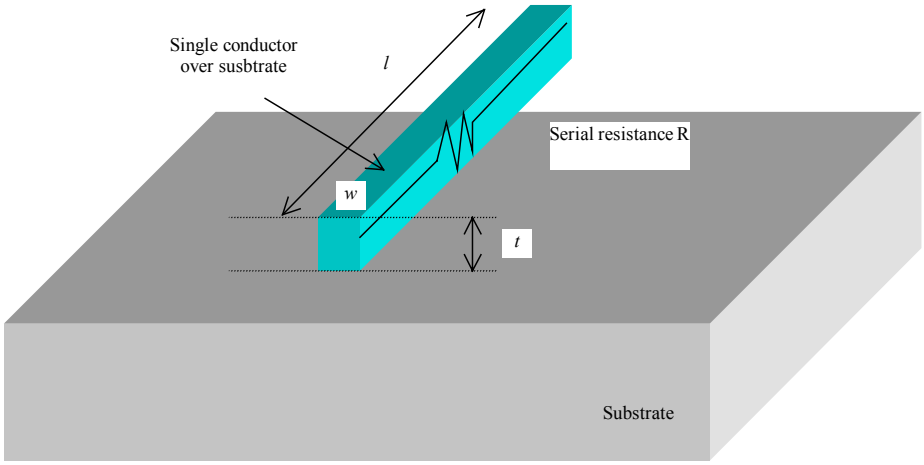


Figure 5-26: Resistance of a conductor

If a conductor with a resistivity has length *l*, width *w*, and thickness *t*, then its serial resistance R (Figure 5-26) can be computed using the following formula:

$$R = \rho \frac{l}{w.t} \qquad \text{(Equ. 5-5)}$$

where

R=serial resistance (ohm)

ρ=resistivity (ohm.m)

w= conductor width (m)

t= conductor thickness (m)

l = conductor length (m)

d = conductor distance (m)

Resistance per square

When designing interconnects, a very useful metric is the "resistance per square". We assume that the width is equal to the length, that is:

$$R_{square} = \rho \frac{w}{w.t} = \frac{\rho}{t} \qquad \text{(Equ. 5-6)}$$

The interconnect material has long been aluminum as it was an easy-to-process material. Unfortunately, the resistivity of this material is quite high. Recently, copper has replaced aluminum for the manufacturing of interconnects, with a significant gain in terms of resistance, as the intrinsic resistivity of copper is almost twice lower than that of aluminum (See table 5-3). Tungsten is used in several CMOS technologies to fabricate contact plugs. Its ability to fill narrow and deep holes compensates its high resistance. Gold is only used to connect the final chip to its packaging, as described in chapter 15.

For a copper interconnect ($1.72\ 10^{-6}\Omega$.cm) and a 0.4µm thickness, the resistance is around 0.043ohm/square. The measured square resistance is higher than this theoretical value as the conductor is not homogenous. The titanium barriers located on both sides of the conductor have an important resistance which reduces the effective section of the conductor. In the CMOS 0.12µm process files, a value of 50mOhm is used.

The square resistance is used to estimate rapidly the equivalent resistance of an interconnect by splitting its layout into elementary squares. The sum of square is then multiplied by R_{square} in order to evaluate the global interconnect resistance. We illustrate this concept in the layout shown in figure 5-27. We assume a resistance per square R_{square} of 50 mΩ. The resistance from A to B can be approximated by 10 squares, that is a resistance of 0.5Ω.

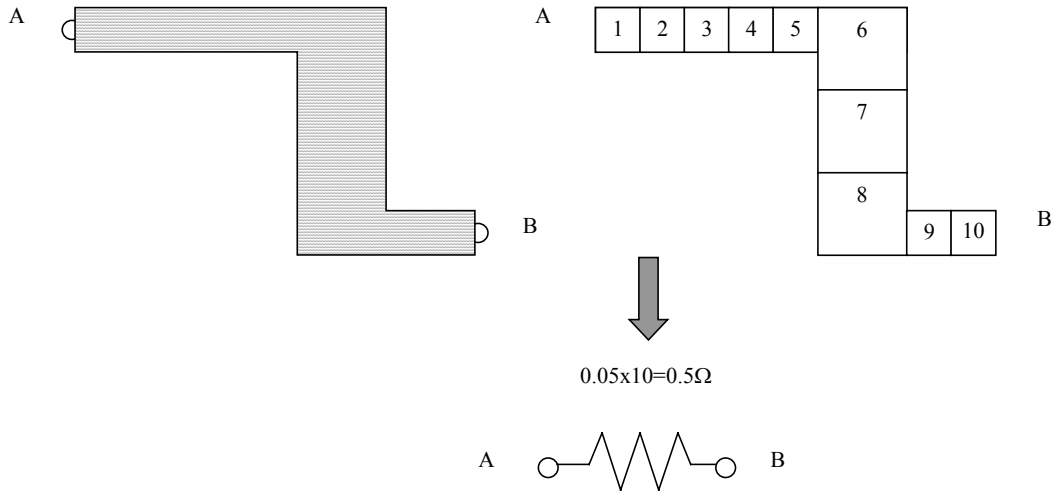


Figure 5-27: applying the concept of resistance per square to a portion of interconnect

The conductor resistivity is usually considered as a constant value. However, it depends on temperature in a complex manner [Hastings]. A usual approximation consists in considering the linear temperature coefficient of resistivity (TCR) expressed in parts per million per degree (ppm/°C). The copper and aluminum materials have similar behaviors, with a TCR around 4000 ppm/°C. The resistance at a temperature T is given by the following equation:

$$R_T = R_{T0}[1 + 10^{-6} TCR(T - T0)] \qquad \text{(equ 5-7)}$$

where

RT=serial resistance at temperature T(ohm)

RT0=serial resistance at reference temperature T0 (ohm)

TCR= temperature coefficient of resistivity (ppm/°C)

T=temperature (°C)

T0=reference temperature (Usually 25°C)

Considering a typical 4 lambda wide metal interconnect, we observe that its cross-section is dramatically reduced with the scale down, due to a decrease of both the lateral and vertical dimensions of the elementary conductors. The elementary resistance Rsquare is increased at each technology generation, as shown in figure 5-29. The introduction of copper in high performance CMOS process has lowered almost by 50% the square resistance, but significantly increased the process complexity and overall fabrication cost.

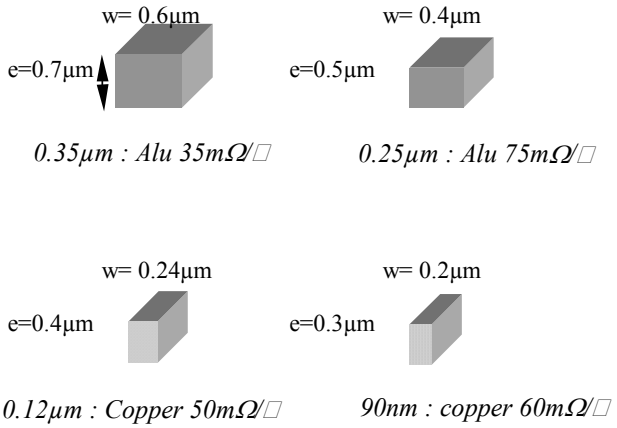


Figure 5-28: Evolution of interconnect resistance with the technology scale down [ITRS]

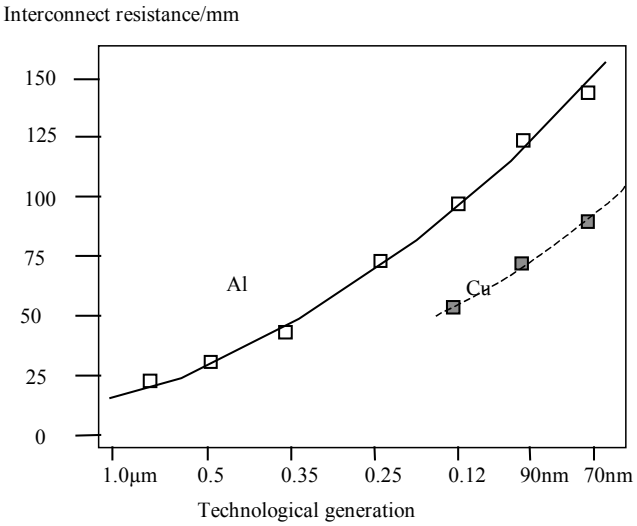


Figure 5-29: Evolution of interconnect resistance with the technology scale down

Via Resistance

Each contact and via has a significant resistance. Typical values for these resistance are given in table 5-4 for three technologies: 0.7µm, 0.12µm and 90nm.

Technology	0.7µm	0.12µm	90nm
Contact resistance	0.5 Ω	15 Ω	20 Ω
Via	0.3 Ω	4 Ω	8 Ω
Upper via	-	1 Ω	3 Ω

Table 5-4: typical resistance of contacts and vias

Globally, the contact resistance and via increase with the technology scale down, due to the continuous reduction of the contact plug section, resulting in a reduced path for current. The contact resistance from active regions to the first metal layer is very important due to the thick oxide (Around 1.0µm) that separates the active MOS device altitude from the metal 1 altitude. A thick oxide is necessary to insert several optional materials such as double gate MOS devices for EEPROM memories, or large storage capacitance for DRAM memories. The upper via resistance is quite small as the size of that via is very large compared to the lower via.

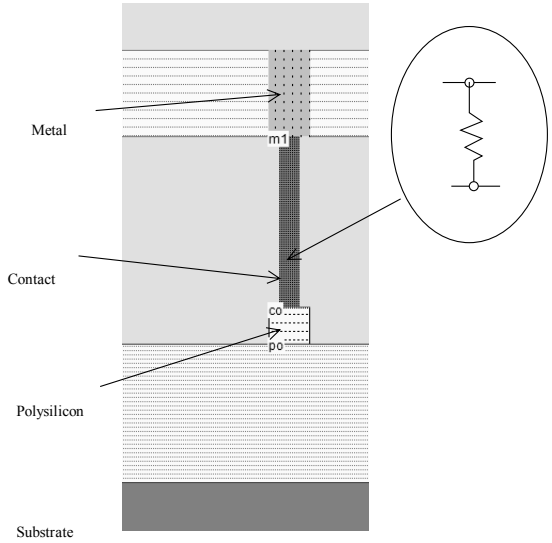


Figure 5-30: The contact is equivalent to a resistance (Contacts.MSK)

7. Signal Transport

The signal transport from one logic cell to another logic cell uses metal interconnects. Depending on the distance between the emitting cell and the receiving cell, the interconnect may be considered as a simple parasitic capacitance or a combination of capacitance and resistance. By default, Microwind only considers the parasitic capacitance. This assumption is valid for short to medium length interconnects . In 0.12µm, interconnects with a length up to 1000µm can be considered as a pure capacitance load *Cl* (Figure 5-31). For interconnects larger than 1000µm, the serial resistance *Rl* should be included into the model. The usual way consists in splitting the capacitance *Cl* into two equivalent capacitances and place the serial resistance in between.

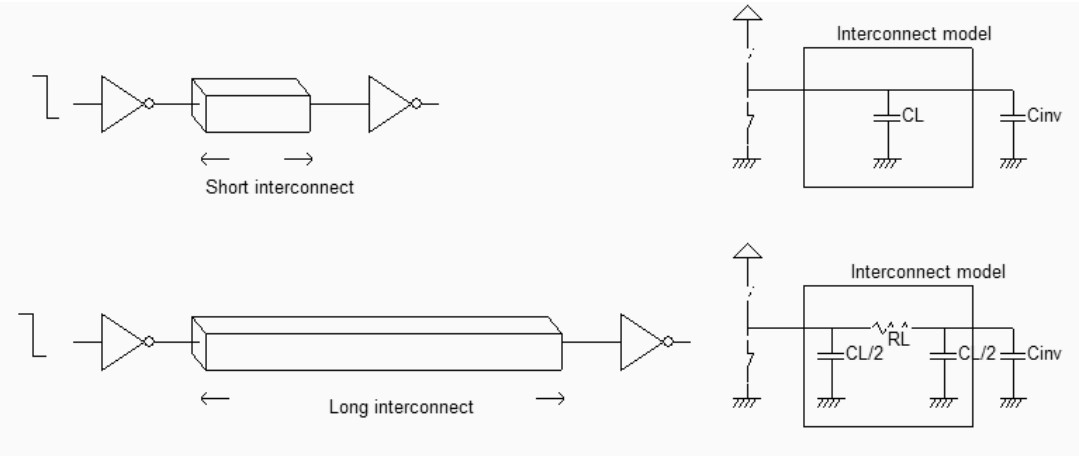


Figure 5-31: C versus RC model for interconnects (RcModels.SCH)

Simulation of the RC effect

The RC delay within interconnects can be simulated using Microwind2 as follows. Consider the circuit shown in figure 5-32. It represents a buffer (Left lower corner) which drives a long interconnect, and then a loading inverter. From a layout point of view, the interconnect is usually straight, but for simplicity's sake, we use here a serpentine to emulate the RC effect without the need of a very large silicon area. The 4mm interconnect is obtained by connecting small portions of metal lines. The layout reported in figure 5-33 is based on 40 bars of metal4, with a length of 100µm. The serial resistance *Rl* is equal to 820 ohm, and the capacitance *Cl* is around 130fF, divided into two parts, shared on each side of the virtual resistance.

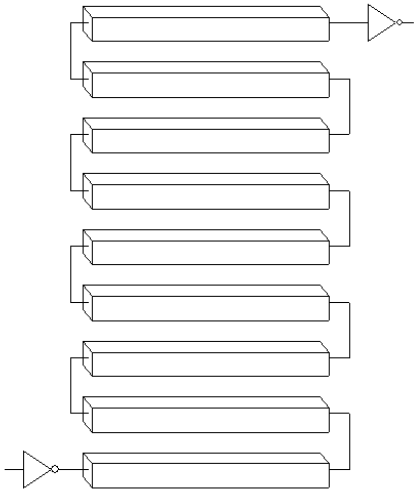


Figure 5-32: Emulation of the RC effect in a 4mm interconnect using serpentine (RcEffect.SCH)

Although both the capacitance and resistance of each node are extracted, the simulator considers by default only the capacitance and ignores the resistance. We add a "virtual" resistance using the resistance icon situated in the palette (Figure 5-33). This resistance is placed directly in the interconnect, as detailed in the layout of figure 5-34, to force

Microwind to handle the equivalent resistance of the interconnect. We assign to the resistor the value of the metal interconnect resistance, appearing as *R(metal)*. This value is updated during the electrical network extraction phase.

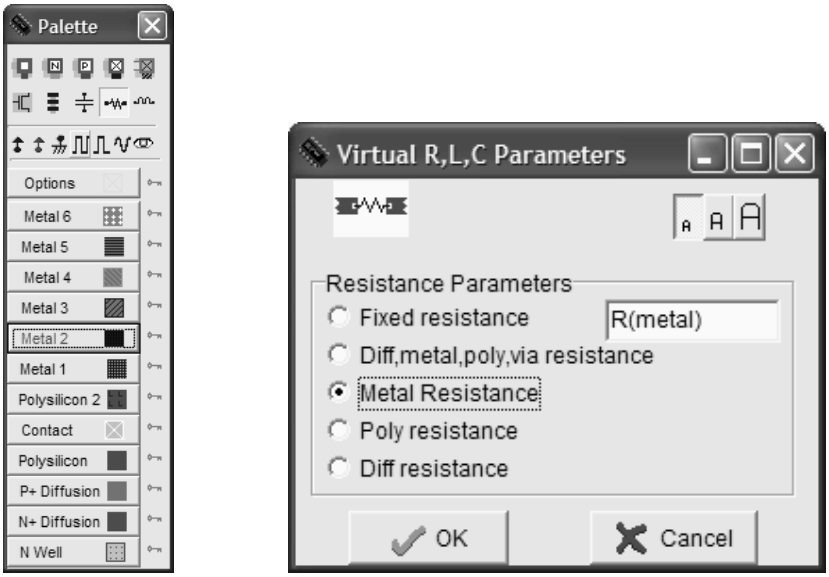


Figure 5-33: Inserting a virtual resistance to take into account the serial resistance of the interconnect (RcModels.SCH)

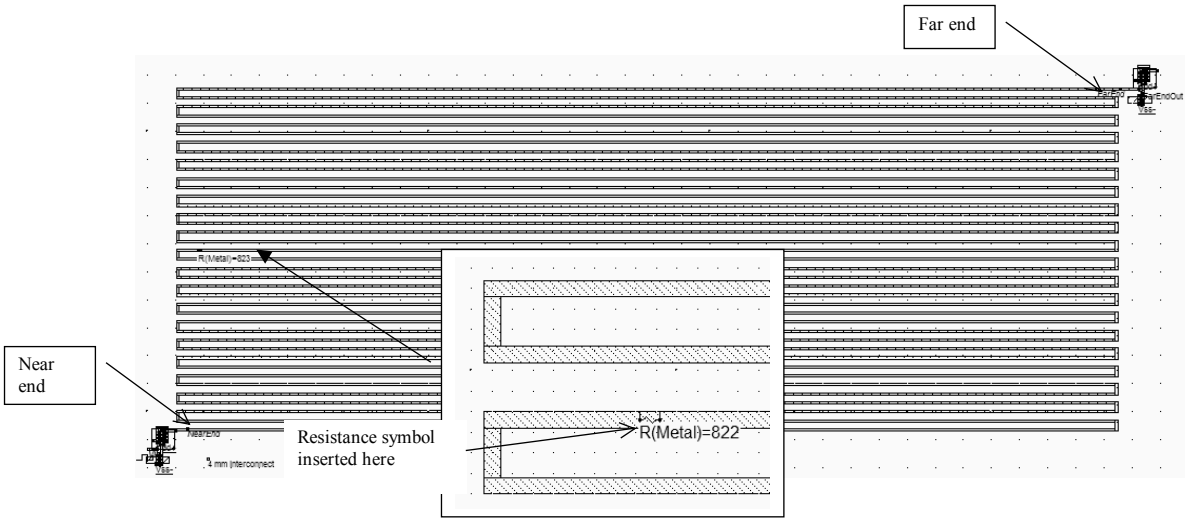


Figure 5-34 : adding a virtual resistance within the layout to simulate the resistance (RcEffect.MSK)

The RC effect of the interconnect appears very clearly in the analog simulation (figure 5-35). The initial phase runs from time 0.0 to time 1.0ns, which is not significant. At time t=1.0ns, the input shifts from a high to a low state. The near end of the line switches within approximately 200ps. The far end of the line reaches VDD/2 after twice this delay. The same delay effects are observed at the fall edge of the clock (t=1.5ns).

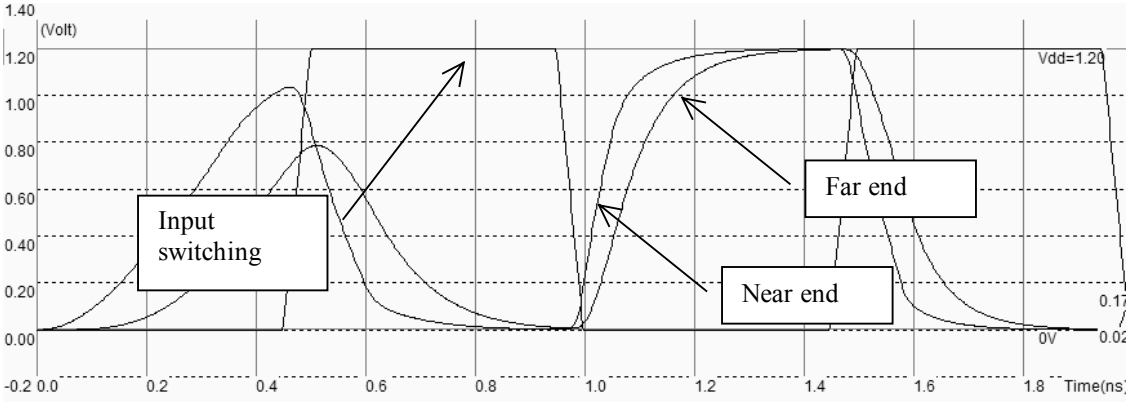
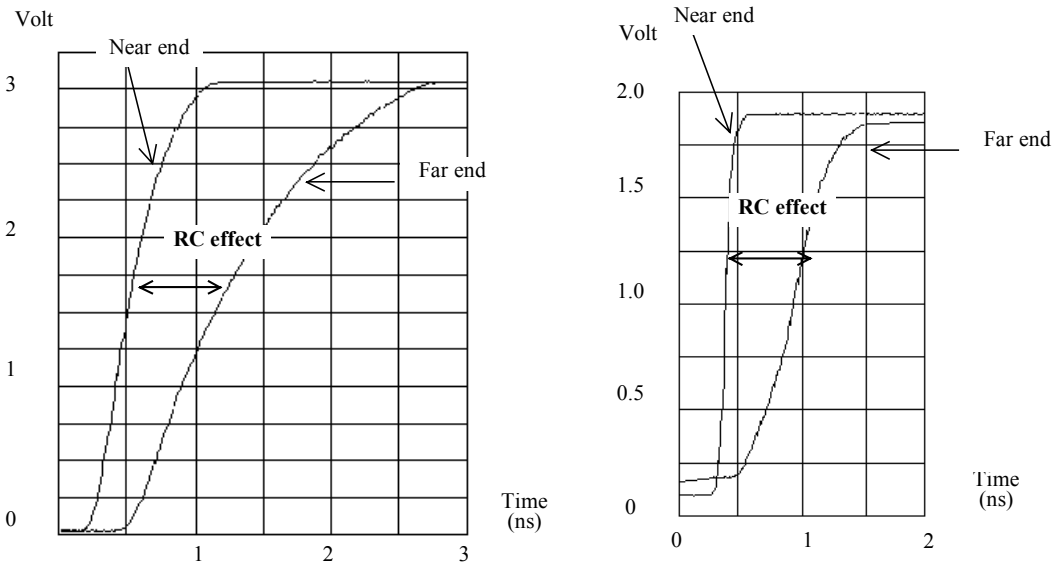


Figure 5-35 : RC delay simulation (RCEffect.MSK)



(a) 10mm, metal 3, 0.35µm technology *(b) 10mm, metal3, 0.18µm technology*

Figure 5-36: Measurement of RC effect in a very long interconnect (10mm) in metal 3, experimented in 0.35µm and 0.18µm technologies [Bendhia]

The RC effect can be measured thanks to an on-chip oscilloscope approach [Bendhia] at the near end and far end of metal interconnects. In the left part of figure 5-36, the measurement concerns a 10mm interconnect routed in metal3, fabricated in a 0.35µm five metal layer process from ST-Microelectronics [ST]. The observed waveforms confirm the important impact of the serial resistance, estimated in this particular case at around 300 ohm. The waveform is similar in 0.18µm technology, 6 metal layers, but the near end of the line is prompt to switch within 100ps, while, the far end of the line needs more than 500ps to pass the VDD/2 limit. Notice the supply voltage difference between these two technologies.

Limit between C and RC models

In 0.12µm technology, the interconnect length limit above which the resistance should not be neglected is around 1mm. This limit can be illustrated by implementing one configuration with the capacitance model, and another configuration with the RC model. To implement a long interconnect, Microwind offers a bus generation command, with an interface reported in figure 5-37. This menu is accessible by the command **Edit → Generate → Metal bus**. Rather than drawing a straight line with a 300, 800 or 2000µm length, we split the metal wire into several portions. This makes the layout more compact, and is equivalent to the straight line. In the example given in figure 5-37, we split the 2mm line into 10 portions of interconnects with a length of 200µm each.

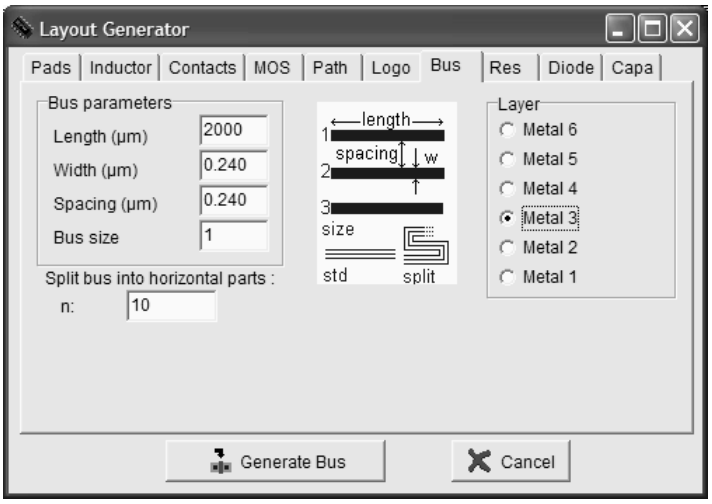


Figure 5-37: Generating a long metal interconnect using the bus generator command

In this study, three interconnect lengths are investigated: 300µm, 800µm and 2mm. Each configuration is implemented with and without the resistance to compare the simulation with the C model alone or the RC model.

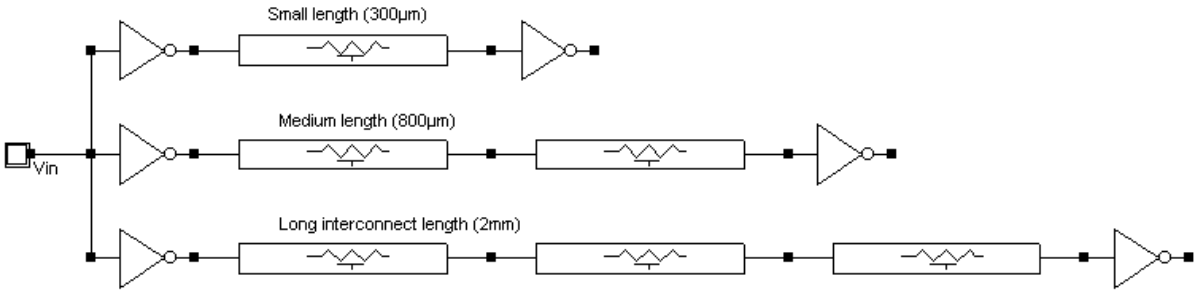


Figure 5-38: Simulation of 3 interconnect configurations to investigate the impact of C/RC models (RcModels.SCH)

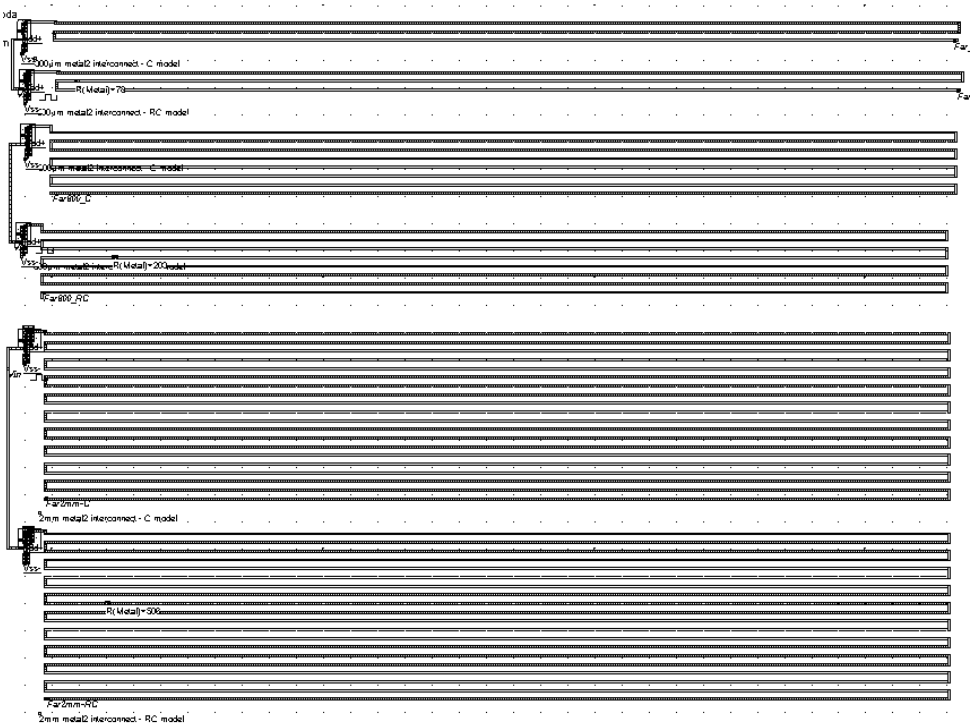


Figure 5-39: Layout of the 3 interconnect configurations (300µm, 800µm and 2mm) to investigate the impact of C/RC models (RCModel.MSK)

Each wire is implemented twice: one version is used to simulate the capacitor model, the other one includes a small resistance fixed approximately half way between the near and far end of the interconnect, to force the simulator to take into account the serial resistance of the interconnect.

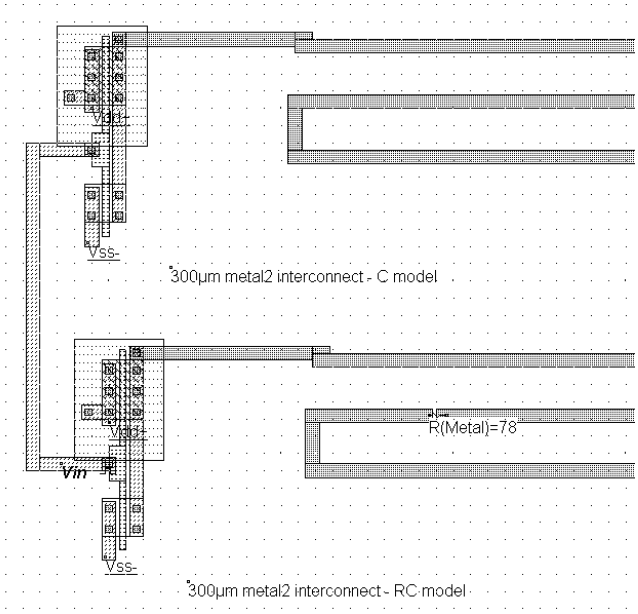


Figure 5-40: Portion of the layout showing the 2 versions of the 300µm configuration, with and without resistance (RCModel.MSK)

In order to handle the resistance effect of the interconnect, we place a virtual resistance approximately in the middle of the metal path. In the menu, we choose **Metal resistance**, which indicates that the extracted metal resistance should be used as the interconnect resistance. In the layout, the text appearing in the R symbol will be *R(metal)*. The corresponding value appears after extraction or simulation. When changing the technology, the value of the resistance will be updated according to the new sheet resistance parameters.

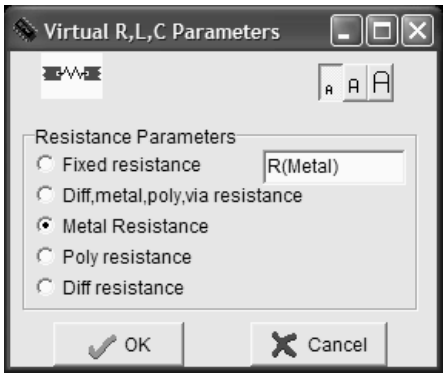
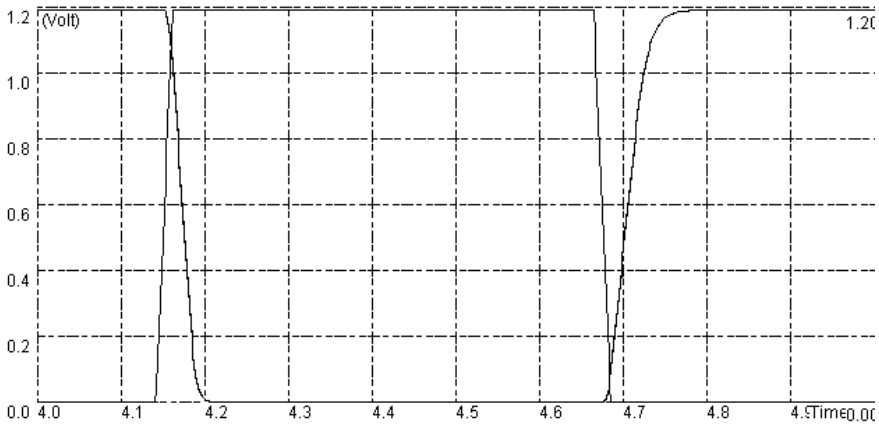


Figure 5-41: Configuring the virtual resistance to handle the metal serial resistance effect in simulation (RCModel.MSK)

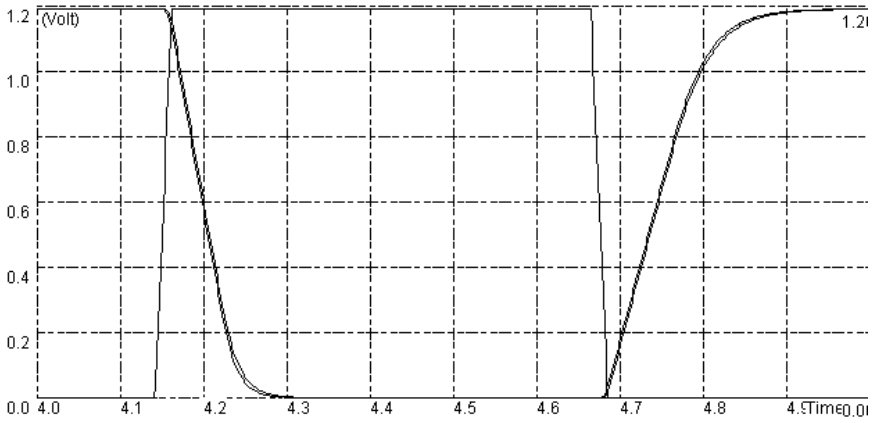
(a) 300µm interconnect

C and RC models give identical results



(b) 800µm interconnect

C and RC models give almost identical results



(c) 2mm interconnect

C and RC models give significantly different results (C is 20% faster than RC)

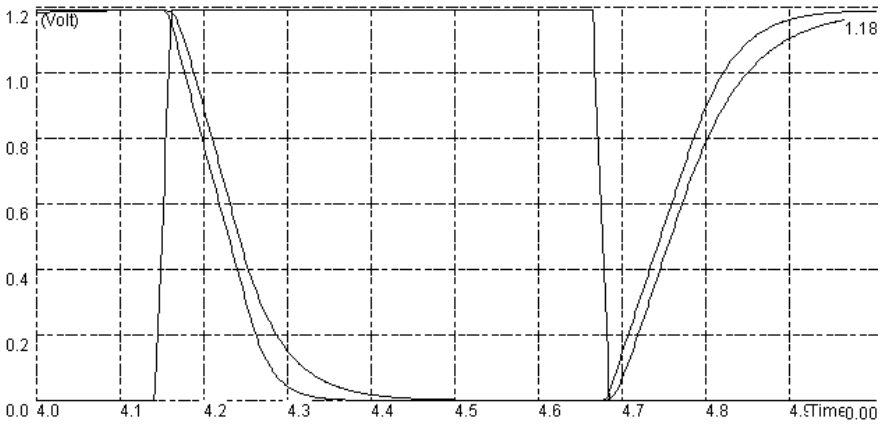


Figure 5-42: Comparative simulation of the C and RC model in signal propagation (RCModel.MSK)

The analog simulation of the signal transport with the 300µm, 800µm and 2mm interconnects are given in figure 5-42. The simulated propagation with C and RC models gives no visible difference below 1mm. Above 1mm, the C model gives optimistic prediction of the delay compared to the RC model. When we plot the delay vs. interconnect length (Figure 5-43), we may see that the RC model is preferable for interconnects with a length greater that 1mm, while the C model is sufficient below 1mm.

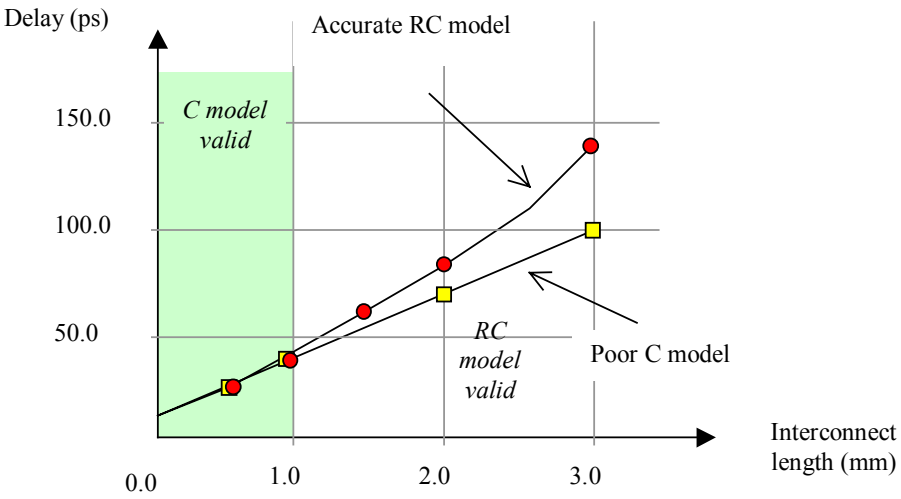


Figure 5-43: Below 1mm, the C model is valid. Above 1mm, the RC model should be considered in 0.12µm CMOS technology

8. Improved signal transport

The basic layout techniques to reduce the signal transport within interconnects are detailed in this paragraph. Two approaches are considered: improving the drive of the switching inverters, and inserting repeaters. We shall also discuss the crosstalk effect and its possible reduction through technology and design improvements.

Increased Current Drive

The simplest approach to reduce the gate delay consists in connecting MOS devices in parallel. The equivalent width of the resulting MOS device is the sum of each elementary gate width. Both nMOS and pMOS devices are designed using parallel elementary devices. Most cell libraries include so-called *x1*, *x2*, *x4*, *x8* inverters. The *x1* inverter has the minimum size, and is targeted for low speed, low power operations. The *x2* inverter uses two devices *x1* inverters, in parallel. The resulting circuit is an inverter with twice the current capabilities. The output capacitance may be charge and discharged twice as fast as with the basic inverter (Figure 5-44), because the *Ron* resistance of the MOS device is divided by two. The price to pay is a higher power consumption. The equivalent *Ron* resistance of the *x4* inverter is divided by four.

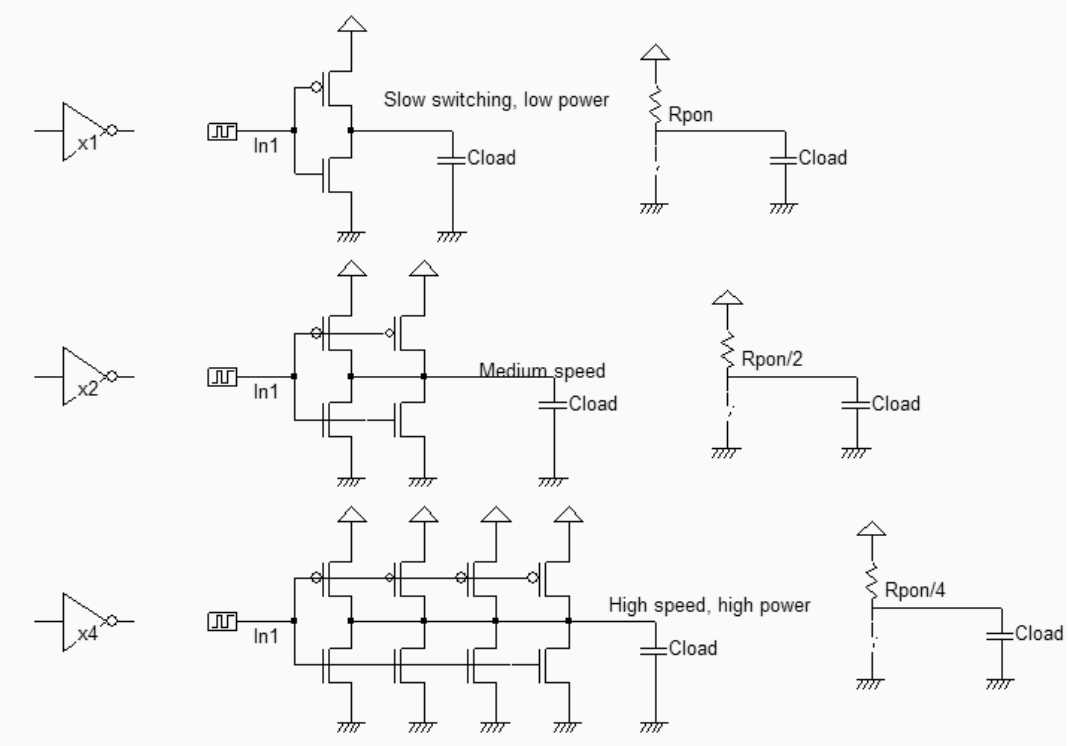


Figure 5-44: The x1, x2 and x4 inverters (Invx124.SCH)

We may use the parametric analyzer included in Microwind to investigate the delay increase with the capacitance load on the output node, for the x1, x2 and x4 inverters. In a first approximation, the capacitance increase is similar to the interconnect length increase. Remember that below 1mm, the interconnect is basically a parasitic capacitance, with a value of 80fF/mm approximately in 0.12µm. What the tool does during the parametric analysis is to modify the output node capacitance step by step according to the desired range of study, to perform the simulation, and to plot the desired delay information.

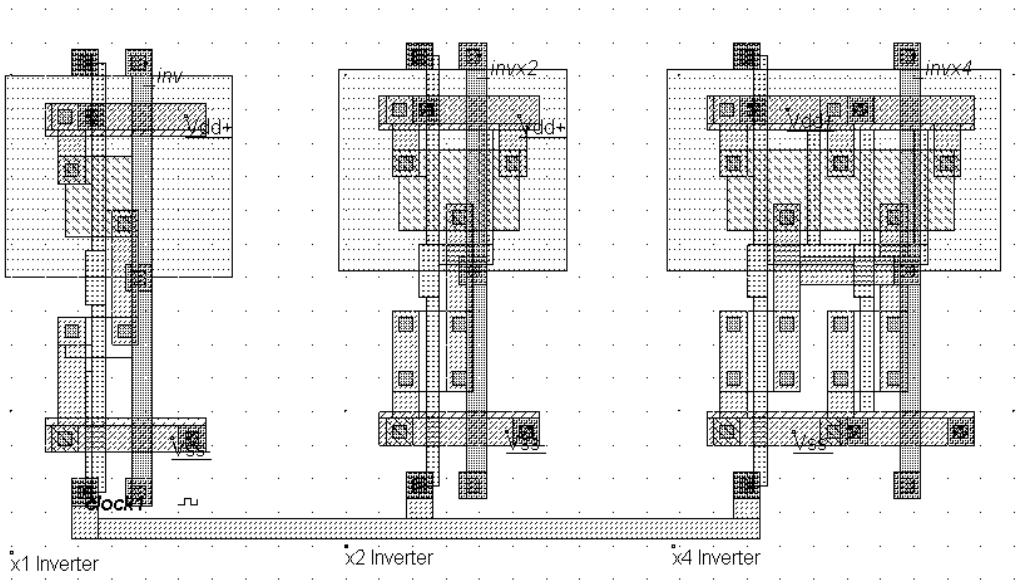


Figure 5-45: Three sizes of inverters used to investigate the delay vs. capacitance load (Invx124.MSK)

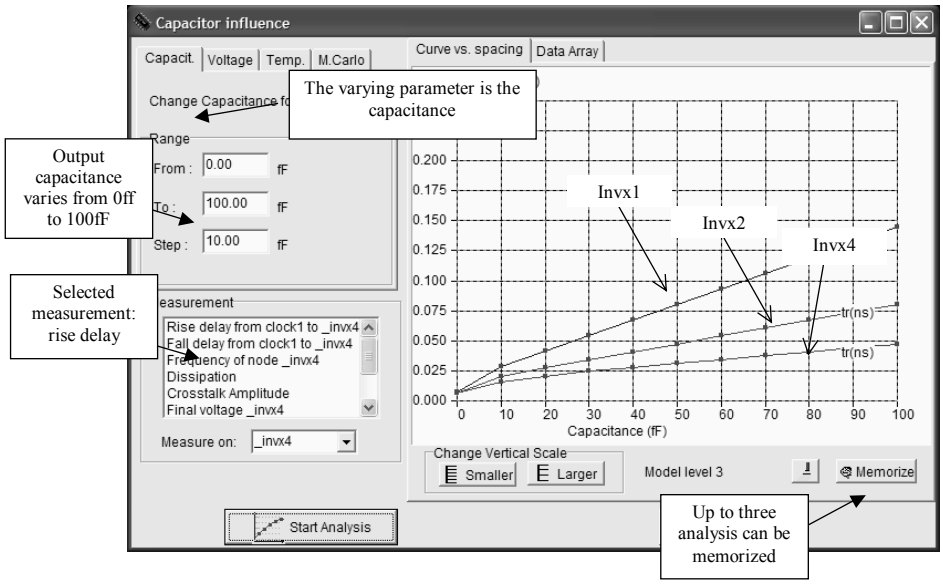


Figure 5-46: Three sizes of inverters used to investigate the delay vs. capacitance load (Invx124.MSK)

From the simulations of figure 5-46, it can be seen that a standard inverter delay (x1) increases rapidly with the capacitance. The inverter with a double drive (x2) has a switching delay divided by 2, the inverter with a quadruple drive (x4) has a switching delay divided by 4. As interconnects may be assimilated to capacitance, driving long interconnects requires large buffers. High drive buffers keep the propagation delay short, at the price of a proportionally higher current consumption. The clock signals, bus, ports and long wires with severe time constraints use such high drive circuits.

A fixed ratio is maintained between p-channel MOS width and n-Channel MOS width to balance the rise and fall time. It is much easier and safer to design the logic cells with similar rise and fall time performances, otherwise the timing analyzer would have to consider rise and fall time cases separately.

9. Repeaters for Improved signal transport

Long distance routing means a huge loading due to a series of RC delays, as shown in figure 5-47. A long line may be considered as a series of RC element where R is the serial resistance and C the ground capacitance. For example, one RC cell represents one millimeter of interconnect. If a very long interconnect is implemented between an emitter and a receiver inverter, the delay is increased according to n^2, where n is the number of RC cells, as given in equation 5-6.

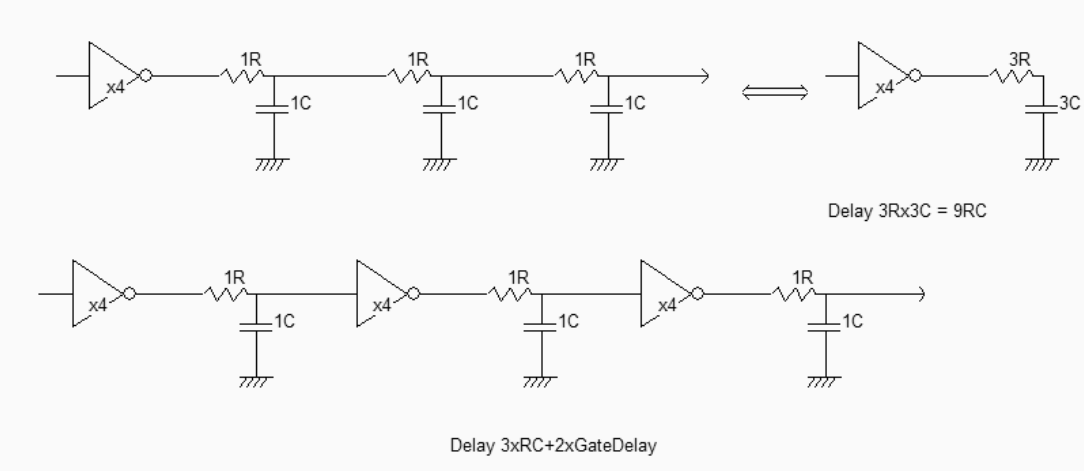


Fig. 5-47 : The propagation delay of a long line with one inverter can be longer than with three inverters (RcLines.SCH)

In the case of a long line driven by a single inverter, the propagation delay on a line modelized by RC cells is given by:

$$t_{dly} = t_{gate} + nR.nC = t_{gate} + n^2 RC \qquad \text{(Equ. 5-6)}$$

The propagation delay on a long line is not linearly dependent on the number of cells *n*, but proportionally dependent on the square of *n*. A good alternative is to use repeaters, by splitting the line into several pieces. Why can this solution be a better one in terms of delay? Because the gate delay is quite small compared to the RC delay. If two repeaters are inserted, the delay becomes:

$$t_{dly} = 3t_{gate} + 3RC \qquad \text{(Equ. 5-7)}$$

Consequently, if the gate delay is much smaller than the RC delay, repeaters improve the switching speed performances, at the price of a higher power consumption. In the case of very long interconnects (Several mm), it is interesting to place repeaters on the path of the interconnects to limit the slowing down effect of the interconnect resistance.

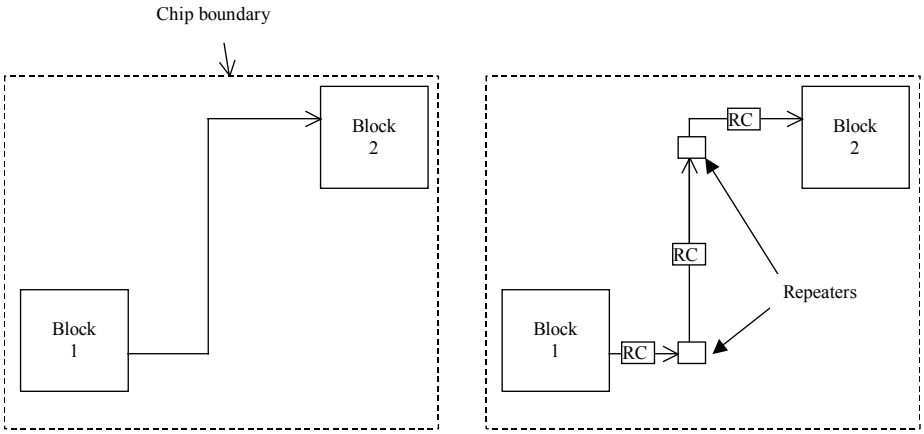


Fig. 5-48 : Inserting repeaters in long lines

The example given in figure 5-48 corresponds to the propagation of a signal from block 1 to block 2, situated at opposite corners of the integrated circuit. The associated model is a set of three RC elements. Each portion of interconnect is several millimeters long.

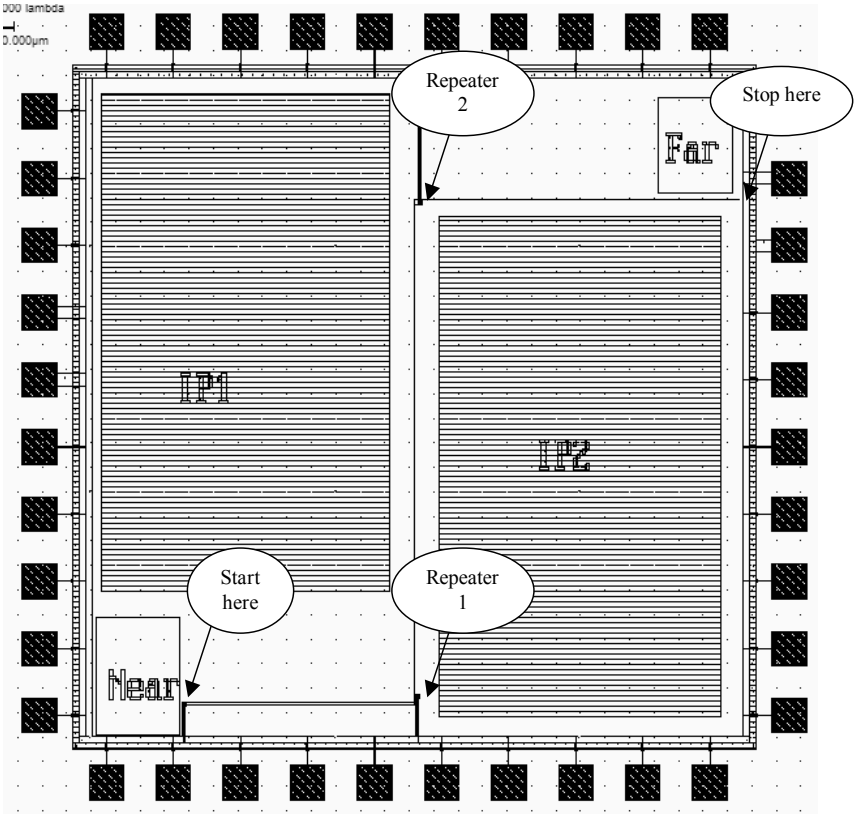


Fig. 5-49 : Propagation with and without repeaters (Repeater.MSK)

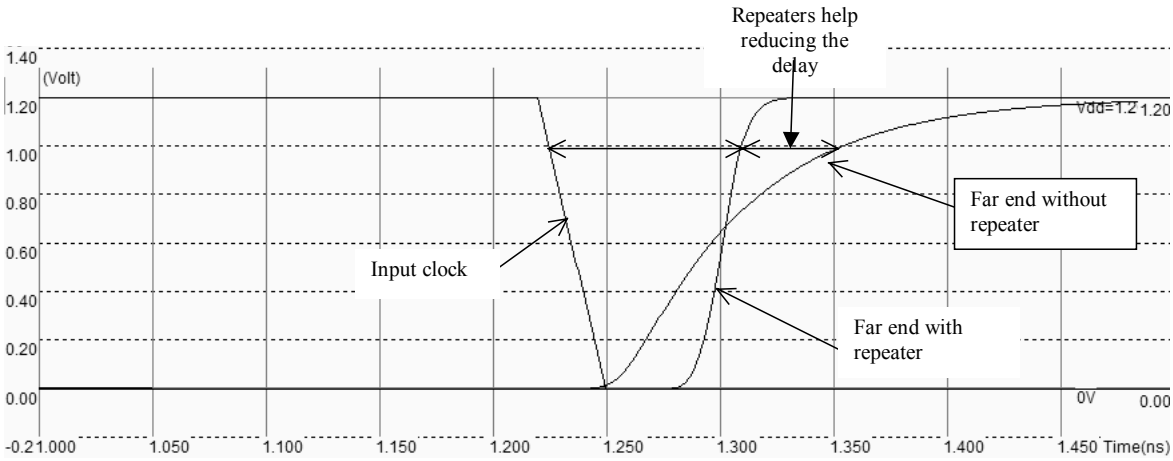


Fig. 5-50 Inserting repeaters in long lines (Repeater.MSK)

10. Crosstalk effects in interconnects

Coupling Increased with scale down

The crosstalk coupling represents the parasitic transient voltage induced by a switching interconnect on a neighbor interconnect. The disturbance may be high enough to create a temporary erroneous state on an interconnect which is supposed to be constant. Over the past recent years, the crosstalk effect has been the focus of active research, in terms of modeling and technology. The main reason for this interest is illustrated in Figure 5-51. The aspect of interconnects has dramatically changed with the technology scale down.

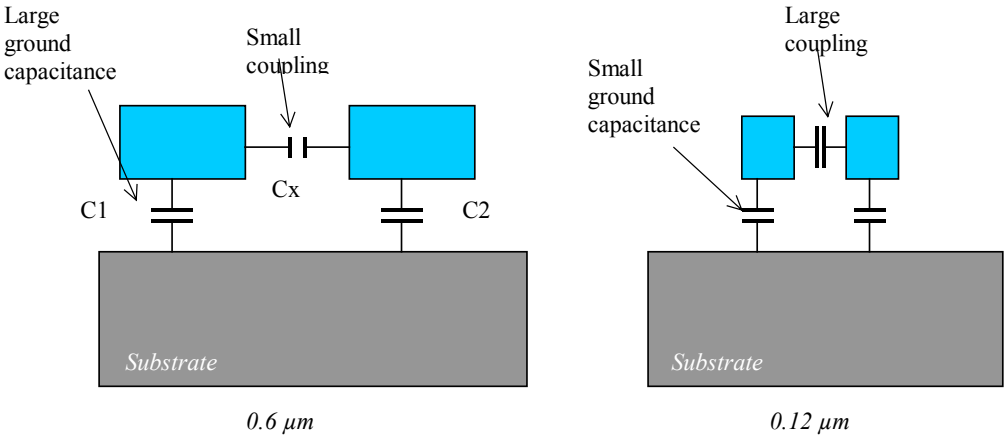
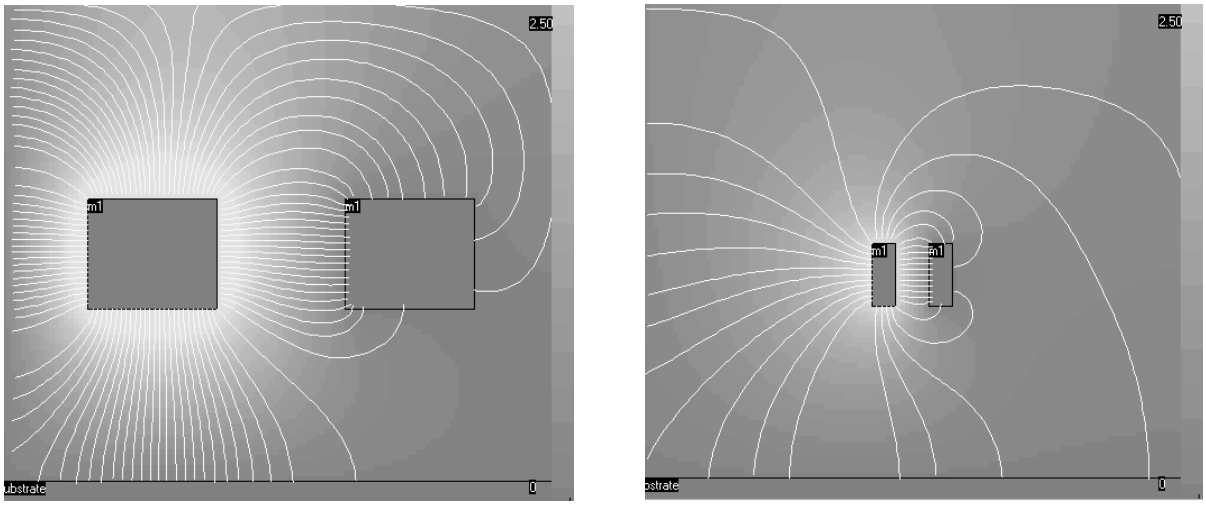


Fig. 5-51 The scale down tends to increase lateral coupling and decrease vertical coupling



(a) 0.6µm (b) 0.12µm

Fig. 5-52 The coupling between adjacent interconnects in 0.6µm and 0.12µm technology showing a very strong coupling increase with the scale down

Using the command **Analysis → Interconnect analysis** with FEM, we compare the coupling effects within two conductors, between the 0.6µm and the 0.12µm technologies. The field lines are computed by a clock on **Compute Field**. In figure 5-52-a, the field lines link the left conductor mainly to ground, with about one third of the lines to the right conductor, which creates the coupling effect. In figure 5-53-b, the number of field lines have been reduced, because of reduced conductor surfaces. Only 15 field lines couple to ground, which means a very low capacitance to ground. However, the coupling field lines are still numerous and very short, which means a very high coupling

Some details about the size and electrical properties of the interconnects are reported in table 5-5. The metal width and spacing are scaled according to the lithography improvement, but the interconnect thickness is not reduced with the same trend. Starting at 0.18µm, copper has been proposed as an alternative to aluminum, for its lower resistivity. To decrease the crosstalk coupling capacitance, low permittivity (Low K) dielectrics have been introduced, with 0.18µm technology.

Technology	Metal layers	Lower metal width (µm)	Metal spacing (µm)	Thickness (µm)	Low K	Interconnect material	Microwind2 file
1.2µm	2	1.8	2.4	0.8		Al	cmos12.rul
0.7µm	2	1.2	1.6	0.7		Al	cmos07.rul
0.6µm	3	0.75	1.0	0.7		Al	cmos06.rul
0.35µm	5	0.6	0.8	0.7		Al	cmos035.rul
0.25µm	6	0.5	0.6	0.6		Al	cmos025.rul
0.18µm	6	0.3	0.4	0.5	3.1	Al, Cu	cmos018.rul
0.12µm	6-8	0.18	0.24	0.4	2.8	Al, Cu	cmos012.rul
90nm	8-10	0.15	0.2	0.35	2.5	Cu	cmos90n.rul
70nm	8-12	0.1	0.14	0.3	2.0	Cu	cmos70n.rul

Table 5-5: Evolution of interconnect parameters with the technology scale down

Simulation of the crosstalk effect

The simulation of the crosstalk effect is based on two inverters, one considered as the affecting signal, the other as the victim signal. The inverters are connected to long interconnects routed with the minimum distance. The victim is connected to a weak inverter, and surrounded by two aggressor lines connected to a very powerful inverter, to create the maximum crosstalk effect (Figure 5-53).

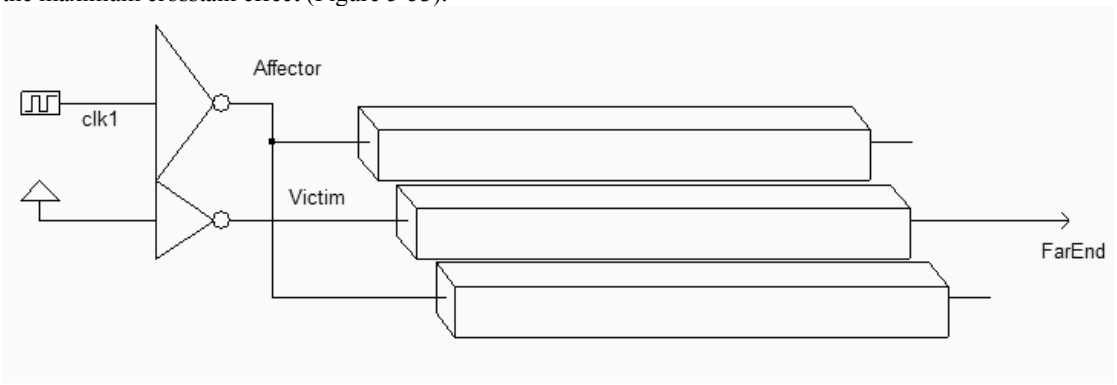


Fig. 5-53 The coupling configuration used to simulate the crosstalk effect (Crosstalk.SCH)

The serpentine shown in the layout of figure 5-54 corresponds to approximately 1mm of interconnect. Virtual resistance symbols are added in the middle of the interconnect to handle the RC effect in the simulation. Notice the unbalanced inverter size to create the worst case conditions for parasitic coupling.

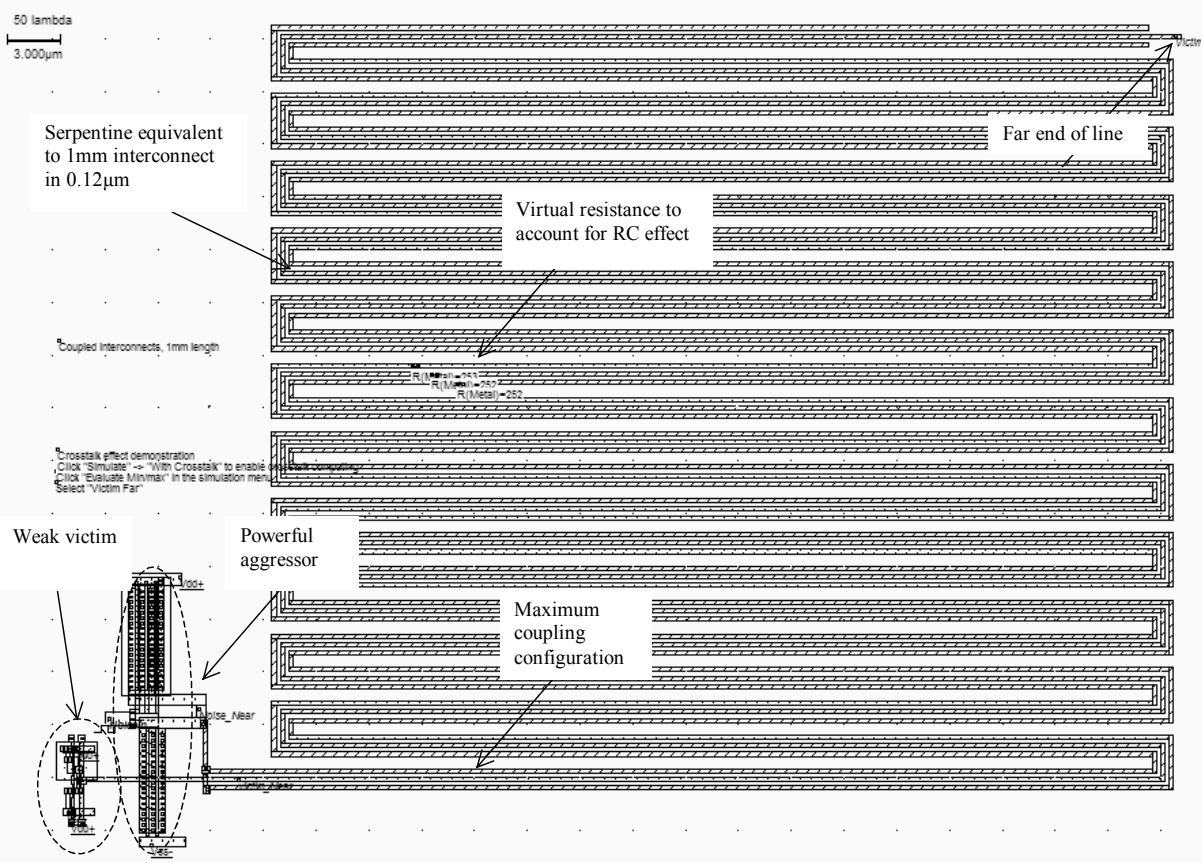


Fig. 5-54 Implementation of strongly coupled lines in worst case configuration (Crosstalk.MSK)

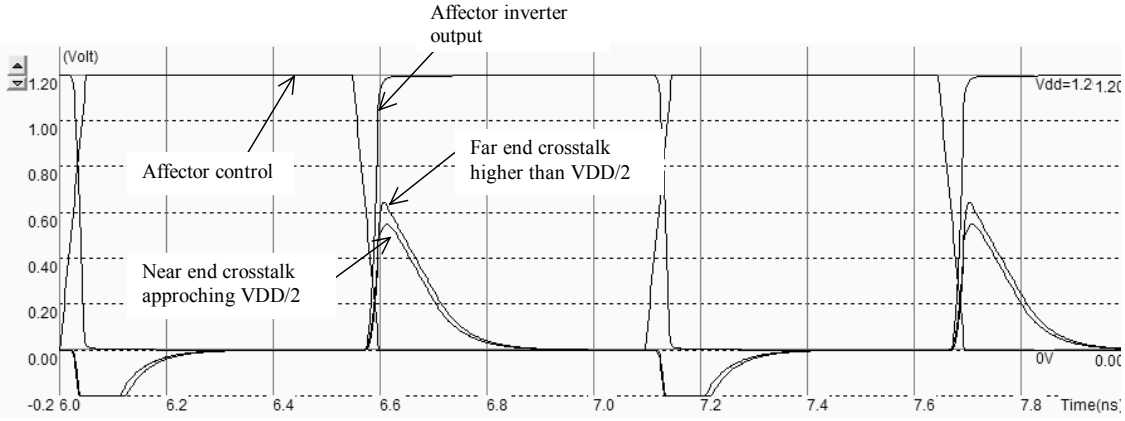


Fig. 5-54 Simulation of the crosstalk coupling in a 1mm interconnect (Crosstalk.MSK)

When the aggressor lines are switching, the coupling is strong enough to increase the voltage at the far end of the victim line, higher than the switching threshold of logic gates (Which is around VDD/2), which may provoke a permanent logic fault (Figure 5-54). The noise is quite impressive. Remember that the line is only 1mm long, which is very common in circuit design. However, the situation where the 1mm interconnect is driven by a very low drive inverter is not usual. Nevertheless, crosstalk is very dangerous and almost uncontrollable when dealing with millions of

interconnects, as may be found in high complexity designs. One solution to avoid crosstalk is to avoid routing long interconnects. The critical routing length is the limit above which a crosstalk fault may occur. This metrics has recently been introduced in design guidelines. In 0.12µm CMOS technology, the critical routing length is 1mm. It means that interconnects longer than 1mm could suffer from crosstalk noise in worst case conditions. In the case of very long routing, repeaters should be used.

Low K Dielectrics

When investigating the maximum crosstalk amplitude versus the technology for a given interconnect length, we observe a severe increase of the coupling effect, as a direct consequence of lithography improvements. In ultra-deep submicron technologies (Lithography lower than 0.18µm) the permittivity of the lateral oxide that fills the spacing between adjacent interconnects is reduced (Low K dielectric with a permittivity of around 3.0), while the oxide that separates vertical layers is kept with a high permittivity (Around 4.0 for SiO2).

The main effect is the decrease of lateral coupling effects. The introduction of low K materials may reduce the coupling effect up to a certain limit. Several CMOS compatible materials exist, which must be compatible with the CMOS process. Low K oxides are usually called SiOLK (For Silicon Oxide Low K<Gloss>). The SiOLK permittivity should ideally be 1, that is corresponding to an air gap between interconnects, still in a research phase. In practice, for 0.12µm technology, SiOLK ε_r is around 3.0.

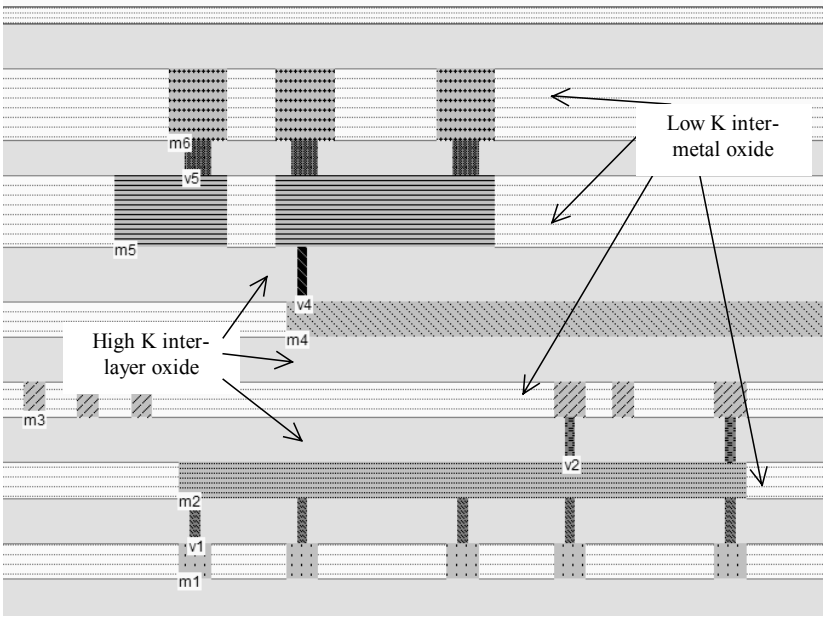


Fig. 5-55: Low dielectric permittivity between lateral metal interconnects reduces the crosstalk effect (Metals.MSK)

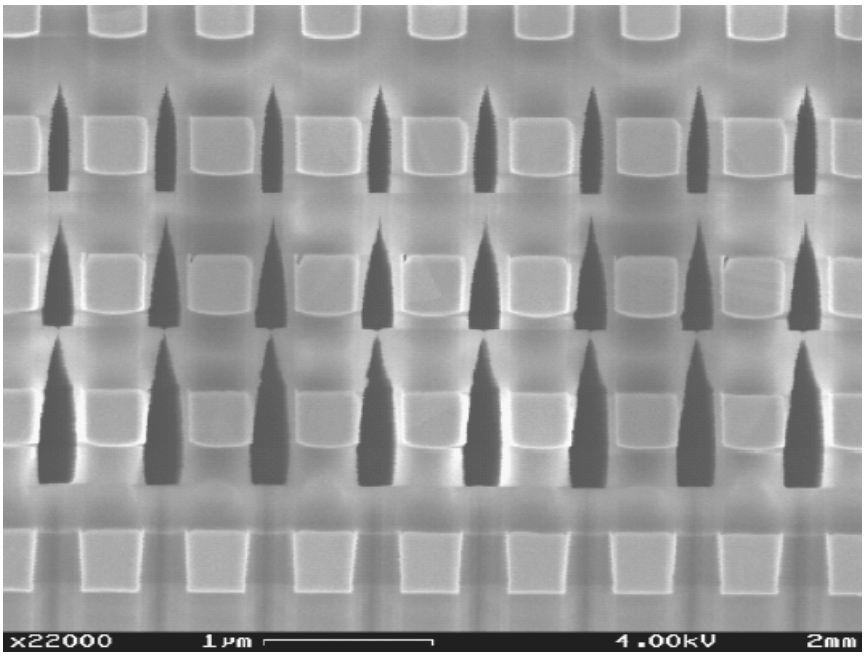


Fig. 5-56: Air-gap between interconnects to cut by half the crosstalk coupling (Courtesy ST-Microelectronics)

11. Antenna Effect

During the fabrication of interconnects, charges may accumulate and endanger the MOS gate by forcing current through the gate oxide [Hastings]. This effect, called antenna effect, appears on long metal interconnects connected to small gate oxide areas, without any path to diffusion. The antenna effect is particularly important in deep submicron technology: without any possible discharge path, the interconnect accumulates sufficient charges to rise its potential to several volts, positive or negative depending on the nature of the chemical process step. Usually, plasma etching charges the interconnect with electrons, corresponding to a negative charge with respect to the substrate ground voltage.

Antenna rule

With 0.35µm process, specific antenna design rules have been introduced. If we consider an interconnect with a length *Li*, a width *Wi*, its surface *Ai* is *Wix*Li. If this interconnect is connected to a MOS device with a length *L* and width *W*, corresponding to a channel surface *A*, the following rule should be verified:

$$Ai \leq R_{antenna}.A \qquad \text{(Equ. 5-10)}$$

where

A=WxL=MOS channel surface (m^2)

Ai=WixLi=interconnect surface (m^2)

$R_{Antenna}$ = antenna ratio (around 100)

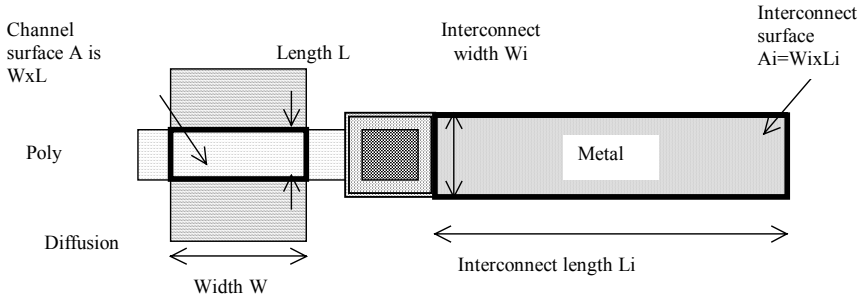


Figure 5-56: The antenna rules related to the surface of the interconnect with respect to the surface of the gate

Design example

An example of valid and invalid interconnect design is given in figure 5-57. The upper layout complies with the antenna rules as the interconnect surface is less than 100 times the gate surface. In contrast, the design (b) is dangerous as the interconnect surface is more than 100 times larger than the gate surface, which has been designed very small. The antenna rules are not yet verified by the design rule checker of Microwind.

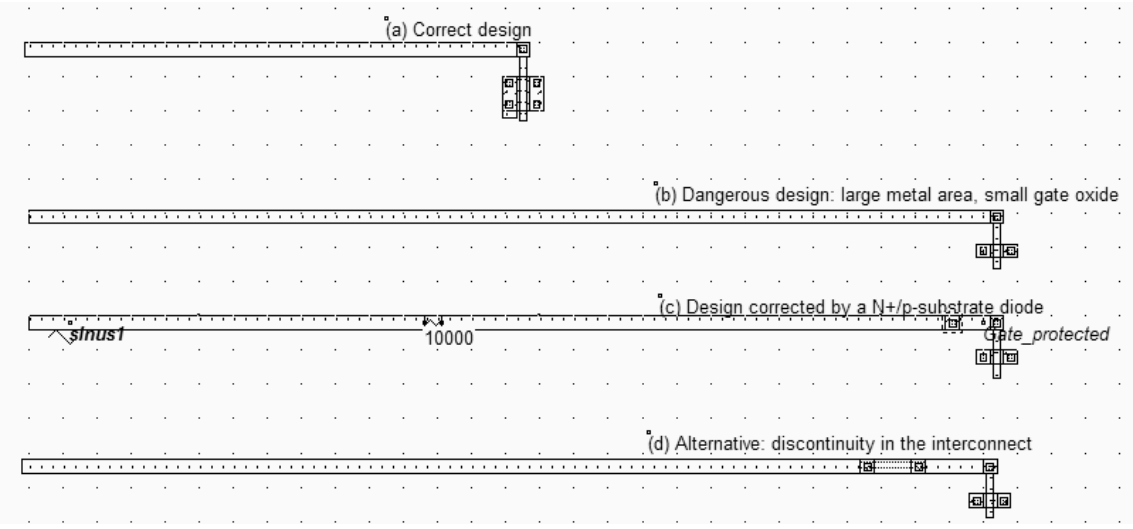


Figure 5-57 Illustration of the antenna design rule (AntennaRules.MSK)

One solution consists in creating a discontinuity by using a bridge with a higher metal, so that when the lower metal is etched, the main part of the interconnect will not be connected to the gate (Figure 5-57-d). An alternative way to avoid the antenna effect is to build a discharge path to evacuate the parasitic charges accumulated during fabrication. A

simple diode is an efficient solution (Figure 5-58). Its parasitic capacitance is quite small, and the diode has no important electrical effect on the nominal signal transport.

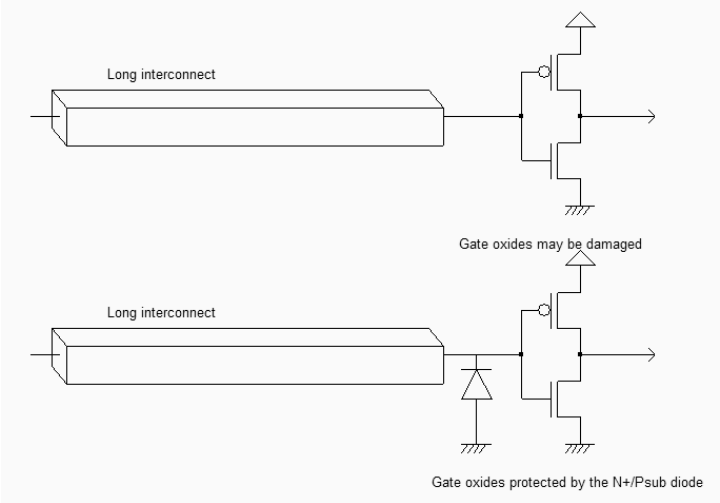


Figure 5-58 Inserting a diode to discharge the interconnect during plasma etching (AntennaRules.SCH)

Simulation

The N+/P-substrate diode inserted near the gate (Figure 5-57-c) turns on when the interconnect voltage is negative, or higher than the reverse Zener voltage, as seen in the simulation reported in figure 5-59. Notice that the BSIM4 model has been used to handle the Zener effect. The Y axis has been changed to -2 to 24V, thanks to the Y scale cursors situated at the left upper part of the voltage chronograms. By default, Microwind does not extract diodes. This is why an option layer surrounds the N+ area near the gate, with a tick in front of **Extract Diode inside box**.

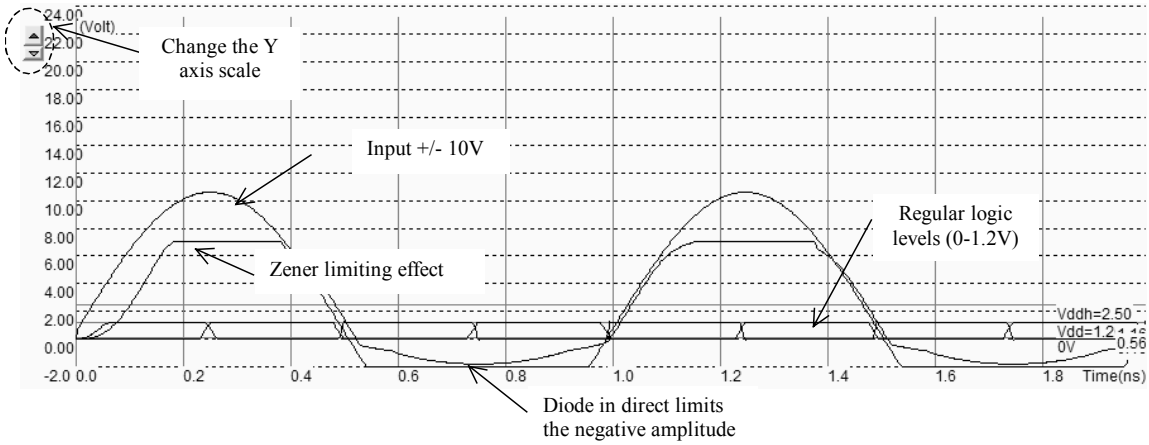


Figure 5-59 The diode clamps the negative charges, an limits the positive amplitudes (AntennaRules.MSK)

The serial resistance used for simulation is very high (10Kohm). Removing that resistance would completely hide the clamping effect of the diode. In reality, the charging of the interconnect is not equivalent to a perfect voltage source, as

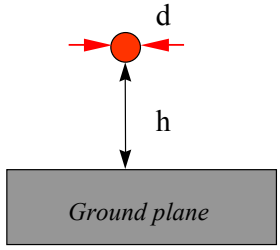
Microwind does with the sinusoidal property. The high serial resistance accounts for the weak charging process, which can be counterbalanced by a small discharge diode.

12. Inductance

The inductance effect is not significant in signal transport because of the high serial resistance of interconnects. This is why we have not paid a lot of attention to the inductance value and its possible consequence on the delay estimation or crosstalk amplitude. A lot of research has been dedicated in the recent years to the extraction and handling of inductance. The inductor is described in details in chapter 13, as a stand-alone passive component for application in radio-frequency circuits. In that case, the inductance is no more a parasitic effect but a voluntary effect. We give here a brief evaluation and illustration of the parasitic inductance effect in deep submicron interconnects.

Parasitic inductance formulation

The wire inductance formulation is based on the estimation of a cylinder for which a very simple formulation exists [Lee]. A well known rule of thumb consists in approximating the serial inductance to 1nH/mm, which is close to reality in the case of bonding wires. The wire has a cylindrical shape and is situated far from the ground plane.



$$L = \frac{\mu 0}{2\pi} ln(4\frac{h}{d}) \qquad \text{(Equ. 5-7)}$$

with

μ_0=1.257e^{-6} H/m for most materials (Al, Cu, Si, $Si0_2$ and Si_3N_4)

d= wire diameter (m)

h = height of the wire vs. ground (m)

In the case of metal interconnects, the formulation 5-7 is adapted with an approximation of the interconnect diameter, based on the conductor width and thickness. The serial parasitic inductance of the conductor appears in the navigator menu, after extraction, together with the capacitance and resistance (Figure 5-60). A metal interconnect exhibits an inductance of around 0.5nH/mm. Notice that the lineic inductance value is also provided in the **interconnect analysis** window, accessible from the **Analysis** menu.

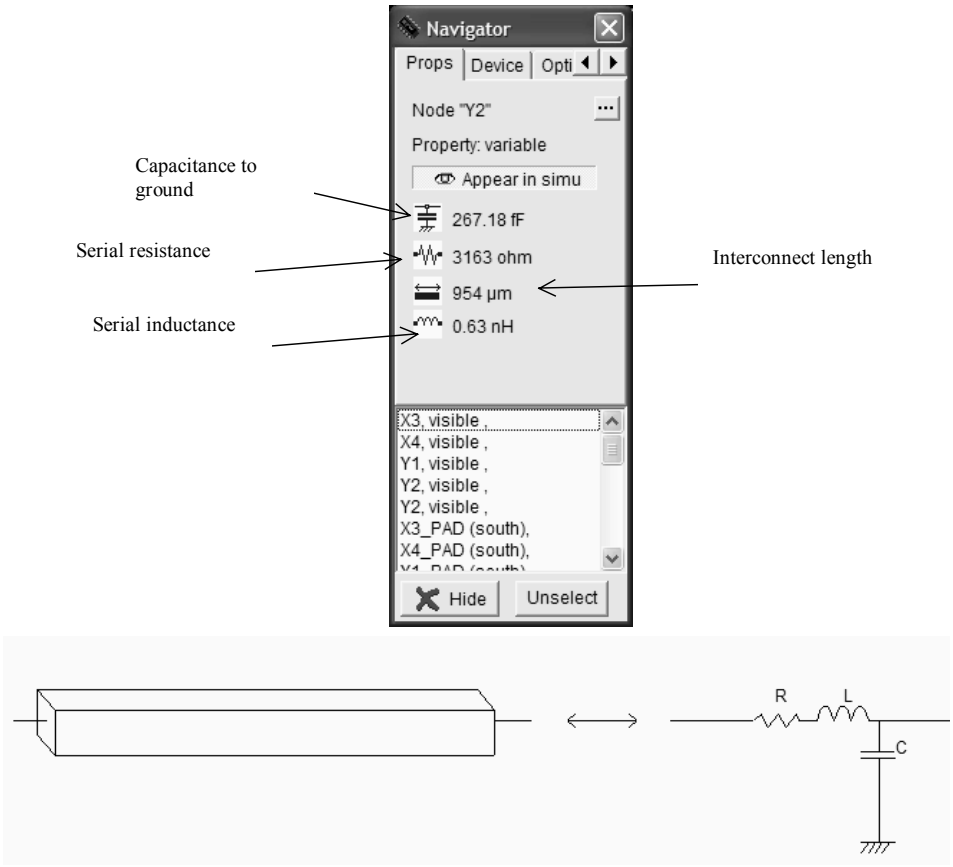


Figure 5-60 An evaluation of the parasitic serial inductance is given in the navigator menu (Rlcg.SCH)

Simulation with/without inductance

Let us compare the signal propagation of a logic signal within a 500µm interconnect with and without the serial inductance. In Microwind, the inductance must be added through a virtual inductor symbol that may be found in the palette. The inductor is placed at the beginning of the line. Handling the simulation of the inductance is not simple in Microwind, which has been optimized for RC networks. Several problems rise at simulation: the initial ringing is very high, which is not realistic at all. This numerical instability is due to the initialization of the inductance, which disturbs the circuit in the early nanoseconds. Secondly, the simulation step by default is not small enough to ensure a correct simulation. In most cases, keeping the default time step of 0.3pS will create the oscillations of all floating nodes with several volts of amplitude. The only solution consists in reducing the simulation time step, at least to 0.03ps or even less. After some nanoseconds, the simulation becomes stable, as displayed in figure 5-61. We see no significant difference between the RC and RLC models.

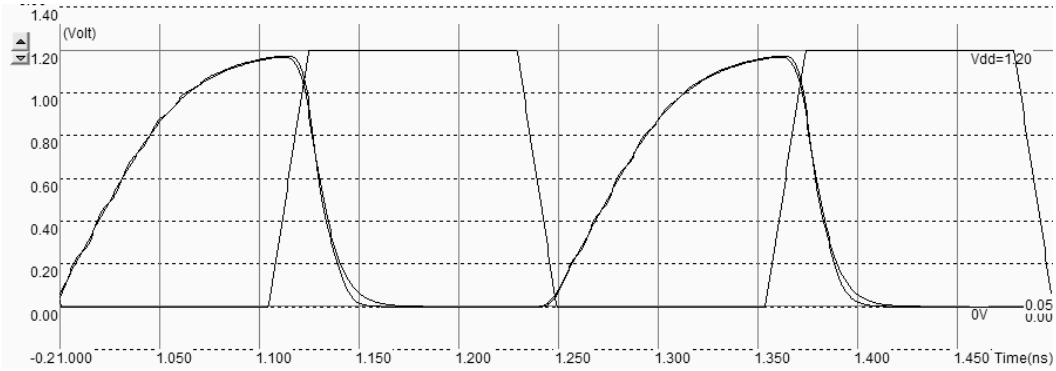


Figure 5-61 Simulation of the signal propagation within a 500µm interconnect with using the RC and RLC models (InductanceInv.MSK)

To observe the inductance effect, we change the configuration and now drive the same RLC line with strong buffers, as shown in figure 5-62. Notice the ringing effect due to the combination of inductance and capacitance at the far end of the interconnect (Figure 5-63). In reality, the resistance, capacitance and inductance are distributed along the wire, which tend to limit the amplitude of the ringing effect, and fastens the damping of the oscillation to a stable logic value.

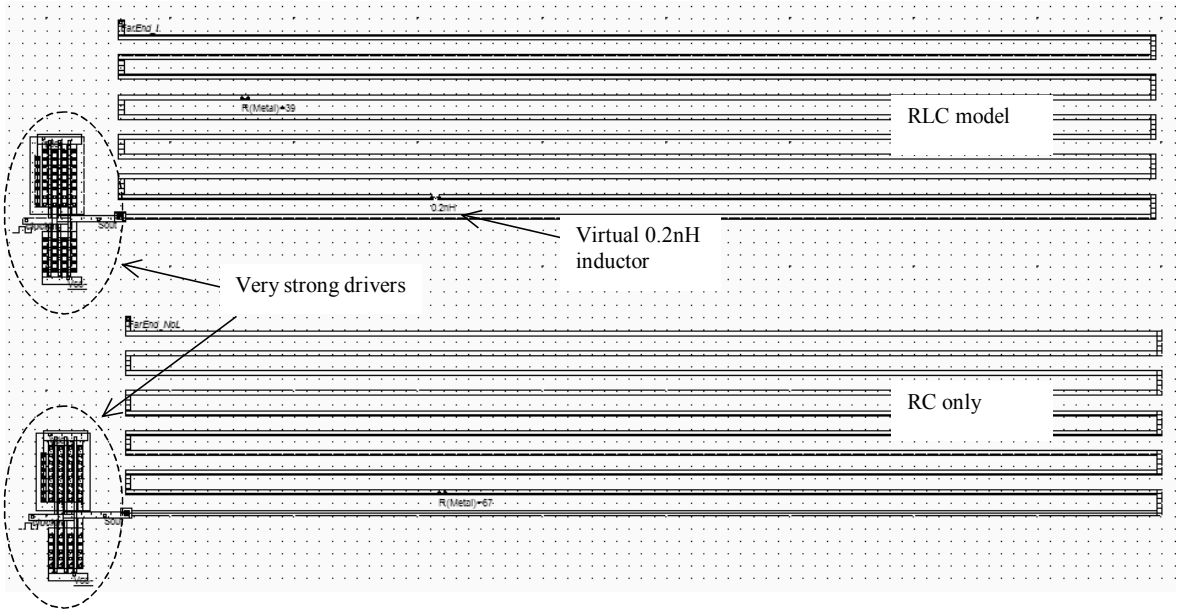


Figure 5-62 Layout of the interconnect with and without serial parasitic inductance (Inductance.MSK)

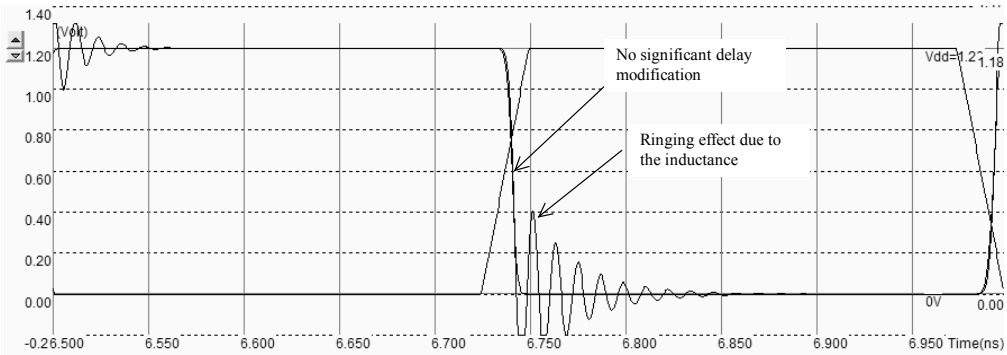


Figure 5-63 Simulation of the signal propagation with and without inductance. This result was obtained with a simulation time step decreased to 0.02ps (inductance.MSK)

Although the ringing effect is very important in this new simulation (Figure 5-63), the switching delay is not altered in a significant way. Consequently, the inductance effect may be neglected in a first order approximation in the analysis of signal propagation. This assumption is valid if the buffer strength is not too large, and the interconnect not too short.

13. Conclusion

In this chapter, we have described some layout techniques for designing metal interconnects between devices. We have given some information regarding the design rules and the electrical parameters for metal interconnects, such as the resistance, capacitance, and later the inductance. The signal propagation has been analyzed from a point of view of RC delay, technology scale down, and parasitic crosstalk effect. We have also made a rapid investigation of the role of the parasitic inductance in the signal transport.

REFERENCES

[Weste] N. Weste, K. Eshraghian "Principles of CMOS VLSI design", Addison Wesley, ISBN 0-201-53376-6, 1993

[Baker] R.J. Baker, H. W. Li, D.E. Boyce "CMOS circuit design, layout and simulation", IEEE Press, ISBN 0-7803-3416-7, 1998

[Hastings] Alan Hastings "The Art of Analog Layout", Prentice Hall, ISBN 0-13-087061-7

[Sakurai] Sakurai T. « Closed-form expressions for interconnection delay, coupling and crosstalk in VLSIs », IEEE Transactions on Electron Devices, vol 40, n°1, pp 118-124, January 1993.

[Delorme] Delorme N., Belleville M., Chilo J. « Inductance and capacitance analytic formulas for VLSI interconnects » Electronic letters, vol 32, n°11, pp 996-997, May 1996.

[Lee] Thomas H. Lee "The design of CMOS radio Frequency Integrated Circuis", Cambridge University Press, isbn 0-521-63922-0

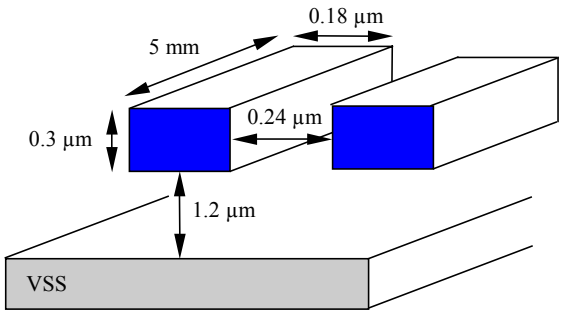
[Bendhia] Sonia Delmas-Bendhia, Fabrice Caignet, Etienne Sicard "A new method for measuring signal integrity in CMOS Ics" Paper in Microelectronics International, Volume 17, N°1, January 2000, pp 17-21.

[ST] www.st.com

[Smith] ASIC (Elmore delay)

EXERCISES

5.1 We consider two coupled lines in metal1, with a 5mm length, in 0.12µm technology (Figure 5-64). What is the equivalent model of the line, in a first order approximation? Calculate the total serial resistance *R*, ground capacitance *Csub* and crosstalk capacitance *C12*, using the command **Analysis →interconnect analysis with FEM** in Microwind. What is the best technological option for interconnect/oxide material as far as signal integrity is concerned (We suppose LowK=2)?

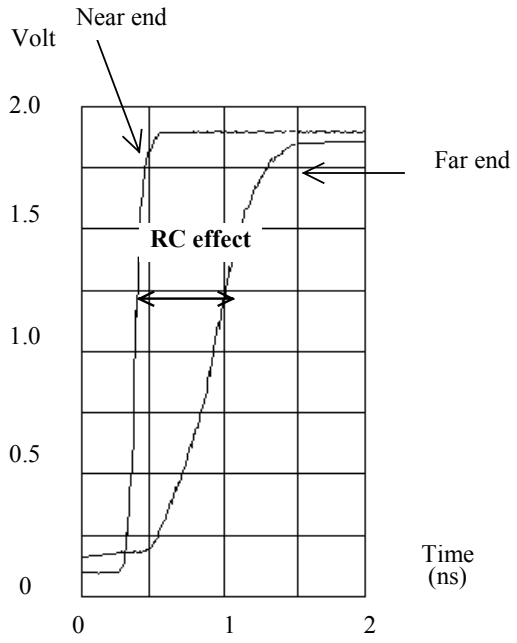


Case	R	C_{12}	C_{sub}
N°1: Al, SiO2			
N°2: Al, Low K			
N°3: Cu, SiO2			
N°4: Cu, Low K			

Figure 5-64: Two coupled lines in metal1, 10mm length

Answer: The model is based on coupled RC lines. The best technological option is <Sonia>

5.2 An experimental measurement of the RC effect has been realized in 0.18µm technology (Figure 5-65). Details on the real case configuration are also provided. Create the corresponding layout for this configuration and compare the simulation with measurements.



Technology	Cmos 0.18µm
RUL file	Cmos018.RUL
Interconnect length	10mm
Interconnect width	4 lambda (0.4µm)
Metal layer	Metal 5
NMOS buffer size	W=32µm, L=0.18µm
PMOS buffer size	W=54µm, L=0.18µm

Figure 5-65: The parameters of the RC propagation test-case in 0.18µm technology

Answer: See Appendix F