POLITECNICO DI TORINO

Master's degree
in Nanotechnologies for ICTs

Master Thesis

RC simulations of advanced metallizations in integrated circuits

Supervisor
Prof. Michele Goano

Candidate
Simone Esposto

Co-supervisors
Ing. Ivan Ciofi
Prof. Marco Vallone

Academic year 2021-2022

Summary

Semiconductor industry is working unceasingly to further improve performance and reduce manufacturing costs of current electronic devices. In last decade the number of transistor per chip increased aggressively from 1 billion in 2010 to roughly 32 billions in 2020 following Moore's predictions [15]. The realization of ever smaller transistors has been possible thanks to the development of new manufacturing equipment and the employment of novel materials. Lithography has always had a leading role in IC manufacturing. The introduction of deep-ultraviolet (DUV) light sources based on excimers with smaller wavelengths allowed to increase the resolution.Nowadays, it is possible to obtain 80 nm features in a single exposure with ArF excimers whose characteristic emission wavelength is 193 nm. In last years new alternatives to push further resolution limits have been used in latest technology nodes, like multiple patterning techniques that will be briefly described in the following and whose employment enabled minimum features size of about 20 nm [22]. For many years ArF 193 nm ruled manufacturing processes for several reasons among which: the switch from dry to immersion lithography that improved the numerical aperture [NA] and consequently, resolution; the effective implementation in high-volume manufacturing industry and the introduction of aforementioned multi-patterning techniques. Unfortunately to further scale electronic devices multi patterning requires more exposure steps and more masks hence higher costs per chip and the only way is to decrease the wavelength of exposure light.At the beginning of 21st century a new lithographic tool based on Extreme Ultraviolet (EUV) light source able to emit radiation with wavelength around 13.5 nm was born. It faced many challenges to reduce the cost/process ratio and only in the last five years it became part of High-Volume-Manufacturing by leading companies as Samsung and TSMC [7]. In microelectronics the front-end-of-line (FEOL), which includes all the processes devoted to the patterning of circuit elements (i.e. transistors,resistors, capacitors etc...) directly on the semiconductor substrate, have owned a central role for many years and research pushed performances further than expectations, however, the increase of chips per unit area have put in serious disadvantage back-end-of-line (BEOL) and this latter is undermining performances of current/future technology nodes. More devices require more interconnects with tighter line-gap periods. Resistance and capacitance are both affected because of smaller and closer lines and these are directly responsible for interconnect delay. This work begins with a short review of scaling challenges that FEOL and BEOL have been experiencing in these years. Successively, new Semi-Damascene modules for BEOL, developed by IMEC, are presented to deal

with its benefits in terms of resistance and capacitance that may be achieved by employing new materials, instead of the well-known copper, thanks also to the possibility to integrate air-gap (AG) scheme, an option that relies on substituting the inter-metal dielectric (IMD) layers with air to drastically reduce parasitic capacitances. In the following, AG is analyzed considering different extensions also in the inter-layer dielectric (ILD) layer and an evaluation of RC-delay improvements is presented. Finally, a study on power delivery networks (PDN) focuses on estimating the real impact on IR-drop of every metal layer that demonstrates to be highly dependent on the current density flowing in the component, denoting the poor information that a PDN resistance breakdown may give for scaling design rules.

Acknowledgements

I am really thankful to all people who have accompanied me during these years and during the making of this project. I would like to thank my family whose support has never missed throughout my studies and all my friends who believed in me. I am really grateful to my university and IMEC that allowed me to take part in this project and I would also like to thank my supervisor who helped me to edit and finalize this thesis. Last, but not least, a great acknowledgment goes to Ivan Ciofi, my team leader, who has been able to guide me during this project outstandingly, but more importantly, his deep dedication taught me the attitude and the commitment needed to do research.

Contents

Chapter 1

Scaling challenges

1.1 Introduction

The constant miniaturization of metal-oxide-semiconductor field-effect-transistors (MOSFETs) leaded to significant improvements in speed and manufacturing costs of integrated circuits (IC) in past decades. In order to scale-down device dimensions, hence to increase chip density, it is necessary to downsize every transistor. The foretelling Moore's law stating that transistors density would have doubled every 1.5/2 years is coming to the end for understandable reasons [15]. Gate length, that initially was the standard to label every new technology node, reached dimensions of the order of 10-20 nm and the underneath channel region is basically made up by less than a hundred Silicon atoms. Shrinking further dimensions is not feasible anymore since the gate length is approaching the physical distance between two single silicon atoms that is around 0.2 nm. This irremovable obstacle pushed for seeking new devices able to process more information at the same time or new ICs architectures. The main problem of MOSFETs is related to a poor channel control of the gate. In short channel MOSFETs drain and source are very close to each other and the gate loses its control on the channel because depletion regions near S and D influence the electric field below the gate. This condition generates a decrease in the threshold voltage with subsequent increase of sub-threshold current. Devices with multiple gates could allow a higher channel control, hence the research moved towards three dimensions devices able to guarantee high performances. In first years of 2000s Samsung demonstrated the design for mass-production of a new 3D device called fin-shaped-field-effect-transistor (FinFET) a tri-gate transistor that replaced most known planar MOSFETs in last technology nodes. This device is composed by a tall and thin silicon layer (the fin) standing on the wafer surface and surrounded by three sides by gate contact. To avoid a too low drain current the device height can be increased as much as

possible, moreover multiple fins are put in parallel to increase the total drain current. Unfortunately shrinking dimensions in FinFETs is becoming challenging due to process limitations in getting very high aspect-ratio (AR) fins and new alternatives as Lateral/Vertical Gate-All-Around FETs (LGAA/VGAA) that can further improve the control on the fin are on the horizon for sub-3 nm nodes. [6] [23].
The aim in scaling design is dictated by three main benchmarks in ICs, Power-Performance-Area (PPA). Although increasing chip density may reduce costs, power and performance have to be carefully investigated. Performance is typically related to switching speed of transistors, but from a broader point of view it must be related to the entire circuit taking into account RC delay also in BEOL whose impact can not be neglected anymore because it is already a limiting factor on the overall speed of the circuit in current technology nodes and it is going to be even more pronounced in future nodes (Fig.1.1) [24].

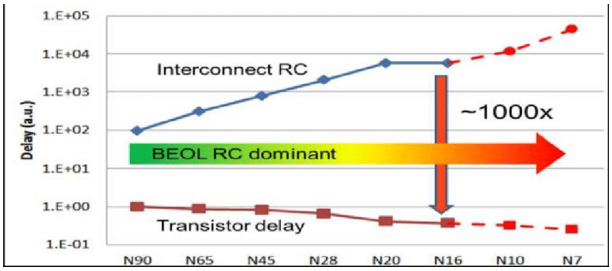


Figure 1.1: Interconnects delay increased by a factor of 100x from node N90 to node N16 [24]

For this reason, interconnects are gaining ever more importance in last years and the challenges to scale them down are the primary subject of this work. In the first chapter latest manufacturing techniques for BEOL are discussed and Cu-based Dual-Damascene modules are reviewed in order to better understand what are the reasons and needs to explore new metal options and new manufacturing methods, such as Semi-Damascene modules, in order to keep on scaling interconnects.

1.2 Back-end-of-line

The back-end-of-line (BEOL) consists of all the processes and circuitry elements devoted to connect the FEOL to the external world, that is to say that every element on FEOL must be connected to each other, following design rules, and they must be able to transmit signals and to be powered. As mentioned before, FEOL scaled-down aggressively and BEOL had to implement multiple metal layers stacked one on the other and properly isolated by dielectric layers whose role is

to isolate different metal layers and different metal lines in the same layer defined as inter-layer dielectric (ILD) and inter-metal layer dielectric (IMD) respectively. Nowadays, BEOL can be composed by 13 up to 15 different metal layers which are connected by vertical lines called vias. Bottom interconnects serve as local connections and top ones are used for semi-global and global connections, namely for connecting far away regions of the chip. Devices continue to scale down, however the number of pins needed to connect them to various type of signals (power,clock, etc...) remain the same, so new interconnections with tighter pitch, the minimum distance between two adjacent lines, are needed [9]. Lithography is surely one limiting factor to achieve higher resolutions, but fortunately a new class of technologies, known with the name of multiple patterning, was born to overcome the inherent limit of lithographic tools.
Fins generation benefited from multiple patterning since it allowed to double/triple fin's density and in 2010 with the 32 nm technology node, these new techniques became established also in BEOL manufacturing, decreasing the minimum obtainable distance between two metal lines. Different approaches are discussed in next sections.

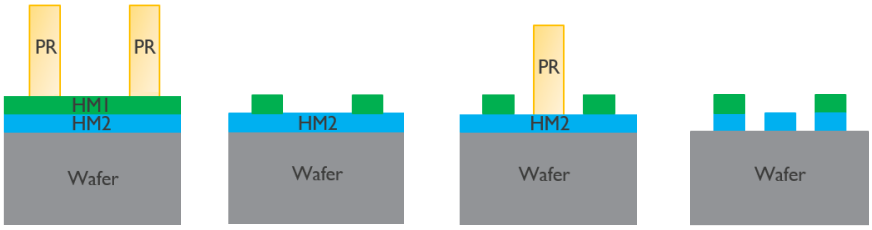
1.2.1 Multiple patterning

Multiple patterning includes a large variety of techniques that can enhance device density to reach sub-resolution feeatures. In this section only most used ones in BEOL are shown.

Litho-Etch-Litho-Etch (LELE)

This first technique relies on multiple exposures phases and the employment of two hard masks, a simple version of the process is shown in Fig. 1.2. After the first lithographic step (1.2a), the pattern is transferred on first hard mask (HM1) (1.2b), then the wafer is coated again and another lithographic process is performed using a mask that is half-pitch shifted w.r.t previous one (1.2c) and the layout is drawn on the second mask (HM2) (1.2d).

The process can be theoretically repeated n times (LE^n) to get even tighter pitches, however this process has intrinsically overlay issues that can arise due to different misalignment errors that in multiple exposures could be very detrimental. Moreover in BEOL, especially at current technology nodes, double/triple patterning is required for several layers and the cost increases for every extra exposure step.



(a) Litho 1 (b) Etch 1 (c) Litho 2 (d) Etch 2

Figure 1.2: Flow process of double patterning with LELE technique: a) lithographic step to pattern HM1 b) etching of HM1 c) lithographic step with half-pitch shift of the mask d) etching of HM2

Self-Aligned Double Patterning

Since the cost and overlay issues trade-off for LELE approach became unfavourable, many companies adopted the new Self-Aligned Double (or even Quadruple) patterning (SADP/SAQP). This technique has the enormous advantage to reduce the number of lithographic steps to a single exposure. Nonetheless, lines are self-aligned with each other thanks to the use of a spacer. In SADP, only lines at double the required pitch are patterned on photo-resist, they are called CORE lines, while GAP lines are obtained by depositing sidewalls on core lines and subsequent etching process (1.3). A simplified flow process is shown in Fig.1.4. The spacer is deposited on the patterned photo-resist (1.4b) and etched (1.4c). Since the result is a "negative" output, the process can be reversed using a sacrificial layer where lines are patterned (1.4d) and then filled again (1.4e). Finally the sacrificial layer is removed and the pattern can be transferred to hard mask (1.4f).

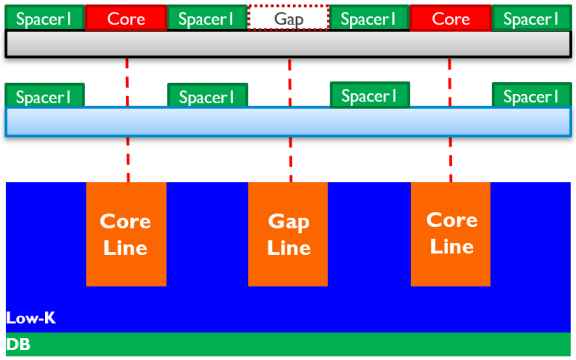
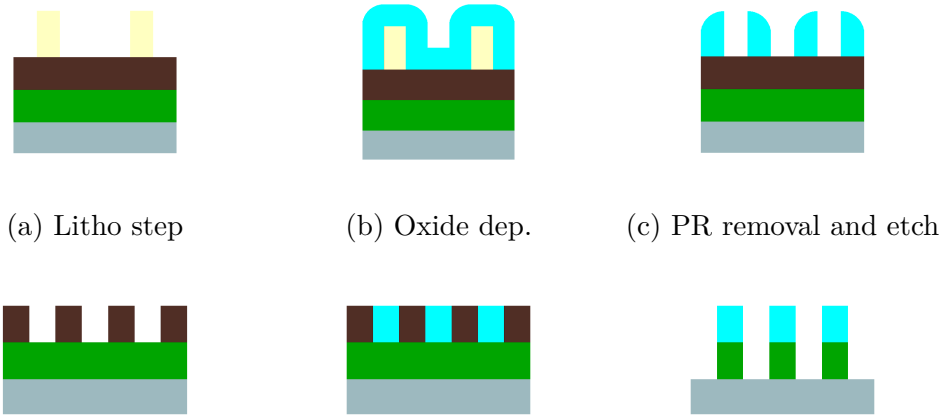


Figure 1.3: First lines (Core) are patterned with a lithographic step and second lines (Gap) are obtained at half-pitch without the need of second exposure



(a) Litho step (b) Oxide dep. (c) PR removal and etch

(d) HM etch (e) Oxide filling (f) HM1 and HM2 etch

Figure 1.4: Simplified flow process of double patterning with SADP technique: a) lithographic step to pattern lines on photoresist (PR) b) oxide deposition to create spacers c) PR removal and directional oxide etching d) oxide and first hard mask (HM1) etching e) oxide filling f) HM1 and HM2 are etched and only the lines protected by oxide are left

1.2.2 Dual-Damascene process

Aluminum has been employed for interconnections until late 90s when first IBM announced a new process based on copper, a metal with low resistivity and higher reliability than Aluminum [10]. Patterning copper was not an easy task because

for high AR lines the removal process was not very effective and it could lead to some residuals of material between metal lines that translate in the possibility of having short circuits, therefore IBM introduced the Dual Damascene process. Copper is not patterned directly but firstly trenches are created in the dielectric and then copper is deposited by electro-plating. The advantage is that you can fill via and lines trenches in one deposition step. If lines are patterned before the via, this latter will be self-aligned with the top metal layer. A simplified flow process is shown in Fig. (1.5). Two dielectrics layers are patterned with line and via layout (Fig. 1.5a). Unfortunately, Cu tends to migrate through dielectrics and copper ions could "poison" silicon chips, hence a barrier must be inserted between the dielectric and the copper (Fig.1.5b). Metal layer is then deposited and chemical mechanical polishing is applied to get a planar structure for the next interconnection layer (Fig.1.5c).

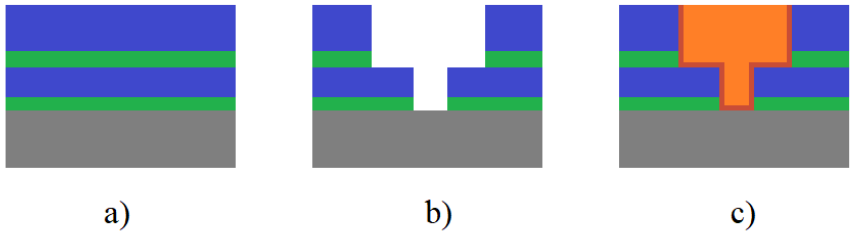
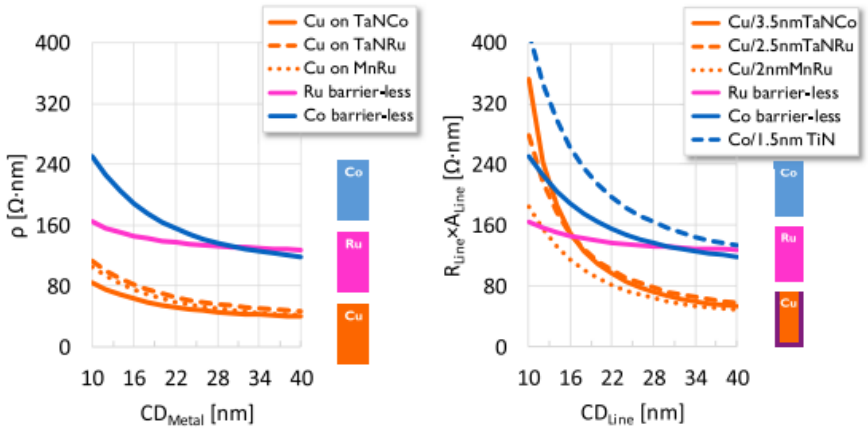


Figure 1.5: Dual-D flow process: a) two dielectric layers spaced out by a an etch stop layer (green) b) trench for lines and via opening c) barrier and copper deposition

Scaling Copper interconnections is becoming challenging because of barrier that is needed to prevent Cu ions diffusion. The barrier can not be less than few nm to be effective, however it reduces the conductive area of the wire significantly when dealing with lines of small dimensions. Barrier scaling reached its limit, moreover different barrier materials have been being studied to allow a better conductivity but it is not enough to avoid an exponential increase of the resistance of Cu lines for small dimensions [4]. Novel materials such as Ruthenium and Cobalt have a worse resistivity than Cu, but they have the peculiarity to not diffuse nonetheless, therefore they do not need a barrier. In this way, they can overcome Cu performance for very small dimensions as shown in following graphs from [4]. In Fig.1.6a metal resistivity is plotted versus line width, here designated as critical dimension (CD). Cu shows outstanding behaviour even at very tight pitches, however when barrier is considered (Fig. 1.6b) the effective resistivity of the line shows an exponential increase starting from 22 nm of CD, while materials like Ruthenium and Cobalt maintains their values.



(a) (b)

Figure 1.6: a) Resistivity versus critical dimension of the line compared for different materials b) effective resistivity computed considering that copper needs a barrier that reduces the effective section of the line [4]

1.2.3 Semi-Damascene process

Alternative metal options are not only beneficial for lower effective resistivity, they enable an easier integration of semi-damascene (Semi-D) modules for first metal layers [9]. Semi-D is a flow process developed by IMEC to face BEOL scaling challenges [16][21]. It relies on creating via first and overfilling it with metal, enough to allow the creation of the next metal layer that will be masked and patterned. Copper patterning is not an easy task because of the lack of suitable wet etching solutions or dry etching gases with volatile by-products, this is why Cu was mainly grown by electro-plating. Semi-D is feasible only for metals that can be directly etched like Cobalt or Ruthenium [18]. A process flow based on Ruthenium is presented in Fig.1.7. Starting from a layer of Ruthenium and a top Hard Mask (Si_3N_4 for example) the first lithographic step is needed to transfer the layout of lines that are directly etched in Ruthenium (Fig.1.7a-1.7b) [21]. Trenches are filled with low-k material and an alternation of Si_3N_4 and SiO_2 is obtained on the inter-metal dielectric (IMD) (Fig.1.7c). Via opening can be performed by exploiting selective etching of Si_3N_4 and next via is overfilled with new layer of metal (Fig. 1.7d-1.7e).

Semi-D has multiple benefits. Higher AR lines are possible and capacitance increase can be wiped out by integrating air gaps (AG) or extended AG (next section) for ILD, instead of SiO_2 or low-k material. Moreover, new metallization options, as mentioned above, enables the possibility to pattern lines directly into the metal, unfeasible with Cu technology.

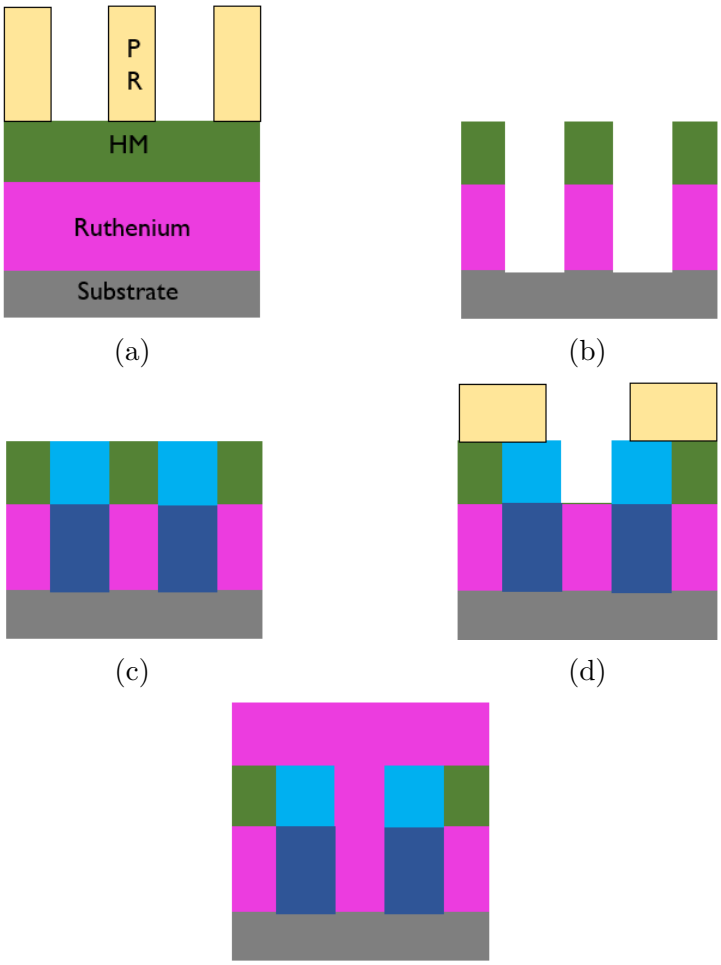
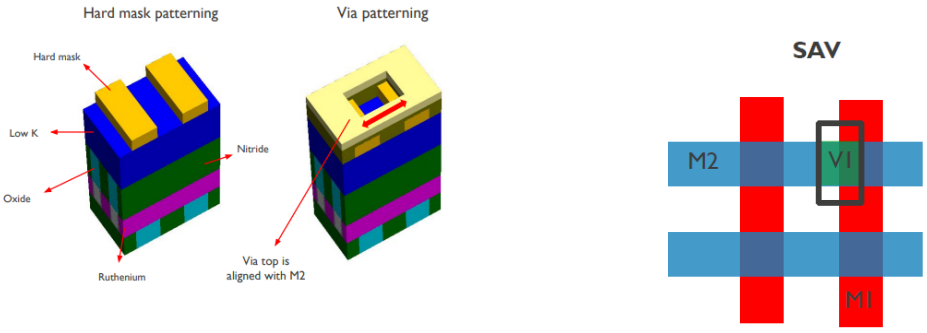


Figure 1.7: Semi-D simplified flow process steps: a)lithographic step on PR b) trench opening in hard mask and metal layer (actual lines are protected by hard mask) c) trenches are filled with low-k material (dark blue) and top of the trenches with oxide (light blue) d) lithographic step for via opening e) via filling and creation of new metal layer

1.2.4 Semi-Dual Damascene integration

In this section Dual-D and Semi-D modules are analyzed from via alignment point of view and a flow process for Semi-D/Dual-D integration will be shown. In the following bottom and top metal layers will be referred to M1 and M2 layers and the via to V1. Via alignment represents major concern at very tight pitches both in terms of performance [5] and reliability [14]. Patterning of small feature may have edge placement errors (EPE) with respect to the IC layout. The risks are: worsen

via-line contacts due to smaller contact area (or in worst case, open circuit) and shorter distance of via edges with neighbour lines that increases time-dependent dielectric breakdown (TDDB) failure.
Misalignment can be partially prevented, and sometimes totally avoided, by employing larger mask features or by exploiting selective etching. In a Dual-D module an SAV process (Self-aligned via) can be enabled by patterning firstly lines (and then vias) on the hard mask in order to force the alignment with M2 layer (Fig.1.8a). The width of via layout (Fig.1.12a) must be decided according to EPE of lithographic tool, so that worst case is still not critical, however SAV can not prevent misalignment across M1 direction. In Fig. 1.12a is shown the actual layout transferred to the substrate and a small misalignment across M1 will result in a smaller contact area, hence a higher contact resistance with bottom metal layer, and a shorter distance with neighbour lines as shown on the left in Fig. 1.9.



(a) Process steps (b) SAV layout

Figure 1.8: a) hard mask is patterned before via opening, via patterning is then constrained in one direction by hard mask lines spacing b) In SAV process via alignment is ensured along M2 but not along M1

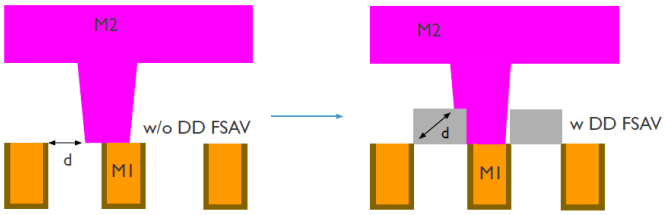


Figure 1.9: left: only SAV process may produce vias close to neighbour lines right: FSAV process ensures a good alignment for both top and bottom of vias

Dual-D module can integrate a fully self aligned via (FSAV) process by recess

etching of Cu. A flow process is shown in Fig. 1.10: via will be self-aligned with bottom M1 layer thanks to the sidewalls created by the recess.In case of misalignment there will be a lower contact area and a thinning of via, both can be prevented by an oversize of via width.

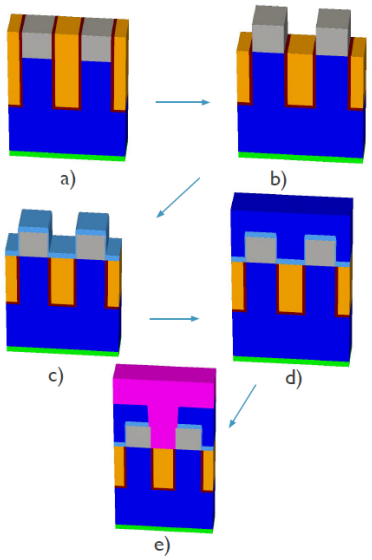
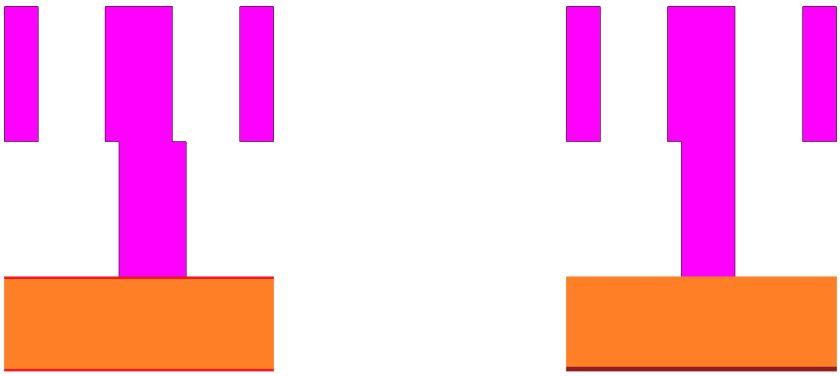


Figure 1.10: FSAV flow process: a) Copper lines are intentionally deposited with higher AR b) previously deposited dielectric shields low-k material (dark blue) from being etched during copper recess c) dielectric barrier is deposited d) low-k material is deposited to create via layer e) via is created by filling via trench with metal, bottom via is aligned with underneath layer

Semi-D is a process that may be thought as the 180° rotation of Dual-D. Misalignment is more critical with top layer, while for bottom one is self-aligned. The IMD layer is composed by an alternation of Si_3N_4 and Oxide that enables the possibility to open the via by selectively etching Si_3N_4 (see process flow in Fig.1.7). In Semi-D via alignment is critical for top metal layer because the via is patterned earlier than lines that are obtained only after having overfilled via and etched the top metal layer. In Fig.1.11 the output of a misaligned process is shown. Via is misaligned across M2 and it shares same problems of Dual-D misalignment, namely smaller contact area and shorter distance with neighbour lines. Oversized via can solve the former maintaining low resistance, however top via edge may approach even more neighbour lines. A good option is a recess etching during lines

patterning showed in Fig. 1.12b. Ruthenium is etched deeply up to via bottom in order to remove the unwanted edge protuberance. The over-etching step may generate a shrinking of the via in case of strong misalignment so an oversized via must be considered to minimize resistance increase.

Figure 1.11: 3D structure from SPX showing top via misalignment



(a) Cut along M1 (b) Recess

Figure 1.12: Via misalignment across M2

If Semi-D is employed in first metal layers, at a certain metal layer there might be the transition from Semi-D to Dual-D modules. The transition must ensure a good via alignment and fortunately it is inherently suitable for FSAV. In Fig.1.13 a simplified process flow shows main process steps. Bottom layer is a Semi-D

module with barrier-less Ruthenium as metal and top layer is a Dual-D module Cu-based with TaN barrier. The combination of a Semi-D and Dual-D inherits the SAV process of Dual-D for top metal layer and SAV process of Semi-D for bottom metal layer.

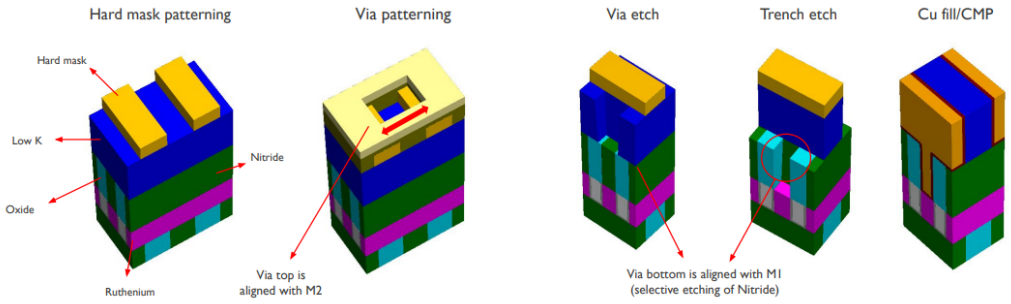


Figure 1.13: Dual-Semi integration: Semi-D module is the layer below, following steps comes from Dual-D process with the difference that via opening is guided by a selective etching of Silicon Nitride (green)

Chapter 2

Semi-Damascene scaling boosters

2.1 Introduction

Once the reasons for Semi-Damascene (Semi-D) modules integration have been clarified, this chapter focuses on the capacitance and performance benefits that may arise from exploiting air-gap (AG) schemes, instead of surrounding metal lines totally by standard dielectric materials.
Semi-D modules involve direct metal etch for lines patterning and it allows to fill inter-metal-dielectric (IMD) trenches with different materials. Low-k options can strongly reduce parasitic capacitances, however a careful reliability analysis must be performed to ensure reasonable operating lifetime [17][3]. Recent nodes rely on interconnects with very tight pitches lower than 20 nm and spacing between lines can reach few nm. Unfortunately spacing is not scaling in the same way of line width, since a lower resistance is generally preferred to control power consumption and heat generation lines are kept as wider as possible, therefore it causes a large increase of capacitance that directly affects RC delay. A possible solution is represented by AG-schemes in which low-k material is substituted by air whose k value close to unity can reduce considerably capacitance. AGs were born during Aluminum technology era but actually they were not intentional. PECVD is the typical technique used to perform depositions at low temperatures and it is difficult to fill trenches with dielectrics because it has high deposition rates and a poor conformity. Moreover the AR of trenches can influence the deposition that will be slower at the bottom for high AR structures and faster at the top causing voids formation. This deposition issue can be actually exploited to get AGs in metal layers where deposition is difficult or in presence of high AR trenches. AG technology is very suitable for Semi-D modules where trenches are formed directly

in the metal and then generally filled with a low-k material, while for Cu Dual-D modules the formation of air cavities becomes quite challenging and it requires AG masks to etch selectively and directionally IMD layers or sacrificial materials as thermal degradable polymers [8] In this chapter AG integration is considered as scaling booster of advanced interconnects and its benefits are evaluated.

2.2 Topography approximation

First of all, AG is analyzed from structure point-of-view. A non-conformal deposition is used to get air cavities and to simplify the 3D modeling some approximations on the topography are necessary. IMEC demonstrated AG for Ru Semi-D module for pitch 32 nm and a TEM picture is shown in Fig. 2.1. In real structures AG extends into the top ILD with a round and pointing shape. This result is actually better from capacitance point of view so it was decided to approximate it with a rectangular profile (Fig. 2.1).

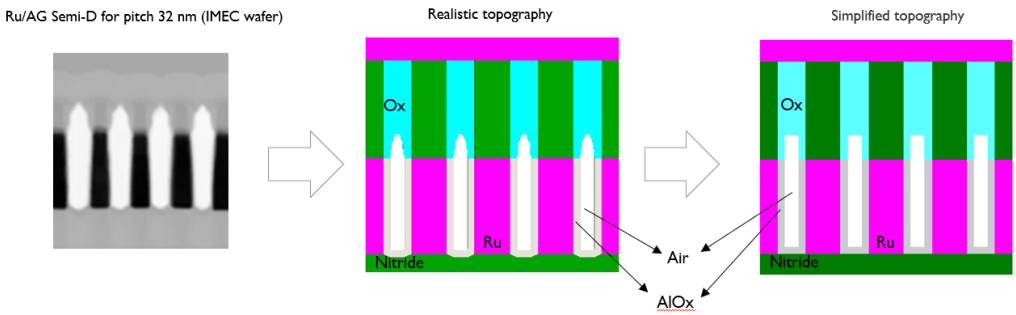


Figure 2.1: From left to right: TEM picture of AG Ru/Semi-D module (IMEC wafer), realistic modeling of AG topography and rectangular approximation

2.3 Capacitance model

Capacitance estimation can be performed using Raphael by Synopsys. The capacitance is highly affected by ILD and IMD materials, this is why low-k or ultra low-k materials are considered. The capacitance of a line is given by multiple contributions defined in Eqs.2.1 and 2.2.

$$C_{in-plane} = C_{Left} + C_{Right} \qquad C_{out-of-plane} = C_{Top} + C_{Bot} \tag{2.1}$$

$$C_{Tot} = C_{in-plane} + C_{out-of-plane} \tag{2.2}$$

where $C_{in-plane}$ is the sum of capacitances given by adjacent metal lines and $C_{out-of-plane}$ is the sum of top and bottom metal layer contributions. The former is depending on spacing between lines and on IMD and the latter depends mainly on via AR/height and ILD. So in order to evaluate the capacitance of one line it is necessary to generate a 3D structure involving three metal layers (Fig. 2.2), therefore it is quite demanding in terms of simulation time.

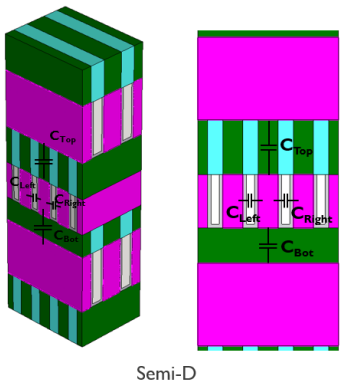


Figure 2.2: Capacitance model on 3D Semi-D module

2.4 Critical Paths

AG scheme is a valid option to minimize capacitance increase, however its integration is still quite challenging and it can not be adopted for every metal layer due to chip-package-interaction (CPI) issues [26], therefore its integration may be limited to one or two non-consecutive layers. In this work [**?**] an ARM core design 64-bit has been simulated to recognize signal paths with negative slack,namely, the required time for a signal to travel through the path is lower than the actual time that it takes. These paths are designated as Critical Paths (CPs): they are grouped for path length and their metal distribution is shown in Fig. 2.3. This analysis reveals that M2 layer is the most used one in short CPs and it has still a good portion for long CPs. It makes this metal layer the perfect candidate for AG integration. As a consequence, in the following, only M2 layer (with bottom and top layers) will be considered for AG integration whose dimensions are listed in Tab. 2.1.

Table 2.1: M_{int} to M3 dimensions

Assumptions						
Layer	Pitch[nm]	Metal	CD[nm]	Spacing[nm]	AR_{Line}	AR_{Via}
M_{int}	18	Ru	10	8	3	2
M1	28	Ru	16	12	3	2
M2	18	Ru	10	8	3	2
M3	28	Ru	16	12	3	2

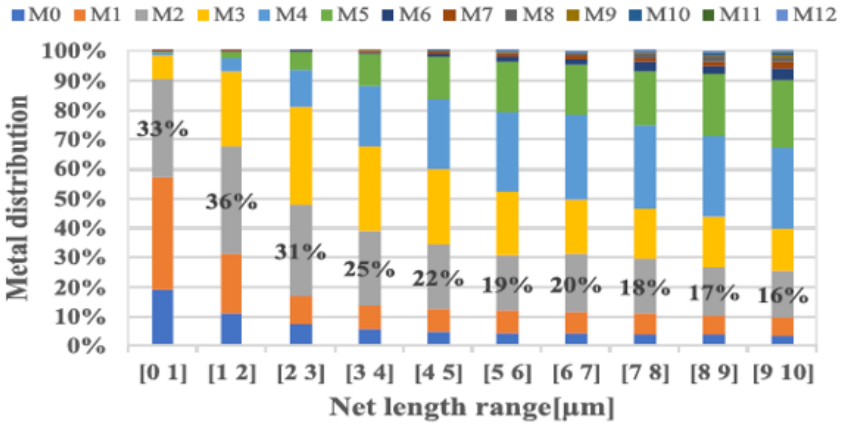


Figure 2.3: Metal layer distribution of CPs grouped for net length.

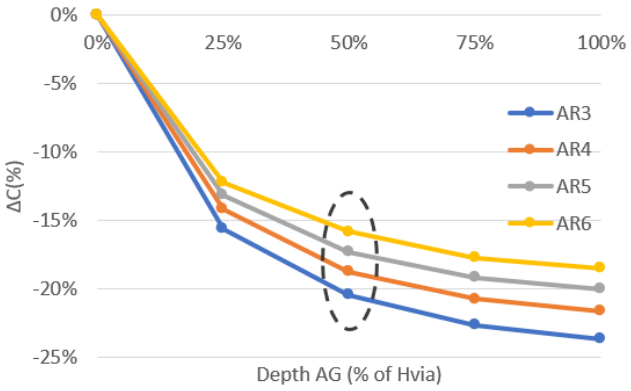
2.5 Ext AG

AG extension may be a future option to enable a further capacitance reduction. The AG extends into the underlying ILD where vias are located. The presence of air, that could be considered as an ultra-low-k dielectric material, reduces considerably the effective dielectric constant of ILD [19].

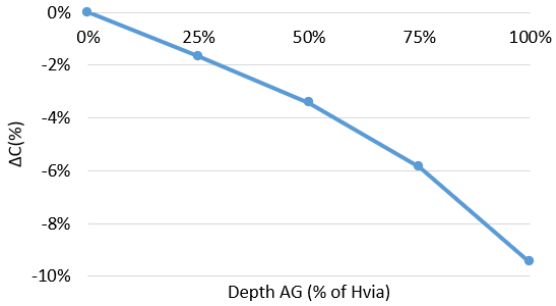
2.5.1 Depth vs AR

Depth of AG influences capacitance variation because, not only in-plane capacitance,also out-of-plane capacitance is affected. Different extensions are considered in fractions of via height, ranging from 0% (the reference AG) to 100% (full extension into bottom layer). In Fig. 2.4a, results show a remarkable decrease of capacitance, 20% for AR3 at 50% of extension. Higher ARs are less sensible to Ext AG because in-plane capacitance is dominating and the extension is very effective on boundary effects, since, due to the presence of fringing fields (included

in the calculation), the capacitance of two plates is higher than a simple parallel plate capacitor model.This results also in lower benefits for extensions over 50% of via height, however greater extensions can affect underlying metal layer almost by 10% (Fig.2.4b).



(a)

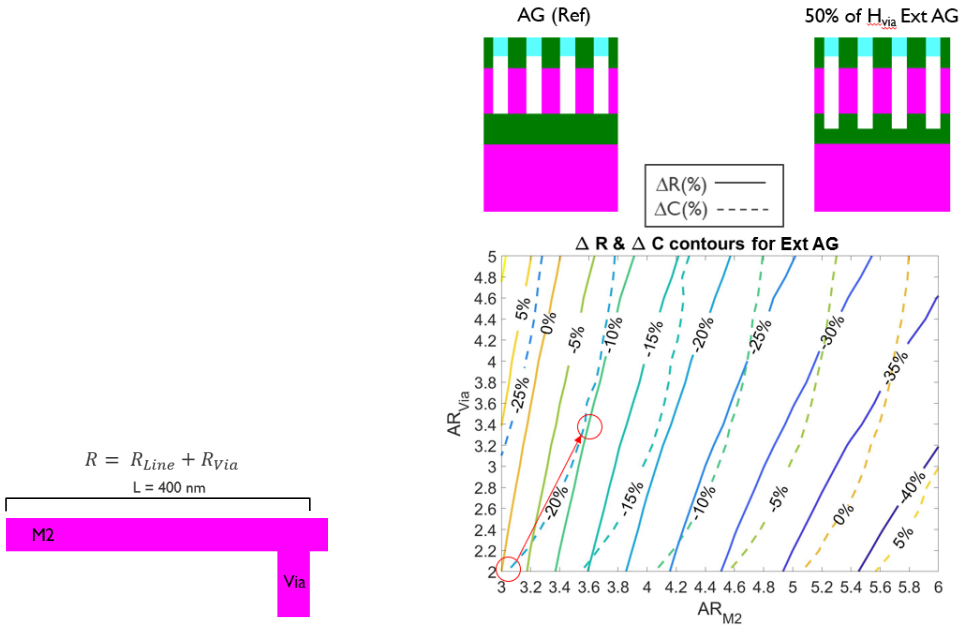


(b)

Figure 2.4: AG extension impact for a) middle layer (M2) and for b) bottom layer (M1)

2.6 RC trade-offs

Interconnect scaling forces lines to be smaller and closer to each other, increasing both resistance and capacitance. Generally capacitance should remain constant during scaling because, assuming a constant line length, capacitance would be dependent on height and distance between neighbour lines only and in a constant scaling approach they both scale in the same way. However lines do not scale as much as spacing does to reduce R increase, hence a price in C must be paid. Scaling boosters as AG extension may be considered to minimize capacitance increase and simultaneously reduce resistance. Here RC simulations are carried out for different AR of the line and underlying via. M2 line is considered with an average length of 400 nm and the total resistance includes both line and via resistance (Fig.2.5a). In Fig. 2.5b, continuous lines and dashed lines describe resistance and capacitance variations of Ext AG option with respect to a standard AG. The bottom extension is fixed to 50% of via height and increasing via AR it can allows to minimize the increase of capacitance of high AR lines. A good RC trade-off can be found by moving along a specific ΔR line to find the intersection where the ΔC is minimized.



(a) (b)

Figure 2.5: a) Total resistance is composed by line and via contribution b) Contour plot for resistance and capacitance variation with via and line aspect-ratios

2.6.1 RC-delay for CPs

RC-delay estimation can be performed by using ring oscillator analysis based on Elmore delay model, it allows to reduce the overall circuit to an RC network where the BEOL-stack participates as a load resistance and capacitance.

RC delay model

To evaluate signal delay of a logic block the ring oscillator (RO) configuration is often used. It relies on a series of N-inverter stages in which the last output is fed back into the first stage (Fig. 2.6).

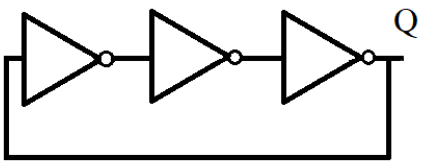


Figure 2.6: 3-stages ring oscillator composed by three inverters port, the output Q is fed back to the input to close the ring and to generate the output oscillation

Using an odd number of stages allows to obtain an oscillating signal whose period is determined by the delay of every stage. According to device we are using and to the measurement equipment the number of stages is decided in order to obtain a reasonable frequency that belongs to our measurable bandwidth (BW). Nowadays, devices are such fast to require several stages to be able to measure their switching period properly. The total period of the signal can be thought as twice the number of stages times the delay of every stage. The factor two takes into account that we have to wait for two cycles of the ring oscillator to obtain a full wave. The circuit can be modelled as an RC network (see Fig. 2.7) where we take into account the internal resistance/capacitance of every stage and the presence of a load that in our case may be BEOL interconnections. This scenario allows us to apply Elmore delay model which states that the net delay of an RC network is the summation on products of capacitance and subsequent resistance of every stage [20]. In the current scenario the delay of every stage is defined as:

$$\tau = C_{int}R_{int} + (R_{int} + R_L)(C_L + FO * C_{pin}) \tag{2.3}$$

where subscripts 'int' and 'load' refer respectively to intrinsic device parameters and BEOL, while C_{pin} is a parasitic capacitance due to the connection with the next stage and FO refers to fan-out, the number of stages connected at the output

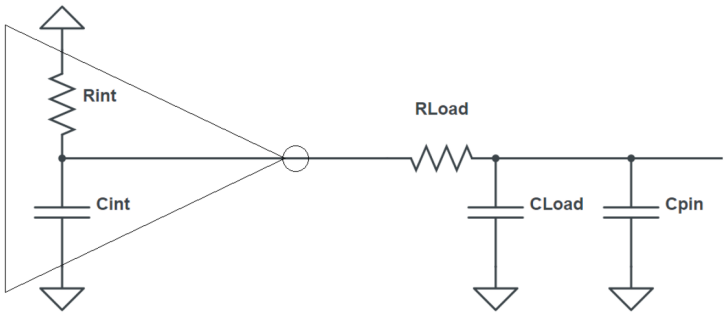


Figure 2.7: RC network of 1 stage

at every stage. This model is very powerful to predict interconnections delay and it can be used to evaluate delay reduction on CPs that include AG-scheme in the most used metal layer, namely M2 [1]. Referring to Fig. 2.8, for $AR = 3$ of the line, frequency boost is around 4%, while AR = 6 lines shows a frequency increase of 10%, primarily to the reduced resistance and secondly thanks to the minimization of capacitance increase due to taller lines.

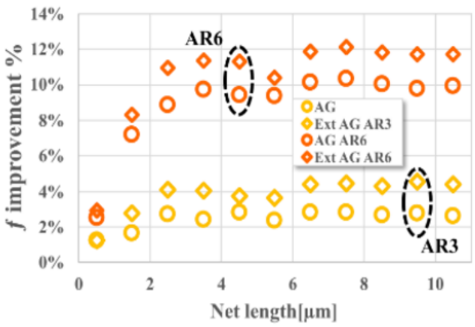


Figure 2.8: Frequency improvements in CPs for different net lengths [1]

2.7 Extended AG integration

AG or Ext AG formation is actually very challenging. Besides the advantages for performances and for TDDB-related reliability, there are many concerns as well that will be discussed in this section. Firstly, the presence of air cavities may have a strong impact on structure stability. In this study [26] AG-scheme integration is analyzed in terms of Chip-Package-Integration (CPI). BEOL stack must withstand mechanical/thermal stresses that may be induced by packaging

processes. The presence of shear stresses is highly risky and may cause line bending and delamination of dielectric layers. Secondly via formation is an important obstacle that must be taken into account. In Fig.2.9, Semi-D process flow shows a simplified AG formation in which via opening may lead to obtain an open cavity that, with subsequent via filling, might be filled as well forming short circuits among lines.

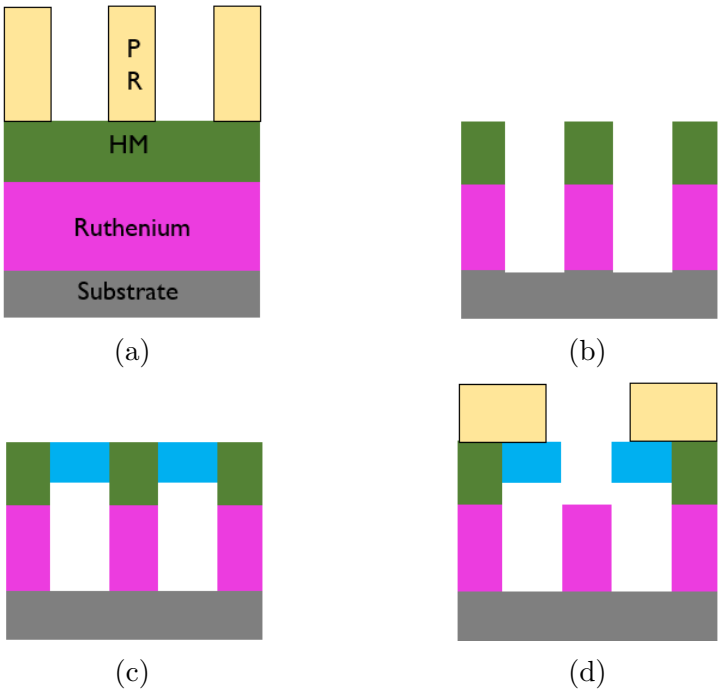


Figure 2.9: Ru/Semi-D AG via formation. After step d) via filling is the critical step, because the via is opened on both sides.

Etch-back (EB)

One possible solution is to deposit the oxide deeper in the trenches by properly controlling deposition parameters [12], hence via formation is preserved, however capacitance increases due to in-plane contribution. Since AG may be obtained by using a temporary sacrificial layer filling the trenches, it could be recessed and oxide would be deposited on the top. Finally the sacrificial is removed and the structure obtained is shown in Fig.2.10. For capacitance simulations the recess is considered to be 5 nm deep.

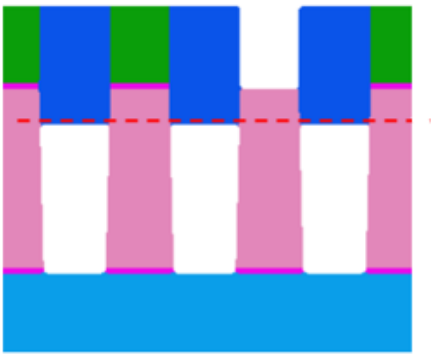


Figure 2.10: Etch-back solution

Exclusion zone (EZ)

Another solution is to employ exclusion/keep-out zones. In this case an extra lithographic step is needed to define exclusion zones in correspondence of vias' locations. In Fig. 2.11, three parallel lines include AG-scheme only in the region where via is not present, while around the via a keep-out zone twice the pitch wide is chosen. A gap-fill with a dielectric (low-k for example) may be carried out to avoid metal "spilling" during via filling.

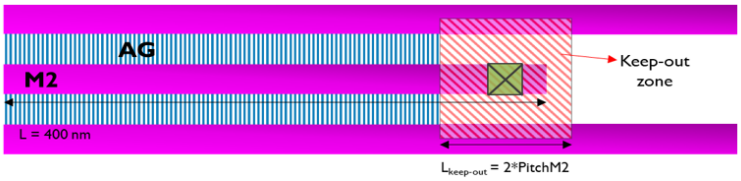


Figure 2.11: Caption

EB is less expensive because it does not need an extra lithographic step. To compare them in term of capacitance (Fig. 2.12), the EB solution is chosen as the reference case for different line lengths, namely from 100 nm to 1000 nm. Assuming that the signal enters the line and it exits only once, two vias are considered for every line. In case of EZ solution a penalty in capacitance is present for very short lines, since a big fraction of the line is surrounded by a low-k material (here considered with k = 3), however for long lines this penalty vanishes while EB option maintains its value constant independently from line length. A combination of them could ensure lower capacitance even for very short lines.

In conclusion, AG-schemes have shown optimal results: they allow to increase lines aspect-ratio to reduce considerably line resistance minimizing capacitance

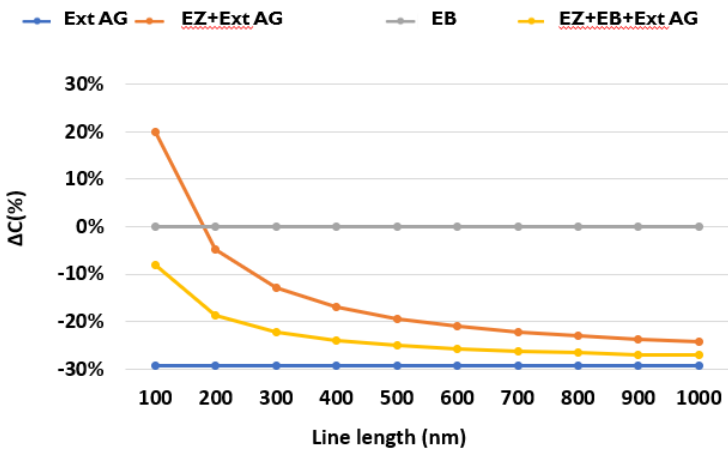


Figure 2.12: Caption

penalty. Moreover RC-delay may be reduced by implementing them in only one layer, hence the integration on multiple layers is expected to be even more beneficial taking into account that etch-back or exclusion zone techniques must be foreseen from the process to preserve via formation. Challenges still remain for chip-package-interaction issues and further studies are needed to make AG-scheme less sensible to shear stresses.

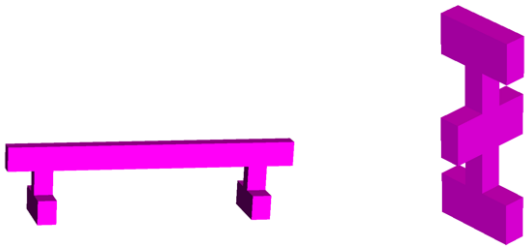
Chapter 3

Power Delivery Network resistance breakdown

3.1 Introduction

In previous chapter Semi-Damascene modules and AG-schemes were analyzed to estimate possible improvements in signal lines. In this final part another side of BEOL is under the spotlight, namely power delivery networks (PDN). As the name suggests, PDN's task is to provide clean power to the circuit and a reference voltage for all electronic elements. The increase of metal layers in microchips is making PDN design more complex due to higher parasitic components (resistive and capacitive/inductive). Design of PDNs aims to reduce and minimize power loss and to guarantee a stable power delivery. In the first part of this chapter a brief review on electromigration phenomenon theory is presented, since it is the main cause of PDN failures. Secondly, a new physics-based model from literature deals with the complexity of determining chip failure probability and IR-drop estimation after many years of device operation. Finally the actual contribution of this analysis provide an evaluation of IR-drop by matching current and resistance extractions.
PDNs are independent networks with respect to signal ones and they develop mostly vertically in the BEOL. For signal lines, impact of via resistance is generally limited, because the line resistance dominates the total net resistance. For PDNs, many lines are just staples, where the current flows vertically, as a consequence via resistance is expected to be more critical (for IR-drop) than for signal lines (a 3D comparison of signal line and PDN is shown in Fig. 3.1).
PDN resistance extraction required an optimization at code level to reduce simulation time and finally real paths of the current are analyzed and compared to evaluate IR-drop and to find common contributions, if there are any. This analysis

aims to demonstrate the importance of current inclusion in design rules.



(a) (b)

Figure 3.1: Example of signal line and PDN with two vias (obtained with SPX)

3.2 Electromigration

3.2.1 Statistical evaluation

The phenomenon of electromigration is the main enemy of IR-drop and with interconnect scaling current density becomes larger and consequently also self-heating. EM must be taken into account during circuit design to guarantee a long operating lifetime. The complexity of integrating the phenomenon in circuit modelling is related to the difficulty of evaluating chip level EM failure probability. EM data are generally obtained for a specific circuit element, a wire or a small structure so critical temperature T and maximum current density j data are extracted for a specific element. Statistically the time-to-failure (TTF) of an element follows an Arrhenius-like trend, known as Black's equation [2]:

$$TTF = \frac{A}{J^2} e^{\frac{-E_a}{kT}} \tag{3.1}$$

where A depends on geometrical parameters, k is the Boltzmann's constant and E_a is the EM activation energy. This latter may assume different values according to the source of EM diffusion. Generally EM is more likely to happen along grain boundaries, in so called triple points where three grains meet, or it may take place at the surface of the conductor or in the bulk. Black's distribution can be used to determine important parameters as max j and critical T, useful for circuit design guidelines of single elements, but less effective for chip design. One

possible approach is described in [13] where, once the failure probability of the i-th element F_i is known, using weakest link statistics, the chip failure probability can be estimated as:

$$F_{Chip} = 1 - \prod_{i=1}^{K}(1 - F_i(t)) \tag{3.2}$$

The chip level probability failure F_{Chip} is actually hard to be obtained since estimating with enough precision the TTF of each element composing the circuit might be rather audacious. Every element has its own geometry and besides that, circuit redundancy and different configurations have to be taken into account. One possibility may be to consider only elements that carry a current higher than maximum limits, however at chip level it might be possible that many elements carrying a lower current may induce a failure. A possible way to simplify (3.2) is to consider equivalent critical EM elements. These latter are normalized elements in terms of j_{max} and of EM failure probability. One critical element may be one element carrying current j_{max} or a group of elements whose failure probability can be associated to one element carrying j_{max}, therefore the chip failure probability can be written as:

$$F_{Chip} = 1 - \prod_{i=1}^{N}(1 - F_n) = 1 - (1 - F_n)^N \tag{3.3}$$

where N is a number that is not strictly an integer because it depends on the normalization factors and F_n is the failure probability decided by design specifications. As low F_n is required as low the number N of equivalent EM elements will be allowed for that specific design.

3.2.2 Physics-based model for PDN

In PDNs, electromigration is very relevant because the current is unidirectional and generally high current densities flow inside each element. The aforementioned statistical mode considers mainly current densities to determine the chip failure probability, however the location of the failure is not determined yet and it is not sure to occur where maximum current densities are detected, but instead in mechanical stress maxima that do not always coincide [25]. To develop a physics-based model, it is necessary to start from theory of thermodynamics of irreversible processes. EM is associated to the three different fluxes J_e, J_m and J_u that are respectively the electron flux, the ion flux and the energy flux, each one induced by three forces X_j with $j = e, m, u$. For electrons and metal ions:

$$X_{e,m} = -\nabla^{e,m}\mu_{ec} \tag{3.4}$$

where μ_{ec} is the electrochemical potential associated to the specific charge Z = +1 for ions and Z = -1 for electrons, and it is defined as:

$$\mu_{ec} = \mu + Ze\psi \tag{3.5}$$

with μ and ψ, the chemical potential and electrostatic potential respectively. The force related to the energy flux is related to a T gradient, nonetheless all fluxes depend on each force and the system can be described with general equations and phenomenological constants $L_{i,j}$ with $i, j = e, m, u$:

$$\begin{cases} J_m = -L_{m,m}\nabla(\frac{\mu^m_{ec}}{T}) - L_{m,e}\nabla(\frac{\mu^e_{ec}}{T}) - L_{m,u}(\frac{\nabla T}{T^2}) \\ J_e = -L_{e,m}\nabla(\frac{\mu^m_{ec}}{T}) - L_{e,e}\nabla(\frac{\mu^e_{ec}}{T}) - L_{e,u}(\frac{\nabla T}{T^2}) \\ J_u = -L_{u,m}\nabla(\frac{\mu^m_{ec}}{T}) - L_{u,e}\nabla(\frac{\mu^e_{ec}}{T}) - L_{u,u}(\frac{\nabla T}{T^2}) \end{cases} \tag{3.6}$$

Simplifications are needed to solve the system: 1) in high conductivity materials the gradient of μ^e_{ec} can be ignored 2) thermodiffusion is neglected 3) in a metal wire the flux of charges is due to electrons only. Recalling that the negative gradient of the electrostatic potential ψ is the electric field E and by using microscopic Ohm's law (3.8) the flux of metal ions J_m can be simplified to this expression (3.9).

$$-\nabla\psi = E \tag{3.7}$$

$$E = \rho J \tag{3.8}$$

$$J_m = -L_{m,m}\Big(\frac{\nabla^m\mu - Z^*e\rho J}{T}\Big) \tag{3.9}$$

where ρ is the resistivity and Z^* is an effective charge of ion atoms that can be defined as the real charge less a corrective term (3.10) (due to transfer of momentum from electrons to ions) that experimentally turns out to be greater than Z, therefore the effective charge is negative and it confirms that the net force acting on atoms is in the same direction of electron flow.

$$Z^* = Z - \frac{L_{m,e}}{L_{m,m}} \qquad with \qquad \frac{L_{m,e}}{L_{m,m}} \gg Z \tag{3.10}$$

Finally the gradient of the chemical potential is related to the variation of the so called hydrostatic stress along the line and the atomic volume Ω, so that flux of metal ions can be recasted into (3.11).

$$J_m = -\frac{D_a C_a}{k_B T}\left(\Omega \frac{d\sigma}{dx} - Z^* e \rho J\right) \tag{3.11}$$

where the phenomenological constant is substituted by the atomic diffusivity D_a and atomic concentration C_a over the Boltzmann's constant. In (3.9) and (3.11), atomic flux confirms the presence of two different forces acting on atoms: the wind force due to electrons moving in the electric field and a stress-related force that is counter-acting to EM diffusion. This model relies on the assumption that stress develops mainly in one direction opposite to electron flow and it is known as Korhonen type 1D model [11]. Under this assumption also the stress time-variation can be extracted:

$$\frac{d\sigma}{dt} = -\frac{d}{dx}\left[\frac{D_a B \Omega}{k_B T}\left(\frac{Z^* e \rho J}{\Omega} - \frac{d\sigma}{dx}\right)\right] \tag{3.12}$$

In this work [25] a PDN of 3nm technology node is considered (schematic shown in Fig.3.2).

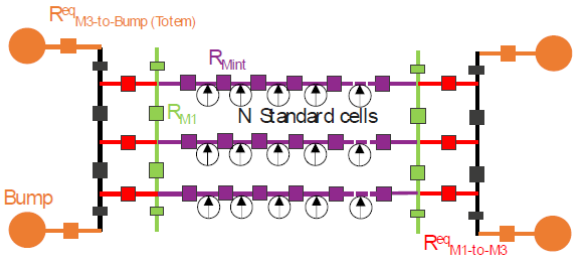


Figure 3.2: Schematic of PDN before voiding.From M1 to M3 and from M3 to M12 equivalent resistances are considered. Source: [25]

The network is subdivided in electromigration unit cells in which the equations (3.11) and (3.12) can be solved and where vacancies can flow freely. Vacancies are stopped by metal barriers that surround Cu in Dual-D modules, so line and

bottom via can be considered as one unit cell. Since two counter forces are acting, a maximum stress σ_{crit} can be defined as the value over which the void formation begins. This value depends on the line features and processing and its distribution can be associated to a log-normal distribution [27]. Values are randomly associated to every EM unit cell and time simulations are performed.In Fig. 3.3 after one year three failures occured, later a steady-state is reached and no failure occured for 10 years.

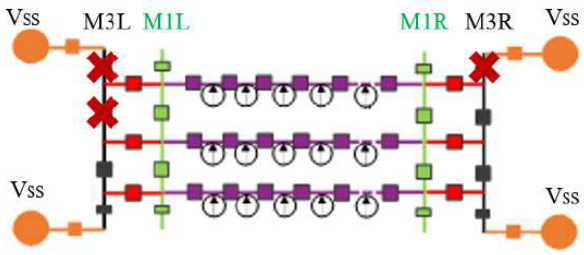


Figure 3.3: Schematic of PDN after voiding and steady-state reached. Source: [25]

An important parameter to evaluate the reliability of a PDN is the IR-drop and in this analysis thanks to the presence of rich circuit redundancy the % of IR-drop impacting on the standard cells is only around 1% after 10 years of operation. Generally to ensure a good device operation IR-drop should remain lower than 10% of V_{dd}. Technological parameters are strictly architecture-specific, however this analysis showed the power of a physics-based model based on hydrostatic stress evaluation, instead of considering only current densities. Moreover, a careful design of circuit redundancy can ensure an high fault-tolerance, although single interconnect failure probability is relatively high. Therefore chip failure probability can not be merely computed considering weakest link statistics as shown in previous section. A careful design, hence, can protect from IR-drop worsening, however an intrinsic IR-drop is due to voltage drop on resistive components and it can not be avoided totally. In the following, the resistance is extracted and analyzed to evaluate the inevitable IR-drop owned by PDNs.

3.3 Resistivity model

Resistance is extracted by using Raphael FX, a 3D simulator by Synopsis. The resistivity model coming from [5] is implemented. It relies on considering the contributions of bulk resistivity and surface resistivity. When electrons are approaching the interface of the metal with a metal barrier or a dielectric they experience surface scattering that decrease the local resistivity. In this model the resistivity is function of the nearest distance from the metal or dielectric interface:

$$\rho(d) = \rho_b + \rho_q \cdot e^{-\frac{d}{\lambda_q}} \tag{3.13}$$

where the first term refers to bulk resistivity and it depends on grain boundaries location and orientation while the second term is the product between the maximum "extra" resistivity given by mobility reduction at interface (so when $d = 0$) and λ_q defines the exponential decay when moving away from the interfaces. These three terms have been obtained experimentally for Copper and Ruthenium in Dual-D and Semi-D modules.
A visualization of local resistivity is given in Fig. 3.4 taken from [5]. Maxima are localized at interface of metal with metal barrier or dielectric cap and minima in the farthest regions from interfaces.

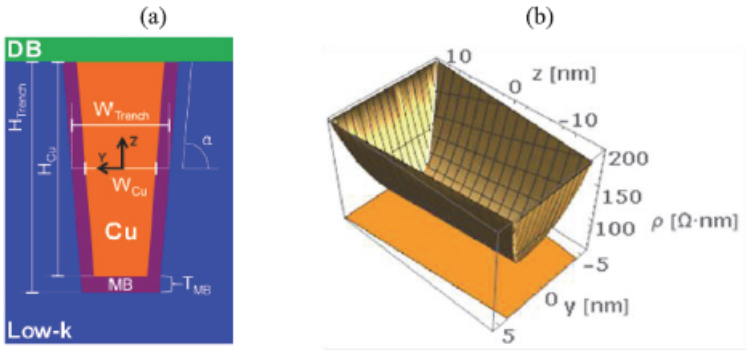


Figure 3.4: (a) Cross section of Cu via in Dual-D module and (b) local resistivity function of nearest distance. Source: [5]

3.4 Via/Line impact

A first rough estimation of via impact may be done considering the resistance contribution in a signal line with an average length and varying the number of vias with respect to a PDN (dimensions assumptions are in Tab.3.5). Signal line and PDN are simulated with different number of vias (from 2 to 4) with different metallization options and results are shown in Fig. 3.5.

Table 3.1: Via and line dimensions

Assumptions	
Pitch [nm]	21
CD [nm]	11
α_{Via}[°]	90
L [um]	1.697

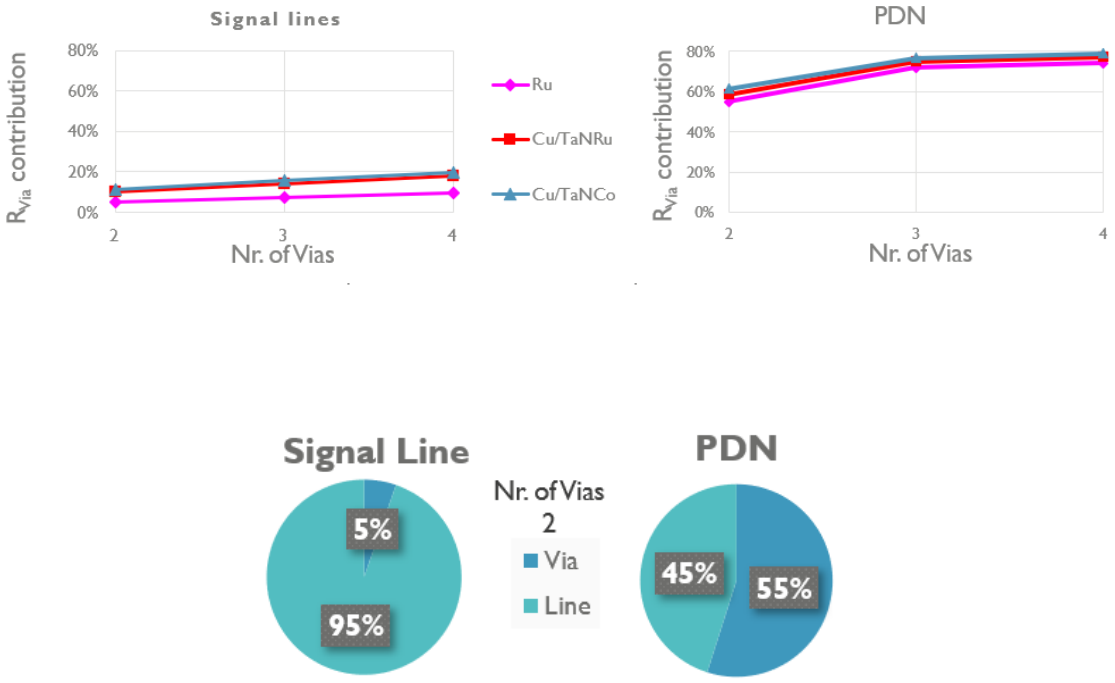


Figure 3.5: Via contribution for different number of vias (top) via contribution in Ru modules with two vias (bottom)

3.5 Approach

First of all, the full structure is generated by Sentaurus Process Explorer (SPX) tool and then Raphael applies the resistivity model to evaluate the total resistance. Two different approaches are possible to extract the total PDN resistance. In the former, the full structure (generally called PDN "Totem") can be totally generated and contacts can be placed at the beginning and end of each PDN path. SPX and RFX simulation time is strictly correlated to the dimension of the structure. In this case the simulation region that RFX has to manage is on the order of μm. In Fig. 3.6 and 3.7 the full structure of a path and simulation time are shown. A

second approach (Fig. 3.8-3.9) may be used to drastically reduce simulation time. PDN is divided in sub-elements, composed by two metal layers and via connection. The total resistance can be extracted by placing contacts on bottom via and on the top of the line where the next via should be placed. RFX does not allow to place contacts at via bottom if the material is the same so it is necessary to use another material (here Air has been used) to obtain an interface on which contact can be placed (see Fig.3.10). Via resistance can be extracted directly by placing contacts on via top and via bottom (Fig. 3.11) so the line resistance (here designated as vertical line resistance) can be estimated by taking the difference between the two structures. This second approach allows to run multiple processes in parallel reducing simulation time by 75% for RFX and by 20% for SPX.

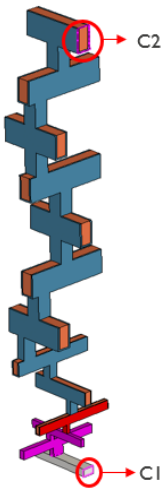


Figure 3.6: 1^{st} sol. (Totem)

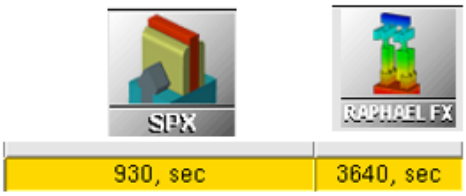


Figure 3.7: Simulation time

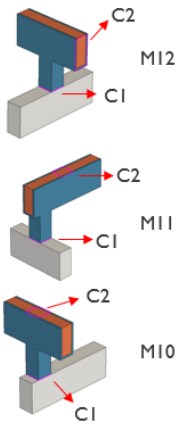


Figure 3.8: 2^{nd} sol.

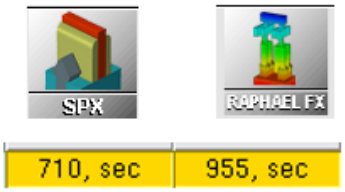


Figure 3.9: Simulation time

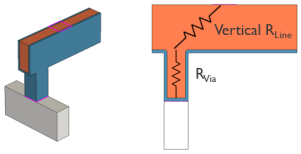


Figure 3.10: via-line extr

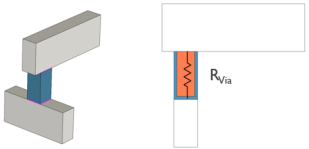


Figure 3.11: Direct Via extraction

3.6 Case study

In order to estimate the actual via contribution three least resistive paths under V_{dd} bumps of a real PDN design are considered. Their layouts are shown in Fig. 3.12 and total length/standard cell (SDC) distance from totem are specified in Tab. 3.2 and Tab. 3.3 respectively. A 3D representation of Path#1 is also shown in Fig. 3.14.

All paths are decomposed in sub-components, whose dimensions come from iN5 assumptions in Tab.3.4 and for each one R is extracted. As it may be expected, via and line contribution is highly dependent on lines length, for short M_{int} (that it means SDC is close to VDD bump) vias are main responsible for total resistance (see Fig.3.13).

Table 3.2: Paths length.

Length [um]	
Path#1	2.55
Path#2	2.11
Path#3	2.35

Table 3.3: M_{int} length

M_{int} [um]	
Path#1	1.01
Path#2	0.63
Path#3	0.47

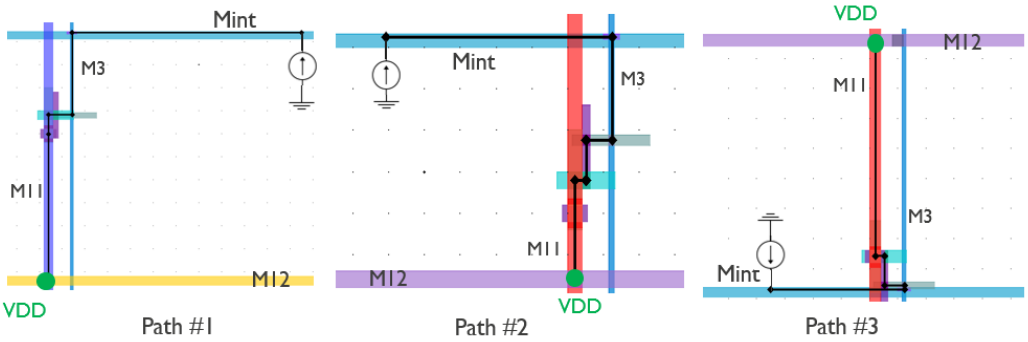


Figure 3.12: Layouts

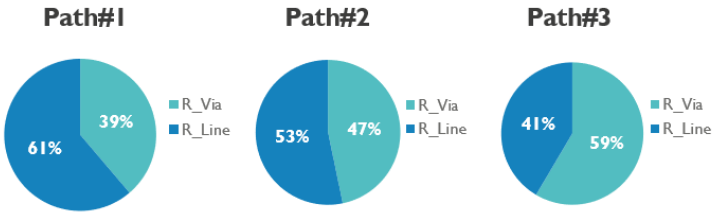


Figure 3.13: Via-Line resistance contribution

3.7 Resistance breakdown

A resistance breakdown for every element of the network may highlight main contributions (for dimensions refers to Tab. 3.4). In Fig. 3.16 the impact of every component can be compared. Interestingly paths have common contributions, namely Mint/M3 layers and V2/V3. The third path has a low M3 contribution because SDC is very close to VDD bump. Considering that lines are mostly staples it is reasonable that long lines (where current mainly flows) become most resistive sections of PDNs. M1 and M3 lines are two parallel paths (as it can be seen from layouts in Fig. 3.12 or in Fig. 3.15) that could be both used by the current. M3 is the most used due to a lower resistance, however the presence of M1

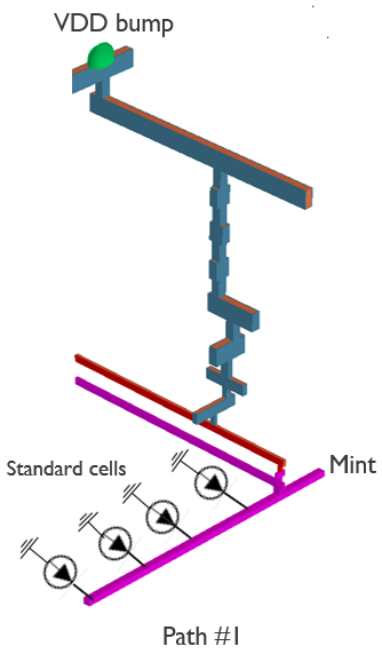


Figure 3.14: Path#1

establishes a circuit redundancy that can be helpful in case of electro-migration (EM) phenomenon. The EM is the movement of atoms in metal lines due to current flowing. It is more likely to happen close to the surface of lines or in correspondence of grain boundaries and it creates voids or hillocks that cause an increase of resistance and in worst case, since new defects may also move, open or short circuits. This effect was one of the reasons why the employment of Aluminum was abandoned in recent technology nodes and it is addressed in following sections.

Figure 3.15: In given paths, current flows in M3 to reach upper layers, however M1 might be an alternative option in case of M3 failure.

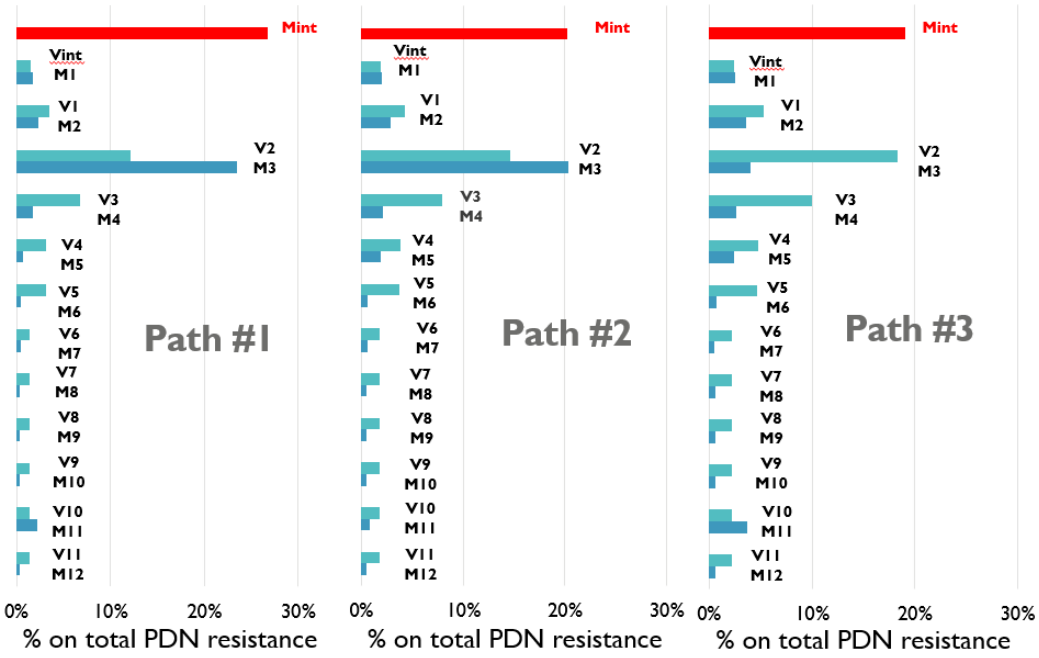


Figure 3.16: Resistance breakdown of $Path\#1$, $Path\#2$ and $Path\#3$

Table 3.4: Dimensions assumptions for every metal layer.

Assumptions							
Layer	Pitch [nm]	CD_{Line} [nm]	CD_{Via} [nm]	H_{Via} [nm]	AR_{Line}	Metal	TMB [nm]
M12/V11	80	40	40	80	2	Cu	5
M11/V10	80	40	40	80	2	Cu	5
M10/V9	80	40	40	80	2	Cu	5
M9/V8	80	40	40	80	2	Cu	5
M8/V7	80	40	40	80	2	Cu	5
M7/V6	80	40	40	80	2	Cu	5
M6/V5	80	40	24	80	2	Cu	5
M5/V4	48	24	24	48	2	Cu	3.5
M4/V3	48	24	15	48	2	Cu	3.5
M3/V2	30	15	11	30	2	Cu	2.5
M2/V1	21	11	15	22	2.5	Ru	0
M1/V0	30	15	11	30	2	Ru	0
Mint	21	32			2.5	Ru	0

3.8 IR-drop evaluation

To estimate an average IR-drop of every element, the current (extracted from *Voltus* and listed in Tab. 3.4) becomes the weighting factor and it is multiplied by the resistance of each component to obtain the voltage drop in every section of the network. Results are summarized in Fig. 3.17 in which resistance and IR-drop breakdowns are compared to evaluate the actual impact of every element. Surprisingly M_{int}, that has got highest resistance, is not impacting equally on IR-drop since a very low current flows through. M3 line confirmed to be the main contributor to voltage drop, except for $Path\#3$ in which M3 line is so short that V2 and V3 gain the first place for IR-drop. Finally, M12/V11 ,whose resistance contribution was summed up because very small, showed an unexpected impact on the total IR-drop around 10% due to the presence of a large current.

Assuming a V_{dd} of 0.7 V, the actual value reaching the cell is reduced by less than 5 %, see Tab. 3.6. The analysis was carried out for paths under V_{dd} bumps, so the path on M12 was actually reduced to the minimum and, since M12 turned out to affect appreciably the total IR-drop, this study does not take into account the worst case of SDC (very far from power bumps), but it highlights main contributions in common paths. Far SDCs may experience a larger IR-drop due to M12 and M3, because these two lines are the only allowed lines to reach SDCs (apart from M1

Table 3.5: Current extracted from Voltus software

$Current[\mu A]$			
Element	$Path\#1$	$Path\#2$	$Path\#3$
V11/M12	245,1	340,2	315,8
M11	102,6	132,9	99,9
V10	95,8	168,1	134,6
V9/M10	95,8	169,5	135,4
V8/M9	95,8	169,5	135,4
V7/M8	95,8	169,5	135,4
M7	95,8	169,5	135,4
V6	95,8	169,5	135,4
M6	95,8	169,5	135,4
V5	96,8	85,1	135,4
V4/M5	97,6	85,1	135,4
M4	97,6	85,1	135,4
V3	97,6	85,1	135,4
M3	38,7	54,1	98,3
V2	34,5	40,5	72,4
V1/M2	34,5	40,5	72,4
M1	38,7	24,3	72,5
V_{int}	53,7	53,5	76,6
M_{int}	9,1	33,9	39,0

that is redundant), however V_{dd} and GND bumps locations depend also on power chip design and it goes beyond this analysis.

Table 3.6: Total IR-drop and impact on V_{dd}

IR-drop		
	Tot [mV]	% of V_{dd}
$Path\#1$	25.5	3.6
$Path\#2$	34	4.9
$Path\#3$	33.2	4.7

3.9 Conclusions

In this analysis, IR-drop was evaluated for paths under V_{DD} bumps, hence very close to power source. The resistance breakdown hid the actual contribution of

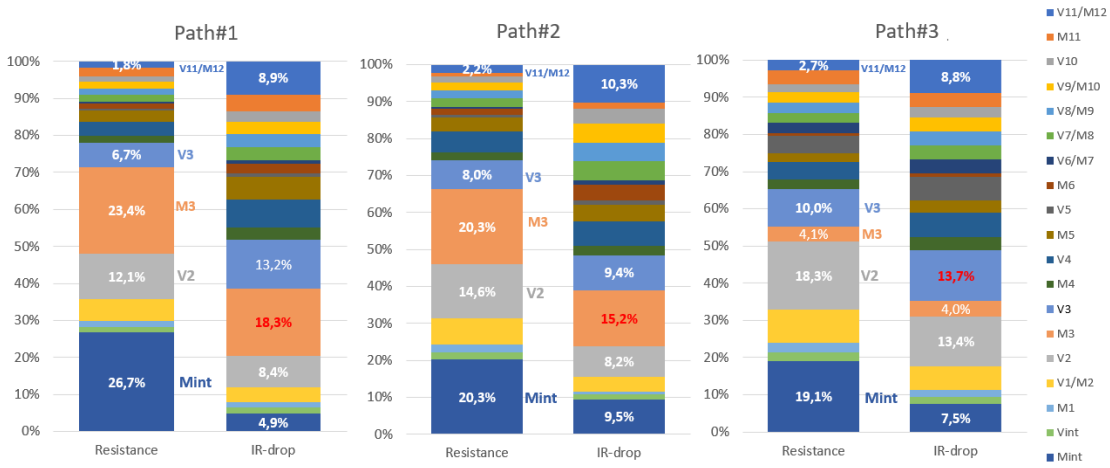


Figure 3.17: IR-drop breakdown obtained by multiplying current flowing through every element for the respective resistance

every component, unless current is used. Results showed that resistance can not be used as the only metrics for PDN design rules, as a consequence, the scaling of every component must be performed by taking into account the current flowing in each part of the PDN. Even if current follows vertical paths major contributors are still lines when standard cells are not located close to power bumps, therefore, despite the limited number of paths analyzed, it is clear that for most paths lines will own the highest IR-drop contribution. Then scaling rules must take into account which metal layers are the most used to reach standard cells to carefully design power delivery networks and keep IR-drop under the safe level set to 10%. Considering electromigration phenomenon, this level might be slightly lower to ensure a good device operation for the entire operating lifetime, however further research is needed because failure rate probability is still highly architecture specific and it must be evaluated case by case.

Appendix A

SPX and Raphael simulations code

A.1 Layout definition

In order to extract interconnect capacitance and resistance with Raphael a .tdr file containing the 3D structure must be generated with SPX. In this latter the full process emulation is carried out based on layout input file. The layout can be provided in .gds or .oasis extension and it can be directly included in the project or it can be generated with SPX tool as well. The following code can generate an .oasis file with layout specifications defined in the design-of-experiments (DOE) table, and so by the user.

```
\caption{Layout code}
\label{}
## Layout definition
## Default unit is nm

######################## Abbreviation ############################
####  The layout is centered
####     l  = lowest value
####     h  = highest value
####     W  = width
####     L  = length

###################### Pre-processing ############################
#This is to avoid getting an error during pre-processing
#set C_LEFT 0.0
#set C_RIGHT 0.0
#set C_TOP 0.0
#set C_BOT 0.0
```

```
#set R_M2 0.0
#set R_M2_OhmPerUm 0.0
#################### Parameter caculation ########################

set Pitch_M1M3  [expr 1.0*@Pitch_M1M3@]
set Pitch_M2    [expr 1.0*@Pitch_M2@]

set M1_CD       [expr  1.00*(@<CD_M1>@ -@CD_M1_var@)]
set M2_CD       [expr  1.00*(@<CD_M2>@ -@CD_M2_var@)]
set M3_CD        [expr  1.00*(@<CD_M3>@ -@CD_M3_var@)]

set L_M1M3      [expr (4.0*$Pitch_M2 + $M2_CD)*10.0]
set L_M2        [expr (2.0*$Pitch_M1M3 + $M1_CD)*10.0]

## Coordinates of M1 M2 M3 ##

set L_M1M3_l         [expr -($L_M1M3/2.0)*1.0]
set L_M1M3_h         [expr  ($L_M1M3/2.0)*1.0]

set L_M2_l         [expr -($L_M2/2.0)*1.0]
set L_M2_h         [expr  ($L_M2/2.0)*1.0]

## This is to define the layout grid ###
layout new  a -dbu 1e-12

#################### layout dimension ############################
layer add 0:0
layer active 0:0
layer name 0:0 M1
set M1_Y_l   [expr  $L_M2_l ]
set M1_Y_h   [expr  $L_M2_h ]

set M1_X_l_l [expr -$Pitch_M1M3 -$M1_CD/2.0 ]
set M1_X_l_h [expr -$Pitch_M1M3 +$M1_CD/2.0 ]
set M1_X_h_l [expr  $Pitch_M1M3 -$M1_CD/2.0 ]
set M1_X_h_h [expr  $Pitch_M1M3 +$M1_CD/2.0 ]

cell object add rectangle
    [list coords [list $M1_X_l_l $M1_Y_h $M1_X_l_h $M1_Y_l]]
cell object add rectangle
    [list coords [list $M1_X_h_l $M1_Y_h $M1_X_h_h $M1_Y_l]]


layer add 1:0
layer active 1:0
layer name 1:0 M2
set M2_Y_l_l [expr -$Pitch_M2 -$M2_CD/2.0 ]
set M2_Y_l_h [expr -$Pitch_M2 +$M2_CD/2.0 ]
set M2_Y_h_l [expr  $Pitch_M2 -$M2_CD/2.0 ]
```

```
set M2_Y_h_h [expr  $Pitch_M2 +$M2_CD/2.0 ]

set M2_X_l   [expr  $L_M1M3_l ]
set M2_X_h   [expr  $L_M1M3_h ]

cell object add rectangle [list coords [list $M2_X_l $M2_Y_l_l $M2_X_h $M2_Y_l_h]]
cell object add rectangle [list coords [list $M2_X_l $M2_Y_h_l $M2_X_h $M2_Y_h_h]]


layout extract n@node@_lyt.oasis -format oasis -cell a

exit
```

Layout file is then used in the next SPX tool to give instructions for lithographic steps, if any, included in the flow process. After parameters and materials definition, the structure generation is performed. Every process step is defined with proper specifications such as materials, deposition/etching rate and time. For simplicity, rate is kept to 1 um/min and time variable is parameterized in such a way etching depth or height deposition can be easily set by using dimensions parameters defined in DOE table. Below it is only an extract of the code representing first layer generation.

```
################## Structure Generation ##########################


############################### Semi-D ###############################
set R n@node@
set F SemiD

spx::route Route= $R
spx::flow  Route= $R Flow= $F
spx::module Route= $R Flow= $F Module= "Wafer Start"
spx::substrate Route= $R Flow= $F Module= "Wafer Start"
    Step= substrate ColumnSizeX= {0.0001 um} ColumnSizeY= {0.0001 um}
        Thickness= {0.02 um}

spx::module Route= $R Flow= $F Module= "M0"
spx::depo_isotropic Route= $R Flow= $F Module= "M0"
    Step= "depo_isotropic" MaterialFilter= {{{Material equals SiN}}}
        Rate= {1 {um min^-1}} Time= "{$H_V2 min}"

spx::depo_isotropic Route= $R Flow= $F Module= "M0"
    Step= "depo_isotropic_01"
        MaterialFilter= {{{Material equals AmorphousCarbon}}}
```

```
            Rate= {1 {um min^-1}} Time= "{$H_V2 min}"

spx::litho_patterning Route=$R Flow= $F Module= "M0"
    Step= "Mask_M0" OffsetX= {0.0 um} OffsetY= {0.0 um} LayerName= {M2}
        Polarity= Positive Thickness= {0.03 um}

spx::depo_isotropic Route= $R Flow= $F Module= "M0"
    Step= "depo_isotropic_02"
        MaterialFilter= {{{Material equals Oxide}}}
            Rate= {1 {um min^-1}} Time= "{$T_SPACER_2 min}"

spx::etch_anisotropic Route= $R Flow= $F Module= "M0"
    Step= "etch_anis"
        MaterialDependent= {{{{{Material equals Oxide}}} {1 {um min^-1}}}}
            Time= "{[expr $T_SPACER_2 + 0.001] min}"

spx::etch_strip Route= $R Flow= $F Module= "M0"
    Step= "etch_strip" MaterialFilter= {{{Material equals Photoresist}}}

spx::etch_anisotropic Route= $R Flow= $F Module= "M0"
    Step= "etch_anis_01"
        MaterialDependent=
            {{{{{Material equals AmorphousCarbon}}} {1 {um min^-1}}}}
                Time= "{[expr $H_V2+0.001] min}"

spx::etch_strip Route= $R Flow= $F Module= "M0"
    Step= "etch_strip_01" MaterialFilter= {{{Material equals Oxide}}}

spx::depo_fill Route= $R Flow= $F Module= "M0"
    Step= "depo_fill"
        MaterialFilter= {{{Material equals Oxide}}} Offset= "{$H_V2 um}"
            ReferenceMaterialsFilter= {{{Material equals SiN}}}

spx::etch_cmp Route= $R Flow= $F Module= "M0"
    Step= "etch_cmp_Ox"
        EtchstopMaterialsFilter= {{{Material equals AmorphousCarbon}}}
            MaterialsFilter= {{{Material equals Oxide}}} Offset= {0 um}

spx::etch_anisotropic Route= $R Flow= $F Module= "M0"
    Step= "etch_anisotropic_02"
        MaterialDependent=
            {{{{{Material equals AmorphousCarbon}}} {1 {um min^-1}}}}
                Time= "{[expr $H_V2+0.001] min}"

spx::etch_anisotropic Route= $R Flow= $F Module= "M0"
    Step= "etch_anisotropic_03"
        MaterialDependent= {{{{{Material equals SiN}}} {1 {um min^-1}}}}
            Time= "{[expr $H_V2+0.001] min}"
```

```
spx::etch_strip Route= $R Flow= $F Module= "M0"
    Step= "etch_strip_02" MaterialFilter= {{{Material equals Oxide}}}

spx::depo_fill Route= $R Flow= $F Module= "M0"
    Step= "depo_fill_01"
        MaterialFilter= {{{Material equals Oxide}}} Offset= "{$H_V2 um}"
            ReferenceMaterialsFilter= {{{Material equals Silicon}}}

spx::etch_cmp Route= $R Flow= $F Module= "M0"
    Step= "etch_cmp"
        EtchstopMaterialsFilter= {{{Material equals SiN}}}
            MaterialsFilter=
                {{{Material equals SiN}} {{Material equals Oxide}}}
                    Offset= {0 um}
```

The 3D structure is then saved in a .tdr file that is used from Raphael FX tool. In the code below the mesh is generated for the entire structure and a refining is applied only at interfaces of metal with the dielectric.

```
       init tdr=n@node|-1@_spx.tdr ;# initialization step to read the tdr

## meshing
## this is material-based regular (coarse) mesh used to settle a maximum size
for the mesh when moving away from the interfaces (see interface mesh next)
refinebox name=Box1
    min= { $top_box $L_M1_l  $L_M2_l }
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  AlOx
refinebox name=Box2
    min= { $top_box $L_M1_l  $L_M2_l }
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  SiCN
refinebox name=Box3
    min= { $top_box $L_M1_l  $L_M2_l }
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  Ru
refinebox name=Box4
    min= { $top_box $L_M1_l  $L_M2_l }
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  Oxide
refinebox name=Box5
    min= { $top_box $L_M1_l  $L_M2_l }
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  SiN
refinebox name=Box6
    min= { $top_box $L_M1_l  $L_M2_l }
```

```
        max= {  $bot_box $L_M1_h $L_M2_h }
            xrefine=0.004 yrefine=0.004 zrefine=0.004  Air


## this is to define the interface (fine) mesh
refinebox name=interface_Al1
    min.normal.size= 0.0005 normal.growth.ratio= 2
        interface.mat.pairs= {Ru AlOx}
refinebox name=interface_Ni
    min.normal.size= 0.0005 normal.growth.ratio= 2
        interface.mat.pairs= {Ru Nitride}
refinebox name=interface_SiN
    min.normal.size= 0.0005 normal.growth.ratio= 2
        interface.mat.pairs= {Ru SiN}
```

The left part is to define contacts. In Raphael FX is sufficient to define a box that includes the metal interface being careful to not incorporate other interfaces. Since the goal is to extract the capacitance here it is necessary to define five different contacts. The bottom one will include all lines in bottom layer (M_{bot}) and the top one lines above (M_{top}), M_l,M_r,M_m refer to left, right and middle lines respectively.

```
##this is to define the contacts
contact name= Mbot @Metal_M1M3@ box
    xlo= $bot_box_l  ylo= $L_M1_l  zlo= $L_M2_l
        xhi= $bot_box_h yhi= $L_M1_h  zhi= $L_M2_h !replace
contact name= Ml @Metal_M2@ box
    xlo= $mid_box_l  ylo= $W_M2_left_l  zlo= $L_M2_l
        xhi= $mid_box_h yhi= $W_M2_left_h  zhi= $L_M2_h !replace

contact name= Mr @Metal_M2@ box
    xlo= $mid_box_l  ylo= $W_M2_right_l zlo= $L_M2_l
        xhi= $mid_box_h yhi= $W_M2_right_h zhi= $L_M2_h !replace

contact name= Mtop @Metal_M1M3@ box
    xlo= $top_box_l  ylo= $L_M1_l  zlo= $L_M2_l
        xhi= $top_box_h yhi= $L_M1_h  zhi=$L_M2_h  !replace

contact name= Mm @Metal_M2@ box
    xlo= $mid_box_l  ylo= $W_M2_Mid_l   zlo= $L_M2_l
        xhi= $mid_box_h yhi= $W_M2_Mid_h   zhi= $L_M2_h !replace
```

Bibliography

[1] Ivan Ciofi Odysseas Zografos Pieter Weckx Julien Ryckaert Pieter Schuddinck Yang Xiang Zsolt Tokei Anita Farokhnejad, Simone Esposto. Evaluation of beol scaling boosters for sub-2nm using enhanced-ro analysis. In *2022 IEEE International Interconnect Technology Conference (IITC)*, 2022.

[2] J.R. Black. Electromigration—a brief survey and some recent results. *IEEE Transactions on Electron Devices*, 16(4):338–347, 1969.

[3] F. Chen, M. Shinosky, B. Li, J. Gambino, S. Mongeon, P. Pokrinchak, J. Aitken, D. Badami, M. Angyal, R. Achanta, G. Bonilla, G. Yang, P. Liu, K. Li, J. Sudijono, Y. Tan, T. J. Tang, and C. Child. Critical ultra low-k tddb reliability issues for advanced cmos technologies. pages 464–475, 2009.

[4] Ivan Ciofi, Philippe J Roussel, Rogier Baert, Antonino Contino, Anshul Gupta, Kristof Croes, Christopher J Wilson, Dan Mocuta, and Zsolt Tokei. Rc benefits of advanced metallization options. *IEEE transactions on electron devices*, 66(5):2339–2345, 2019.

[5] Ivan Ciofi, Philippe J Roussel, Yves Saad, Victor Moroz, Chia-Ying Hu, Rogier Baert, Kristof Croes, Antonino Contino, Kevin Vandersmissen, Weimin Gao, Philippe Matagne, Mustafa Badaroglu, Christopher J Wilson, Dan Mocuta, and Zsolt Tokei. Modeling of via resistance for advanced technology nodes. *IEEE transactions on electron devices*, 64(5):2306–2313, 2017.

[6] Uttam Kumar Das and Tarun Kanti Bhattacharyya. Opportunities in device scaling for 3-nm node and beyond: Finfet versus gaa-fet versus ufet. *IEEE transactions on electron devices*, 67(6):2633–2638, 2020.

[7] Nan Fu, Yanxiang Liu, Xiaolong Ma, and Zanfeng Chen. Euv lithography: state-of-the-art review. *J. Microelectron. Manuf*, 2(2):1–6, 2019.

[8] L.G. Gosset, A. Farcy, J. de Pontcharra, Ph. Lyan, R. Daamen, G.J.A.M. Verheijden, V. Arnal, F. Gaillard, D. Bouchu, P.H.L. Bancken, T. Vandeweyer, J. Michelon, V. Nguyen Hoang, R.J.O.M. Hoofman, and J. Torres. Advanced

cu interconnects using air gaps. *Microelectronic Engineering*, 82(3):321–332, 2005. Proceedings of the ninth european workshop on materials for advanced metallization 2005.

[9] Naoto Horiguchi and Zsolt Tokei. A view on the logic technology roadmap. *Semiconductor Digest*, 2(06), 2020.

[10] C.W. Kaanta, S.G. Bombardier, W.J. Cote, W.R. Hill, G. Kerszykowski, H.S. Landis, D.J. Poindexter, C.W. Pollard, G.H. Ross, J.G. Ryan, S. Wolff, and J.E. Cronin. Dual damascene: a ulsi wiring technology. pages 144–152, 1991.

[11] M. A. Korhonen, P. Bo/rgesen, K. N. Tu, and Che-Yu Li. Stress evolution due to electromigration in confined metal lines. *Journal of Applied Physics*, 73(8):3790–3799, 1993.

[12] A. Leśniewska, O. Varela Pedreira, M. Lofrano, G. Murdoch, M. van der Veen, A. Dangol, N. Horiguchi, Zs. Tőkei, and K. Croes. Reliability of a dme ru semidamascene scheme with 16 nm wide airgaps. In *2021 IEEE International Reliability Physics Symposium (IRPS)*, pages 1–6, 2021.

[13] Baozhen Li, Paul McLaughlin, Jeanne Bickford, Peter Habitz, Dileep Netrabile, and Timothy Sullivan. Statistical evaluation of electromigration reliability at chip level. *Device and Materials Reliability, IEEE Transactions on*, 11:86 – 91, 04 2011.

[14] W Liu, Y. K Lim, J. B Tan, W. Y Zhang, H Liu, and S. Y Siah. Study of tddb reliability in misaligned via chain structures. 2012.

[15] Gordon E Moore. Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85, 1998.

[16] Gayle Murdoch, Zsolt Tokei, Sara Paolillo, Olalla Varela Pedreira, Kris Vanstreels, and Christopher J Wilson. Semidamascene interconnects for 2nm node and beyond. pages 4–6, 2020.

[17] J. Noguchi, T. Saito, N. Ohashi, H. Ashihara, H. Maruyama, M. Kubo, H. Yamaguchi, D. Ryuzaki, K.-I. Takeda, and K. Hinode. Impact of low-k dielectrics and barrier metals on tddb lifetime of cu interconnects. pages 355–359, 2001.

[18] Sara Paolillo, Danny Wan, Frédéric Lazzarino, Nouredine Rassoul, Daniele Piumi, and Zsolt Tőkei. Direct metal etch of ruthenium for advanced interconnect. *Journal of vacuum science and technology. B, Nanotechnology microelectronics*, 36(3):3, 2018.

[19] Seongho Park, Sue Ann Bidstrup Allen, and Paul A Kohl. Air-gaps for high-performance on-chip interconnect part ii: modeling, fabrication, and characterization. *Journal of Electronic Materials*, 37(10):1534–1546, 2008.

[20] Arya Tah and Debasish Kumar Rakshit. Study of the frequency characteristics of a ring oscillator. *International Journal of Computing and Network Technology*, 1(03), 2013.

[21] Danny Wan, Sara Paolillo, Nouredine Rassoul, Bogumila Kutrzeba Kotowska, Victor Blanco, Christoph Adelmann, Frederic Lazzarino, Monique Ercken, Gayle Murdoch, Jürgen Bömmels, Christopher J. Wilson, and Zsolt Tökei. Subtractive etch of ruthenium for sub-5nm interconnect. pages 10–12, 2018.

[22] Yayi Wei. *Advanced processes for 193-nm immersion lithography.* SPIE, Bellingham, Wash, 2009.

[23] D Yakimets, M. Garcia Bardon, D Jang, P Schuddinck, Y Sherazi, P Weckx, K Miyaguchi, B Parvais, P Raghavan, A Spessot, D Verkest, and A Mocuta. Power aware finfet and lateral nanosheet fet targeting for 3nm cmos technology. pages 20.4.1–20.4.4, 2017.

[24] Geoffrey Yeap. Smart mobile socs driving the semiconductor industry: Technology trend, challenges and opportunities. pages 1.3.1–1.3.8, 2013.

[25] Houman Zahedmanesh, Ivan Ciofi, Odysseas Zografos, Mustafa Badaroglu, and Kristof Croes. A novel system-level physics-based electromigration modelling framework: Application to the power delivery network. In *2021 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)*, pages 1–7, 2021.

[26] Houman Zahedmanesh and Kris Vanstreels. Mechanical integrity of back-end-of-line with ru nanowires and airgaps. *Microelectronics Reliability*, 110:113700, 2020.

[27] Z. Zhang, M. Kraatz, M. Hauschildt, S. Choi, A. Clausner, E. Zschech, and M. Gall. Strategy to characterize electromigration short length effects in cu/low-k interconnects. In *2021 IEEE International Reliability Physics Symposium (IRPS)*, pages 1–5, 2021.