

# From Feature Scale Simulation to Backend Simulation for a 100nm CMOS Process

Fuad Badrieh\*, Helmut Puchner\*, Clemens Heitzinger<sup>°</sup>,

Alireza Sheikholeslami<sup>°</sup>, and Siegfried Selberherr<sup>°</sup>

\*Cypress Semiconductor, San Jose, CA 95134, USA

<sup>°</sup>Institute for Microelectronics, TU Vienna, A-1040 Vienna, Austria

fib@cypress.com

## Abstract

*The current challenge for TCAD is the prediction of the performance of groups of devices, backends, and – generally speaking – large parts of the final IC in contrast to the simulation of single devices and their fabrication. This enables one to predictively simulate the performance of the final device depending on different process technologies and parameters, which the simulation of single devices cannot achieve.*

*In this paper we focus on the simulation of backend, interconnect capacitance, and time delays. To that end topography simulations of deposition, etching, and CMP processes in the various metal lines are used to build up the backend stack. The output of the feature scale simulations is used as input to a capacitance extraction tool, whose results are made available directly to the circuit designer.*

*We discuss the utilized simulation tools and their integration. The topography simulations were performed by our tool called ELSA (enhanced level set applications) and the subsequent simulations by RAPHAEL. Finally simulation results for a 100nm process are presented, where the influence of void formation between metal lines profoundly impacts the performance of the whole interconnect stack.*

## 1. Introduction

Interconnects are becoming more important with shrinking technologies. Timing delays due to metal lines contribute more and more to the overall delay. It is therefore imperative that the capacitance and resistance of interconnect lines are modeled accurately.

In order to model capacitance correctly, one has to know the metal profile, e.g., bottom and top critical dimensions and metal slope, the profile of the deposited layer with and without CMP, and the profile of the void, if it is formed. Generally these three profiles depend heavily on deposition process conditions, on metal thickness and line-to-line space, and less strongly on metal width.

Most resistance and capacitance extraction (RCX) tools assume rectangular metal profiles, either planar or con-

formal dielectrics, and have very simplistic void models. Even if the metal slope is modeled, it is mostly assumed constant and independent of space. While these ideal assumptions may be sufficient for older technologies, they are insufficient for today's technologies like the example from a 100nm process considered in Section 5, where interconnects have a large number of special features which are nowhere close to ideal.

Voids in ILD (inter-layer dielectric) materials may serve a beneficial purpose. They lower the overall capacitance of a given metal layer and thus this approach is economically advantageous compared to using low- $k$  materials. The impact of the voids on metal capacitance, especially at small space, is tremendous.

In this work we present the results from a rigorous simulation approach. In order to provide predictive simulations of the overall capacitance, the shapes and positions of the voids must be simulated accurately. Here topography simulations serve as the input to the following extraction of the capacitance. The extracted capacitance for the specific technology at hand is then made accessible to the circuit designer.

## 2. Feature Scale Simulation

Starting from a plain substrate, the principal topography simulation steps of etching trenches, depositing thin films, and CMP (chemical mechanical planarization) are carried out in order to build up the backend stacks (cf. Figure 1). Their simulation is performed by the ELSA simulator which is based on an advanced level set algorithm for describing moving boundaries [1–3]. It provides dynamic narrow banding and extends the speed function so that the signed distance function is retained throughout the simulation. The transport of the particles above the wafer can be simulated both in the diffusion and radiosity regime. A surface coarsening algorithm significantly speeds up the radiosity simulations, which is important for the backend simulation where a large number of simulations is required (cf. Section 3).

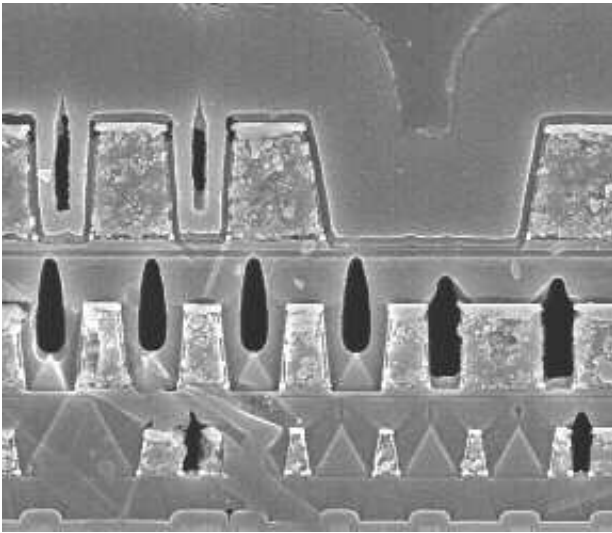


Figure 1. SEM image of a whole backend stack comprised of three Al metals and a Ti-nitride local interconnect.

## 2.1. Etching

For simulating the etching processes, we used a multi-region extension of the level set algorithm. The main moving boundary, namely the wafer surface, is superposed by regions with different speed functions. These are the masks at the top and the etch stop at the bottom or regions at the bottom which are not affected because of the selectiveness of the etch process. In all these regions the speed function equals zero.

The transport of particles above the wafer is handled by a radiosity model taking into account the specular reflection of high energy particles.

## 2.2. Deposition

In the processes considered, the transport of particles happens in the radiosity regime. The deposition processes are governed by luminescent reflection. The films deposited are silicon nitride and silicon dioxide films (cf. Section 5).

The silicon nitride films were deposited by PECVD from silane and  $\text{NH}_3$  and were not doped. The reaction is  $\text{SiH}_4 + \text{NH}_3 \rightarrow \text{SiNH} + 3\text{H}_2$  [4]. For simulation purposes this was considered the essential reaction. The silicon dioxide films were deposited by pyrolytic decomposition of TEOS in an LPCVD process.

When modeling topography processes it is generally possible to write down quite complicated reaction paths. However it is not straightforward to determine the vital reactions and their constants. Therefore it is mandatory to reduce the possible reaction paths to an essential minimum which reproduces the observed phenomena.

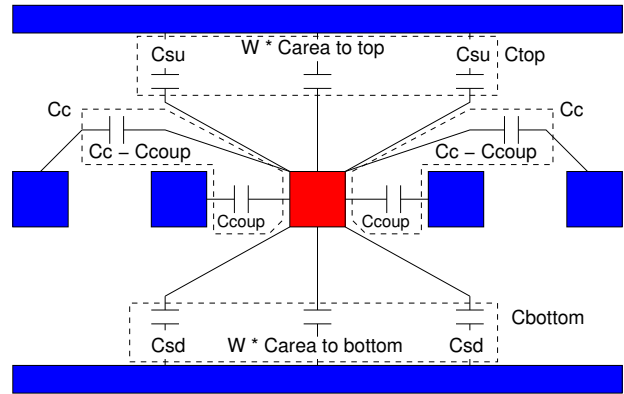


Figure 2. Two-dimensional schematics of a signal line (middle) surrounded by grounded lines and planes. The total capacitance equals the sum of twice the coupling capacitance plus top and bottom capacitances.

## 2.3. CMP

On the feature scale level the simulation of CMP can be performed in a straightforward manner, since the thickness of the remaining part of the layer is known. In this simulation step the boundary is modified so that all parts above the given thickness are removed and the remaining points are joint accordingly.

## 3. Backend Simulation

The coordinates generated by the topography simulator are then fed into a field solver, in this case RAPHAEL [5], which builds the structure, calculates electric fields and charge densities, and solves for various capacitances.

ELSA generally reports a large number of coordinates for describing ILD and void coordinates depending on the resolution required. Since this number of surface elements determines the number of RAPHAEL grid points, a significant reduction is necessary. Otherwise simulation times would be orders of magnitude larger than those for equivalent simplified structures. Hence we use a surface coarsening algorithm for reworking the output and reducing the number of points supplied to the field solver [2]. For this application simulation time is an important factor, since thousands of simulations are necessary for characterizing a technology (see below and cf. Section 5).

### 3.1. Capacitance Types

In order to model capacitances, we assume that a signal line is at high voltage and surrounded by two lines on the left, two lines on the right, a plane underneath, and a plane above. The surrounding lines and planes are assumed to be at ground voltage. For example an M2 signal line could be surrounded by M2 grounded lines above the M1 plane and underneath the M3 plane. This skeleton is shown in Figure 2.

There is a coupling capacitance to an adjacent line ( $C_{\text{coup}}$ ), a coupling capacitance to a next to adjacent line (which usually is very small), a capacitance to the top plate ( $C_{\text{top}}$ ), and a capacitance to the bottom plate ( $C_{\text{bot}}$ ). The top capacitance is further split into a top area capacitance ( $C_{\text{areat}}$ ) and a top side wall capacitance ( $C_{\text{su}}$ ). Similarly the bottom capacitance is split into a bottom area capacitance ( $C_{\text{areab}}$ ) and a bottom side wall capacitance ( $C_{\text{sd}}$ ). The bottom (top) capacitance equals width  $W$  times bottom (top) area capacitance plus twice bottom (top) side wall capacitance.

### 3.2. Capacitance Models

Capacitance simulations are performed as a function of three variables: First, they depend on metal combination, e.g., M4 above M2, or M3 above M1 and underneath M5. For a six layer technology the combinations can be above a hundred.

Second, they depend on line-to-line space. Simulations start from the minimum allowed space (e.g.,  $0.14\mu\text{m}$ ) and end in the range of a few microns (e.g.,  $6\mu\text{m}$ ). Generally initial space increments are fine to capture strong dependence of capacitance on space and final increments are coarse to capture capacitance saturation.

Third, capacitance simulations depend on line width. Simulations start from the minimum allowed width to approximately 60 times the minimum width. Since the dependence on the width is well behaved, only a few intermediate widths are usually needed.

Once the set of metal combinations, spaces, and widths is finalized, the field solver simulations are started. In each case technology coordinates are read from the topography simulator and a set of capacitances is calculated and stored in a database. After all simulations are performed, one has accumulated a large range of capacitance parameters for different combinations of metals, spaces, and widths. All capacitances are reported per unit length, e.g.,  $\text{fF}/\mu\text{m}$ .

## 4. Interfacing to Circuit Design

In order to ensure that the above database is usable for design, a TCAD (technology computer aided design) flow is needed as an intermediate between the database and the circuit designer. First the designer instantiates a capacitance element for, e.g., M3 above substrate. The designer has to specify the metal name and surrounding planes (in this case M3 and substrate) as well as metal width, length, and space. Second the designer netlists the schematics, i.e., creates a SPICE program file.

Finally the database is read depending on designer options, locates the capacitor of interest, interpolates or extrapolates for space and width, and replaces the capacitance instance with a numeric value for the capacitance at that node. Now the designer can run the SPICE netlist and observe circuit performance. If timing requirements are not met, the designer changes the properties of the capac-

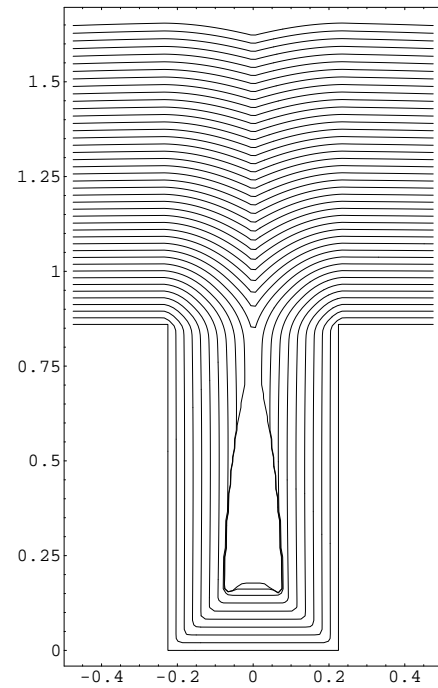


Figure 3. Simulation of a deposition process and void formation.

itance, e.g., makes it narrower for lower capacitance, and repeats the process.

## 5. Example and Simulation Results

An SEM image of a typical backend stack simulated is shown in Figure 1. The films deposited are nitride and silicon dioxide films and the interconnect lines are made of aluminum. The backend stacks considered are part of a  $100\text{nm}$  process.

We applied ELSA to the following real world example (Figure 3). Here we have M3 lines, i.e., top metal in this case, above the M2 plane. M3 lines are at  $0.50\mu\text{m}$  width, and at a line-to-line space which varies between  $0.45\mu\text{m}$  and  $1.9\mu\text{m}$ . Figure 4 shows the RAPHAEL structure constructed from ELSA coordinates at minimum space. Notice that at minimum space the metal lines are vertical and voids are maximum in size.

Next we build the structure at  $1.5\mu\text{m}$  space. Figure 5 shows the corresponding RAPHAEL structure again built from ELSA simulation results. Notice that in this case, however, metal lines have a slope, which duplicates the real process geometry. Voids are quite small in this case. Finally we build the structure at  $1.9\mu\text{m}$  space. Figure 6 shows the resulting RAPHAEL structure. Metals have slopes here as well, but voids no longer exist.

The middle line capacitance was simulated for the above three structures and compared to CBCM (charge based capacitance measurement). The results are shown

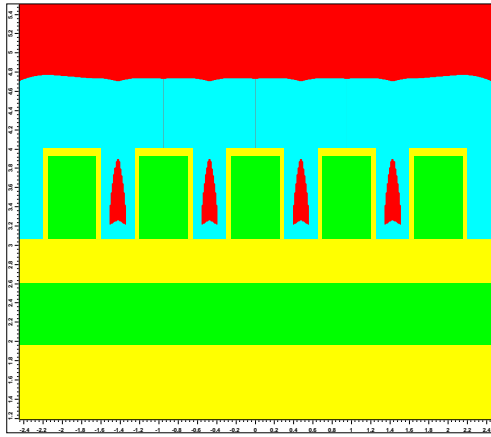


Figure 4. M3 lines ( $0.5\mu\text{m}$  top width) at  $0.45\mu\text{m}$  space above the M2 plane. Cap oxide was deposited (thin layer surrounding M3 lines) before the top nitride (thick layer) was deposited. Voids formation occurs between the lines (cf. Figure 3).

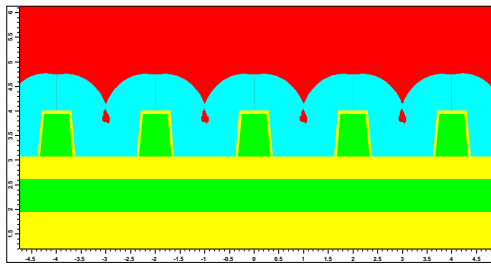


Figure 5. M3 lines ( $0.5\mu\text{m}$  top width) at  $1.50\mu\text{m}$  space above the M2 plane. In this case the voids formed are quite small.

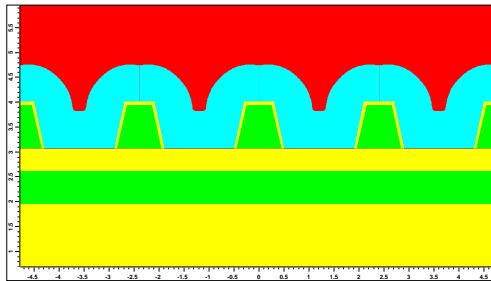


Figure 6. M3 lines ( $0.5\mu\text{m}$  top width) at  $1.90\mu\text{m}$  space above the M2 plane. Since line-to-line space is large enough, no voids form between lines.

in Figure 7. We observe very good agreement with an error of less than 5%.

## 6. Conclusion

The topography simulator ELSA and the RCX tool RAPHAEL, together with a few auxiliary tools, were combined to provide simulations of capacitances, and thus timings, in backend stacks. This simulation flow starts ab initio with silicon surfaces and takes into account etch-

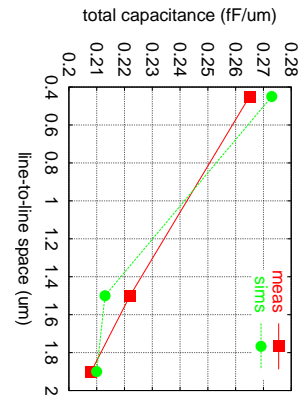


Figure 7. This figure compares the M3 middle line capacitance as a function of line-to-line spaces between simulations and measurement.

ing, deposition, and CMP steps. From these the complex structures of interconnect lines are built resulting in many combinations depending on metal combination, line-to-line space, and line width. The interconnect structures serve as input to the field solver whose capacitance simulations are stored in a database. The circuit designer accesses the results of this simulation flow and uses them in SPICE.

Void formation has a significant influence on the capacitances and hence a rigorous simulation approach like the one presented is indispensable for today's technologies. Furthermore using voids in a controlled and reproducible manner can be an economically advantageous substitute for low- $k$  materials.

The simulations show very good agreement with CBCMs (charge based capacitance measurements). These simulations were used during the development of a 100nm CMOS process and proved to be of great help for circuit design.

- [1] J. Sethian, *Level Set Methods and Fast Marching Methods*. Cambridge: Cambridge University Press, 1999.
- [2] C. Heitzinger, J. Fugger, O. Häberlen, and S. Selberherr, "On Increasing the Accuracy of Simulations of Deposition and Etching Processes Using Radiosity and the Level Set Method," in *Proc. 32th European Solid-State Device Research Conference (ESSDERC 2002)* (G. Baccarani, E. Gnani, and M. Rudan, eds.), (Florence, Italy), pp. 347–350, University of Bologna, Sept. 2002.
- [3] C. Heitzinger, J. Fugger, O. Häberlen, and S. Selberherr, "Simulation and Inverse Modeling of TEOS Deposition Processes Using a Fast Level Set Method," in *Proc. Simulation of Semiconductor Processes and Devices (SISPAD 2002)*, (Kobe, Japan), pp. 191–194, Business Center for Academic Societies, Japan, Sept. 2002.
- [4] G. Schumicki and P. Seegebrecht, *Prozess-technologie*. Springer, 1991.
- [5] Avant! Corporation, TCAD Business Unit, Fremont, CA, USA, *Raphael, Interconnect Analysis Program, Version 2000.2, Reference Manual*, July 2000.