A million wafer, virtual fabrication approach to determine process capability requirements for an industry-standard 5nm BEOL two-level metal flow

W. F. Clark, A. Juncker
COVENTOR SARL,
Villebon sur Yvette,
France
wfc@coventor.com

E. Paladugu, D. Fried
COVENTOR Inc.
Waltham, MA USA

C. J. Wilson, G. Pourtois, M. Gallagher, A. De Jamblinne, D. Piumi, J. Boemmels, Z. S. Tokei, D. Mocuta
IMEC, Kapeldreef 75
3000 Leuven, Belgium

Abstract Process simulations provide vital insights to identify the key process steps to dedicate wafer resources for improvement or to determine investment on tool capability. We considered this problem in the context of an industry-like 5nm Back-End-of-Line flow being developed in IMEC and modeled the approximately 150 step process flow in COVENTOR SEMulator3D®. For the first time a one-million wafer Design of Experiments was conducted to sample a 10-dimensional variable space and derive the failure points for each process parameter. A vector based algorithm was used to search the parameter space and derive a hyper-surface to represent the absolute yield limits. The virtual wafers were run to identify process sensitivities and spec limits for expected process variations. This work highlights that process optimization is needed to improve the capability of many processes to the order of 1nm and this methodology should be used to screen standard libraries for process sensitivities.

(Keywords: BEOL, Integration, Process Window, Process Step , Simulation, COVENTOR, N5)

I. Introduction

A. BEOL Integration Flow and Design Clip

The N5 based flow considered in this study consists of a 21nm half-pitch Spacer Assisted Double Patterning (SADP) local-level (M0) and a 16nm half-pitch Spacer Assisted Quadruple Patterning (SAQP), Triple Lithography Etch (LE3) Block and LE3 via second tight-pitch level (M1) [1]. As has been presented in previous FEOL studies [2, 3] our BEOL process was modelled in COVENTOR SEMulator3D® resulting in an approximately 150 process step flow (Fig. 1).

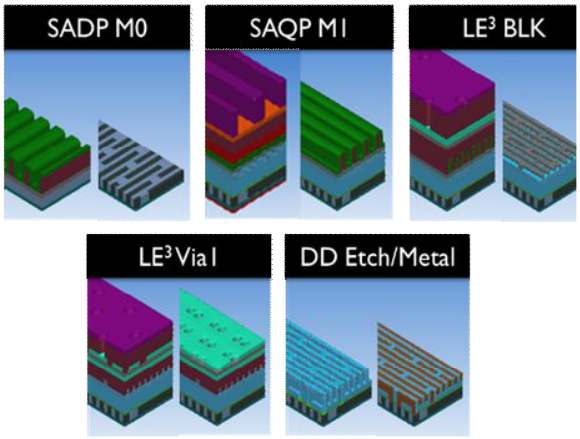

Fig. 1. Two-level metal integration flow at end of key modules.

We chose to focus on the tight-pitch level (M1) and simplified the M0 SADP by emulating the flow with Boolean mask operations to improve the runtime. We also only considered Block A and Via A to represent the cases of Block and Via B's and Block and Via C's. This simplification resulted in the 10 variables summarized in table 1.

The runtime of a single virtual build was approximately 300s running on a 4 four-core laptop computer. The design clip comprised an area of 0.26um^2 and was built with 1.0nm voxel resolution. The extended experiment was run on Linux based system comprised of 400 cores and used 500 software licenses with and load leveling. An incremental rebuild feature is incorporated into SEMulator3D which allows new runs to build on existing runs therefore the total run time scales sub-linearly with additional builds. Virtual metrology measurements including critical dimension (CD) and 3-dimensional design checks were used to verify the build against the design requirements. The failure criteria was based on electrical connectivity (Fig. 3) with a 30%

978-1-5090-0818-6/16/$31.00 ©2016 IEEE

change in the via contact resistance which was determined by interface contact area measurements.

Table 1. Process parameters, simulated process space, and expected N5 capability

Process Step	Nominal (nm)	High (nm)	Low (nm)	Expected N5 process 3s (nm)
Core CD bias	0	5	-5	2
Core Overlay in Y	0	10	-10	4
Spacer 1 thickness	16	20	12	1
Spacer 2 thickness	16	20	12	1
Block A CD	0	5	-5	2
Block A overlay in X	0	10	-10	4
Block A overlay in Y	0	10	-10	4
Via A CD	0	5	-5	2
Via A overlay in X	0	10	-10	4
Via A overlay in Y	0	10	-10	4

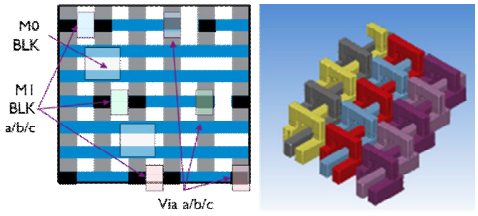

Fig. 2(left). Reduced design clip of the simulated intertwined structure; 6 via populations formed by the intersection of the SAQP and SADP defined lines; (right) 3D representation of the structure cut (colors are isolated conductor nets)

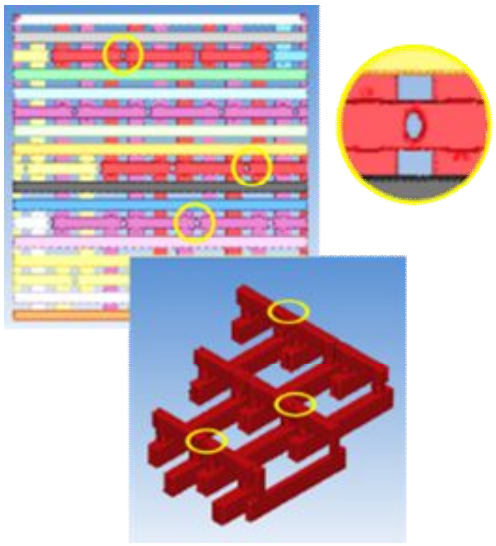

Fig. 3. Target Structure showing shorted chains due to an incomplete line cut (colors are isolated conductor nets). Electrical shorts decrease the total number of nets while opens increase the number of nets.

B. Simulation Methodology

A brute force approach would uniformly probe the total space of variables: e.g. this is demonstrated for two parameters in left side of Fig. 4. Such a dimensional sampling of even the reduced parameter set in table 1 requires performing more than 6.4×10^9 evaluations. To circumvent this problem, a vector based sampling algorithm has been implemented as in the right side of Fig. 4. Starting from unitary variations along the main variable axes, the algorithm identified the combination of variables that met the failure criteria. The apex of the function was determined using a bisection. Once they were identified for the initial primary vectors, they were linearly combined to sample a different direction. This gradually refined the resolution of the probed 10-dimensional space and resulted in a significant reduction of the number of iterations needed to establish the shape of the hyper surface. With each vector combination cycle, the resolution of the hyper-surface was gradually improved. Each combination was probed independently from the others, allowing effective parallelization.

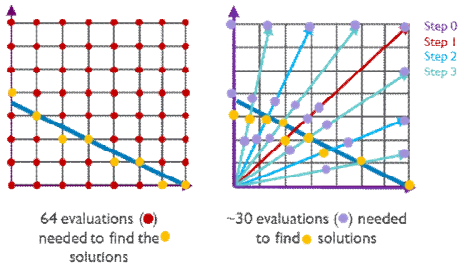

Fig. 4. Simulation methodology: (left) A brute force probing to derive the functional surface of two parameters requiring 64 evaluations; (right) The vector sampling methodology requiring only 30 evaluations.

II. Analysis and Discussion

The surface generated in the 10-dimensional space can be thought of having 10 major axes each of which intersects twice with it. These points of intersection correspond to the maximum magnitude a parameter can have before the structure fails while all other parameters are held at a constant, nominal value (left axis of Fig. 5). However, these values cannot be thought of as upper/lower limits for each parameter as there is variability in the other parameters that requires each single parameter be held to tighter tolerances than the limit of each considered alone.

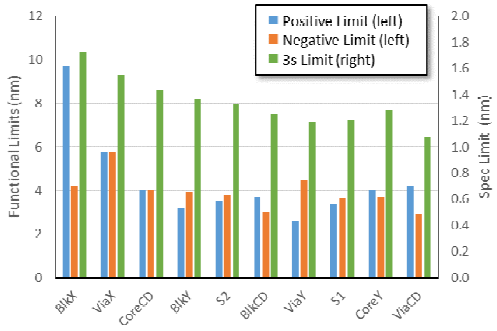

Fig. 5. The left axis shows the extent of the functional surface in each of the 10 directions for positive or negative values. These correspond to the maximum functional value of each parameter while all others are held at 0 offset from their nominal values. The right axis shows the 3s spec limit for each parameter assuming all other parameters are free within the expect range from table 1.

The data points generated in 10-dimensional space represent the functioning outer limits of the test structure and can be thought of as a surface. Visualizing a 10-dimensional surface is a challenge. However it can be plotted easily when holding all other parameters constant (e.g. Fig. 6 and Fig. 7 for the via CD bias and x and y overlay). Bisecting the surface at different via CD bias conditions clearly changes the possible process window and demonstrates a negative bias on via CD can increase process window in the y-overlay (at the cost of x-overlay).

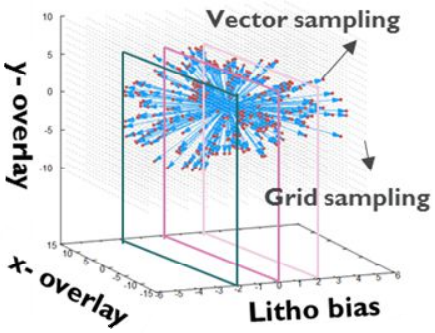

Fig. 6. 3D representation of the process window surface derived for via CD bias, x overlay and y overlay with all other parameters assumed nominal. Dots represent a point on the surface and the arrows the vector.

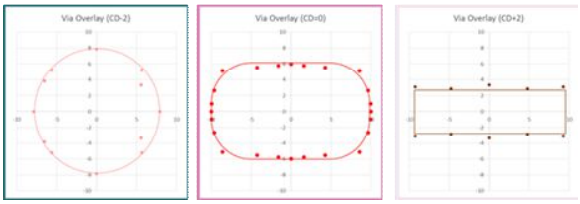

Fig. 7. Surface cross-sections at different via CD bias conditions demonstrating the via self-alignment and therefore increased overlay window in x.

The 10-dimensional-space can also be explored by generating virtual wafers and determining whether each wafer is within the functional space or outside of it. A wafer was generated by assuming that each of the 10 process parameters, $\mathbf{x_i}$, can have a random, Monte-Carlo value with a normal distribution with mean $\boldsymbol{\mu_i}$, and standard deviation σi. Each wafer then is represented by a vector $\mathbf{x}=[\mathbf{x_1}, \mathbf{x_2}, .. \mathbf{x_{10}}]$ which in turn can be represented by its unit vector $\mathbf{x}\hat{}$, and magnitude $|\mathbf{x}|$. The unit vector of the wafer is compared with the unit vectors of the generated functional surface to find which has the most similar direction in 10-dimensional-space. To establish 3σ specs, each parameter can be investigated independently while the others parameters were kept within an expected capability range (right axis of Fig.5.)

These tighter specs compared to the functional limits are expected given the design of the simulated structure. The process sensitivities were also calculated (Fig. 8). This method included variability in all the parameters while a single parameter was allowed to widen its range. It was logical the structure is most sensitive to via CD bias and least sensitive to Block overlay in x as the block is oversized in the x-direction. The reduced sensitivity to via overly in x (as observed in Fig. 7) was also visible.

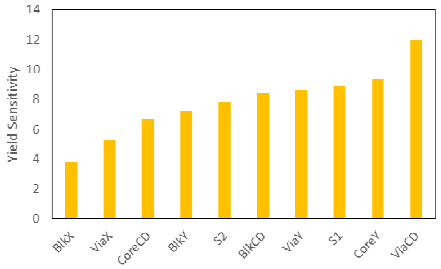

Fig. 8. Sensitivity of the yield to a 2x change in variability of each process parameter.
(Data are normalized to the nominal yield.)

III. Conclusions

We demonstrated a methodology to extract yield sensitivities for a full 150 step BEOL flow using predictive process simulations. We used a vector algorithm search method to sample a 10-dimensional variation space to a manageable one million wafers. We showed that to enable challenging interconnect design constructs work is needed to improve process capability to ±1nm. Improving lithography

overlay variation below the expected ±4nm alone will not account for the required improvement. Hence improved capability in the deposition and lithography-etch processes in the multi-patterning modules is needed. This study also makes a compelling case that in order to continue scaling, innovations in patterning, self-alignment and selective depositions are required.

IV. Acknowledgments

The authors gratefully acknowledge the support of the entire IMEC 5/7nm logic and advanced interconnect development teams, the IMEC IT team and the SEMulator3D development team. This work was guided by management support including An Steegen and Mike Jamiolkowski.

V. References

[1] J. Ryckaert et al: "DTCO at N7 and beyond: patterning and electrical compromises and opportunities", in Proc. SPIE 9427, IX, 94270C, March 2015

[2] S. Narasinha et al: “22nm High-Performance SOI Technology Featuring Dual-Embedded Stressors, Epi-Plate High-K Deep-Trench Embedded DRAM”, December 2012, pp. 52-55.

[3] B. Cipriany et al: "22nm Technology Yield Optimization Using Multivariate 3D Virtual Fabrication"; in: Proc. SISPAD 97-100, September 2013