SYNOPSYS®

WHITE PAPER

Meeting the Major Challenges of Modern Memory Design

Authors

Anand Thiruvengadam
Segment Management Director

Ricardo Borges
Director of Product Line Management

Guy Cortez
Product Management Manager

Farzin Rasteh
Principal Engineer

Jim Schultz
Product Management Manager

Overview

Memory lies at the heart of every electronics application, and demand is growing all the time. Users want ever greater capacity, throughput, and reliability. At the same time, time to market (TTM) goals and competitive pressures mandate that memories be developed in ever shorter project schedules. These requirements put enormous pressure on designers of discrete memory chips, memory dies in 2.5D/3D configurations, and memories embedded within system-on-chip (SoC) devices.

Among the many challenges of memory design, three stand out: scaling performance and capacity; ensuring silicon safety and reliability; and reducing development turnaround time (TAT). This white paper discusses these challenges, describes the requirements for viable solutions, and introduces the Synopsys approach to addressing the challenges. It presents four major areas of innovation in memory development: design technology co-optimization, memory design shift left, digitization of memory design, and design for reliability.

Scaling Challenges

Unlike in the past, the design and development of new memories is not independent of process development. With today's deep submicron technologies, much closer cooperation between the design and process teams is necessary to provide the required improvements in memory density and performance. There are several factors and trends driving this evolution:

- The slowing of Moore's Law means that memory designers can no longer count on regular, predictable benefits from scaling alone
- The end of Dennard scaling has led to early design/architectural optimization, detailed optimization of physical layout design rules, and new process recipes
- The slowing of supply voltage scaling and the increasing effect of leakage currents have limited reductions in device power at new nodes
- Bitline and wordline parasitics have an increased effect in DRAM arrays
- The need for sufficiently high storage capacitor values drives higher aspect ratio capacitor structures and the use of materials with higher dielectric constants
- DRAM scaling has become more challenging due to cell capacitance, cell contact resistance, and row hammer effects
- Scaling of DRAM and NAND periphery is increasingly impacted by process variability, which reduces the design margin for the sensing circuits
- The number of layers in 3D NAND devices has grown to around 200 and is projected to increase to more than 500, driving innovation in high aspect ratio etching processes and process techniques to improve channel conductivity

All these effects have produced a technology-design gap that has resulted in suboptimal devices and process recipes, suboptimal memory performance, and late-stage design changes that increase TTM. Minimizing this gap requires co-optimization of materials, processes, and device structures with target designs to ensure directional correctness, and this need will grow even stronger with emerging memory technologies.

Design Technology Co-Optimization

The collaboration between technology developers and designers required to address the scaling challenges is known as design technology co-optimization (DTCO). A memory DTCO flow must simulate the impact of technology choices and process variability on critical high precision analog circuits in the memory periphery, such as the sense amplifiers. This flow must include the following phases:

- Transistor modeling—technology computer-aided design (TCAD) simulates the fabrication process with its variability sources, followed by simulation of the transistor electrical characteristics and generation of data for subsequent extraction of a SPICE model
- Parasitic extraction—a 3D representation of the circuit is created, using as inputs a description of the interconnect process flow and a layout of the circuit element (for example a sense amplifier), and is fed to a parasitic field solver that extracts a circuit netlist and annotates it with RC parasitics
- SPICE simulation—the SPICE model and annotated netlist are simulated, to assess impact of variation on design metrics

This flow develops a virtual process development kit (PDK) that enables early and rapid design exploration before wafers in the new process are available. The tight fusion of TCAD and SPICE technology provides design enablement with high-quality models that can be further refined when wafers are available and fabrication data can be gathered. The layout can be created early from virtual PDKs, with power, performance, and area (PPA) assessed from both pre-layout and post-layout netlists.

Synopsys provides a memory DTCO solution that meets all these requirements, and more. As shown in Figure 1, the heart of this flow is Synopsys PrimeSim™ SPICE, a high-performance simulator for analog, RF, and mixed-signal designs including memories. The transistor modeling phase uses Synopsys Sentaurus™ Process, which simulates the transistor fabrication steps, Synopsys Sentaurus Device, which simulates transistor performance, and Synopsys Mystic to extract SPICE models from the TCAD output. The SPICE netlist is generated by Synopsys Process Explorer process emulation and the Synopsys Raphael™ FX resistance and capacitance extraction tool.

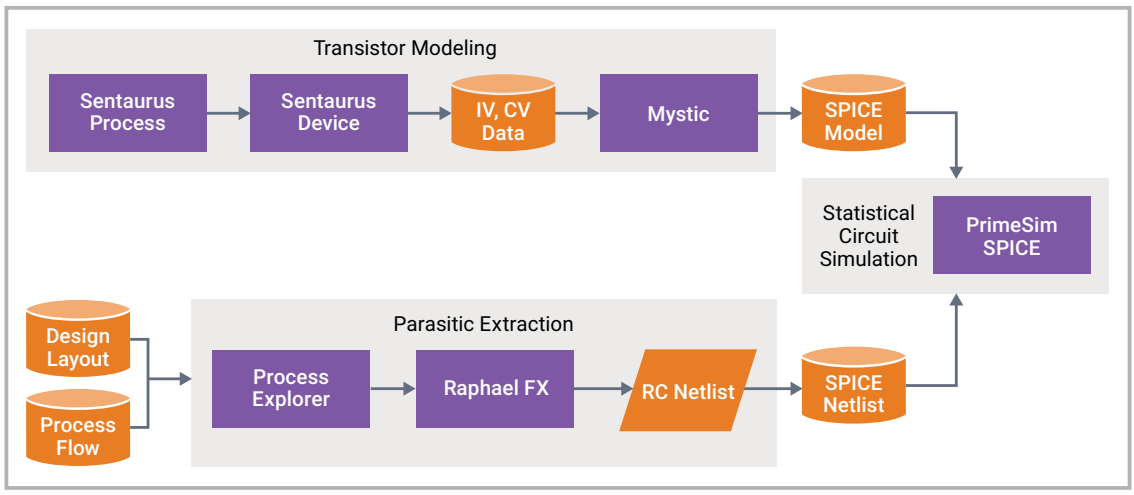

Figure 1: Synopsys DTCO flow for memory development

The DTCO solution also includes a data-to-design workflow that enables fab data to be directly consumed by Synopsys PrimeSim simulators for quick design PPA assessments. As shown in Figure 2, this allows process technologists and design engineers to skip the compact model extraction step, which can be cumbersome and time consuming for non-standard process technologies. Design engineers can perform a more complete PPA assessment with the traditional DTCO flow or the data-to-design workflow with early layout and post-layout simulations using Synopsys Custom Compiler, Synopsys PrimeWave Design Environment, and Synopsys PrimeSim simulators.

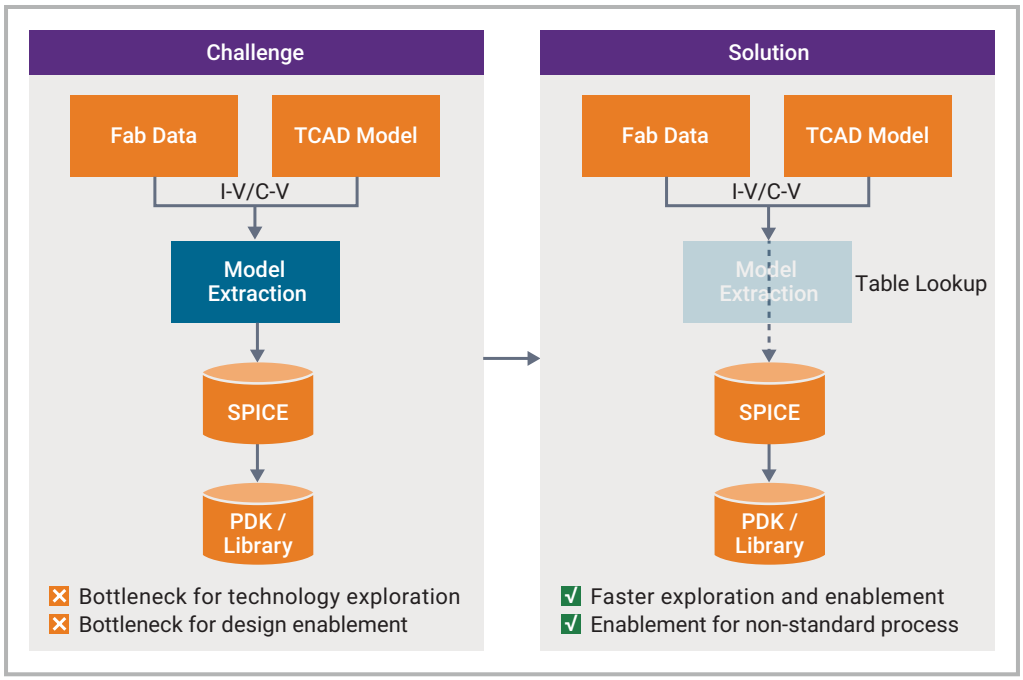

Figure 2: Synopsys data-to-design flow with TCAD-to-SPICE direct link

Schedule Challenges

As noted earlier, demand for more memory is a common theme for many semiconductor-driven products. Artificial intelligence (AI) and machine learning (ML) algorithms rely on fast, plentiful memory for real-time performance, and storage at all levels is key to data-intensive applications. General-purpose memory devices are giving way to customized chips for applications such as AI, servers, and automotive to meet specific performance, power, and bandwidth requirements. The need to produce derivative designs and variants quickly is increasing TTM pressures.

In response to these demands, memory devices are becoming larger and more complex, with aggressive power, performance, and area (PPA) goals. Memories are increasingly grouped into multi-die configurations such as multi-chip modules (MCMs) and 2.5D/3D structures, posing significant challenges to design, analysis, and packaging. For example, the complete memory array, including the interconnections between the dies and the power distribution network (PDN), must be considered while designing the most advanced memories to optimize for PPA and ensure silicon reliability.

Shift Left for Memory Design

The only way to address the schedule challenge is to "shift left" the memory design and verification process to perform better analysis earlier, avoid surprises late in the flow, and minimize iterations. This shift eliminates the key bottlenecks in memory development that impact overall TAT and TTM: macro cell characterization, the pre-layout to post-layout simulation gap, and custom layout.

Macro cell characterization requires Monte Carlo simulations, which traditionally have been a significant but manageable portion of the analysis stage for memory designs. However, with contemporary devices, the cost in time and resources for brute-force Monte Carlo is prohibitive, increasing TAT for memory development. Billions of simulation runs are typically required to achieve the desired high sigma characterization and ensure design robustness. Fortunately, this is a domain where ML can make a big difference.

As shown in Figure 3, highly accurate surrogate models of the design can be built and trained to predict high sigma circuit behavior, thus greatly reducing the number of runs required. Published case studies have shown that this approach can achieve speedups of 100-1000X over traditional methods while delivering accuracy within 1% of golden SPICE results.

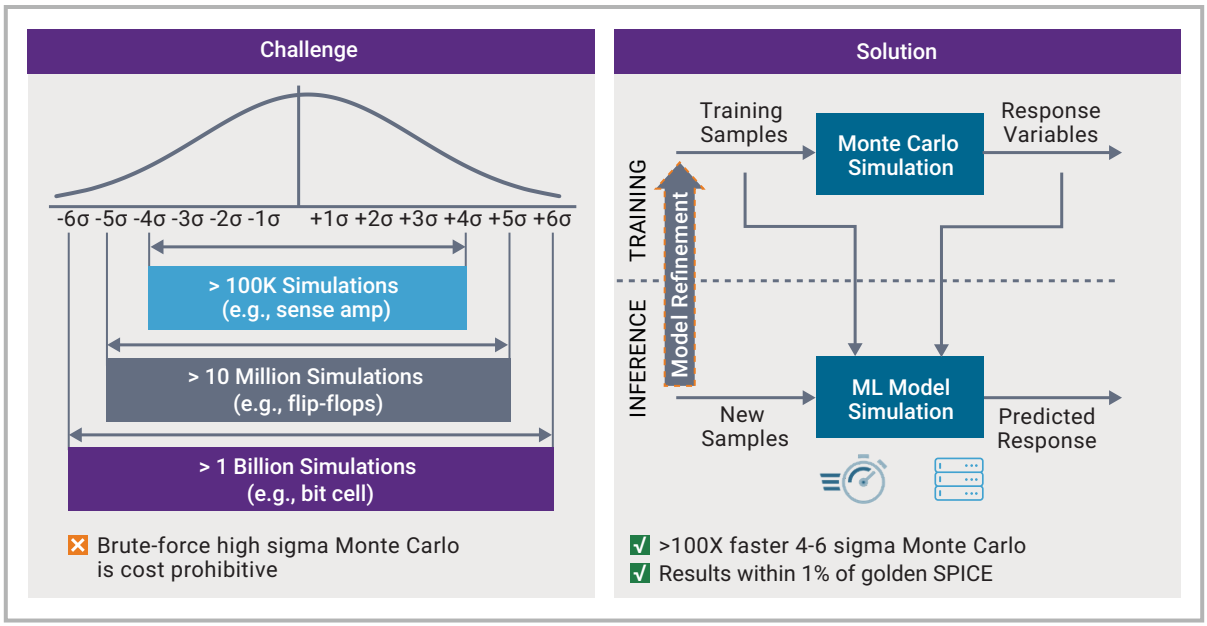

Figure 3: Elimination of brute-force Monte Carlo simulations

A major source of iterations that lengthen TAT and increase TTM is the gap between pre-layout and post-layout simulations. The goal is to pre-fetch the impact of parasitics on design specifications such as timing, power, noise, and stability as accurately as possible before layout to avoid unpleasant surprises when parasitics are extracted from the layout. With traditional flows, such surprises are common, resulting in repeated layout and simulation.

As shown in Figure 4, the solution is an early parasitic analysis workflow that allows for accurate estimation of net parasitics both for pre-layout and partial-layout designs. Published case studies indicate that using an early parasitic analysis workflow to pre-fetch parasitics reduced the gap between pre-layout and post-layout timing for designs from 20-45% down to 0-20%. Early parasitic analysis workflows can be further improved with the use of ML to predict interconnect parasitics.

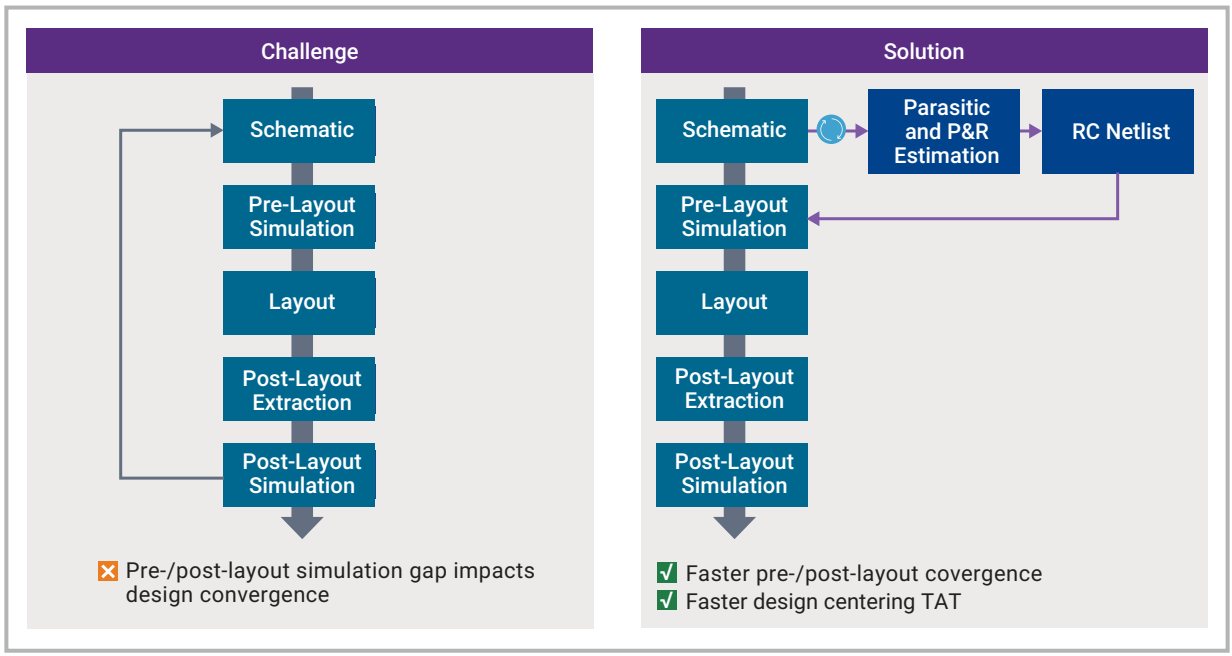

Figure 4: Elimination of pre-layout and post-layout gaps

To further shift left, there is opportunity to reduce the time and effort for the custom layout stage. The same sub-circuit topologies recur frequently in memory designs. As shown in Figure 5, the ability to reuse existing layouts created by expert designers is possible through the creation and application of templates that extract placement and routing patterns. Junior designers can create new layouts from those templates using whatever device size they need, saving time and leveraging the expert wisdom and experience embodied in the original layout.

Published case studies have shown that creating and using templates achieves more than 50% faster layout TAT for critical analog circuits in memories and produce more consistent layout quality regardless of the engineers' experience. The next frontier in layout design is the use of ML techniques to automate analog layout placement and routing, driving further improvements in layout productivity.

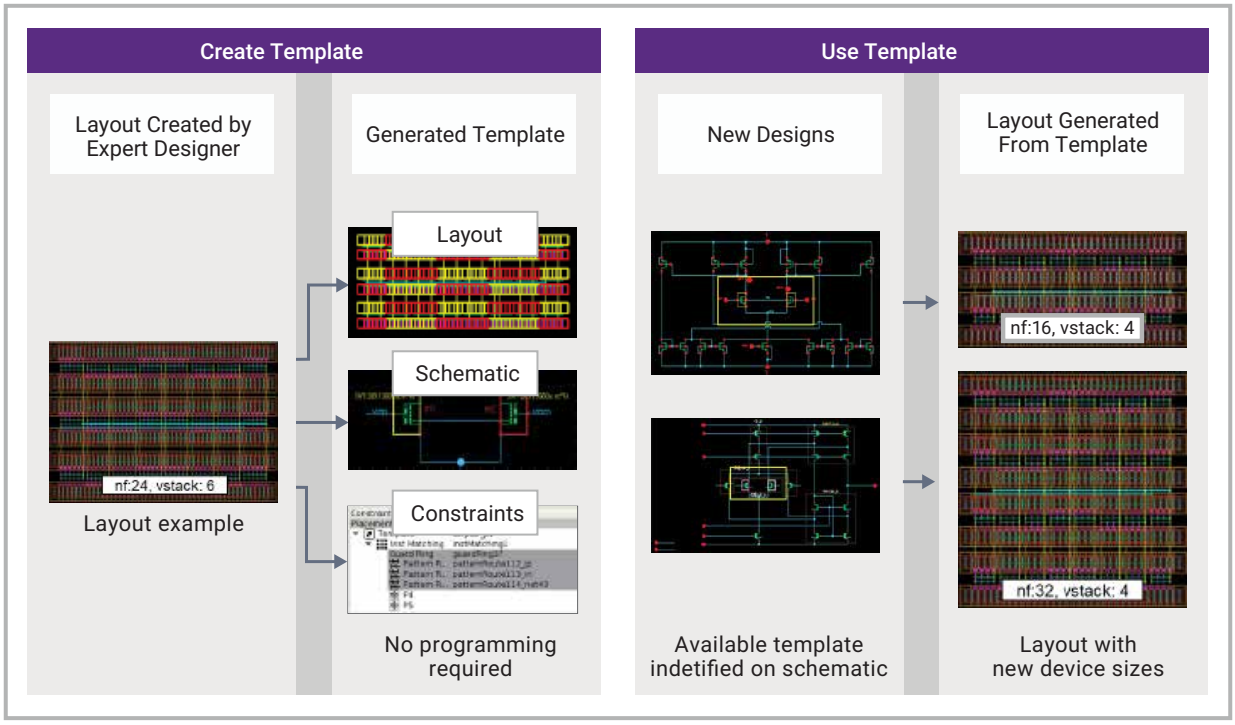

Figure 5: Use of templates to leverage expert knowledge

The Synopsys Custom Design Family satisfies these requirements and removes key memory design and verification bottlenecks. Synopsys PrimeSim, in conjunction with Synopsys PrimeWave Design Environment, provides a unified workflow across best-in-class circuit simulation technologies. This includes ML-driven high sigma Monte Carlo analysis and early parasitic analysis, eliminating the hassles and modeling inconsistencies inherent in point tool flows. The Synopsys Custom Compiler™ design and layout solution includes full support for template-based design reuse.

Digitization of Memory Design

The demands for more memory and more application-specific variants of memory chips are leading to "digitization" with the adoption of design and verification techniques well proven in the digital domain. While the core memory array continues to be developed using traditional techniques, much of the circuitry on the boundary of the array, beyond the sense amps, is closer to custom digital design than analog/mixed-signal (AMS). Applying digital techniques to the memory of the periphery is both logical and practical.

An effective solution for memory development digitization must have several elements. The digital design environment and the AMS design suite must be closely linked, enabling seamless design/place/route of the digital blocks in the memory periphery from within the custom design environment. Once the design is complete, timing-aware place-and-route of the periphery logic both automates a traditionally manual process and replaces tedious analysis loops with an integrated flow. By taking static timing into account during placement and routing, convergence to the desired PPA goals for the memory is faster and more predictable. Human layout experts can spend their time focusing on the core array and not on the periphery.

On the verification front, a digital-on-top flow enables efficient verification of memory datapaths using co-simulation and digital testbenches. By using a digital abstraction of memory datapaths and selectively switching to analog views for critical blocks and time periods during simulation, datapath verification TAT can be greatly improved. There are other benefits of this hybrid flow, including the use of AMS-level noise and signal integrity analysis coupled with static timing. The result is signoff quality characterization, validation, and verification of the complete memory chip, both core array and peripheral logic.

Synopsys provides a complete, robust solution for the design and verification of memories, including digitization of key stages in the flow. Co-design of the digital and AMS portions is provided by Synopsys Custom Design Family and Synopsys Digital Design Family. Designers can take advantage of digital implementation techniques where possible, while not sacrificing the hand-optimized layouts for memory cells and sense amps.

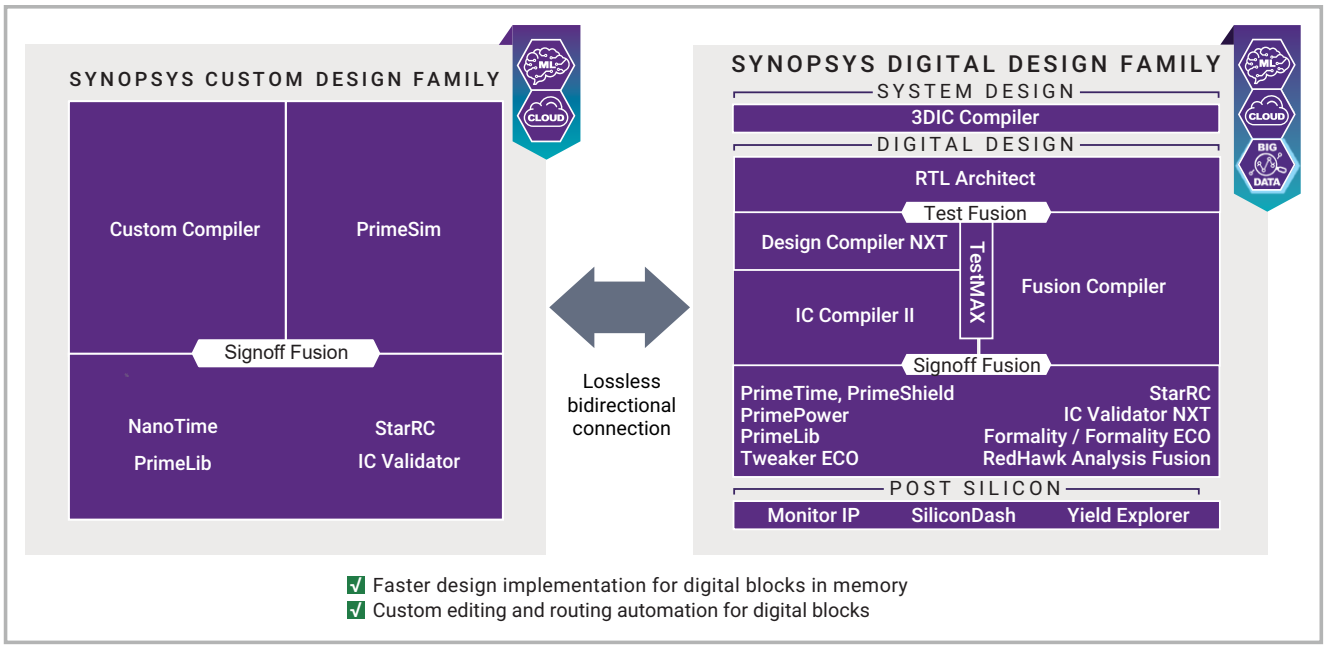

Figure 6: Synopsys memory design and verification solution

Using Synopsys Custom Compiler, layout engineers can define the floorplan for the memory chip and then manually place critical cells or nets. As shown in Figure 7, the rest of the periphery logic can be automatically placed and routed using Synopsys Fusion Compiler™ or Synopsys IC Compiler™ II. This reduces the layout time from days to hours without reducing the QoR. In addition, clock tree synthesis automates the traditionally tedious process of clock hookup. When engineering change orders (ECOs) must be applied to the design, this step takes minutes rather than hours. In addition, Synopsys is working closely with major memory vendors to ensure that the timing-aware place-and-route process takes reliability and aging effects into account.

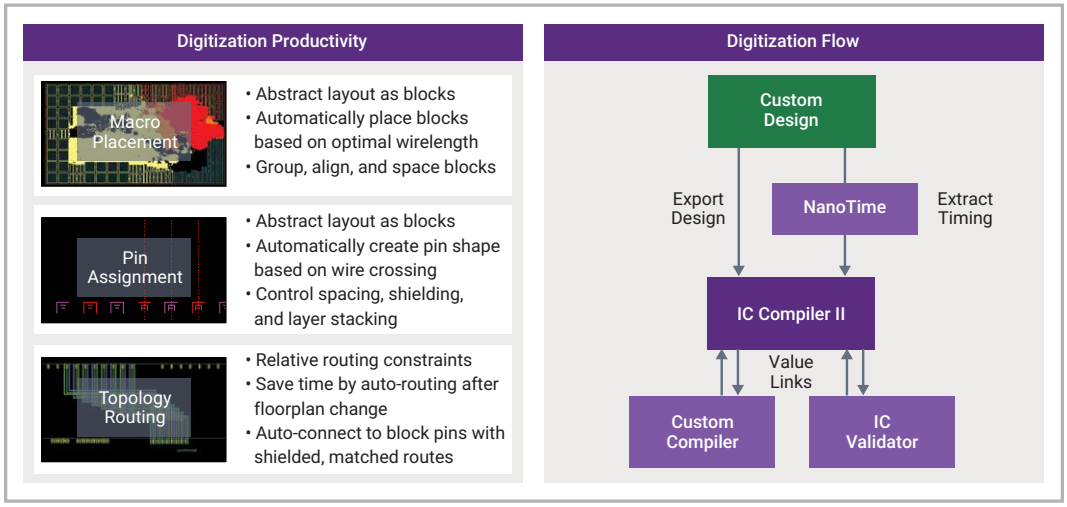

Figure 7: Synopsys digitization of memory development

For verification, Synopsys PrimeSim provides a unified workflow of next-generation simulation technologies, from gold-standard SPICE to FastSPICE. Co-simulation with Synopsys PrimeSim and Synopsys VCS® combines analog and digital simulators to offer high performance mixed-signal analysis. As shown in Figure 8, it also offers the capability to swap digital and analog views back and forth dynamically within one memory co-simulation. That allows running the simulation fast by using digital views for blocks and swapping some of those views with SPICE only during the simulation periods when more accuracy is needed. This capability, called Real Time View Swapping (RTVS), enables verification engineers balance accuracy and runtime speed.

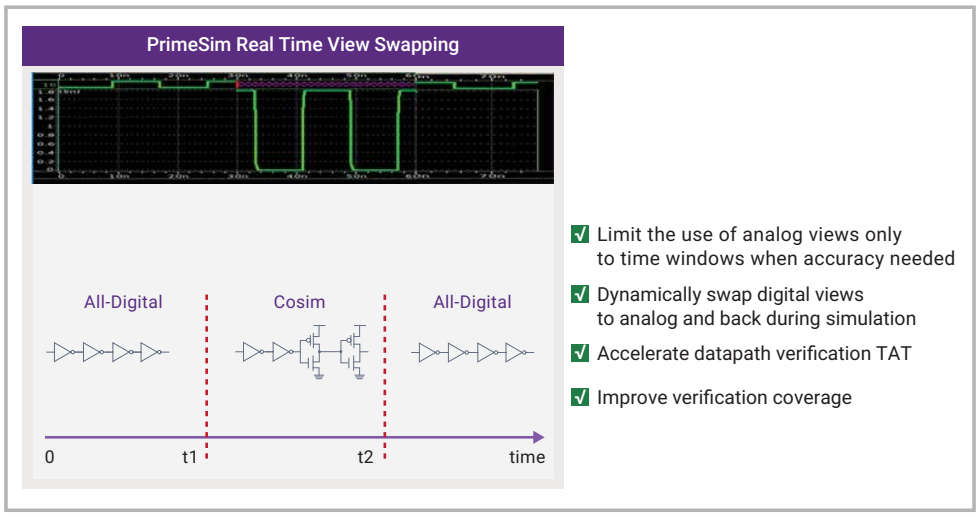

Figure 8: Synopsys PrimeSim RTVS capability

Safety and Reliability Challenges

As discussed previously, designing and verifying memories is a major portion of many projects. Safety-critical applications such as autonomous vehicles, space-borne systems, implanted medical devices, and nuclear power plants are no exception. The ICs powering these applications contain a lot of memory, and the memory technology used must meet the same high reliability and functional safety standards as the rest of the electronics.

Functional safety requires building in safety mechanisms to detect faults in electronic devices and respond appropriately, as well as calculating that this detection and response produce a high degree of fault coverage. This approach is mandated by safety standards such as ISO 26262 for road vehicles. Reliability demands that the chances of a fault occurring be reduced as much as possible by design and manufacturing.

Both safety and reliability must span the entire silicon lifecycle, from design and verification through lab bring-up all the way to production use in the field. In the case of memory designs, the early and late stages of the lifecycle present the greatest challenges for reliability. Early chip failures (sometimes called infant mortality) shake out marginal devices, followed by a period (perhaps years) of low-risk operation. As silicon aging effects start to kick in, reliability goes down and failures become more common.

Ensuring Memory Safety and Reliability

To ensure reliability throughout the silicon lifecycle, the memory development process must include robust static and dynamic analysis to identify and mitigate potential failures across the silicon lifecycle before tapeout:

- Early life
  - Static analog and digital circuit checks
  - Analog fault simulation
- Normal life
  - High-sigma Monte Carlo analysis
  - Static power/signal net resistance checks
- End of life
  - Dynamic electromigration/IR drop (EMIR) analysis
  - Silicon aging analysis

An industry-leading solution meeting all these requirements is available with the comprehensive and flexible Synopsys PrimeWave Reliability Environment. It delivers a unified workflow around all the analysis technologies of Synopsys PrimeSim Reliability Analysis and the engines of Synopsys PrimeSim to improve productivity and ease of use. As shown in Figure 9, the process starts with Synopsys PrimeSim CCK, which extends traditional electrical rules checking (ERC) into the analog domain.

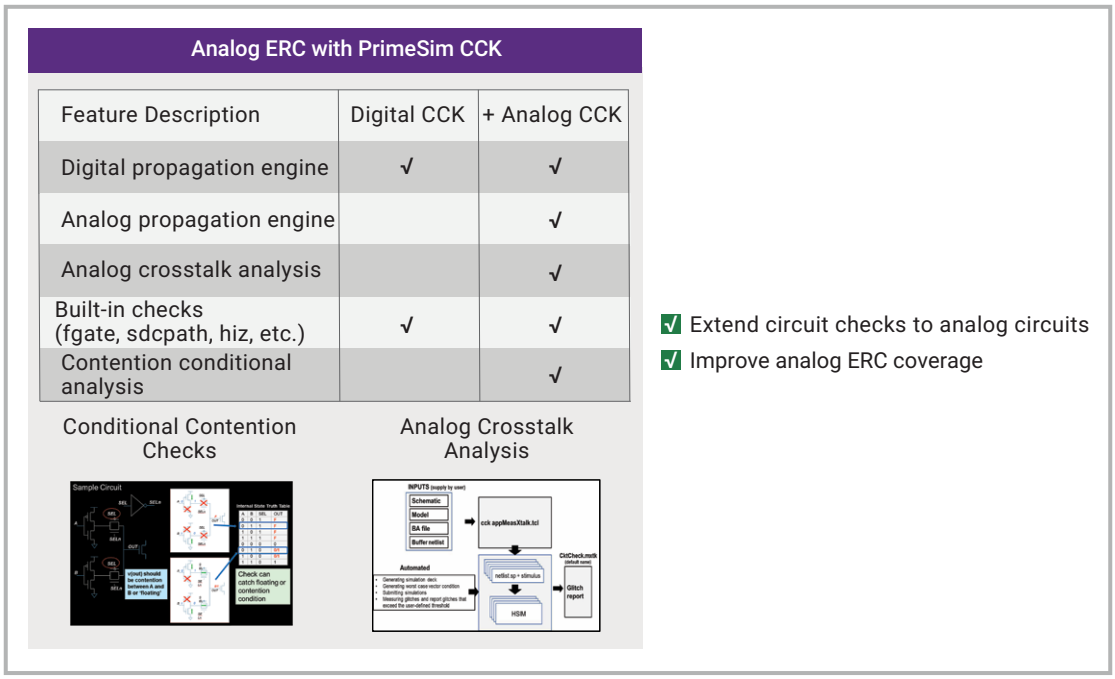

Figure 9: Synopsys solution for analog electrical rules checking

As shown in Figure 10, Synopsys PrimeSim Custom Fault is the industry's leading analog fault simulation solution. It complements digital fault simulation to make functional safety and test coverage analysis practical for complete chips. It satisfies even the demanding requirements of ISO 26262 and other safety standards for complex and comprehensive failure modes, effects, and diagnostic analysis (FMEDA).

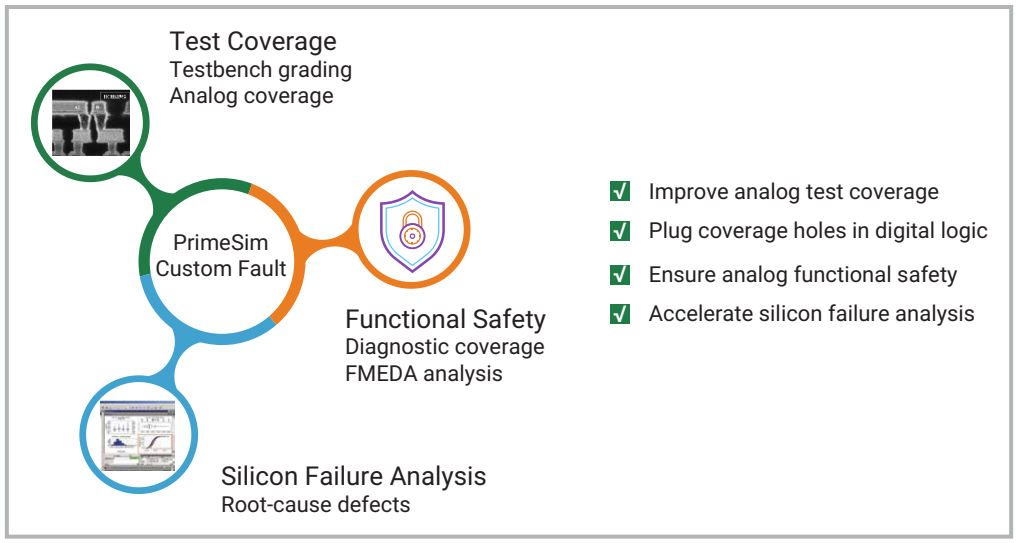

Figure 10: Synopsys solution for analog fault simulation

Synopsys PrimeSim Design Robustness provides high-sigma (typically 4-7) Monte Carlo analysis. It uses ML techniques to run more efficiently while delivering accuracy to within 1% of Synopsys PrimeSim HSPICE, the industry's "gold standard" circuit simulator. ML reduces the number of runs by orders of magnitude over the traditional brute-force Monte Carlo simulation approach.

Power/ground integrity analysis is provided by Synopsys PrimeSim SPRES, which is fast enough to run early in the memory development process. Similarly, Synopsys PrimeSim EMIR provides both high performance and foundry-certified signoff accuracy. As shown in Figure 11, this analysis covers the power distribution network (PDN) as well as the signals in the memory design. If issues are uncovered, what-if analysis and debug hints make it easier to find and fix the source of potential faults as the silicon ages.

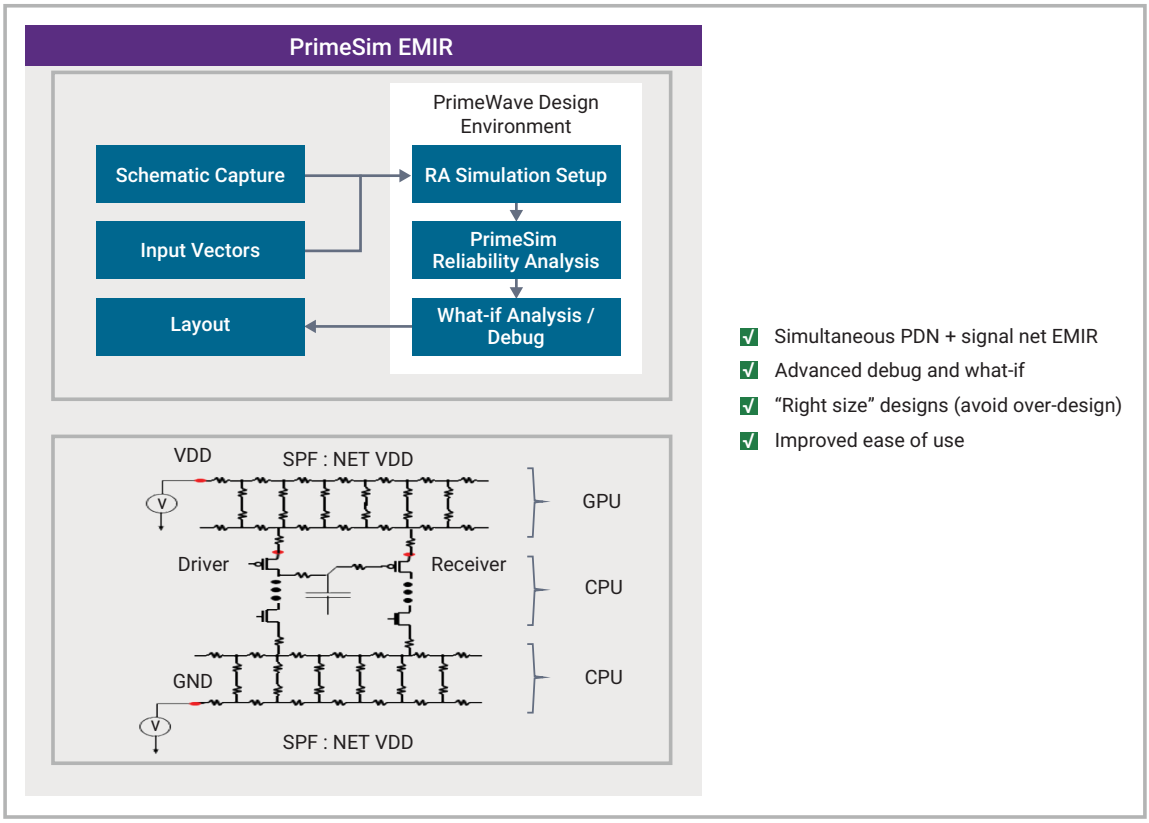

Figure 11: Synopsys solution for analog EMIR analysis

Synopsys PrimeSim MOSRA checks for reliability risks due to silicon aging effects. It also offers high performance with foundry-certified accuracy. When combined with the other Synopsys PrimeSim Reliability Analysis technologies within the Synopsys PrimeWave Reliability Environment, as shown in Figure 12, memory designers can be sure that their chips will be functionally safe and reliable throughout a long and productive silicon lifecycle. This solution is complementary to the Synopsys integrated Silicon Lifecycle Management (SLM) family of products.

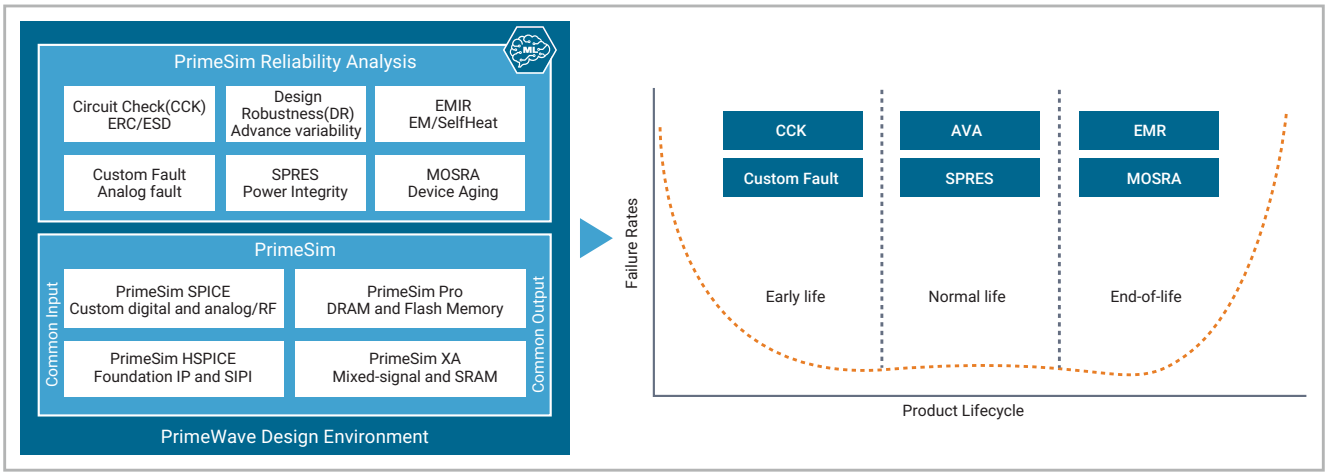

Figure 12: Synopsys solution for reliability across the silicon lifecycle

Summary

Demands for more memory, faster memory, and more reliable memory are growing rapidly. At the same time, TTM and project schedules are shrinking. Old methods such as designing independently of process development and requiring manual analysis loops no longer suffice. Synopsys provides a comprehensive solution for memory development that meets key challenges using differentiated technologies. The key techniques of design technology co-optimization, memory design shift left, digitization, and design for safety and reliability ensure a robust solution that will meet the needs of every memory designer.