MEPTEC Luncheon Wednesday, November 13, 2019

SoC Design Methodology Challenges for Advanced Process Nodes

Tom Dillinger

semiwiki.com info@vlsidesignmethodology.com



Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges
- Final Thoughts



Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges
- Final Thoughts



Moore's Law and Dennard scaling

- "Moore's Law": forecasts transistor density over time
 - assumption is that cost per transistor also improves







Moore's Law and Dennard scaling

- Dennard scaling
 - defines how transistor current, power density, circuit density, and performance scale with transitions in lithographic processes
 - original scaling assumptions were based on:
 - field-effect devices
 - constant electric field across gate oxide, device node-to-substrate junctions
 - a "constant voltage" derivative was introduced later, due to the need to maintain the supply voltage VDD above scaled values

Dennard Technology Scaling (1972)

Device or Circuit Parameter	Scaling Factor			
Device dimension tox, L, W	1/k			
Doping concentration N _a	k			
Voltage V	1/k			
Current /	1/k			
Capacitance eA/t	1/k			
Delay time per circuit VC/I	1/k			
Power dissipation per circuit VI	1/k ²			
Power density VI/A	1			

Table I: Scaling Results for Circuit Performance (from Dennard)





ITRS Process Roadmap

- Moore's Law and Dennard scaling led to the "ITRS roadmap"
 - specific process "nodes" were defined, based on a lithographic measure
 - "half" nodes were also developed, as a "performance kicker" to production parts

• the half-node goal was to minimize the physical layout changes, and focus on analysis of critical timing paths (e.g., 130nm \rightarrow 115 nm, a 0.9X lithography "shrink")

Characteristic	1992	1995	1998	2001	2004	2007
Feature size (microns)	0.50	0.35	0.25	0,18	0,12	0.10
Gates per chip (millions)	0.3	0.8	2.0	5.0	10.0	20.0
Bits per chip						
DRAM	16M	64M	256M	16	4G	16G
SRAM	4M	16M	64M	256M	1G	4G
Wafer processing cost (\$/cm²)	\$4 00	3.90	3.80	3.70	3.60	3.50
Chip size (mm')					0.00	
logic	250	400	600	800	1,000	1,250
memory	132	200	320	500	700	1,000
Wafer diameter (mm)	200	200	200-400	200-400	200-400	200-400
Defect density (defects/cm ²)	0.10	0.05	0.03	0.01	0.004	0.002
Levels of interconnect (for logic)	3	4-5	5	5-6	6	6.7
Maximum power (watts/die)						
high performance	10	15	30	40	40-120	40-200
portable	3	4	4	4	4	4
Power supply voltage						
desktop	5	3.3	2.2	22	1.5	1.5
portable	3.3	22	3.2	1.5	1.5	1.5

The first ITRS roadmap (1993)

Revisions/updates were released on a regular basis.

The roadmap definition was overhauled in 2014.



ITRS Process Roadmap



- reductions in the lithography exposure wavelengths supported Moore's Law
- another major innovation was the introduction of Chemical-Mechanical-Polishing (CMP), which planarized the wafer surface after (damascene) metal deposition and patterning
- the number of available metal interconnect levels in subsequent process nodes increased dramatically (over the ITRS projections)



And, then things began to slow down...

- R*C interconnect delays began to dominate logic stage timing → Dennard performance scaling (based on transistor optimizations) was reduced
- lithographic fidelity was increasingly difficult to achieve
 - a source wavelength of 193nm (with liquid immersion) exposure through the optical reduction path was being used to expose minimum features on the wafer photoresist
 - numerous "optical correction" methods were applied to the original layout design to generate the mask data: "serifs", sub-resolution assist features, and ultimately "computational lithography" (with a non-uniform illumination source)





"Moore's Law is over"... well, not quite

- The traditional Power/Performance/Area (PPA) advantages of Dennard scaling are significantly reduced.
 - the transition to FinFET device topologies at the 22/20nm process node provided a PPA boost
 - subsequent FinFET process nodes are back to (interconnect-limited) performance gains, due to resistance increases from wire scaling
- New physical layout design requirements have enabled additional process node evolution and related transistor density scaling



- unidirectional devices and wires
- restrictions on device length, contact and via sizes
- multipatterning decomposition of a mask layer into multiple subsets

 (e.g., 2 "colors" of mask data utilizing a Litho-Etch-Litho-Etch sequence)

Source: Intel



Cost/transistor scaling is over...

- Although transistor density continues to improve with the 10/7/5/3nm process node roadmap, the transition to multipatterning lithography and (more recently) EUV lithographic exposure adds significantly to the wafer processing cost.
- The traditional cost/transistor "scaling" that has been associated with Moore's Law is no longer applicable.



Heterogeneous packaging technology integrating multiple die offers an opportunity to maintain "system-level" cost reductions.



Source: TSMC

"More than Moore" heterogeneous packaging

- Multi-die package integration offers an opportunity to leverage process scaling for critical parts of a system design, maintaining other functionality in older (more cost-effective) process nodes.
- Heterogeneous packaging also offers an opportunity to bring a large amount of memory "closer to" the logic (e.g., HBM "stacks").





Source: Samsung

- 2.5D (with silicon interposer)
 - interposer size originally limited to 1X max reticle exposure field, moving to 1.5X and 2X max reticle



Chip-on-Wafer-on-Substrate (CoWoS) 2.5D packaging cross-section. (CoWoS is a registered trademark of TSMC.)

Source: TSMC



- fan-out wafer-level (or panel-level) packaging
 - extends redistribution fanout beyond die area
 - through vias from redistribution metal on reconstituted base die carrier enables a stacked silicon die configuration



Source: TSMC



- "embedded" bridges between die
 - connections between die provided by embedded silicon "chiplet", embedded within the top of the organic package substrate
 - requires design collaboration on interface definition, pad definition, and pad placement in the rows of pads between adjacent edges of die
 - e.g., 3mm max dimension for embedded bridge







Source: Intel

- "bumpless" thermo-compression bonding between die
 - face-to-back or face-to-face orientations possible "3D stacking"

Thermo-compression bonding ("bumpless") for TSMC SoIC 3D stacking



front-side of the top die is attached to a carrier, and thinned to expose the Cu "nails" (through silicon vias)

 Y.H. Hu, et al., "Cu-Cu Hybrid Bonding for 3D IC Stacking",
 2012 IEEE International Interconnect Technology Conference





front-side to back-side TCB attach topology is shown

front-side to front-side is also an option



future roadmap merges "front-end" die-to-die bonding with "back-end"
 2.5D or fan-out redistribution package technologies



Source: TSMC



MEPTEC – KGD Workshop

- Be sure to attend the MEPTEC KGD Workshop
 - <u>https://www.kgdworkshop.org/</u>
 - Thursday, December 12th





Application Markets \rightarrow Foundry "platforms"

- The advanced silicon and packaging technology offerings need to meet the requirements of the key application markets.
- The foundries have focused on design enablement for various "platforms":
 - Automotive, Mobile/Consumer, High-Performance Computing (HPC), IoT (including "Industrial" IoT)
 - Mil-Aero and medical markets get special attention
 - specific cell libraries and IP developed for low-power or high-performance design targets; additional transistor Vt options available for LP or HP
- New transistor and memory alternatives in development
 - stacked "nano-sheets" replacing vertical FinFET's
 - (Spin-Transfer Torque) Magneto-Resistive RAM storage cells (MRAM)
 - silicon photonics





Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges
- Final Thoughts



SoC Design Methodology Challenges -- Multipatterning

- Multipatterning decomposition was introduced at the 20nm node (for layout pitch < 80nm using 193i lithography)
- Significant EDA algorithm development required

(e.g., routers must avoid cyclic errors, DRC, parasitic extraction)

 Methodology question: foundry-assigned or user-assigned decomposition? (user-assigned emerged, as the electrical characteristics of decomposition masks "A" and "B" may differ)



Multipatterning decomposition "cyclic error"

(odd number of shapes at minimum spacing for two-color assignment)



SoC Design Methodology Challenges -- Multipatterning

- Subsequent nodes added more stringent design requirements e.g., unidirectional wires
- Multipatterning evolved to "pre-assigned" tracks
 - Routes and cell pins aligned with the track assignment
- New assignment rules for the router to manage "wide" wires (1.5X, 2X, 3X)







SoC Design Methodology Challenges – Design Rules

- Sub-micron process nodes included "recommended" layout design rules, as part of "Design for Yield" optimizations
 - RDR's rapidly became "required" design rules
 - EDA composition tools were required to implement "fill" algorithms for lithographic and process uniformity
 - Chemical-Mechanical Polishing requires a uniform metal distribution
 - photolithographic exposure, material etch rates, and silicon crystal growth process steps require extremely high uniformity





Implementing Engineering Changes on a design with metal fill requires specific methodology flow development.



SoC Design Methodology Challenges – Design Rules

- Resolution of end-to-end spacing for wires has become especially problematic at advanced nodes
 - new "cut masks" were introduced, with sufficient area to be resolved
 - requires alignment of distinct wire ends to add cut mask shape routers need to support new rules





SoC Design Methodology Challenges – Litho Checks

- Optical proximity corrections evolved from serifs to added SRAF data to computational lithography (aka, "Inverse Lithography Tech", or ILT)
- Layout design rules were added to try to ensure a successful exposure solution, a combination of mask data and the illumination pattern
 - new "forbidden pitch" rules were added as a restriction
- The design methodology required an additional step –

Lithography Process Checking (LPC)

- a combination of "pattern checks" and illumination model analysis
- very computationally-intensive



SRAF: "Sub-Resolution Assist Features"

wafer image

imec

mask data w/ILT

SoC Design Methodology Challenges – LDE

- With advancing process nodes, additional steps were used to improve transistor performance:
 - the addition of "stressor" materials to improve carrier mobility
 - the use of "shallow trench isolation" between devices
 - also, spacing between (high implant) wells and devices were scaled
- As a result, (second-order) *layout dependent effects* were identified and needed to be modeled



SoC Design Methodology Challenges – LDE

- How to characterize library cells with LDE?
 - Design enablement teams need to include a representative "context" layout cell prior to electrical parasitic extraction (and DRC) to estimate the final placed environment.
 - For improved model accuracy, the cell layout design may itself include additional surrounding (non-functional) shapes data for LDE measures.





SoC Design Methodology Challenges – Self-Heating Effect

- FinFET and FD-SOI processes have a restricted thermal flow from the transistor
 - transistor "self-heating" raises the temperature of neighboring contacts and wires
 - a new analysis flow is required to estimate the delta-T temperature increase for thermally-sensitive analysis (esp. electromigration)





SoC Design Methodology Challenges – Thermal maps

- The increased transistor density (and clock frequencies) results in an increased (local) power dissipation.
- A thermal analysis model is required to evaluate the temperature increase across the die (a "thermal map").



Die thermal map and power dissipation model

Source: Ansys

• A mechanical analysis is **required** for the die + package attach to evaluate the risk of material fatigue and fractures/delamination.



SoC Design Methodology Challenges – power/clock domains

- The number of distinct power and clock domains on leading SoC's has increased dramatically.
 - Designers provide power domain information separately from the functional description a "UPF" side file.
 - Methodology flows need to interpret the UPF to insert (and verify) the appropriate power domain control and interface logic.



- Equivalency verification of the (design + UPF) vs. final implementation is a new, critical flow.
- The electrical switching

 "noise transient" when a
 power domain transitions
 from a sleep-to-active
 state requires unique
 analysis.



SoC Design Methodology Challenges – design data management

- Current SoC designs incorporate ~20B devices, and trillions of extracted parasitic elements.
- EDA vendors are addressing how to work with this data volume:
 - new, multi-threaded algorithms
 - (cloud-enabled) partitioned algorithms executing across many servers
 - map-reduce methods for data aggregation and results reporting





Source: Ansys

SoC Design Methodology Challenges – "The Productivity Gap"

- Functional validation is the largest resource investment in any design methodology.
- A "productivity gap" refers to the increase in scope for system models, without a corresponding increase in design abstraction + simulation performance.
- A myriad of new system-level modeling languages have been introduced, with support to synthesize a hardware implementation model – e.g., MATLAB, SystemC (with a time reference), C/C++, TensorFlow/Caffe.
- Emulation and prototyping products have been developed for improved throughput, with the capacity for large system-level model validation.
- A "digital twin" model includes an executable representation of sensor + actuator I/O, enabling complex system simulation e.g., an ADAS or autonomous vehicle.





Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges
- Final Thoughts



Heterogeneous Packaging Methodology Impacts

- EDA vendors often have separate products for chip and package physical design.
- Vendors are working on a consistent "connectivity" model between platforms.



Source: Cadence



Heterogeneous Packaging Methodology Impacts

- Foundries and OSAT's are releasing "design kits", with runsets and techfiles for:
 - physical design rule checking (DRC)
 - parasitic extraction (stack-up materials electrical parameters)
- EDA vendors provide model extraction for (time-domain and frequency-domain) model generation and simulation
 - extraction includes: R, C, L, and k elements
 - simulation is applied to the VDD + GND network for noise analysis
 - simulation is applied to signal interfaces (both parallel and serial busses)



Detailed models are extracted for simultaneous switching noise analysis on VDD/GND networks, and transmit/receive circuit behavior for highspeed signal interfaces.



Heterogeneous Packaging Methodology Impacts

- Cross-vendor sourcing requires qualification of materials compatibility
 - die attach metallurgy
 - underfill encapsulation
- By and large, a "solved" problem...



Example: Intel EMIB with AMD and HBM memory stack die



Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges

A "Top 10" list... with apologies to David Letterman

• Final Thoughts



10

• Split manufacturing

- There is no U.S. "trusted foundry" committed to process development beyond the 14/12 nm node.
- How do advanced SoC and heterogeneous package designs achieve "trusted" status?
- "Split manufacturing" uses FEOL fabrication overseas, incoming validation efforts, and BEOL fabrication at a trusted foundry (using existing interconnect processes).
- Active area of methodology research a comprehensive proof the non-trusted silicon does not include any malicious functionality (e.g., a "hardware Trojan")





9

• Failure analysis and diagnostic strategies for new device types + structures

- Advanced process nodes will be incorporating:
 - spin-transfer torque magneto-resistive memory cells (MRAM)
 - integrated silicon photonics structures
- There is little failure mechanism data, focused test strategies, diagnostic analysis methods, and/or long-term reliability history for these new structures.
- Designers need assess how to integrate these features in (high-rel) applications.





8

- electromagnetic coupling model extraction and analysis
 - Designers are familiar with (short-range) capacitive "crosstalk" between nets.
 - Designers are not as familiar with (long-range) inductive coupling, and how the coupled electromagnetic energy may impact distant design blocks.
 - esp. design blocks that include inductors, such as VCO's and PLL's
 - New methodologies are being developed to:
 - identify (and prioritize) potential sources of electromagnetic coupling
 - recommend design modifications to reduce coupling
 - provide (frequency-domain and time-domain) extracted models of interconnects to merge with critical circuits for simulation
 - Still early, although preliminary results are enlightening



Plot of power-supply noise induced clock jitter

- horizontal axis is time
- vertical axis is the simulated clock jitter each cycle (ranges from +50 to -120 fsec)
- yellow curve is with C_coupling only model
- red curve is with full EM coupling

Source: Nvidia



7

- testing of "resilient" system designs (for cost-sensitive markets)
 - Automotive Class 5 designs must include fail-safe functionality.
 - EDA vendors provide "test insertion" features for structured DFT architectures e.g., scan-based test, MBIST.
 - Verification of TMR voting insertion has been extensively studied, to ensure proper logic behavior.
 (Borg and LaBol, "Varification of Triple Medular Bodundancy (TMB) for

(Berg and LaBel, "Verification of Triple Modular Redundancy (TMR) for Reliable and Trusted Systems", NASA.)

- Yet, how will reliable systems be (economically) tested at wafer and package-level to detect the presence of a manufacturing defect?
- Active research area: e.g., design of voting logic that also detects and propagates (traditional) test faults



- testing of "resilient" system designs (for cost-sensitive markets)
 - examples of modifications to TMR logic for improved fault detection Arifeen, et al, *"A Fault Tolerant Voter for Approximate Triple Modular Redundancy"*, Electronics, 29 January 2019.





- 6
- burn-in screening parameters for advanced process nodes
 - What are the appropriate burn-in procedures for 5/3 nm processes?
 - static burn-in sufficient? dynamic burn-in required?
 - What are the infant fail mechanisms at these advanced nodes?
 - new structures: Gate All-Around (GAA) devices, MRAM arrays
 - What are the infant fail mechanisms for advanced 3D packages?
- *new HTOL qualification procedures required?*
 - new materials used for underfill have unique CTE and modulus of rigidity

behavior



5

- reliability analysis
 - FIT rate calculation is currently based on "independent" failure rates
 - example: electromigration is the summation of (infinitesimally small) failure projections for billions of wires and vias still valid?
 - SER rates for new structures (e.g., MRAM arrays)
- "aging" models for device and interconnect behavior not typically incorporated into design flows
 - example: NBTI/PBTI/HCI device degradation and delta_R mechanisms are very dependent upon operating history
 - design techniques for "adaptive aging" required?



An SRAM bit cell is extremely sensitive to a combination of pFET NBTI and nFET PBTI, resulting in a much weaker write current state transition.



4

- electrical/thermal/mechanical analysis of "complex" systems
 - requires close collaboration among engineering disciplines, avoiding team *silos*
 - examples span a huge range of model complexity: die delamination issues, package/board solder joint fatigue, full electromagnetic analysis of cellular 4G/5G antenna radiation patterns



- 3
- improving designer productivity through high-level (abstract) system modeling/simulation/synthesis/equivalency verification
 - A number of design description languages are available, for different applications e.g., (untimed) C, SystemC, MATLAB, TensorFlow/Caffe.
 - Specific simulation tools are available for optimum throughput. (Testcase development and model coverage measurements are a bit ad hoc.)
 - EDA vendors provide a "high-level" synthesis flow from these languages to RTL.



3 (continued)

- improving designer productivity through high-level (abstract) system modeling, simulation, synthesis, and equivalency verification
 - Yet, these designs require additional designer input ("side files") to guide:
 - assignment and scheduling of operations to allocated on-chip resources
 - implementation of clock and power domains
 - insertion of test logic
 - Once this design detail is incorporated into the physical chip implementation, it is **extremely** difficult to apply an engineering change:
 - to the abstract model, within minimal impact to the physical design
 - to the implementation (ensuring functional equivalency to the abstract model)

UPF side file for power/clock domain definitions





- 2
- utilization of "digital twin" models and simulation data as an alternative for post-silicon system assurance and field testing
 - Design methodologies for complex systems are increasingly using simulation acceleration technology (e.g., emulation or prototyping).
 - Digital models are being developed for (libraries of) sensors, actuators, radar/lidar, antennas, GPS, electromechanical and fluid systems:

a full system "digital twin" can be simulated

• Testcases can be developed to represent input to the sensing subsystems.





- 2 (continued)
- utilization of "digital twin" models and simulation data Example: ISO 26262 ASIL Class 5 vehicles
 - "Achieving Class 5 autonomous vehicle qualification requires driving over 8B miles... that's 1,000 cars averaging 25mph, driving 24/7/365 for 40 years."
 - "How will vehicle qualification be done in a manner that covers all different types of traffic, weather patterns and road conditions? V2V /V2I communications, too."
 - "How will situations be tested that involve potentially dangerous results with other vehicles and/or pedestrians?"
 - "How is the accuracy of the transmission/engine/braking/steering subsystem models confirmed?"





Source: Qualcomm

- 2 (continued)
- utilization of "digital twin" models and simulation data as an alternative for post-silicon system assurance and field testing

Other examples:

- factory automation operations
- drone-based materials delivery
- robotic surgery



• any mission-critical system (esp., utilizing AI-based inference methods)

A key challenge:

"How will coverage of the qualification test suite be measured?" And, ultimately:

"Will government agencies accept (partial) 'digital twin' validation as sufficient for regulatory approval?"

("All models are wrong, but some are useful." - George Box)



1

- system partitioning and implementation definition methodologies "Early" product planning and definition involves:
 - PPA assessment (for a monolithic die/metal stack, in an advanced process node)
 - Cost analysis for advanced heterogeneous die packaging alternatives
 - evaluation of available IP for heterogeneous integration



- Reliability estimation against FIT targets
 - electromigation analysis, SER estimates, aging simulations
 - thermal/mechanical analysis of the proposed die/package/board/enclosure
- Security analysis, ensuring the system functionality achieves the appropriate level of protection against intrusion attempts
 - malicious external methods (software, electromagnetic)
 - internal hardware "trojans" (from non-trusted sources)
- Safety design and verification
 - design, verification, and test to meet redundancy/resiliency requirements

- 1 (continued)
- system partitioning and implementation definition methodologies
- There are currently no EDA tools that readily assist system architects with this planning and partitioning task – aka "pathfinding".
- System architects will need much more extensive (and early) engagement with a broad cross-section of the engineering team.
 - PPA is no longer the norm
- There is a broad opportunity to address the tool needs to assist with the initial definition of the functional implementation, incorporating electrical, thermal, and mechanical analysis, with alternatives for SoC process fabrication and advanced packaging technology.
 - PPA, Cost, Reliability, Security, Safety \rightarrow PPACRSS ("Packers")



Bonus Challenge

- managing statistical variation in chip/package/board fabrication
 - Increasingly, application markets are requiring "high-sigma" verification.
 - Advanced process nodes offer a VERY large number of devices and interconnects.
 - Logic circuits may be more tolerant of stat variation (3-sigma may be sufficient).
 - SRAM and flip-flop circuits on-chip require more stringent circuit verification.
 - Detailed parasitic extraction with (sampled) statistical simulation used to verify high-sigma circuit operation.
 - Package and board variation models required for high-speed signal interface simulation and BER calculation.







Outline

- Introduction to Silicon Fabrication Process Nodes
- SoC Design Methodology Challenges
- Heterogeneous Packaging Methodology Impacts
- Future Challenges
- Final Thoughts



Final Thoughts

- There are tremendous opportunities ahead in the electronics market:
 - autonomous driving (with V2V and V2I)
 - robotics applications
 - 5G network communications (mobile, IIoT)
 - design and utilization of "mega" data centers
 - (hybrid) cloud-local computation models
 - ML/AI algorithms (applied to applications requiring < 100% accuracy)
- These opportunities will enable new materials and technologies to flourish:
 - 5nm/3nm devices (GAA, SiGe, Ge)
 - non-volatile storage (e.g., on-chip MRAM)
 - silicon photonics, optoelectronic conversion
 - 2.5D and 3D system-level packaging configurations
 - "in-memory" computing (off-loading a traditional von Neumann CPU)
 - analog computing?



Final Thoughts (continued)

- Design methodologies and project planning need to change (somewhat drastically) to meet the corresponding challenges:
 - early analysis of product reliability
 - evaluating thermal/mechanical/electrical interdependencies
 - establishing suitable design margins for product aging
 - improved "pathfinding" tools for die and packaging solutions
 - greater dependence on functional model validation and coverage measures
 - less dependence on product bring-up resources
 - more dependence upon wafer/package test coverage strategies
 - "PPACRSS" will be the new optimization objective.



Thank you.

Tom Dillinger info@vlsidesignmethodology.com

