Driving, Driven, Along for the Ride: Evolutions of EDA, Manufacturing and Design

Andrew B. Kahng
CSE and ECE Departments
UC San Diego

http://vlsicad.ucsd.edu

Bonus: A few ML in EDA links ☺

https://vlsicad.ucsd.edu/GSRC/metrics/
https://vlsicad.ucsd.edu/Publications/Conferences/356/c356.pdf
https://vlsicad.ucsd.edu/Publications/Conferences/360/c360.pdf
http://vlsicad.ucsd.edu/ASPDAC18/ASP-DAC-2018-Keynote-Kahng-POSTED.pptx
https://vlsicad.ucsd.edu/NEWS18/2018_CASS_ML_Kahng-v5-ACTUAL.pdf
https://dac.com/sites/default/files/files/2018/malenda_agenda_v2.pdf

Agenda

- **The Moore's Law Road**

“Moore’s Law” = Scaling of Cost and Value

- **Moore, 1965**: “The complexity for minimum component costs has increased at a rate of roughly a factor of two per year”

RELATIVE MANUFACTURING COST/COMPONENT
1962
1965
1970
Min cost per transistor
NUMBER OF COMPONENTS PER INTEGRATED CIRCUIT

LOG2 OF THE NUMBER OF COMPONENTS PER INTEGRATED FUNCTION
YEAR

- **Moore’s Law is a law of cost reduction (1% = 1 week)**

ITRS “Greatest Threat is Design Cost” (2001)

DESIGN

SCOPE

Design technology (DT) enables the *conception*, *implementation*, and *validation* of microelectronics-based systems. Elements of DT include *tools*, *libraries, manufacturing process characterizations,* and *methodologies*. DT is the link that transforms ideas and objectives of the electronic systems designer into manufacturable and testable representations. The role of DT is to enable profits and growth of the semiconductor industry via cost-effective production of designs that fully exploit manufacturing capability.

For the 2001 ITRS, the Design ITWG has developed the new System Drivers Chapter, along with models for clock frequency, layout density, power dissipation, etc. in support of the Overall Roadmap Technology Characteristics. New analyses of design cost and design productivity have been introduced into this Design Chapter. DT challenges and needs have been mapped, where possible, to the 2001 ITRS System Drivers. Readers of this Chapter are encouraged to also review previous editions of the ITRS Design Chapter, which provide excellent and still-relevant summaries of DT needs.

The main message in 2001 is this: *Cost of design is the greatest threat to continuation of the semiconductor roadmap.* Cost determines whether differentiating value is best achieved in software or in hardware, on a programmable commodity platform or on a new IC. Manufacturing non-recurring (NRE) costs are just reaching one million dollars (mask set + probe card); design NRE costs routinely reach tens of millions of dollars, with design shortfalls being responsible for silicon re-spins that multiply manufacturing NRE. Rapid technology change shortens product life cycles and makes time-to-market a critical issue for semiconductor customers. Manufacturing cycle times are measured in weeks, with low uncertainty. Design and verification cycle times are measured in months or years, with high uncertainty.

"Moore's Law" = Scaling of Cost and Value

- **Moore, 1965**: "The complexity for minimum component costs has increased at a rate of roughly a factor of two per year"

RELATIVE MANUFACTURING COST/COMPONENT

1962

1965

1970

Min cost per transistor

NUMBER OF COMPONENTS PER INTEGRATED CIRCUIT

LOG_2 OF THE NUMBER OF COMPONENTS PER INTEGRATED FUNCTION

YEAR

- **Moore's Law is a law of cost reduction** **(1% = 1 week)**

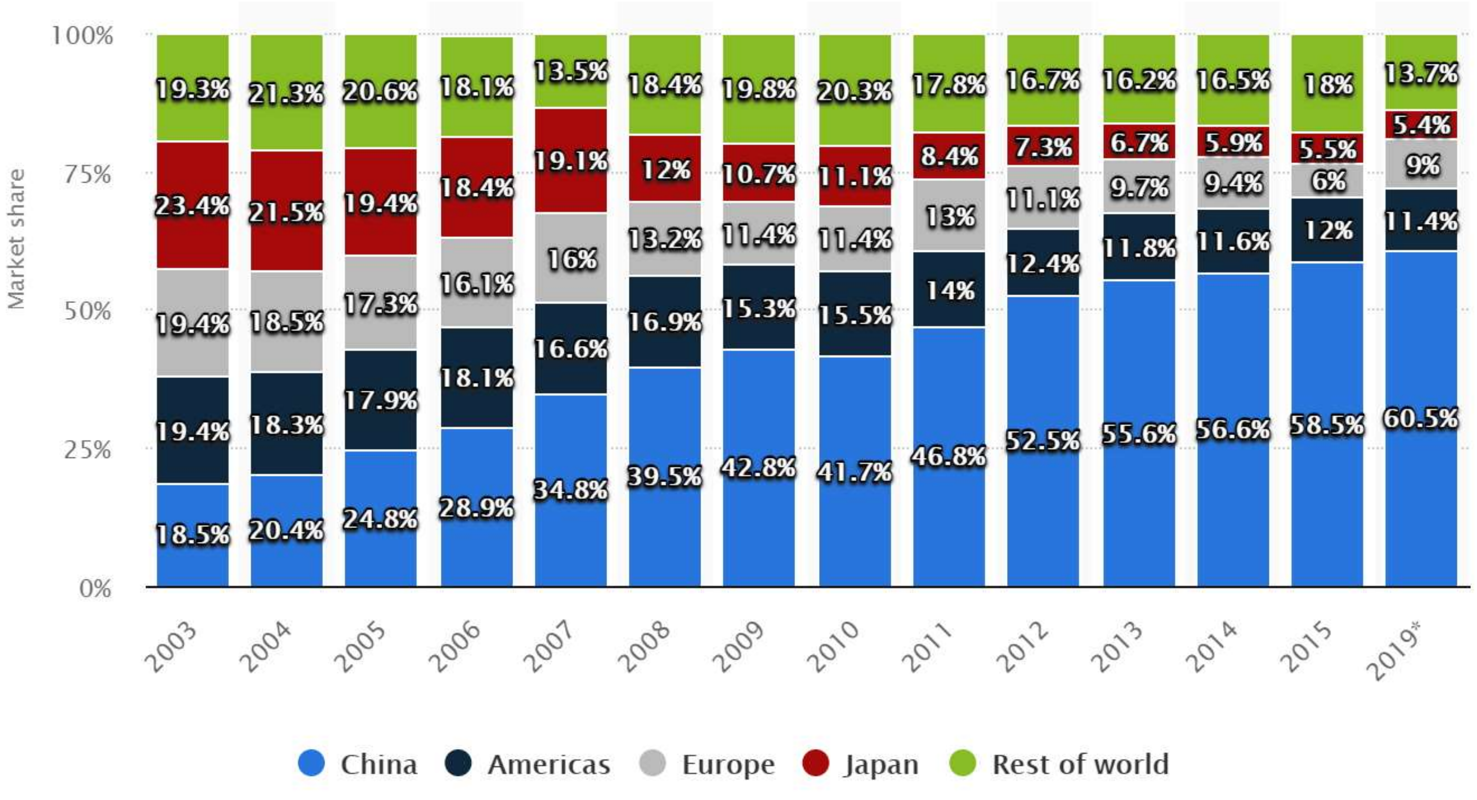
EDA, Manufacturing, Design: Who Drives?

Agenda

- **The Moore's Law Road**
- **Macro Trends** *(or, how we got here)*

Macro Trend 1

- "Race to the end of the roadmap"
  - Scaling levers vanishing (4T cells, buried P/G, backside power – really?)
  - + new world order (rise of Asia; USA off leading edge)



https://www.statista.com/statistics/238228/global-semiconductor-market-share-by-region/

Scaling Effects: Signoff

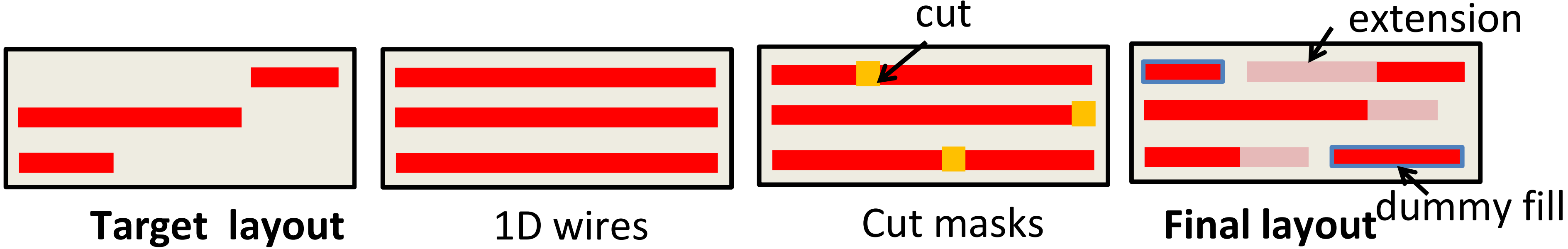
DAC-2015

90*nm* | 65*nm* | 45/40*nm* | 28*nm* | 20*nm* | 16/14*nm* | 10*nm* ≤7*nm*

- Temp inversion
- Noise
- MCMM
- Maxtrans
- PBA
- EM
- AOCV / POCV
- Dynamic IR
- Fixed-margin spec
- Phys-aware timing ECO
- Multi-patterning
- MIS
- Cell-POCV
- Min implant
- MOL, BEOL R ↑
- LVF
- BTI
- BEOL, MOL variations
- Signoff criteria with AVS
- SOC complexity
- Fill effects
- Layout rules

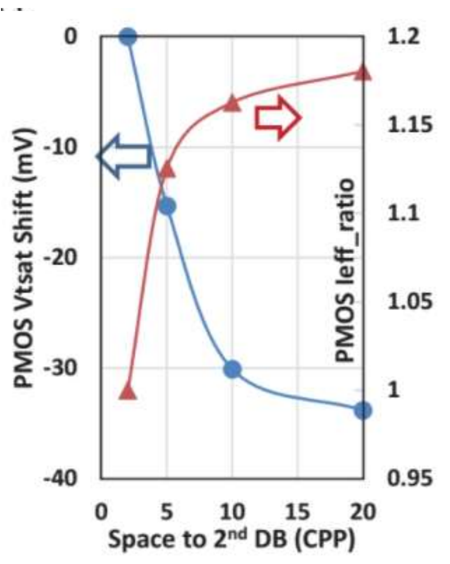
Scaling Effects in BEOL

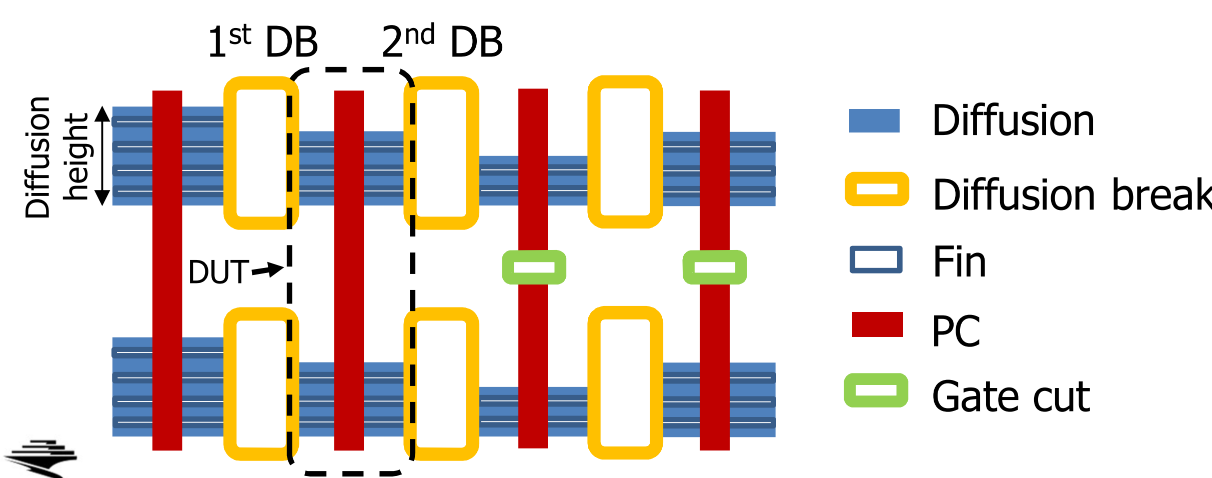
- Self-aligned multiple patterning + Cutmask
- Make a "sea of wires"
- Make "cuts"
- Cut shapes and locations determine **dummy wires** and **end-of-line extensions** of wire segments
- **Final layout ≠ Target layout**
  - → Timing and power not the same as originally designed !
  - → Need more margin !

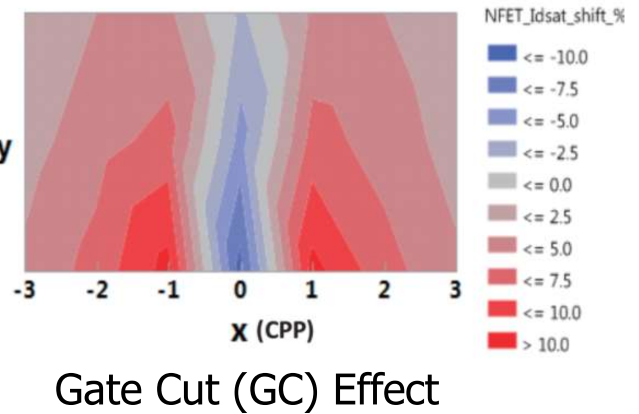


Scaling Effects in FEOL

- **Neighbor diffusion effect (NDE)**
  - Diffusion step = neighboring diffusion height change
  - Drive strength, leakage vary according to horizontal fin spacing
- **2nd Diffusion Break (DB)**
  - Vt shift is function of spacing to **2nd** diffusion break
- **Gate Cut (GC)**
  - Idsat shifts according to gate-cut distance to device
- **Worst corner must consider NDE + 2nd DB + GC**
  → **More margin added besides PVT (!)**







Gate Cut (GC) Effect

Macro Trend 2

- “Race to the end of the roadmap”
  - Scaling levers vanishing
  - + new world order
- **Design technology, capability gaps “rediscovered”**

SEMATECH: Design Productivity Gap (1993)

Potential Design Complexity and Designer Productivity

Logic Transistors per Chip (M)

Productivity (K) Trans./Staff - Mo.

Logic Tr./Chip
Tr./S.M.

Equivalent Added Complexity

58%/Yr. compounded Complexity growth rate

21%/Yr. compound Productivity growth rate

10,000 1,000 100 10 1 0.1 0.01 0.001

100,000 10,000 1,000 100 10 1 0.1 0.01

1981 1983 1985 1987 1989 1991 1993 1995 1997 1999 2001 2003 2005 2007 2009

Year	Technology	Chip Complexity	Frequency	3 Yr. Design Staff	Staff Cost*
1997	250 nm	13 M Tr.	400	210	90 M
1998	250 nm	20 M Tr.	500	270	120 M
1999	180 nm	32 M Tr.	600	360	160 M
2002	130 nm	130 M Tr.	800	800	360 M

* @ $150K / StaffYr. (In 1997 Dollars)

Source: ©SEMATECH

ITRS: “Greatest Threat is Design Cost” (2001)

DESIGN

SCOPE

Design technology (DT) enables the *conception*, *implementation*, and *validation* of microelectronics-based systems. Elements of DT include *tools*, *libraries, manufacturing process characterizations,* and *methodologies*. DT is the link that transforms ideas and objectives of the electronic systems designer into manufacturable and testable representations. The role of DT is to enable profits and growth of the semiconductor industry via cost-effective production of designs that fully exploit manufacturing capability.

For the 2001 ITRS, the Design ITWG has developed the new System Drivers Chapter, along with models for clock frequency, layout density, power dissipation, etc. in support of the Overall Roadmap Technology Characteristics. New analyses of design cost and design productivity have been introduced into this Design Chapter. DT challenges and needs have been mapped, where possible, to the 2001 ITRS System Drivers. Readers of this Chapter are encouraged to also review previous editions of the ITRS Design Chapter, which provide excellent and still-relevant summaries of DT needs.

The main message in 2001 is this: *Cost of design is the greatest threat to continuation of the semiconductor roadmap.* Cost determines whether differentiating value is best achieved in software or in hardware, on a programmable commodity platform or on a new IC. Manufacturing non-recurring (NRE) costs are just reaching one million dollars (mask set + probe card); design NRE costs routinely reach tens of millions of dollars, with design shortfalls being responsible for silicon re-spins that multiply manufacturing NRE. Rapid technology change shortens product life cycles and makes time-to-market a critical issue for semiconductor customers. Manufacturing cycle times are measured in weeks, with low uncertainty. Design and verification cycle times are measured in months or years, with high uncertainty.

ITRS: Design Capability Gap (2013)

Transistor density (#transistors/mm²)

10000000
1000000
100000
10000
1000

1995 2000 2005 2010 2015

Year

Actual data

Ideal: 2x per node

Real: <2x per node

"GAP"

A NEW IC DESIGN GAP: available density scaling vs. realized density scaling

Non-ideal A-factor → larger cells, wires for reliability

Uncore in architecture → small, distributed functions

DARPA: Failure of EDA? (2018)

DARPA Has EDA failed to keep up with Moore's Law?

Transistors

Cost ($M)

10B 100M 1M 10K 100

100 75 50 25 0

Transistors per chip

Total HW Cost

Design Cost

Verification Cost

1980 85 90 95 2000 05 10 15

Technology Node 180 130 65 45 32 22 14

A. Olofsson, http://www.ispd.cc/slides/2018/k2.pdf

Macro Trend 2

- "Race to the end of the roadmap"
  - Scaling levers vanishing
  - + new world order
- Design technology, capability gaps rediscovered
  - EDA revenues stable at < 3% of semiconductor revenues

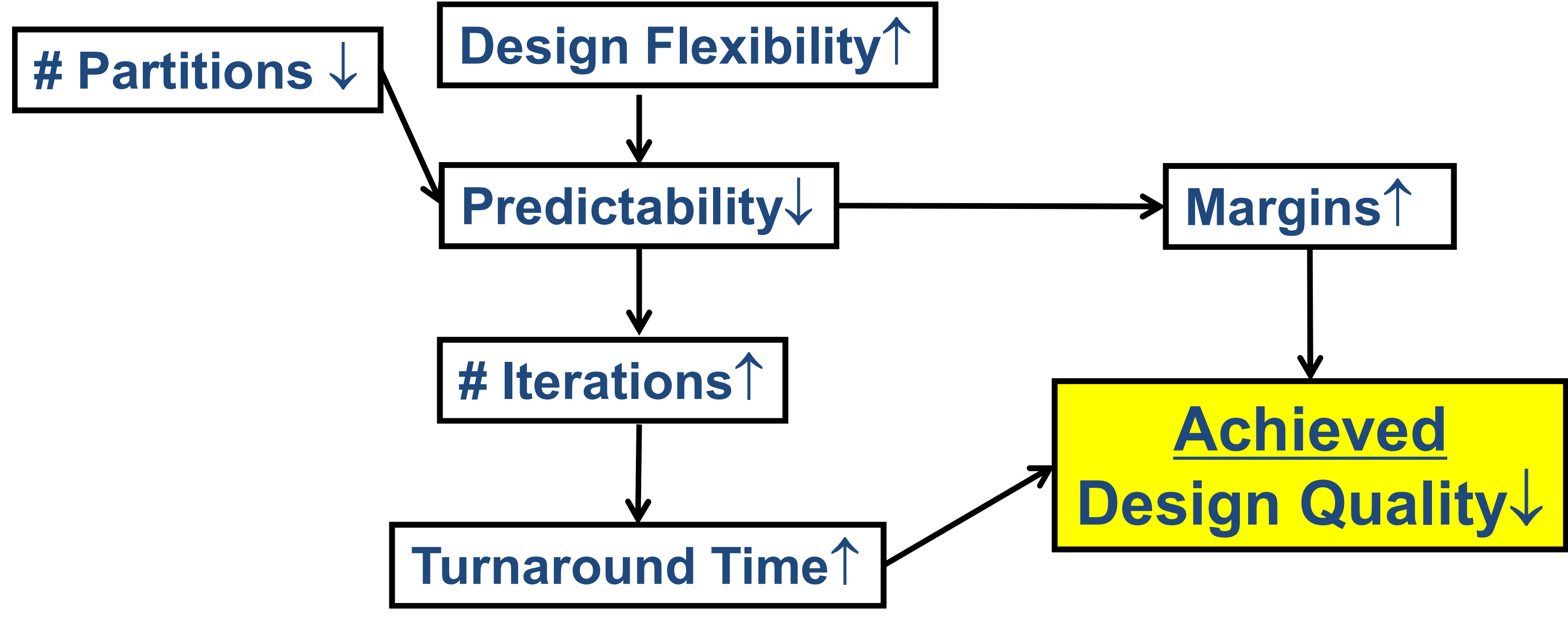
Was EDA Just “Along for the Ride”?

- Intractable optimizations → heuristics piled on heuristics
- “Noise” or “Chaos” when EDA tools “try hard”
- Unpredictability → added margin and schedule

14nm PULPino: Δarea = 6% from Δfreq = 10MHz !

Barriers to Design Starts: Time, Expertise, Risk

Was EDA Just "Along for the Ride"?



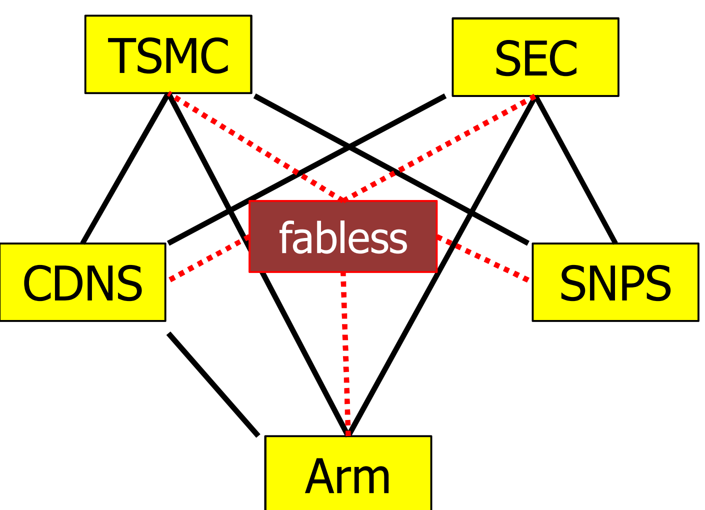
Today: in "local minimum" of design technology, methodology, quality

Macro Trend 3

- "Race to the end of the roadmap"
  - → new world order
  - + loss of scaling levers
- Design technology, capability gaps rediscovered
  - EDA revenues stable at < 3% of semiconductor revenues
- **Extreme consolidation**
  - **Two leading foundries**
  - **Two dominant EDA companies**
  - **One litho tool supplier**
  - **GPU, mobile SOC, FPGA, … clearly delineated**

The Rules Have Changed

- **Semiconductor wars fought with fewer arms manufacturers, dealers** but at faster pace, on higher-value fronts
- **Point-to-point, encrypted, mutually exclusive relationships** collaborations → mutual survival
  - **"Second supply" is history**
  - **"Pre-competitive" is history** (but, everyone pays IMEC)
  - **Everything hard-coded**
  - ***Life is good?***

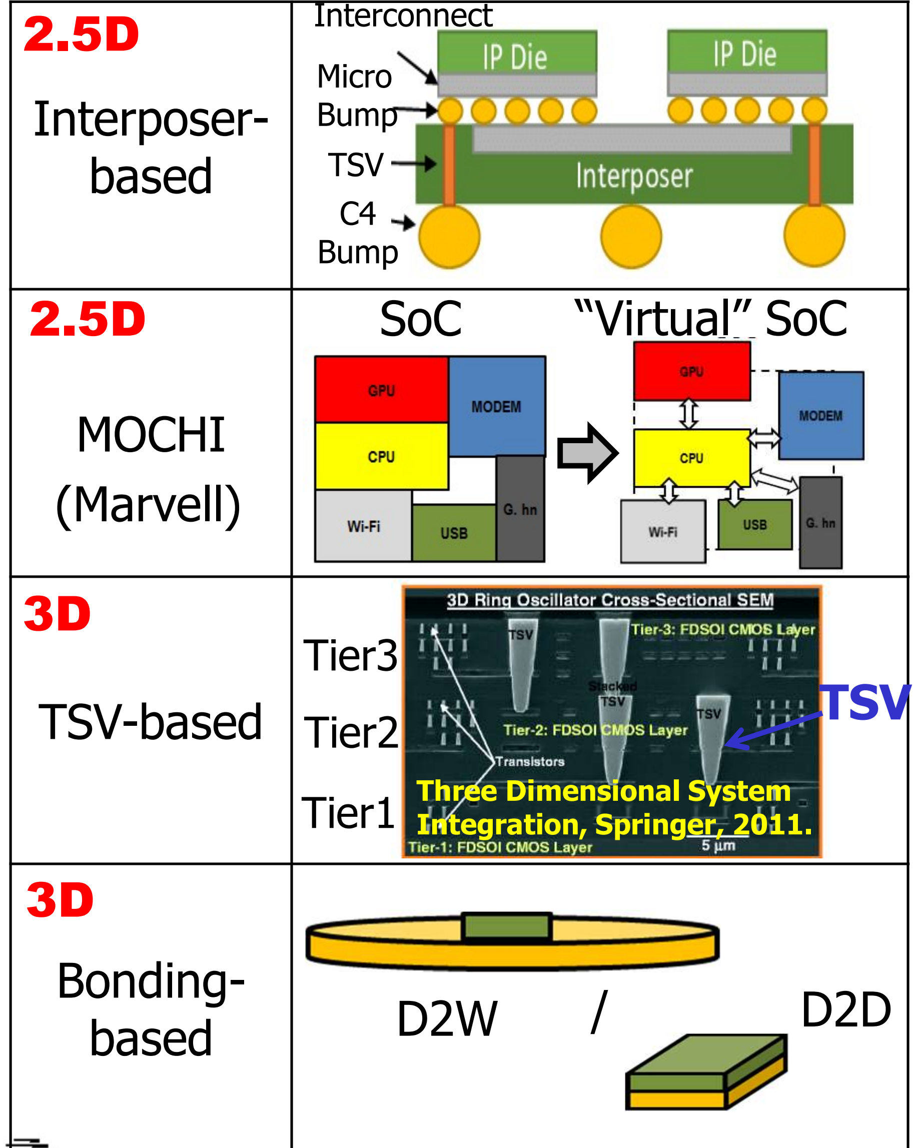


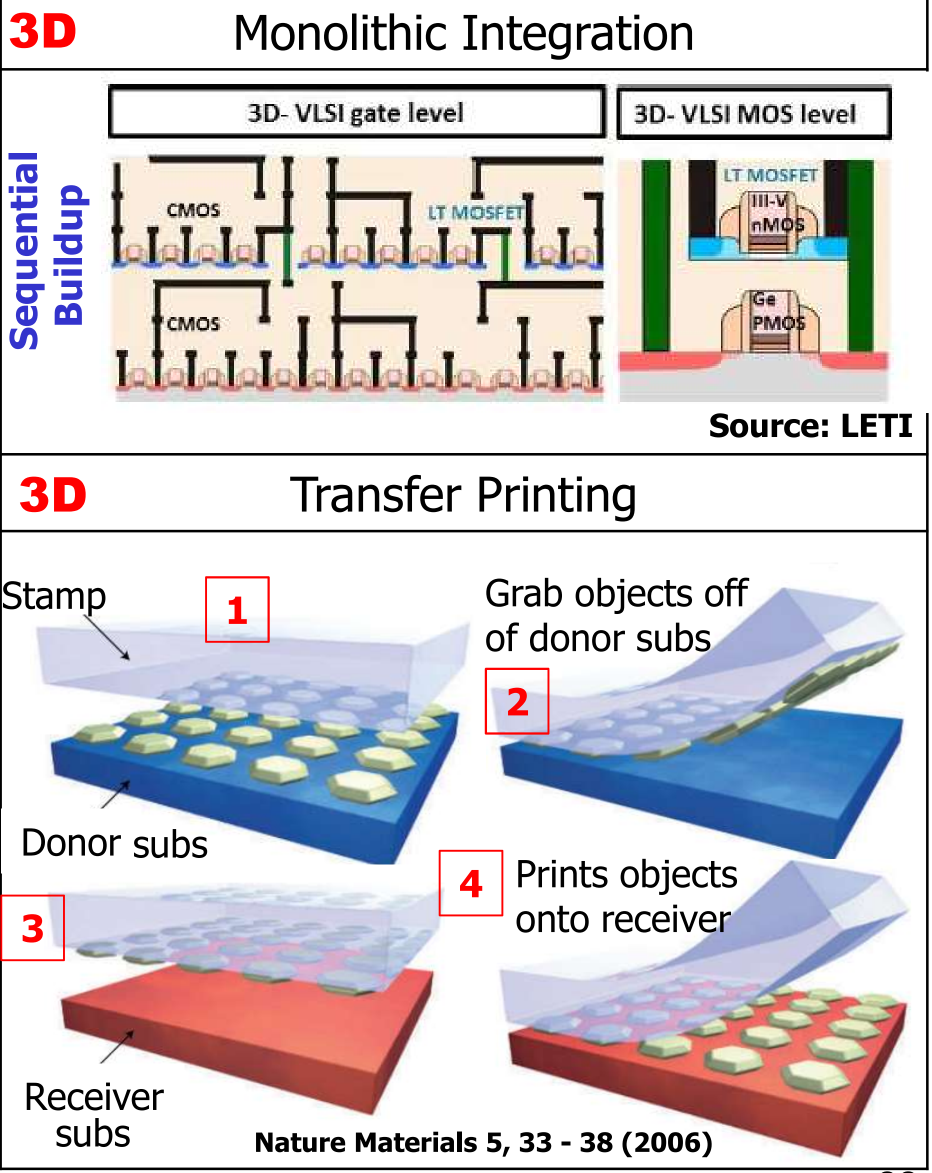
- **Foundry is the borg**
  - **IPs, design services, integration, packaging** + bespoke EDA?

Macro Trend 4

- "Race to the end of the roadmap"
  - → new world order
  - + loss of scaling levers
- Design technology, capability gaps rediscovered
  - EDA revenues stable at < 3% of semiconductor revenues
- Consolidation
  - Two leading foundries
  - Two dominant EDA companies
  - One litho tool supplier
  - GPU, mobile SOC, FPGA, … clearly delineated
- **More Than Moore (finally! but…)**
  - **Beyond-die, beyond-CMOS, heterogeneous integration…**
  - **But still at "1000 points of light" stage, plus caveats**

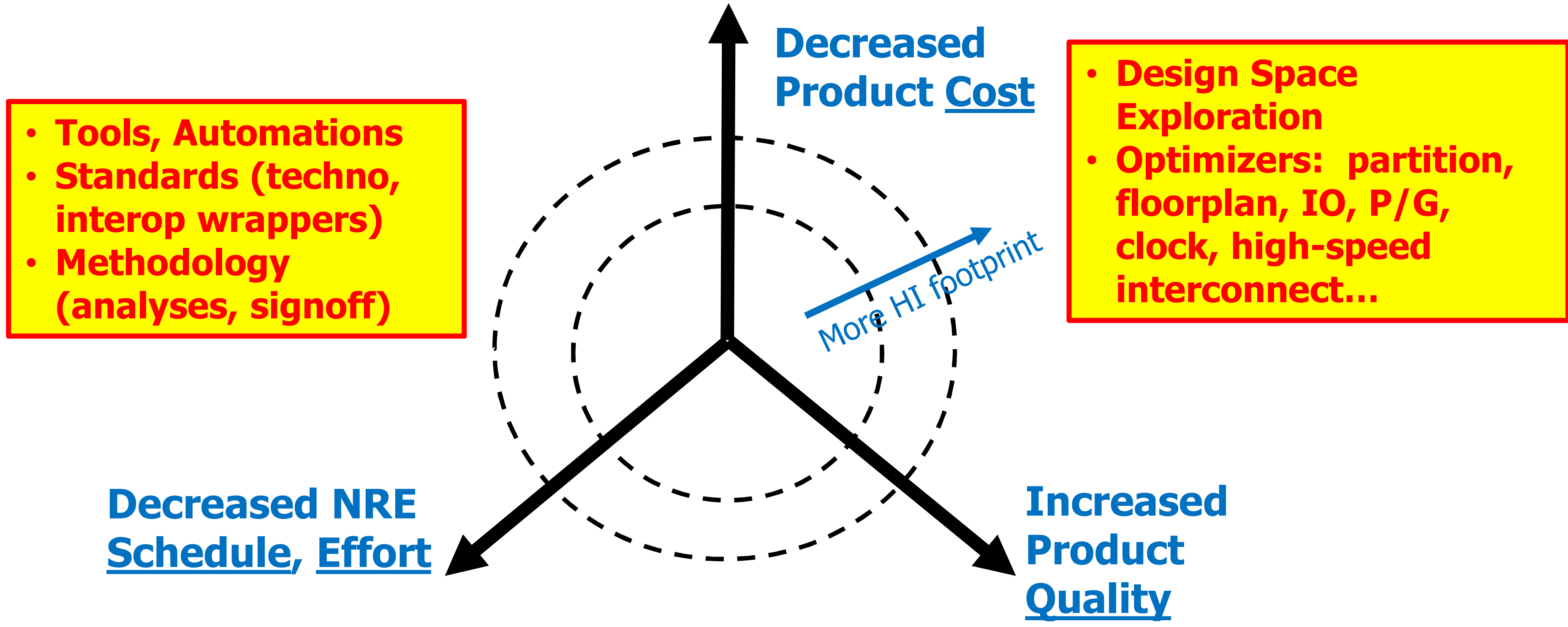
"More Than Moore": 2.5D/3D Integration





Heterogeneous Integration Inevitable, But …

- 2.5D, 3D: adoption is cost- and value-driven
- Unresolved co-design (beyond the die) challenges



- Can evolution of EDA, Design, Manufacturing for HI learn from the IC experience? who drives, standards, …

Agenda

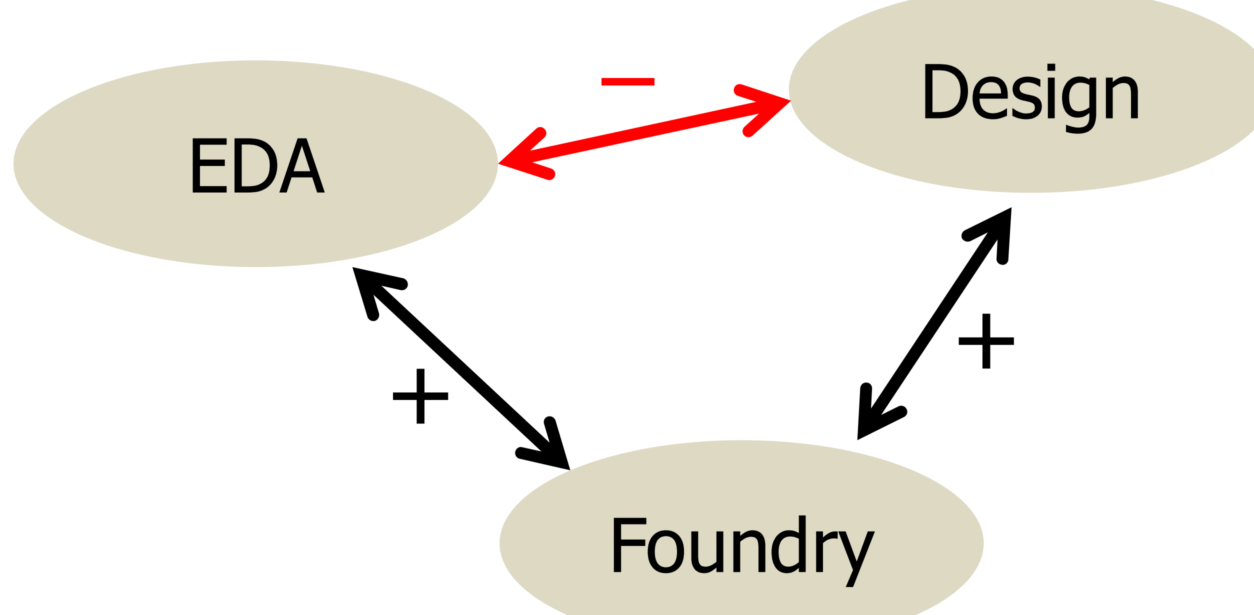
- **The Moore's Law Road**
- **Macro Trends** ***(or, how we got here)***
- **Driving, or Driven?**

Past ~20 Years

- Lived: DFM, DDK, virtual fab, statistical, DAM, MAD
  - SSTA
  - Design-aware manufacturing exposure push, poly bias, yield ramp
  - Manufacturing-aware design CMP-aware PEX, spread-fatten-fill, LFD / PV bands, ...
- Learned: Foundry and Design hold many cards
  - DAM + MAD actually live here
  - Levers that change Paretos / entitlements in an instant IPs, layout methodology, signoff criteria, process
  - Design: tests, signoff corners, custom cells, …
  - Foundry: wiggle room for process → SSTA a non-starter
  - → Poor ROI for EDA $ for A-B expts, payoff only if Foundry, Design fail
- Era of DTCO (design-technology co-optimization)
  - Effective, symbiotic collaboration between foundry, EDA (and IP) Encouraging growth of lead times; Design can be “odd man out” (!)

Observation: Alignability

- **Design and Foundry:** ***alignable*** yield ramp, 1-off node, services++
- **Foundry and EDA:** ***alignable*** DDKs, DFM rules, DTCO org's
- **EDA and Design:** ***not fully alignable (!)***
  - **EDA wants to know designs, tool usage of Design**
  - **Design knows EDA will leverage to sell to competitors**
  - **Design:EDA :: Customer:Supplier** contention over CAD $$$



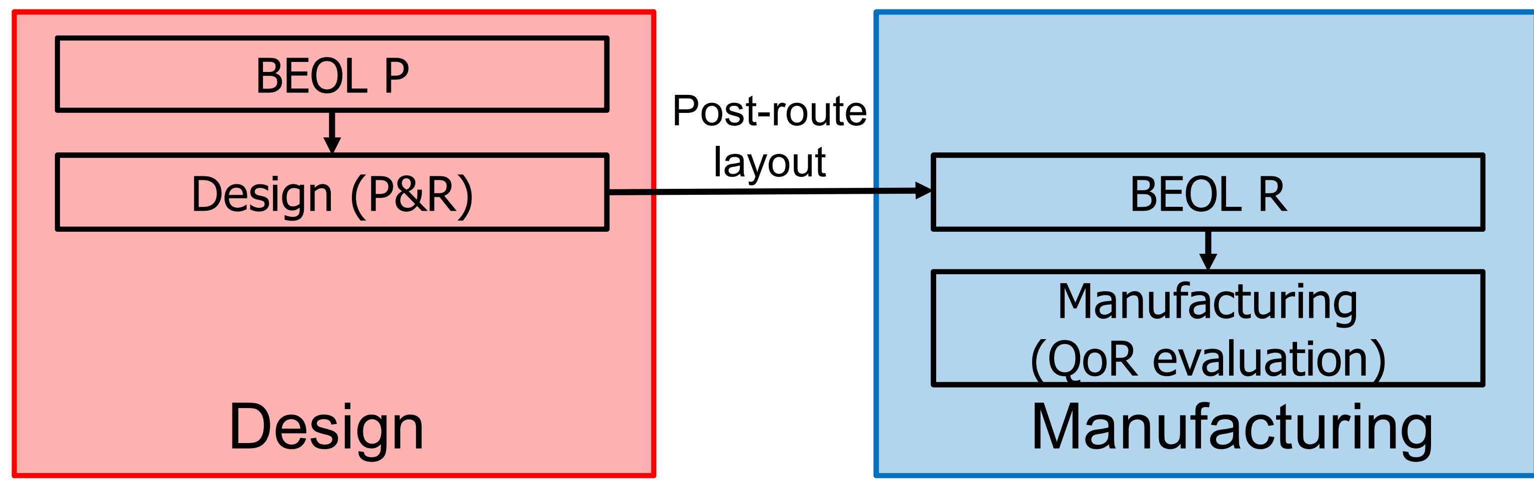
Who Will Drive? scaling of cost and value !

EDA
Design
Foundry
−
+
+

- **Foundry?**
  - Can drive EDA and entire semi supply chain
  - Can offer IPs, design services, bespoke EDA
  - One day: takes system design and spec, handles everything from there ?
- **Design?**
  - "Fruit company", "GPU company", "FPGA company" all can command custom nodes, business models
  - One day: drives process and EDA from application and architecture?
- **Foundry + Design? (MAD + DAM)**

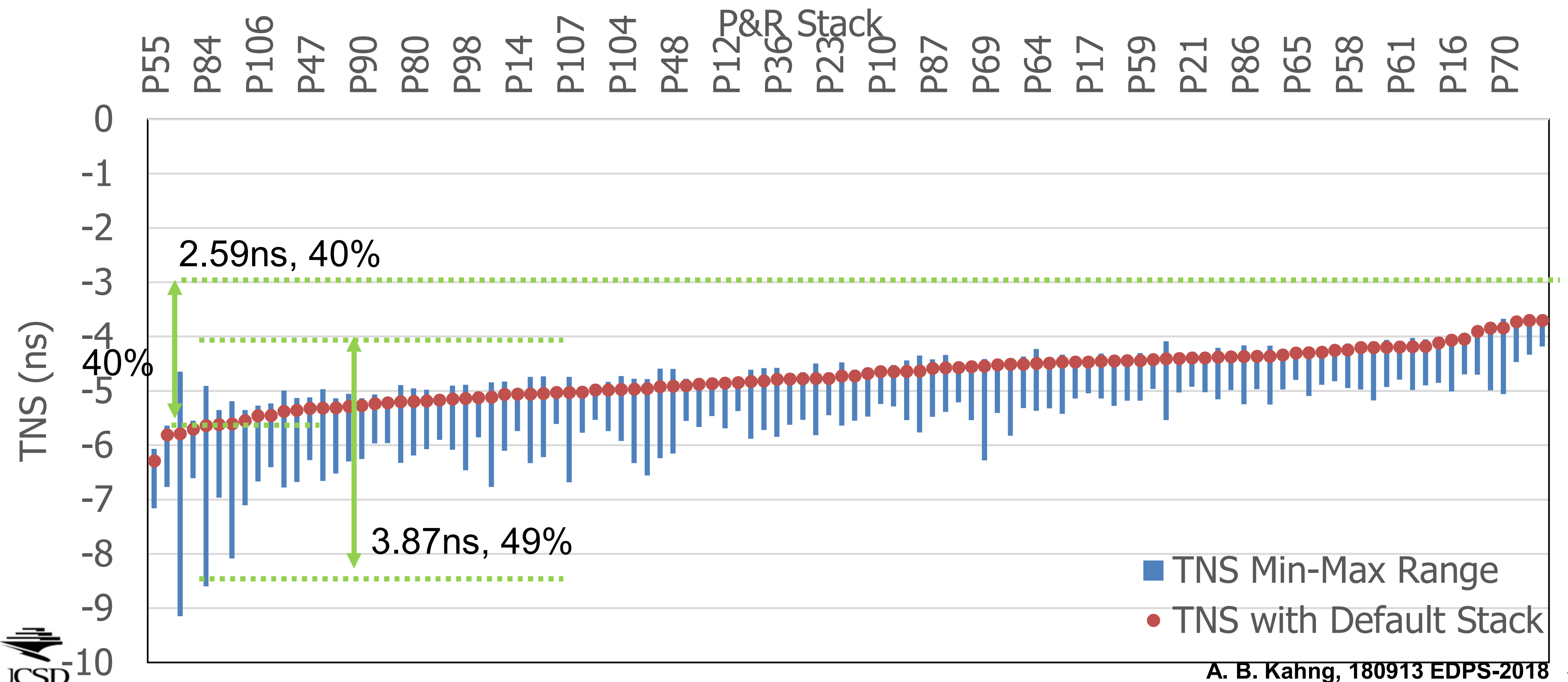
DAM & MAD Study: Experimental Setup

- #Stacks = 3x3x3x4=108, each w/ unique number
  - BEOL P{stack}: stack used for P&R
  - BEOL R{stack}: stack used for PEX/STA
- Layer type: two 1X, two 1.5X and four 2.5X layers
- DC = {0.5, 0.6, 0.7} for each layer type
- AR = {1.5, 1.75, 2, 2.25} uniform for all layers
- Design: LDPC (HP=0.8ns), both X1 and X4 cells



DAM & MAD Study: Experimental Results

- X-axis: P{stack}
- Y-axis:
  - For a given P{stack}, TNS range using all R{stack}s (blue)
  - For a given P{stack}, TNS using R{default} (red)
- DAM+MAD=60% difference in TNS



ISQED17

DAM & MAD Study: Experimental Results

- Y-axis:
  - For a given P{stack}, power using all R{stack}s (orange)
  - For a given P{stack}, power using R{default} (black)
- DAM+MAD=7% difference in power
- Weak correlation

P&R Stack

P55 P108 P84 P91 P106 P100 P47 P79 P90 P7 P80 P20 P98 P67 P14 P68 P107 P49 P104 P88 P48 P4 P12 P39 P36 P77 P23 P37 P10 P63 P87 P89 P69 P29 P64 P24 P17 P97 P59 P52 P21 P45 P86 P51 P65 P6 P58 P35 P61 P38 P16 P31 P70 P5

TNS (ns): -3 -4 -5 -6 -7 -8 -9 -10 -11 -12 -13 -14 -15

Power (mW): 5.9 5.7 5.5 5.3 5.1 4.9 4.7 4.5

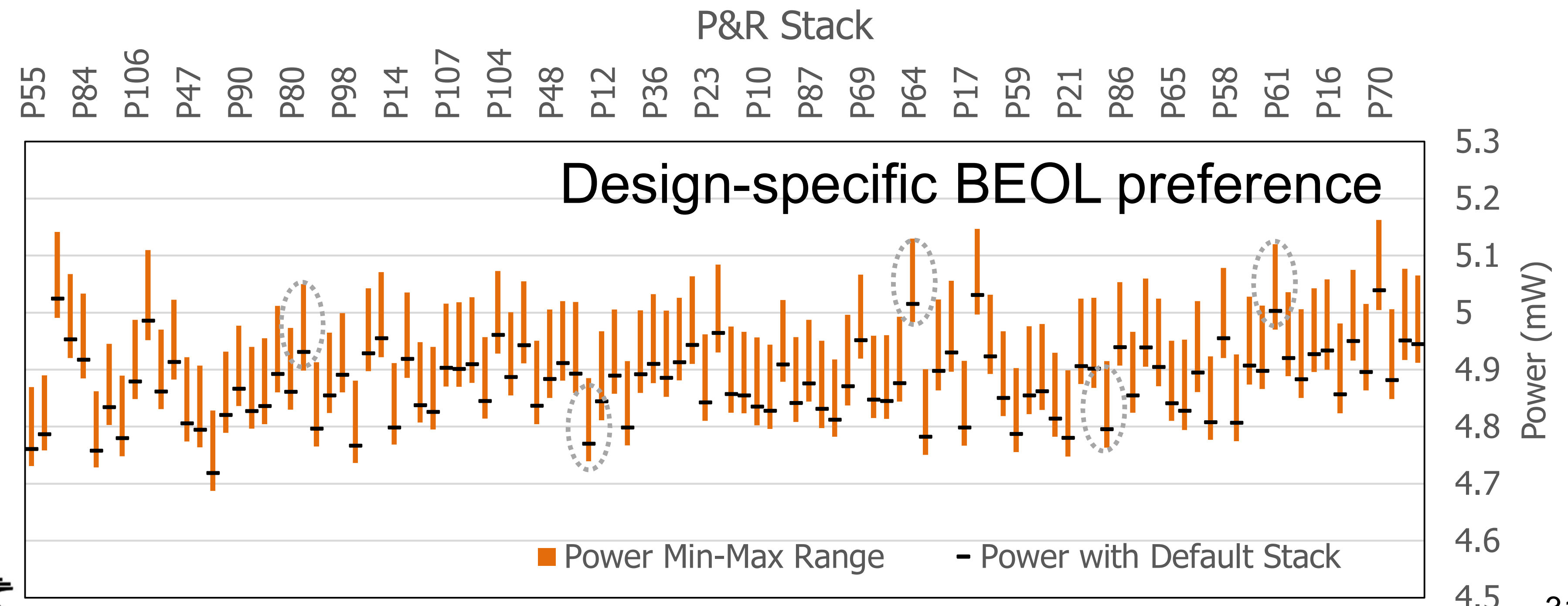
Weak correlation of timing and power

TNS Min-Max Range
Power Min-Max Range
TNS with Default Stack
Power with Default Stack

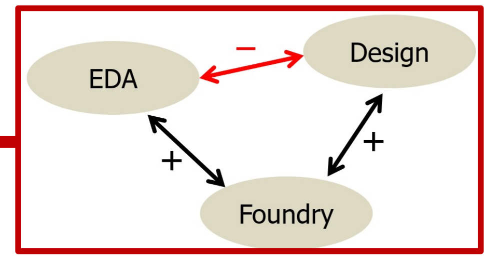
ISQED17

DAM & MAD Study: Experimental Results

- Towards N7/N5, are there potential benefits of new DAM/MAD methodologies?
- A: Possibly, yes.
  - Up to 60%/7% difference in TNS/power
  - One optimal design-specific stack for manufacturing may be preferred regardless of the BEOL stack assumed during P&R



Who Will Drive? scaling of cost and value !



- Foundry?
  - Can drive EDA and entire semi supply chain
  - Can offer IPs, design services, bespoke EDA
  - One day: takes system design and spec, handles everything from there ?
- Design?
  - "Fruit company", "GPU company", "FPGA company" all can command custom nodes, business models
  - One day: drives process and EDA from application and architecture?
- Foundry + Design? (MAD + DAM)
- EDA?
  - No. EDA as we know it really is a supplier industry.
  - But has lots more value that could be unlocked !

+Value: 1 Week = 1 Percent

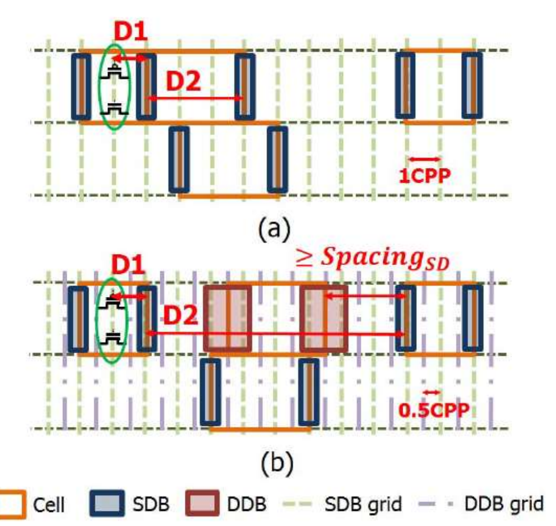
- DARPA IDEA program kickoff in June 2018
  - Part of DARPA Electronics Resurgence Initiative
- No humans, 24-hour TAT = Design-based equivalent scaling writ large

> IDEA will create a no-human-in-the-loop hardware compiler for translating source code to layouts of System-On-Chips, System-In-Packages, and Printed Circuit Boards in less than 24 hours

A. Olofsson, DARPA
ISPD-2018 keynote

+Value: Diffusion Break-Aware Leakage Opt

- A diffusion break (DB) isolates two neighboring devices
  - Single Diffusion Break (SDB) and Double Diffusion Break (DDB)
  - The distance to 2nd DB has an impact on delay and leakage power.
- 2nd DB-aware leakage optimization and placement methodology by relocation, gate sizing, Vt swapping and DB swapping.

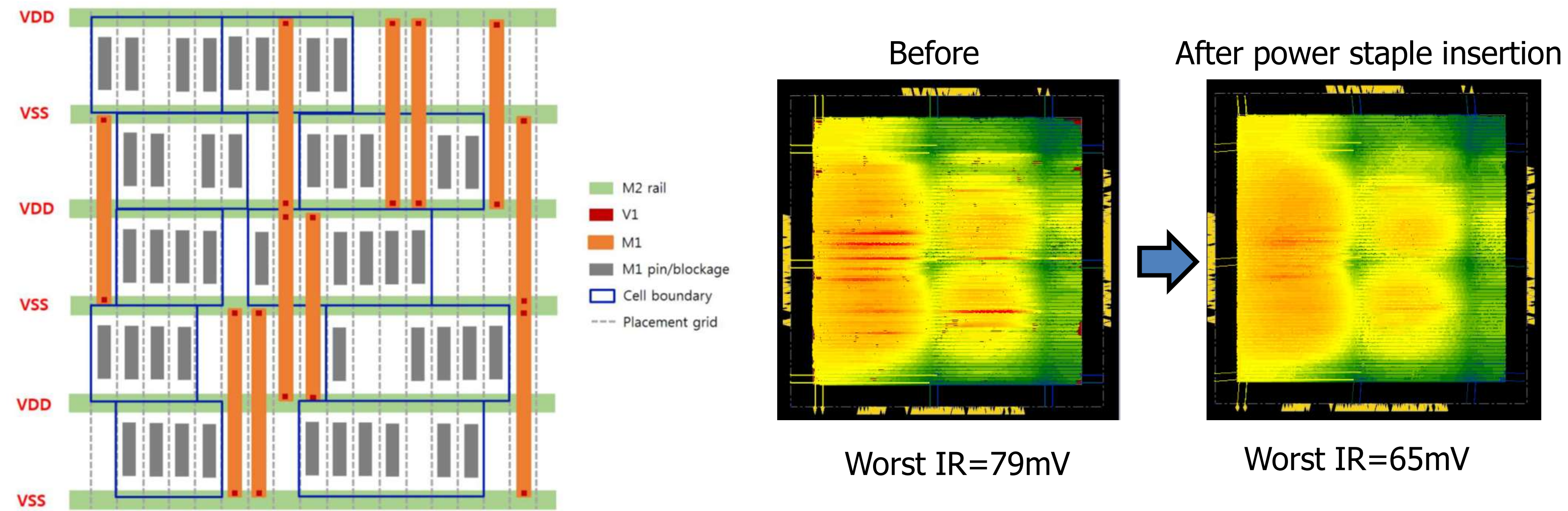


Design	Type	2nd DB-Unaware		2nd DB-Aware		Our Result		
		WS (ns)	Leak(mW)	WS (ns)	Leak(mW)	WS (ns)	Leak(mW)	Recovery
AES	Type-I	0.001	0.228	0.002	0.300	-0.002	0.237	87.5%
MPEG	Type-I	0.002	0.219	0.003	0.261	-0.001	0.245	38.1%
JPEG	Type-I	0.001	0.667	0.001	0.840	-0.002	0.735	60.7%
VGA	Type-I	0.001	1.329	0.002	1.905	0.000	1.453	78.5%
AES	Type-II	0.000	0.189	0.001	0.194	-0.002	0.184	>100%
MPEG	Type-II	0.002	0.222	0.002	0.258	-0.001	0.243	41.7%
JPEG	Type-II	0.001	0.704	0.001	0.762	-0.002	0.716	79.3%
VGA	Type-II	0.004	1.330	0.004	1.589	-0.005	1.364	90.5%

submitted

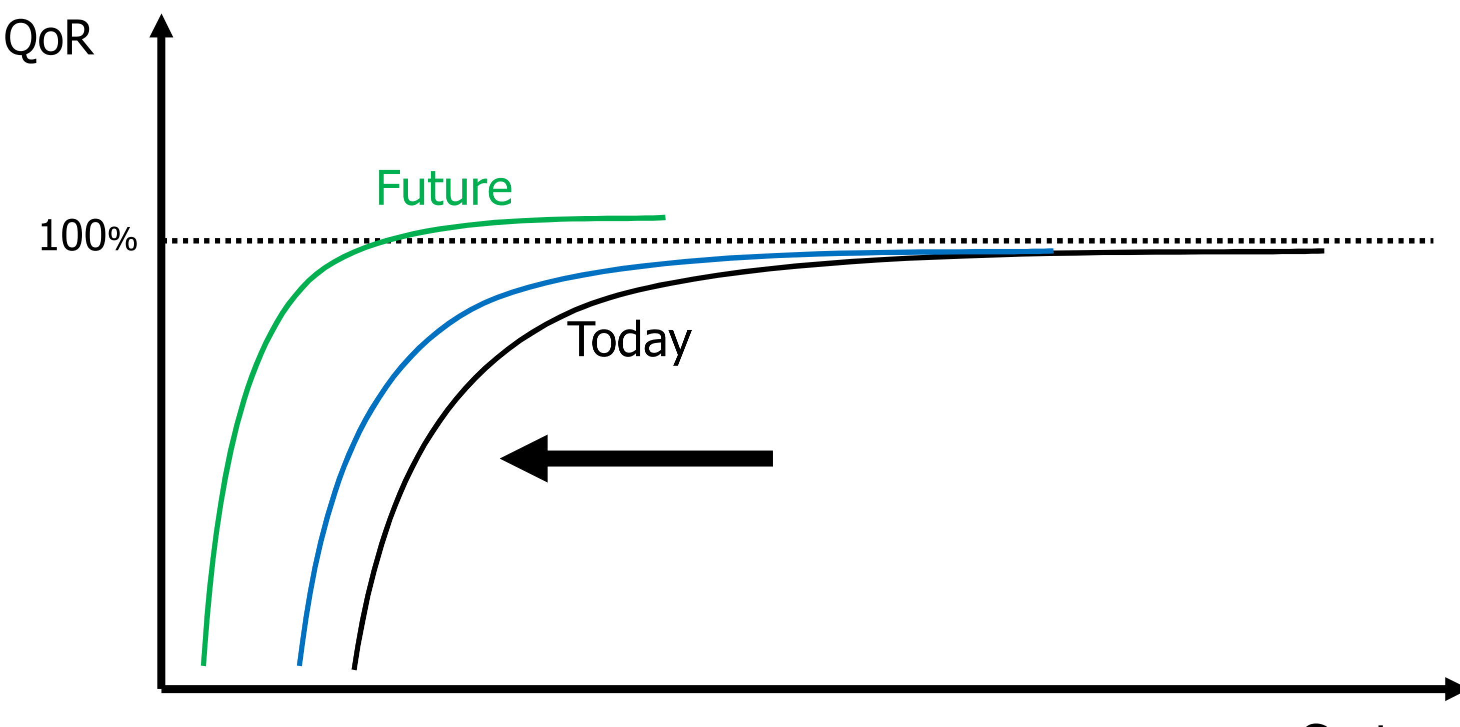
+Value: IR Drop Mitigation by Power Stapling

- Power staples are short pieces of wires and vias connecting two or more adjacent power rails to mitigate the IR drop.
- Dynamic programming (DP)-based detailed placement optimization to improve power staple insertion
  - Single-row and Double-row optimization



+++Value: Improved Design Quality vs. Cost

- **“Design-Based Equivalent Scaling” won’t be driven by EDA, but will depend on EDA**
- **Shared goal (EDA, Manufacturing, Design)**
  = shift Quality vs. Cost curve up and to the left
  not clear how this matches EDA business model …



Agenda

- The Moore's Law Road
- Macro Trends *(or, how we got here)*
- Driving, or Driven?
- Conclusion

Evolutions of EDA, Manufacturing, Design

- Foundry may increasingly be the "driver"
  - Well-positioned for future heterogeneous integration
  - Has established point-to-point linkages to Design, EDA
    but those 1-1 + encrypted + mutex relationships aren't helpful
  - Positioned to take on more of Design, EDA scope
- Both MAD-DAM (Design + Foundry) and Design-based equivalent scaling (EDA) have headroom
  - Especially, many design quality improvements still on table
  - Increasingly attractive as benefits of scaling dwindle
- EDA delivers/enables more "wingman", not "along for the ride"
  - "Design-based equivalent scaling"
    - Schedule and NRE reduction (1 week = 1 percent; no-humans)
    - Better QOR earlier in technology lifecycle
  - Learn from history of Macro Trends + why they happened → do more good in the beyond-Moore future scaling of design capability: system-level, design space exploration, pathfinding

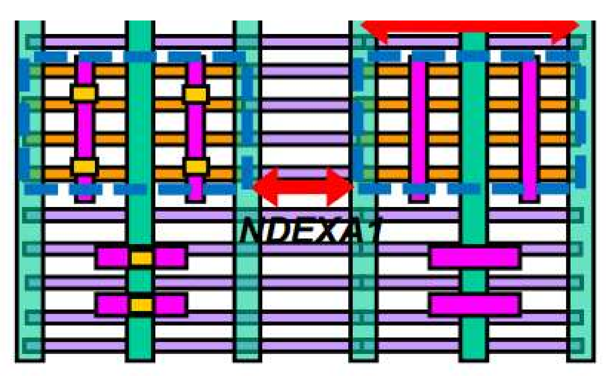
THANK YOU !

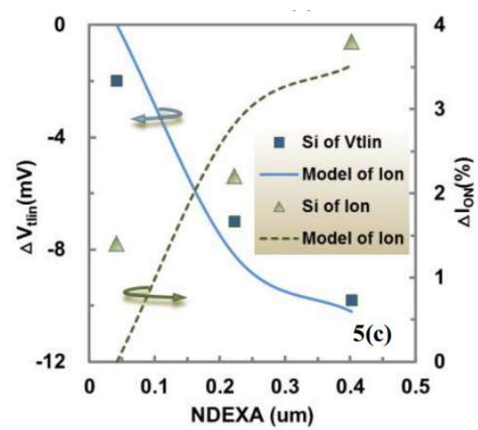
Support from NSF, DARPA, Qualcomm, Samsung, NXP, Mentor Graphics and the C-DEN center is gratefully acknowledged.

BACKUP

Motivation

- Neighboring diffusion area creates stress on the channel [1]
  - NDEXA: width of neighboring diffusion spacing
  - Threshold voltage and on current change accordingly



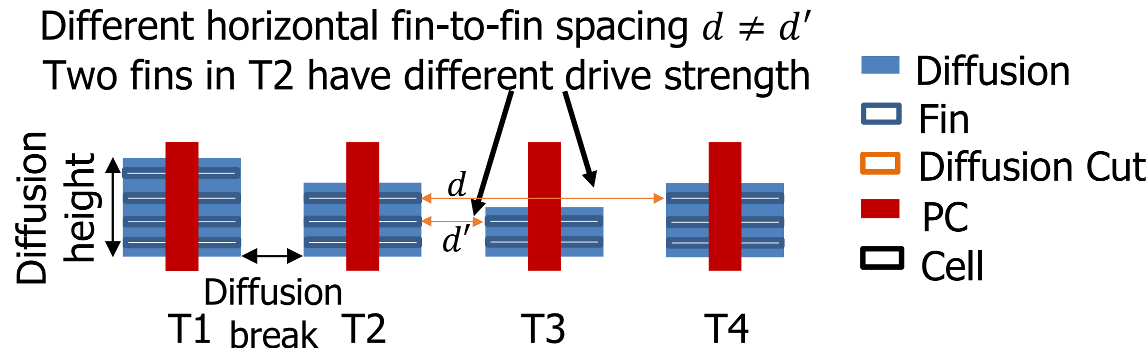


PMOS/NMOS may show different characteristics w.r.t. process

[1] D. C. Chen, G. S. Lin, T. H. Lee, et al., *Compact Modeling Solution of Layout Dependent Effect for FinFET Technology*, Proc ICMTS, 2015, pp. 110-115.

Neighbor Diffusion Effect (NDE)

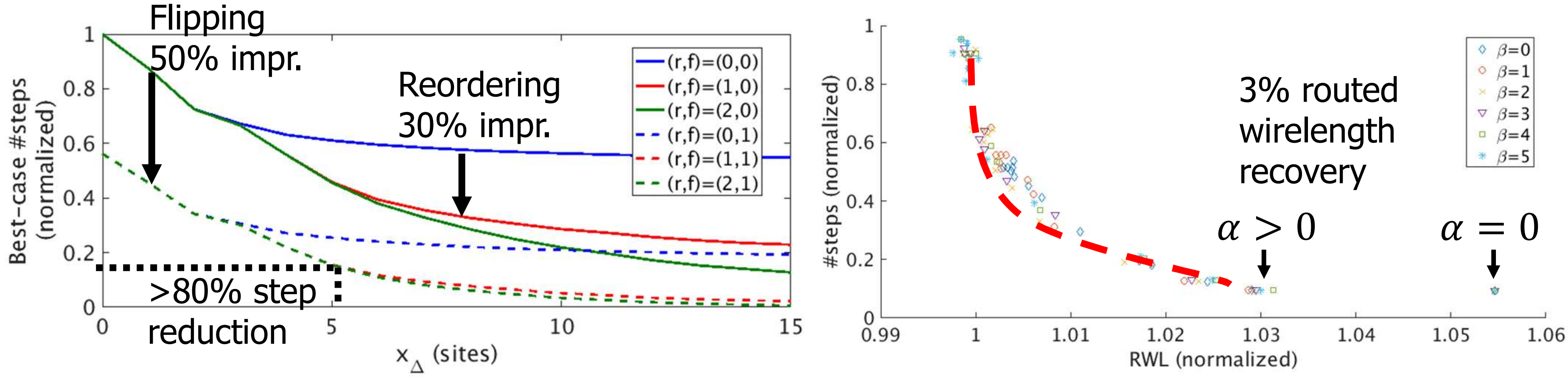
- Diffusion step
  - Neighboring diffusion area height change



- Model-hardware correlation issue
  - Inter-cell NDE cannot be captured in standard cell characterization

Sensitivity to Ranges, Flipping and Coefficients

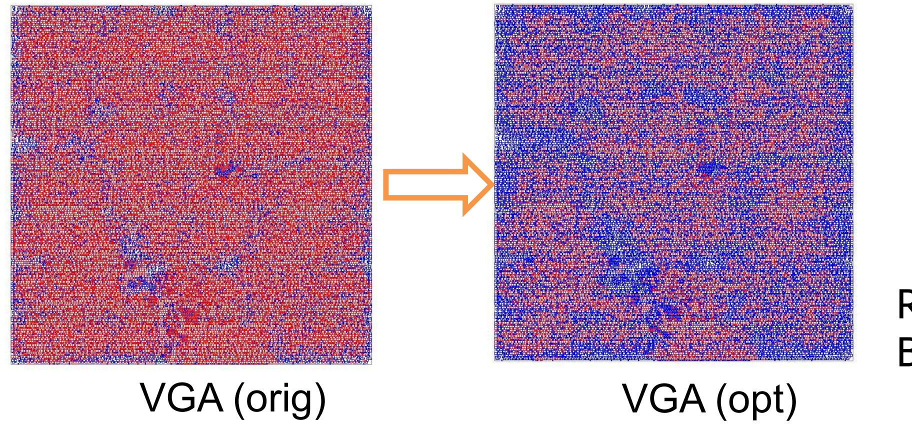
- Sensitivity to disp. range (x_Δ), reordering range (r), and flipping (f)
  - >80% step reduction
  - Reordering = 30% more step reduction; flipping = 50% step reduction
- Sensitivity to displacement coefficients α, and flipping coefficients β
  - Clear tradeoff between displacement, flipping and step reduction
  - 3% RWL overhead with maximum step reduction



Experimental Results

- SR: up to 84.0% step reduction
- DR: up to 90.6% step reduction

Design	Initial		Single-row (SR)	Double-row (DR)		
	#steps	DH%	#steps	#steps	Runtime	Est. yield impr.
AES	7973	4.3%	1278 (-84.0%)	750 (-90.6%)	37s	+0.71%
M0	6588	8.4%	1612 (-75.5%)	842 (-87.2%)	38s	+0.57%
JPEG	34760	8.3%	9275 (-73.3%)	4555 (-86.9%)	156s	+2.86%
VGA	50766	24.8%	27054 (-46.7%)	11816 (-76.7%)	195s	+3.59%
MPEG	9994	23.0%	5071 (-49.3%)	2402 (-76.0%)	25s	+0.75%

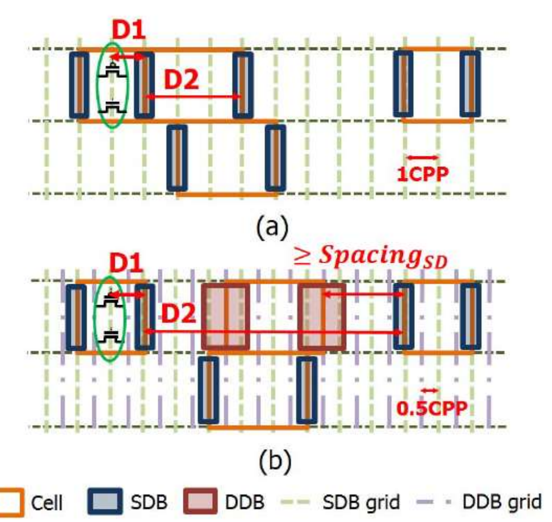


Red: cells w/ steps
Blue: cells w/o steps

ASP-DAC19

+Value: Diffusion Break-Aware Leakage Opt

- A diffusion break (DB) isolates two neighboring devices
  - Single Diffusion Break (SDB) and Double Diffusion Break (DDB)
  - The distance to 2nd DB has an impact on delay and leakage power.
- 2nd DB-aware leakage optimization and placement methodology by relocation, gate sizing, Vt swapping and DB swapping.



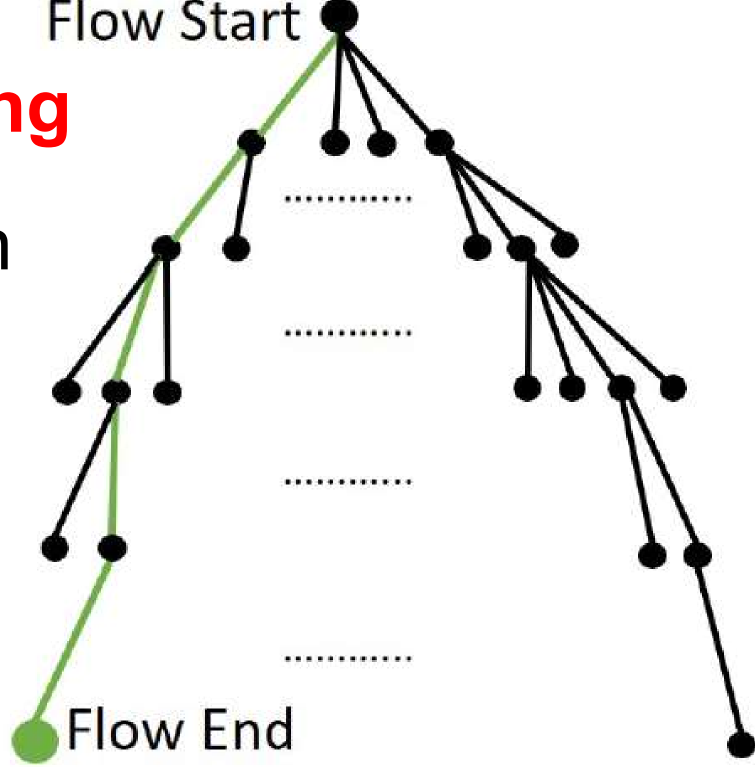
Design	Type	2nd DB-Unaware		2nd DB-Aware		Our Result		
		TNS (ns)	Leak(mW)	TNS (ns)	Leak(mW)	TNS (ns)	Leak(mW)	Recovery
AES	Type-I	0.000	0.228	0.000	0.300	-0.012	0.237	87.5%
MPEG	Type-I	0.000	0.219	0.000	0.261	-0.003	0.245	38.1%
JPEG	Type-I	0.000	0.667	0.000	0.840	-0.010	0.735	60.7%
VGA	Type-I	0.000	1.329	0.000	1.905	0.000	1.453	78.5%
AES	Type-II	0.000	0.189	0.000	0.194	-0.015	0.184	>100%
MPEG	Type-II	0.000	0.222	0.000	0.258	-0.004	0.243	41.7%
JPEG	Type-II	0.000	0.704	0.000	0.762	-0.014	0.716	79.3%
VGA	Type-II	0.000	1.330	0.000	1.589	-0.011	1.364	90.5%

Bonus: Extracted from an ML in EDA talk ☺

Solution 2: ML for Time, Effort Reduction

Four Stages of Machine Learning

1. Mechanization and Automation
2. Orchestration of Search and Optimization
3. Pruning via Predictors and Models
4. From Reinforcement Learning through Intelligence



Huge space of tool, command, option trajectories through design flow

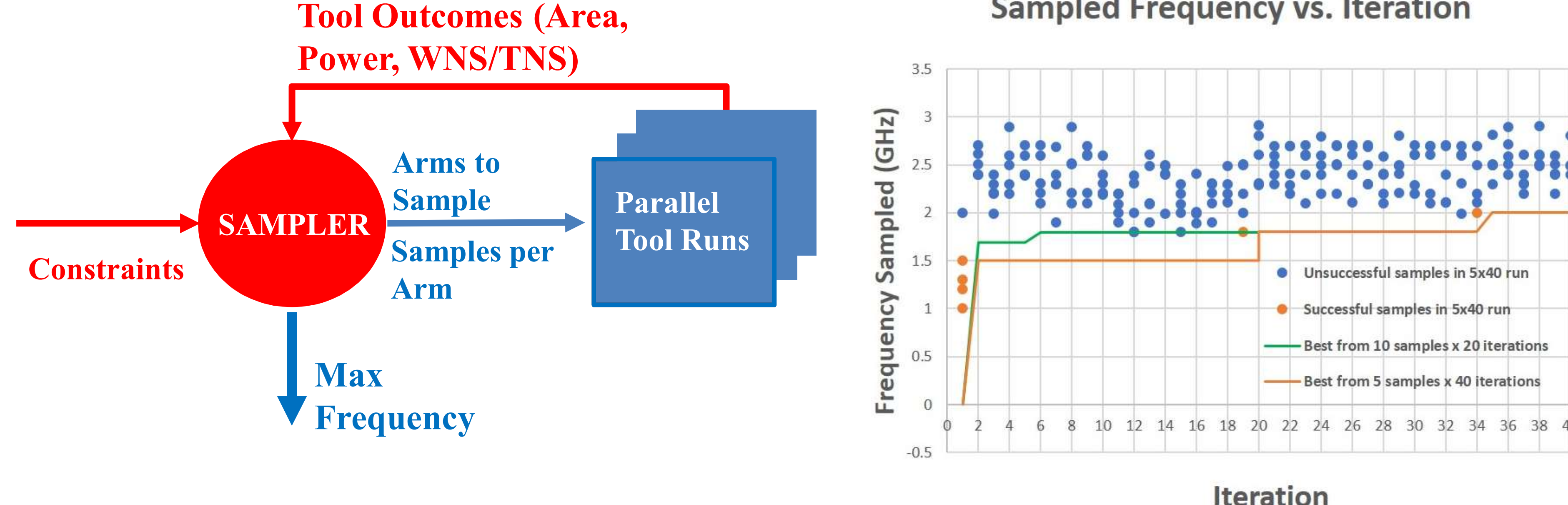
Stage 1. Mechanization and Automation

- **Create "robot IC design engineers"**
  - Observe and learn from humans
  - Search for command sequences in design tools
  - Path to "no humans" requires understanding of why, where humans are needed…

- **Path forward is through pain points**
  - **Automation of manual DRC violation fixing**
  - **Automation of manual timing closure steps**
  - **Placement of memory instances in P&R block**
  - **Package layout automation**

Example: Multi-Armed Bandit

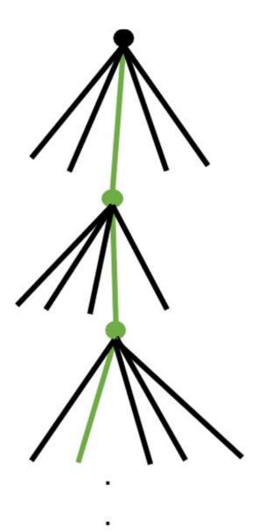
- **Multi-Armed Bandit Problem**: Given slot machine with N arms, maximize reward obtained using T pulls
  - Well-studied in context of Reinforcement Learning
- **IC Design: "arm"** = target frequency; **"pull"** = run flow



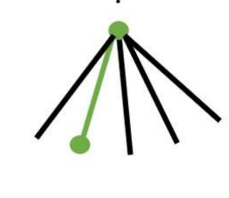
vlsicad.ucsd.edu/MAB

Stage 2. Orchestration of Search, Optimization

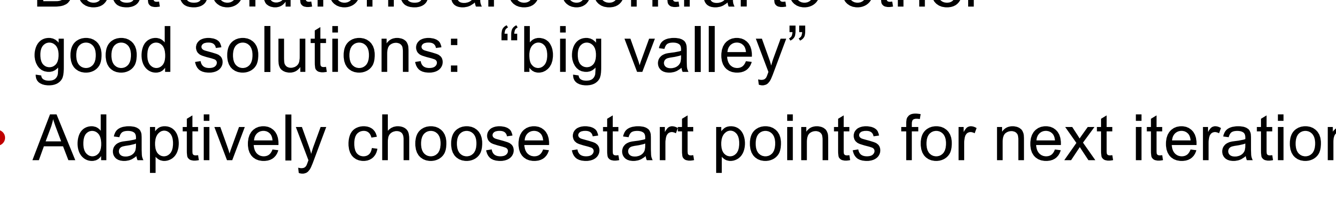
- **How to optimally orchestrate N robot engineers (s.t. risk, resource limits)**
  - Concurrent search of N flow trajectories
  - Explore, identify good flow options efficiently
  - Constraint: compute and license resources



- Example: “Go with the winners”
  - Launch multiple optimization threads
  - Periodically identify promising thread
  - Clone promising thread and terminate others



- Example: “Adaptive multi-start”
  - Best solutions are central to other good solutions: “big valley”
  - Adaptively choose start points for next iteration



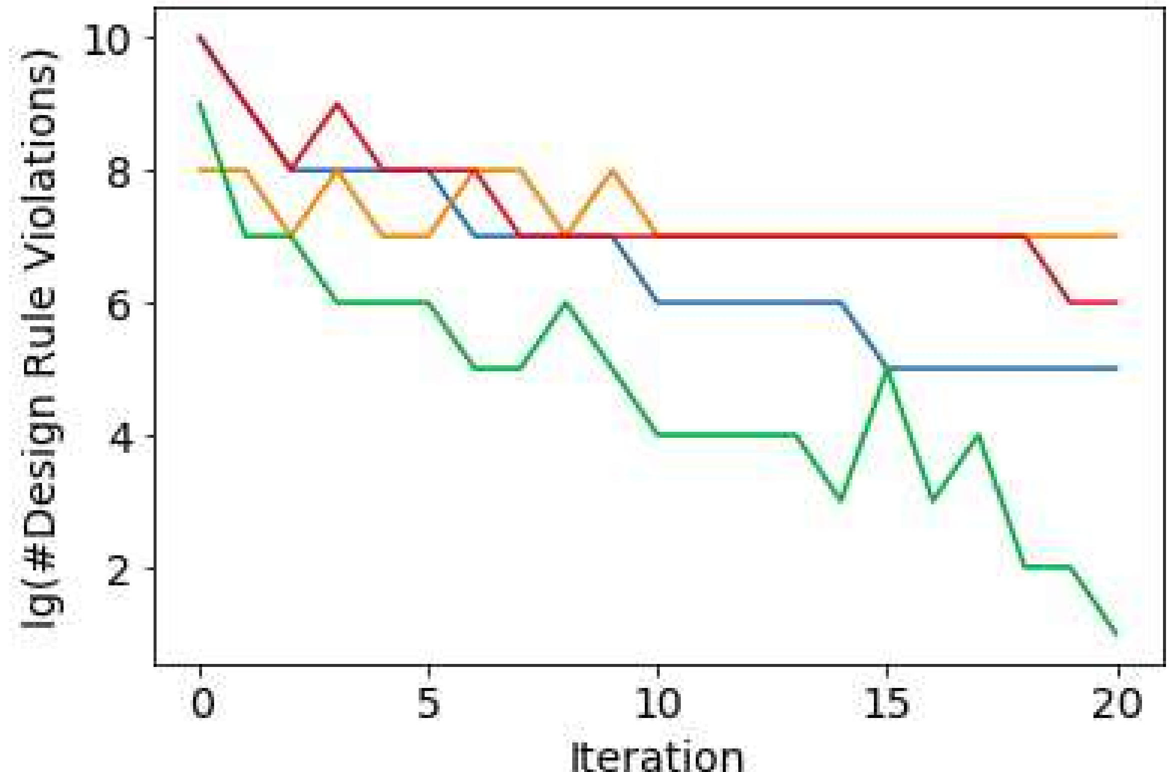
Stage 3. Modeling and Prediction

- Prediction of tool- and design-specific outcomes over longer and longer subflows
  - Wiggling of longer and longer ropes
- Enables **pruning** and termination → avoid wasted design resources
  - Better outcome within given resource budget
- **Complementary requirement: New heuristics and tools that are inherently more predictable and modelable.**
  - **No more chaos !**

Example: Predicting Doomed Runs

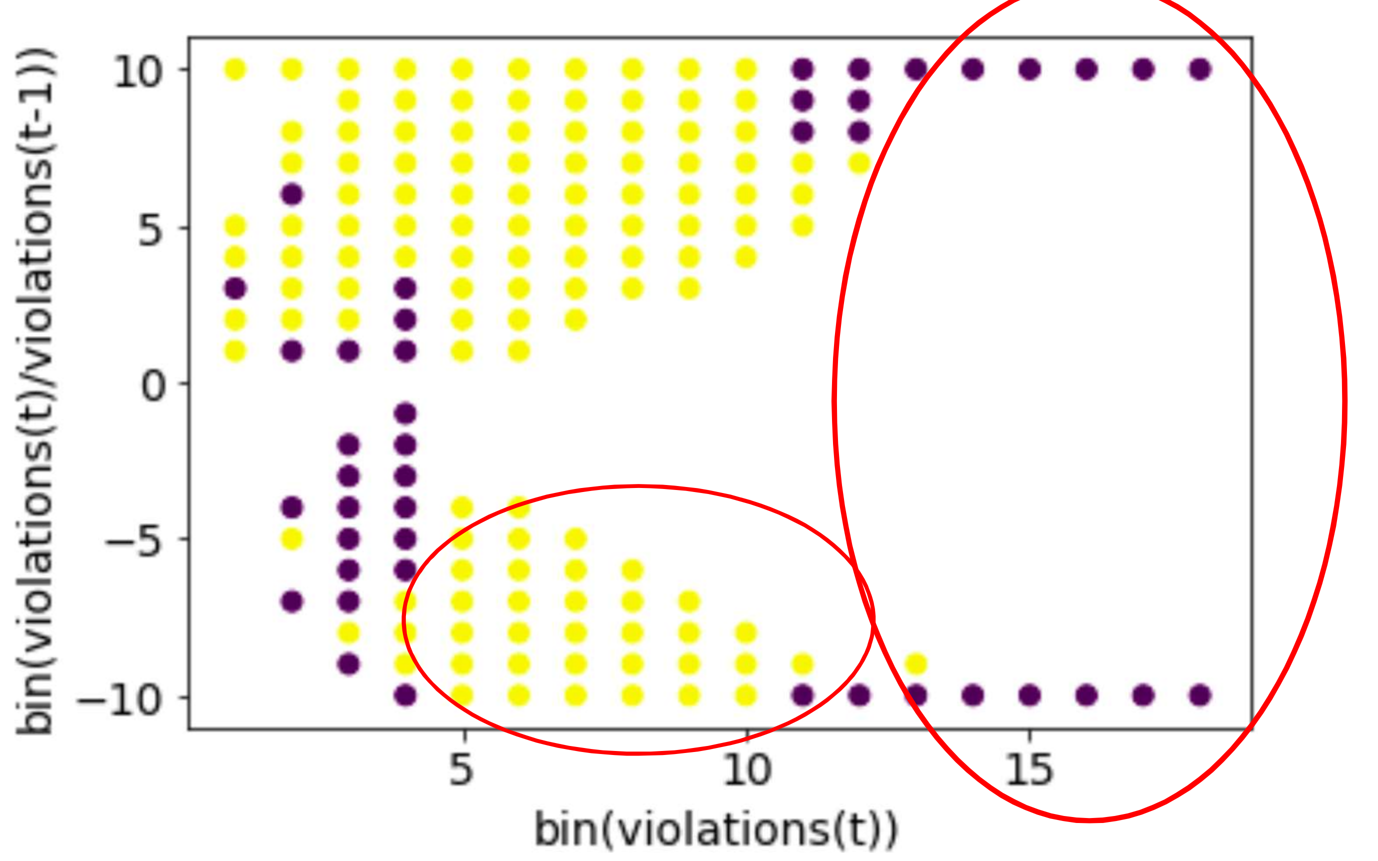
- Example: long P&R runs end up with too many post-route DRVs
- Approach: track and project metrics **as time series**
- Markov decision process (MDP): terminate "doomed runs" early

- Shown: 4 example progressions of #DRVs (commercial router)

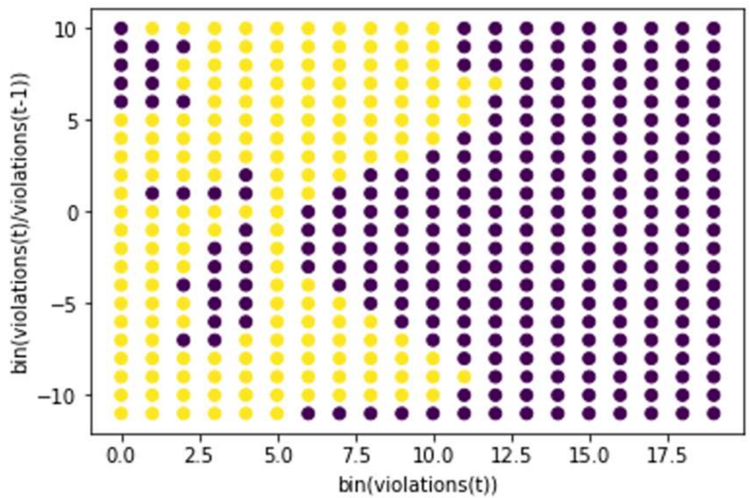


Markov Decision Process = "Strategy Card"

- **State space** from Fibonacci binning
- **Actions** – ***GO*** or ***STOP***
- **Rewards** at each state – e.g., small negative reward for *non-stop* state, large positive reward for *stop* with low #DRVs, etc.
- Automatically trained MDP "strategy card": **Yellow = GO, Purple = STOP**



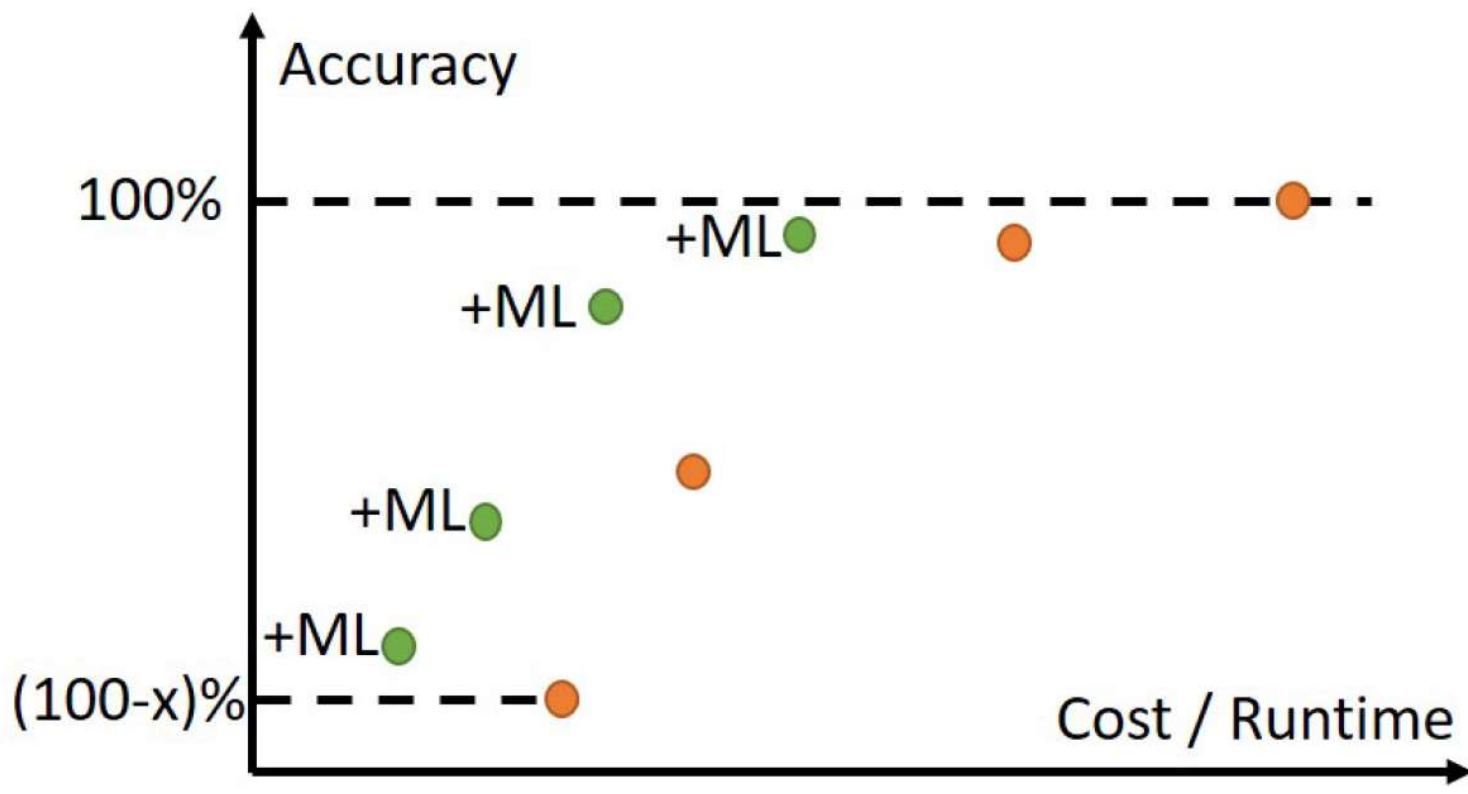
Strategy Card “Completion”



- **TYPE 1 Prediction Error:** MDP STOPs a run that will eventually succeed
- **TYPE 2 Prediction Error:** MDP predicts GO at each iteration, but run fails

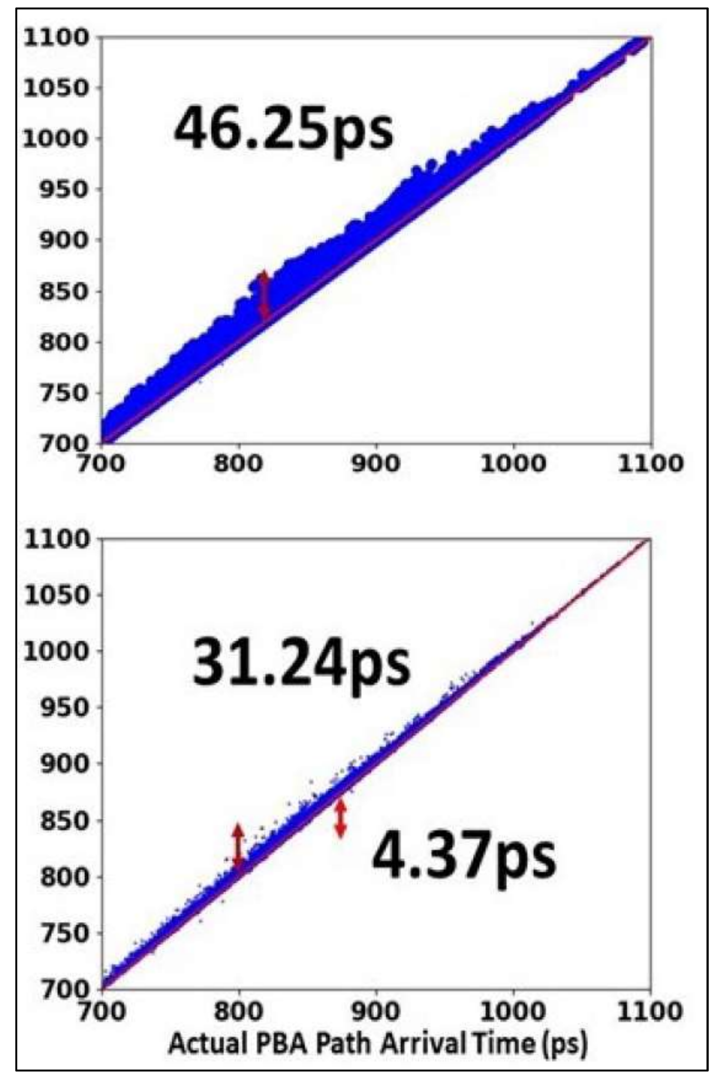
Errors	**Training (1200 logfiles)**			**Testing (3442 logfiles)**		
N = 200	Training Error	#TYPE 1	#TYPE 2	Testing Error	#TYPE 1	#TYPE 2
1 STOP	29.17%	251	99	38.3%	1317	3
2 consec STOPs	10.5%	27	99	9.0%	307	3
3 consec STOPs	**8.5%**	**3**	**99**	**4.6%**	**154**	**3**

Also: Improve Analysis Correlation (Accuracy)

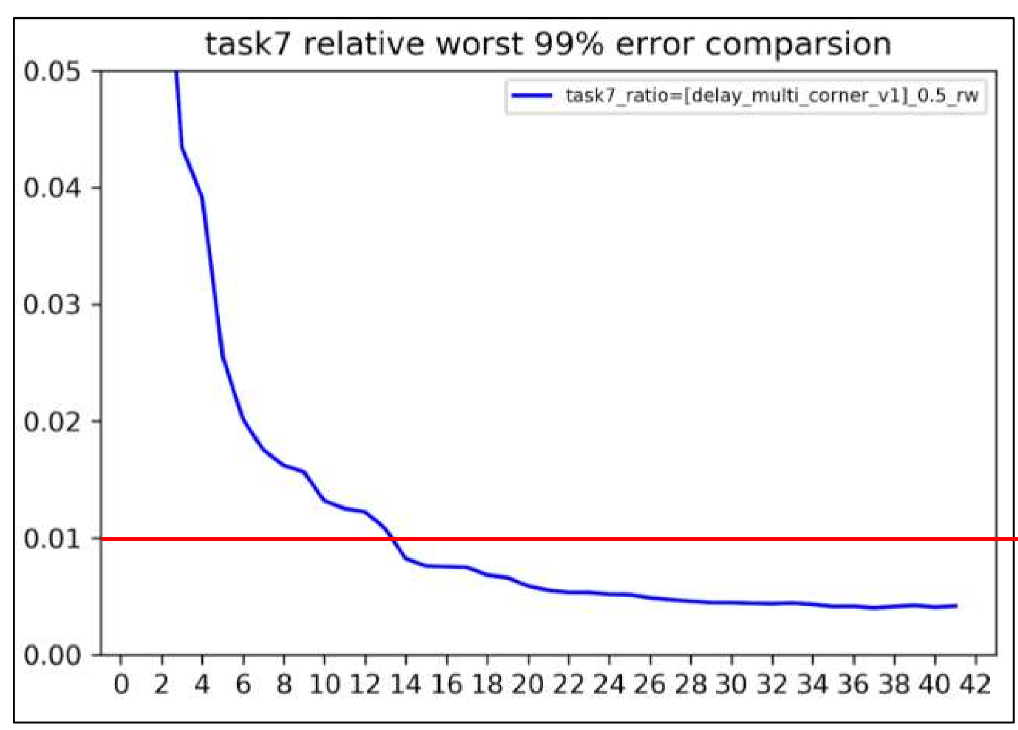


ML shifts the Accuracy-Cost Tradeoff Curve (for free) !

Examples: PBA-lite / K Corners Suffice ?



Reduced GBA pessimism vs. PBA



14 corner analyses →

<1% path delay error at non-analyzed corners

Stage 4. Learning → "Intelligence"

Many challenges on the road ahead…

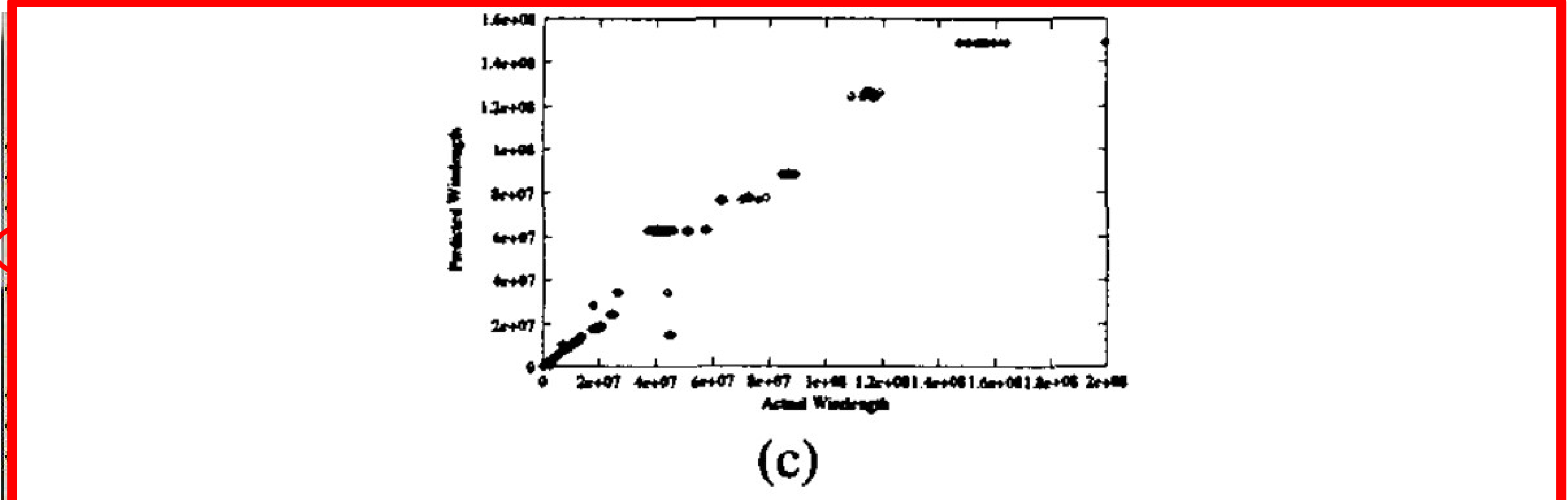
- Latency of IC design tools/flows
  - Can't "play the IC design game" 100M times in 3 days
- "Small data" context
  - Data points are expensive
  - Huge implementation space
  - Tool versions, design versions, technology all changing (pictures of cats and trees don't change)
  - Open: bridging real (top-secret!) and artificial (fake!) – e.g., with "eye charts"
- Model parameters identified using domain expertise

ML in IC Design Requires Infrastructure !

- **Support for ML in IC design**
  - Standards for model encapsulation, model application, and IP preservation when models are shared
- **Standard ML platform for EDA modeling**
  - Enablement of design metrics collection, tool/flow model generation, design-adaptive tool/flow configuration, prediction of tool/flow outcomes
  - This recalls "METRICS" http://vlsicad.ucsd.edu/GSRC/metrics
- **Modelable algorithms and tools**
  - Smoother, less chaotic outcomes than present methods
- **Datasets to support ML**
  - Real designs, Artificial designs and "Eyecharts"
  - Shared training data – e.g., analysis correlation, post-route DRV prediction, optimal sizing
  - Plus challenges and incentives: "Kaggle for ML in IC design"

(This is “METRICS” !)

[DAC00, ISQED01]



(c)

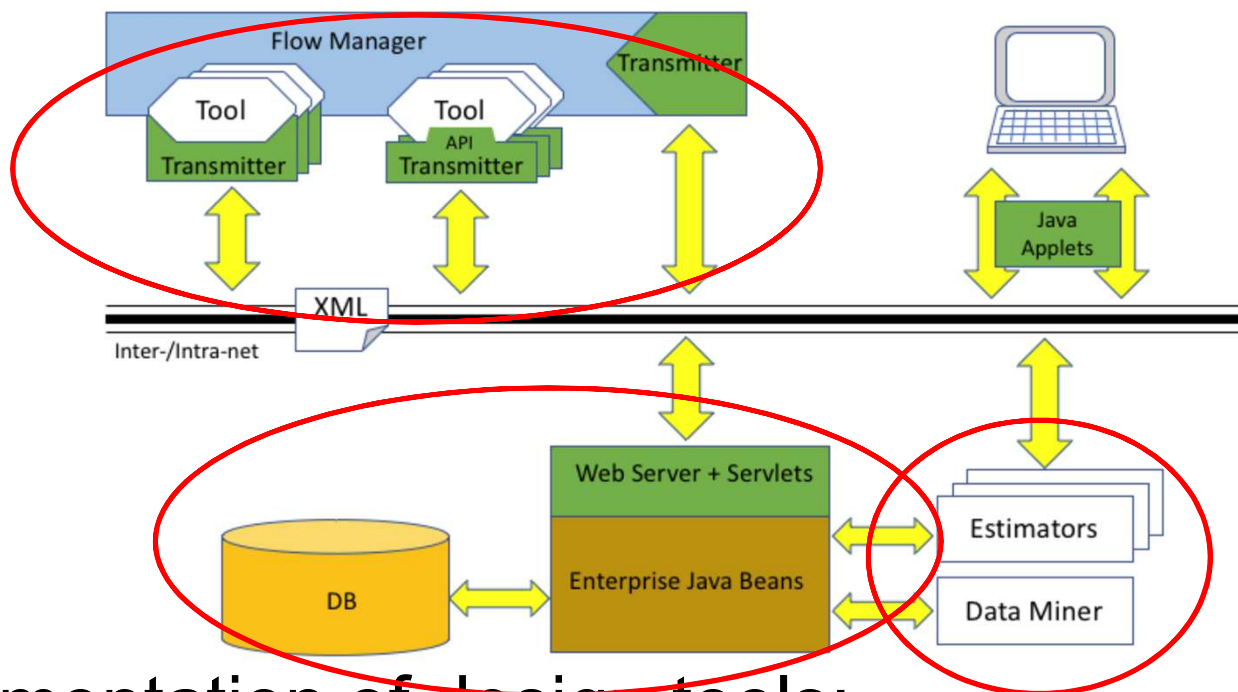
Figure 9: Predicted QPlace wirelength versus actual wirelength: (a) random, (b) distinct, and (c) representative cases.

“Measure to Improve”

- Goal #1: Predict outcome
- Goal #2: Find sweet spot (field of use) of tool, flow
- Goal #3: Dial in design-specific tool, flow knobs

http://vlsicad.ucsd.edu/GSRC/metrics

Original METRICS Architecture



- Instrumentation of design tools:
  - Wrapper scripts to extract data from outputs and logfiles,
  - Callable API codes that allow direct interaction from within the design tools
- METRICS server: central data collection (Oracle8i)
- Data mining process: analyzes existing data to improve existing design flow (CUBIST, etc.)

Lessons Learned From METRICS

- Collaboration and support from EDA needed
  - Constantly changing behaviors, logfile outputs, etc. best handled through direct integration with METRICS API
- Common METRICS vocabulary essential
  - Same semantics of crosstalk delay, vertical overcongestion, etc. across similar tools
- Must be able to adapt/evolve: recalibrate to new process, specialize to particular type of design, etc.
- **METRICS should seamlessly integrate with and drive the design flow itself**
- **Good news today:**
  - **(1) This is critical to do. (2) Social barriers are gone. (3) Many commodity building blocks for METRICS 2.0. (4) Open source is a viable path to all of this.**