

The Challenges of Physical Implementation for Advanced Technology

Miao Liu
Product Engineering Group Director
Cadence Design System, Inc.

Semiconductor Evolution

Left Brain:
Computation



Right Brain:
Sensing

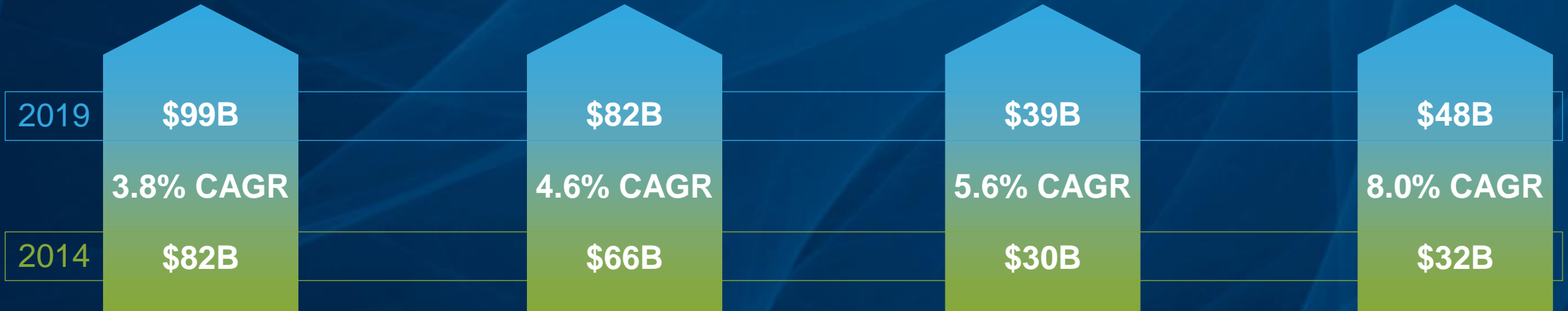
Complex Digital
Advanced Networks Digital

High-performance Custom/
analog and analog

*Mobile, Data Center
Network Infrastructure*

Automotive, Industrial

Vertical markets – drivers of intelligence and devices



Mobile



Battery life
Size
Connectivity
Video

Datacenter/Cloud



High performance
Energy efficiency
High throughput

Automotive



Sensor fusion
Mixed signal
Reliability

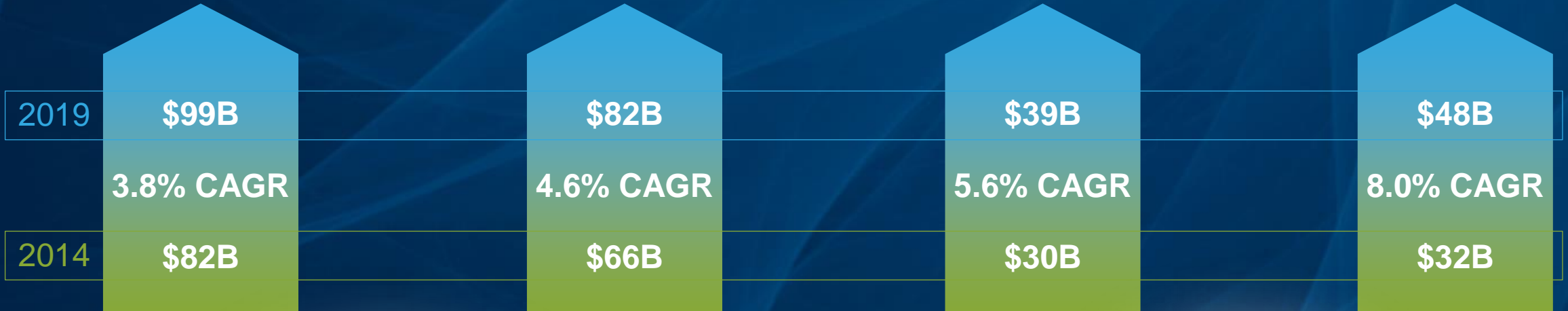
Industrial



Low power
Analog design
Connectivity

source: Gartner Semiconductor Forecast 2015 Q4

Drivers of intelligence and devices



Mobile



Battery
Size
Connectivity
Video

**Complex Digital
Advanced node**

High performance
Energy efficiency
High throughput

Datacenter/Cloud



Automotive



Sensor fusion
Mixed signal
Reliability

**High-performance
analog and digital**

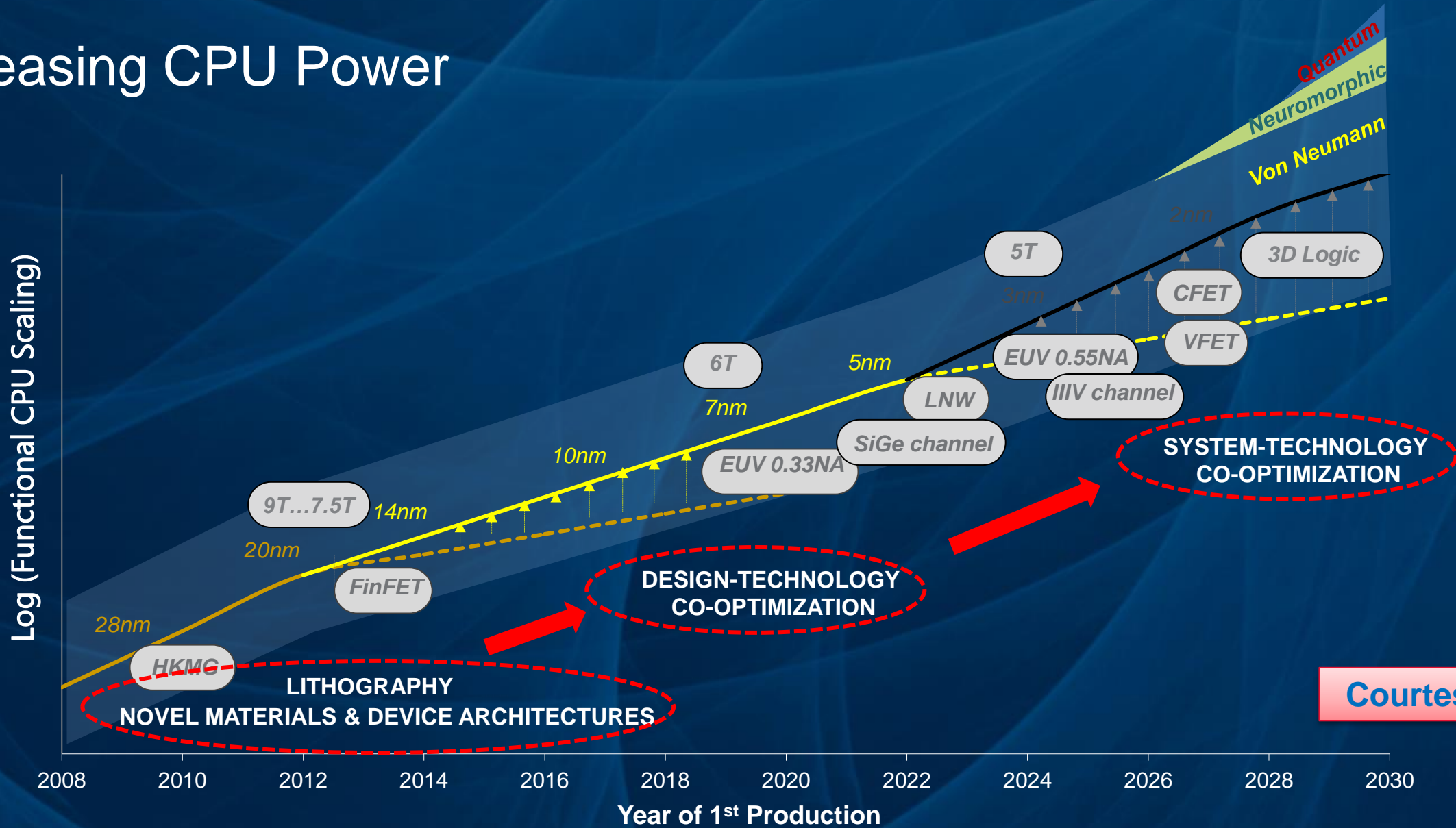
Low power
Analog design
Connectivity

Industrial



source: Gartner Semiconductor Forecast 2015 Q4

Increasing CPU Power



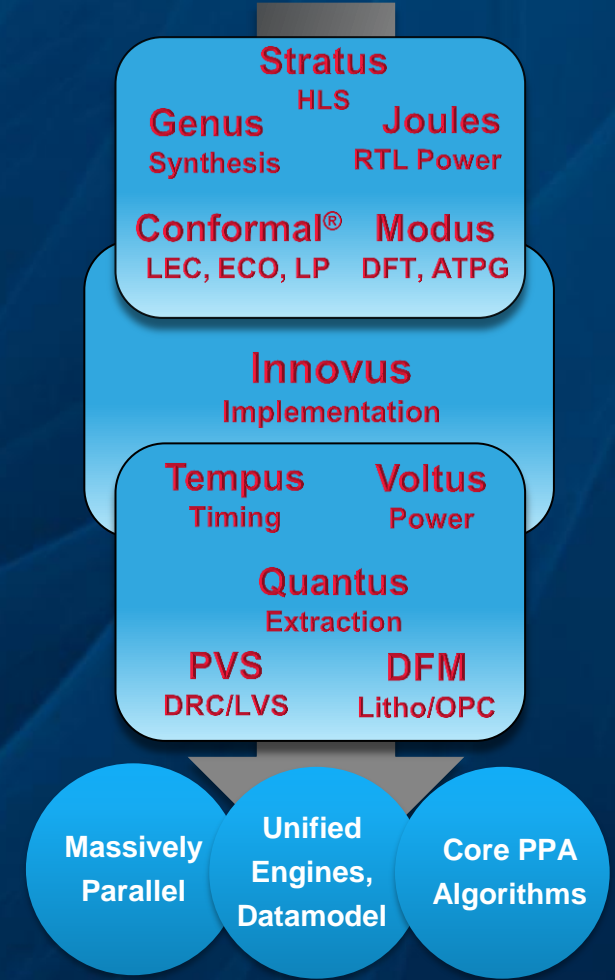
Courtesy imec

LNW = lateral nanowire, CFET = Complementary FET (n-on-p), VFET = Vertical FET (Nanowire), 6T = 6 Track Cells, 5T = 5 Track Standard Cells

Revolutionizing the Digital Flow

Innovation continues...

Cadence® Full-Flow Digital Solution

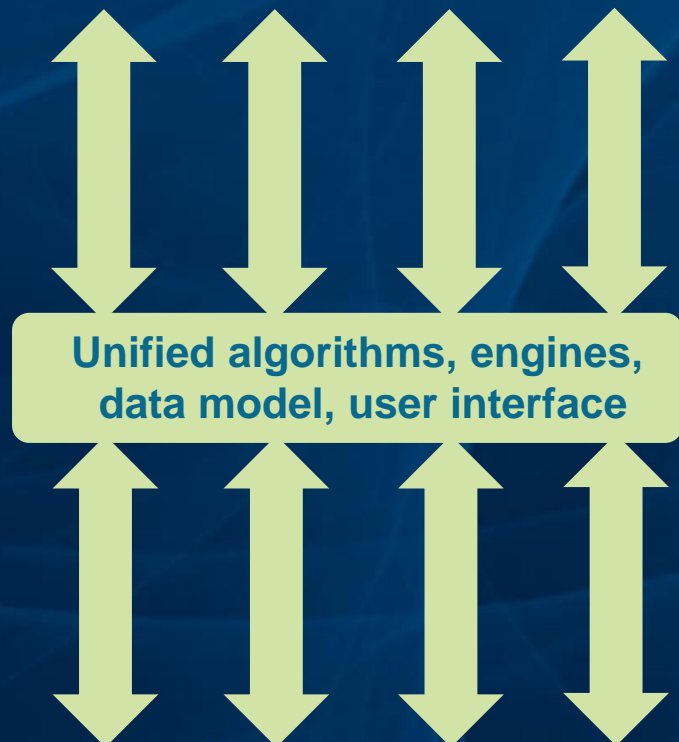
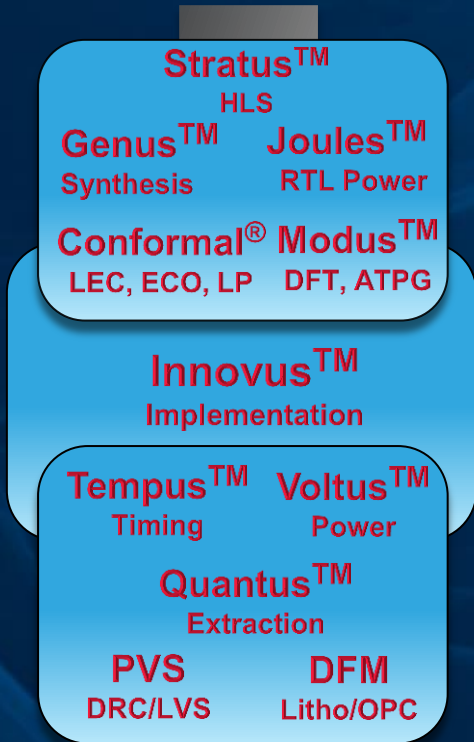


Differentiating Full-Flow Technologies

Design Convergence

Full-Flow Correlation and Ease-of-Use

Cadence® Full-flow Digital Solution



Integration, ECO, and avoidance

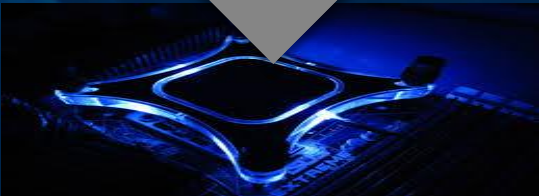
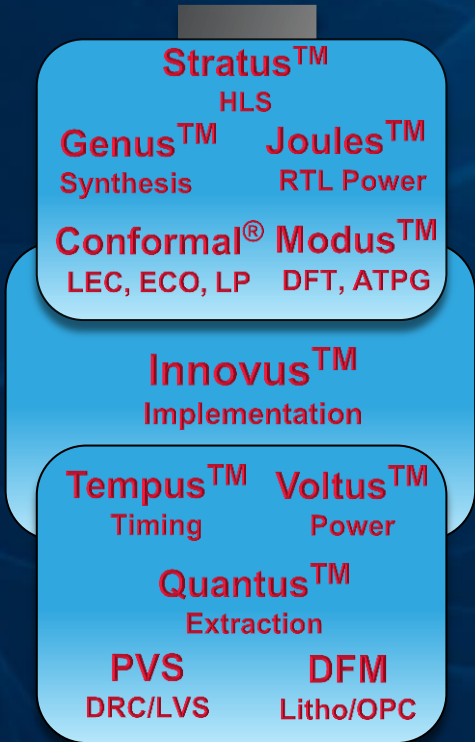
- Genus + Modus solutions:
 - Physically aware test insertion
- Innovus + Tempus solutions:
 - Physically aware timing ECO
- Innovus + Voltus solutions:
 - Physically aware EMIR ECO fixing
- Tempus + Voltus solutions:
 - Variation-aware STA and power signoff
- Full digital and mixed-signal signoff:
 - Tempus, Voltus, and Quantus™ solutions, Cadence® Physical Verification Solution

Tight integration and correlation throughout the flow are critical to best PPA

Full-Flow High Performance Design

**10 – 20%
Better PPA**

Cadence® Full-flow Digital Solution



★ **Genus™ Synthesis Solution, architecture-level PPA optimization**

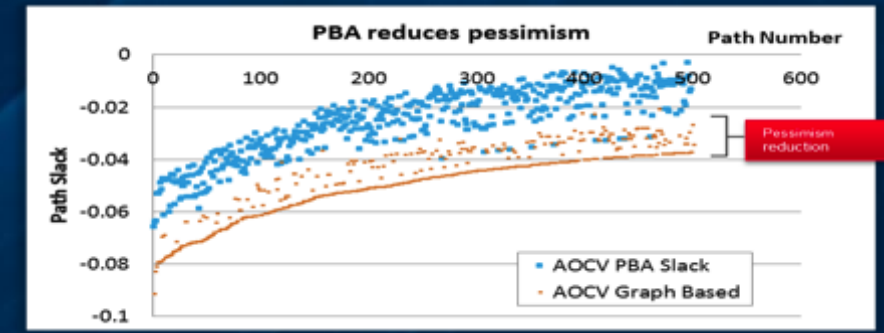
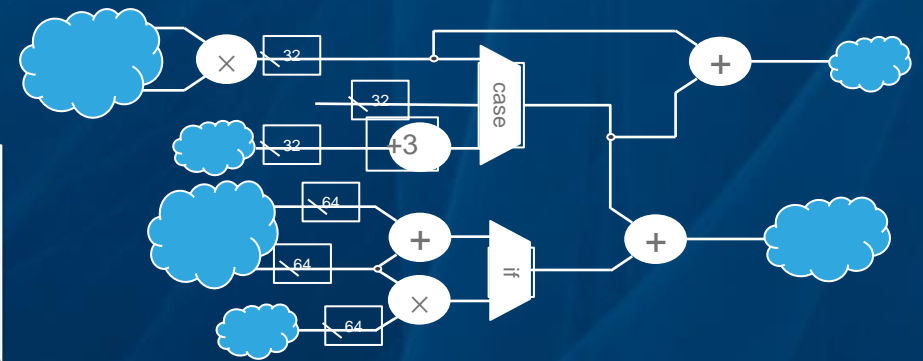
- Globally focused datapath architecture selection

★ **Innovus™ Implementation System, block-level PPA optimization**

- New GigaPlace™, GigaOpt™ power-driven solutions, CCOpt™ FlexH solution

★ **Tempus™ Timing Signoff Solution, path-based PPA analysis and ECO**

- Reduced pessimism improves PPA



Full-Flow Massively Parallel Architecture

Up to 10X
TAT/Capacity Gain

Cadence® Full-Flow Digital Solution

- 3–5X faster runtime
- 2X fewer iterations unit ↔ block level
- Scalable beyond 10M instances flat

Genus™ **Joules™**
Synthesis RTL Power

- 20X speedup for RTL-level time-based power analysis and stimulus processing
- 1M+ instances / hour prototype synthesis

- Automated ECOs of 100s/1000s of cells
- Not possible by hand

Conformal® **Modus™**
LEC, ECO, LP DFT, ATPG, Diagnosis

- Up to 3X digital test time reduction
- Up to 2.6X codec wirelength reduction
- Near linear up to 32 CPUs/machines

- Up to 10X faster design closure
- Scalable to 100s of CPUs
- Distributed graph/path based STA

Innovus™
Implementation

- Any design can be run in few hours
- Scales up to 100s of CPUs/machines
- Capacity for designs up to 2B inst.

- Scalable to 100s of CPUs and machines
- Multi-corners and incremental extraction
- Smallest netlist → faster simulation

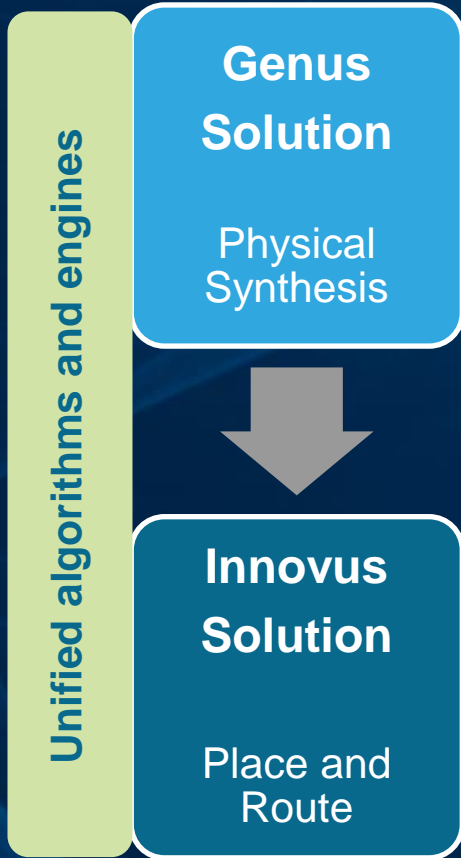
Tempus™ **Voltus™**
Timing Power

- 5–10X TAT and capacity gain
- Multi (MMMC) scenario acceleration
- Handles 5–10M+ instance blocks

Quantus™ **DFM**
Extraction Litho/OPC

- Near-linear scalability up to 100s of CPUs/machines
- Automated fixing with min. design changes
- ECO-mode after design changes

Genus Synthesis Solution Summary



Massively Parallel

- Timing-driven distribution across multiple machines and CPUs
- **Up to 5X faster turnaround time scaling to 10M+ instances**



Tightly Correlated

- Unified routing, extraction, and delay calculation with Innovus™ Implementation System
- **Timing, wirelength, and congestion within 5%**



Architecture Level PPA

- Globally focused datapath architecture selection
- **Up to 10-20% reduction in datapath area and power**



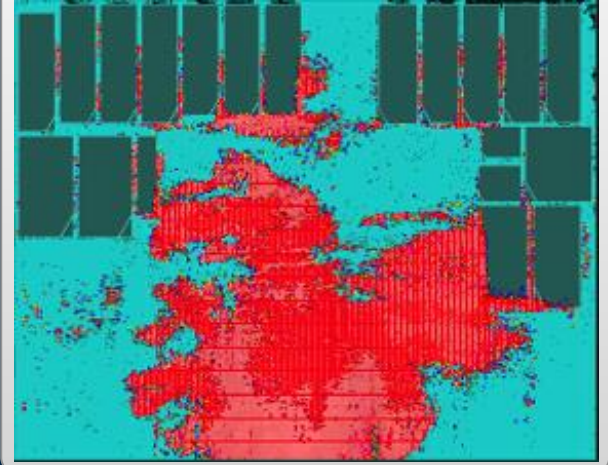
RTL Designer Focus

- Generate full timing and physical context for any subset of a design
- **2X fewer iterations between unit level and block level**

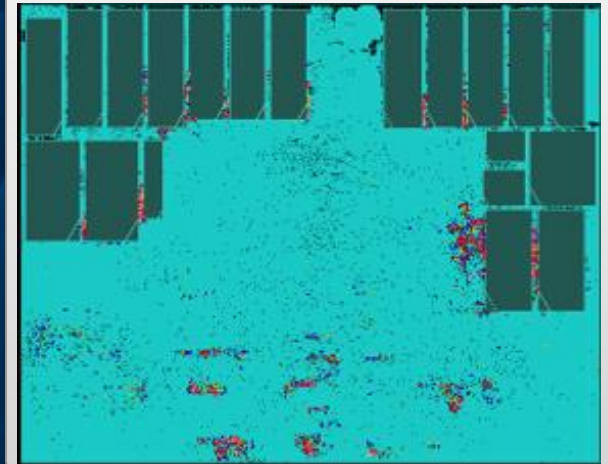
Cisco adopts Cadence Physical Synthesis Genus Eliminates Congestion

- Block Details:
 - Networking SoC
 - ~1M Instance block
 - MUX heavy
- Initial results:
 - 17.79% H
 - 15.67% V
 - Not Routeable!
- Using Physical Aware Structuring (PAS)
 - 0.09% H
 - 0.73% V
 - Routeable, met timing & area goals

Initial Congestion Map:



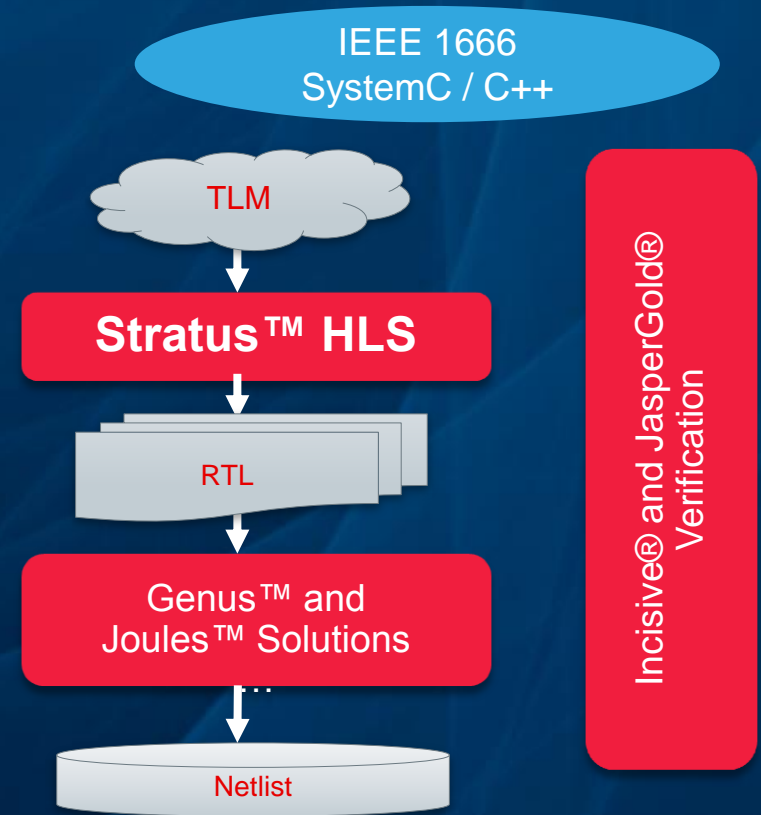
Phys. Synth. Congestion Map:



Stratus High-Level Synthesis (HLS)



- 10X productivity and IP reuse
 - Separates functionality from implementation
 - Re-targetable high-level IP
- 5X faster, better verification
 - Fast high-level simulation models
 - Consistent comprehensive verification platform
- 20% better quality of results
 - Power/area/performance
 - Beat hand-coded RTL results
- Broad applicability
 - Can be utilized across an entire system-on-chip (SoC) design
 - Addresses real-world challenges of ECO, low power, routing congestion, and IP reuse



Stratus Case Study: Socionext

First to market with HEVC/H.265 encoder via physically aware HLS

Challenge:

- Be first to market with 4K/p60 HEVC video encoder chip

Solution:

- Use **TLM interface IP**, **architectural exploration**, and **physically aware HLS** flow

Results:

- Identified and completely **removed congestion** for 79 modules via HLS
- For most congested modules **35% area reduction** and **69% net length reduction**
- First silicon was successful

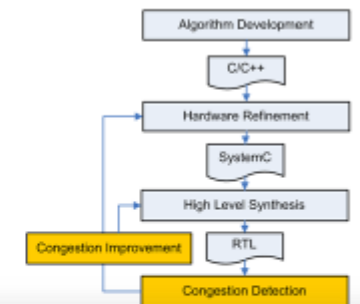
Congestion Improvement Results

Applied to 79 modules and improved the congestion in 6 modules



Physically Aware High-level Synthesis Design Flow

- Introduce 2 additional steps before P&R phase to target congestion
 - Congestion detection
 - Congestion improvement
- Integrate with existing physical aware logic synthesis tool to detect congestion and find its root cause in SystemC
- Apply to a real design to confirm the efficiency

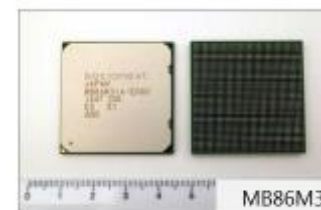


4K/p60 HEVC Video Encoder Achieved First Silicon Success

The new device can perform real-time encoding with a single chip, enabling a dramatic reduction in the equipment's size and power consumption. The new device has only 1/20th the volume and 1/50th the power of a multiprocessor system, enhancing performance 1,000 times.

MB86M31 Features

- Real-time HEVC/p60 encoding of 1 channel x 4K (3840 x 2160) images
- Real-time HEVC/p60 encoding of 4 channels x full HD (1920 x 1080) images
- Input format: YUV 4:2:2 10bit progressive / interlace
- Original HEVC codec core (co-developed with Fujitsu Laboratories Ltd.)
- Package: FCBGA-1764 35mm x 35mm



MB86M31



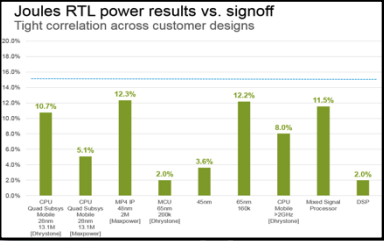
Evaluation Board MB86M31-EVB

Joules RTL Power Solution – RTL Power Analysis

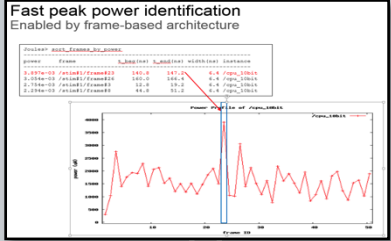
Within 15% of signoff and up to 20X faster time-based power



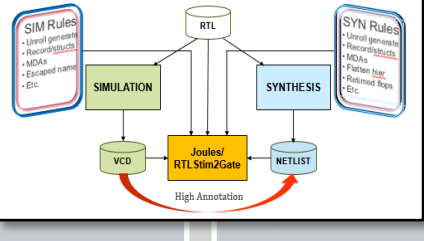
RTL power accuracy
Within 15% of signoff



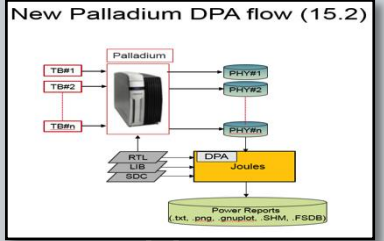
Stimuli management
Up to 20X faster frame-based power



Power bridge for RTL2GDS flow
Eliminate GL sims.



System-level power
~ 15 hours savings for 1ms activity window



Credible synthesis
(Genus™ solution inside)

- SDC and timing
- Clock tree
- P&R buffers
- DFT
- CPF/UPF

Multiple stimulus, customizable framing

- Peak power frame identification
- Aggregate SoC-level power using block-level activity

Connecting verification, implementation, and signoff

- Apply RTL activity on netlist
- Generate peak activity for signoff

Direct read of Palladium® PHY DB

- Save time writing, reading switching activity files (VCD, SAIF, etc.)
- Power profile for AnTuTu and other system-level tests

Joules Solution is Fast and Accurate, and Finds Power Bugs

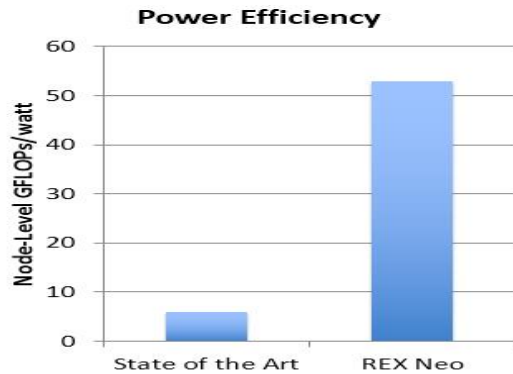
CDNLive 2016 Silicon Valley – Best Paper Award – FED Track

Introducing the REX Neo architecture

Complete computing solution designed for 21st century realities



- 256 core MIMD Compute Chip
- 2D Mesh Network on Chip
- Software managed memory system
- Global Flat Address Space

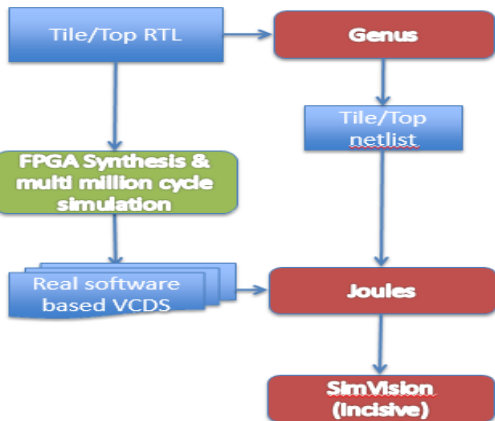


© 2016 REX Computing

11

System-level power analysis flow with Joules

Cycle-based analysis with application software running



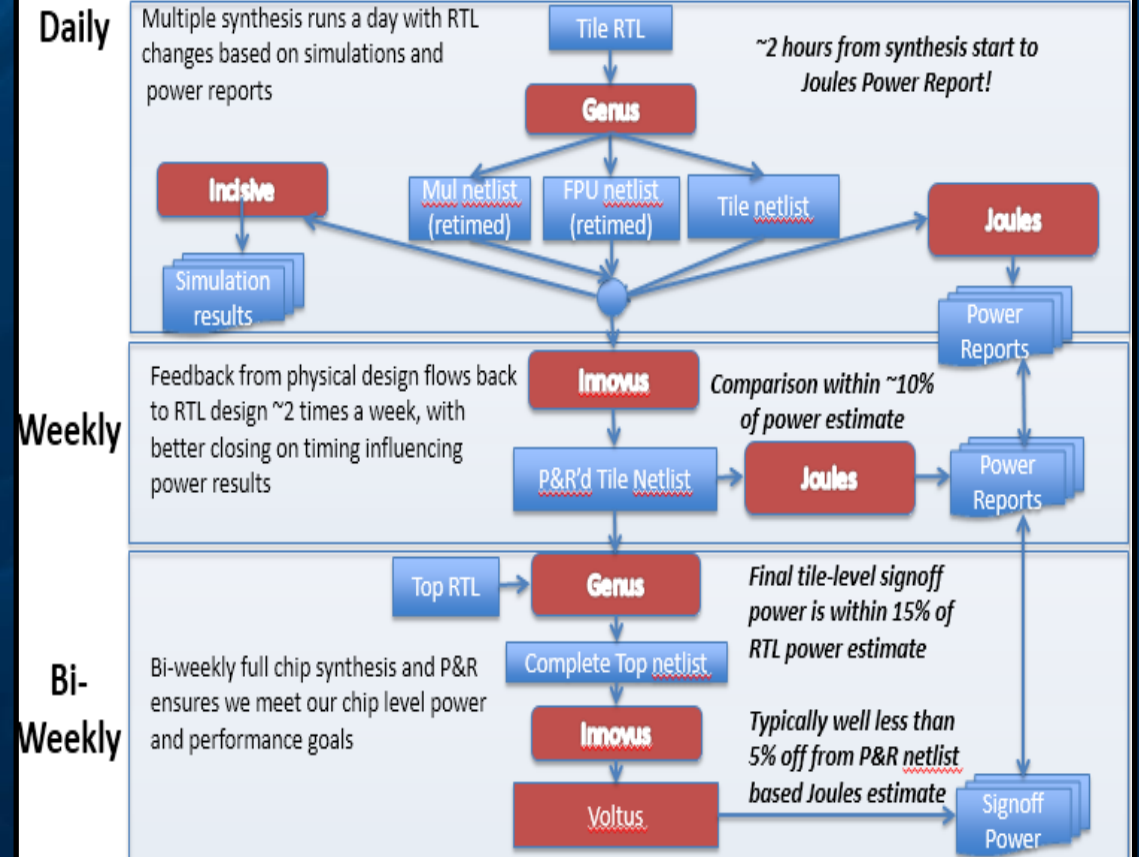
- Not possible without Joules
 - No good way to determine activity factors without running software
- Flow instrumental in finding design/power issues exposed by application software
 - Wasted power scenarios detected

© 2016 REX Computing

24

Joules RTL power is within 15% of signoff power

Fast, accurate and useful



© 2016 REX Computing

22

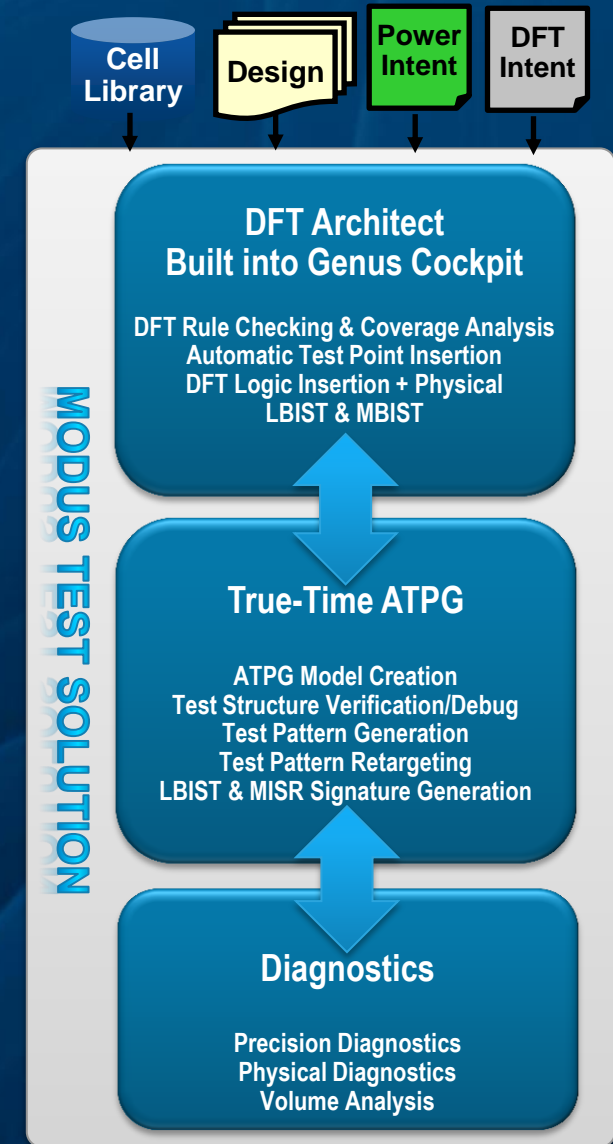
Modus Test Solution – Overview

- **Solution highlights**

- Comprehensive manufacturing test offering for logic and embedded memory IP and stacked die 2.5/3D IC
- Single-pass logic synthesis, test insertion, and pattern generation, that is physically, clock, and power domain-aware
- Reduced cost of defect detection with higher quality test patterns, accurate silicon defect diagnostics, and volume analysis

- **Solution components**

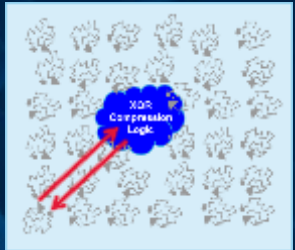
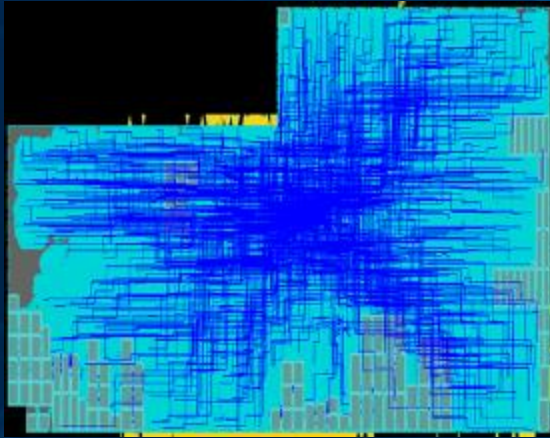
- ✓ *Modus DFT Architect tightly integrated into Genus Synthesis Solution*
- ✓ *Modus DFT Architect Logic BIST*
- ✓ *Modus Advanced Memory BIST*
- ✓ *Modus Test True-Time ATPG*
- ✓ *Modus Diagnostics Precision and Volume*
- ✓ *Modus Hierarchical Test*



Modus Test Solution – 2D Elastic Innovation



Current Industry XOR Compression

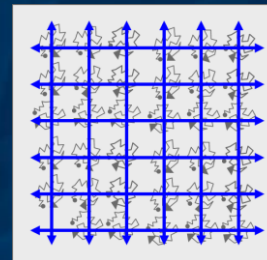
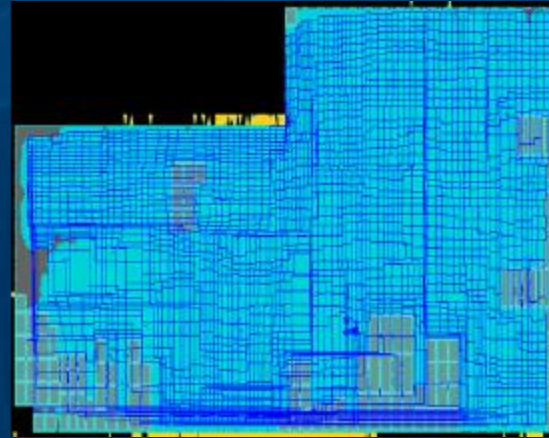


Compression logic is tangled web across the chip

3.6X reduction in test time



Modus 2D Elastic Compression



Compression logic forms an elegant grid across the chip

1.7X reduction in test time



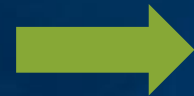
2.6X reduction in compression wirelength



2X reduction in test time

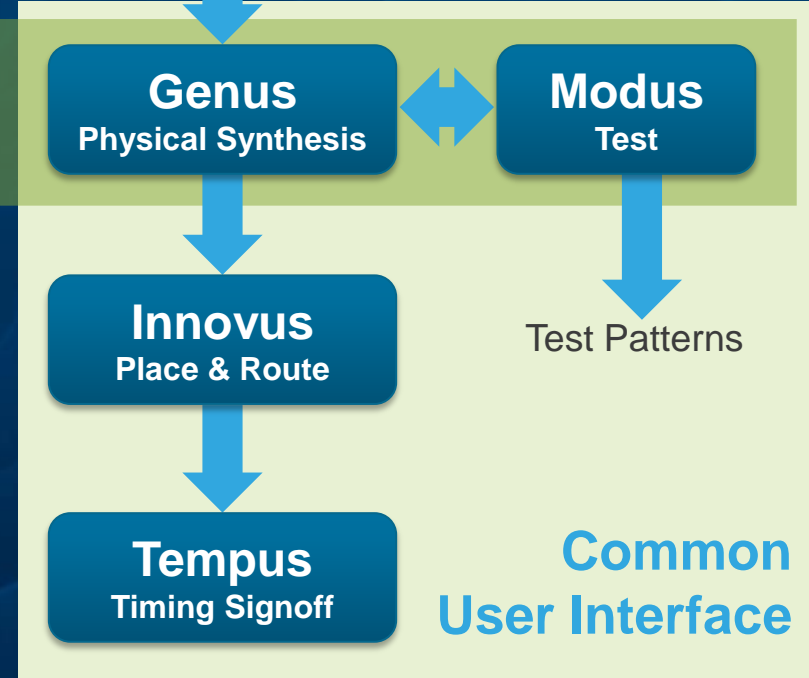


Half the test time



Half the wirelength

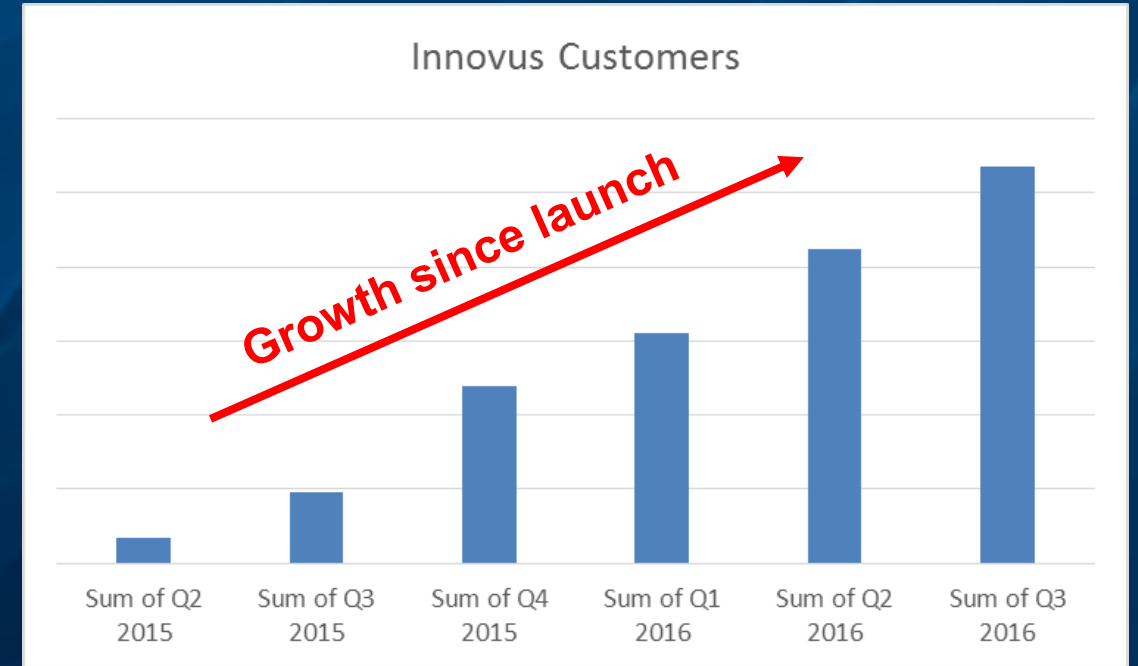
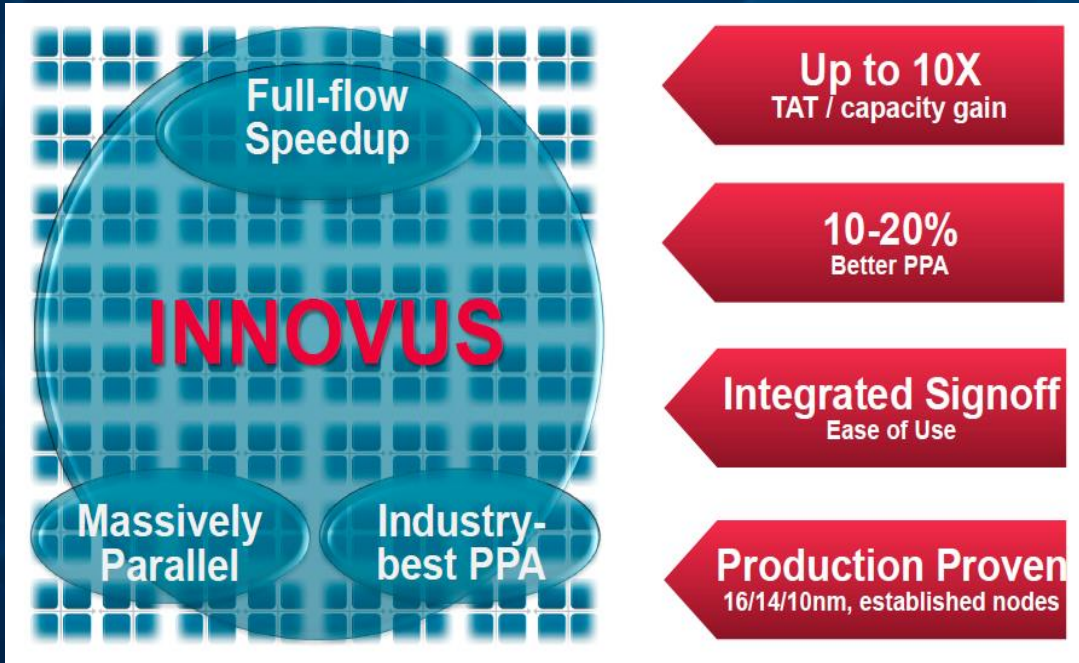
Verilog/SystemVerilog/VHDL



Innovus™ Implementation System
Launched Two Year ago ...



Innovus 2015- 2016



- Q4 2016: Record bookings quarter for Innovus.
- 7nm momentum continues to build. Over 40 customer interactions. Testchip projects at [all major semiconductors companies](#)

Innovus Adoption in DSP Block

Wireless, Datacenter and Mobile Chips
28 and moving to FinFET designs

Challenges

- Aggressive PPA targets
- Difficult to meet timing with Hierarchical Block closure
- Large number of Hard Macros

Design Details

Design Size	3.7M
Technology	TSMC 28 HPC
Metal Scheme	7 routing layers
Hard Macros	135
Target Frequency	1.16 Ghz



Flat trial run
Test capacity
PostRoute timing met

QoR Push
New FP with 10% area reduction.
Signoff Timing met

Area and Perf. Push
20% Area reduction
10% perf. improvement



Full Flat implementation runtime < 50 hrs on 16 cpus

Frequency pushed to **1.2 Ghz** with **20% Area Reduction**

Innovus deployed at 16FF+

Graphics Processor Unit (GPU) Design

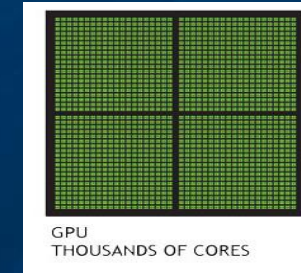
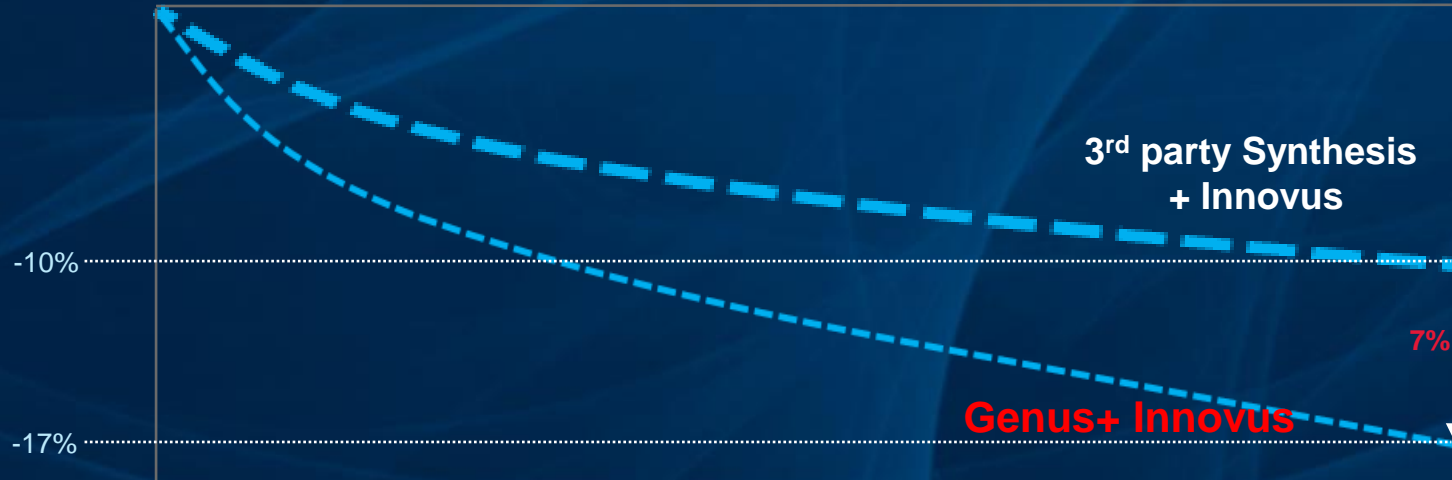
Advanced node, highly-parallel

GPU: 5M+ instances, 1GHz+

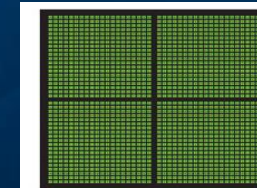
3rd party Synthesis
+ P&R

3rd party Synthesis
+ Innovus

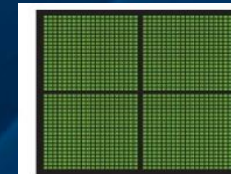
Genus+ Innovus



GPU THOUSANDS OF CORES



GPU THOUSANDS OF CORES



GPU THOUSANDS OF CORES

Shrinking
Die Size

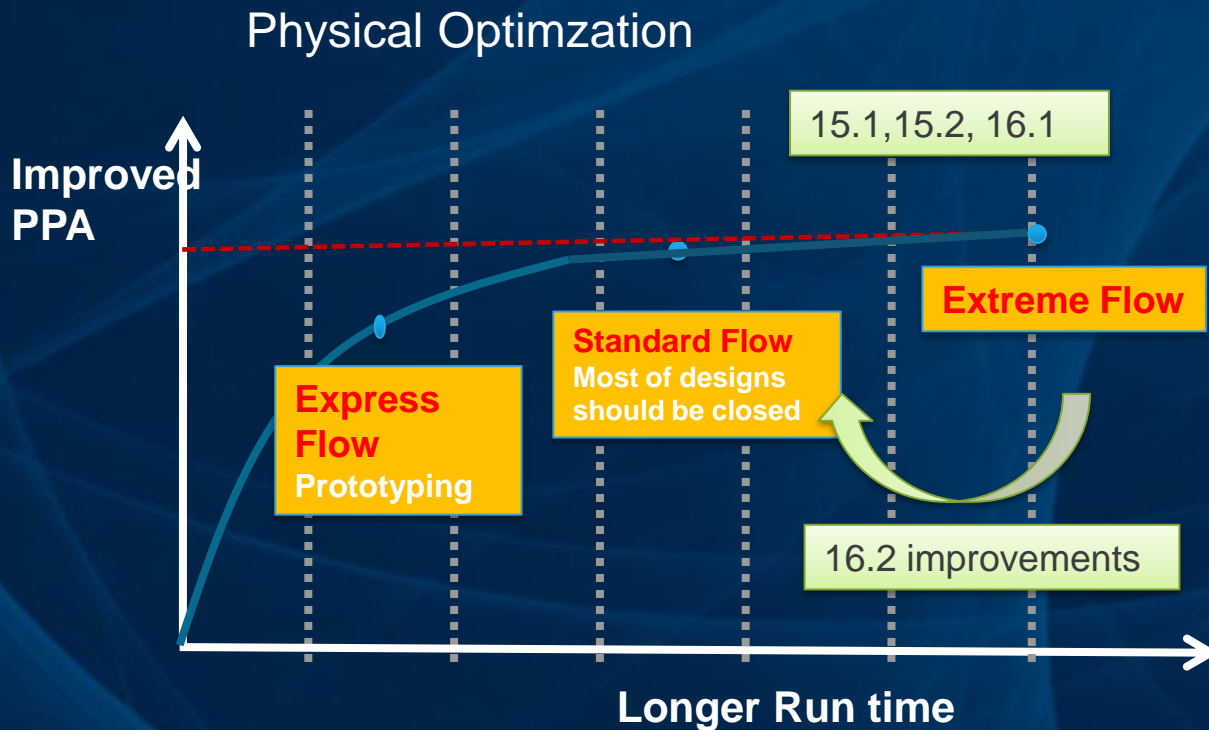
10-17% Die Size reduction
with Genus + Innovus

Digital full flow TAT <2days
(4X faster than 3rd party flow)

Predictable timing and density with
deterministic results independent of # cores

Innovus 16.2 – Improved TAT with better PPA !

- Releases prior to 16.2
 - 30%~40% TAT for final 2~5% PPA



16.2/16.21

Full flow TAT and PPA improvements

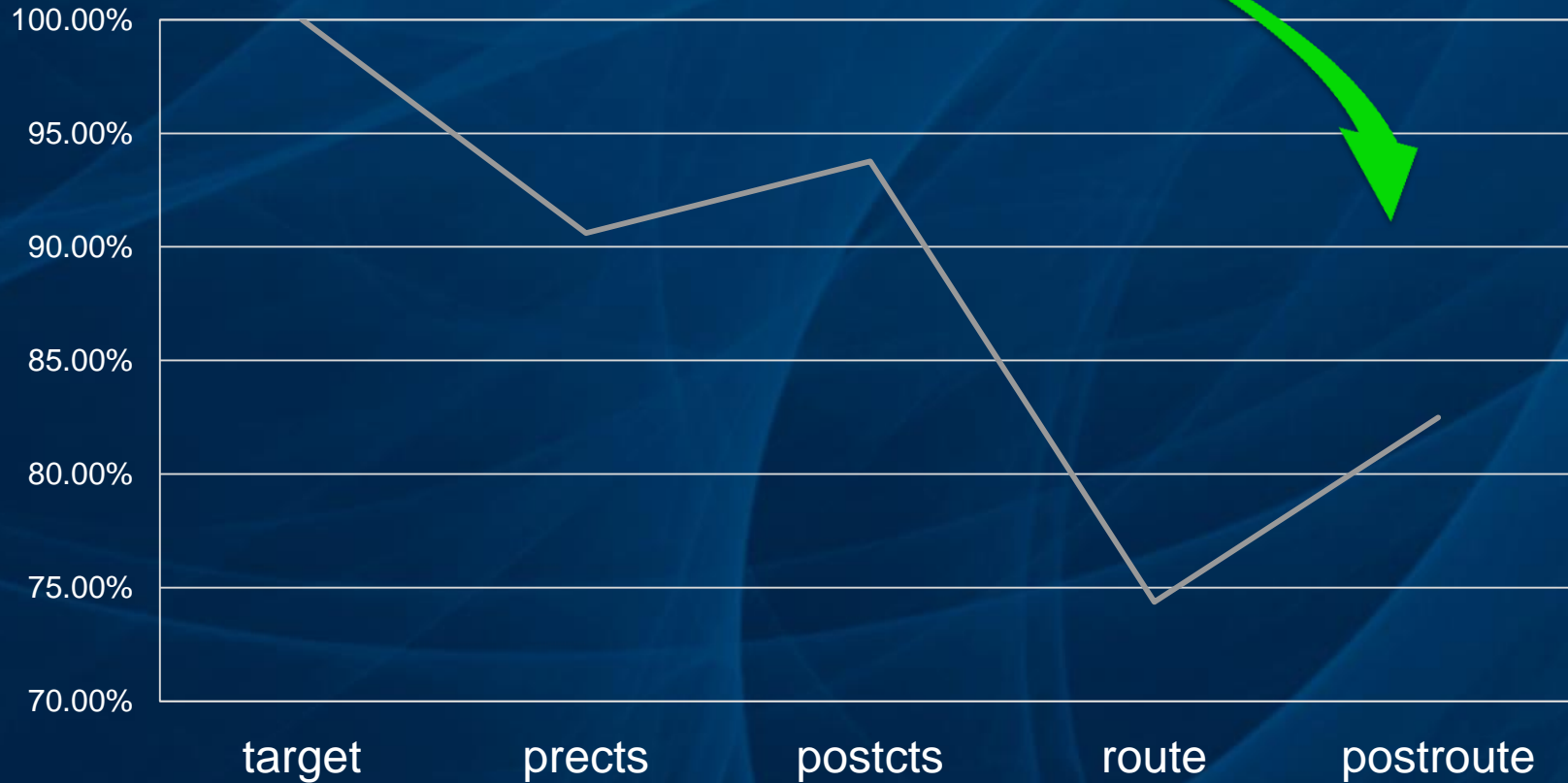
Upto 30% runtime improvement for best PPA flow.

Improved PPA (setup and density) with the same runtime.

QoR improvement of Extreme flows will be available in Standard without impact to TAT

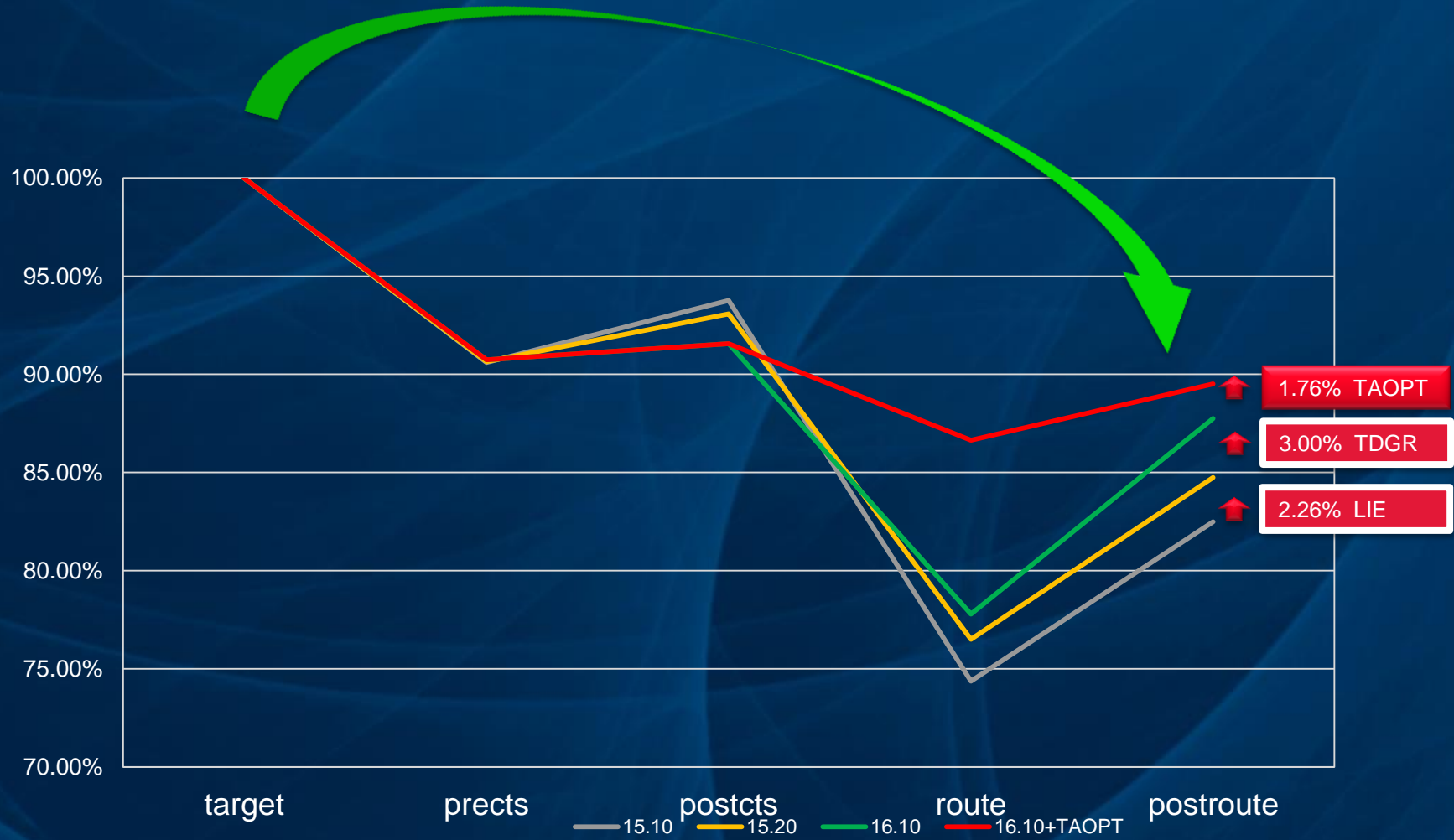
Timing Convergence

Based on 25 Designs (16nm and above)



Timing Convergence

Based on 25 Designs (16nm and above)

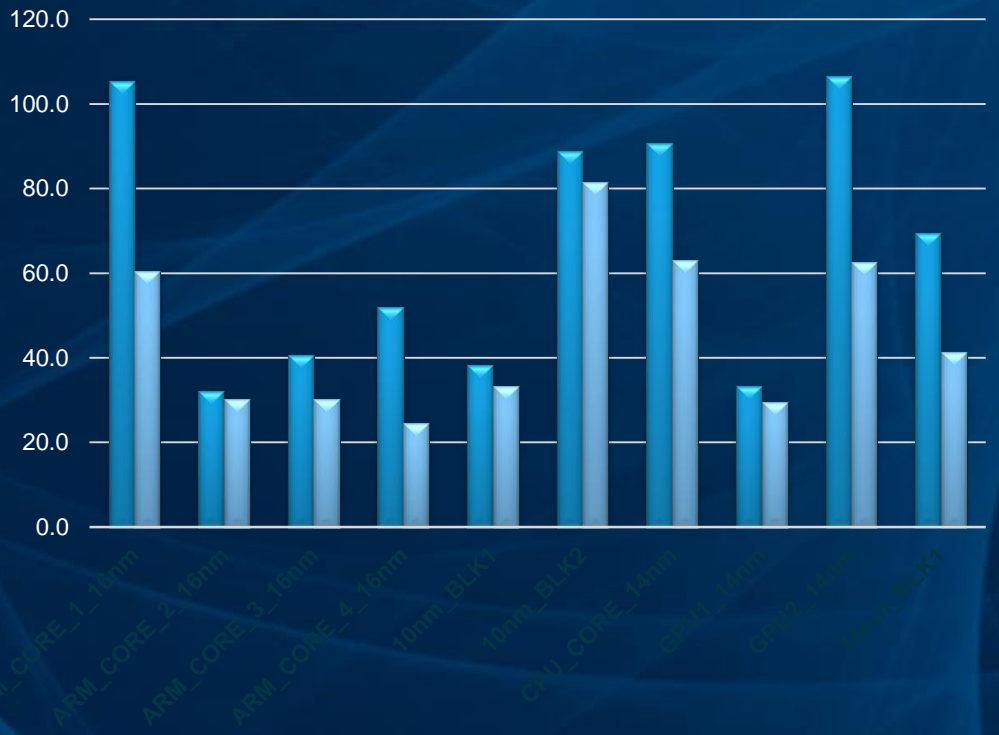


Degradation after postCTS

- 2.05% (16.1+TAOPT) ↑
- 3.81% (16.1) ↑
- 8.33% (15.2) ↑
- 11.3% (15.1) ↑

Innovus : 16.2/16.21 “Out-of-the-Box” Runtime & QoR Improvement

Performance improvement ■ 16.1 ■ 16.2_perf_flow



30-50% Runtime improvement production designs

Same or Better QoR !

- ✓ Early Clock flow enabled
- ✓ Enhanced Global View Pruning
- ✓ Convergent Gigaplace
- ✓ Early GlobalRouter for Clock Routing
- ✓ Efficient Power Optimization flow
- ✓ Global and Detailed Routing Engine improvements

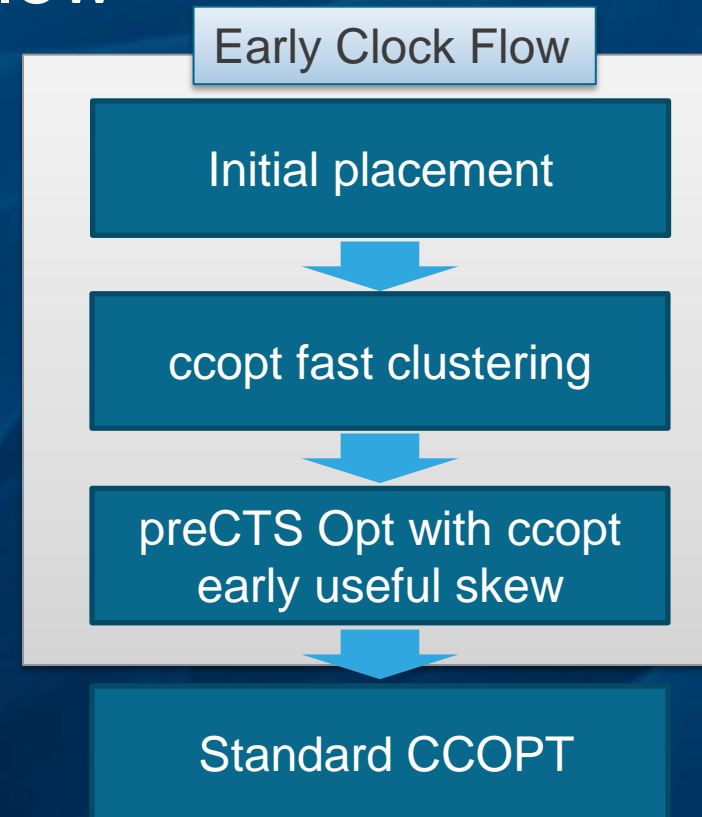
Innovus Early Clock Optimization Flow

Early Clock Flow

VS

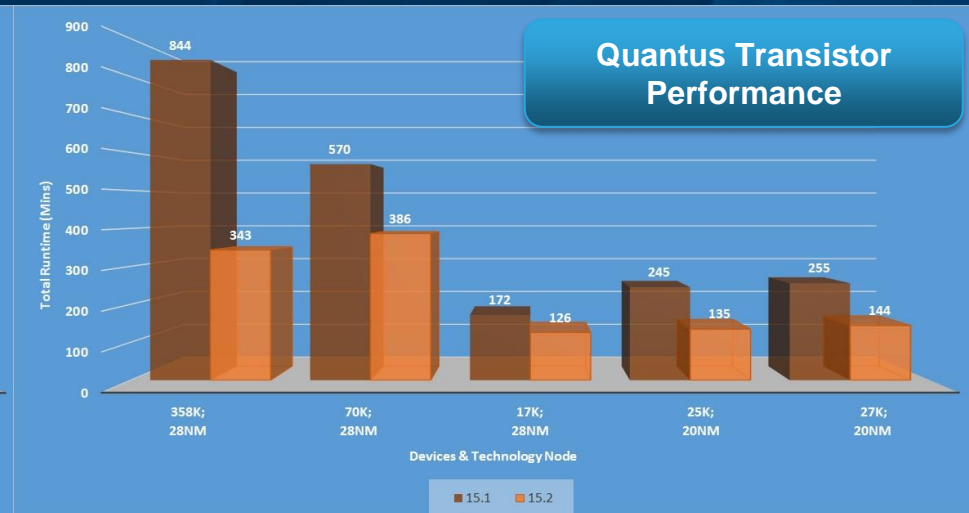
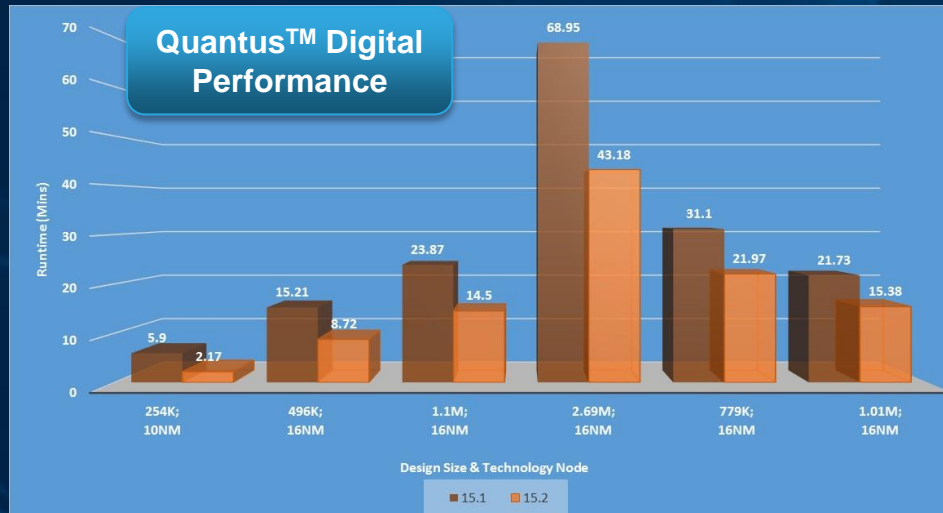
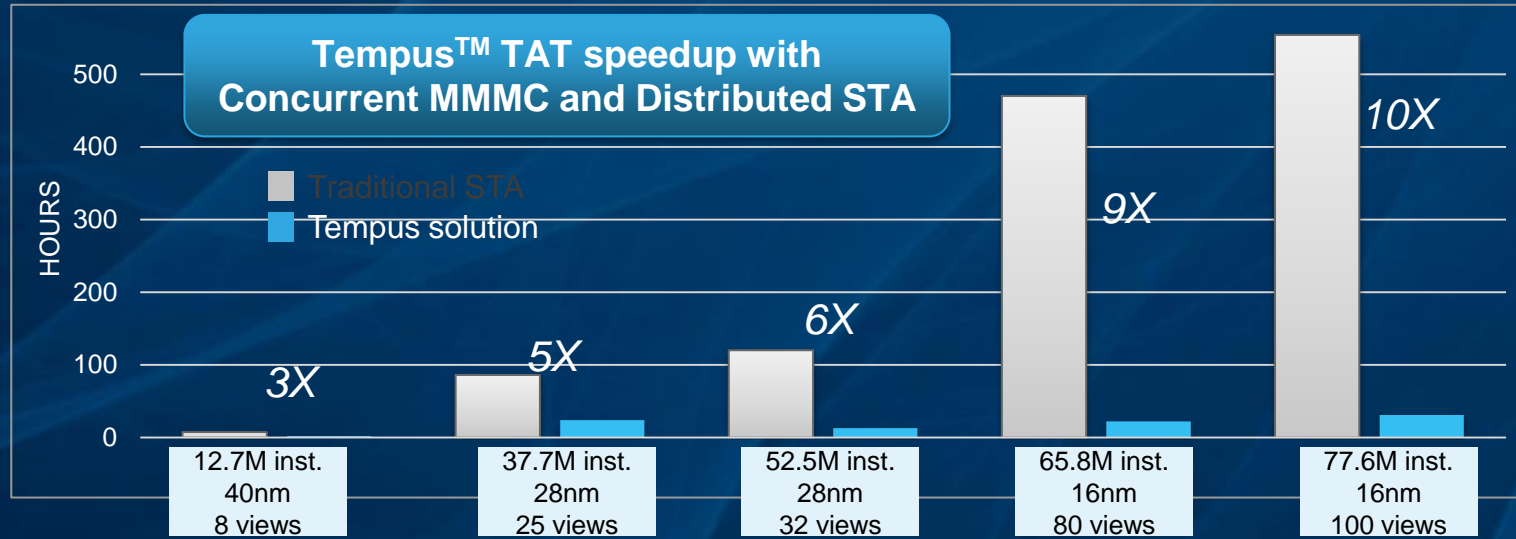
Place_opt + CCOPT Flow

- 5% better TNS.
- Great improvement in congestion hot spot after CCOPT.
- 5~10% better clock area
- 5% better flow TAT



- Early clock-based place_opt_design:
- clock-aware congestion estimation,
 - clock gating path violation estimation
 - ccopt useful skew technology

Tempus and Quantus Performance



Up to 1.8X Faster and Delivers Linear Scalability For Digital Designs on 8 and 16 CPUs

Up to 1.8X Faster For Transistor Designs on 8 CPUs

Voltus Distributed Processing Solution

Industry-best TAT and capacity



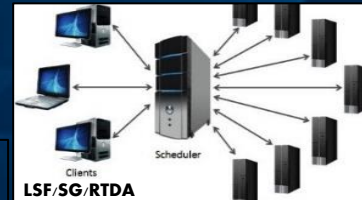
Voltus™ Distributed Processing (DP)

Runs on a Single Machine
Scales up to 32 CPUs

Multi-Threading (MT)

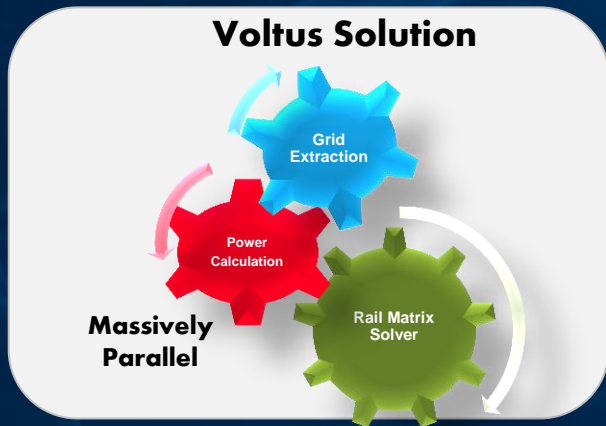


Single, "Super" Computer

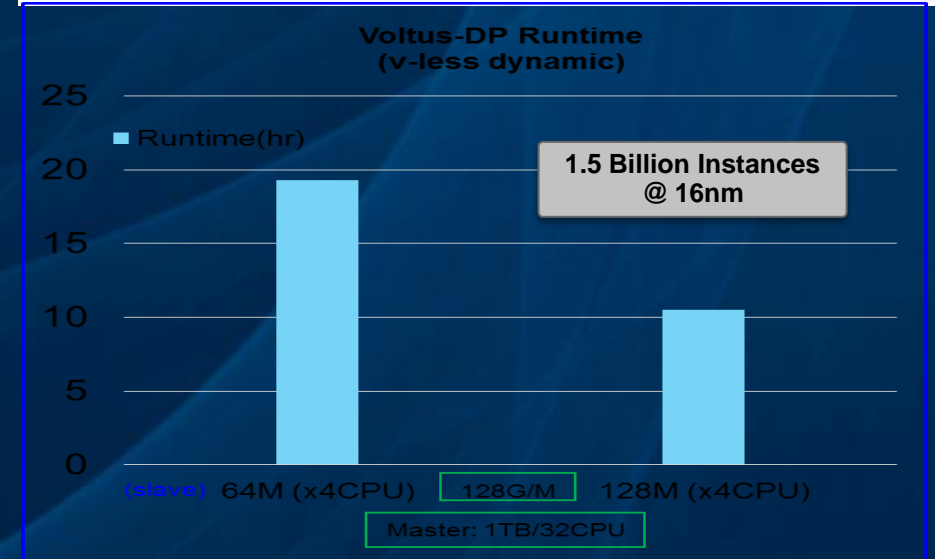
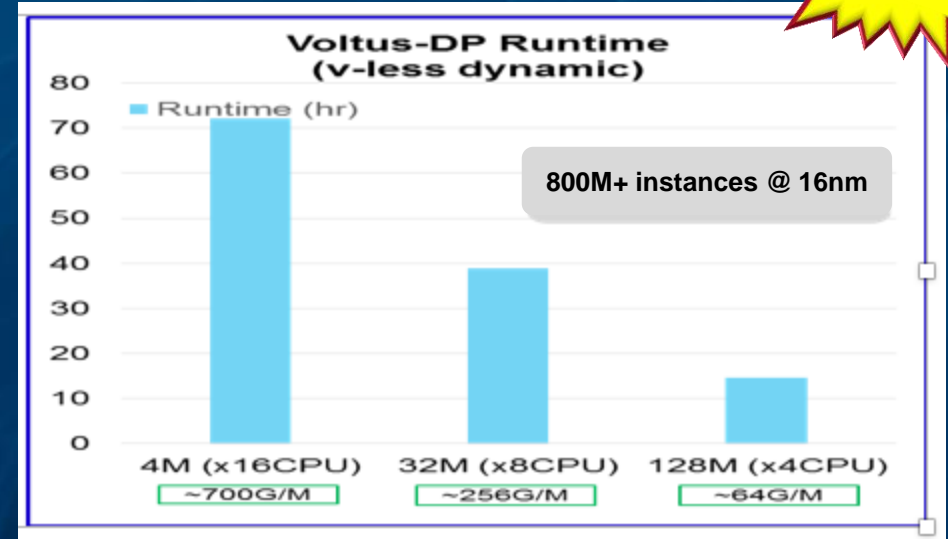


Over Multi Computers

Clients
LSF/SG/RTDA



Distributed Processing (DP)



Significant performance and capacity gains on large designs with no accuracy loss

Tempus ECO: Faster Signoff Closure

Eliminates tapeout schedule risk

Tempus + Innovus Solutions

Unified algorithms, engines, datamodel, user interface

Innovus™ Implementation System

Physically aware ECO

Hold, DRV, setup, leakage

2-3 Iterations

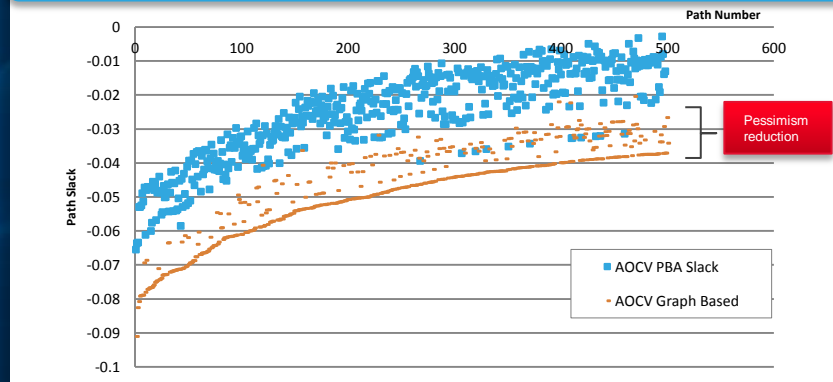


Tempus™ Solution

Distributed MMMC delay calculation and STA

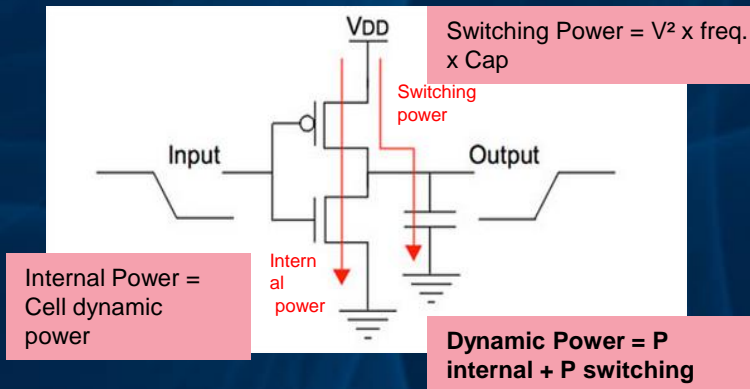
Timing closed

PBA Reduces Pessimism



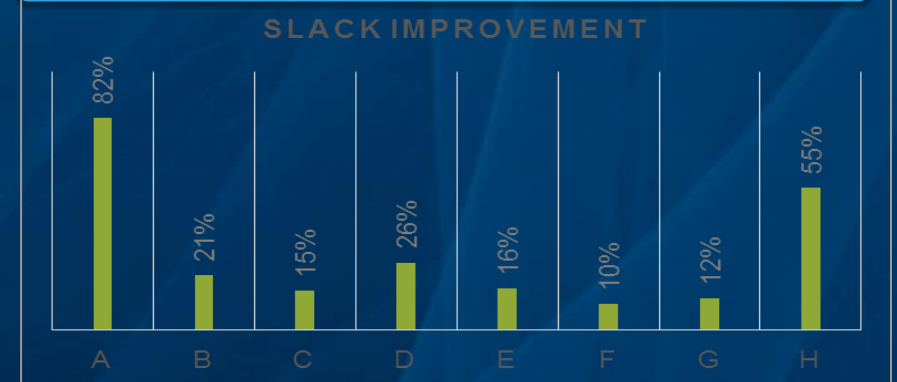
2-3% pessimism reduction

Dynamic Power Optimization



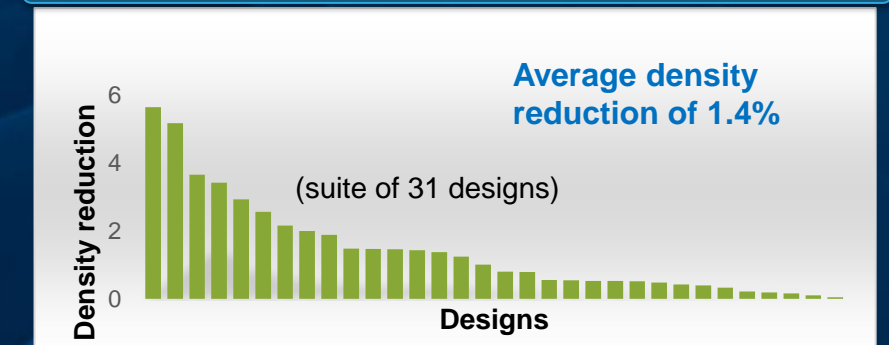
Up to 5% power reduction

Useful Skew Improves Performance



Up to 100ps of slack improvement in some designs

Area Reduction DRV/Setup/Hold Aware

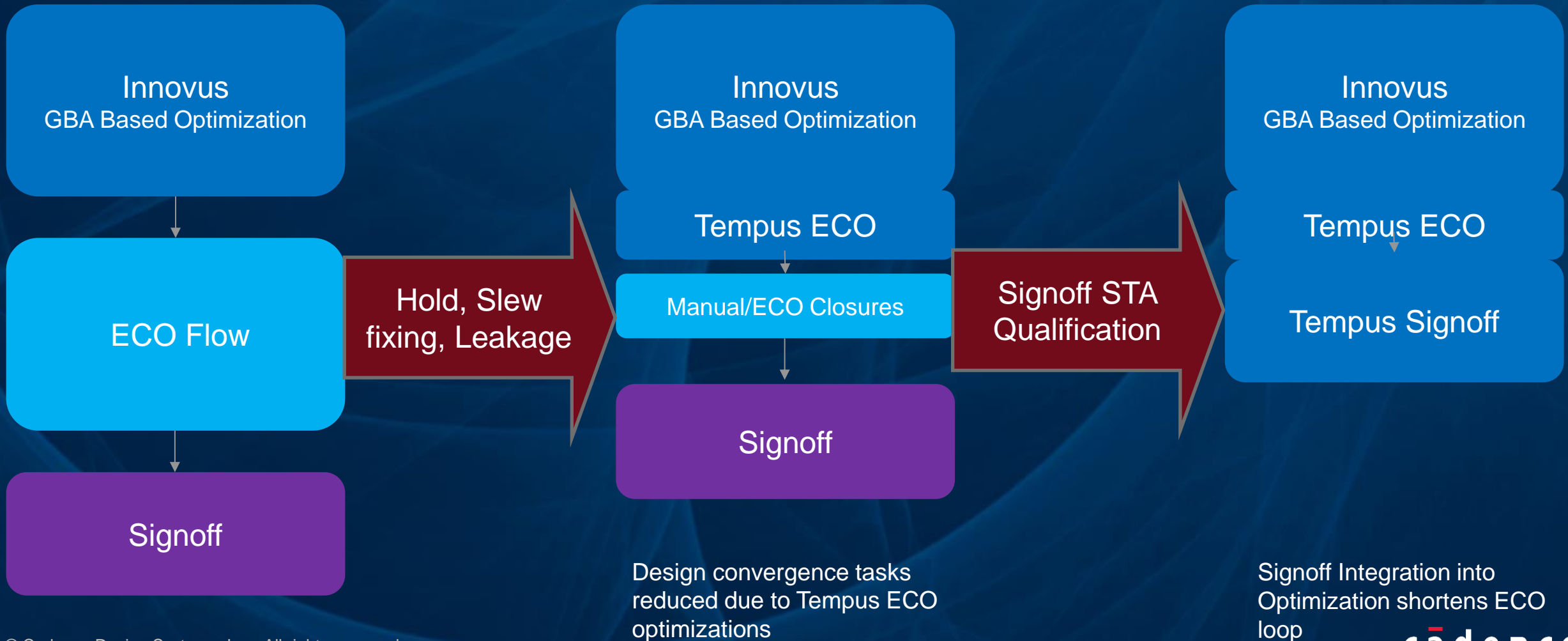


Clock skewing for setup timing closure

Dummy cell insertion for hold fixing

I/O path logic preservation during leakage/area recovery

Signoff Transition Advantages

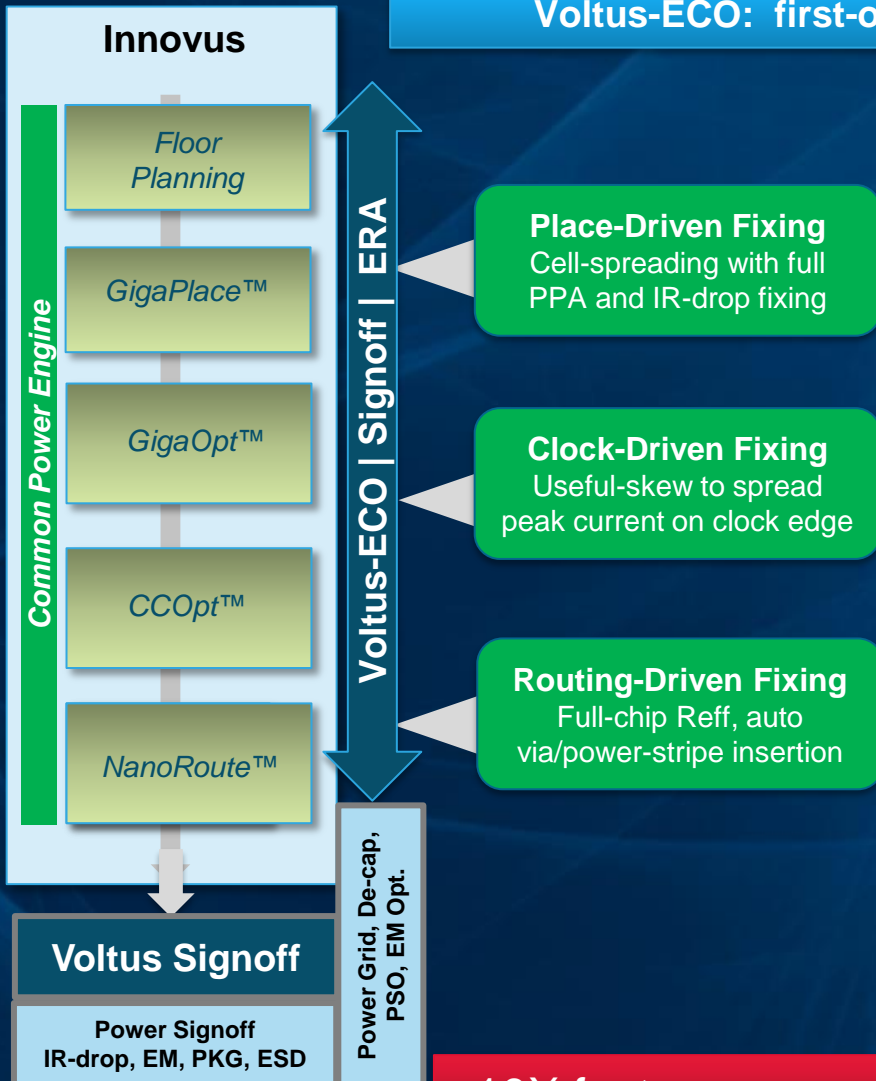


Voltus Innovus Integration: Voltus ECO - faster power closure

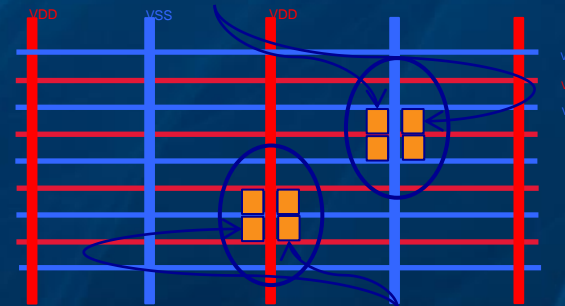
Eliminates tapeout schedule risk

Voltus-ECO: first-of-its-kind IR Fixing

Innovus spreads out the aggressors to reduce IR drop



Aggressors sharing same VDD rail -> high VDD drop



Aggressor sharing same VSS rail -> high VSS jump

Customer Design – 1M instances

**Voltus + Innovus IR Fixing
in 3 iterations**

IR-Drop violation (Victim instances)				IR-Drop voltage (Worst value: mV)			
before	1 st	2 nd	3 rd	before	1 st	2 nd	3 rd
1334	555	125	87	110.542	89.636	82.041	81.943

Victim instance numbers down by 93.5%

Worst IR-drop down by 25.9%

10X faster power closure versus manual, error-prone, iterative traditional flow

Voltus: 14nm Power Grid Optimization

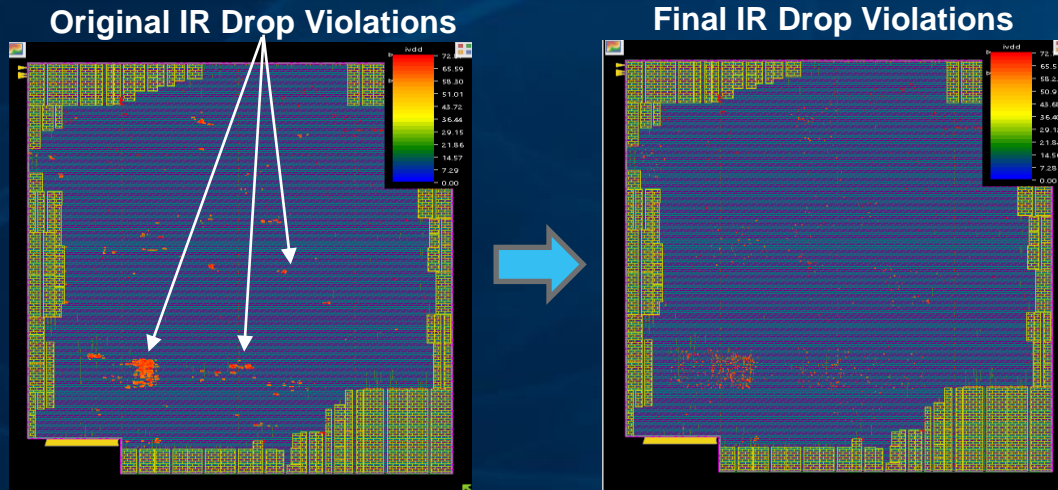
Consumer electronics leader

GPU design core
2.15M instances, frequency: ~1GHz
Ultra low power 0.85V

Design Needs:

- Very dense power grid causes congestion/utilization issues
- Challenging to reduce power-mesh to free-up routing resources
- Need in-design power-strapping flow to fix local IR drop violations

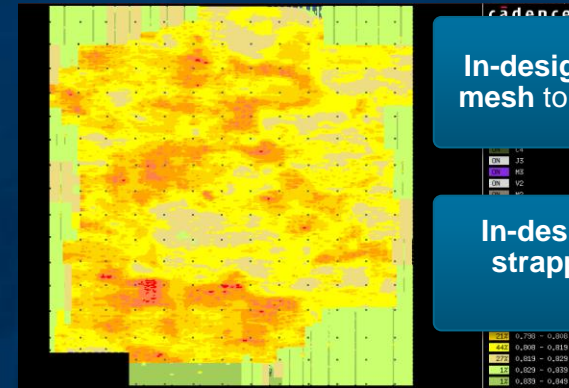
Voltus + Innovus IR-aware placement flow



Iterations	IR Drop Violations	Instances Moved	Run Time
Original	2,241	N.A.	35min
1st	1,646	6,550	20min
2nd	1,428	3,813	20min

Fixed 35%+ violations in 2 iterations

Voltus + Innovus in-design power grid opt.



In-design flow to reduce power mesh to free up routing resources

In-design flow to add power-strapping to fix local IR drop violations

C5 layer = 50% reduction in power-mesh

C6 layer = 25% reduction in power-mesh

Dynamic IR drop still within 85mV target

Voltus + Innovus in-design power grid optimization
 → Frees up routing resources → Die area savings

Tempus 200+ tape-outs

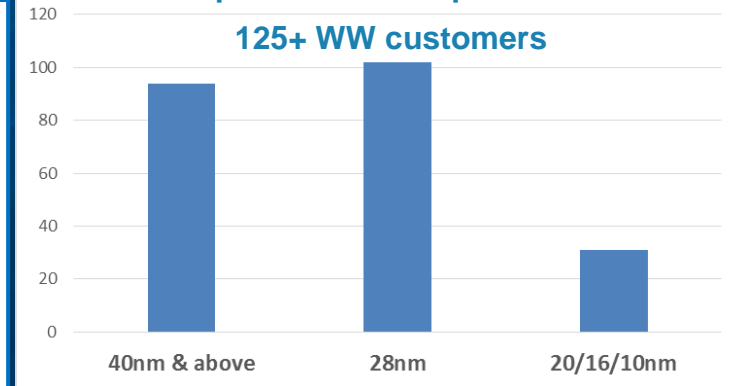
Cadence Tempus Timing Signoff Solution Surpasses 200 Tapeout Milestone Within Two Years of Product Inception

SAN JOSE, Calif., 11 Jan 2016

Highlights:

- Solution becomes the most rapidly adopted signoff tool in Cadence history
- Production deployment across mature process nodes and advanced FinFET nodes
- Customers experience 5-10X faster signoff timing closure and significant PPA gains

200+ tape-outs across process nodes



...could handle more than 50M cells quickly and efficiently...Tempus the right timing platform to address our signoff analysis and closure needs...**expect continued success in taping out complex designs at 28nm and beyond.**

HITACHI



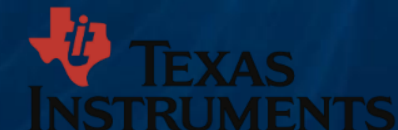
Tempus for all of our SoCs that enable **smart TV, set-top boxes and media connectivity**. Its **runtime performance, coupled with integration within Innovus**...allowed us to significantly reduce the time we spend in timing signoff and, ultimately, **time to market**.

...several successful tapeouts of our **datacenter interconnect solutions**... **distributed multi-mode, multi-corner (MMMC) timing analysis & closure** to get our products out the door and into the fab to meet our customers' aggressive schedules.



Successful design of **industry's first production 16nm FinFET SoC**... a network processor running at **speeds up to 2.6 GHz** ...employing 32 processor cores and a 64-bit architecture

As we move to more advanced process nodes, timing closure becomes more difficult...great to see Cadence taking on this challenge by offering **new technology designed to tackle tough design closure issues**.



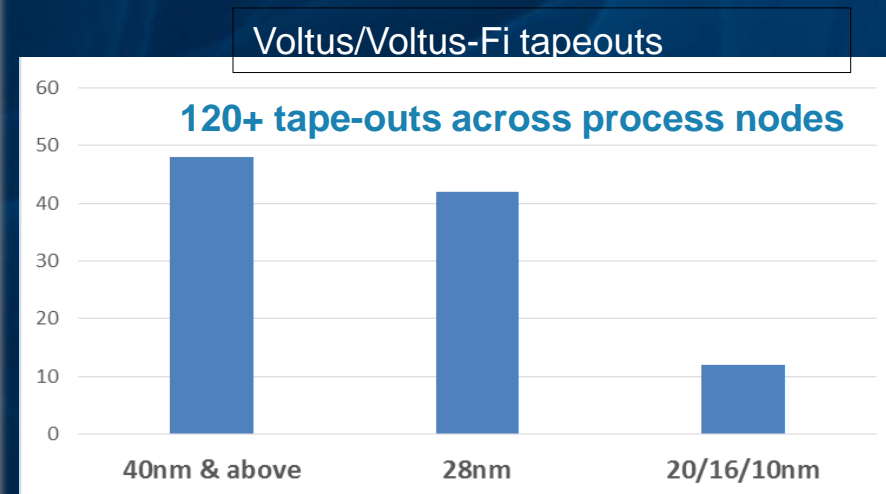
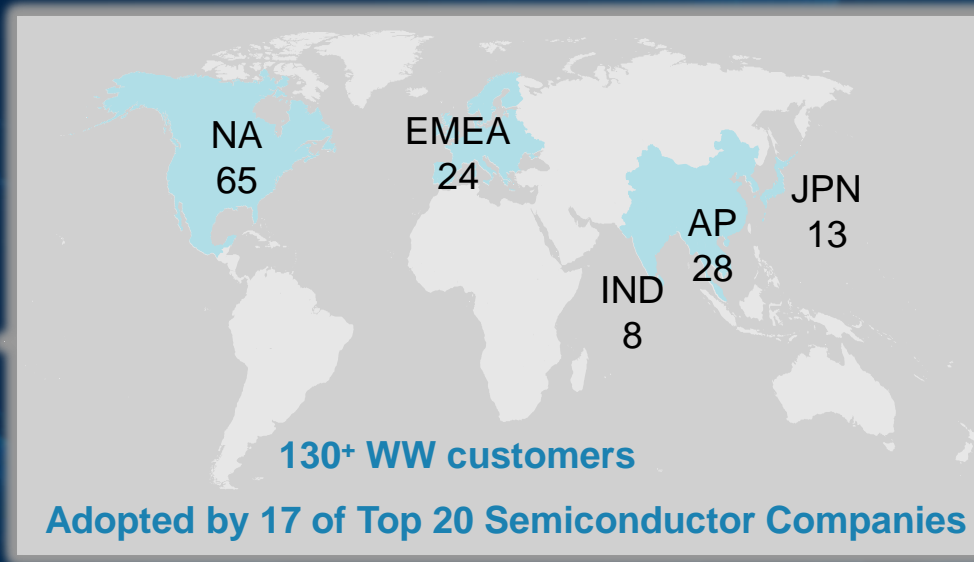
Voltus/Voltus-Fi customer adoption

“Static and dynamic signoff power analysis for a design ~**380M instances** at TSMC28nm... 21hours total runtime on a 1TB machine with **32 CPUs**... About **8X runtime improvement** over previous method with equivalent accuracy”

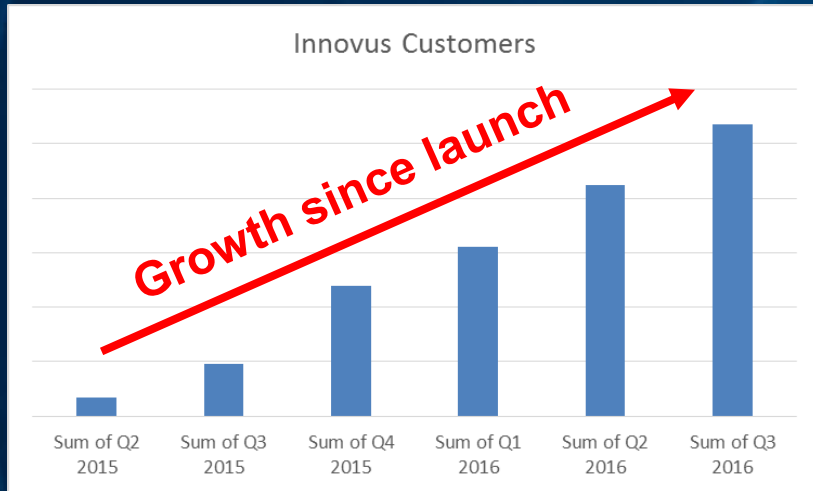


“We are teaming up early on with Cadence to validate the Voltus technology and we are **impressed by its performance gains**. This type of enhanced productivity is **invaluable** to help us meet our time-to-market goals.”

“Phison improves **time-to-market by 40%** with Voltus-Fi ... delivered 3 Si-proven designs in 12 months .. Engineers have been able to find design weaknesses such as potential voltage drop and EM failures, **preventing** costly silicon re-spins.



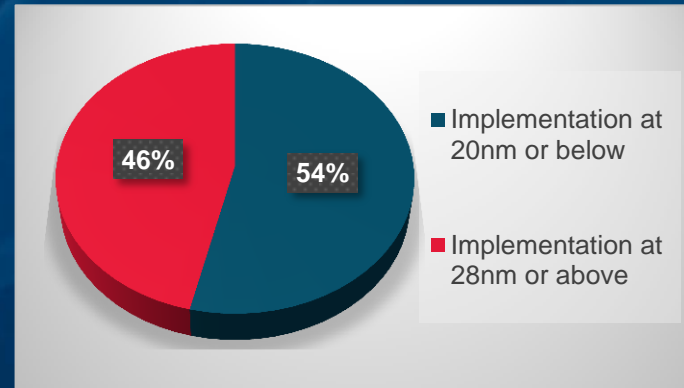
Cadence Digital Flow Market Leadership established @ 7nm



2015

- Innovus/Genus Launched
- Best TAT and PPA
- Integrated Signoff
- Foundry qualified
- Rapid market adoption

16nm ~40-50% designs Innovus



2016

- Gigaplace (Innovus)
- NanoRoute cut metal optimization
- Power Optimization (Genus/Innovus)
- Layer Aware Optimization
- Accurate SOCV and IR modelling (Tempus/Voltus)

10nm ~70-80% designs using Innovus

2017

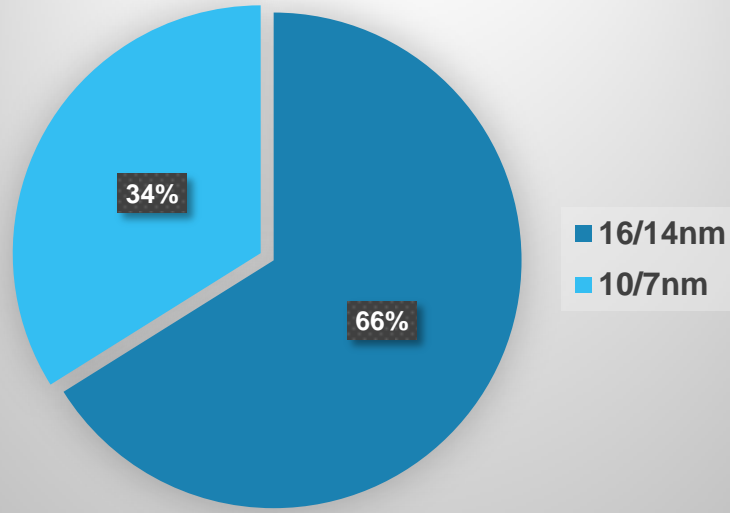
- Architectural exploration
- Power aware Physical Synthesis
- Metal Trim Routing
- Via Pillars
- In-Design Timing and IR Signoff

~ Preferred flow for 7nm Testchips

Market Leadership established at Advanced nodes

Cadence Digital Solution : Unique full flow capabilities @ 7nm

Innovus 10/7nm implementation



Market Leadership @ 10nm.
Extending to 7nm



Massively Parallel architecture
Can handles block capacity up to 10M instances
Suited for next-gen FinFET technology



Power Convergence
(Joules-Genus-Innovus-Voltus)
Full-flow power optimization on next-gen FinFET
Accurate and Convergent power Flow



Shared Engines
Synthesis-Implementation-Signoff
Enabling faster convergence

Advanced
Packaging



Complex
Digital



High Performance
Mixed Signal



Final Thoughts

Human versus artificial brain

Human

10^{14} neurons

100 meters/sec

Natural evolution



Artificial

10^8 neurons

Speed of light

“Designable”

SoC – the center of intelligence

... **still 1,000,000X
to go!**



Cadence Commitment to innovation

> 35% revenue invested in R&D



Towards the intelligent future

ENABLING SYSTEM DESIGN

From Chip to End Product

