



Emulation Driven Power Estimation for Real World Applications

DVCOn US 2025

Arti Dwivedi, Michael Young – Cadence Design Systems

Yash Bhagwat - nVidia

cādence®

Power Predictability Why Do We Care?

Risk and cost from un-managed power is very high

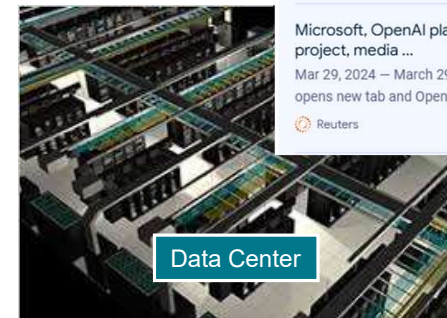
- **Over Spec. Design**

- If average power dissipation too high (over Spec.)
 - Poor battery life
 - Reduced reliability
 - May be “too hot to handle” (ex: Cell Phone)
- If peak power dissipation too high (over Spec.)
 - Voltage drops (due to too-small power rails)
 - Blown bond-wires
 - Hot spots
 - Electro-migration failure (field failures)

- **Under Spec. Design**

- If average & peak power less than estimated (under Spec.)
 - Extra cost due to overbuilding of ASIC, Package, Battery

- **Operating Cost**



Microsoft Pays \$800M More in Data Centers Energy Costs

Microsoft CFO Amy Wood shared that the firm will pay more than \$800 million in extra energy costs to operate its data...

Data Center Knowledge

Microsoft, OpenAI plan \$100 billion data-center project, media ...

Mar 29, 2024 — March 29 (Reuters) - Microsoft (MSFT. O) , opens new tab and OpenAI are working on plans for a data...

Reuters



Low Power: Everyone's Concern

Different drivers in different verticals

Mobile/Handheld

Battery life (energy)
Unit cost (chip package)
Reliability / Safety



Consumer/Digital Home

Unit cost (package, fans)
Energy certification
Reliability / Safety



Hyperscale/Data Center

Power efficiency / Green
Total cost of ownership
Reliability / Safety



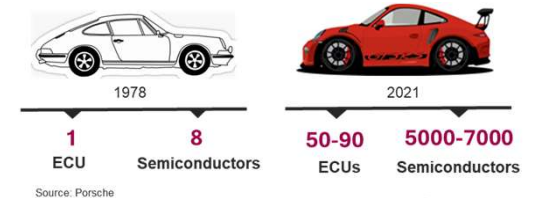
Automotive/EV

Power efficiency/Miles
Total cost of ownership
Reliability / Safety



Low-power requirements drive different design decisions

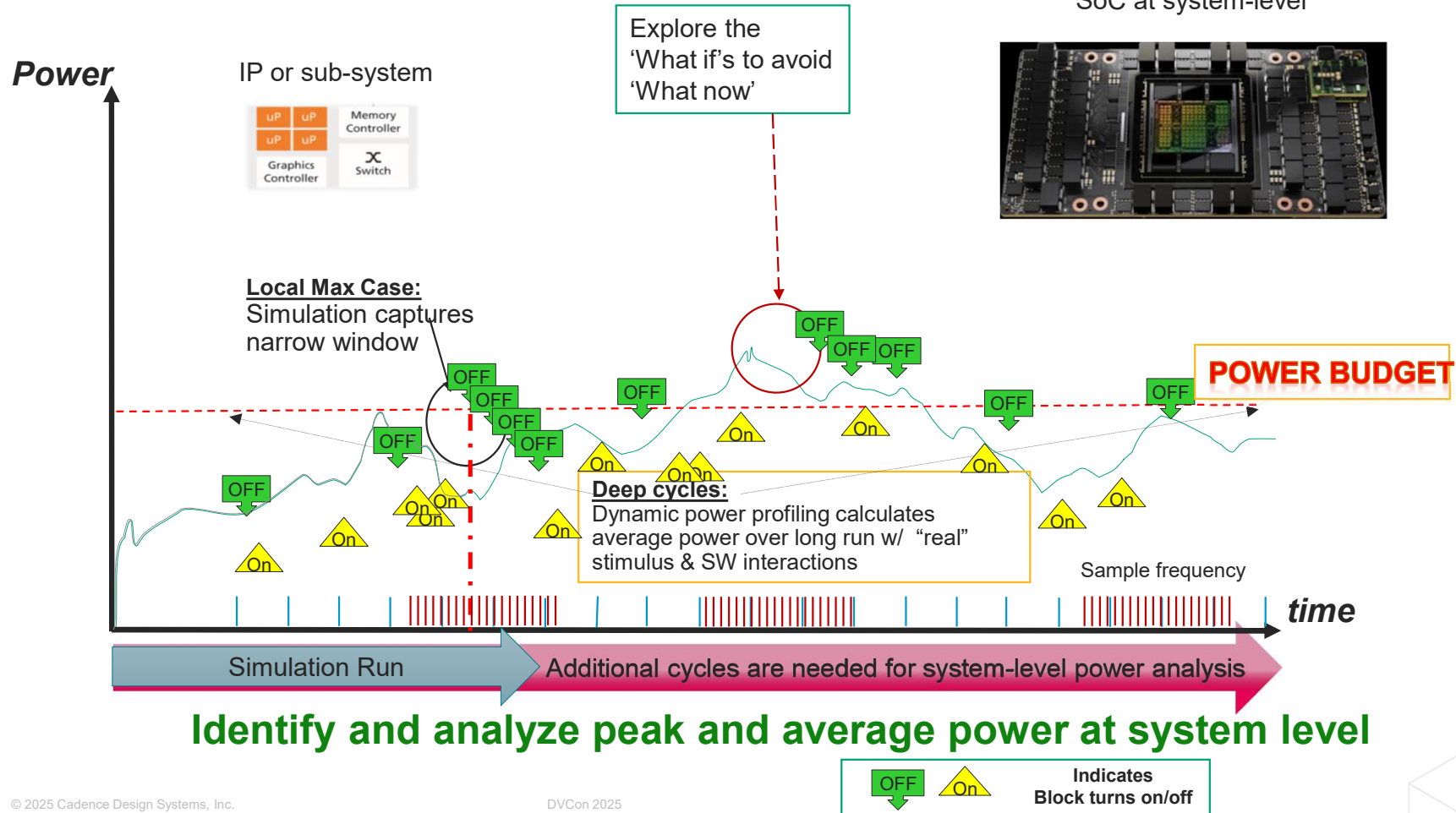
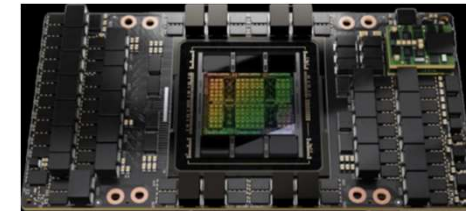
- Design architecture and low-power design techniques
- IP make vs buy and power vs performance
- Unit cost and manufacturing process



SoC Power Analysis Requires “Deep” Cycles

@100MHz for 20 secs → 2 Billion cycles

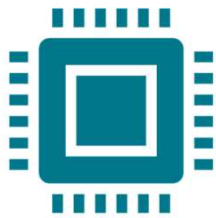
SoC at system-level



SOC Power Analysis for Real Use Cases

Arti Dwivedi

SOC Power for Real Use Cases: Requirements



Billion+ gates SOC capacity

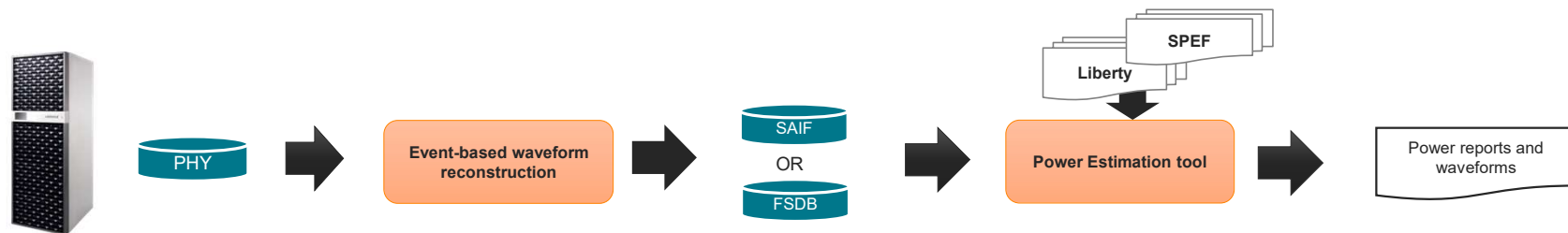


Fast turnaround for power estimation of real use cases spanning million to billion cycles

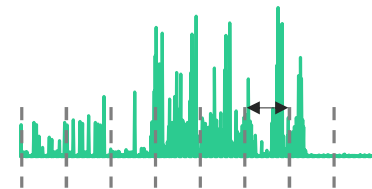


Actionable insights into power hot-spots for debug

Traditional Power Methodology Challenges

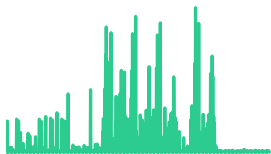


- **SOC power for real use cases not supported**
 - Limited design size capacity: million+ gates
 - Limited vector cycle capacity: few million cycles
- **Days/weeks of turnaround to get power number of an emulation use case**
 - Power estimated by rolling power of smaller time windows
- **Generating emulation activity for long tests limited by size and throughput**



Palladium Dynamic Power Methodology at RTL

Billion+ Gates
Million+ cycles



- Profile RTL flop activity in **Software** after PHY dumping
- Cycle capacity limited by disk requirement of large number of cycles

Billion+ Gates
Billion+ cycles



- Profile RTL flop activity *in Hardware at Emulation-speed*
- **Billion+ cycles in hours**
- **100X-1000X speed up** with capacity overhead

Early RTL Activity Profiling for Real World Scenarios

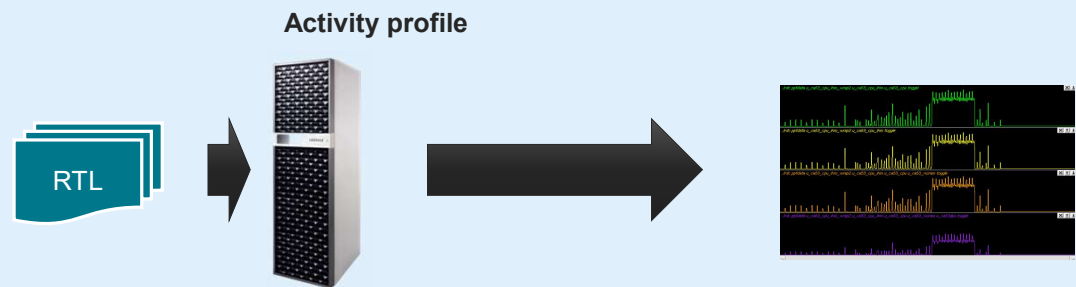
Analyze flip-flop activity profiles early at RTL

Billion+ gates and billion+ cycles in hours

- *Identify redundant activity and hot-spots per hierarchy*
- *Identify interesting windows for power analysis*

Hardware Native Toggle Count (HW-NTC)

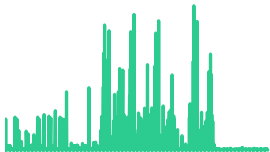
- At-speed hardware activity profile generation
- 100X-1000X improvement in turnaround time over offline approach, 1.5X-2X capacity overhead
- Ability to generate waveforms for hierarchies of interest



Software Native Toggle Count (SW-NTC) also available.
Enables Offline activity analysis using Palladium PHY DB.

Palladium Dynamic Power Methodology at Gate

Billion+ Gates Million+ cycles



- Estimate gate power in **Software** after PHY dumping
- **Accuracy within 5%** of sign-off power tools with no capacity overhead
- Turnaround time improvement from **weeks in signoff tools to hours**

Billion+ Gates Billion+ cycles

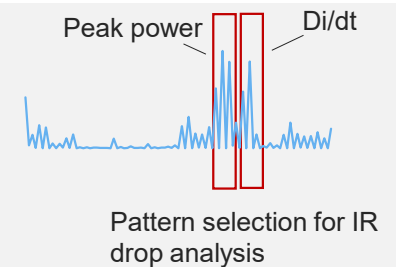


- Estimate gate power in **Hardware** at Emulation-speed
- **Billion+ cycles** of power estimation in **hours**
- **Accuracy within 10%** of sign-off power
- **100X-1000X speed up** with capacity overhead

Gate Power Analysis with SOC Capacity for Real Work Loads

- **Analyze accurate gate level power early for real use cases**

- *Analyze power efficiency of software-hardware interactions*
- *Debug power hot-spots for real workloads*
- *Identify critical windows for power sign off and IR drop analysis*



- **Accuracy w.r.t sign-off power**

- 10% with hardware power estimation
- 5% with software power estimation

- **Accurate dynamic power per cycle for long vectors**

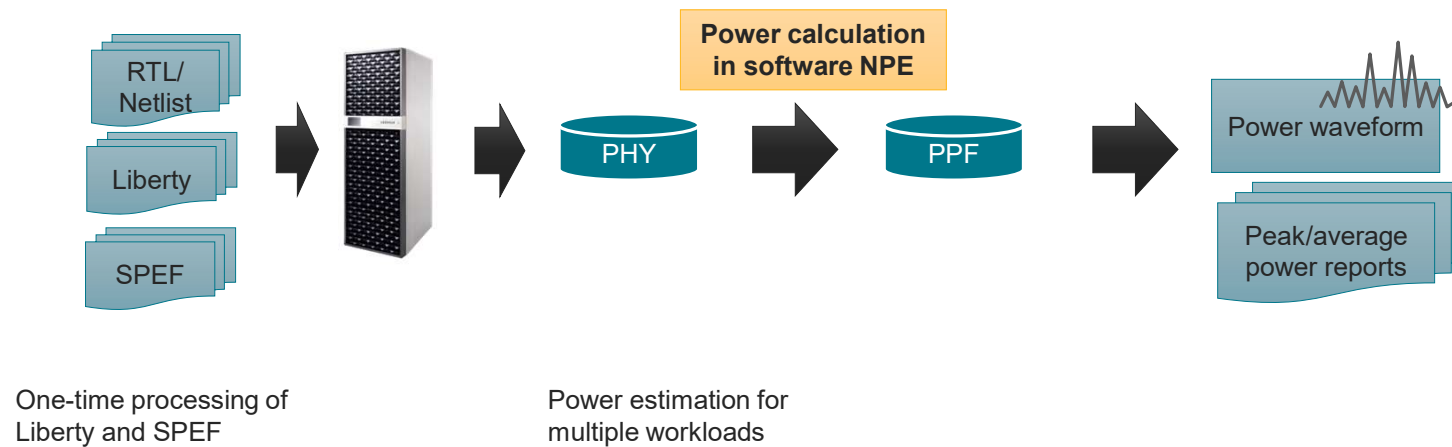
- No loss of accuracy when generating peak power profile for long vectors

- **Peak and average power reports by category, hierarchy for hot-spot analysis**

Gate Level Power Estimation in *Software*

Software Native Power Estimation (SW-NPE)

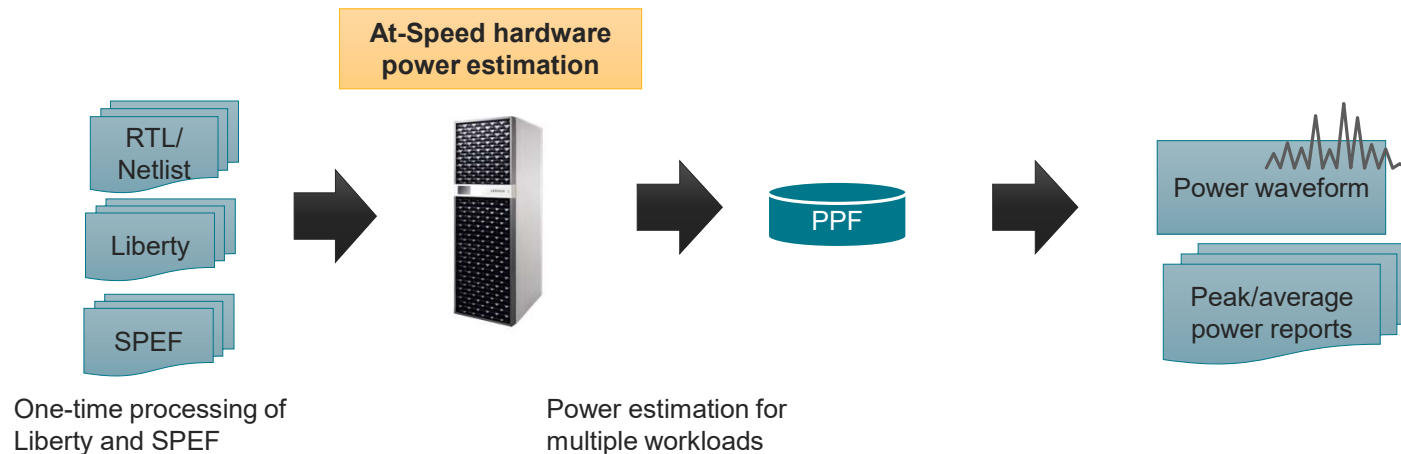
- Billion+ gates and million+ cycles in hours within 5% of sign-off
- Offline power estimation using PHY database
- Improves turnaround time of power estimation from weeks in sign-off tools to hours



Gate Level Power Estimation in *Hardware*

Hardware Native Power Estimation (HW-NPE)

- Billion+ gates and billion+ cycles in hours
 - At-speed power calculation in hardware, 100X-1000X improvement in turnaround time with capacity overhead
 - Not possible to estimate billion+ cycles power in traditional power tools
- Accuracy within 10% of sign-off power tools



Palladium DPA Gate Power Reporting

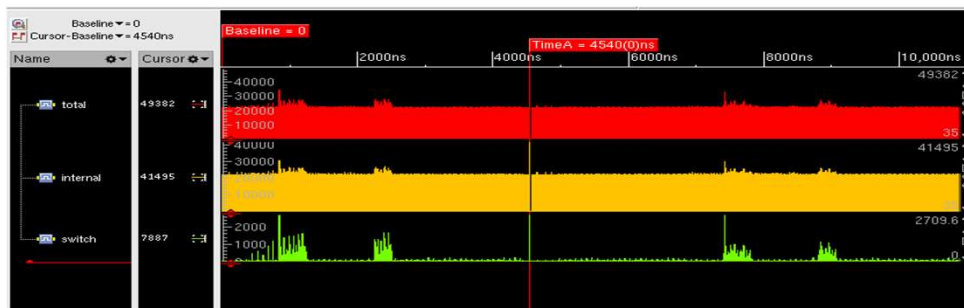
Power by category, instance & Power waveforms

- Palladium enables identification of blocks, windows for power sign-off
- Power per instance and by category to identify power hot-spots in SOC

Instance	Internal(mW)	Switching(mW)	Total(mW)	PinCap(pf)	WireCap(pf)
tb_tgen_axi	2.2904e+01(97.2%)	6.7061e-01(2.8 %)	2.3575e+01	3.5017e+01	5.8735e+01
Seq	5.0934e+00(21.6%)	6.9426e-03(0.0 %)	5.1003e+00		
Com	7.2339e-03(0.0 %)	6.6350e-01(2.8 %)	6.7074e-01		
Bbox	1.7804e+01(75.5%)	1.6171e-04(0.0 %)	1.7804e+01		

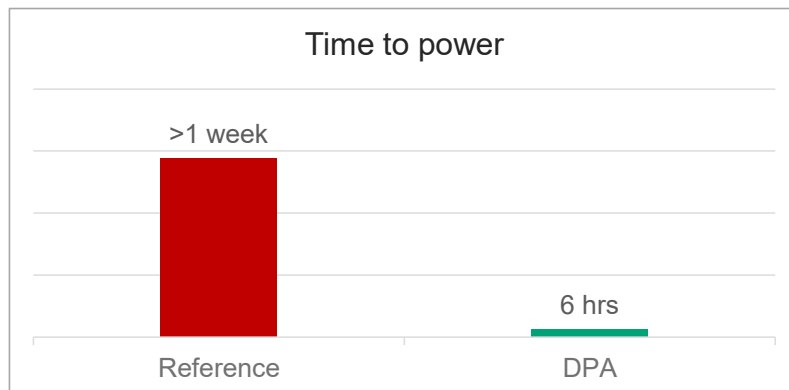
Instance	Internal(mW)	Switching(mW)	Total(mW)	PinCap(pf)	WireCap(pf)
.	1.7645e+01(85.4%)	3.0143e+00(14.6%)	2.0660e+01	2.8580e+00	2.5260e-01
dut	1.7645e+01(86.8%)	2.6896e+00(13.2%)	2.0335e+01	2.8580e+00	1.5980e-01
dut.cnt_u0	4.1759e-01(50.6%)	4.0849e-01(49.4%)	8.2608e-01	1.4400e-01	2.9300e-02
dut.cnt_u1	3.8836e-01(58.9%)	2.7073e-01(41.1%)	6.5909e-01	1.4400e-01	3.3400e-02
dut.cnt_u2	3.8836e-01(59.0%)	2.6998e-01(41.0%)	6.5834e-01	1.4400e-01	2.9800e-02
#1	7.9526e-03(11.1%)	6.3697e-02(88.9%)	7.1649e-02	4.1000e-02	5.4000e-03
dut.cnt_u3	3.8836e-01(59.0%)	2.6987e-01(41.0%)	6.5823e-01	1.4400e-01	2.9800e-02
#2	7.9526e-03(11.1%)	6.3697e-02(88.9%)	7.1649e-02	4.1000e-02	5.4000e-03
dut.mem_u	1.6054e+01(100.0%)	2.4420e-03(0.0 %)	1.6056e+01	2.2820e+00	3.7000e-02
#3	1.6054e+01(100.0%)	2.4420e-03(0.0 %)	1.6056e+01	2.2820e+00	3.7000e-02

- Power waveforms for total, internal and switching power to identify critical windows



DPA Success: Software NPE

- Accuracy
 - Power estimation accuracy within 5% of sign-off
 - Accuracy of critical window identification improved
- Turnaround time
 - Power of 5ms GPU frame in 6hrs



GPU Frame-level Power Analysis with DPA

Source: CadenceCONNECT



For a 300MG design, 5ms frames, comparing Average Power estimation.

Long pattern average power

- Report 72 instance
- Total time: 5.5hr to generate ppfddata and 30min to "tca -getNative"
- Accuracy Analysis (Correlation with power sign-off tool + SAIF)
 - Average Accuracy=98.54%

Pickup windows of interest

- Report 72 instance
- Total time: 5.5hr to generate ppfddata and 20min to "tca -getTopNPeak N -filter {IntervalAverage 100}"
 - Generating power curve by power sign-off tool to find Wol is almost infeasible for execution (>1w)
- Accuracy of Wol (Correlation with power sign-off tool + FSDB)
 - Average Accuracy=95.21%

cadence CONNECT
TAIWAN

- In this work, we adopt DPA on GPU frame-level power profiling to achieve finishing millisecond frame power calculation in 6 hours.
- In average, DPA has over 95% competitive accuracy compared to power sign-off tool.
- In performance part, DPA beats traditional method that generating power curve by power sign-off tool to find Wol. (About >1week → hours)



Nvidia Power Analysis Case Study



Early Power Estimation on Palladium

Yash Bhagwat, Hardware Emulation | DVCon 02/24/2025

Agenda

- Challenges in power simulation

- Can we go to emulation early?

- One emulation tool for all power

- Sampling and accuracy

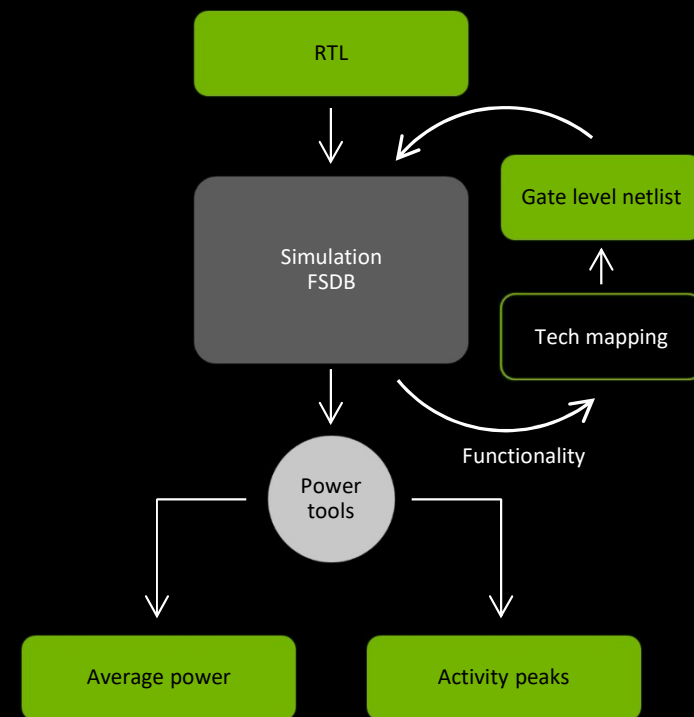
- Big software workload problem

Power simulation

- RTL simulation with power estimation infra
 - Window of interest
 - Track perf and power
- Incorporate gate-level netlists from synthesis
 - Actual physical implementation
 - Improved accuracy and reliability
- Long process
 - Simulate with waveform dumping
 - Run through multiple tools for
 - Name mapping & event propagation
 - Average power estimation
- Late power feedback has a big risk
 - Break timing
 - Slip schedule

Concerns with short tests

- Hundred Thousand cycles
- No application-level power
- Time to power ~10 hours



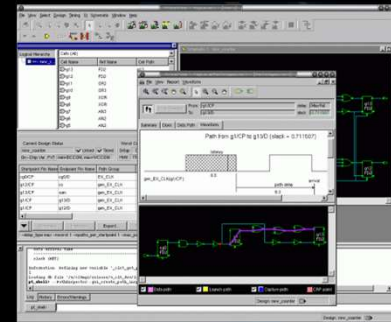
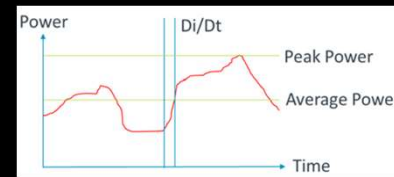
Emulation for Power tools

Palladium

- Bring up a unit or sub-system for power at an early stage
 - Min config to run app checkpoints
 - Multi million gate design
 - Compile emulation model in an hour
- Run a variety of long tests with the same power infra
 - Functionality check (monitors?)
 - Track performance in the window of interest
 - Capture FSDB waveforms at speed (5-10min for 5M cycles)
- Post process using multiple tools (licenses)
 - Name mapping
 - Power tool for average power estimation (not cycle based)

Concerns with long tests

- Massive compute for FSDB processing
- Power tools struggle with volume
- Time to power ~2 days (even without peaks)



“

When you get little, you want more.

— Anonymous

”

One tool for power

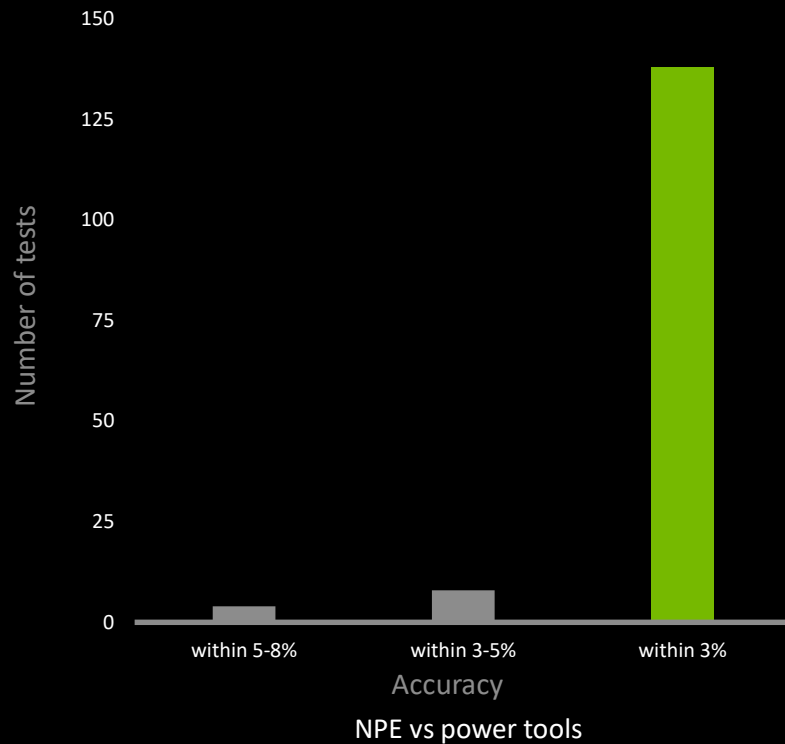
Native Power Estimation

- Power estimation handled by emulation compiler
 - Reads in energy/capacitance info (SPEF and liberty)
 - Supports encrypted IP throughout the flow
 - One time processing of power data
- Heterogenous RTL and gate design
 - Compute block(s) power using synthesized gates
 - Flat netlist with primary I/O preserved
 - Other RTL components and testbench can stay same (speed)
- Emulation tool can now estimate power!
 - Run with raw wave capture (lite and fast unlike FSDB)
 - Reconstruct every cycle to compute power
 - Time to power comes down to ~20min for long tests (vs ~2days earlier)
- Pre-silicon power study
 - Power profile for every netlist milestone
 - Study app power and measure power fixes



Benchmarks

Average power



- Power estimation for full suite of long tests
- Time to power in minutes (LSF)
- Good accuracy levels across checkpoints
- Configurable sampling based on user clocks
 - Trade-off between speed and accuracy
- Difference comes from
 - Netlist cells (boundary optimization)
 - Sampling frequency
 - Clock tree power

Concerns with extra-long tests (real software)

- Billion cycles
- Impractical to dump waves at this scale
- Nothing to post process power ??

Real application power

Hardware NPE

- **Specialized hardware to compute power**
Toggle infra managed by emu compiler
 - Capacity impact depends on block size
 - Single database supports SW-NPE as well as HW-NPE
 - Check for memory inference and critical paths
- **Compiler models power characteristics**
SPEF files for available cells and wire load model
 - Cell report available at compile time
- **Run with power streaming**
Select block hierarchy for capturing power
- **Turnaround time is critical**
Run software, zoom into billions of cycles to measure
 - Measure average power
 - Extract smaller power windows of interest, full RTL power analysis



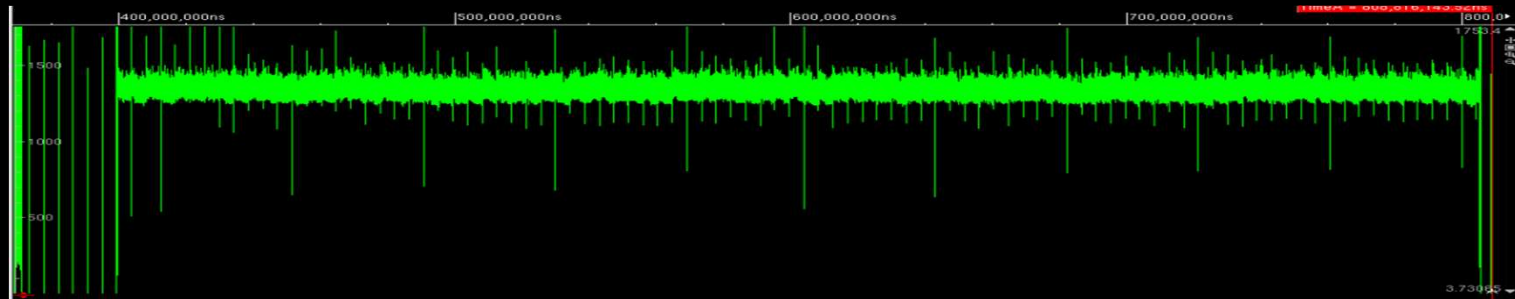
Application power numbers using Hardware NPE

Estimate power of apps (billion+ cycles) in ~1 hour !!

	Cycles	App run	App run with power streaming	Time to power (end-to-end)
App1	1.26B	23min	40min	42min
App2 (multi-threaded)	3.83B	38min	65min	85min

Impractical to dump waves at this scale using other methods

Get visibility for every cycle amidst billions of cycles of power



“

Open collaboration and partnership
are the keys to driving progress and
innovation.

— Jensen Huang, NVIDIA

”



2025
DVCON
UNITED STATES
SAN JOSE, CA, USA

—
Q/A

cādence®



cādence®

© 2025 Cadence Design Systems, Inc. All rights reserved worldwide. Cadence, the Cadence logo, and the other Cadence marks found at <https://www.cadence.com/go/trademarks> are trademarks or registered trademarks of Cadence Design Systems, Inc. Accellera and SystemC are trademarks of Accellera Systems Initiative Inc. All Arm products are registered trademarks or trademarks of Arm Limited (or its subsidiaries) in the US and/or elsewhere. All MIPI specifications are registered trademarks or service marks owned by MIPI Alliance. All PCI-SIG specifications are registered trademarks or trademarks of PCI-SIG. All other trademarks are the property of their respective owners.