

# Optimizing Power and Performance Yield in Mobile Application Processors

Sreeram Chandrasekar, Sourav Banerjee, Ajoy Mandal, Mohit Sharma

*Texas Instruments Inc, Bangalore, India*

Aniket Singh

*BITSPilani, Goa, India*

# Introduction

- Apps Processors for Mobile/Hand-held devices present a complex challenge
  - Ever increasing performance needs (Multi-tasking, Imaging, Video, Graphics)
  - Stringent power goals (Battery life)
- No longer Performance “Or” Power
  - Need to consider multiple dimensions:
    - DVFS/Adaptive Voltage Scaling (AVS): Adjust voltage for performance/power
    - Transistor  $V_T$ : Performance vs Leakage
    - Hardware vs software, and their relative efficiencies
    - Power-off during idle, at software and hardware levels
    - Dynamic power reduction: Clock gating, Frequency/ voltage selection based on use case
- Need to look at Performance and Power yield
  - Maximize performance yield across process conditions
  - For a given use case, minimize power over the entire  $V_T$  spread
  - Consider impact of process variability on transistor and interconnect (SSTA)
  - Ensure robustness at Ultra low power/Low voltage use cases (SSTA)

# Finding the Operating Point

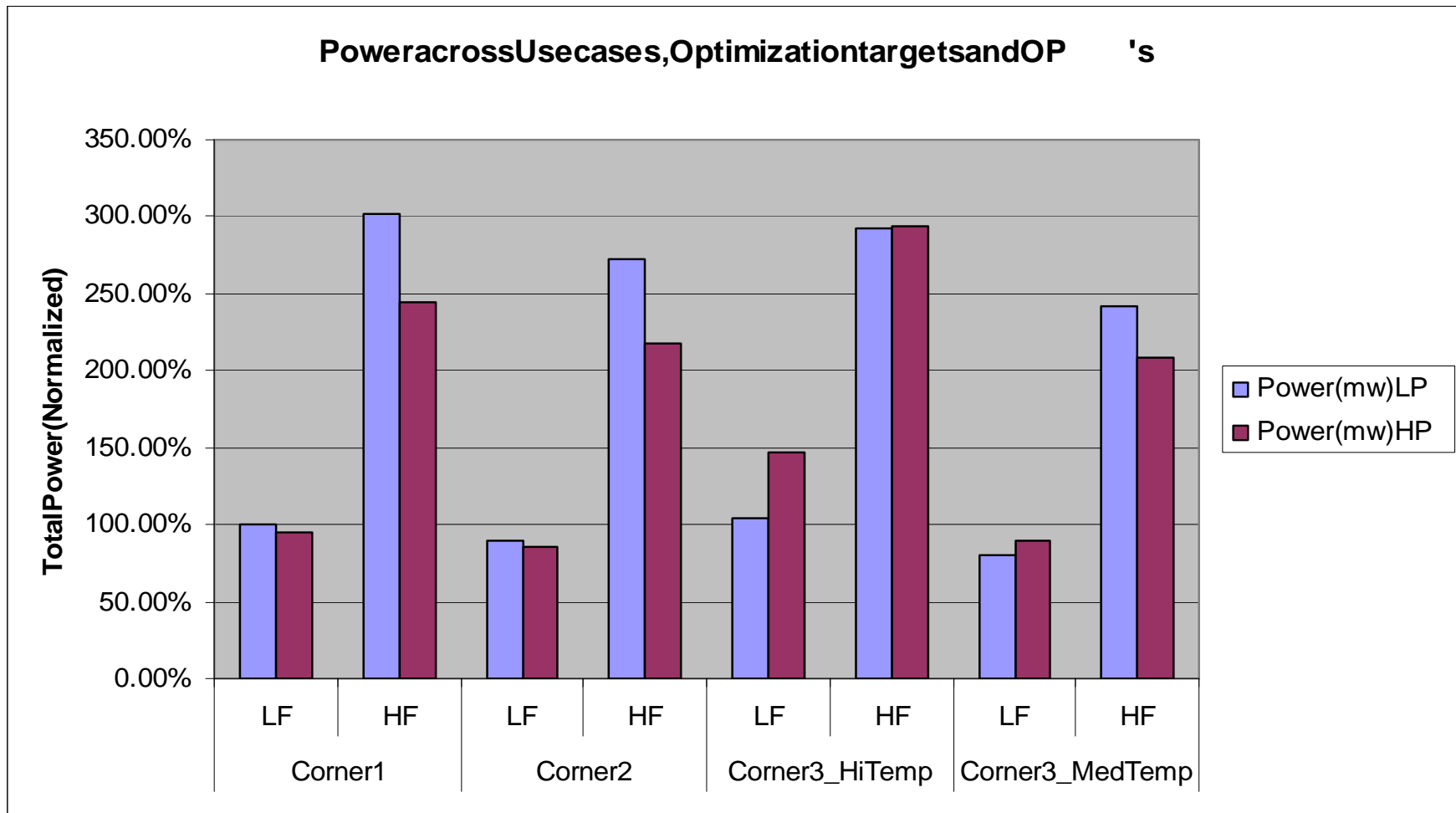
- Our design: HDVideoProcessor
- Use cases
  - LowFrameRateFullHDVideo(“LF”)
  - 2XHigherFrame(HF)rate:Needs2X performance
- Operating points (OPPs)
  - LF or HF can be achieved at different OPPs
    - By controlling Voltage, Frequency, System Idle time,  $V_T$  of transistors
- Key Power Considerations
  - Complete fast and shutdown (save leakage)
  - Low voltage and run slower (save dynamic power)
- Final Goal
  - Maximize power yield for a target use case
  - Find such performance target/operating voltage to optimize design

Optimization Target	Use Case	Performance	Voltage
“LP” (LowPerf) Achieve LF at Nom voltage, and HF at Voltage Overdrive	LF	1X	100%
	HF	2X	Overdrive ( $>>100\%$ )
“HP” (HighPerf): Achieve HF at Nom voltage and LF at reduced voltage	LF	1X	Reduced ( $<<100\%$ )
	HF	2X	100%

# Total Power Vs System Performance

- Impact of total power at various operating points were considered
- Depending on the system performance lack at the frequency/ $V_T$ -class combination, benefits of various power reduction techniques were quantified
- Leakage power reduction
  - Power-off during idle, at software and hardware level
  - Voltage Scaling to adapt to different process
    - Hot silicon runs low voltage, for given performance point
  - Mixed  $V_T$  cell library
    - Lower  $V_T$  cells on timing critical paths
    - Higher  $V_T$  cells on non-critical paths
- Dynamic power reduction
  - Clock gating
  - Process based voltage adjustment
  - Frequency/voltage reduction based on use case

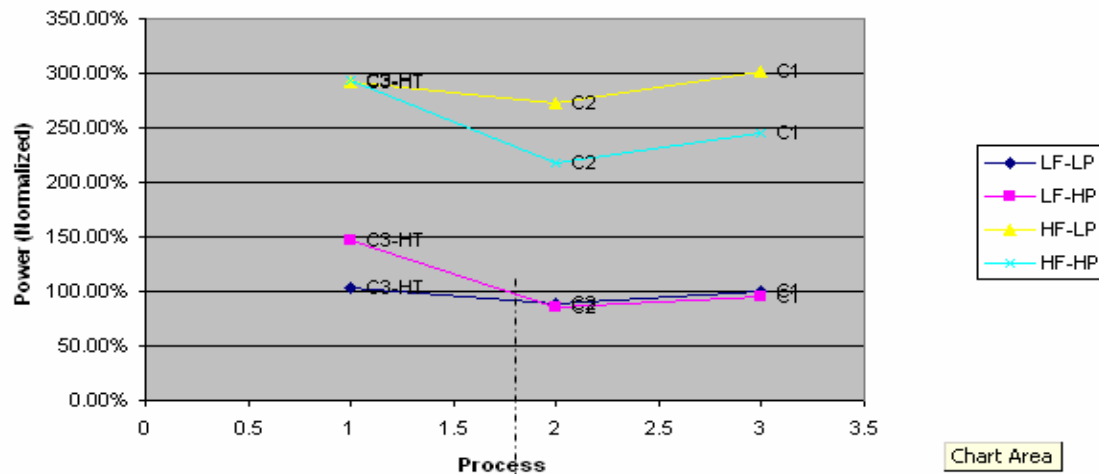
# Results:TotalPower



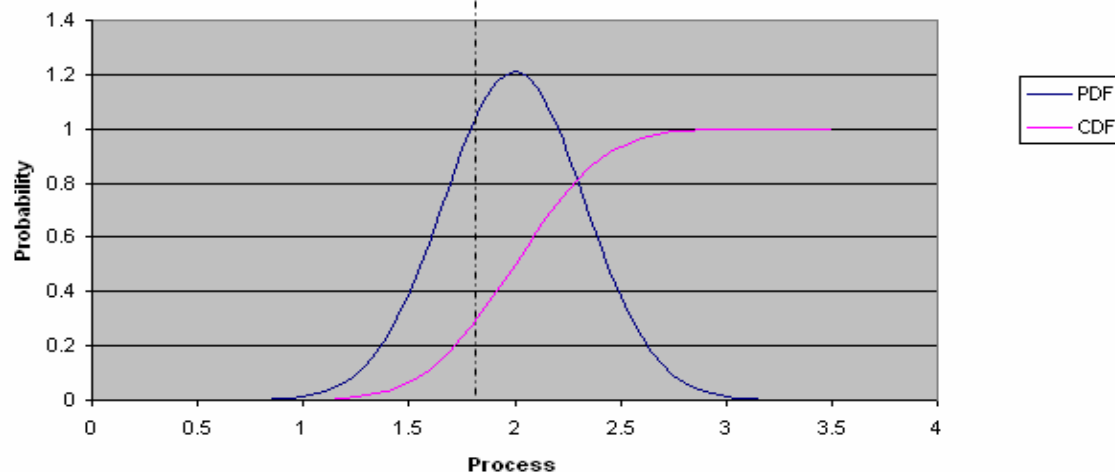
- Corners cover Slowest/least leakage to fastest/ most leaky corners
- Slow corners also get highest voltage, for a fixed frequency point
- Take-Away: Pick the Optimization target based on dominant use case

# PowerYieldAnalysis

Power across Use cases, Optimization targets and OP's



Power Yield (For Process centered around corner C2)



- Helps choose the design target that gives better yield for the dominant use case
- Total power across PVT used to predict power yield
  - Impact of AVS, DVFS, Clock gating, power gating etc. comprehended

## OPP Selection:

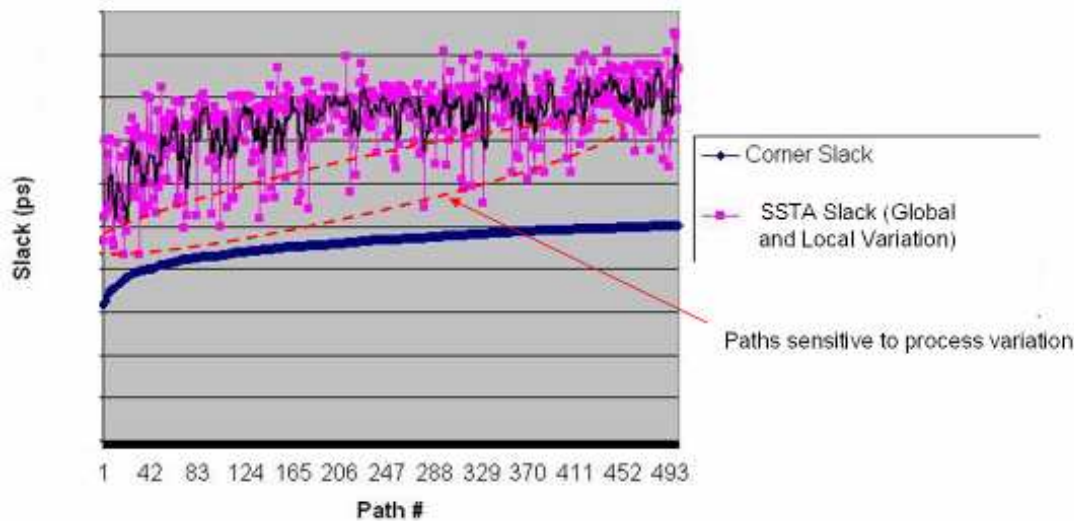
- High Performance Video Systems:
  - HF is the dominant use case
  - Clearly, HP is better, almost 100% of the parts
- Low-Power Video Use cases
  - LF is the dominant use case
  - LP design implementation better by 30%

# Robustness/Performance Scaling

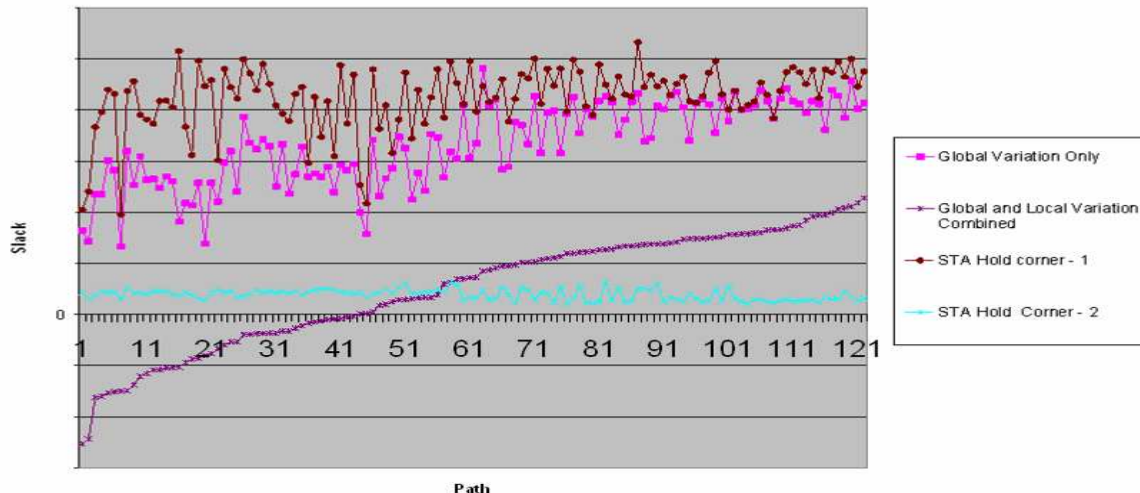
- Widerange of operating points in Mobile Apps Processor
  - High performance and Low Power scenarios
  - Need to ensure
    - robust operation (low voltage holds safety, for example)
    - performance scaling (overdrive voltages, speed binning, yield etc.)
- In 45nm, global and local variation effects are significant
  - Transistor Global and Local Variation
  - Interconnect Mis-track
  - Effect more pronounced in 32/28nm
  - Modeling all these effects in traditional OCV based derates is inaccurate
- SSTA for Variation Aware Timing Analysis
  - In 45nm, We adopted a combination of SSTA and Marginal timing methodology
  - Violations validated using Statistical SPICE simulations
  - Detected and fixed low voltage hold issues
  - Built more performance scaling margin by addressing variability sensitive paths

# SSTA: Setup Outlier Analysis

Corner STA Vs Statistical STA Slack Profile



Variation Impact on Hold Slack



- SSTA Setup Analysis
  - Slack outliers identified
    - Paths sensitive to variation
  - Outliers fixed, to ensure all paths scale uniformly
    - Mostly seen on low drive cells / poor transition times
    - Higher local variation on small transistors
- SSTA Hold Analysis
  - Hold found to be most impacted by variation at low voltage conditions
    - Increased impact of local variation / VT variation
    - Not just local variation, but combination of global and local variation
  - Several hold failing paths were detected using SSTA
    - Passed all traditional corner STA analyses

# Summary

- Today's mobile application processors drive high performance and low power requirements simultaneously
- Combined analysis of performance and power across process conditions, and use cases required
- Power Yield prediction was done to arrive at optimal OPP selection
  - Impact of DVFS, AVS, VT-classes, P, V, T comprehended
  - Design optimization target selected to suit the end application
- In 45nm, SSTA was used for timing closure
  - Setup: Mainly to improve performance robustness
  - Hold: to identify and fix low voltage hold robustness
  - Hold path passing all corners STA found to fail due to combination of global+local variation
  - Violations validated using Stat-SPICE simulation
- For early design closure and new technology nodes, Corner STA margin tuned based on SSTA