Announcements

- Homework 2
  - Due today
- Please let me know when you have updated your proposal abstracts
  - Midterm reports next Thursday
# Class Material

- **Last lecture**
  - Lowering power
- **Today's lecture**
  - Dynamic voltage scaling
  - Clock gating

---

## Power /Energy Optimization Space

<table>
<thead>
<tr>
<th></th>
<th>Constant Throughput/Latency</th>
<th>Variable Throughput/Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Energy</strong></td>
<td>Design Time</td>
<td>Sleep Mode</td>
</tr>
<tr>
<td><strong>Leakage</strong></td>
<td>Stack effects + Multi-$V_T$</td>
<td>Sleep T's Multi-$V_{DD}$ Variable $V_T$ + Input control</td>
</tr>
</tbody>
</table>

---
A shared-well technique is appropriate for random placement of cells.
Standard-Cell Dual-Supply-Voltage

- A $V_{DDH}$ circuit is assigned only to a critical path
- A $V_{DDL}$ circuit is used in a non-critical path and for driving a large capacitive load

Shared-Well Dual-Supply-Voltage

- Both circuits can be placed in the same N-well
- Cell layout becomes complex
- An intrinsic negative back-biasing of PMOS degrades speed

Shimazaki, ISSCC'03
ALU Block Diagram

Low Swing Bus & Level Converter

- Delay of INV1 does not increase
- INV2 is placed near 9:1 MUX to increase noise immunity
- Level conversion is done by a domino 9:1 MUX
The dual-supply technique expands the power-delay optimization space.

Adaptive Supply Voltages

Exploit Data Dependent Computation Times To Vary the Supply

from [Nielsen94]

(IEEE Transactions on VLSI Systems)
Processors for Portable Devices

- Eliminate performance ↔ energy trade-off.

Typical MPEG IDCT Histogram
**Processor Usage Model**

*Desired Throughput*  
Compute-intensive and low-latency processes  
Maximum Processor Speed  
System Idle  
Background and high-latency processes  

**System Optimizations:**  
- Maximize Peak Throughput  
- Minimize Average Energy/operation

---

**Common Design Approaches (Fixed VDD)**

*Compute ASAP:*  
Excess throughput  
Always high throughput  

*Clock Frequency Reduction:*  
Energy/operation remains unchanged...  
while throughput scaled down with $f_{CLK}$
Scale $V_{DD}$ with Clock Frequency

Constant supply voltage

Energy/operation

Throughput ($\propto f_{CLK}$)

Reduce $V_{DD}$, slow circuits down.

~10x Energy Reduction

Burd
ISSCC’00

CMOS Circuits Track Over $V_{DD}$

Normalized max. $f_{CLK}$

$V_{DD}$

Delay tracks within +/- 10%

Burd
ISSCC’00
Dynamic Voltage Scaling (DVS)

1. Vary $f_{CLK}$, $V_{DD}$
2. Dynamically adapt

- Dynamically scale energy/operation with throughput.
- Always minimize speed $\rightarrow$ minimize average energy/operation.
- Extend battery life up to 10x with the exact same hardware!

Operating System Sets Processor Speed

- DVS requires a voltage scheduler (VS).
- VS predicts workload to estimate CPU cycles.
- Applications supply completion deadlines.

$$\frac{CPU \text{ cycles}}{\Delta \text{ time}} = F_{\text{DESIRED}}$$
Converter Loop Sets $V_{DD}$, $f_{CLK}$

- Feedback loop sets $V_{DD}$ so that $F_{ERR} \rightarrow 0$.
- Ring oscillator delay-matched to CPU critical paths.
- Custom loop implementation $\rightarrow$ Can optimize $C_{DD}$.

Recent DVS-Enabled Microprocessors

- **Xscale**: 180nm 1.8V bulk-CMOS [Intel00]
  - 0.7-1.75V, 200-1000MHz, 55-1500mW (typ)
  - Max. Energy Efficiency: ~23 MIPS/mW

- **PowerPC**: 180nm 1.8V bulk-CMOS [Nowka02]
  - 0.9-1.95V, 11-380MHz, 53-500mW (typ)
  - Max. Energy Efficiency: ~11 MIPS/mW

- **Crusoe**: 130nm 1.5V bulk-CMOS [Transmeta03]
  - 0.8-1.3V, 300-1000MHz, 0.85-7.5W (peak)

- **Pentium M**: 130nm 1.5V bulk-CMOS [Intel03]
  - 0.95-1.5V, 600-1600MHz, 4.2-31W (peak)
Design Over Wide Range of Voltages

- Circuit design constraints. (Functional verification)
- Circuit delay variation. (Timing verification)
- Noise margin reduction. (Power grid, coupling)
- Delay sensitivity. (Local power distribution)

Design verification complexity similar to high-performance processor design @ fixed $V_{DD}$

Delay Variation & Circuit Constraints

- Cannot use NMOS pass gates – fails for $V_{DD} < 2V_T$.
- Functional verification only needed at one $V_{DD}$ value.
Relative Delay Variation

- Timing verification only needed at min. & max. $V_{DD}$.

Delay relative to ring oscillator

Four extreme cases of critical paths:

- Gate
- Interconnect
- Diffusion
- Series

All vary monotonically with $V_{DD}$.

Delay Sensitivity

$$\frac{\Delta \text{Delay}}{\text{Delay}} \approx \frac{\partial \text{Delay}}{\partial V_{DD}} \cdot \frac{\Delta V_{DD}}{\text{Delay}(V_{DD})}, \quad \Delta V_{DD} = I(V_{DD}) \cdot R$$

- Design of local power grid (for timing constraints) only need to consider $V_{DD} \approx 2V_T$. 

Burd
ISSCC'00
Design for Dynamically Varying VDD

- Static CMOS logic.
- Ring oscillator.
- Dynamic logic (& tri-state busses).
- Sense amp (& memory cell).

Max. allowed $|dV_{DD}/dt| \rightarrow$ Min. $C_{DD} = 100nF \ (0.6 \mu m)$
Circuits continue to properly operate as $V_{DD}$ changes

Static CMOS Logic

- $V_{in} = 0 \Rightarrow V_{out} = V_{DD}$
- $r_{dsPMOS} \leq 4\text{ns}$
- $0.6 \mu m$ CMOS: $|dV_{DD}/dt| < 200V/\mu s$
- Static CMOS robustly operates with varying $V_{DD}$.
Ring Oscillator

- Output $f_{\text{CLK}}$ instantaneously adapts to new $V_{\text{DD}}$.

Dynamic Logic

- Cannot gate clock in evaluation state.
- Tri-state busses fail similarly → Use hold circuit.

$0.6\mu\text{m CMOS: } |dV_{\text{DD}}/dt| < 20\text{V/}\mu\text{s}$

- False logic low: $\Delta V_{\text{DD}} > V_{\text{TP}}$
- Latch-up: $\Delta V_{\text{DD}} > V_{\text{be}}$

Simulated with $dV_{\text{DD}}/dt = 20\text{V}/\mu\text{s}$
Measured System Performance & Energy

- Dynamic operation can increase energy efficiency > 10x.

V_{DD}-Hopping

Application slicing and software feedback guarantee real-time operation.

Two hopping levels are sufficient.
Reducing Active Power: System-Level

Don’t switch capacitance if you don’t need to!

- Application-specific processing
- Preservation of data correlations
- Locality of reference
- Distributed processing
- Demand-driven / Data-driven computation

Clock gating

Requires careful skew control ...
Well handled in today’s EDA tools
Clock-gating efficiently reduces power

Without clock gating

With clock gating

Power [mW]

90% of F/F’s were clock-gated.

70% power reduction by clock-gating alone.

Courtesy M. Ohashi, Matsushita, ISSCC 2002

Circuit-Level Activity Encoding

from [Stan94]
(1994 International Workshop on Low-power Design)
Number Representation

- Sign-extension activity significantly reduced using sign-magnitude representation