Power Trends & Challenges

- System power dissipation is becoming more critical
  - On-going integration increasing overall system power
  - “Green” systems and “Green” companies imperative
  - Power management no longer limited to mobile applications

- Underlying technology adding to the challenge
  - Device leakage power increasing by process node (mitigated somewhat by FinFET and FDSOI)
  - Voltage no longer scales with process node
  - Package thermal transfer not improving
  - Tools focus on execution

Intelligent Power Management Required!
Complete power management includes:
- Using the right process, libraries and IP
- Leveraging a power-aware design methodology
- Minimizing overhead while ensuring power integrity
Power and Performance Variations for a JPEG Encoder Core

- Wide variation in leakage power seen in same technology depending on target frequency
- Faster doesn’t always imply higher leakage
- Dynamic power (per MHz) is similar and primarily depends on power supply voltage
Library Selection is Complex

- Multiple library variables impact design PPA
  - Number of tracks
  - Channel length
  - VT
  - VDD
  - Availability of power-optimized cells like multi-bit flip-flops
# Standard Cell Libraries for Low Power Design

## Features

<table>
<thead>
<tr>
<th>Low Power Baseline</th>
<th>Features</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Characterized over wider VDD range</td>
<td>Multi-VDD operation</td>
</tr>
<tr>
<td></td>
<td>Multi-Vt components</td>
<td>Leakage minimization</td>
</tr>
<tr>
<td></td>
<td>Multi-bit flipflops</td>
<td>Dynamic power reduction</td>
</tr>
<tr>
<td></td>
<td>Decoupling capacitors</td>
<td>Power integrity</td>
</tr>
<tr>
<td>Multi-VDD</td>
<td>Level shifters</td>
<td>Level translation between multi-VDD islands</td>
</tr>
<tr>
<td></td>
<td>Isolation cells</td>
<td>Control the logic inputs for interface between OFF and ON domains</td>
</tr>
<tr>
<td></td>
<td>Retention flops</td>
<td>Saves states in shut-down mode</td>
</tr>
<tr>
<td></td>
<td>Power gating switches</td>
<td>Enables island power-down</td>
</tr>
<tr>
<td>Long Channel Devices</td>
<td>Long channel standard cells</td>
<td>Lowers leakage on non-critical paths</td>
</tr>
<tr>
<td>Gate Length Bias</td>
<td>Automated biasing</td>
<td>Lowers leakage on non-critical paths</td>
</tr>
</tbody>
</table>
Using higher VDD with a high VT and/or long channel library helps reduce leakage for the same performance.
Impact of Multi-Vt Cells For Power

Cortex A7 core system in 28nm

- ULVT cells for 1GHz operation at SS, 0.81V, -40C
- Multi-Vt cells for leakage power recovery

<table>
<thead>
<tr>
<th></th>
<th>Leakage Power (mW)</th>
<th>Active Power (mW)</th>
<th>Total Power (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ULVT</td>
<td>430</td>
<td>404</td>
<td>834</td>
</tr>
<tr>
<td>Mixed VT</td>
<td>191</td>
<td>374</td>
<td>565</td>
</tr>
<tr>
<td>Reduction</td>
<td>55%</td>
<td>7%</td>
<td>32%</td>
</tr>
</tbody>
</table>
SRAM Power Management in Memory Compilers

- Low dynamic power
  - Multi-Vt peripheral logic
  - Dual rail operation
  - Ultra low voltage operation using custom logic rule bit cell
  - Memory segmentation

- Low leakage power
  - LL bit cell
  - High Vt peripheral logic
  - Multiple sleep modes
    - Light sleep – 40% leakage power reduction with memory data retention and fast wake-up, uses array biasing; one cycle recovery
    - Deep sleep – 70% leakage power reduction with memory data retention and integrated power switches, uses periphery shutdown, ten cycle recovery
    - Shutdown – 90% leakage power reduction with integrated power switches, fifty cycle recovery time, no data retention
## Single Port Optimized Low Voltage/Low Power SRAM

### 40LP Logic Rule Approach

<table>
<thead>
<tr>
<th>Memory architecture</th>
<th>Word</th>
<th>Bit</th>
<th>Mux</th>
<th>Area (um^2) pre-shrink</th>
<th>Read Power</th>
<th>Write Power</th>
<th>Leakage Power</th>
<th>Leakage Power – shutdown mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF / 1.21V / 125C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 40LP SP HD design using 6T Bit Cell</td>
<td>2048</td>
<td>64</td>
<td>4</td>
<td>50,039</td>
<td>19.38</td>
<td>23.04</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8192</td>
<td>64</td>
<td>16</td>
<td>189,764</td>
<td>43.78</td>
<td>43.03</td>
<td>2.52</td>
<td></td>
</tr>
<tr>
<td>eSilicon 40LP Custom Low Power / Low Voltage SRAM using Logic Rule Bit Cell</td>
<td>2048</td>
<td>64</td>
<td>2</td>
<td>113,000</td>
<td>14.75</td>
<td>10.99</td>
<td>0.435</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>8192</td>
<td>64</td>
<td>4</td>
<td>419,000</td>
<td>41.06</td>
<td>29.04</td>
<td>1.595</td>
<td>0.033</td>
</tr>
<tr>
<td>FF / 0.85V / 125C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2048</td>
<td>64</td>
<td>2</td>
<td>113,000</td>
<td>6.76</td>
<td>5.28</td>
<td>0.185</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>8192</td>
<td>64</td>
<td>4</td>
<td>419,000</td>
<td>19.62</td>
<td>14.11</td>
<td>0.684</td>
<td>0.017</td>
<td></td>
</tr>
</tbody>
</table>

**Standard HD architecture (top) vs LP/LV architecture (bottom)**
ASIC Example – Network Processor
Reducing Power at the Same Performance

- Technology: 28nm
- 394Mb memory subsystem
- Customization
  - Standard Vt memory array operated at nominal VDD
  - Migrated memory peripheral logic to high Vt
  - Re-characterized for overdrive operating voltage

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Array Leakage (mW)</th>
<th>Periphery Leakage (mW)</th>
<th>Total Leakage (mW)</th>
<th>Array Leakage (mW) Overdrive</th>
<th>Periphery Leakage (mW) Overdrive</th>
<th>Total Leakage (mW) Overdrive</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP SRAM</td>
<td>231</td>
<td>3726</td>
<td>3957</td>
<td>304</td>
<td>2844</td>
<td>3075</td>
</tr>
<tr>
<td>2P RF</td>
<td>2653</td>
<td>1211/14/70</td>
<td>2/69</td>
<td>9250</td>
<td>11903</td>
<td></td>
</tr>
<tr>
<td>5P SRAM</td>
<td>262</td>
<td>1966</td>
<td>2227</td>
<td>313</td>
<td>1500</td>
<td>1767</td>
</tr>
</tbody>
</table>

Nominal VDD with SVt = 21 Watts
Overdrive VDD with HVt = 16.7 Watts

Result = Same Performance
Static Power Savings = 20%
Power-Aware Design Methodology

- Dynamic Power
- Static Power
- Libraries & Process
- Power Integrity

Power Management System
Architectural Optimization

- **Example:** Memory organization

Total system power management begins at the architectural level
Memory Selection and Optimization

- Typical chip statistic: 80% of the memory area and power are contributed by 20% of the memories
- Memory selection is critical for area and power optimization
- EDA tools focus on logic optimization but leave memory optimization to user
- Traditional memory selection is manual and often ad-hoc
- Memories are changed primarily for functional reasons

![Leakage Power vs. Frequency](image)

1kx32 Single port memory choices
Memory Selection & Optimization Using Generic Memory Models

- eSilicon generic memory model (GMEM) provides automated memory optimization
  - User RTL is based on parametrized generic memory models
  - eSilicon tools select memory based on synthesis results and memory compiler constraints
Signoff for Power

- **Standard signoff (28nm)**
  - Voltage regulator tolerance: 5%
  - IR drop: 5%
  - Timing: SS, VDD – 10%, -40C
  - Power: FF, VDD + 5%, 125C

- **Aggressive signoff (28nm)**
  - Voltage regulator tolerance: 3%
  - IR drop: 3%
  - Timing: SS, VDD – 6%, 0C
  - Power: FFG, VDD + 3%, 105C
  - Lower and re-center VDD

VDD = 0.9V
- IR drop: 45mV
- Voltage regulator tolerance: 45mV
- Timing: SS, 0.81V, -40C
- Power: FF, 0.945V, 125C

VDD = 0.864V
- IR drop: 27mV
- Voltage regulator tolerance: 27mV
- Timing: SS, 0.81V, 0C
- Power: FFG, 0.891V, 105C
Easier to close timing
- Process is faster at 0C compared to -40C by about 8%
- Can use more transistors with higher Vt
- Can use additional margin

Lower Power
- FFG is more realistic as local variations in wafer average out
- Leakage is significantly reduced by lowering VDDmax and temperature
- Leakage reduction: 63%
- Active power reduction 13%
Voltage Scaling and Binning

- Voltage scaling is the most effective method of reducing power for FF parts
  - Scaling can be continuous or discrete
  - Discrete scaling is equivalent to binning
- Binning allows parts to be separated based on their process corner
  - Different voltages for each bin ensure that performance is met while power is optimized
Voltage Scaling for Performance and Power

- The FF, $V_{\text{max}}$, 125C (Fast) part gives the maximum performance.
- The TT, $V_{\text{nom}}$, 25C (Typical) part loses performance due to all three components $P$, $V$, and $T$.
- The worst performance is from the SS, $V_{\text{min}}$, -40C (Slow) part.
## Voltage Scaling for Performance and Power

<table>
<thead>
<tr>
<th>Target Period</th>
<th>Fast</th>
<th>Typical</th>
<th>Slow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate and wire delays</td>
<td>-ΔV</td>
<td>-ΔV</td>
<td>+ΔV</td>
</tr>
<tr>
<td>FF, $V_{\text{max}}$, 125°C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT, $V_{\text{nom}}$, 25°C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS, $V_{\text{min}}$, -40°C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Performance can be recovered by increasing voltage on a part
- If target frequency is lower than that of the Fast part, power can be recovered by lowering the voltage of the Fast part
- Optimum performance and power can be achieved by centering the part around the target frequency
• Foundry process is well-controlled
• $3\sigma$ signoff is extremely conservative
• Process and yield management can yield power and performance improvements
  • Shift process by $1\sigma$
  • Discard parts outside $2\sigma$ (4.6% yield loss)

Dark blue is less than one standard deviation away from the mean. For the normal distribution, this accounts for 68.2% of the set, while two standard deviations from the mean (medium and dark blue) account for 95.4%, and three standard deviations (light, medium, and dark blue) account for 99.7%.
Power Management Case Study

- Customer requires maximum total power dissipation to not exceed 40W
- Tapeout ready database power at FFG, 1.05V, 105C is 66.58W

Assumptions:
- Typical voltage at 1V
- Voltage regulator tolerance is ±50mV
- IR Drop is 50mV
- Operating frequency of 600MHz

How do we meet customer’s power target?

<table>
<thead>
<tr>
<th></th>
<th>Total (W)</th>
<th>Dynamic (W)</th>
<th>Leakage (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logic</td>
<td>35.02</td>
<td>25.74</td>
<td>9.28</td>
</tr>
<tr>
<td>Memory</td>
<td>20.78</td>
<td>7.76</td>
<td>13.02</td>
</tr>
<tr>
<td>CAM</td>
<td>3.24</td>
<td>1.17</td>
<td>2.06</td>
</tr>
<tr>
<td>SerDes (AVDD)</td>
<td>7.48</td>
<td>6.19</td>
<td>1.29</td>
</tr>
<tr>
<td>IO ring + rest</td>
<td>0.08</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>66.58</strong></td>
<td><strong>40.93</strong></td>
<td><strong>25.65</strong></td>
</tr>
</tbody>
</table>
Any strategy for power reduction to achieve 40W requires:
- Use of two power supplies
  - Power supply for SerDes (1V typical)
  - Power supply for Core
- Lower frequency operation

Core supply voltage tolerance should be as low as possible:
- Customer is unwilling to reduce voltage tolerance below 50mV

Minimum core voltage is 0.81V (memory VDDmin)

Use IR drop based on actual data from power analysis:
- 30mV simulated at 125C with >100W chip power (FF, 1.05V, 125C)
- 15mV assumed for 105C with < 50W chip power (linear scaling)
Power Management Solution

- Operating Frequency is 450MHz
- Separate power supplies for SerDes and Core
- SerDes VDD = 1V ±50mV
- Two bins, bins separated at TT
- Process skewed by one sigma
- Bin 1 (TT to SS-)  
  - Core VDD = 0.93 ±50mV  
  - Power at TT, 0.98V, 105C 35.96W  
  - Worst timing at SS-, 0.865V, -40C
- Bin 2 (TT+ to TT) No yield loss  
  - Core VDD = 0.875 ±50mV  
  - Power at FFG-, 0.925V, 105C 40.2W  
  - Worst timing at TT, 0.81V, -40C
- Bin 2 (TT+ to TT) 2.1% yield loss  
  - Core VDD = 0.875 ±50mV  
  - Power at TT+, 0.925V, 105C 36.35W  
  - Worst timing at TT, 0.81V, -40C
Power Management System Summary

- A complete system approach from concept to production
  - Lower dynamic and static power
  - Power integrity ensured
- Leveraging the best process, IP, libraries, power aware tools and power management methodology
- Power management solutions complement traditional EDA solutions

Intelligent Power Management Delivered