Towards Variation-Aware System-Level Power Estimation of DRAMs: An Empirical Approach

Karthik Chandrasekar
TU Delft

Christian Weis$, Benny Akesson*, Norbert Wehn$ & Kees Goossens#
Problem Statement

- Below 90nm, impact of process-variations on power is significant.
- Studied well for processors but NOT for DRAMs (worst-case IDD).
- Unfortunately, this cannot be ignored! DRAMs influence system power.
- How severe is the impact of variations on DRAM power?

* Normal Distribution ($\mu, \sigma$) for 11,000 DDR3 1Gb-533-x8 devices at 70nm

- Unfortunately, there are NO known models or measures publicly available.
Contributions [What do we propose?]

- (1) To empirically estimate the impact of variation on DRAM power
  - Using Monte-Carlo simulations on a baseline SPICE DRAM cross-section.
  - Extending for different DRAM frequencies, capacities and data-widths.

- (2) To empirically characterize a given DRAM at boot-time
  - By assessing the actual performance of a given DRAM (dictates power).

- (3) To derive actual power consumption
  - Using distributions from the Monte-Carlo simulations.
  - Using the actual performance from characterization.
Estimation of impact of variation on DRAM power consumption
Baseline DRAM Bank cross-section model
- NGSPICE - BSIM Model Cards - 45nm LP-PTM (Bulk CMOS)
- 1Gb DDR3-533MHz - x8 data-width
- Core Timings adhered to: CL-RCD-RP : 7-7-7 cc (JEDEC specified)
- 1T1C DRAM cell & 512 cells per local bitline/wordline
- Column Select Lines (CSLs) & Master Wordlines
- Primary & Secondary Sense Amplifiers & IO Buffers
**SPICE Simulation Results**

- Functionality and Timings Verified: ACT-RD-PRE & Core-Timings
- Monte-Carlo: Device & Interconnect variations (ITRS/IBM) introduced
- 1000 MC runs in NGSPICE: 4 important device parameters
  - Global - Channel Length (l), Mobility (μ), Oxide Thickness (T_{OX}) & Local - V_{TH} (LER)
**Impact of Variation on Currents (IDDs)**

- **Baseline, Monte-Carlo & Extensions:** 3 parameters (Freq, Cap, Width)
  - Frequency [F] (533 to 800): All currents change except leakage
  - Capacity [C] (1Gb to 2Gb): Changes in all currents
  - Data-Width [W] (x8 to x16): Data transfer currents change

<table>
<thead>
<tr>
<th>Config</th>
<th>Baseline 1Gb-533-x8</th>
<th>Freq 1Gb-800-x8</th>
<th>Capacity 2Gb-533-x8</th>
<th>Width 1Gb-533-x16</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{DD}$ Type</td>
<td>$\mu$ mA</td>
<td>$\sigma$ %</td>
<td>$\mu$ mA</td>
<td>$\sigma$ %</td>
</tr>
<tr>
<td>$I_{DD0}$</td>
<td>98.4</td>
<td>2.4</td>
<td>112</td>
<td>2.6</td>
</tr>
<tr>
<td>$I_{DD1}$</td>
<td>104</td>
<td>2.3</td>
<td>118</td>
<td>2.2</td>
</tr>
<tr>
<td>$I_{DD2N}$</td>
<td>37.7</td>
<td>4.8</td>
<td>42.7</td>
<td>4.5</td>
</tr>
<tr>
<td>$I_{DD3N}$</td>
<td>41.5</td>
<td>5.7</td>
<td>56.7</td>
<td>4.5</td>
</tr>
<tr>
<td>$I_{DD4R}$</td>
<td>118</td>
<td>2.9</td>
<td>153</td>
<td>3.5</td>
</tr>
<tr>
<td>$I_{DD4W}$</td>
<td>123</td>
<td>2.7</td>
<td>159</td>
<td>4.1</td>
</tr>
<tr>
<td>$I_{DD5}$</td>
<td>146</td>
<td>2.1</td>
<td>161</td>
<td>2.4</td>
</tr>
<tr>
<td>$I_{DD2P0}$</td>
<td>8.4</td>
<td>13.7</td>
<td>8.4</td>
<td>13.7</td>
</tr>
<tr>
<td>$I_{DD6}$</td>
<td>8</td>
<td>14</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>
Characterization of a given DRAM at boot-time
Characterization of a given DRAM at boot-time

- To identify actual performance and thereby, power of a given DRAM.
- Performance defined by: Core-timings & frequency- Functional Speed (FS)
- Get IDD & FS values for $\mu$ and $\pm 6\sigma$ variations (distribution) from SPICE.

- @Boot-time, for all CT, F & banks, run IDD1 test: get FS/bank (< 100ms)
- Use Min(FS) of all - Fastest for IDDs (+1$\sigma$ conservative) - actual worst-case
- Optionally, Use Max(FS) of all - Slowest for CT & F (-1$\sigma$ conservative)
Results & Analysis

- Mean Vs. Datasheet (DS): Micron 2Gb DDR3-800-x16 device (50nm)
  - Combining influence of all three system parameters (F, C & W)
  - $\mu$ is 30% < DS for active & 86% < DS for leakage (60% @ 70nm)

- Application energy differences:
  - Up to 58% worst-case diff. using $\mu$ and 28% using $+2\sigma$

<table>
<thead>
<tr>
<th>Current</th>
<th>$DS$ mA</th>
<th>$\mu$ (mA)</th>
<th>$\mu$ vs DS</th>
<th>$2\sigma$ mA</th>
<th>$2\sigma$ vs DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDD0</td>
<td>110</td>
<td>98</td>
<td>-12.2</td>
<td>102.8</td>
<td>-6.96</td>
</tr>
<tr>
<td>IDD1</td>
<td>125</td>
<td>112.2</td>
<td>-11.4</td>
<td>117.3</td>
<td>-6.53</td>
</tr>
<tr>
<td>IDD2N</td>
<td>42</td>
<td>33.5</td>
<td>-25.4</td>
<td>36.9</td>
<td>-13.8</td>
</tr>
<tr>
<td>IDD3N</td>
<td>45</td>
<td>34.6</td>
<td>-30.1</td>
<td>38.7</td>
<td>-16.1</td>
</tr>
<tr>
<td>IDD4R</td>
<td>270</td>
<td>232.2</td>
<td>-16.2</td>
<td>247.3</td>
<td>-9.15</td>
</tr>
<tr>
<td>IDD4W</td>
<td>280</td>
<td>246.7</td>
<td>-13.5</td>
<td>260</td>
<td>-7.67</td>
</tr>
<tr>
<td>IDD5</td>
<td>215</td>
<td>193.6</td>
<td>-11.1</td>
<td>202.1</td>
<td>-6.34</td>
</tr>
<tr>
<td>IDD2P0</td>
<td>12</td>
<td>6.62</td>
<td>-81.1</td>
<td>8.77</td>
<td>-36.7</td>
</tr>
<tr>
<td>IDD6</td>
<td>12</td>
<td>6.45</td>
<td>-85.9</td>
<td>8.67</td>
<td>-38.4</td>
</tr>
</tbody>
</table>
Conclusions

Problems:

- Process-variations significantly impact DRAM power consumption.
- Vendors only provide worst-case current measures.
- Every DRAM device has individual performance and current measures.

Solutions:

- We proposed variation-aware power estimation for DRAMs.
- We proposed boot-time DRAM characterization methodology.
- We identified actual DRAM performance and realistic power measures.

- Added new variation-aware datasheets to the open-source DRAM power tool - **DRAMPower**. Available at: [www.drampower.info](http://www.drampower.info)
Thank You!

Questions?

k.chandrasekar@tudelft.nl

www.drampower.info