

Parametrical Characterization of leakage power in Nanoscale Technologies

D. Kudithipudi, E. John

University of Texas at San Antonio

(PUBLIC session)

Rapid progress in deep sub-micron technologies has led to a new era of nanotechnology. The need for enhanced performance along with low-power consumption will necessitate the usage of nanotechnologies in the next generation embedded applications. This evolution has compelled designers to understand the behavior and effects of various process parameters on such systems. In general, with decreasing technology sizes, the sub-threshold leakage and gate-oxide leakage power dissipation tend to form a significant portion of the total power consumption. Specifically, the usage of nanotechnology is associated with constraints arising from sub-threshold leakage power in CMOS circuits. Sub-threshold leakage current is affected by various parameters such as supply voltage, technology size, process parameters and temperature. Furthermore, on-chip caches occupy up to 50% of the processor and therefore contribute a significant portion to the leakage power in embedded applications. Therefore, it is necessary to characterize the cache leakage power in such applications to develop effective nanoscale embedded systems.

To this end, we have studied various parameters affecting leakage power in 70nm and 100nm sizes. Hotleakage, a temperature-aware leakage power model was used to perform architectural-level simulations. The workloads used for these simulations were obtained from the MiBench suite. The inputs to this simulator were executables compiled by SimpleScalar's gcc compiler for ALPHA instruction set.

There was a one-order increase in baseline leakage power consumption, for 70nm technology, as compared to 100nm. When the cache sizes were increased from 16k to 128k for L1, and 256k to 2M for L2, the dissipated leakage power demonstrated an increase of 50% in both set-ups. The benchmarks with greater numbers of conditional and unconditional branches, consumed more power as compared to other workloads. Furthermore, increasing the temperature had exponential effects on gate leakage power.

We also implemented leakage reduction techniques on L1 and L2 caches. The effects of various parameters (L1 and L2 cache sizes, latency, temperature, and oxide thickness) on leakage power were evaluated using gated-Vss, a leakage control technique. The power savings in 70nm for simulations performed on 16k L1 cache, 2M L2 cache with a latency of 10cc was 59% at 80°C, and increased to 64% at 100°C. Increasing the latency of L2 cache increased the total leakage power dissipation. Gated-Vss performed better in L1 caches, despite using additional execution cycles. Leakage savings were lesser in thinner oxides, and hence systems with higher misprediction and dense memory operations demonstrated optimum savings.

In summary, these results support evidence that leakage power in L1 and L2 caches increase significantly with increasing cache sizes and decreasing technology sizes, with greater consumption of bitline leakage power in complex programs. We further characterized the leakage control techniques in nanoscale technologies.