18.2 A 1.1GHz 12µA/Mb-Leakage SRAM Design in 65nm Ultra-Low-Power CMOS with Integrated Leakage Reduction for Mobile Applications

Intel, Hillsboro, OR

Growing demand for high-performance multi-media processor in handheld devices continues to drive the need for large embedded SRAM with low power consumption. The low-power requirement has often been addressed by compromising SRAM performance through the adoption of lower supply voltage and slower performing transistors with low leakage. Supply voltage scaling in large SRAM has not kept pace with technology scaling due to reduced SRAM cell stability and write margin [1]. Many circuit techniques have been proposed [2, 3, 4] to expand design margins for low-voltage operation while reducing leakage current at standby mode. In this work, we report an SRAM design in a 65nm ultra-low-power (ULP) CMOS process [5]. Co-optimization between process technology and circuit design allowed us to achieve fast SRAM access at nominal operating voltage. A large design window in operating voltage along with integrated leakage reduction schemes provides flexibility for both active power management and leakage reduction through dynamic voltage scaling.

The 65nm ULP CMOS process, featuring uni-axial strained silicon transistors, provides a significantly improved trade-off between transistor drive current and leakage [5]. This is essential to achieve fast SRAM access at substantially lower operating voltage and leakage. Standby leakage of the SRAM cell is optimized by simultaneously reducing gate, subthreshold and junction leakage components to sub-100pA/µm at 1.2V, while maintaining competitive cell read current and meeting the retention voltage requirement for a 1Mb SRAM. The low damage junction engineering employed in this technology results in substantial junction leakage reduction while still maintaining good short channel control. Gate oxide thickness optimization and gate nitridation are applied to reduce gate leakage. Well and pocket implants and source/drain spacers are optimized simultaneously to reduce subthreshold leakage. To achieve minimum supply voltage ($V_{min}$) needed for the target transistor I on/Ioff performance [5], the supply voltage is lowered to 0.5V. The measured average cell leakage is reduced to ~2pA/bit. PMOS transistors, providing a significantly improved trade-off between transistor drive current and leakage [5], is 1.2V while the array supply voltage is lowered to 0.5V.

A 44Mb SRAM chip containing the 1Mb ULP SRAM macro with column redundancy, electrical-programmable fuse, and programmable sleep bias, is built using a 65nm 8-metal layer land-grid-array (LGA) package with flip-chip technology. Figure 18.2.4 shows the measured average cell leakage from an SRAM array with all three leakage reduction schemes turned on. By lowering supply to the retention voltage of 0.5V and maintaining the PMOS n-well voltage at the nominal voltage of 1.2V, the measured average cell leakage is reduced to ~2pA/bit. PMOS reverse-body-bias reduces the cell leakage by ~30% at the retention voltage of 1.2V, which is mainly due to the reduction in PMOS sub-threshold leakage. Figure 18.2.5 shows the standby leakage of the 1Mb SRAM operating in standby, low-power and high-speed modes. In standby mode, the 1Mb SRAM macro draws ~12µA of leakage current. The leakage current increases to 22µA for low-power operation at 0.7V and 74µA for high-speed operation at 1.2V. For high-speed operation, by turning on the NMOS sleep feature and floating bitlines in unselected sub-arrays, the array leakage current of the 1Mb SRAM macro is reduced by ~70%.

Figure 18.2.6 shows the measured average leakage from the 1Mb SRAM across a wide range of supply voltages. It achieves 1.1GHz frequency at a nominal voltage of 1.2V and 250MHz at 0.7V, which represents the highest reported frequency for the same class of standby power consumption for mobile applications. The die micrograph of the 44Mb ULP SRAM is shown in Fig. 18.2.7.

Acknowledgements: The authors are grateful to many members of Advanced Memory Design team and PTD technical staff for their contributions to this work.

References:
Figure 18.2.1: Configuration and features of 1Mb SRAM with integrated sleep transistor.

Figure 18.2.2: Read path schematic, simulation timing diagram, and single-cycle throughput pipeline of 128kb sub-array.

Figure 18.2.3: SRAM array leakage reduction features, and designs of sleep, bias transistors and power lines in a 128kb sub-array.

Figure 18.2.4: Measured SRAM cell leakage at various leakage reduction schemes.

Figure 18.2.5: Standby leakage current of 1Mb SRAM macro.

Figure 18.2.6: Maximum access frequency vs. supply voltage of 1MbULP SRAM macro.

Continued on Page 606
Figure 18.2.7 Die micrograph of 44Mb ultra-low-power SRAM.