

## 7.5 A 28nm 0.6V Low-Power DSP for Mobile Applications

Gordon Gammie<sup>1</sup>, Nathan Ickes<sup>2</sup>, Mahmut E Sinangil<sup>2</sup>, Rahul Rithe<sup>2</sup>, J. Gu<sup>3</sup>, Alice Wang<sup>1</sup>, Hugh Mair<sup>1</sup>, Satyendra Datla<sup>1</sup>, Bing Rong<sup>1</sup>, Sushma Honnavara-Prasad<sup>1</sup>, Lam Ho<sup>1</sup>, Greg Baldwin<sup>1</sup>, Dennis Buss<sup>1</sup>, Anantha P Chandrakasan<sup>2</sup>, Uming Ko<sup>1</sup>

<sup>1</sup>Texas Instruments, Dallas, TX,

<sup>2</sup>Massachusetts Institute of Technology, Cambridge, MA,

<sup>3</sup>Texas Instruments (now with MaxLinear), Dallas, TX

A multimedia applications processor is fabricated using a 28nm low-power process technology for ultra-low-power applications. Based on a 4-issue, 32-register version of the TMS320C64x+ VLIW DSP, this System on Chip (SoC) includes 32kB L1 and 128kB L2 caches, and I2S, SPI, UART, MultiMediaCard, and external memory interfaces (Fig. 7.5.1). The design incorporates over 600K instances of custom low-voltage logic cells and 43 instances (1.6 Mb) of 6T SRAM. Utilizing ultra-low-voltage (ULV) optimized standard-cell libraries and 6T SRAM macros, and demonstrating a new statistical static timing analysis (SSTA) methodology, the SoC scales as designed from high performance at 1.0V down to ultra-low power at 0.6V.

The 28nm low-power (LP) technology (Fig. 7.5.2) used for this DSP SoC design is a custom process with a dual-gate poly/SiON gate stack, double patterning at gate, high-NA 193i lithography and epitaxial S/D SiGe for pMOS performance enhancement. Typical strain techniques are also used for nMOS performance enhancement. The ultra-low-K dielectric dual damascene metal stack includes a thick top Cu level and an Al level that can be used for power and signal routing. The integration techniques described in [1] are utilized to support SoC components (e.g. multi- $V_t$ , multi-channel length, analog and I/O transistors, capacitors and diodes) along with a custom 0.12 $\mu\text{m}^2$  6T SRAM bitcell with minimal added process cost.

Performance of logic circuits operating in or near sub-threshold is highly sensitive to any variation of the threshold voltage ( $V_t$ ), and some circuits can cease to function at the extremes of  $V_t$  variation. To quantify the functionality of each combinational cell, the static noise margin is determined using the procedure reported in [2]. To maintain sufficient reliability and performance at ULV, a custom digital cell library is developed, using a variation-driven design flow. Figure 7.5.3(a) shows typical functional failure modes for logic cells. Through selective adjustment of transistor sizes, a beta ratio of 1 is found to provide optimal performance and ensure functionality over a wide voltage range. To maintain proper clock duty cycle, clock-tree cells require a higher beta ratio (stronger pMOS), with the optimal ratio ranging from 1.5 at high voltage to 2.0 at ULV. The increase in beta ratio is driven by the increased difference in drive current between pMOS and nMOS transistors at ultra-low supply voltages. A beta ratio of 1.5 is chosen for clock buffers to ensure performance goals at high voltage are met. To alleviate flip-flop failures from data slip through or reverse conduction, an inverter is inserted between the master latch and pass-gate in order to delay the turn-on of the master stages and avoid reverse current flow. In the SRAM, a hierarchical bit-line is used to improve read stability, with wordline boosting to ensure write-ability, and a pre-read during write to avoid half-select-related disturbances (Fig. 7.5.3(b)) [3].

Timing closure is a challenging problem for ULV designs. As shown in Fig. 7.5.4(a) for a representative library cell, the local  $3\sigma$  delay variation is larger than the  $3\sigma$  global corner delay by 1.5 $\times$ . As the supply voltage decreases from 1.0 to 0.5V, the  $3\sigma$  global corner delay increases by 15 $\times$  and the standard deviation of the local delay increases by 100 $\times$ . The relative impact of local variation varies with the drive strength: as drive strength increases, the standard deviation of local variation decreases and the PDF becomes more Gaussian (Figure 7.5.4(b)). During synthesis, the use of low-drive-strength cells on critical paths was reduced by applying delay-derating factors. The derating factors were calculated for each library cell, and are proportional to the magnitude of the local variation impact at ULV. Additionally, all clock cells are restricted to drive strengths of 8 $\times$  or higher. The area impact of these drive-strength increases was less than 5%.

Predicting path-level delay distributions for timing closure is another challenge of ULV design. Modeling delay with Gaussian PDFs, as traditional statistical static timing analysis (SSTA) tools do, leads to a consistent underestimate of the actual delay distribution by 10 to 70%. For this design, setup and hold time margins at 0.5V are verified using a ULV SSTA design methodology based on Nonlinear Operating Point Analysis for Local Variations (NLOPALV). NLOPALV models cell delay distributions to within 5% [4], and can be used in conjunction with existing static timing analysis (STA) tools to predict path delays to within 8% [5]. To manage run-time, the analysis is performed in four passes of successively increasing accuracy. In the first pass, non-critical paths are identified and discarded using standard STA with  $3\sigma$  cell delays. Summing  $3\sigma$  cell delays gives highly pessimistic results compared to the actual  $3\sigma$  variation of the overall paths. 92% of setup paths and 95% of hold paths are eliminated in this pass. Three subsequent passes reduce the pessimism by running NLOPALV on the capture clock tree only, capture and launch clock trees, and finally the entire timing path including the datapath. Figure 7.5.5(a) shows the number of paths analyzed decreases dramatically with each pass. The plots in Figure 7.5.5(b) show the distribution of analyzed path delays at the end of each pass. This 4-pass process integrates with and makes use of conventional timing closure tools. This analysis found 87 hold violations at 0.5V before hold-fixing, and verified that no hold violations remained after hold-fixing. The setup analysis shows that the design achieves 14MHz at 0.5V.

The DSP SoC design is fabricated and demonstrated (Fig. 7.5.6) to be operational from 587MHz at 1.0V (113mW) down to 3.6MHz at 0.34V (720 $\mu\text{W}$ ) when operating from external memory (caches disabled). At the ULV target voltage of 0.5V, the maximum frequency is 43.4MHz, as compared to 14MHz from SSTA at worst-case conditions, end of life, and with margins. The on-chip caches are functional for supply voltages above 0.6V. When executing from cache, the chip scales from 145mW at 331MHz (1.0V) down to 5.9mW at 14.4MHz (0.6V). For lower voltage and reliable ULV cache operation in production, redundancy and repair should be implemented. Active and leakage power scale by 60 $\times$  and 8.5 $\times$ , respectively, when executing from cache, and by 1240 $\times$  and 39 $\times$  when executing from external memory. The measured leakages are representative of early development silicon with transistors not yet at final leakage targets for the technology. The minimum energy-per-cycle occurs at 0.75V (cache on), or 0.5V (external memory) and is expected to reduce slightly as leakage is reduced.

### References:

- [1] K. Benaissa, G. Baldwin, S. Liu, P. Srinivasan, F. Hou, B. Obradovic, S., Yu, H. Yang, R. McMullan, V. Reddy, C. Chancellor, S. Venkataraman, H. Lu, S. Dey, and C. Cirba,, "New Cost-Effective Integration Schemes Enabling Analog and High-Voltage Design in Advanced CMOS SOC Technologies", *VLSI Technology Symposium*, pp. 221-222, June 2010.
- [2] J. Kwong and A. Chandrakasan, "Variation-Driven Device Sizing for Minimum Energy Sub-threshold Circuits", *International Symposium on Low Power Electronics and Design*, pp. 8-13, Oct. 2006.
- [3] M. Sinangil, H. Mair, and A. Chandrakasan "A 28nm High-Density 6T SRAM with Optimized Peripheral Assist Circuits for Operation down to 0.6V", *ISSCC Dig. Tech. Papers*, in press, Feb. 2011.
- [4] R. Rithe, S. Chou, J. Gu, A. Wang, S. Datla, G. Gammie, D. Buss, A. Chandrakasan, "Cell Library Characterization at Low Voltage using Non-linear Operating Point Analysis of Local Variations", *International Conference on VLSI Design*, in press, 2011.
- [5] R. Rithe, J. Gu, A. Wang, S. Datla, G. Gammie, D. Buss, and A. Chandrakasan, "Non-Linear Operating Point Statistical Analysis for Local Variations in Logic Timing at Low Voltage", *Design Automation and Test in Europe Conference (DATE)*, pp. 965-968, March 2010.

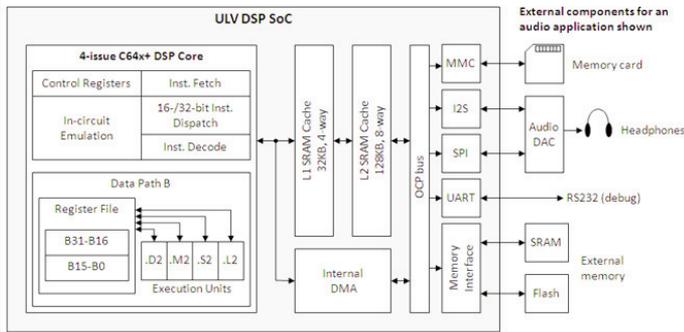
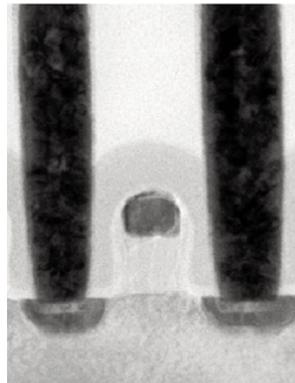


Figure 7.5.1: Series (C64x+) based DSP system with peripherals for low-power applications.



<b>Process</b>	28nm Low Power CMOS
<b>Lithography</b>	high-NA 193i
<b>Performance</b>	pMOS: Epitaxial S/D SiGe nMOS: Strain techniques
<b>Gate stack</b>	dual-gate poly/SiON double patterning
<b>Metal stack</b>	ULK dual damascene Thick top Cu level and Al level
<b>Logic</b>	SVT, LVT, multi-channel length
<b>Bit</b>	0.12 $\mu\text{m}^2$ 6-T SRAM bitcell

Figure 7.5.2: 28nm low-power process.

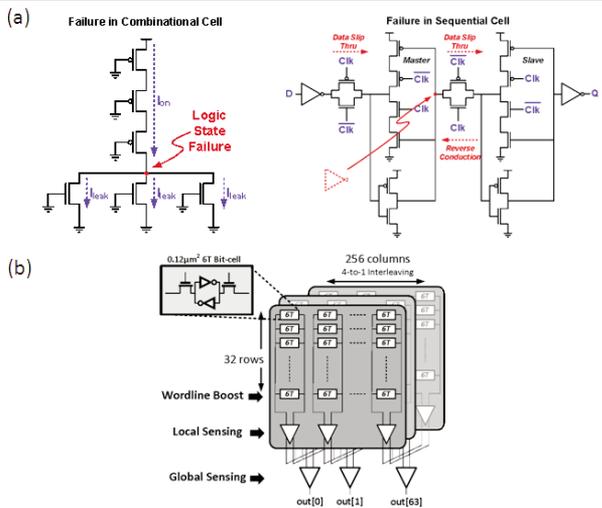


Figure 7.5.3: (a) Typical functional failure modes for logic cells. (b) ULV SRAM with 6T bit-cell, hierarchical sensing, wordline boosting, and pre-read during write.

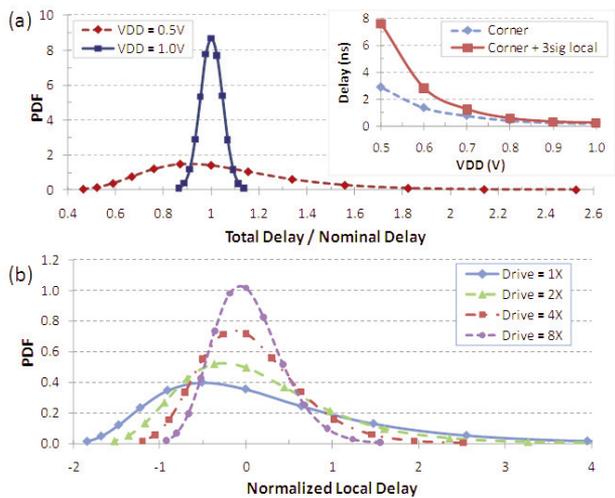


Figure 7.5.4: (a) Delay PDF of a representative library cell at  $V_{DD} = 1.0\text{V}$  and  $V_{DD} = 0.5\text{V}$ . The horizontal axis is normalized the respective  $3\sigma$  global corner delays. (b) Impact of drive strength on PDF.

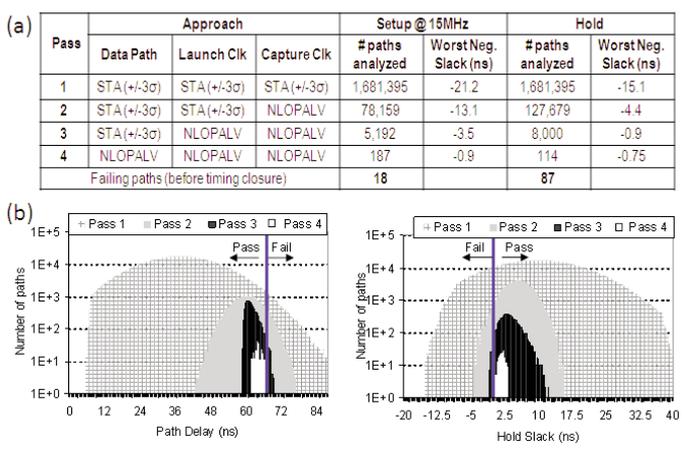


Figure 7.5.5: (a) SSTA 4-pass progressive NLOPALV methodology (b) Distribution of paths for hold and setup analysis using the progressive application of NLOPALV.

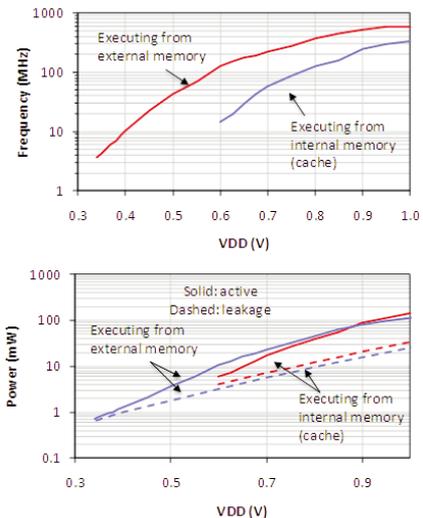


Figure 7.5.6: Measured power and performance, comparing operation from internal and external memories.

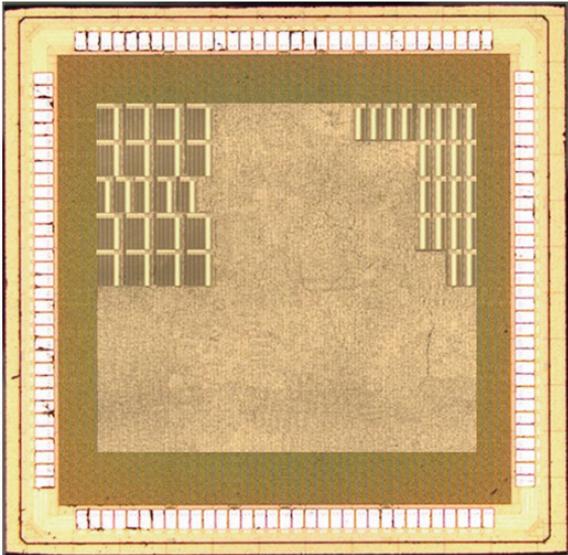


Figure 7.5.7: Chip micrograph.