

Leakage-Biased Domino Circuits for Dynamic Fine-Grain Leakage Reduction*

Seongmoo Heo and Krste Asanović
MIT Laboratory for Computer Science
200 Technology Square
Cambridge, MA 02139
{heomoo,krste}@lcs.mit.edu

Abstract

A Leakage-Biased Domino circuit family is proposed that maintains high speed in active mode but which can be rapidly placed into a low-leakage inactive state by using leakage currents themselves to bias internal nodes. A 32-bit Han-Carlson domino adder circuit is used to compare LB-Domino with conventional single and dual Vt domino circuits. For equal delay and noise margin, the LB-Domino technique gives two decades reduction in steady-state leakage energy compared to a dual-Vt technique.

Introduction

Energy dissipation has emerged as the primary design constraint for many systems, from portable electronics to high-performance microprocessors. Until recently, the dominant cause of energy dissipation in digital CMOS has been dynamic switching of load capacitances. Continuing reductions in feature size reduce capacitance and supply voltage and hence dynamic switching energy per operation but, to maintain performance, threshold voltages must also be scaled down with supply voltage. Unfortunately, lowering the threshold voltage increases static leakage current exponentially, and within a few process generations it is predicted energy dissipation from static leakage current could be comparable to dynamic switching energy [3, 4].

A number of techniques have been proposed to combat this increase in leakage power. These approaches can be divided into two categories. The first category focuses on the static design-time selection of slow transistors on non-critical paths. These techniques include: conventional transistor sizing, lower Vdd [8, 10], stacked gates [14, 24, 21], longer channels [7], higher threshold voltages [20, 9, 19, 23, 1], and thicker T_{ox} ; we collectively refer to these as statically-selected slow transistors (SSSTs). Once SSST techniques have been applied, most leakage current is concentrated on critical paths. For example, in a recent embedded PowerPC 750 design, the lowest threshold transistors accounted for only 5% of the total transistor width but around 50% of the total leakage [11].

Critical path transistors cannot be permanently slowed down to reduce leakage without affecting circuit performance. The second category of leakage reduction techniques dynamically switch the fast transistors into a low-leakage state during idle periods. Techniques to deactivate fast transistors dynamically include body biasing, sleep transistors, and sleep vectors; we collectively refer to these as dynamically-deactivated fast transistors (DDFTs). Body-biasing [13, 17, 18, 6, 5] can reduce leakage to low levels, but incurs a large energy overhead to charge highly-capacitive wells and takes significant time to apply, and so can only be profitably

applied for long idle times. Sleep transistors [13, 15, 23, 22, 16] placed in series with the power supply can reduce leakage to low levels but these impact circuit speed and add to circuit area. In addition, switching the sleep transistors can also have large energy overheads due to charging and discharging of the virtual power supply nodes. Sleep vector techniques [24, 21] drive input vectors into a circuit such that leakage currents are minimized. These techniques are fast, but it is difficult to find a good sleep vector that will propagate a low-leakage state throughout a circuit. Adding sleep vector circuitry to force intermediate nodes to the desired values can increase circuit delay and transition energy overhead.

Domino logic is often used on critical paths, and several DDFT techniques have been proposed to reduce leakage on idle domino blocks. Dual-Vt domino [23] requires additional input gating to force the internal nodes into a sleep state which reduces performance and increases active energy. Also, as shown below, the high-Vt keepers increase active energy once noise margin is equalized [23]. MHS-Domino [1] modifies a clock-delayed keeper circuit to force internal dynamic nodes into a low leakage state. However, the internal node is pulled down through a PMOS leaving the possibility of an intermediate voltage on the dynamic node of the first stage of a domino chain if the data inputs are not high. This can cause short-circuit current in the static output inverter until the leakage through the input transistors finally pulls the dynamic node to ground.

This paper presents a new DDFT circuit family, *Leakage-Biased Domino* (LB-Domino). LB-Domino uses sleep transistors only on non-critical paths and uses the leakage current itself to bias internal critical paths into a minimal leakage state — leakage currents are used to apply the optimal sleep vector. This technique has little impact on active energy or delay when applied to conventional domino circuitry. LB-Domino provides a low-leakage state which can be rapidly entered and exited with low transition energy overhead. This enables fine-grain leakage reduction, where small sub-circuits can be deactivated for short periods of time.

Leakage-Biased Domino

An LB-Domino buffer is shown in Figure 1. This example is a footless domino buffer without a clock transistor in the dynamic pull-down stack, but the LB technique can also be applied to footed domino stacks. Only two small sleep transistors are added to a conventional CMOS domino gate: a high-Vt PMOS in series with the keeper power supply and a high-Vt NMOS in series with the static output logic pulldown. When the sleep signal is deasserted, the circuit operates as a conventional domino gate with minimal performance degradation because there are no additional series transistors in the critical evaluate path.

To place the circuit into sleep mode, the clock signal is left high after an evaluate cycle and the sleep signal is asserted (`sleep=1` and `sleepb=0`). If the data input was high, `node1` would have been discharged. If the data input was low, `node1` is high but the leakage through the NMOS dynamic pull-down stack will slowly

*This work was partly funded by DARPA PAC/C award F30602-00-2-0562 and by NSF CAREER award CCR-0093354.

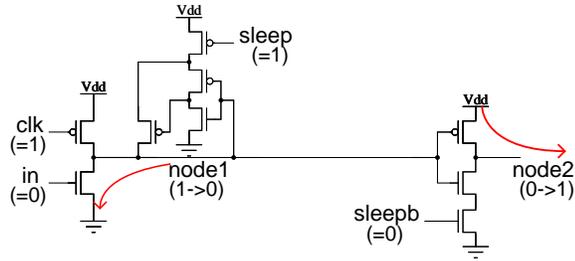


Figure 1: A leakage-biased domino buffer.

discharge the node to ground (the precharge and keeper pull-up transistors are high-Vt devices with significantly lower leakage than the pull-down stack). The NMOS sleep transistor is added to prevent any short-circuit current in the static output logic while the dynamic node discharges to ground. The static output, node2, will rise as the static pull-up turns on. As the leakage current of one domino gate causes its output node to rise, this will cause the NMOS transistors in the pulldown stacks of the following domino gates to turn on, accelerating the discharge of their internal dynamic nodes. In this way, LB-Domino gates bias themselves into a low-leakage state where the internal dynamic nodes are discharged low and static nodes are charged high regardless of input vector state.

When the internal dynamic node is discharged, the main leakage is across the high-Vt PMOS precharge transistor which is turned off by the clock signal remaining high. The leakage path of the static output includes at least two series NMOS transistors, one of which is a high-Vt device. A conventional precharge cycle is used to move from sleep mode back to active mode.

Compared with MHS-Domino, LB-Domino has a simpler sleep mechanism that is compatible with, but does not require, a clock-delayed keeper. LB-Domino also avoids short-circuit current in the static output inverter of the first gate of a domino chain.

Evaluation Methodology

The carry generation circuit of a 32-bit Han-Carlson adder [12] was used to evaluate LB-Domino. The carry generation circuit is pure domino with six levels of alternating dynamic and static logic. The basic propagate-generate cells are shown in Figure 2. Four variants of the design were compared. The first uses only low-Vt transistors (LVT), while the second is a dual-Vt (DVT) design where only evaluation phase transistors are low-Vt. The third variant is an LB-Domino (LB) design based on the DVT but with high-Vt sleep transistors added to the keeper feedback circuits and the static logic pulldowns. The fourth variant (LB2) is another LB-Domino design which only uses high-Vt for the precharge transistors and for the added sleep transistors.

For all four designs, the input and output noise margin of all dynamic circuits was set to 10% of the supply voltage and the precharge/evaluation delays were equalized to within 1% error

Table 1: Processes.

| Process | 180 nm | 70 nm |
|---------------------|--------------|--------------|
| High Vt (NMOS/PMOS) | 0.46V/-0.45V | 0.39V/-0.40V |
| Low Vt (NMOS/PMOS) | 0.27V/-0.23V | 0.15V/-0.18V |
| Vdd | 1.8V | 0.9V |
| Temperature | 100 °C | 100 °C |

Table 2: Input vectors.

| | a | b | ci |
|----------|------------|------------|----|
| Vector 1 | 0x00000000 | 0x00000000 | 0 |
| Vector 2 | 0xffffffff | 0x00000000 | 0 |
| Vector 3 | 0xffffffff | 0xffffffff | 1 |

through transistor sizing. The circuits were designed for an existing TSMC 180 nm process and a projected 70 nm process obtained from the BPTM project [2] (Table 1). All simulations used HSPICE.

Since both active energy and leakage power are dependent upon inputs, three different input vectors were considered (Table 2): vec1 doesn't discharge any dynamic nodes, vec3 discharges all dynamic nodes, and vec2 discharges half and leaves half high.

Results

Figures 3 and 4 show the delay and active energy consumption for 180 nm and 70 nm processes respectively. The active energy of DVT is greater than that of LVT because the high-Vt keeper transistors must be sized up to give equal noise margin and equal precharge delay. For the same reason, the active energy of LB is greater than that of DVT. However, LB2 can meet the delay constraints with only a small increase in active energy over the LVT design because it uses only a small number of high-Vt transistors.

Figures 5 and 6 show the steady-state leakage power for 180 nm and 70 nm processes respectively. The leakage power of DVT is very sensitive to input values. For vec3 with clk=1, all the high-Vt transistors in DVT are turned off and the lowest leakage power is obtained. On the other hand, for vec1 with clk=1, the leakage

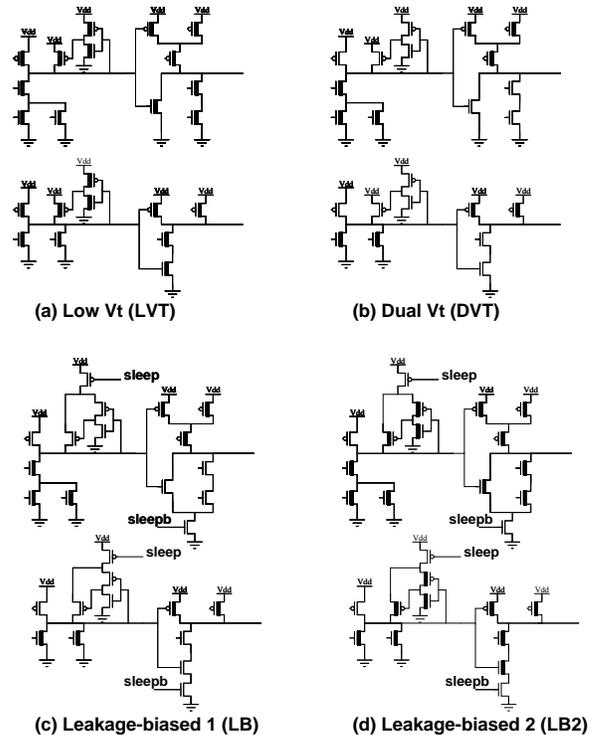


Figure 2: Cells for a 32-bit Han-Carlson adder. Low-Vt transistors are shaded.

is comparable to the LVT design. The sleep-state leakage power of LB and LB2 is independent of input vector because leakage currents bias the internal nodes into the lowest leakage state over some transition time. The LB schemes have worst-case sleep-state leakage currents that are around two decades lower than the LVT and DVT designs. For the 180 nm process, the LB scheme is preferred for circuits that spend enough time in sleep mode as it has lower leakage than LB2, but for circuits that are more active, LB2 has lower active energy and reasonable steady-state leakage. For the 70 nm process, LB2 is always better than LB since it has lower active energy and lower steady-state leakage than LB.

Figures 7 and 8 show how energy consumption evolves over time when the circuit is put into a sleep state for 180 nm and 70 nm processes respectively. The energy curves show the energy consumption when the circuit sleeps for the specified time, including the cost to transition the circuit into and out of the sleep state (e.g., the energy to switch the gates of the sleep transistors). For LVT and DVT schemes, the sleep energy is just linearly proportional to sleep time as leakage currents are constant. The sleep energy curve of LB shows a very different characteristic. There is a large jump in energy after a short sleep time (around 20 ns for 180 nm and around 1 ns for 70 nm). At this point, the static output of the first domino stage charges up to the threshold voltage, and causes the following stage to move rapidly to the low-leakage state. This process quickly ripples through the chain of domino gates. The energy stored in any precharged dynamic nodes is lost and must be restored during precharge when the circuit is next woken up, hence the steep rise in effective sleep energy dissipation. After this point, the energy curve has a very shallow slope due to the lowered leakage currents.

For short sleep times, the LB schemes require more total energy than simply idling an LVT or DVT circuit. But for longer sleep times the energy cost of discharging the internal dynamic nodes is amortized and the lower sleep leakage current yields lower overall energy. For LB and LB2, the cross-over point is around 2 μ s in the 180 nm process for the worst case (v_{ec1}). However, the cross-over point in the 70 nm process is under 10 ns because active energy scales down faster than leakage power.

Conclusion

As leakage currents become more significant, the leakage currents themselves can be used to bias nodes into low-leakage states. When used to dynamically deactivate critical path circuits in projected 70 nm process technologies, LB-Domino provides two decades reduction in steady-state leakage current compared with low-Vt or dual-Vt domino at equal delay and noise margin. LB-Domino has sub-cycle deactivation and reactivation latencies, and because leakage currents are used to bias the circuit, LB-Domino also has low transition energy overheads. Using LB-Domino to place circuits into a sleep state can yield net energy savings even for sleep times of under 10 ns. This makes dynamic fine-grain circuit deactivation practical, where small pieces of an active system can be powered-down for short periods of time to save leakage energy.

References

- [1] M. W. Allam, M. H. Anis, and M. I. Elmasry. High-speed dynamic logic styles for scaled-down CMOS and MTCMOS technologies. In *ISLPED*, pages 155–160, 2000.
- [2] Device Group at UC Berkeley. Predictive technology model. Technical report, PTM, 2001.
- [3] A. Chandrakasan, W. J. Bowhill, and F. Fox. *Design of High Performance Microprocessor Circuits*. IEEE Press, 2000.
- [4] V. De and S. Borkar. Technology and design challenges for low power and high performance. In *ISLPED*, pages 163–168, 1999.
- [5] A. Keshavarzi et al. Effectiveness of reverse body bias for leakage control in scaled dual Vt CMOS ICs. In *ISLPED*, pages 207–212, August 2001.
- [6] H. Makino et al. An auto-backgate-controlled MT-CMOS circuit. In *Symp. on VLSI Circuits*, pages 42–43, 1998.
- [7] J. Montanaro et al. A 160-MHz, 32-b, 0.5-W CMOS RISC microprocessor. *IEEE JSSC*, 31(11):1703–1714, November 1996.
- [8] K. Usami et al. Automated low-power technique exploiting multiple supply voltages applied to a media processor. *IEEE JSSC*, 33(3):463–471, March 1998.
- [9] L. Wei et al. Design and optimization of low voltage high performance dual threshold CMOS circuits. In *DAC*, pages 489–494, 1998.
- [10] M. Takahasi et al. A 60-mW MPEG4 video codec using clustered voltage scaling with variable supply-voltage scheme. *IEEE JSSC*, 33(11):1772–1778, November 1998.
- [11] S. Geissler et al. A low-power RISC microprocessor using dual PLLs in a 0.13 μ m SOI technology with copper interconnect and low-k BEOL dielectric. In *ISSCC*, February 2002.
- [12] S. K. Mathew et al. Sub-500-ps 64-b ALUs in 0.18 μ m SOI/bulk CMOS: Design and scaling trends. *IEEE JSSC*, 36(11):1636–1646, November 2001.
- [13] S. Mutoh et al. 1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS. *IEEE JSSC*, 30(8):847–854, August 1995.
- [14] S. Narendra et al. Scaling of stack effect and its application for leakage reduction. In *ISLPED*, pages 195–200, August 2001.
- [15] S. Shigematsu et al. A 1-V high-speed MTCMOS circuit scheme for power-down application circuits. *IEEE JSSC*, 32(6):861–869, June 1997.
- [16] S. V. Kosonocky et al. Enhanced multi-threshold (MTCMOS) circuits using variable well bias. In *ISLPED*, pages 165–169, August 2001.
- [17] T. Kuroda et al. A 0.9-V, 150-MHz, 10-mW, 4 mm², 2-D discrete cosine transform core processor with variable threshold-voltage (VT) scheme. *IEEE JSSC*, 31(11):1770–1779, November 1996.
- [18] T. Kuroda et al. Variable supply-voltage scheme for low-power high-speed CMOS digital design. *IEEE JSSC*, 33(3):454–462, March 1998.
- [19] T. McPherson et al. 760 MHz G6 S/390 microprocessor exploiting multiple Vt and copper interconnects. In *ISSCC*, pages 96–97, 2000.
- [20] W. Lee et al. A 1-V programmable DSP for wireless communications. *IEEE JSSC*, 32(11):1766–1776, November 1997.
- [21] J. P. Halter and F. Najm. A gate-level leakage power reduction method for ultra-low-power CMOS circuits. In *CICC*, pages 457–478, 1997.
- [22] T. Inukai. Boosted Gate MOS (BG MOS): Device/circuit cooperation scheme to achieve leakage-free giga-scale integration. In *CICC*, pages 409–412, 2000.
- [23] J. T. Kao and A. P. Chandrakasan. Dual-threshold voltage techniques for low-power digital circuits. *IEEE JSSC*, 35(7):1009–1018, July 2000.
- [24] Y. Ye, S. Borkar, and V. De. A technique for standby leakage reduction in high-performance circuits. In *Symp. on VLSI Circuits*, pages 40–41, 1998.

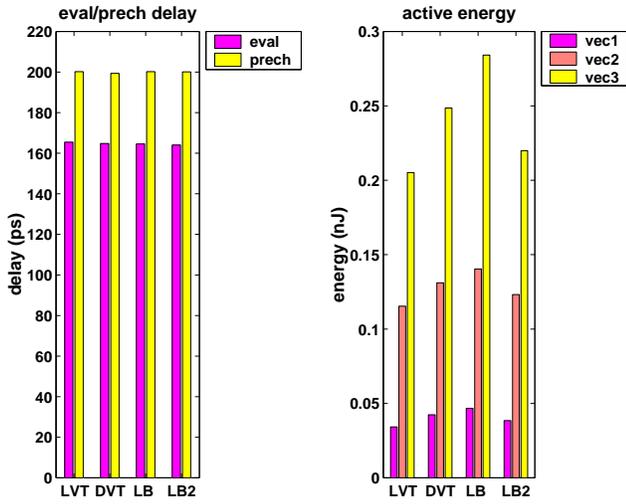


Figure 3: Delay and active energy consumption : 180 nm process.

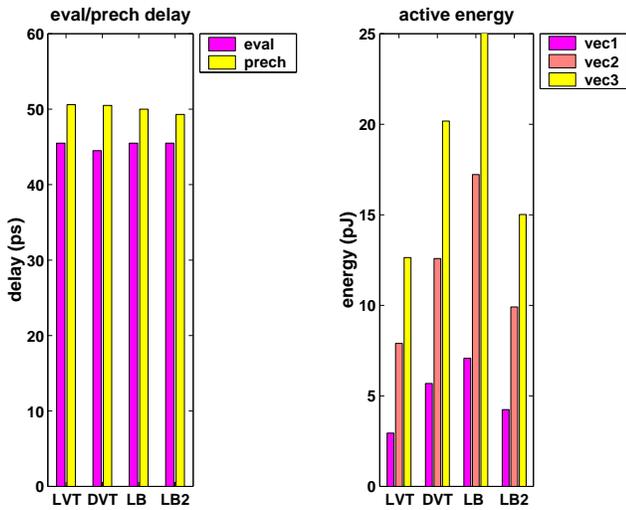


Figure 4: Delay and active energy consumption : 70 nm process.

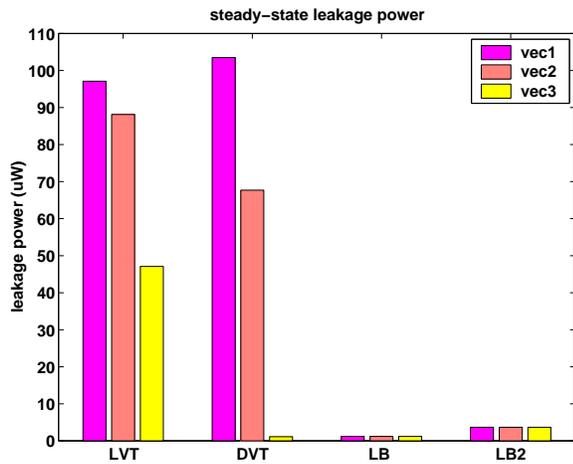


Figure 5: Steady-state leakage power : 180 nm process. `clk` is high for all and `sleep` is asserted for LB and LB2.

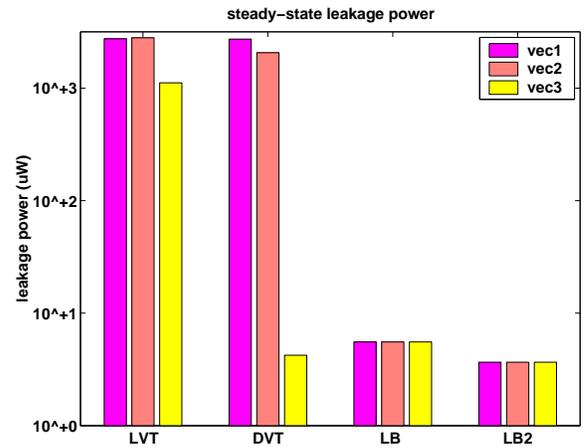


Figure 6: Steady-state leakage power : 70 nm process. `clk` is high for all and `sleep` is asserted for LB and LB2. Note that y-axis is log-scale.

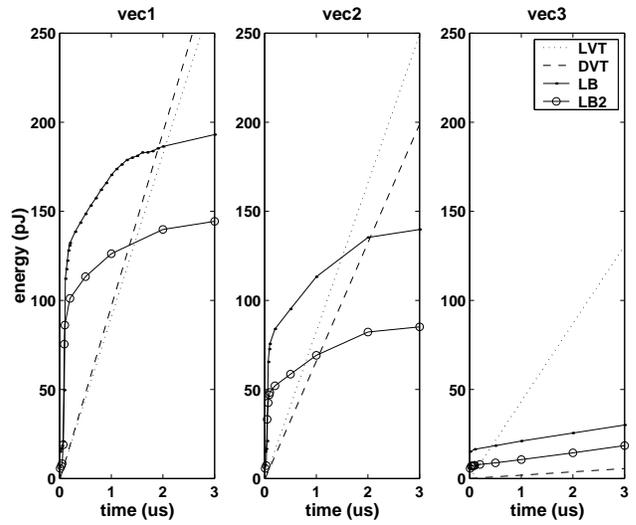


Figure 7: Cumulative sleep energy : 180 nm process.

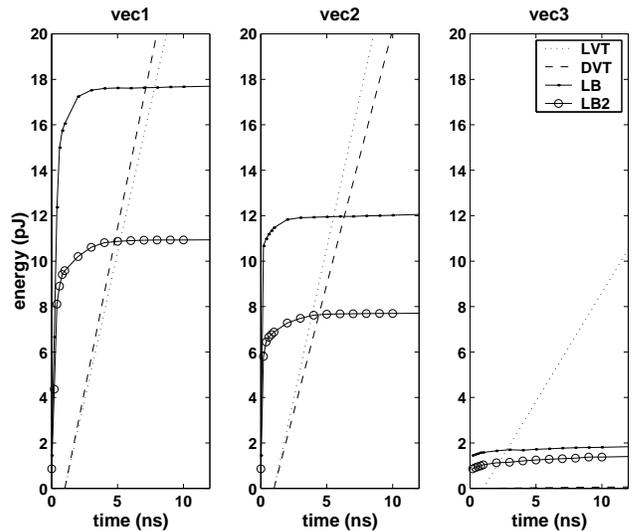


Figure 8: Cumulative sleep energy : 70 nm process.