A Low-Leakage Dynamic Multi-Ported Register File in 0.13µm CMOS

Atila Alvandpour, Ram Krishnamurthy, K. Soumyanath, and Shekhar Borkar Microprocessor Research Labs, Intel Corporation, Hillsboro, OR 97124, U.S.A

Abstract

Increasing leakage currents combined with reduced noise margins are seriously degrading the robustness of dynamic circuits. This paper describes a dynamic implementation of a 256X32b 4-read/write-port Register-File for ~6GHz operation at 1.2V in a 0.13 μ m technology. The pre-charged local bit-lines utilize an efficient conditional keeper-technique, where a large fraction of the keeper is turned *ON* only if the dynamic output remains *High* in the evaluation phase. Using this technique, we are able to improve upon all-low-Vt performance by 4%, while maintaining Dual-Vt usage. Thus, the robustness is improved by 96% and the active leakage power is reduced by 5X.

1. Introduction

High fan-in compact dynamic gates are often employed in performance-critical units of microprocessors and other high performance VLSI circuits. The use of wide dynamic gates is strongly impacted by reduced noise margin and increasing leakage currents in sub-0.13µm low-Vt devices. Traditionally, dynamic floating nodes have been avoided by employing a static path trough a pull-up and/or pull-down device referred to as "keeper". For small leakage currents, weak keepers were sufficient to maintain the voltage level of pre-charged nodes without a significant impact on the performance of the dynamic gates. However, in the presence of increasingly larger leakage currents the keepers must be sized to compensate for the leakage currents. This significantly degrades the performance of dynamic circuits. Fig.1 shows an example of a K-bit wide dynamic gate, a K-bit wide MUX, with the standard keeper PK_0 . To ensure correct operation during the evaluation phase (clock *High*), two worst-case conditions must be fulfilled:

1-Worst-Case noise (where dynamic output remains *High*): Vss (Low) + a DC noise on gates of M_{11} , M_{12} , ..., M_{1K} , and Vcc (High) on the gates of M_{21} , M_{22} , ..., M_{2K} . 2- Worst-Case delay (High-to-Low transition): Vss on all the pull-down transistors, except the stack transistors M_{11} , M_{21} , which are turned on to pull-down the output node.

Thus the task for the keepers is not only to compensate for "Ioff"-leakage currents but also higher sub-threshold currents due to the potential worst-case noise on the inputs of all the pull-down devices. The same circuit (Fig. 1) effectively shows the local read-path of conventional Register-Files, where read-select signals (M_{11} , M_{12} , ..., M_{1K}) select one of the storage cells (through M_{21} , M_{22} , ..., M_{2K}) to be evaluated. Multi-port Register-Files are performance-critical processor components with single-cycle read/write latency and high throughput requirement. A large number of read-select entries per port, enforces the use of wide dynamic MUX-structures. The elevated Ioff in sub-0.13 μ m technologies necessitates alternative dynamic techniques to achieve low read path delays, while simultaneously meeting robustness requirements.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*ISLPED*⁷⁰¹, August 6-7, 2001, Huntington Beach, California, USA. Copyrigth 2001 ACM 1-58113-371-5/01/0008...\$5.00.



Fig.1: A K-bit wide dynamic gate (MUX) with a standard keeper.

This paper describes design of a low-leakage, and robust 256X32b 4-read/wirte-port Register-File for ~6GHz operation at 1.2V in a 0.13µm technology. Utilizing an efficient Conditional *Keeper* technique, CKP [1, 2], allows a Dual-Vt Register-File to operate with 4% higher performance than that of an all-Low-Vt Register-File, while the noise-margin or robustness is improved by 96%, and the active leakage power is reduced by ~5X. This technique is a high-density low-leakage alternative to pseudo static techniques for higher leakage tolerance [3]. The paper is organized as follows: In Sec. 2, we describe the Multi-Port Register-File and its organization; particularly the local read operation on which the CKP technique has been utilized. In Sec. 3, the concept of CKP technique is reviewed. Simulation results and comparisons are presented in Sec. 4, followed by conclusions in Sec. 5.

2. A 256X32b 4-read/write ported Register File

A 256x32b register-file in 1.2V, 0.13µm dual-Vt CMOS technology with copper interconnect [4] is described for 6GHz operation. Single-ended read-select and bit-line signaling is used to reduce wiring congestion, enabling 4-read, 4-write port capability in a dense layout occupying 356µmx89µm.





Fig. 2 shows organization of the register file. 8-bit read/write address per port is decoded in previous cycle and fed as read/write select signals into the 256x32b array. Each local bit-line, LBL, (1 per port) supports single-ended read on 16 cells with two-way merge via a static NAND.

Fig. 3 shows the register-file bit-cell, with symmetric loading of 2 read ports on each side of storage cell for optimal cell stability [5]. Matched pass transistors on each side of the storage cell enable single-ended write. Fig. 4 shows the local read organization. Two access transistors per word (M1 and M2) read the data in the storage cells, forming a dynamic 16-way MUX on the local read path.



Fig. 3: The Register file multi-ported bit-cell



Fig. 4: The local read-path per port

Fig.5 shows, the 8-way global bit-line (GBL) MUX, which selects the NAND outputs to deliver a 32b word per read-port. Fig. 6 shows the read/write 2Φ domino-timing plan with fully time-borrowable Φ 2 boundary, where both the LBL, and GBL read operations are sequentially performed within a clock cycle.



Fig. 5: The register file global read-path



Fig. 6: The timing-scheme of the read operation

LBL and GBL dynamic MUXes are susceptible to noise due to the increasingly large active leakage and potential input noises during the evaluation phase, when pre-charged dynamic nodes should stay high. LBL is particularly more sensitive than GBL due to a smaller stored charge (0.1x compared to GBL) and a wider dynamic MUX structure (16-way for LBL vs. 8-way for GBL). The goal of this work has been to increase the robustness of the LBL operation without any considerable impact on power consumption and performance of the Register File.

As was discussed in Sec. 1, larger standard keepers can increase the robustness of the read path. However, this results in significant contention, which degrades the performance and increases the power consumption. In the next section, we describe a technique that results in a significant robustness improvement and leakage reduction without any performance/area penalty or any modification of the standard bit-cell topology.

3. Review of Conditional Keeper Technique, CKP

Fig.7 shows one topology of the CKP technique [1, and 2]. It employs two keepers: A fixed keeper *PK1*, and a conditional keeper *PK2*. At the onset of the evaluation phase (Clock *Low-to High*), *PK1* is the only active keeper. After a delay-time, $T_{\text{keeper}}=T_{\text{Delay element}} + T_{\text{NAND}}$, the keeper *PK2* is activated only if the dynamic output is still *High* (Fig. 8).



Fig. 7: A 16-bit MUX for LBL read operation utilizing the CKP.

Knowing the worst-case time for a potential output High-to-Low transition, the highest performance can be achieved when PK2 is activated close to or later than the worst-case clock-to-output transition, T_{MAX}. The fixed keeper, PK1 (Fig.7) ensures sufficient robustness during the T_{keeper} , which can be a small fraction of the clock phase. Depending on the required robustness of the actual gate during the T_{keeper} , different sizecombinations of PK1 and PK2 can be used. Compared to the standard keeper (PK_0) , the conditional keeper-technique can meet higher robustness at comparable performance, where $W(PK1) \sim W(PK_0)$, and the additional PK2 is activated conditionally with negligible impact on performance (W is the width of the devices at fixed length). To meet higher performance at comparable robustness, the keepers can be sized $W(PK_0) =$ such that W(PKI)+W(PK2),where $W(PK1) < W(PK_0)$.



Fig. 8: Different states of the keepers in Fig.8 during the evaluation.

Another optional but important advantage of the keeper-circuit in Fig.7 is that an inversion of the input signal to *PK2* provides a domino-compatible dual-output. This, when needed, can save a significant amount of area and power consumption, as a singlerail wide gate offers the same function of its dual-rail counterpart.

CKP technique [1] is an enhanced version of [2] verified at the worst-case I_{OFF} corner of a 0.13µm technology. A special case of the general concept in [2], has been later published in [6], where the standard keeper was removed. In the next section we show that the standard keeper is required to maintain robustness during the time (2-3 inversion delays) required to activate the conditional keeper.

4. Simulation and Comparisons

We have replaced the standard keepers on the local bit-lines (Fig. 4) with the CKP technique (Fig.7). As was described in the previous section, by employing the CKP technique, higher performance or higher robustness can be achieved depending on the total strength of the keepers, *PK1*, and *PK2* and their gain ratio $\lambda = W(PK1)/W(PK2)$.

4.1 Robustness Analysis

The robustness analysis for the 16-bit MUX with the CKP technique should be considered for two different cases, during two different time slots:

- 1- During the time, $T_{\text{keeper}}=T_{\text{Delay element}} + T_{\text{NAND}}$, where the *PK1* is the only active keeper.
- 2- After T_{keeper} , when both *PK1*, and *PK2* are active (when the dynamic output should remain High).

<u>Case 1</u>:

If the total strength of the keepers is equal to that of the standard keeper in the conventional design, $W(PK_0) = W(PKI) + W(PK2)$, then $W(PKI) < W(PK_0)$. Thus, the DC noise margin is reduced during the short time, T_{keeper} . However, the impact of noise on circuits depends on the magnitude of the noise as well as the time at which the noise is applied. To compare the noise tolerance of the proposed technique, worst-case leakage and input DC noise condition was provided. DC robustness has been evaluated as unity-gain noise margin, UGNM, where the DC-noise level at the output of the static NAND is equal to the DC noise level at the inputs of the dynamic MUX. A DC noise higher than the UGN level would be amplified at the output of the NAND gate, and thus, such noise level should be avoided.

To analyze the robustness of the LBL operation with CKP technique, we first find the reference UGN-level by applying a variable DC noise at the inputs of the conventional 16-bit wide MUX with the standard keeper PK_0 , which is sized such that it meets the target robustness (UGN-level) at or above the target noise-floor. The UGN reference-level is later applied on the MUX with the CKP technique. The following criterion has been used to avoid robustness degradation during the T_{keeper} , where the keeper PKI is active only: Applying the reference UNG-level on the inputs (worst-case condition), PKI is sized such that during the T_{keeper} , the noise level at the output of the NAND gate does not exceed the reference UNG-level (Fig. 9).



Fig.9: Simulation waveforms at worst case leakage corner and an applied DC-noise on the inputs of the dynamic gates. The keeper PK_1 meets the target robustness during T_{keeper} , as the noise on the output of the NAND gate (Fig. 8) does not exceed the final output noise of the NAND gate following the dynamic circuit with the standard keeper (Fig. 5)

Case2:

After T_{keeper} , both PK1, and PK2 are conditionally ON and the DC noise robustness of the MUX with CKP is equal to that of the conventional MUX with standard keeper *PK*0, as $W(PK_0) = W(PKI) + W(PK2)$. Following the above criterion, and at $W(PK_0) = W(PKI) + W(PK2)$, the robustness of CKP technique is fairly comparable to the conventional technique. This robustness-criterion has been followed for all the comparisons between the CKP and the standard keeper technique in the next section.

4.2 Simulation Results

The 0.13µm technology offers two threshold voltages for each device. We have simulated the local read operation of the Register-File, where performance, robustness, and energy/transition of all-low-Vt, (LVT) and dual-Vt (DVT) 16-bit conventional MUXes (STD) have been compared to those utilizing the CKP technique. In the dual-Vt case, the read-select pass-transistors (M1 in Fig.4 and 7) as well as the keepers are high-Vt devices, while the rest of the devices are low-Vt.

Simulations are performed at worst-case Ioff corners of the 0.13µm technology, at Vcc=1.2V, 110°C. In order to verify the sensitivity of the delay and robustness of the CKP technique to any potential clock-to-data race, the $T_{\text{keeper}}=T_{\text{Delay element}} + T_{\text{NAND}}$ was swept over a wide range at a fixed worst-case clock-to-output transition, T_{MAX} . The X-axis in Fig.10-12 shows the ratio $T_{\text{keeper}}/T_{\text{MAX}}$.

4.2.1 All-Low-VT Design

Fig.10 shows the bit-line evaluation delay for the 16-bit LVT MUXes. For the conventional MUX (STD), the standard keeper is sized such that UGN-level is marginally above the specified noise-floor. This gives the fastest bit-line evaluation at the target robustness. At this point, the CKP-MUX utilized a keeper-ratio $\lambda = W(PK1)/W(PK2) = 0.25/0.75$ which was sufficient to meet the ISO robustness criterion at as large $T_{\text{keeper}}/T_{\text{MAX}}$ as 1.25. The figure (10) shows that the CKP technique results in up to 19% higher performance at a comparable level of robustness. Further, the performance improvement is relatively insensitive to clockto-data race. The figure suggests that the best performance improvement can be achieved by activating the conditional keeper close to or after a potential worst-case output transition, with the optimum design point at $T_{\text{keeper}}/T_{\text{MAX}} = 1.25$ which result in maximum performance without violating our described robustness criterion. **Optimum Design point**



Fig.10: CKP performance improvements as a function of Tkeeper.

The additional power consumption due to the circuit overhead and larger clocked load is efficiently compensated by the reduced contention during the output High-to-Low transition (Fig 11). This resulted in an energy/transition comparable to that of the conventional case.

Comparison parameters (Normalized to standard all-LVT)	Low VT Standard	Low VT CKP	Dual VT Standard	Dual VT CKP
Bit-line evaluation delay (only)	1	0.81	1.12	0.93
Total local read operation delay	1	0.90	1.05	0.96
Robustness	1	1	1.96	1.96
Contention	1	0.24	0.95	0.17
Leakage power	1	1	0.21	0.21

Table 1: Comparison summary



Fig.11: Contention power consumption normalized with that of the conventional technique with the standard keeper.

4.2.2 Low-Vt / Dual-Vt Simulation results

The previous all low-Vt simulations were performed at lowest acceptable noise margin to achieve the highest performance, where CKP technique resulted in 19% faster bit-line evaluation. There are two main techniques to increase the robustness to a higher level: 1-Upsizing the keepers for the all-Low-Vt conventional and CKP-based LBL MUXes. 2-Utilizing the High-Vt devices. Fig. 12 shows the delay-robustness trade-off for both cases. The figure shows that "keeper-upsizing" is an inefficient technique for increasing the robustness as it results in significant delay penalties. Still, the CKP all-low-Vt MUX maintains its relative performance benefit. However, Fig.12 shows that utilizing the HVT devices is a much more efficient way to increase the robustness.

Comparing the low-VT result with the dual-VT result (the two point) we show clearly that large delay-robustness trade-offs are involved for relatively small performance improvements.

Since the CKP technique is independent of Vt-levels, for the dual-Vt case, it results also in 17% less bit-line evaluation time. The interesting observation is that this performance improvement is about the same as the all-low-Vt performance improvement (compared to the dual-Vt design). Another important consequence of the use of dual-Vt scheme is that the leakage power consumption is also significantly reduced.

Table 1, summarizes the performance/robustness comparisons, normalized to the standard all-low-Vt design. Where, the dual-Vt LBL MUX with CKP technique results in ~2X higher robustness, and 5X less active leakage power consumption at a comparable level of performance (4% faster). The all-low-Vt LBL MUX with CKP results in 19% faster MUX operation, and 10% faster total LBL read operation.

The total LBL performance improvement is partially screened by the driver delay, and the merge-delay, which are fixed delaytimes for all the circuit alternatives. For highest performance the CKP technique enables a total local read delay of 86ps. This allows the Multi-ported Register File to operate at 5.8GHz clock.



Fig.12: The evaluation -delay of the 16-bit LBL MUXes vs. the DCnoise robustness for different keeper sizes and different Vt assignments

5. Conclusions

In this paper we have described a dynamic implementation of a 256X32b 4-read/write-port Register-File, for ~6GHz operation at 1.2V in a 0.13 μ m technology. The pre-charged local bit-lines utilize an efficient conditional keeper-technique, CKP, where a large fraction of the keeper is turned *ON* only if the dynamic output remains *High* in the evaluation phase. Using this technique, we are able to improve the dual-Vt-based circuit-performance upon all-low-Vt-based one by 4%, while reducing the active leakage currents by 5X, and increasing the noise-robustness by ~2X. Alternatively, up to 19% higher performance at comparable robustness has been observed for all-low-Vt-based wide MUXes utilizing CKP technique.

References

[1] Atila Alvandpour, Ram Krishnamurthy, K. Soumyanath, Shekhar Borkar, "A Conditional Keeper Technique for Sub-0.13µm Wide Dynamic Gates", 2001 Symposium on VLSI Circuits, June 12-16, Kyoto, Japan.

[2] Atila Alvandpour, Per-Larsson Edefors, and Christer Svensson," A Leakage Tolerant Multi-Phase Keeper for Wide Domino Circuit", IEEE Intl Conf. on Electronics, Circuits, and System Sept. 5-8, 1999 pp: 209 – 212.

[3] Ram Krishnamurthy, Atila Alvandpour, Ganesh Balamurugan, Naresh Shanbhag, K. Soumyanath, Shekhar Borkar," A 0.13µm 6GHZ 256X32b Leakage-tolerant Register File", 2001 Symposium on VLSI Circuits, June 12-16, Kyoto, Japan.

[4] S.Tyagi et al, 2000 IEDM Tech. Digest, pp. 567-570.

[5] M.Golden et al, 1999 VLSI Circuits Symposia, Digest, pp. 105-108.
[6] Allam, M.W. Anis, M.H. Elmasry, M.I. "High-speed dynamic logic styles for scaled-down CMOS and MTCMOS technologies", Low Power Electronics and Design, July 26-27 2000, pp. 155-160.