

# Standby Power Management for a 0.18 $\mu$ m Microprocessor

Lawrence T. Clark  
Intel Corporation  
5000 W Chandler Blvd  
Chandler, AZ, 85226-3699  
(480) 554-1458  
lawrence.clark@intel.com

Neil Deutscher  
Intel Corporation  
5000 W Chandler Blvd  
Chandler, AZ, 85226-3699  
(480) 552-2357  
neil.f.deutscher@intel.com

Shay Demmons  
Intel Corporation  
5000 W Chandler Blvd  
Chandler, AZ, 85226-3699  
(480) 552-3151  
shay.p.demmons@intel.com

Franco Ricci  
Intel Corporation  
5000 W Chandler Blvd  
Chandler, AZ, 85226-3699  
(480) 554-3611  
franco.ricci@intel.com

## ABSTRACT

Static power dissipation is a concern for battery powered hand-held devices since it can substantially impact the battery life. Here, the use of reverse body bias to limit  $I_{off}$  on the high performance, low power XScale™ microprocessor core is described. The scheme utilized is amenable to implementation on a low-cost (non-triple well) process and has limited regulation requirements. The regulation requirements and circuits are described, as is the performance of the method. A measured current reduction factor of over 25 is achieved with this method of reverse body bias. Implications of the use of body bias leakage control for active power and performance, as well as system level implications are also discussed.

**Categories & Subject Descriptors:** B.7.1 [Integrated Circuits]: Types and Design Styles – Advanced Technologies, Microprocessors and microcomputers

**General Terms:** Measurement, Performance, Design

**Keywords:** Low power, microprocessors, body effect

## 1. INTRODUCTION

Drain to source leakage current ( $I_{off}$ ) has been increasing with each sub-micron process generation, where source to drain leakage must be traded against active performance. It is a natural byproduct of transistor scaling since, to maintain performance,  $V_t$  must be lowered to compensate for scaling in  $V_{dd}$ , and shorter  $L_c$  produces increasing drain induced barrier lowering (DIBL). The need to scale  $V_t$  is particularly acute for processes supporting hand-held devices

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, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ISLPED'02, August 12-14, 2002, Monterey, California, USA  
Copyright 2002 ACM 1-58113-475-4/02/0008...\$5.00

since it is the primary device characteristic controlling low voltage performance. From a process development perspective, this requires careful optimization of the  $V_t$  selected to allow sufficiently low standby power balanced with active power dissipation for maximum battery life. This optimization is dependent on the expected system usage model, i.e., it strongly depends on the time the system spends in standby vs. active operation. While  $I_{off}$  varies exponentially to 8-12x the room temperature value at 125C, the focus here is on meeting the stringent hand-held requirements at room temperature. This is due to “hand-held” implying near room temperature ambient (at least most of the time) since the systems must be in human contact. Therefore device battery life is primarily specified there.

The market for hand-held battery powered devices requires standby currents in the 10's to 100's of  $\mu$ A. For IC's intended for hand-held products this leads to a leakage current component under 100pA/ $\mu$ m [1] driven even lower by high integration system on a chip (SOC) designs. That 1nA/ $\mu$ m is excessive for a battery-powered device is apparent, since 7 meters of gate width (approximately that in the microprocessor core described here) contributes over 3mA at room temperature. Reducing this leakage to 100  $\mu$ A requires  $V_t$ 's over 500mV, independent of the supply voltage chosen. So-called non-state retentive sleep modes, eg. MTCMOS approaches [2] are interesting in that they promise greatly reduced standby power. Unfortunately, these approaches also incur substantial penalties: Since they do not retain state, the present condition must be stored away before sleep and restored upon resuming active operation. Consequently, a low standby power storage medium must be provided. Additionally, the data movement requires substantial time and if the storage is off-chip, a potentially large power penalty that must be amortized by the static power savings achieved in the time in sleep. This can preclude frequent use, eg. between keystrokes. It also tends to imply that the software and OS must be aware of the modes, since some indication of the expected duration of the sleep mode is required for effective utilization.

A secondary, but important leakage effect is gate induced drain leakage (GIDL) at the gate-drain edge. It is most prevalent in the NMOS transistors, being about two orders of magnitude weaker for PMOS devices. For a gate having a 0V bias with the drain at  $V_{dd}$ ,

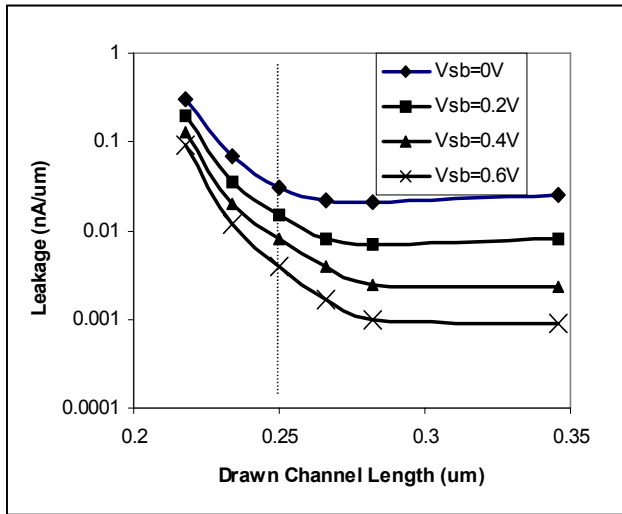
significant band bending occurs in the drain region, allowing electron-hole pair creation mechanisms. Essentially, the gate voltage attempts to invert the drain region, but since the holes are rapidly swept out, a deep depletion condition occurs [3]. Reverse body bias exacerbates this and GIDL is consequently the limiting mechanism to increasing back-bias for  $I_{off}$  control, i.e., eventually most of the current contribution is due to GIDL and increases faster with body bias than the  $I_{off}$  is decreased. The onset of this mechanism can be lessened by limiting the drain to gate voltage.

Diode area components from both the source-drain diodes and the well diodes are generally negligible with respect to  $I_{off}$  and GIDL components despite having an approximately 1000x/100C temperature coefficient.

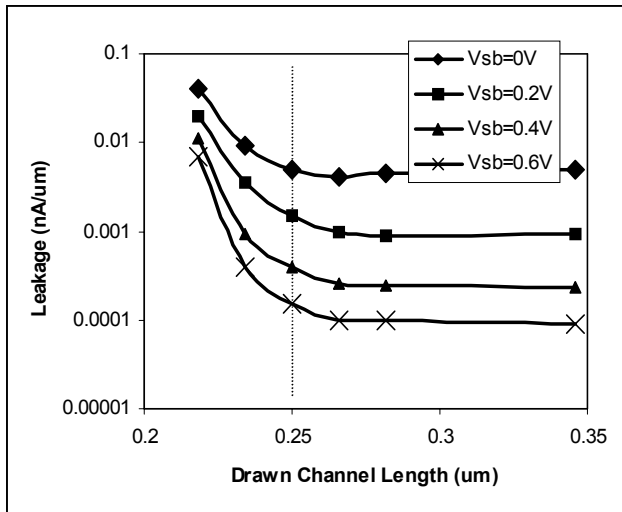
## 2. $I_{off}$ CONTROL VIA BACK BIAS

One technique which has been applied successfully, c.f. [4] is the application of reverse body bias for leakage control. This has a number of benefits: 1. It is a circuit design approach and therefore requires no changes to the process; 2. It does not affect the active performance [5] or alternatively, it allows improved active power at the same standby current level vs. a device not so equipped (see section 6); 3. It allows state retentive power-down, i.e., data is not lost. This latter consideration is important in that it allows the leakage control to be transparent to the operating system and application software. Body bias has been successfully applied to limit leakage on a 1.8V microprocessor implemented in a dual well 0.25 $\mu$ m process [6]. Here, the substrate is biased negatively by use of a charge pump to apply body bias to NMOS transistors while the PMOS bulk is switched to the 3.3V IO voltage. A large number of local switches are distributed across the device to apply the body bias and provide a low impedance bulk connection, at the expense of routing the controls and supplies throughout the layout. High (>1V) body bias is provided with a  $V_{dd}-V_{ss}$  of 1.8V.

Figure 1 shows transistor-level leakage characteristics as a function of back-bias and drawn channel length for 0.25 $\mu$ m NMOS and PMOS devices. Drain voltage is 1.0V, which is representative of a conservative  $V_{dd}$  in standby mode. The vertical lines in the plots show the nominal drawn channel length for this process. The leakage reduction with reverse body bias is lower at shorter channel lengths due to increased DIBL. Therefore, a combination of body bias and modest channel length increase results in the lowest leakage. Note that the longer channel implies a performance penalty. Finally, NMOS leakage is substantially higher than PMOS leakage. This effect is exaggerated somewhat on the particular data shown, but demonstrates that body bias may have differing value for NMOS and PMOS transistors.



(a)



(b)

Figure 1. Measured  $I_{off}$  vs.  $V_{sb}$  and  $L_e$  for (a) NMOS and (b) PMOS transistors on 0.25 $\mu$ m process.

## 3. LEAKAGE CONTROL CIRCUIT CONFIGURATION

The microprocessor is built using an Intel 0.18 $\mu$ m process technology supporting dual gate oxides and improved metal pitches over those used on the desktop microprocessors as described in [7]. The core supports operating voltages from 0.75V to 1.5V, with the lower values targeted at providing good performance (greater than 150MHz) at less than 50mW total power dissipation [7] at 25C. The circuit configuration used comprises Figure 2. Power is applied to the  $V_{dd}$  and  $V_{ss(gnd)}$  pins. Power supply IR drop is avoided by having only one power carrying supply gated by a series transistor. Large N channel devices M5 provide  $V_{ss}$  to the active circuitry during normal operation. Similarly large P channel devices M7 provide clamping of the N well ( $V_{ddsup}$ ) to  $V_{dd}$ . These transistors must have an oxide thickness suitable to exposure to high voltages as indicated. The NMOS clamp transistors are less than 2% of the total transistor width of the microprocessor, while being over 10 times the drive of the PMOS clamping transistors. Adequate decoupling capacitance is provided to avoid a large instantaneous voltage drop across the N transistors in active operation.

Given a choice of applying power through either N or P devices, NMOS was chosen for larger gate overdrive available at low voltages and their larger inherent current drive. The PMOS (N-well

clamp) devices can be much smaller as they provide only the diode, N-well leakage, and AC coupling current components. The microprocessor is capable of exceeding 800MHz with power provided through the large NMOS devices. The area penalty to support the reverse body bias powerdown mode is minimal, as the power supply clamping transistors are provided in the pad ring only, occupying otherwise empty space within the supply pins. Additionally, the bulk connections are routed sparsely throughout the circuitry, to limit the density impact. This is possible due to the low currents that need to be supplied.

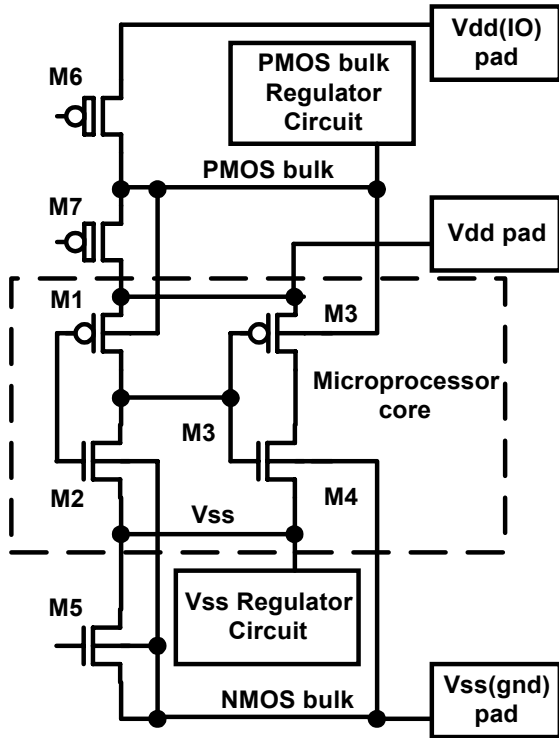


Figure 2. Circuit to apply reverse body bias.

For non-state retentive “sleep” as well as in reverse body biased “standby” modes, transistors comprising M5 in Fig. 2 are in cutoff. In the former case, the core  $V_{ss}$  is allowed to float towards  $V_{dd}$ , and power consumption is limited to the leakage current through the NMOS clamp devices. Obviously, since body bias is not applied to the clamps, it is desirable to have a high  $V_t$  for the clamp devices to optimize this operating characteristic. Generally, the higher  $V_t$  of the thick gate devices intended for IO is adequate, while the entire IO voltage applied to the gate allows the best current drive. The active current of the PLL greatly exceeds the 100uA target standby current described above, necessitating that the PLL be disabled while in standby. Leaving standby mode requires the PLL to restart and lock. Since this takes less than 20 $\mu$ s, the mode is usable often, e.g., between keystrokes.

To apply body bias to the NMOS devices,  $V_{ss}$  is raised towards  $V_{dd}$  while regulating  $V_{ss}$  to avoid losing state from too little  $V_{dd}-V_{ss}$  voltage. The limits to how much  $V_{dd}-V_{ss}$  may be collapsed are discussed in a following section. Raising the NMOS source voltage rather than decreasing the NMOS body voltage has a number of advantages: Firstly, it eliminates the requirement of a twin tub or triple well process. This also eliminates the need for charge pump

circuitry. Secondly, some decrease in  $I_{off}$  can be derived from limiting  $V_{ds}$  (which raises the  $V_t$  corresponding to the DIBL component of  $V_t$ ) as well as limiting GIDL. Lastly, since gate current is strongly affected by the gate to source voltage, it can afford substantial decrease in that component on processes with thin oxides. The voltage regulator to provide  $V_{ss}$  current to the IC core is shown in Fig. 3. Also shown in the figure, another regulator provides a high voltage to the PMOS bulk (N-well) node, to reverse body bias the PMOS transistors. The  $V_{ss}$  regulator output voltage must have a constant output voltage over at least 3 decades of current demand. Specifically, when entering the low power state directly from high frequency operation the die may be hot. While it may cool to ambient in a few package thermal time constants, insufficient  $V_{dd}-V_{ss}$  voltage during this time would cause loss of state. Consequently, the regulator must be capable of providing the entire IC leakage component at high temperatures, a value that may exceed 30mA for the microprocessor considered here at 110C at some process corners. Stability must be ensured at all conditions and overshoot on  $V_{ss}$  must be limited, particularly when entering the body bias state, which is essentially a voltage step on  $V_{ss}$ .

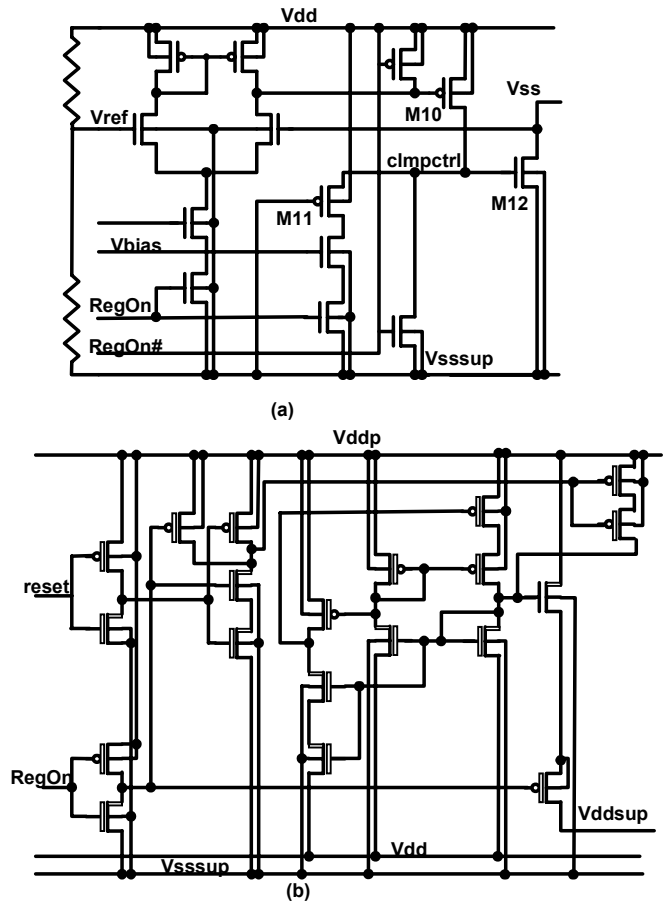


Figure 3. Leakage control circuit. (a)  $V_{ss}$  regulator (b) N-well ( $V_{ddsup}$ ) regulator

#### 4. REGULATOR DESIGN

Details of the  $V_{ss}$  regulator are illustrated in Fig. 2. The amplifier compares the voltage on  $V_{ss}$  with a reference voltage. The reference

is generated by a resistor stack, which allows it to vary with power supply variations. In this manner, higher supplies allow larger body bias. The resistor stack current is under 0.1uA and thus may be left continuously connected in all modes. The regulator is a three-stage amplifier with an NMOS output transistor M12. Three stages were required due to the bias conditions required and the very low current requirements to keep the regulator power consumption a negligible (at typical process less than 5%) contributor to the total standby power.

Since the transistor can only source current,  $V_{ss}$  rises passively, driven by the core leakage. Power would not be saved by actively raising  $V_{ss}$  and consequently, no power is wasted on circuitry to drive  $V_{ss}$  high. In the event that the standby powerdown mode is prematurely terminated, e.g., to service an interrupt before  $V_{ss}$  reaches equilibrium, the power required to charge and discharge  $V_{ss}$  is saved. The output transistor is sized to provide the full IC leakage current at high temperature and the worst-case process corner. The first stage is a differential OTA, while the second buffer stage provides increased voltage output range and current drive to the gate of M12. The first and second stages combined use less than 4uA at typical operating conditions. At such low current levels, gain is limited, which improves stability, as discussed below. Slew rate also suffers, which makes the step response poor. To address this, the buffer stage devices, includes the diode connected transistor M11 which, combined with the sizes chosen, delivers the DC characteristic shown in Fig. 4. This keeps transistor M12 from completely cutting off and thereby limits the required slew rate of the buffer stage driving the large capacitive load presented by M12.

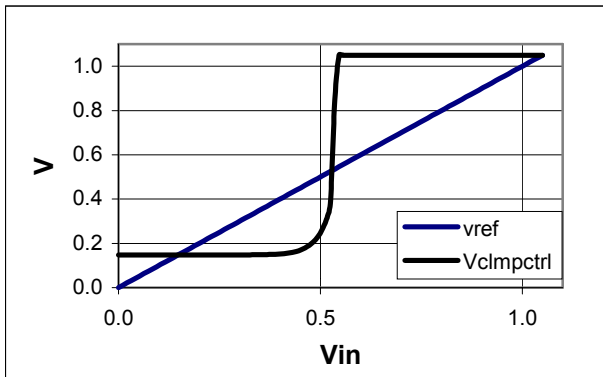


Figure 4. Voltage swing limiting on node clmpctrl.

To back-bias the PMOS devices, two schemes may be used. At low IO voltages, e.g., 1.8V, the PMOS transistor bulk node may be directly connected to this voltage (See M6 in Fig. 2). For IC's supporting higher IO voltages, this is not desirable, as the incremental leakage reduction does not offset the greater charge switched in raising the well voltage. The design thus utilizes a regulator as shown in the figure, which derives a constant voltage from the IO supply  $V_{dd(IO)}$ . It is worth noting that as long as circuit configurations that accumulate the gates of the PMOS transistors are avoided, high voltages may be applied to the bulk without oxide stress or damage. The regulator utilizes a bootstrapped voltage reference driving a wide NMOS vertical drain transistor in a source follower configuration as shown in Fig. 3. This transistor can provide the well and diode leakage current while operating in the subthreshold region of operation, which results in a negligible voltage drop from the reference voltage to  $V_{dssup}$  in operation. The

vertical drain configuration allows the thin gate oxide device to tolerate high drain to gate voltages as in [9].

Since the entire circuit, amplifier and the IC leakage must comprise a stable system, adequate phase margin must be maintained. This is particularly important since overshoot at the  $V_{ss}$  node, even momentarily, may cause loss of state. Essentially, the IC leakage presents the load current for output transistor M12 with the AC equivalent circuit a rather complicated function of the body transconductance presented by the core NMOS transistors, the decoupling capacitance due to  $V_{ss}$  capacitance and intentional decoupling devices, and an "auxiliary" RC due to the on NMOS devices and the gates on the far side of those devices, as well as the amplifier nodes. Fortunately, the circuit poles may be approximated by the dominant terms, greatly simplifying the analysis. For the  $V_{ss}$  node, this comprises just the output conductance of transistor M12, while the amplifier pole is dominant. The former pole is at approximately 670kHz calculated from the small signal parameters while the latter is at 9kHz. The simulated Bode plot for the entire  $V_{ss}$  system is shown in Fig. 5. The dominant poles are shown as calculated from the transistor  $g_{ds}$  as well as the capacitance of the nodes. The low gain of the amplifier is advantageous, leading to a low unity gain bandwidth and greater than 60 degrees of phase margin at the typical process. The highly capacitive nature of the  $V_{ss}$  node ensures a low-pass characteristic that does not require high amplifier speed to stabilize.

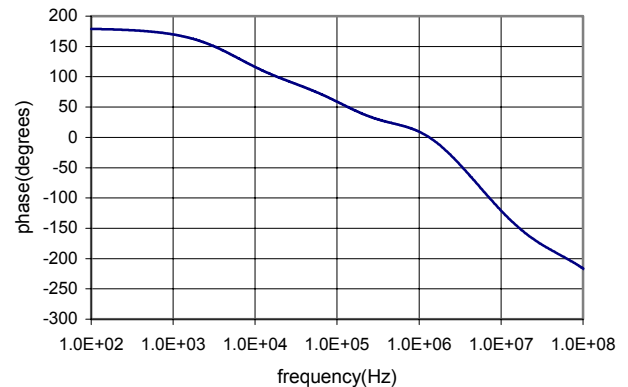
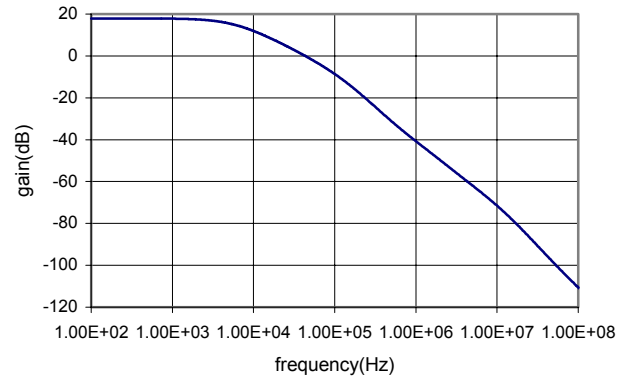


Figure 5. Bode Plot of  $V_{ss}$  Regulator.

## 5. LIMITING CIRCUIT OPERATION

The reverse body bias standby mode is designed to retain the state of the microprocessor so all memory, such as latches, need to be able to hold a '0' or a '1'. The application of a reverse body bias is by increasing the N-well voltage ( $V_{\text{ddsup}}$  in Fig. 3) and  $V_{\text{ss}}$  as shown in Fig. 2. This will reduce the  $I_{\text{off}}$  leakage current by increasing  $V_t$  of the transistors and by decreasing the source to drain voltage. As  $V_{\text{dd}}$  and  $V_{\text{ss}}$  collapse together eventually the transistors in saturation are pushed into subthreshold, as the reverse body bias increases  $V_t$  and the increase in  $V_{\text{ss}}$  decreases  $V_{\text{gs}}$ . In subthreshold these 'on' transistors rapidly weaken following the subthreshold slope.

In a memory element the voltage level of a node is maintained by an 'on' transistor being able to supply enough current to overcome the leakage of all the attached 'off' transistors. In normal operation this is not a problem due to the large  $I_{\text{on}}$  to  $I_{\text{off}}$  ratio. As transistors reach subthreshold, the current drops rapidly with  $V_{\text{ss}}$  due to

$$I_{\text{ds,sat}} = 0.5 \frac{\mu C_{\text{ox}} Z}{L} (V_{\text{gs}} - V_t)^2,$$

becoming

$$I_{\text{ds,sub}} \propto \exp\left(\frac{-V_t}{S/\ln 10}\right),$$

as the gate overdrive ( $V_{\text{gs}} - V_t$ ) is reduced below 0, where  $S$  is the subthreshold swing parameter. Ideally,  $V_{\text{dd}} - V_{\text{ss}}$  can be lowered to drive all of the transistors into subthreshold operation, since the  $I_{\text{on}}/I_{\text{off}}$  ratio will scale for all transistors. Essentially, assuming a 70-80mV/decade transistor subthreshold characteristic, over three decades of current difference between on and off transistors will be maintained with 250mV of  $V_{\text{ds}}$ . Consequently, the limit is determined by the worst case N to P width ratio, typically occurring in domino nodes, where large ratios nearing 100 may occur. Nodes that may lose state without affecting machine operation may be considered non-critical. State loss depends upon many factors such as the type of latch, the transistor ratios, the logic state being held, the local transistor  $V_t$ 's and the temperature. Criticality is also determined by the function and potential redundancies in the logic. To determine stability,  $V_t$  variation in  $V_t$ , as well as the ratio of P to N width are the primary considerations.

This can be illustrated in the fail point as a function of the PMOS body voltage and  $V_{\text{ss}}$  (corresponding to the NMOS source-body voltage) shown in Fig. 6. Points higher on the vertical axis have higher PMOS  $V_t$  and further right have higher NMOS  $V_t$ . Measured parts retained state to the left of the curve and lost state to the right of the curve after application of that level of reverse body bias. The lower portion of the curve is where the PMOS transistors have limited reverse body bias applied to them, so their  $I_{\text{off}}$  leakage has not been reduced very much. The nmos transistors have substantial reverse body bias applied to them, so "on" devices are in subthreshold. This lower curve represents a memory element trying to hold a '0' being flipped to a '1'. As  $V_{\text{ddsup}}$  increases the PMOS transistors leakage is reduced, so that the amount of reverse body bias that can be applied to the nmos transistors can be increased, continuing until a maximum value of  $V_{\text{ddsup}}$  and  $V_{\text{ss}}$  is reached.

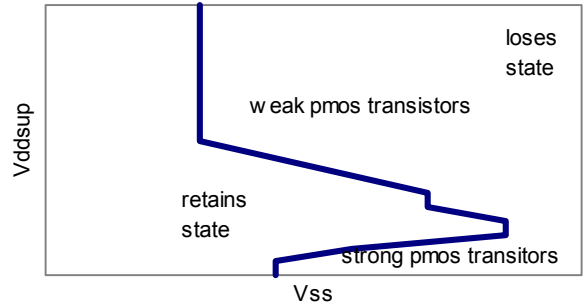


Figure 6. Non-retaining failure points of  $V_{\text{ddsup}}$  and  $V_{\text{ss}}$

The top part of the curve represents the converse case where the PMOS transistors are weakened with respect to the NMOS. With a large  $V_{\text{ddsup}}$  applied 'on' devices are in subthreshold and are eventually unable to supply enough current to overcome leakage from nmos transistors. This top part of the curve represents a memory element holding a '1' flipping to a '0'.

## 6. Results

The simulated behavior of the  $V_{\text{ss}}$  regulator is shown in Fig. 7. When the leakage current from the microprocessor is low the voltage on  $V_{\text{ss}}$  will not rise to the reference voltage since the regulator doesn't actively drive its output. When the leakage current is large enough the regulator clamps at the reference voltage, about 0.73V. Measured results with  $V_{\text{dd}}$  at 1.2V closely match the simulated values. This voltage level is low enough that memory elements in the part will not lose state but significant reduction in  $I_{\text{off}}$  is achieved.

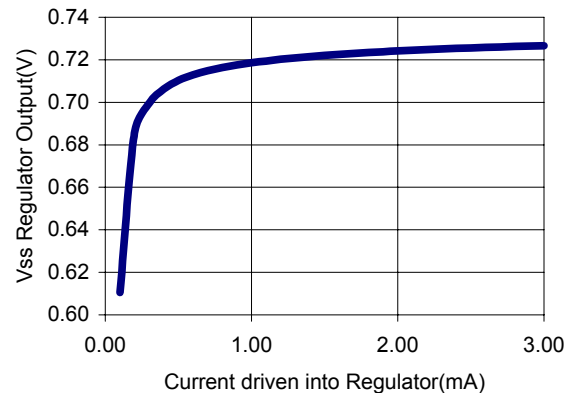
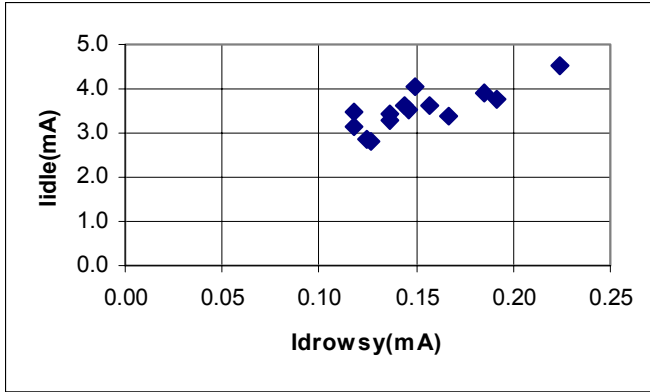


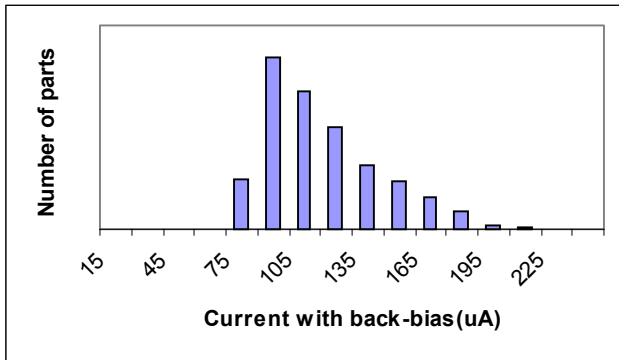
Figure 7. Behavior of  $V_{\text{ss}}$  Regulator.

There is a substantial reduction factor in the leakage current when the body bias is applied. Figure 8 shows the XScale microprocessor™ current in the Idle mode (no body bias applied) compared to the Drowsy mode (body bias applied). The results show that  $I_{\text{off}}$  is reduced by an average factor of 25.



**Figure 8. Standby current of the microprocessor with and without body bias.**

The effect of this reduction factor on the active power of the IC can be estimated by considering the  $V_i$  increase required to match the  $I_{off}$  reduction and the requisite  $V_{dd}$  increase to achieve the same performance. This was obtained by simulating a circuit metric calibrated to the measured frequency voltage performance of the microprocessor. A  $V_i$  increase of 110mV (to a typical value of 500mV) results in the same reduction. At this value of  $V_p$  the same frequency at  $V_{dd}=0.75V$  is obtained by an increase to 0.86V, demonstrating an active power savings of 24%. Figure 9 shows the distribution of the current with reverse body bias applied across one wafer.



**Figure 9. Standby current of the microprocessor with body bias.**

## 7. Summary

Technology scaling has increased the primary subthreshold source to drain leakage component to make  $I_{off}$  reduction by circuit means attractive. The efficacy of applying back bias for the purpose of limiting standby power consumption for battery powered integrated circuits has been shown. The scheme implemented controls static leakage while maintaining performance and active power. Circuits used to implement reverse body bias control that operate over a wide range of conditions have been described and power reduction effects at the transistor, test-chip and whole-chip levels shown.

## 8. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Kim Wagner and Bruce Fishbein as well as the entire Xscale design team.

## 9. REFERENCES

- [1] S. Thompson, IEDM '99 tutorial.
- [2] S. Mutoh, et. Al., IEEE J. of Solid-State Circuits, 30 (1995) 847.
- [3] S. Wolf, Silicon Processing for the VLSI Era: Volume 3 – The Submicron MOSFET, 1995
- [4] H. Mizuno, et al., J. Solid-State Circuits, 34 (1999) 1492..
- [5] S. Thompson, et al., Proc. 1997 Symp. VLSI Tech. Symp., June 10-12 1997, 69.
- [6] K. Ishibashi, 2001 VLSI Circuit Symp. Short Course.
- [7] M. Bohr, et al., Proc. IEDM 1998, 197.
- [8] L. Clark, et al., JSSCC, 36 (2001) 1599.
- [9] L. Clark, Proc. 1999 VLSI Circuit Symp., 61.