Reduction of Instantaneous Power by Ripple Scan Clocking

Kirti Joshi Eric MacDonald
The University of Texas at El Paso
Department of Electrical and Computer Engineering
El Paso, TX 79968
Email: kmjoshi@utep.edu, emac@utep.edu

Abstract

The exponential increase in the number of transistors implemented in integrated circuits in each new generation of CMOS technology is causing an explosion not only in functional power consumption but in test power consumption as well. Although most research has focused mainly on reducing average power or the total energy consumed during test, instantaneous power consumption is also increasing and posing a serious threat for the ability of a chip to be tested in a manufacturing test floor - or worse in field testing using built-in-self-test (BIST) where battery-powered applications lack the supply voltage robustness of automated test equipment (ATE). In this paper, a flip-flop design is proposed that is the cornerstone of a novel scan clocking architecture inspired by the need to reduce instantaneous power during scan.

1. Introduction

The increasing levels of integration on circuits in contemporary CMOS technology is allowing for unprecedented levels of functionality to be included on die at an exponential rate. To effectively test the ever-increasing number of transistors requires the use of test techniques like scan testing which provide access and visibility to the millions of internal nodes in the integrated circuit. The necessary state of the sequential elements required to exercise hard-to-test nodes can be easily introduced into the integrated circuit by scanning test vectors in through scan chains which are activated in a special test mode. However, as the state is scanned into the chip, many more flip-flops can simultaneously change state than is normally possible in functional mode where a large fraction of flip-flops are idle in any given cycle. This large increase in switching activity can result in large power consumption as nodes are charged / discharged resulting in either overheating or supply voltage noise – either of which can result in incorrect device operation. Managing the power at test has been the intense focus of significant research for the past decade – mainly focusing on average power or total energy consumed. Most techniques have been reported that either minimize the clock tree activity by implementing gating or minimize the data transitions by appropriate test vector selection or by fixing output voltages of flip-flops during scan. However these techniques can result in extended test times or the need for additional test vectors to obtain comparable fault coverage.

In this paper a novel approach is presented that reduces the instantaneous power by spreading clock edges throughout the scan clock cycle. Consequently, the clock and combinational logic transitions during scan are not coincident and power consumption is evenly distributed throughout the cycle. This does not significantly affect the total energy consumed but does reduce the power at any point in time and thereby avoids power supply voltage compression. Furthermore, the technique avoids the possibility of hold time violations based on a daisy-chain configuration of the clock that follows the reverse order of the scan data. Scan chains are in danger of hold times violations as the data path in many cases contain little logical delay and - if clock skew is present - can result in a data skipping a victim flip-flop and thus corrupting the scan data. By eliminating hold time violations, this architecture presents the possibility of operating scan chains at the lowest theoretical limit of V_{dd} which is the sum of the NFET and PFET threshold voltages (in contemporary CMOS technologies is between 600 and 700 mV). At this low voltage, power (both average and instantaneous) is reduced quadratically with regards to the supply voltage in addition to the further reduction in instantaneous power provided by the non-coincident clock edges of this architecture. Running traditional scan chains at this low voltage is often not possible as the clock trees are

Proceedings of the 23rd IEEE VLSI Test Symposium (VTS'05)
1093-0167/05 $ 20.00 **IEEE**

balanced at the functional operating voltage. At low voltage the discrepancy between the exponential increase in device delay versus the relatively insensitive wire (RC) delay leads to significant clock skew between different branches of the clock tree. This is due to difference in delay scaling through each branch depending on the fraction of delay that is device and RC driven. This variation of delay through the clock tree therefore increases the probability of hold time violations by introducing clock skew.

In section 2, the previous work relating to reducing average power and instantaneous power is discussed. Section 3 describes the proposed ripple scan clocking architecture which is compared to traditional scan testing and presents a novel scan element capable of implementing the architecture. Section 4 explains the experimental evaluation of the proposed ideas which are based on transistor level simulations and finally, section 5 provides concluding remarks.

2. Previous Work

Many papers have been published regarding reducing power in the general sense during scan with an aim to maintain the viability of scan testing - where power consumption is generally much higher than during functional mode operation [1-11]. This increase in power consumption can result in overheating the device during test or cause incorrect operation due to power supply compression (i.e. internal ground bounce and / or voltage supply droop). Power supply compression results in increases in delay through logic paths due to the temporary reduction in V_{dd} levels. Furthermore, with the wide adoption of Built-In-Self-Test (BIST) – where scan chains are controlled and monitored by an internal test state machine allowing for logical testing in the field - the reduction of scan power consumption is even more important as it may affect the useful life of battery-powered electronics.

However few reports have focused specifically on reducing instantaneous power - the main culprit of power supply compression - rather than power in the general sense. In [12], a clever concept is introduced where the duty cycle of the clock is adjusted for flip-flops included in a given scan chain. Each scan chain has potentially a different delay adjustment so that the flip-flops in each chain are provided a rising edge on the clock at nominally the same time and clock skew is minimized along the chain. This avoids hold time violations among the flip-flops of any given chain and scan chains (shift registers in the general sense) are particularly susceptible to hold violations as the data paths generally have little logical delay. Additionally, the clock for each chain is purposely skewed to avoid the simultaneous switching of all flip-flops on the die and thus the instantaneous power is spread through out the full clock cycle. Average power and the total energy consumed during scan testing remains the same, but the supply voltage compression is dramatically reduced, thus avoiding incorrect device operation during test.

One drawback of this design is the complexity of the clocking required to provide this intentional skew between scan chains while minimizing skew along each scan chain. A second concern is the potential impact on skew during functional operation. Neither concern is insurmountable; however both require special attention during the integration and clock synthesis stages of the chip development. The proposed approach described in this paper avoids these difficulties while simultaneously reducing instantaneous power and furthermore outright eliminates the possibility of hold times during scan shifting. Consequently, it makes possible scanning at ultra-low voltages where clock skew can become a major challenge and where power (both average and instantaneous) are reduced by a quadratic factor.

3. Proposed Scan Cell and Clocking Architecture

3.1. Proposed Scan Clocking Architecture – Ripple Scan Clocking

The essence of the new architecture is a scan clock that is introduced at the end of the chain and is daisy chained in the reverse flow of the data to the beginning of the chain. In each scan element, a rising edge on the clock results in the capture of the data at the scan input and the launch of that data to the scan output. Subsequently the scan clock is presented with a small delay to the preceding scan element that supplied the data to the current flip-flop. Consequently, hold time violations are eliminated as good clock skew is intentionally introduced. This is repeated until the clock reaches the first scan element of the scan chain.

For comparison purposes, Fig. 1 illustrates traditional muxed scan architecture. In this configuration each flip-flop has a functional input and output (*DI* and *Q,* respectively) as well as clock. To enable scan, two additional inputs are required: a scan input *SI* and a scan enable *SE.* When the scan enable is active, the scan element uses the scan input as the source of data to be captured and launched at each rising edge of the clock. As the scan chains are connected in a shift register configuration, this allows

for the straight forward introduction of new state into the flip-flops to facilitate testing the combinational logic as well as the scanning out of the results of the test. When the scan enable is de-asserted, the scan elements operate in functional mode as a simple D flip-flop. Alternatively, Level Sensitive Scan Design (LSSD) can be used as well.

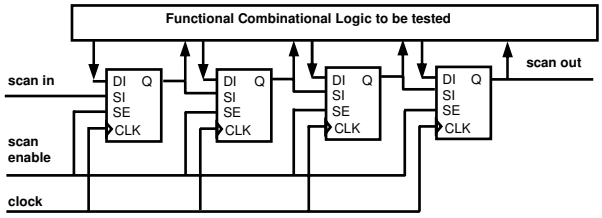

Figure 1. Traditional Scan Architecture

The proposed architecture is portrayed in Fig. 2 and illustrates how each flip-flop now has the scan enable input replaced by a scan clock input. A new output is introduced that is simply a buffered version (slightly delayed) of the scan clock input. By daisy chaining the scan clock from one flip-flop to the next – each stage slightly delayed from the previous stage – no two flip-flops along the chain will see a coincident rising clock edge and thus instantaneous power is evenly distributed throughout the duration of the clock rippling backwards through the chain. Furthermore, the possibility of a clock launching the next data bit from any flip-flop before the subsequent stage has captured the current cycle data (hold time violation) is eliminated.

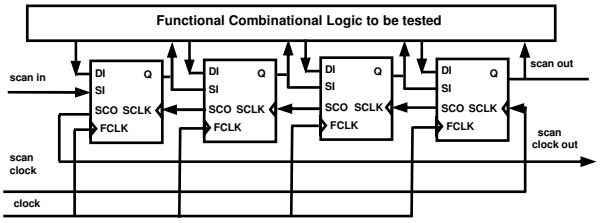

Figure 2. Proposed Ripple Scan Clocking Architecture

In scan mode, the functional clock is held in the inactive state by asserting and holding a logical true value and the scan clock controls the capture and launch function of the scan element with each rising edge presented to the scan clock input. Conversely, in functional mode the scan clock is held in the inactive state (logical true) and surrenders control of the scan element to the functional clock. This provides the capability to first scan in a test vector through the scan chain via scan clock cycles and then implements the test of the combinational logic by holding the scan clock inactive while providing a single rising edge on the functional clock. The results of the test are then contained in the chain and the scan clock can be used to unload the results through the scan out port while simultaneously scanning in the next test vector.

Routing complexity is approximately equivalent between the two architectures as the scan enable signal which is routed through a buffer tree in a traditional scan implementation is replaced by the daisy-chaining of the scan clock in the new architecture. Any area increase in the new scan element due to the buffering of the scan clock between the scan in to scan out ports is compensated for by the elimination of the need for the scan enable buffer tree in the traditional architecture as well as hold time violation fix buffers for the scan chains. Otherwise all other signals are equivalent in terms of wire length and required area.

Although hold time violations are eliminated in this architecture, one trade-off is the limit of the maximum scan frequency that the chains can be scanned - which directly affects test time. The minimum period must be longer than the duration that the clock requires to propagate from the end of the chain to the beginning and ideally would be just less than the targeted scan frequency to spread the clock edges throughout the period. The total propagation delay can be calculated as shown in Equation 1

$$T_{sc} = N \bullet D_{se} + \Sigma T_{wire} + T_{su} \qquad \text{Eqn 1}$$

where N is the number of scan elements in the chain, Dse is the delay of the clock from the scan clock input to the scan clock output in each scan element, $Twire$ represents the individual wire delays between each scan element and Tsu is the time required for the tester or BIST engine to provide the next data value. The main component to this equation is the first product of the scan chain length and the delay through each scan element. The scan element delay can be varied based on the buffer sizing to some degree for fine tuning, however the total propagation delay is dominated by the scan chain length, N. Possible solutions to this include implementing pass through flip-flops and buffering the clock between every n-1 and n^{th} flip-flop. Additionally, the techniques described in [8] could also improve test durations.

One other interesting aspect of this approach is that if the last scan clock output is made visible to the ATE or BIST, the signal can be monitored as a self-timed done signal for that scan cycle – indicating that the next bit can now be scanned in. In a traditional

sequential architecture, the scan clock frequency is determined during the development phase by analyzing the slowest scan path for the worst-case manufacturing process. Alternatively, in the proposed architecture the next scan bit can be scanned in as soon as the clock completes the full ripple if the scan clock output is accessible through a primary output. Thus, faster devices can benefit from this by automatically increasing the scan frequency without fear of setup time violations. Exploiting this self-timed aspect of ripple clock scan can eliminate the wasted time between clock cycles due to the often unnecessary design margin introduced with traditional sequential circuit operation.

3.2. Proposed Scan Element Capable of Implementing Ripple Scan Clocking

The traditional scan element design consists of a simple multiplexer followed by a Master and a Slave D–Latch. The cell has four input signals – Scan In *(SI)*, Data In *(DI)*, Scan Enable *(SE),* Clock *(CLK)* and a single output signal *Q*. Figure 3 illustrates the design of a traditional scan element. In the scan mode, the *SE* signal is pulled high and the scan data is sampled at the rising edge of the clock and launched at the output *Q*. In functional mode, the *SE* signal is pulled low, data is sampled from *DI* at the rising edge of the clock and is launched at the output *Q*.

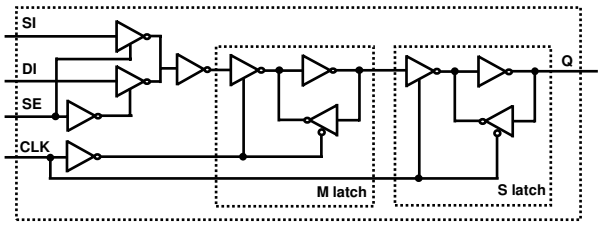

Figure 3. Traditional Scan Element Design

The proposed ripple scan cell design consists of the same number of input signals as discussed in the traditional scan element. However, the scan enable signal is now replaced with a Scan Clock *(SCLK)* input signal and an additional Scan Clock Out *(SCO)* output signal is required. The *SCO* signal is just a delayed version of the original scan clock and may serve as a delayed scan clock to another scan cell which may be connected in a daisy-chain fashion. Figure 4 illustrates the design of the proposed ripple scan element design.

In the scan mode, the functional clock *(FCLK)* is statically pulled high and the scan clock controls the operation of the flip-flop. When the scan clock transitions from high to low, the M-latch becomes transparent to the scanned in data. Subsequently, at the rising edge of the scan clock, the S-latch becomes transparent and the data is launched at the output *Q*. In the functional mode, the scan clock is statically pulled high and the functional clock governs the operation of the flip-flop. The *DI* input signal reaches the output only at the rising edge of the functional clock.

The transistor count is roughly equal between the two cells with the primary difference being the internal buffers required to delay the scan clock. However, in the traditional flip-flop a scan enable buffer tree – which is not required for the ripple scan element – would need to be included in the traditional design.

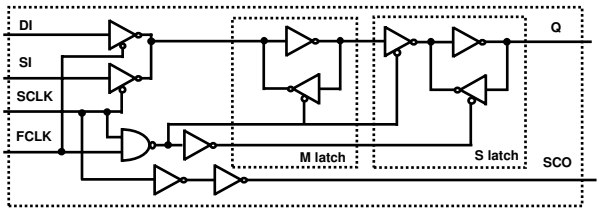

Figure 4. Proposed Ripple Scan Element Design

3.3. Ultra-Low Voltage Application of Ripple Scan Clocking

Although ripple scan clocking can be used to reduce instantaneous power and thus minimize the noise on the power supply, it also enables the possibility of running scan chains at ultra low voltage. This not only improves instantaneous power but average power as well with both forms decreased by a quadratic factor. By reducing and maintaining a low voltage during scan and the associated logic test, Ultra-Low Voltage Testing (ULVT) techniques can be implemented as described in [13-16]. ULVT has been shown to provide greater sensitivity to specific defects that are difficult to detect and these techniques have never been reported in conjunction with BIST.

Running scan chains at ultra low voltages is possible with traditional scan chain architectures but requires special attention to the clock balancing which becomes more difficult across a larger voltage supply range – particularly when the range is taken to the extreme. As a result of eliminating hold time violations with the ripple scan architecture, this concern is fundamentally eliminated. One challenge with lowering the voltage is the overhead of generating the lower voltage. However, in low power CMOS design, the trend of dynamic voltage scaling has already been established [17-18] and thus enables the intriguing possibility of ultra-low voltage testing in the field.

Table 1. Simulation Results

ISCAS89 Benchmark Circuits	S27	S298	S344	S713	S444	S838	S1423	S5378
Scan Chain Length	3	14	15	19	21	32	74	179
Instantaneous Power (mW) for Traditional Scan (Simulated @ 3.3V)	6.38	21	22.7	31.1	33.8	52.8	122	287
Instantaneous Power (mW) for Ripple Scan (Simulated @ 3.3 V)	2.27	5.68	6.61	17.8	4.81	24.9	17.5	36.5
Ratio of Power	2.8	3.7	2.3	1.7	7	2.1	7	7.9
Instantaneous Power (mW) for Ultra-Low Voltage Ripple Scan (Calculated @ 1.0V)	0.208	0.522	0.607	1.635	0.442	2.287	1.607	3.352

4. Simulation Results

To experimentally validate the effectiveness of ripple scan clocking, benchmark circuits were simulated using Cadence Spectre with CMOS models from TSMC's 0.25 micron technology at 3.3 V. To compare the proposed and traditional scanning methods, eight ISCAS89 benchmark circuits represented as columns in Table 1 were synthesized with scan chains of both styles and simulated. Simulations included scanning in vectors of alternating 1's and 0's – a pattern which would result in a reasonable level of transitions during scan. Only eight of the smaller benchmarks were simulated as larger transistor-level simulations were not possible due to disk limitations.

As illustrated in Figs. 5 and 6, the maximum instantaneous power derived from the depicted supply current is reduced for the ripple scan case. The even distribution of power is evident in the relatively flat current profile of the ripple scan vs. the large spikes shown in the traditional case. Depending on the circuit, the improvement ranged between 1.7 and 7.9 as shown in Table 1. In general, a trend exists in which increases in scan chain length result in improvements in the reduction of maximum instantaneous power. This is explained by the fact that the proposed technique only allows one flop to change state at a time as opposed to the traditional approach which is accumulative. The reason that the trend is not seen more clearly is likely because of the staggered presentation of the scan vector along the chain. As the ripple clock traverses the chain, the logic is presented with many more intermediate vectors which would result in additional hazard-related power consumption and this component would be heavily data and logic dependent.

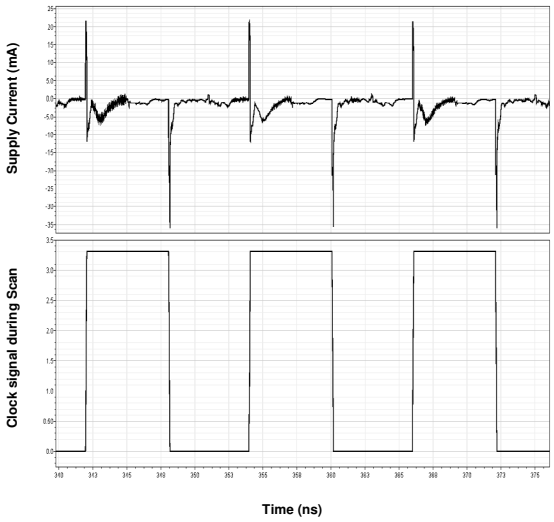

Figure 5. Scan Test Waveforms for Traditional Scan Architecture

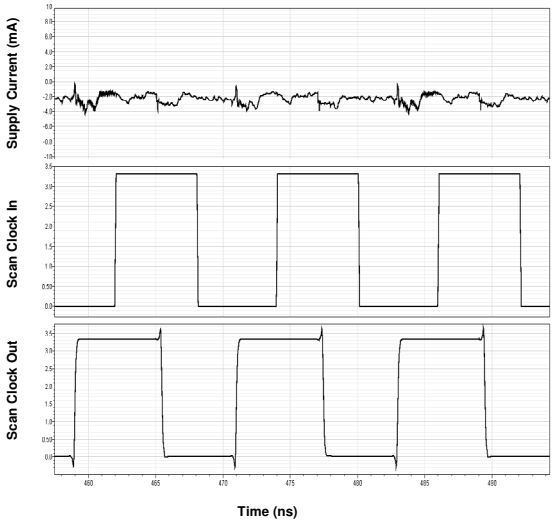

Figure 6. Scan Test Waveforms for Proposed Ripple Scan Architecture

5. Conclusion

Reduction in the average power and total energy consumption during test has been the major focus of research in the past. Although these parameters are important, very little attention has been given to the reduction of the instantaneous power which may result in power supply compression and hence incorrect logical operation. The paper has proposed a new scan cell architecture which reduces the instantaneous power during scan by distributing the logical activity throughout the scan clock cycle and at the same time eliminates any possibility of hold time violations along the scan chains. This further encourages the use of scanning at ultra low voltage levels which effectively reduces both the average and instantaneous power by a quadratic factor. The proposed design is simple and has the same number of signals as compared to the traditional scan design.

Acknowledgement –

The authors would like to thank Rob Aitken for input on the timing implications of the proposed architecture.

6. References

[1] E. Larson, Z. Peng "Test Scheduling and Scan Chain Division Under Power Constraint", *Asian Test Symposium*, pp. 259-264, Nov. 2001.

[2] R. Chou, K. Saluja, V. Agrawal, "Scheduling Test for VLSI Systems under Power Constraint", *Transactions on Very Large Scale Integration Systems*, Vol. 5, No. 2, pp. 175-185, June 1997.

[3] E. Saxena, K. Butler, L. Whetsel, "An Analysis of Power Reduction Techniques in Scan Testing*", International Test Conference*, pp. 670-677, Oct., 2001.

[4] T. Huang, K Lee, "A Token Scan Architecture for Low Power Testing", *International Test Conference*, pp. 660-669, 2001.

[5] Y. Bonhomme, P. Girard, L. Guiller, C. Landraut, S. Pravossoudovitch, "A Gated Clock Scheme for Low Power Scan Testing of Logic ICs or Embedded Cores", *Asian Test Symposium,* pp. 253-258, Nov. 2001.

[6] N. Nicolici, M. Al-hashimi, "Scan Latch Partitioning into Multiple Scan Chains for Power Minimization in Full-Scan Sequential Circuits", *Design, Automation, and Test in Europe (DATE) Conference*, pp. 715-722, Mar. 2000.

[7] R. Sankaralingam, B. Pouya, N. Touba, "Reducing Power Dissipation during Test using Scan Chain Disable", *VLSI Test Symposium*, pp. 319-324, April, 2001.

[8] L. Whetsel, "Adapting Scan Architectures for Low Power Operation", *International Test Conference*, pp. 863-872, Oct. 2000.

[9] O. Sinaoglu, I. Bayraktaroglu, A. Orailoglu, "Scan Power Reduction Through Test Data Transition Frequency Analysis", *International Test Conference*, pp. 844-850, 2002.

[10] S. Wang, "Generation for Low Power Dissipation and High Fault Coverage Patterns for Scan-Based BIST", *International Test Conference*, pp. 834-843, 2002.

[11] M. Santos, I. Teixeria, J. Teixeria, S. Manich, R. Rodriquez, J. Figueras, "RTL Level Preparation of High-Quality / Low Energy / Low Power BIST", *International Test Conference*, pp. 814-823, 2002.

[12] T. Yoshida, M. Watari, "MD-SCAN Method for Low Power Scan Testing*", Asian Test Symposium,* pp. 80-85, Nov. 2002.

[13] J. Chang, E. McCluskey, "Quantitative Analysis of Very-Low-Voltage Testing", *VLSI Test Symposium*, pp. 332-337, 1996.

[14] J. Chang, E. McCluskey, "Detecting Delay Faults by Very-Low-Voltage Testing", *International Test Conference*, pp. 367-376, 1996.

[15] H. Hoa, E. McCluskey, "Very-Low Voltage Testing for Weak CMOS Logic ICs", *International Test Conference*, pp. 275-284, 1993.

[16] E. MacDonald, N. Touba, "Very Low Voltage Testing of SOI Integrated Circuits", *VLSI Test Symposium*, pp.25-30, 2002.

[17] K. Nowka, G. Carpenter, E. MacDonald, H. Ngo, B. Brock, T. Nguyen, "A 32-bit PowerPC System-on-a-Chip with Support for Dynamic Voltage and Frequency Scaling,", *IEEE Journal of Solid-State Circuits,* Vol. 37, No.11, pp 1441-1447, Nov 2002.

[18] Clark, L.T.; Hoffman, E.J.; Miller, J.; Biyani, M.; Luyun Liao; Strazdus, S.; Morrow, M.; Velarde, K.E.; Yarch,M.A." An Embedded 32-b Microprocessor Core for Low-Power and High-Performance Applications". Solid-State Circuits, IEEE Journal of, Volume: 36, Issue: 11, Nov. 2001 Pages: 1599 – 1608.