

A SIMULATION STUDY TO QUANTIFY THE ADVANTAGES OF SILICON-ON-INSULATOR (SOI) TECHNOLOGY FOR LOW POWER

David Donaghy¹, Linda Brackenbury², Steve Hall¹

INTRODUCTION

In this work, we combine the advantages of silicon-on-insulator (SOI) technology with asynchronous design techniques to assess the overall benefits in reducing power. A 16-bit self-timed adder is employed as the demonstrator circuit. At the technology level, fully-depleted SOI offers advantages compared with bulk CMOS. These arise due to lower subthreshold slope, lower vertical electric field hence enhanced channel mobility and reduced junction capacitance. The superior off-state current of SOI enables the threshold voltage to be lowered enhancing the current drive properties of SOI technology and enabling further power reduction without the loss of performance (1). At the architectural level of design, the adoption of asynchronous timing rather than a global clock reduces power. Here, the synchronous clock is replaced by local handshake signals between blocks. Although asynchronous control tends to be larger than in synchronous systems, significant power savings should result as the clock generation, drivers, and distribution are consuming around one third of the power in large, complex, high performance systems (2).

ASYNCHRONOUS PROTOCOL

An asynchronous approach encourages a modular design, whereby circuit blocks operate independently of and concurrently with other blocks at their fastest natural rate. Fig. 1 shows the commonly used bundled data method of communicating between blocks. Valid input data to the block is indicated by a 'Request In' signal. The data remains valid until the block signals 'Acknowledge In' to the driving block. After allowing the block to operate, 'Request Out' signals the readiness of the output data and this data must remain constant until 'Acknowledge Out' from the receiving block. A four-phase protocol where the activation of the Acknowledge causes the Request line to be lowered which in turn causes the Acknowledge line to be deactivated, is highly suited to an ALU block. Here, the Request In signal can act as a Start signal for the arithmetic operation while Completion is used to form Request Out. In generating a Request Out signal, it is necessary to detect when the block has completed its operation. Techniques commonly used include a matched delay, a dummy-bit data path or self-timing. The latter is possible in blocks such as the adder where the completion time is data dependent.

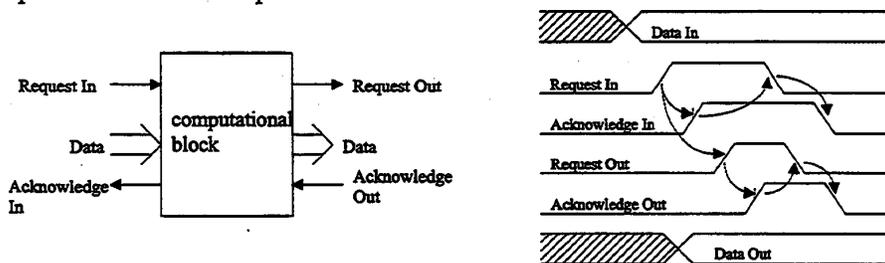


Fig. 1: Asynchronous Protocol

¹ Dept. of Elec. Eng. & Electronics, University of Liverpool, Brownlow Hill, Liverpool, L69 3GJ

² Dept. of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL

ADDER DESIGN

A circuit common to many synchronous and asynchronous designs is a fixed-point adder for integer operations. Thus although many different adder types can be constructed (3), almost all involve techniques for speeding up the propagation of the carry path from the least to the most significant bit as this determines the adder performance. In most systems, the adder block is on the critical path requiring that its performance is comparable to that of the other blocks. The AMULET1 microprocessor adopted a simple ripple adder approach for its 32-bit dynamic adder since analyses showed an average carry path length of around 12 bits (4). The schematic for one bit of the 16-bit static ripple adder for SOI & bulk technology is shown in fig. 2. The sum is formed by performing the XOR of the 2 data inputs (A & B) and the carry in (Cin). A multiplexer is used to form the carry out. If the two data inputs are equal, the carry out (Cout) is known regardless of Cin and is equal to A and B. If A and B are not equal, Cout is equal to Cin and will not be correct until Cin arrives. The states of A & B are also used to drive the carry valid logic for the stage (two Nand gates). If A and B are equal, not only is Cout known but it is valid for the stage and assuming the Start signal is active, the stage indicates the carry is valid by pulling its Vout low. The carry valid logic propagates in an exactly similar way to the propagation in the carry path. Completion of the addition operation is denoted when all stages indicate that their carry is valid. It is only necessary to use the valid signal from every other stage due to the propagation delay through the completion circuit (fig. 4). Once start is lowered, all valid signals go high, the completion signal is removed and the adder is ready to restart.

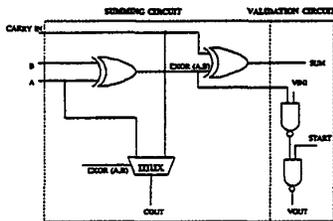


Fig.2: 1 bit of the adder

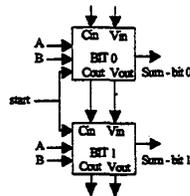


Fig.3: Connection of 2-bits

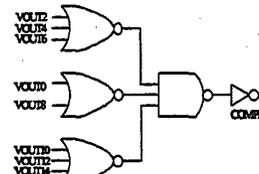


Fig.4: Completion circuit of the adder

DEVICE MODELS AND CIRCUIT SIMULATION

The bulk circuit simulation results were obtained by using Mitec 2 μm models in conjunction with Cadence 4.4.3., with 1.5 μm channel lengths. The devices were fully-depleted with accumulation mode PMOS and enhancement mode NMOS. The front oxide thickness was set equal to that of the SOI device, 20nm, and the threshold voltage (V_T) was reduced via the "vto" model parameter (5). Note that this allows a direct assessment of the intrinsic advantages (mobility and capacitance) of SOI compared to bulk. It does however, result in a much reduced 'off-current' for the bulk device. The model parameters for the SOI simulations were extracted from TMA Medici v2.0.2 simulations and the good agreement with HSPICE can be seen in Figs.5 & 6. The SOI adder simulation results were obtained using SOISPICE 4.4 under HSpice 99.2. To reduce any unfair advantage between the two technologies the leakage currents when the gate to source voltage was equal to 0V were kept approximately the same (6pA for PMOS & NMOS). The threshold voltages of SOI-PMOS devices undergo an increase as the supply voltage (source voltage) is reduced because the effective substrate bias is reduced. Therefore, the SOI-PMOS V_T 's were adjusted to be always equal to -0.35V for each value of supply voltage by using the VFBB parameter (6). Simulations were conducted with aspect ratios of all devices set at unity for maximum power reduction. Further simulations were conducted after a degree of optimisation for power/throughput whereby aspect ratios of the transistors that were not in the critical time delay path were fixed at unity and other blocks were 'speeded up' by judicious increase in appropriate aspect ratios.

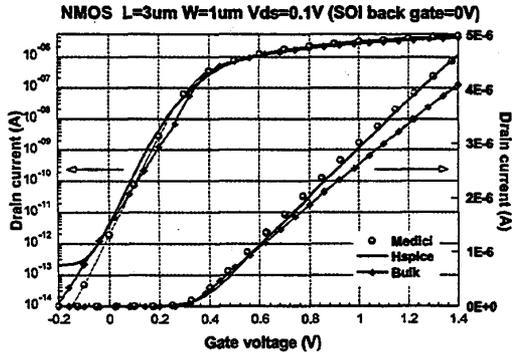


Fig.5: NMOS Device model verification

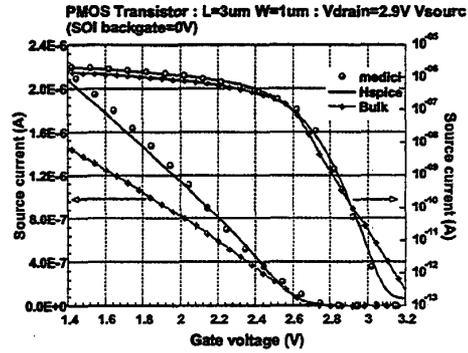


Fig.6: PMOS Device model verification

The simulations were conducted with 'carry in' equal to logic 1, all 16 bits of input A equal to logic 1 and all 16 bits of input B equal to logic 0. This represents a worst-case condition. The delay through the adder was taken between the time when the inputs, A & B, were correct and the complete signal going high, taking the 50% points as the reference. The energy per operation refers to the energy consumed by the adder during the longest carry operation. The energy consumed was logged through the use of an integrator circuit (7) and was measured between the time when the carry in and A_{0-15} were all going high, B_{0-15} were going low and the complete signal was switching from low from high. The energy required to reset the adder so that it is ready for its next operation was also included in the calculation.

RESULTS & ANALYSIS

As expected the unity aspect ratio adder on SOI technology offers reduced energy consumption and superior time delay over bulk, as shown in Fig.7. The SOI adder has a delay of 5ns at 5V compared to bulks 14ns - a factor of 3 difference. For a reduced supply voltage of 1.5V the SOI adder maintains its relative performance advantage over the bulk (16ns while the bulk is 47ns). The SOI adder uses 40% less energy than bulk at 5V and this increases to 52% at 1.5V. For the optimised bulk adder the delay reduces and the energy consumed by the adder increases in respect of the unity adder, Fig. 8, a delay of 14ns at 5V before optimisation being reduced to 8.1ns. The SOI delay improves with supply voltage because delay is proportional to load capacitance and inversely proportional to the transistor drive. The SOI enhancement NMOS transistor drive advantage over bulk reduces with decreasing supply voltage however, because the accompanying reduction in the vertical field has more leverage on bulk mobility than that of SOI. The contribution of bulk conduction in the case of accumulation mode SOI-PMOS is thought to result in the enhanced current drive at lower voltages.

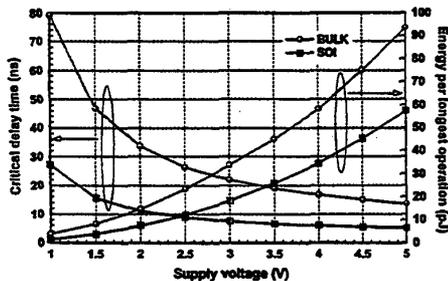


Fig.7: Energy per operation and critical propagation delay vs. supply for the unity adder

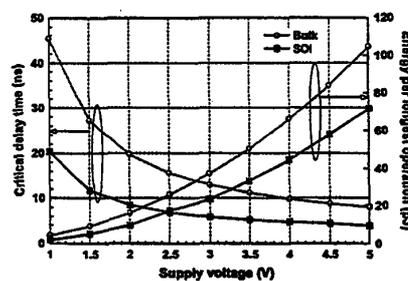


Fig.8: Energy per operation and critical delay vs. supply voltage for the optimised aspect ratio aspect adder

The delay and energy results of the two adders were compared. The delay of the bulk adder was reduced by 40-45% (supply voltage range 1.5V to 5V) due to optimisation while the energy increased from 10-15% over the same voltage range (Figs.9 and 10). The SOI adders delay is reduced by 20-25% while it uses 20-25% more energy. There is more leverage to optimise the bulk design than compared to the SOI design.

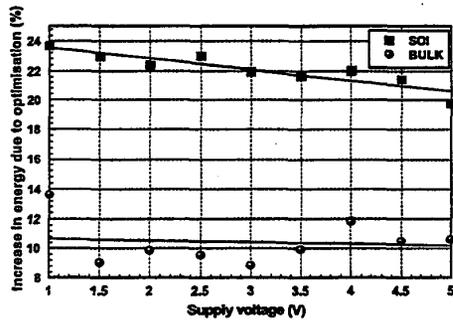


Fig9: Increase in energy due to optimisation

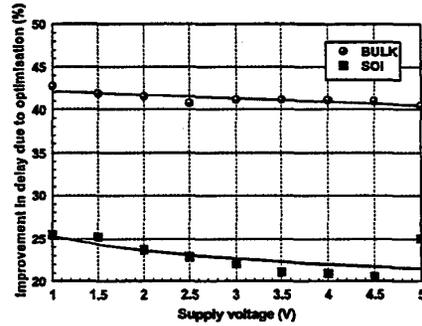


Fig 10: Improvement in delay due to optimisation

The energy-delay data for the 4 versions of the adder can be seen in figure 11. As the plots have no “dip” the optimum supply voltage for the adder still has not been reached. The energy-delay product for bulk is higher than for SOI; while the SOI optimised and unity aspect ratio adders have nearly identical energy-delay product. The energy consumed by the adder was divided up into 3 categories and is summarised in Fig.12: firstly the energy dissipated in the first XOR gate in each bit, secondly the energy consumed by the adder when there is carry ripple in the circuit (‘carry’), and finally the energy required to reset the adder (‘reset’). The energy required to reset the adder for the two supply voltages stated consumes about a fifth of the total energy for both SOI and bulk adders. The energy used by the first XOR gate in the SOI adder is a higher proportion of the total energy compared to bulk (approx. 40% for SOI compared to approx. 30% for bulk). It is worth noting that only 40% of the energy used by the SOI adder (50% for bulk) is actually being used for the carry ripple and validation circuitry.

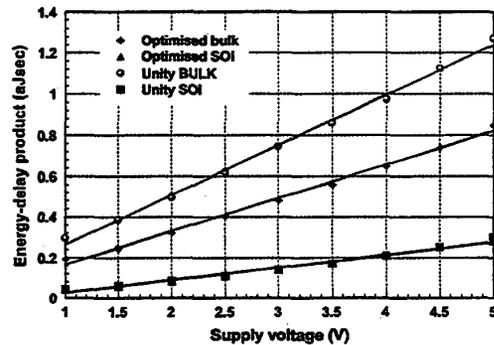


Fig11: Energy-delay plots for the 16-bit self timed adder

The results for the SOI unity aspect ratio adder at 5V are a little confusing as there are an XOR gate and 2 inverters in the XOR circuit while there is an XOR, a multiplexer, an inverter and two NAND gates in the Carry circuit. When all the A inputs switch at the same time (A0-A15 in Fig.14) a large current spike is apparent in the current drawn from the supply. The magnitude of this spike is greater in SOI than in bulk due to the higher drive of the SOI transistor, 14mA compared to 9mA (Fig. 14), although the energy dissipated during switching is 67pJ for bulk and 56pJ for SOI. The magnitude of this current spike for the SOI case is much reduced during the carry-ripple phase of the adder: 1.25mA compared to 14mA.

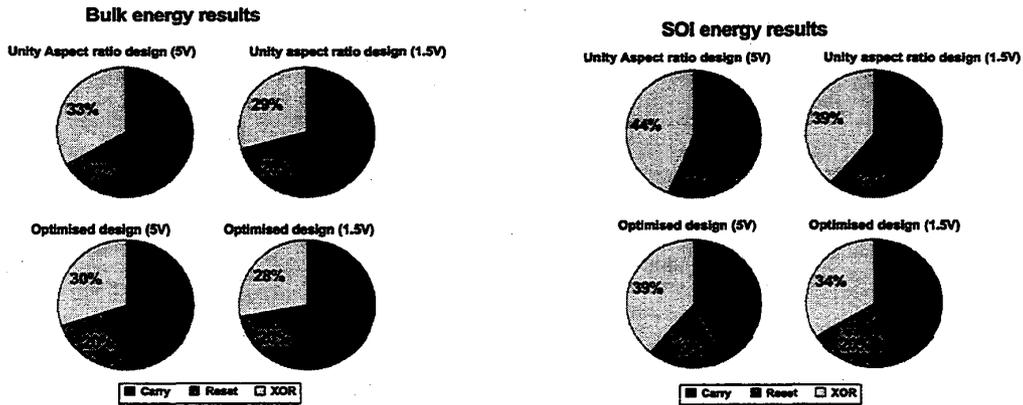


Fig. 12: A breakdown of the energy consumed by the adder

The adder only consumes 22pJ of energy in this mode. Figs. 14, 15 & 16 indicate the reason for the large current spike. Referring to Fig.13, consider the propagation of data through one bit of the adder and subdivide each bit into three sections: namely the first XOR gate; the multiplexer and second XOR gate; and finally the validation circuit. All three stages draw current from the supply during the same period of time, resulting in a large current spike. The "INV" data in figure 14 relates to the inverter in the first XOR gate. As all of the B inputs are grounded the other inverter in the first XOR gate draws negligible current. The "XOR" data in figure 14 refers to current entering the logic circuit after the two inverters in the exclusive OR gate.

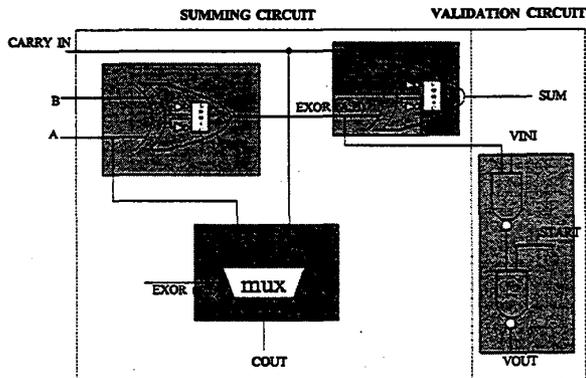


Fig13: 1 Bit of the adder displaying the different sections of the adder used in figs.14-16

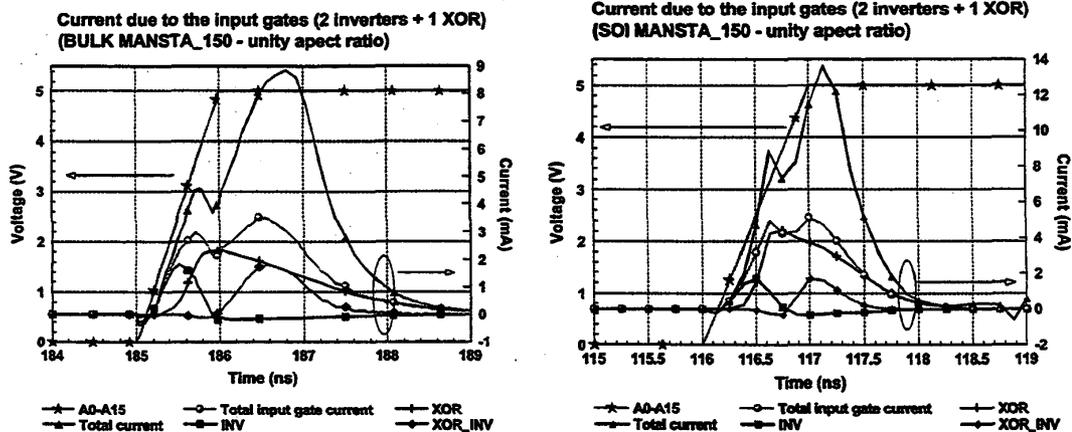


Fig. 14: Total current & input gate current consumed by the bulk (left) and SOI (right) unity aspect ratio adder.

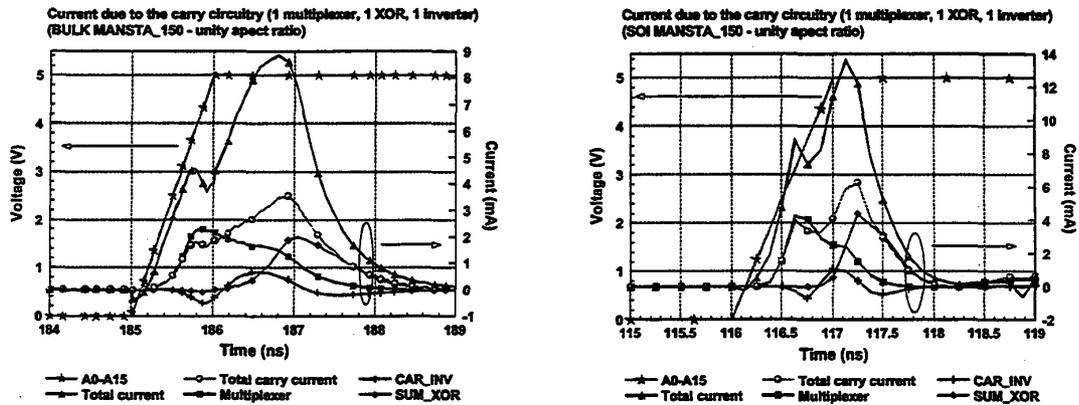


Fig. 15: Total current & total carry current consumed by the bulk (left) and SOI (right) unity aspect ratio adder.

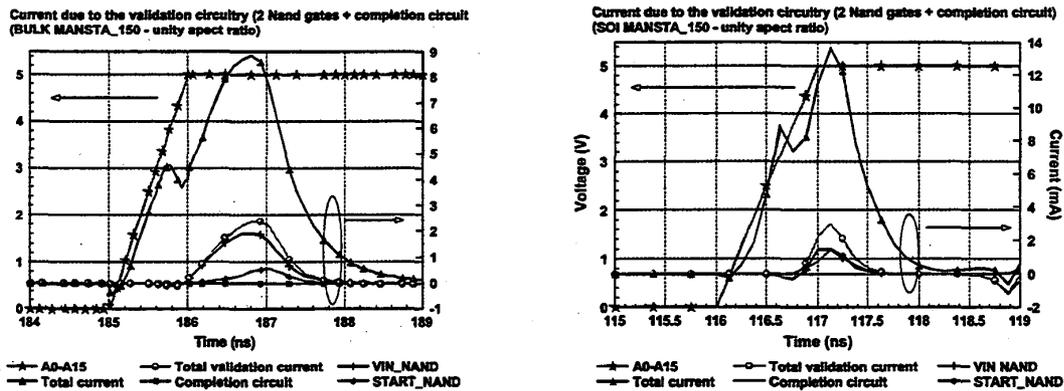


Fig. 16: Total current & total validation current consumed by the bulk (left) and SOI (right) unity aspect ratio adder.

The above displays a need for SOI circuits to have the capability of handling large current spikes, which may not occur in bulk circuits. This will relate to wider metal tracks and more vias & contacts than required in bulk CMOS design. Hence, the SOI designer needs to be extra careful when using global signals in SOI circuits, which activate circuits at the same time. The latter is a good case for using asynchronous design for SOI compared to synchronous design.

References

- (1) J. P. Colinge, "The Development of CMOS/SIMOX technology", *Microelectronic Engineering*, Vol. 28, 1995, pp.423-430.
- (2) D. W. Dobberphul et al, "A 200MHz 64-bit dual issue CMOS microprocessor", *IEEE Journal of Solid-State Circuits*, Vol.27, 1992, pp 1555-1565.
- (3) I. Flores, "The logic of computer arithmetic", Prentice Hall, 1963.
- (4) J.D. Garside, "A CMOS VLSI implementation of an asynchronous ALU", *IFIP Conf. on Asynch. Design Methodologies*, Manchester, UK, March 1993.
- (5) Cadence Spice Reference Manual, Dec. 1998.
- (6) Prof. J.G. Fossum, "SOISPICE-4 User's Guide",
- (7) S. M. Kang, "Accurate simulation of power dissipation in VLSI circuits", *Journal of Solid State Circuits*, Vol. SC-21, No. 5, Oct. 1986, pp 889-891.