IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 4, Issue 1, Ver. I (Jan. 2014), PP 38-43
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Reducing Power in Processor Unit via Centralized Dynamic Resource Size Management

Teena P John, Mrs.G.Dhanalakshmi.
ME VLSI Design Student, Sri Ramakrishna Engineering College, Coimbatore (Member of IEEE)
Assistant Professor, Department of ECE (UG), Sri Ramakrishna Engineering College, Coimbatore

Abstract:*In Microprocessor power minimization is a major concern. Many circuit and micro-architectural innovations have been proposed to reduce power in many individual processor units. The strategy is used to dynamically and simultaneously adjust the reorder buffer to reduce power dissipation in the data path without significant impact on the performance. Therefore a challenge is to find a centralized approach which can address power issues for one unit with the least amount of redesign and verification efforts and the least hardware overhead. Therefore it proposes such a centralized approach that attempts to simultaneously reduce power in processor unit with highest dissipation in reorder buffer. It is based on an observation that utilization varies significantly, during cache miss period. Some of the superscalar processors, such as the Intel processor implement physical registers using the Reorder Buffer (ROB) slots. One of the main dynamic instruction scheduling artifacts used in such datapath designs is the Reorder Buffer (ROB), which is used to recover to a precise state when interrupts or branch mispredictions occur. One of the power reduction technique used in ROB is dynamic resizing.*

Index Terms: *Cache miss driven, centralized low power technique, dynamic resource resizing, SRAM unit.*

I. Introduction

Leakage and dynamic power has grown significantly and is a major challenge in microprocessor design. In particular, in deep sub-micrometer technology (65 nm and below) high power dissipation become a major concern. For many individual processor units, several power reduction techniques have been proposed in the literature. Attempts have been made to either design new power-efficient units or to make current architectures more power-aware. However, the prior efforts have resulted in approaches which require considerable redesign and verification efforts.

Power consumption is a key parameter in most electronics applications. The most obvious applications for which power consumption is critical are battery-powered applications, such as home thermostats and security systems, in which the battery must last for years. Low power also leads to smaller power supplies, less expensive batteries, and enables products to be powered by signal lines (such as fire alarm wires) lowering the cost of the end-product. As a result, low power consumption has become a key parameter of microprocessor designs.

In recent years, several efforts have sought such a centralized algorithm through two major approaches: either dynamically adapting the data-path resources for power conservation or dynamically adapting the voltage and frequency level at a fine granularity or for the entire chip level. . The challenge is thus to find a centralized and simple mechanism which can reduce power for more than one unit (and ultimately the entire chip) and comes with the least amount of redesign and verification efforts, the lowest possible design risk and the least hardware overhead. Therefore the aim is to reduce power in the reorder buffer (ROB), with minimal hardware changes.

II. Related Work

There is a significant body of work on the design of the reorder buffer. Indeed several techniques have been proposed to reduce their power expenditure. Yet, as we shall show, much improvement can still be achieved by our technique.

A. Reorder Buffer

The reorder buffer is a multiported SRAM structures with the following many functions:

1) setting entries for up to IW instructions in each cycle.
2) releasing up to IW entries during commit stage in a cycle.
3) flushing entries during the branch recovery.

III. Existing Work

A.Introduction

Contemporary superscalar microprocessors use extensive execution reordering to maximize performance. One of the main dynamic instruction scheduling artifacts used in such data path designs is the Reorder Buffer (ROB), which is used to recover to a precise state when interrupts or branch mispredictions occur. The ROB is essentially implemented as a circular FIFO queue with head and tail pointers. Entries are made at the tail of the ROB in program order for each of the co–dispatched instructions. Instructions are committed from the head of the ROB to the architectural register file, thus preserving the correct order of updates to the program state.

B.Conventional ROB

The conventional ROB memory cell is composed of two cross-coupled CMOS inverters with two pass transistors connected to complementary bit-lines. Fig.3.1 shows this well known architecture of conventional ROB SRAM Cell, where the gate of access transistors are connected to the word-line (wl) to perform the access write and read operations thought the column bit-lines (bl and blbar). Bit-lines act as input/output nodes carrying the data from SRAM cells to a sense amplifier during read operation, or from write circuitry to the memory cells during write operations. The stability and robustness of a SRAM cell is usually evaluated analyzing both its dynamic and static behavior during the typical operations, write, read, and hold. In this paper we analyzed the leakage currents, standby currents at different Vdd and temperature. Further there effects are analyzed separately at 65nm respectively.

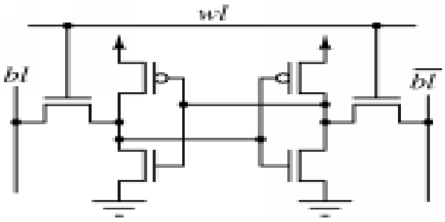

FIG1.CONVENTIONAL ROB

C.Reducing Performance Degradation

The elimination of the ROB read ports for reading out source operand values results in "disappearance" of these values for several cycles. In particular, the value is initially available through forwarding network in the write back cycle. As operand read ports are eliminated in the ROB, the value is not available again until it commits from the ROB to the ARF, at which point the second forwarding of the same value occurs. Certainly, the issue of some instructions that could not read this value from the proposed ROB structure will be delayed. To compensate for any resulting performance degradation, we use a set of latches, called retention latches, to cache results recently written to the ROB. Each retention latch is capable of holding a single-precision result, so double precision results would need two latches. Instructions dispatched since the writing of a result to the ROB can still access the value as long as they can get it from the retention latches. In case the lookup for source operand fails to find it within the retention latches, the operand value will be eventually obtained through the second forwarding of the same value when it is written to the ARF from the ROB.

D.DynamicDatapath Resource Allocation Strategy

The superscalar data path that employs the issue bound operand fetch policy. Here, even if input registers for an instruction contain valid data, these registers are not read out at the time of dispatch. Instead, when all the input operands of an instruction waiting in the DB are valid and a function unit of the required type is available, all of the input operands are read out from the register file (or as they are yet to be written to the register file, using bypassing logic to forward data from latter pipeline stages) and the instruction is issued. The dispatch/issue logic can be implemented using a global scoreboard that keeps track of instructions and register/FU availability. Alternatively, an associative logic can be used to update the status of input registers for instructions.

The ROB is partitioned in a similar fashion. The DB and the ROB are each implemented as a number of independent partitions. Each partition is a self–standing and independently usable unit, complete with its own precharger, sense amps, input/output drivers. In fact, the numbers of entries in a DB and ROB partition are 8 and 16 respectively in our studies. The maximum number of partitions available is 8 for the DB (for a total of 64 entries) and 8 for the ROB (for a total of 128 entries). A number of the DB (or ROB) partitions can be strung up together to implement a larger DB (or ROB). The connection running across the entries within a partition can be connected to a common through line. Similarly, by turning off the bypass switches associated with a partition, it can be deallocated from the current DB (or ROB). Unallocated partitions are turned off to save power/energy. Entries can be allocated in the DB across one or more active partitions in any order; an associative searching mechanism is used to look up free entries. Entries within the active ROB are used in a circular FIFO, and their allocation/deallocation strategies are not as straightforward as that for the DB.

IV. Proposed Work

A.Introduction

The proposed work aims at reducing the dynamic and static power dissipation of the ROB by adaptive resizing. Reducing the size of each of these units will require different hardware modifications (at extra power/area cost) which are determined by the resizing scheme. In order to minimize the hardware cost and complexity, a simple resizing scheme that goes from normal to half size and back. While this is not the most optimal resizing scheme for individual units it is the simplest one in terms of hardware modifications.

Based on a significant increase in ROB occupancies during cache miss periods it increases their size during such periods. During normal periods, we keep the ROB, size at half its possible size.

B.Dynamically Resizing ROB

The circuit level implementation of an SRAM is shown in figure. To read or write an entry each cycle all the bit lines must be precharged high (fired). For write operations, the high voltage on the bit lines induces a logic 1 into a cell for which the word line is fired. To read the content of an entry, one of bit line or will be conditionally discharged. The sense amplifier detects such a difference and will drive it to the output buffer. Bit lines thus must run across the entire ROB or register file height. As we may have multiple accesses to the same cell in a cycle, the read bit line and the write bit line thus need to be separated. Hence, if we are to read entries at each cycle and write to entries in the same cycle, we must have bit lines (register file is single ended). In our design of register files, we have (data width is 64 bits) bit lines for register files and bit lines for ROB. Precharging and discharging the large number of bit lines are the major sources of power dissipation in these two structures.

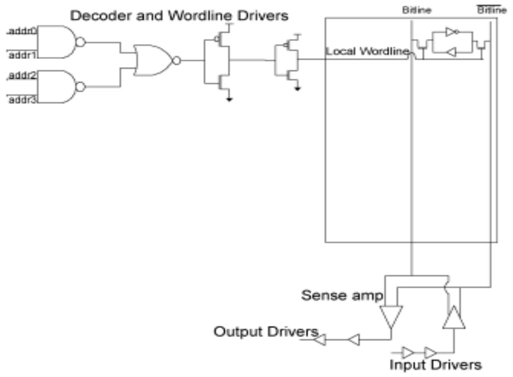

FIG 2. ROB SRAM CIRCUIT

The bit lines are the major consumers of power. It should be noted that most of the leakage of bit lines is due to the leakage currents of memory cells, which flow through the two off pass transistor to the bit lines. Accordingly, by eliminating the leakage in memory cells, we can eliminate the bit line leakage. The ROB and register file utilization is relatively low. Hence, one approach to reduce power dissipation in these two units would be to turn of the unused entries and their associated wordline drivers using circuit techniques such as gated-Vdd or gated Vss and eliminating the leakage power dissipation virtually completely.

The transition from sleep to active mode adds a one-cycle delay to the ROB or register file access which has significant performance impact. The algorithms proposed in the previous section reduce the performance impact of frequently activating and deactivating the entries. Resizing the ROB could be achieved by partitioning it into several independent units with separate sense amps, prechargers, and input and output drivers as explained. ROB is partitioned into eight units. This requires eight times more sense amps precharge lines and input and output drivers compared to the non-partitioned structure. The cost in terms of power and area is not negligible. To avoid adding to the complexity of ROB and register files, we use the divided bit line technique proposed for SRAMs to reduce the bit line capacitance and hence its dynamic power.

The 34 bit line segments for the ROB results in minimal area and power over- head. To downsize the ROB, the select signal of the lower partition is being AND together with the downsize signal. Doing that, no write and read can be done to/from the partition and the entire partition can be turned off safely. To turn off the en- tire partition we use the gated Vdd technique to suppress the voltage in all memory cells of the partition and eliminating its leakage almost completely. We also use a similar technique to eliminate leakage in the wordline driver of the disabled partition. Beginning of a cache miss period triggers upsizing the unit by deasserting the downsize signal and turning on the disabled partition. The overhead of downsizing and upsizing is one cycle (gated-Vdd overhead). The end of a cache miss period triggers downsizing the ROB. Note that the downsize signal is asserted only when the segment is empty.

The benefits of such resizing is in reducing both dynamic and leakage power. Leakage is suppressed by turning off the entire segment of memory cells and wordline driver. Dynamic power is reduced due to a smaller equivalent capacitance on the bit lines.

V. Simulation Results and Discussion

A. Introduction

This chapter deals with simulation result of the reorder buffer using 65nm and 45nm technology. Simulation results and power analysis was obtained using tanner tool.

B.65nm TECHNOLOGY

The power reduction can be obtained for both the existing and proposed work using 65nm technology.

1. Simulation of Conventional ROB

The simulation diagram of conventional ROB is shown below. It consists of 64bit. From the schematic diagram the conventional ROB circuit is called 64 times. In conventional ROB no control signal is applied. In this all the bits are enable during the operation. The data’s are stored in memory. The first rows in the schematic consist of 64bit.Totally the schematic consist of 20 rows. In this the power reduction can be obtained without dynamic allocation control.

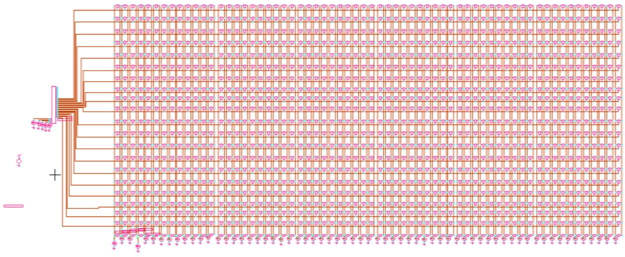

FIG 3. SIMULATION DIAGRAM OF CONVENTIONAL ROB

2. Simulationof ROB SRAM Circuit

The simulation result of rob SRAM consist of 64 bits. Here the bits can be divided into 32 bit. When the control signal is 1 all the 64 bits can be enable. If the control signal is 0 half will be enable and half will be disable. This uses nmos logic to work. In this work the power reduction can be obtained with dynamic allocation control.

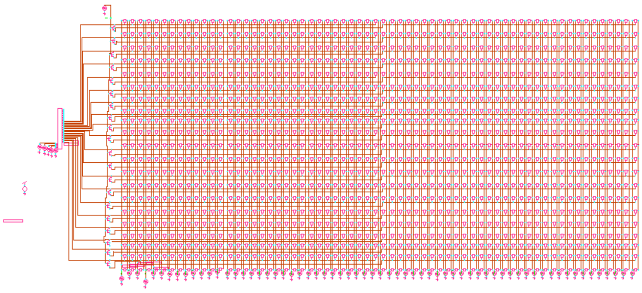

FIG 4. SIMULATION DIAGRAM OF ROB SRAM CIRCUIT

C.45nm TECHNOLOGY

Similar way the power reduction can be obtained for both existing and proposed work using 45nm technology.

1. Simulation of Conventional ROB

The simulation diagram of conventional ROB is shown below. It consists of 64bit. From the schematic diagram the conventional ROB circuit is called 64 times. In conventional ROB no control signal is applied. In this all the bits are enable during the operation. The data‘s are stored in memory. The first rows in the schematic consist of 64bit.Totally the schematic consist of 20 rows. This can be obtained without dynamic allocation control.

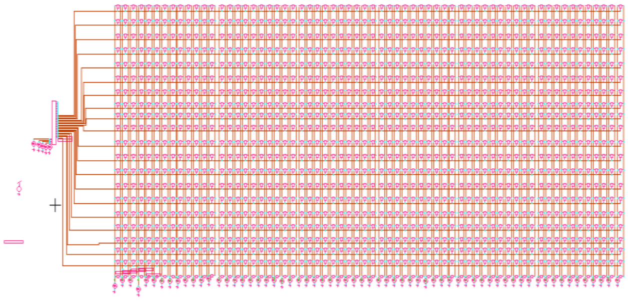

FIG 5. SIMULATION DIAGRAM OF CONVENTIONAL ROB

2. Simulation of ROB SRAM Circuit

The simulation result of rob SRAM consist of 64 bits. Here the bits can be divided into 32 bit. When the control signal is 1 all the 64 bits can be enable. If the control signal is 0 half will be enable and half will be disable. This uses nmos logic to work. This can be obtained using dynamic allocation control.

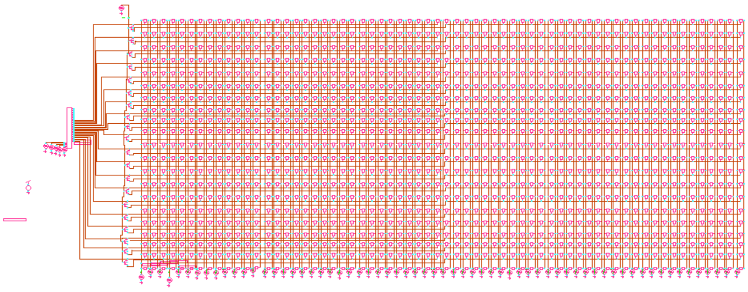

FIG 6. SIMULATION DIAGRAM OF ROB SRAM CIRCUIT

D.Simulation Results and Comparison

In this section we are going to discuss about the simulation results of existing work and the proposed work using 65nm and 45nm technology which is simulated using Tanner EDA tool.

The ROB SRAM circuit is power wise efficient when compared to conventional ROB in both technology. The power reduction can be obtained in the circuit.

TABLE 1
COMPARISON RESULT
65 nm TECHNOLOGY

CONVENTIONAL ROB	87mW
ROB SRAM CIRCUIT	68mW

Table 1 shows the comparison result of the conventional ROB and ROB SRAM circuit using 65nm technology.

TABLE 2
45nm TECHNOLOGY

CONVENTIONAL ROB	6mW
ROB SRAM CIRCUIT	5mW

Table 2 shows the comparison result of the conventional ROB and ROB SRAM circuit using 45nm technology.

VI. Conclusion

In SRAM based reorder buffer it reduces both leakage and dynamic power in processor units. It proposed to dynamically adjust the size of these units during cache miss period. In conventional SRAM circuit the power is reduced during 65nm technology. And in proposed SRAM ROB circuit the power is reduced during 45nm technology. Therefore the power reduction can be obtained. The power reduction of proposed work obtained for ROB SRAM circuit for both 65nm and 45nm is given as 68mW and 5mW.

References

[1]. HoumanHomayoun , “reducing power in all major CAM and SRAM based processor units via centralized ,dynamic resource size management”IEEE transactions on Very large scale integration system,vol.19,no.11,nov 2011.

[2]. A. Karandikar and K. K. Parhi, “Low power SRAM design using hierarchical divided bit-line approach,” in Proc. Int. Conf. Comput. Des., 1998, pp. 82–88.

[3]. I. Bahar and S. Manne, “Power and energy reduction via pipeline balancing,” in Proc. Int. Symp. Comput. Arch. (ISCA), 2001, pp. 218–229.

[4]. H. Homayoun, S. Pasricha, M. Makhzan, and A. Veidenbaum, “Dynamic register fileresizingandfrequency scaling to improve embedded processor performance and energy-delay efficiency,” presented at the 45th Des. Autom. Conf., Anaheim, CA, 2008.

[5]. D.Ponomarev, G. Kucuk, and K. Ghose, “Energy-efficient design of the reorder buffer,” presented at the Int. Workshop Power Tim. Model, Opt. Simulation (PATMOS), Seville, Spain, 2002.