IMPLEMENTATION OF PRECISE INTERRUPTS IN PIPELINED PROCESSORS

James E. Smith

Department of Electrical and Computer Engineering
University of Wisconsin-Madison
Madison, WI 53706

Andrew R. Pleszkun

Computer Sciences Department
University of Wisconsin-Madison
Madison, WI 53706

Abstract

An interrupt is precise if the saved process state corresponds with the sequential model of program execution where one instruction completes before the next begins. In a pipelined processor, precise interrupts are difficult to achieve because an instruction may be initiated before its predecessors have been completed. This paper describes and evaluates solutions to the precise interrupt problem in pipelined processors.

The precise interrupt problem is first described. Then five solutions are discussed in detail. The first forces instructions to complete and modify the process state in architectural order. The other four allow instructions to complete in any order, but additional hardware is used so that a precise state can be restored when an interrupt occurs. All the methods are discussed in the context of a parallel pipeline structure. Simulation results based on the CRAY-1S scalar architecture are used to show that, at best, the first solution results in a performance degradation of about 16%. The remaining four solutions offer similar performance, and three of them result in as little as a 3% performance loss. Several extensions, including virtual memory and linear pipeline structures, are briefly discussed.

1. Introduction

Most current computer architectures are based on a sequential model of program execution in which an architectural program counter sequences through instructions one-by-one, finishing one before starting the next. In contrast, a high performance implementation may be pipelined, permitting several instructions to be in some phase of execution at the same time. The use of a sequential architecture and a pipelined implementation clash at the time of an interrupt; pipelined instructions may modify the process state in an order different from that defined by the sequential architectural model. At the time an interrupt condition is detected, the hardware may not be in a state that is consistent with any specific program counter value.

When an interrupt occurs, the state of an interrupted process is typically saved by the hardware, the software, or by a combination of the two. The process state generally consists of the program counter, registers, and memory. If the saved process state is consistent with the sequential architectural model then the interrupt is *precise*. To be more specific, the saved state should reflect the following conditions.

(1) All instructions preceding the instruction indicated by the saved program counter have been executed and have modified the process state correctly.

(2) All instructions following the instruction indicated by the saved program counter are unexecuted and have not modified the process state.

(3) If the interrupt is caused by an exception condition raised by an instruction in the program, the saved program counter points to the interrupted instruction. The interrupted instruction may or may not have been executed, depending on the definition of the architecture and the cause of the interrupt. Whichever is the case, the interrupted instruction has either completed, or has not started execution.

If the saved process state is inconsistent with the sequential architectural model and does not satisfy the above conditions, then the interrupt is *imprecise*.

This paper describes and compares ways of implementing precise interrupts in pipelined processors. The methods used are designed to modify the state of an executing process in a carefully controlled way. The simpler methods force all instructions to update the process state in the architectural order. Other, more complex methods save portions of the process state so that the proper state may be restored by the hardware at the time an interrupt occurs.

1.1. Classification of Interrupts

We consider interrupts belonging to two classes.

(1) *Program interrupts*, sometimes referred to as "traps", result from *exception conditions* detected during fetching and execution of specific instructions. These exceptions may be due to software errors, for example trying to execute an illegal opcode, numerical errors such as overflow, or they may be part of normal execution, for example page faults.

(2) *External interrupts* are not caused by specific instructions and are often caused by sources outside the currently executing process, sometimes completely unrelated to it. I/O interrupts and timer interrupts are examples.

For a specific architecture, all interrupts may be defined to be precise or only a proper subset. Virtually every architecture, however, has some types of interrupts that must be precise. There are a number of conditions under which precise interrupts are either necessary or desirable.

(1) For I/O and timer interrupts a precise process state makes restarting possible.

(2) For software debugging it is desirable for the saved state to be precise. This information can be helpful in isolating the exact instruction and circumstances that caused the exception condition.

(3) For graceful recovery from arithmetic exceptions, software routines may be able to take steps, re-scale floating point numbers for example, to allow a process to continue. Some end cases of modern floating point arithmetic systems might best be handled by software; gradual underflow in the proposed IEEE floating point standard [Stev81], for example.

(4) In virtual memory systems precise interrupts allow a process to be correctly restarted after a page fault has been serviced.

(5) Unimplemented opcodes can be simulated by system software in a

way transparent to the programmer if interrupts are precise. In this way, lower performance models of an architecture can maintain compatibility with higher performance models using extended instruction sets.

(6) Virtual machines can be implemented if privileged instruction faults cause precise interrupts. Host software can simulate these instructions and return to the guest operating system in a user-transparent way.

1.2. Historical Survey

The precise interrupt problem is as old as the first pipelined computer and is mentioned as early as Stretch [Buch62]. The IBM 360/91 [Ande67] was a well-known computer that produced imprecise interrupts under some circumstances, floating point exceptions, for example. Imprecise interrupts were a break with the IBM 360 architecture which made them even more noticeable. All subsequent IBM 360 and 370 implementations have used less aggressive pipeline designs where instructions modify the process state in strict program order, and all interrupts are precise.[1] A more complete description of the method used in these "linear" pipeline implementations is in Section 8.4.

Most pipelined implementations of general purpose architectures are similar to those used by IBM. These pipelines constrain all instructions to pass through the pipeline in order with a stage at the end where exception conditions are checked before the process state is modified. Examples include the Amdahl 470 and 580 [Amdh81,Amdh80] and the Gould/SEL 32/87 [Ward82].

The high performance CDC 6600 [Thor70], CDC 7600 [Bons69], and Cray Research [Russ78, Cray79] computers allow instructions to complete out of the architectural sequence. Consequently, they have some exception conditions that result in imprecise interrupts. In these machines, the advantages of precise interrupts have been sacrificed in favor of maximum parallelism and design simplicity. I/O interrupts in these machines are precise, and they do not implement virtual memory.

The CDC STAR-100 [HiTa72] and CYBER 200 [CDC81] series machines also allow instructions to complete out of order, and they do support virtual memory. In these machines the use of vector instructions further complicates the problem, and all the difficulties were not fully recognized until late in the development of the STAR-100. The eventual solution was the addition of an *invisible exchange package* [CDC81]. This captures machine-dependent state information resulting from partially completed instructions. A similar approach has more recently been suggested in MIPS [Henn82] where pipeline information is dumped at the time of an interrupt and restored to the pipeline when the process is resumed. This solution makes a process restartable although it is arguable whether it has all the features and advantages of an architecturally precise interrupt. For example, it might be necessary to have implementation-dependent software sift through the machine-dependent state in order to provide complete debug information.

The recently-announced CDC CYBER 180/990 [CDC84] is a pipelined implementation of a new architecture that supports virtual memory, and offers roughly the same performance as a CRAY-1S. To provide precise interrupts, the CYBER 180/990 uses a history buffer, to be described later in this paper, where state information is saved just prior to being modified. Then when an interrupt occurs, this "history" information can be used to back the system up into a precise state.

1.3. Paper Overview

This paper concentrates on explaining and discussing basic methods for implementing precise interrupts in pipelined processors. We emphasize scalar architectures (as opposed to vector architectures) because of their applicability to a wider range of machines. Section 2 describes the model architecture to be used in describing precise interrupt implementations. The model architecture is very simple so that the fundamentals of the methods can be clearly described. Sections 3 through 6 describe methods for implementing precise interrupts. Section 3 describes a simple method that is easy to implement, but which reduces performance. It forces instructions to complete in architectural order which sometimes inhibits the degree of parallelism in a pipelined system. Section 4 describes a higher performance variation where results may be bypassed to other instructions before the results are used to modify the process state. Sections 5 and 6 describe methods where instructions are allowed to complete in any order, but where state information is saved so that a precise state may be restored when an interrupt occurs. The descriptions of these methods assume that the only state information is the program counter, general purpose registers, and main memory. The methods are also discussed in the absence of a data cache. Section 7 presents simulation results. Experimental results based on these CRAY-1S simulations are presented and discussed. Section 8 contain a brief discussion of 1) saving additional state information, 2) supporting virtual memory, 3) precise interrupts when a data cache is used, and 4) linear pipeline structures.

[1] Except for the models 95 and 195 which were derived from the original model 91 design. Also, the models 85 and 165 had imprecise interrupts for the case of protection exceptions and addressing exceptions caused by store operations.

2. Preliminaries

2.1. Model Architecture

For describing the various techniques, a model architecture is chosen so that the basic methods are not obscured by details and unnecessary complications brought about by a specific architecture.

We choose a register-register architecture where all memory accesses are through registers and all functional operations involve registers. In this respect it bears some similarity to the CDC and Cray architectures, but has only one set of registers. The load instructions are of the form: Ri = (Rj+disp). That is, the content of Rj plus a displacement given in the instruction are added to form an effective address. The content of the addressed memory location is loaded into Ri. Similarly, a store is of the form: (Rj + disp) = Ri, where Ri is stored at the address found by adding the content of Rj and a displacement. The functional instructions are of the form Ri = Rj op Rk, where op is the operation being performed. For unary operations, the degenerate form Ri = op Rk is used. Conditional instructions are of the form P = disp : Rj op Rj, where the displacement is the address of the branch target; op is a relational operator, =, >, <, etc.

The only process state in the model architecture consists of the program counter, the general purpose registers and main memory. The architecture is simple, has a minimal amount of process state, can be easily pipelined, and can be implemented in a straightforward way with parallel functional units like the CDC and Cray architectures. Hence, implementing precise interrupts for the model architecture presents a realistic problem.

Initially, we assume no operand cache. Similarly, condition codes are not used. They add other problems beyond precise interrupts when a pipelined implementation is used. Extensions for operand cache and condition codes are discussed in Section 8.

The implementation for the simple architecture is shown in Fig. 1. It uses an instruction fetch/decode pipeline which processes instructions in order. The last stage of the fetch/decode pipeline is an issue register where all register interlock conditions are checked. If there are no register conflicts, an instruction issues to one of the parallel functional units. Here, the memory access function is implemented as one of the functional units. The operand registers are read at the time an instruction issues. There is a single result bus that returns results to the register file. This bus may be reserved at the time an instruction issues or when an instruction is approaching completion. This assumes the functional unit times are deterministic. A new instruction can issue every clock period in the absence of register or result bus conflicts.

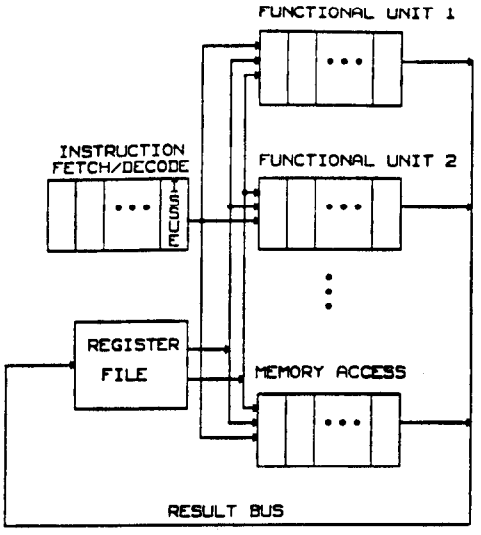


Figure 1. Pipelined implementation of our model architecture. Not shown is the result shift register used to control the result bus.

Example 1

To demonstrate how an imprecise process state may occur in our model architecture, consider the following section of code which sums the elements of arrays A and B into array C. Consider the instructions in statements 6 and 7. Although the integer add which increments the loop count will be issued after the floating point add, it will complete

			Comments	Execute Time
0		R2 <- 0	Init. loop index	
1		R0 <- 0	Init. loop count	
2		R5 <- 1	Loop inc. value	
3		R7 <- 100	Maximum loop count	
4	L1:	R1 <- (R2 + A)	Load A(I)	11cp
5		R3 <- (R2 + B)	Load B(I)	11cp
6		R4 <- R1 +f R3	Floating add	6cp
7		R0 <- R0 + R5	Inc. loop count	2cp
8		(R0 + C) <- R4	Store C(I)	
9		R2 <- R2 + R5	Inc. loop index	2cp
10		P= L1 : R0 != R7	cond. branch not equal	

before the floating point add. The integer add will therefore change the process state before an overflow condition is detected in the floating point add. In the event of such an overflow, there is an imprecise interrupt.

2.2. Interrupts Prior to Instruction Issue

Before proceeding with the various precise interrupt methods, we discuss interrupts that occur prior to instruction issue separately because they are handled the same way by all the methods.

In the pipeline implementation of Fig. 1, instructions stay in sequence until the time they are issued. Furthermore, the process state is not modified by an instruction before it issues. This makes precise interrupts a simple matter when an exception condition can be detected prior to issue. Examples of such exceptions are privileged instruction faults and unimplemented instructions. This class also includes external interrupts which can be checked at the issue stage.

When such an interrupt condition is detected, instruction issuing is halted. Then, there is a wait while all previously issued instructions complete. After they have completed, the process is in a precise state, with the program counter value corresponding to the instruction being held in the issue register. The registers and main memory are in a state consistent with this program counter value.

Because exception conditions detected prior to instruction can be handled easily as described above, we will not consider them any further. Rather, we will concentrate on exception conditions detected after instruction issue.

3. In-order Instruction Completion

With this method, instructions modify the process state only when all previously issued instructions are known to be free of exception conditions. This section describes a strategy that is most easily implemented when pipeline delays in the parallel functional units are fixed. That is, they do not depend on the operands, only on the function. Thus, the result bus can be reserved at the time of issue.

First, we consider a method commonly used to control the pipelined organization shown in Fig. 1. This method may be used regardless of whether precise interrupts are to be implemented. The precise interrupt methods described in this paper are integrated into this basic control strategy, however. To control the result bus, a "result shift register" is used; see Fig. 2. Here, the stages are labeled 1 through n.

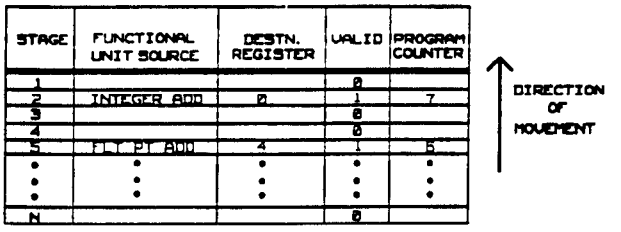


Figure 2. Result Shift Register

where n is the length of the longest functional unit pipeline. An instruction that takes i clock periods reserves stage i of the result shift register at the time it issues. If the stage already contains valid control information, then issue is held until the next clock period, and stage i is checked once again. An issuing instruction places control information in the result shift register. This control information identifies the functional unit that will be supplying the result and the destination register of the result. This control information is also marked "valid" with a validity bit. Each clock period, the control information is shifted down one stage toward stage one. When it reaches stage one, it is used during the next clock to control the result bus so that the functional unit result is placed in the correct result register.

Still disregarding precise interrupts, it is possible for a short instruction to be placed in the result pipeline in stage i when previously issued instructions are in stage j, j > i. This leads to instructions finishing out of the original program sequence. If the instruction at stage j eventually encounters an exception condition, the interrupt will be imprecise because the instruction placed in stage i will complete and modify the process state even though the sequential architecture model says i does not begin until j completes.

Example 2

If one considers the section of code presented in Example 1, and an initially empty result shift register (all the entries invalid), the floating point add would be placed in stage 6 while the integer add would be placed in stage 2. The result shift register entries shown in Fig. 2 reflect the state of the result shift register after the integer add issues. Notice that the floating point add entry is in stage 5 since one clock period has passed since it issued. As described above, this situation leads to instructions finishing out of the original program sequence.

3.1. Registers

To implement precise interrupts with respect to registers using the above pipeline control structure, the control should "reserve" stages $i < j$ as well as stage j. That is, the stages $i < j$ that were not previously reserved by other instructions are reserved, and they are loaded with null control information so that they do not affect the process state. This guarantees that instructions modifying registers finish in order.

There is logic on the result bus that checks for exception conditions in instructions as they complete. If an instruction contains a non-masked exception condition, then control logic "cancels" all subsequent instructions coming on the result bus so that they do not modify the process state.

Example 3

For our sample section of code given in Example 1, assuming the the result shift register is initially empty, such a policy would have the floating point add instruction reserve stages 1 through 6 of the result shift register. When, on the next clock cycle, the integer add is in the issue register, it would normally issue and reserve stage 2. However, this is now prohibited from happening because stage 2 is already reserved. Thus, the integer add must wait at the issue stage until stage 2 of the result shift register is no longer reserved. This would be 5 clock periods after the issue of the floating point add.

A generalization of this method is to determine, if possible, that an instruction is free of exception conditions prior to the time it is complete. Only result shift register stages that will finish before exceptions are detected need to be reserved (in addition to the stage that controls the result).

3.2. Main Memory

Store instructions modify the portion of process state that resides in main memory. To implement precise interrupts with respect to memory, one solution is to force store instructions to wait for the result shift register to be empty before issuing. Alternatively, stores can issue and be held in the load/store pipeline until all preceding instructions are known to be exception-free. Then the store can be released to memory.

To implement the second alternative, recall that memory can be treated as a special functional unit. Thus, as with any other instruction, the store can make an entry in the result shift register. This entry is defined as a *dummy* store. The dummy store does not cause a result to be placed in the registers, but is used for controlling the memory pipeline. The dummy store is placed in the result shift register so that it will not reach stage 1 until the store is known to be exception-free. When the dummy store reaches stage 1, all previous instructions have completed without exceptions, and a signal is sent to the load/store unit to release the store to memory. If the store itself contains an exception condition, then the store is cancelled, all following load/store instructions are cancelled, and the store unit signals the pipeline control so that all instructions issued subsequent to the store are cancelled as they leave the result pipeline.

3.3. Program Counter

To implement precise interrupts with respect to the program counter, the result shift register is widened to include a field for the program counter of each instruction (see Fig. 2). This field is filled as the instruction issues. When an instruction with an exception condition appears at the result bus, its program counter is available and becomes part of the saved state.

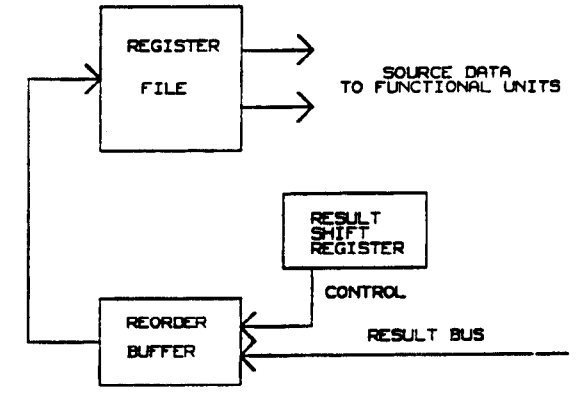
4. The Reorder Buffer

The primary disadvantage of the above method is that fast instructions may sometimes get held up at the issue register even though they have no dependencies and would otherwise issue. In addition, they block the issue register while slower instructions behind them could conceivably issue.

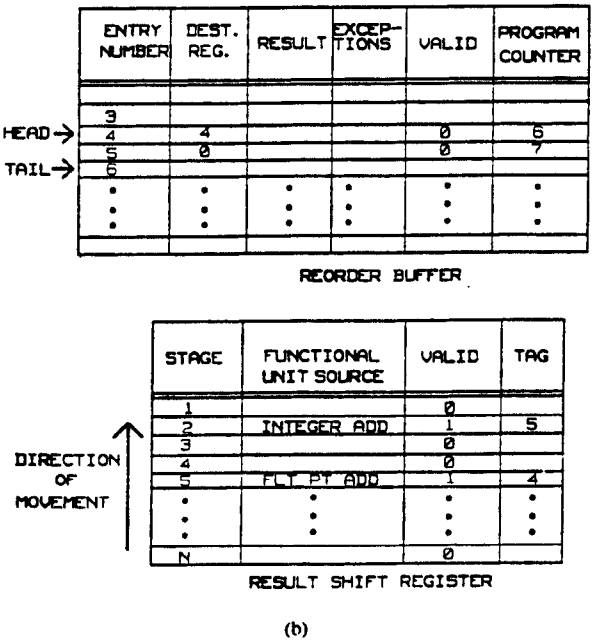
This leads us to a more complex, but more general solution. Instructions are allowed to finish out of order, but a special buffer called the *reorder buffer* is used to reorder them before they modify the process state.

4.1. Basic Method

The overall organization is shown in Fig. 3a. The reorder buffer, Fig. 3b, is a circular buffer with head and tail pointers. Entries between the head and tail are considered valid. At instruction issue time the next available reorder buffer entry, pointed to by the tail pointer, is given to the issuing instruction. The tail pointer value is used as a tag to identify the entry in the buffer reserved for the instruction. The tag is placed in the result shift register along with the other control information. The tail pointer is then incremented, modulo the buffer size. The result shift register differs from the one used earlier because there is a field containing a reorder tag instead of a field specifying a destination register.



(a)



(b)

Figure 3. (a) Reorder Buffer Organization. (b) Reorder Buffer and associated Result Shift Register.

When an instruction completes, both results and excepuon conditions are sent to the reorder buffer. The tag from the result shift register is used to guide them to the correct reorder buffer entry. When the entry at the head of the reorder buffer contains valid results (its instruction has finished) then its exceptions are checked. If there are none, the results are written into the registers. If an exception is detected, issue is stopped in preparation for the interrupt, and all further writes into the register file are inhibited.

Example 4

The entries in the reorder buffer and result shift register shown in Figure 3b reflect their state after the integer add from Example 2 has issued. Notice that the result shift register entries are very similar to those in the Figure 2. The integer add will complete execution before the floating point add and its results will be placed in entry 5 of the reorder buffer. These results, however, will not be written into R0 until the floating point result, found in entry 4, has been placed in R4.

4.2. Main Memory

Preciseness with respect to memory is maintained in manner similar to that in the in-order completion scheme (Section 3.2). The simplest method holds stores in the issue register until all previous instructions are known to be free of exceptions. In the more complex method, a store signal is sent to the memory pipeline as a "dummy" store is removed from the reorder buffer. Stores are allowed to issue, and block in the store pipeline prior to being committed to memory while they wait for their dummy counterpart.

4.3. Program Counter

To maintain preciseness with respect to the program counter, the program counter can be sent to a reserved space in the reorder buffer at issue time (shown in Figure 3b). While the program counter could be sent to the result shift register, it is expected that the result shift register will contain more stages than the reorder buffer and thus require more hardware. The length of the result shift register must be as long as the longest pipeline stage. As will be seen in Section 7, the number of entries in the reorder buffer can be quite small. When an instruction arrives at the head of the reorder buffer with an exception condition, the program counter found in the reorder buffer entry becomes part of the saved precise state.

4.4. Bypass Paths

While an improvement over the method described in Section 3, the reorder buffer still suffers a performance penalty. A computed result that is generated out of order is held in the reorder buffer until previous instructions, finishing later, have updated the register file. An instruction dependent on a result being held in the reorder buffer cannot issue until the result has been written into the register file.

The reorder buffer may, however, be modified to minimize some of the drawbacks of finishing strictly in order. For results to be used early, bypass paths may be provided from the entries in the reorder buffer to the register file output latches, see Fig. 4. These paths allow data being held in the reorder buffer to be used in place of register data. The implementation of this method requires comparators for each reorder buffer stage and operand designator. If an operand register designator of an instruction being checked for issue matches a register designator in the reorder buffer, then a multiplexer is set to gate the data from the reorder buffer to the register output latch. In the absence of other issue blockage conditions, the instruction is allowed to issue, and the data from the reorder data is used prior to being written into the register file.

There may be bypass paths from some or all of the reorder buffer entries. If multiple bypass paths exist, it is possible for more than one destination entry in the reorder buffer to correspond to a single register. Clearly only the *latest* reorder buffer entry that corresponds to an operand designator should generate a bypass path to the register output latch. To prevent multiple bypassing of the same register, when an instruction is placed in the reorder buffer, any entries with the same destination register designator must be inhibited from matching a bypass check.

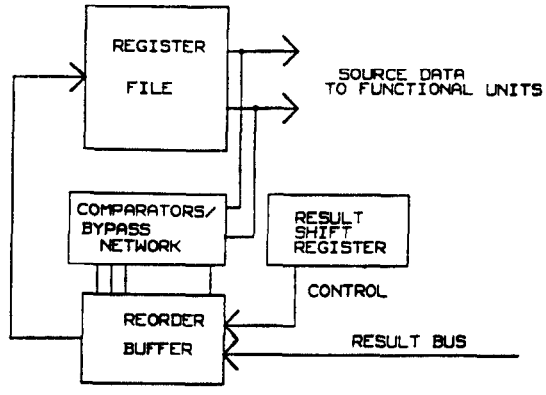


Figure 4. Reorder Buffer Method with Bypasses.

When bypass paths are added, preciseness with respect to the memory and the program counter does not change from the previous method.

The greatest disadvantage with this method is the number of bypass comparators needed and the amount of circuitry required for the multiple bypass check. While this circuitry is conceptually simple, there is a great deal of it.

5. History Buffer

The methods presented in this section and the next are intended to reduce or eliminate performance losses experienced with a simple reorder buffer, but without all the control logic needed for multiple bypass paths. Primarily, these methods place computed results in a working register file, but retain enough state information so a precise state can be restored if an exception occurs.

Fig. 5a illustrates the history buffer method. The history buffer is organized in a manner very similar to the reorder buffer. At issue time, a buffer entry is loaded with control information, as with the reorder buffer, but the value of the destination register (soon to be overwritten) is also read from the register file and written into the buffer entry. Results on the result bus are written directly into the register file when an instruction completes. Exception reports come back as an instruction completes and are written into the history buffer. As with the reorder buffer, the exception reports are guided to the proper history buffer entry through the use of tags found in the result shift register. When the history buffer contains an element at the head that is known to have finished without exceptions, the history buffer entry is no longer needed and that buffer location can be re-used (the head pointer is incremented). As with the reorder buffer, the history buffer can be shorter than the maximum number of pipeline stages. If all history buffer entries are used (the buffer is too small), issue must be blocked until an entry becomes available. Hence the buffer should be long enough so that this seldom happens. The effect of the history buffer on performance is determined in Section 7.

Example 5

The entries in the history buffer and result shift register shown Fig. 5b correspond to our code in Example 1, after the integer add has issued. The only differences between this and the reorder buffer method shown in Fig. 3b are the addition of an "old value" field in the history buffer and a "destination register" field in the result shift register. The result shift register now looks like the one shown in Fig. 2.

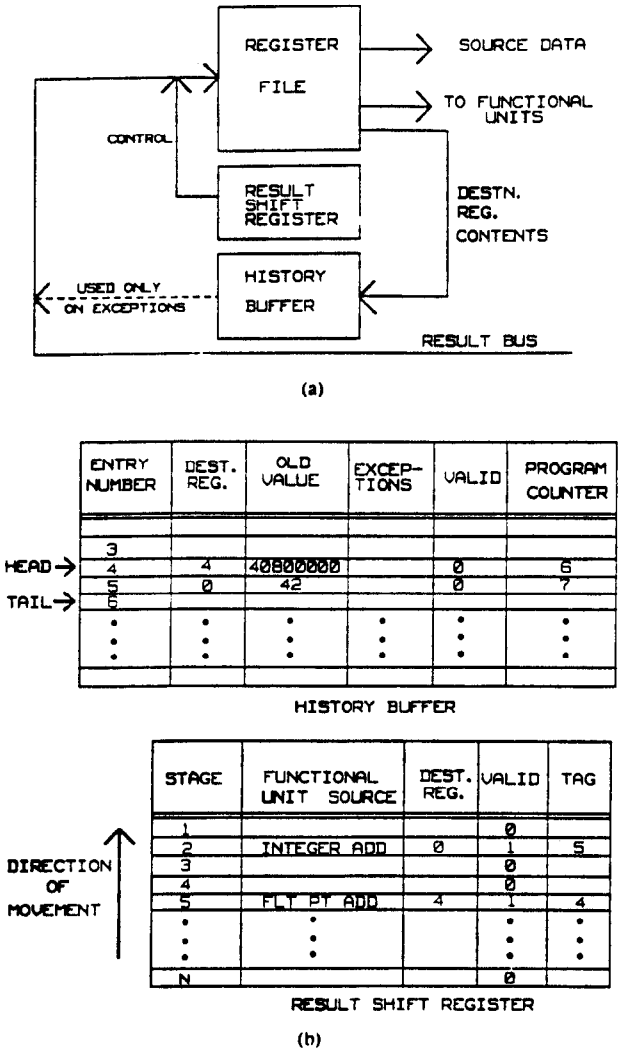


Figure 5. (a) History Buffer Organization. (b) History Buffer and associated Result Shift Register.

When an exception condition arrives at the head of the buffer, the buffer is held, instruction issue is immediately halted, and there is a wait until pipeline activity completes. The active buffer entries are then emptied from tail to head, and the history values are loaded back into their original registers. The program counter value found in the head of the history buffer is the precise program counter.

To make main memory precise, when a store entry emerges from the buffer, it sends a signal that another store can be committed to memory. Stores can either wait in the issue register or can be blocked in the memory pipeline, as in the previous methods.

The extra hardware required by this method is in the form of a large buffer to contain the history information. Also the register file must have three read ports since the destination value as well as the source operands must be read at issue time. There is a slight problem if the basic implementation has a bypass of the result bus around the register file. In such a case, the bypass must also be connected into the history buffer.

6. Future File

The future file method (Fig. 6) is similar to the history buffer method, however it uses two separate register files. One register file reflects the state of the architectural (sequential) machine. This file will be referred to as the *architectural file*. A second register file is updated as soon as instructions finish and therefore runs ahead of the architectural file (i.e. it reflects the future with respect to the architectural file). This *future file* is the working file used for computation by the functional units.

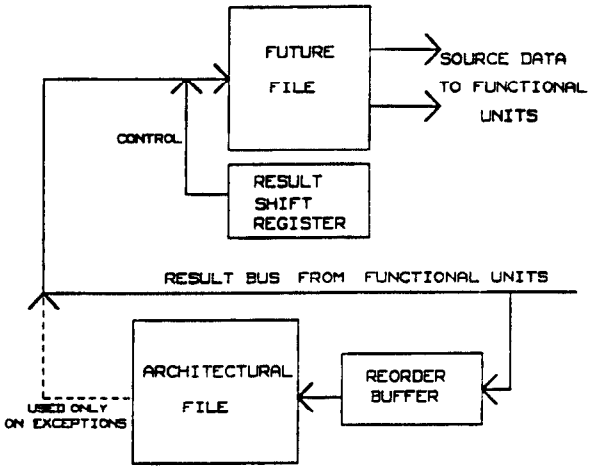


Figure 6. Future File Organization.

Instructions are issued and results are returned to the future file in any order, just as in the original pipeline model. There is also a reorder buffer that receives results at the same time they are written into the future file. When the head pointer finds a completed instruction (a valid entry), the result associated with that entry is written in the architectural file.

Example 6

If we consider the code in Example 1 again, there is a period of time when the architecture file and the future file contain different entries. With this method, an instruction may finish out of order, so when the integer add finishes, the future file contains the new contents of R0. The architecture file however does not, and the new contents of R0 are buffered in the reorder buffer entry corresponding to the integer add. Between the time the integer add finishes and the time the floating point add finishes, the two files are different. Once the floating point finishes and its results are written into R4 of both files, R0 of the architecture file is written.

Just as with the pure reorder buffer method, program counter values are written into the reorder buffer at issue time. When the instruction at the head of the reorder buffer has completed without error, its result is placed in the architectural file. If it completed with an error, the register designators associated with the buffer entries between the head and tail pointers are used to restore values in the future file from the architectural file.[2]

The primary advantage of the future file method is realized when the architecture implements interrupts via an "exchange" where all the registers are automatically saved in memory and new ones are restored (as is done in CDC and Cray architectures). In this case, the architectural file can be stored away immediately; no restoring is necessary as in history buffer method. There is also no bypass problem as with the history buffer method.

[2] The restoration is performed from the architectural file since the future file is register file from which all execution takes place.

7. Performance Evaluation

To evaluate the effectiveness of our precise interrupt schemes, we use a CRAY-1S simulation system developed at the University of Wisconsin [PaSm83]. This trace-driven simulator is extremely accurate, due to the highly deterministic nature of the CRAY-1S, and gives the number of clock periods required to execute a program.

The scalar portion of the CRAY-1S is very similar to the model architecture described in Section 2.1. Thus, casting the basic approaches into the CRAY-1S scalar architecture is straightforward.

For a simulation workload, the first fourteen Lawrence Livermore Loops [McMa72] were used. Because we are primarily interested in pipelined implementations of conventional scalar architectures, the loops were compiled by the Cray FORTRAN compiler with the vectorizer turned off.

In the preceding sections, five methods were described that could be used for guaranteeing precise interrupts. To evaluate the effect of these methods on system performance, the methods were partitioned into three groups. The first and second group respectively contain the in-order method and the simple reorder buffer method. The third group is composed of the reorder buffer with bypasses, the history buffer, and the future file. This partitioning was performed because the methods in the third group result in identical system performance. This is because the future file has a reorder buffer embedded as part of its implementation. And the history buffer length constrains performance in the same way as a reorder buffer: when the buffer fills, issue must stop. All the simulation results are reported as for the reorder buffer with bypasses. They apply equally well for the history buffer and future file methods. The selection of a particular method depends not only on its effect on system performance but also the cost of implementation and the ease with which the precise CPU state can be restored.

For each precise interrupt method, two methods were described for handling stores. Simulations were run for each of these methods. For those methods other than the in-order completion method, the size of the reorder buffer is a parameter. Sizing the buffer with too few entries degrades performance since instructions that might issue could block at the issue register. The blockage occurs because there is no room for a new entry in the buffer.

Table 1 shows the relative performance of the In-order, Reorder Buffer, and Reorder Buffer with bypass methods when the stores are held until the result shift register is empty. The results in the table indicate the relative performance of these methods with respect to the CRAY-1S across the first 14 Lawrence Livermore Loops; real CRAY-1S performance is 1.0. A relative performance greater than 1.0 indicates a degradation in performance. The number of entries in the reorder buffer was varied from 3 to 10.

Table 1. Relative Performance for the first 14 Lawrence Livermore Loops, with stores blocked until the results pipeline is empty.

Number of Entries	In-order	Reorder	R w/ BP
3	1.2322	1.3315	1.3069
4	1.2322	1.2183	1.1743
5	1.2322	1.1954	1.1439
8	1.2322	1.1808	1.1208
10	1.2322	1.1808	1.1208

The simulation results for the In-order column are constant since this method does not depend on a buffer that reorders instructions. For all the methods, there is some performance degradation. Initially, when the reorder buffer is small, the In-order method produces the least performance degradation. A small reorder buffer (less than 3 entries) limits the number of instructions that can simultaneously be in some stage of execution. Once the reorder buffer size is increased beyond 3 entries, either of the other methods results in better performance. As expected, the reorder buffer with bypasses offers superior performance when compared with the simple reorder buffer. When the size of the buffer was increased beyond 10 entries, simulation results indicated no further performance improvements. (Simulations were also run for buffer sizes of 15, 16, 20, 25, and 60.) At best, one can expect a 12% performance degradation when using a reorder buffer with bypasses and the first method for handling stores.

Table 2 indicates the relative performance when stores issue and wait at the same memory pipeline stage as for memory bank conflicts in the original CRAY-1S. After issuing, stores wait for their counterpart dummy store to signal that all previously issued register instructions have finished. Subsequent loads and stores are blocked from issuing.

Table 2. Relative Performance for the first 14 Lawrence Livermore Loops, with stores held in the memory pipeline after issue.

Number of Entries	In-order	Reorder	R w/ BP
3	1.1560	1.3058	1.2797
4	1.1560	1.1724	1.1152
5	1.1560	1.1348	1.0539
8	1.1560	1.1167	1.0279
10	1.1560	1.1167	1.0279

As in Table 1, the In-order results are constant across all entries. For the simple reorder buffer, the buffer must have at least 5 entries before it results in better performance than the In-order method. The reorder buffer with bypasses, however, requires only 4 entries before it is performing more effectively than the In-order method. Just as in Table 1, having more than 8 entries in the reorder buffer does not result in improved performance. Comparing Table 1 to Table 2, the second method for handling stores offers a clear improvement over the first method. If the second method is used with an 8 entry reorder buffer that has bypasses, a performance degradation of only 3% is experienced.

Clearly there is a trade-off between performance degradation and the cost of implementing a method. For essentially no cost, the In-order method can be combined with the first method of handling stores. Selecting this 'cheap' approach results in a 23% performance degradation. If this degradation is too great, either the second store method must be used with the In-order method or one of the more complex methods must be used. If the reorder buffer method is used, one must use a buffer with at least 3 or 4 entries.

8. Extensions

In previous sections, we described methods that could be used to guarantee precise interrupts with respect to the registers, the main memory, and the program counter of our simple architectural model. In the following sections, we extend the previous methods to handle additional state information, virtual memory, a cache, and linear pipelines. Effectively, some of these machine features can be considered to be functional units with non-deterministic execution times.

8.1. Handling Other State Values

Most architectures have more state information than we have assumed in the model architecture. For example, a process may have state registers that point to page and segment tables, indicate interrupt mask conditions, etc. This additional state information can be precisely maintained with a method similar to that used for stores to memory. If using a reorder buffer, an instruction that changes a state register reserves a reorder buffer entry and proceeds to the part of the machine where the state change will be made. The instruction then waits there until receiving a signal to continue from the reorder buffer. When its entry arrives at the head of the buffer and is removed, then a signal is sent to cause the state change.

In architectures that use condition codes, the condition codes are state informaton. Although the problem condition codes present to conditional branches is not totally unrelated to the topic here, solutions to the branch problem are not the primary topic of this paper. It is assumed that the conditional branch problem has been solved in some way, e.g. [Ande67]. If a reorder buffer is being used, condition codes can be placed in the reorder buffer. That is, just as for data, the reorder buffer is made sufficiently wide to hold the condition codes. The condition code entry is then updated when the condition codes associated with the execution of an instruction are computed. Just as with data in the reorder buffer, a condition code entry is not used to change processor state until all previous instructions have completed without error (however condition codes can be bypassed to the instruction fetch unit to speed up conditional branches).

Extension of the history buffer and future file methods to handle condition codes is very similar to that of the reorder buffer. For the history buffer, the condition code settings at the time of instruction issue must be saved in the history buffer. The saved condition codes can then be used to restore the processor state when an exception is detected. Since the future file method uses a reorder buffer, the above discussion indicates how condition codes may be saved.

8.2. Virtual Memory

Virtual memory is a very important reason for supporting precise interrupts; it must be possible to recover from page faults. First, the address translation pipeline should be designed so that all the load/store instructions pass through it in order. This has been assumed throughout this paper. Depending on the method being used, the load/store instructions reserve time slots in the result pipeline and/or re-order buffer that are read no earlier than the time at which the instructions have been checked for exception conditions (especially page faults). For stores, these entries are not used for data; just for exception reporting and/or holding a program counter value.

If there is an addressing fault, then the instruction is cancelled in the addressing pipeline, and all subsequent load/store instructions are cancelled as they pass through the addressing pipeline. This guarantees that no additional loads or stores modify the process state. The mechanisms described in the earlier sections for assuring preciseness with respect to registers guarantee that non-load/store instructions following the faulting load/store will not modify the process state; hence the interrupt is precise.

For example, if the reorder buffer method is being used, a page fault would be sent to the reorder buffer when it is detected. The tag assigned to the corresponding load/store instruction guides it to the correct reorder buffer entry. The reorder buffer entry is removed from the buffer when it reaches the head. The exception condition in the entry causes all further entries of the reorder buffer to be discarded so that the process state is modified no further (no more registers are written). The program counter found in the reorder buffer entry is precise with respect to the fault.

8.3. Cache Memory

Thus far we have assumed systems that do not use a cache memory. Inclusion of a cache in the memory hierarchy affects the implementation of precise interrupts. As we have seen, an important part of all the methods is that stores are held until all previous instructions are known to be exception-free. With a cache, stores may be made into the cache earlier, and for performance reasons should be. The actual updating of main memory, however, is still subject to the same constraints as before.

8.3.1. Store-through Caches

With a store-through cache, the cache can be updated immediately, while the store-through to main memory is handled as in previous sections. That is, all previous instructions must first be known to be exception-free. Load instructions are free to use the cached copy, however, regardless of whether the store-through has taken place. This means that main memory is always in a precise state, but the cache contents may "run ahead" of the precise state. If an interrupt should occur while the cache is potentially in such a state, then the cache should be flushed. This guarantees that prematurely updated cache locations will not be used. However, this can lead to performance problems, especially for larger caches.

Another alternative is to treat the cache in a way similar to the register files. One could, for example, keep a history buffer for the cache. Just as with registers, a cache location would have to be read just prior to writing it with a new value. This does not necessarily mean a performance penalty because the cache must be checked for a hit prior to the write cycle. In many high performance cache organizations, the read cycle for the history data could be done in parallel with the hit check. Each store instruction makes a buffer entry indicating the cache location it has written. The buffer entries can be used to restore the state of the cache. As instructions complete without exceptions, the buffer entries are discarded. The future file can be extended in a similar way.

8.3.2. Write-Back Cache

A write-back cache is perhaps the cache type most compatible with implementing precise interrupts. This is because stores in a write-back cache are not made directly to memory; there is a built-in delay between updating the cache and updating main memory. Before an actual write-back operation can be performed, however, the reorder buffer should be emptied or should be checked for data belonging to the line being written back. If such data should be found, the write-back must wait until the data has made its way into the cache. If a history buffer is used, either a cache line must be saved in the history buffer, or the write-back must wait until the associated instruction has made its way to the end of the buffer. Notice that in any case, the write-back will sometimes have to wait until a precise state is reached.

8.4. Linear Pipeline Structures

An alternative to the parallel functional unit organizations we have been discussing is a linear pipeline organization. Refer to Fig. 7.

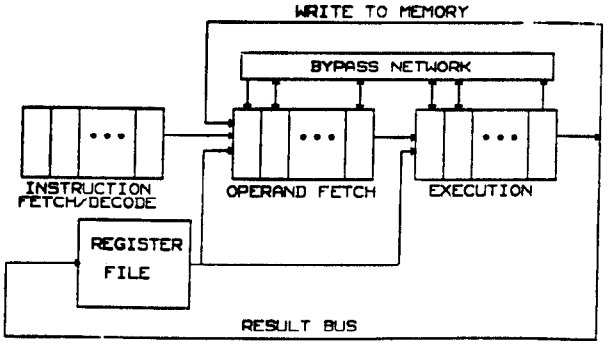


Figure 7. Example of a linear pipeline implementation.

Linear pipelines provide a more natural implementation of register-storage architectures like the IBM 370. Here, the same instruction can access a memory operand and perform some function on it. Hence, these linear pipelines have an instruction fetch/decode phase, an operand fetch phase, and an execution phase, any of which may be composed of one or several pipeline stages.

In general, reordering instructions after execution is not as significant an issue in such organizations because it is natural for instructions to stay in order as they pass through the pipe. Even if they finish early in the pipe, they proceed to the end where exceptions are checked before modifying the process state. Hence, the pipeline itself acts as a sort of reorder buffer.

The role of the result shift register is played by the control information that flows down the pipeline alongside the data path. Program counter values for preciseness may also flow down the pipeline so that they are available should an exception arise.

Linear pipelines often have several bypass paths connecting intermediate pipeline stages. A complete set of bypasses is typically not used, rather there is some critical subset selected to maximize performance while keeping control complexity manageable. Hence, using the terminology of this paper, linear pipelines achieve precise interrupts by using a reorder buffer method with bypasses.

9. Summary and Conclusions

Five methods have been described that solve the precise interrupt problem. These methods were then evaluated through simulations of a CRAY-1S implemented with these methods. These simulation results indicate that, depending on the method and the way stores are handled, the performance degradation can range from between 25% to 3%. It is expected that the cost of implementing these methods could vary substantially, with the method producing the smallest performance degradation probably being the most expensive. Thus, selection of a particular method will depend not only on the performance degradation, but whether the implementor is willing to pay for that method.

It is important to note that some indirect causes for performance degradation were not considered. These include longer control paths that would tend to lengthen the clock period. Also, additional logic for supporting precise interrupts implies greater board area which implies more wiring delays which could also lengthen the clock period.

10. Acknowledgement

One of the authors (J. E. Smith) would like to thank R. G. Hintz and J. B. Pearson of the Control Data Corp. with whom he was associated during the development of the CYBER 180/990. This paper is based upon research supported by the National Science Foundation under grant ECS-8207277.

11. References

[Amdh81] Amdahl Corporation, *Amdahl 470V/8 Computing System Machine Reference Manual*, publication no. G1014.0-03A, Oct. 1981.

[Amdh80] Amdahl Corporation, "580 Technical Introduction," 1980.

[Ande67] D.W. Anderson, F.J. Sparacio, F.M. Tomasulo, "The IBM System/360 Model 91: Machine Philosophy and Instruction-Handling," *IBM Journal of Research and Development*, V 11, January 1967, pp. 8-24.

[Bons69] P. Bonseigneur, "Description of the 7600 Computer System," *Computer Group News*, May 1969, pp. 11-15.

[Buch62] W. Bucholz,ed., *Planning a Computer System*, McGraw-Hill, New York, 1962.

[CDC84] Control Data Corporation, *CDC Cyber 180 Computer System Model 990 Hardware Reference Manual*, pub. No. 60462090, 1984.

[CDC81] Control Data Corporation, "CDC CYBER 200 Model 205 Computer System Hardware Reference Manual," Arden Hills, MN, 1981.

[Cray79] Cray Research, Inc., "CRAY-1 Computer Systems, Hardware Reference Manual," Chippewa Falls, WI, 1979.

[Henn82] J. Hennessy et. al., "Hardware/Software Tradeoffs for Increased Performance," *Proc. Symp. on Architectural Support for Programming Languages and Operating Systems*, April 1982, pp. 2-11.

[HiTa72] R. G. Hintz and D. P. Tate, "Control Data STAR-100 Processor Design," *Compcon 72*, IEEE Computer Society Conference Proc., Sept. 1972, pp. 1-4.

[McMa72] F. H. McMahon, "FORTRAN CPU Performance Analysis," Lawrence Livermore Laboratories, 1972.

[PaSm83] N. Pang and J. E. Smith, "CRAY-1 Simulation Tools," Tech. Report ECE-83-11, University of Wisconsin-Madison, Dec. 1983.

[Russ78] R.M. Russell, "The CRAY-1 Computer System," *Comm. ACM*, V 21, N 1, January 1978, pp. 63-72.

[Stev81] David Stevenson, "A Proposed Standard for Binary Floating Point Arithmetic," *Computer*, V 14 N 3, March 1981, pp. 51-62.

[Thor70] J.E. Thornton, *Design of a Computer - The Control Data 6600*, Scott, Foresman and Co., Glenview, IL, 1970

[Ward82] William P. Ward, "Minicomputer Blasts Through 4 Million Instructions a Second," *Electronics*, Jan. 13, 1982, pp. 155-159.