Techniques for Extracting Instruction Level Parallelism on MIMD Architectures

Gary Tyson and Matthew Farrens
Computer Science Department
University of California, Davis
Davis, CA 95616 tel: (916) 752-9678, fax: (916) 752-4767
email: tyson@cs.ucdavis.edu, farrens@cs.ucdavis.edu

Abstract

Extensive research has been done on extracting parallelism from single instruction stream processors. This paper presents some results of our investigation into ways to modify MIMD architectures to allow them to extract the instruction level parallelism achieved by current superscalar and VLIW machines. A new architecture is proposed which utilizes the advantages of a multiple instruction stream design while addressing some of the limitations that have prevented MIMD architectures from performing ILP operation. A new code scheduling mechanism is described to support this new architecture by partitioning instructions across multiple processing elements in order to exploit this level of parallelism.

1. Introduction

The compiler and code scheduler for a *multi-issue* architecture requires a high degree of sophistication in order to realize the full potential for parallel execution. It must be able to assign independent instructions to operational units in a manner that minimizes the number of cycles in which no instructions can be issued. The task of the scheduler in a multi-issue system is further complicated by the fact that while the latency of operational units and memory remain fixed, the number of instructions that must be scheduled in a period is increased by the width of the issue stage.

Several studies [JoWa89, TjFl70] indicate that compilers using simple scheduling techniques are capable of identifying 2-3 independent instructions per cycle. Other studies [AuSo92, BuYP91] suggest that even more parallelism can be found if the compiler's scheduler is capable of performing extensive code motion.

In this paper, we will present a brief overview of single and multiple instruction stream approaches to multiple issue processor design. We will then introduce the basics of a multiple instruction stream/multiple issue architecture we have developed, show how code is scheduled for several loops, and present an analysis of the performance of this architecture.

2. Multiple instruction issue architectures

Several distinct approaches have been taken in the development of multiple issue architectures. The *Very Long Instruction Word (VLIW)* approach increases the resources available to (and the demands on) the compiler. It is responsible for *all* scheduling, including the assignment of null operations to functional units that cannot be assigned a useful task during a given cycle. VLIW advocates believe that since the compiler has the most complete information about the entire program, it is best suited to deal with the inclusion of additional resources (e.g. ALUs, FPUs and I/O units), and is best able to increase instruction execution bandwidth in areas of the code that were previously performance limited by resource constraints. However, VLIW does not support out-of-order execution, and any change of the hardware description requires all code to be recompiled in order for the program to work correctly.

Superscalar advocates believe that this backward code incompatibility is unacceptable, and that a hardware scheduler can just as efficiently allocate resources to the list of instructions ready for execution. Originally, some researchers believed that a hardware based scheduler was capable of generating an efficient schedule of resources *independent* of the original compiler schedule. However, the hardware implementation of the scheduler is restricted to selecting instructions from a fixed-size window of available instructions, and thus does not have the breadth of information available to it that the compiler does at compile time. This lack of knowledge prevents it from scheduling code as efficiently as a VLIW. As research into these two techniques has progressed, advocates of each approach have learned to appreciate the interaction of the static, compiler induced schedule with the more dynamic capabilities found in a runtime analysis of resource usage.

A third approach to issuing multiple instructions takes advantages of the characteristics found in the Von Neuman computational model. *Decoupled* architectures attempt to exploit the independent nature of *control flow*, *memory access* and *data manipulation* operations that comprise conventional computations by splitting a task into distinct pieces and executing them on separate pieces of hardware. Since these hardware units communicate via FIFO queues, the instruction streams are allowed to slip with respect to one another, providing dynamic support for out-of-order execution. This approach attempts to take advantage of the best that VLIW and superscalar have to offer; the compiler partitions the tasks in a manner similar to VLIW, and the queues provide the same dynamic

scheduling benefits found in superscalar.

These decoupled systems differ from VLIW and superscalar designs in the manner in which the independently issued instructions interact. VLIW and superscalar processors can be thought of as very tightly coupled shared memory systems; they share not only addressable memory but also register space. This shared register approach differs from the explicit message passing (via FIFO ordered queues) found in decoupled machines. Furthermore, in order to transmit data among operational units by writing and then reading the contents of a register, the clocks on VLIW and superscalar processors must be synchronized. This requirement is relaxed with an explicit message passing approach. [Smit82]

The greater flexibility found in a decoupled design allows both single and multiple instruction stream descriptions of a task. The **ZS-1** [SDVK87] and **WM** [Wulf92] systems operate in a decoupled manner while receiving instructions from a single instruction stream. Their architectural component descriptions are similar to those of *Split Register* superscalar designs [Site93, SCHP92]. The **PIPE** machine, in contrast, [GHLP85] consists of two PIPE processors [CGKP87] which run asynchronously, each with their own instruction stream, and cooperate on the execution of a single task.

3. Exploiting ILP on a MIMD architecture

Parallelism in a single instruction stream architecture resides primarily at the instruction level, and is a well-studied problem [JoWa89, Wall91]. Extracting parallelism on a MIMD architecture, on the other hand, has traditionally been accomplished by partitioning the program into data independent portions and assigning them to separate processing elements, ignoring any other parallelism that might exist. Little research has been done on exploiting instruction level parallelism across processors on a multiple instruction stream machine.

There are a number of reasons why this approach merits further investigation, however. Superscalar machines do not scale well - expanding the number of processing elements available necessitates a corresponding increase in the size of the hardware window over which code scheduling occurs, significantly increasing the scheduling complexity. Compilers for VLIW machines can help circumvent this problem, but do not support out-of-order execution well.

Exploiting instruction level parallelism on MIMD architectures can overcome both these problems. The instruction issue stage of each processor can perform in a simple single-issue, in-order manner, avoiding much of the hardware complexity required to support out-of-order issue in a single instruction stream approach. Out-of-order issue is also supported on a MIMD because the processors are run independently; therefore, any independent instructions executed on different processors can issue in any order without necessitating any hardware support. This is fundamentally different than multiple issue in a VLIW machine because a strict ordering of instructions is not imposed by the compiler unless a dependence exists. Furthermore, by incorporating multiple program counters, a MIMD machine provides the architecture with more dataflow information by enriching the specification of the object language; taken to its extreme this would allow a dataflow machine description of the program.

Separating a program into multiple single issue instruction streams additionally allows the decentralization of the hardware resources, since there is no central instruction window from which instructions are issued. Similarly, there is no central register file to be overloaded with contention among the processing elements, which allows for easier expandability in a MIMD approach.

While a MIMD approach to code scheduling clearly possesses certain advantages, historically these architectures have suffered from severe limitations. Data transfer latencies have been high, and the bandwidth required to support high-throughput, low contention data transfer has been unavailable because of pin limitations and/or board-level interconnects. Even if maximum data transfer rates can be made acceptable, the need to provide synchronization points can cause unacceptable performance loss. Using main memory to handle data transfers between processors can also lead to an unacceptable dependence on memory latency. These problems help explain why current MIMD designs do not exploit ILP.

Increasing the number of transistors that can be fabricated per square centimeter provides the means by which many of the interprocessor communication problems can be eliminated. Placing several of these processing elements on the same die circumvents the pin limitations on bandwidth, and supports high on-chip data transfer rates. In addition, using FIFO queues in a manner similar to that used by decoupled machines provides a clean way to handle synchronization. These facts led to the design of the MISC architecture, a decoupled MIMD machine that is designed to support and exploit instruction level parallelism.

4. The multiple instruction stream computer (MISC)

The MISC architecture was designed to handle many of the dynamic characteristics of program execution by allowing the compiler to convey more information to the hardware during code translation. Variable operational unit latencies (primarily memory loads) create difficulties for code scheduling in VLIW and superscalar processors, due to the sequential instruction flow imposed by translating a dataflow intermediate representation to a single instruction stream architecture. Superscalar designs can remove some of the restrictions imposed by single stream scheduling by regenerating some the dataflow information at the issue stage of the pipeline, but not without considerable hardware issue logic. Furthermore, software pipelining [Lam88] and loop unrolling schemes

[WeSm87]

have difficulty in efficiently scheduling instructions with variable latency dependencies.

MISC avoids these scheduling problems by allowing operations with indeterminate latencies to transfer data between PEs. The inherent asynchronous relationship among the PEs can compensate for the variability of the latency without affecting the execution rate of non-dependent instructions.

The MISC processor has been described in detail in [TyFP92]. A brief overview of MISC will be presented here, focusing on aspects of the architecture that will be featured in the code scheduling discussion later in the paper. MISC is a direct descendant of the PIPE project, but unlike the two processor PIPE design, the MISC system is capable of balancing the processor load of instructions performing control flow, memory access and execute operations among multiple processors. As its name indicates, MISC is composed of multiple *Processing Elements* (PEs) which cooperate in the execution of a task.

The example MISC configuration used throughout this paper consists of four processing elements, a bank selected data cache (DCache) and a set of internal data paths used to transmit data among PEs and the DCache. The component design of this MISC configuration is illustrated in figure 1. Each PE executes in an asynchronous manner from other PEs and the DCache. The internal data paths are used to facilitate communication between elements (PEs and/or DCache). Each data path is controlled by a single element; for instance, the internal data path labeled PBUS1 is controlled (written) solely by PE1. Each PE has its own bus (PBUS{1-4}), and the data cache controls two busses (CBUS1 and CBUS2). Each PE is capable of transmitting a message directly to any other processor (including itself), or of broadcasting a message to all processors.

The processing elements (figure 1) are collectively responsible for the execution of a single task, with each PE having its own independent instruction stream and its own instruction cache. Each PE is identical and maintains all state information required to function as an independent processor - in fact, the MISC hardware is capable of running four completely unrelated tasks in parallel. However, is assumed that a single task will be partitioned (by the compiler) into four instruction streams that cooperate in the execution of that task. Each PE contains a 5-stage pipeline, 32 General Purpose Registers (GPRs) which are available for data storage that persists over multiple references, a FIFO processor queue (PQ) for each PE in the system (including itself) to store data transfers between PEs, a *Program Counter* (PC), a *Vector Register* (VREG), and 2 memory queues (MQs) which contain data requested from memory. The size of each of the queues (PQs and MQs) is not given here, however their size is an architecturally visible component; the compiler must know the size of each queue (they need not all be the same size) in order to schedule code correctly and avoid deadlocks due to resource depletion.)

All instructions in MISC are 32 bits in length and have three 6-bit source operand fields and a 6-bit destination operand field. In addition, many instructions allow for two of the source fields to be replaced by a 12-bit constant. Each source field can address any of the 32 GPRs, a PQ, a MQ, the PC, the VREG, or a small signed constant (-16 to 15). The destination specifier may address a

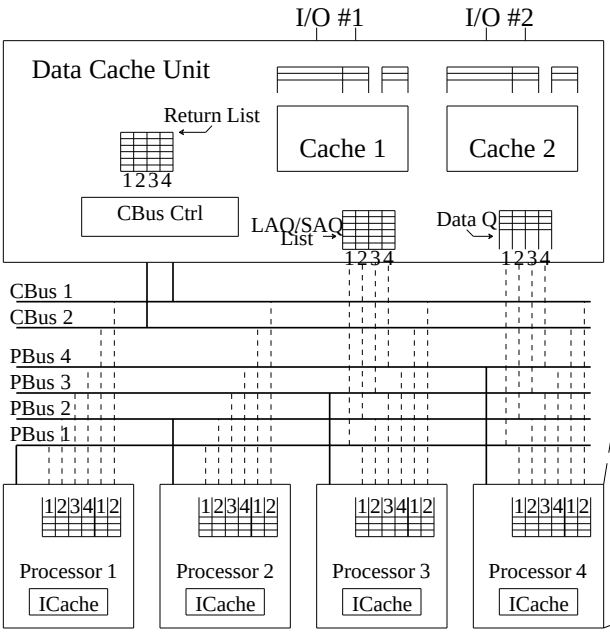
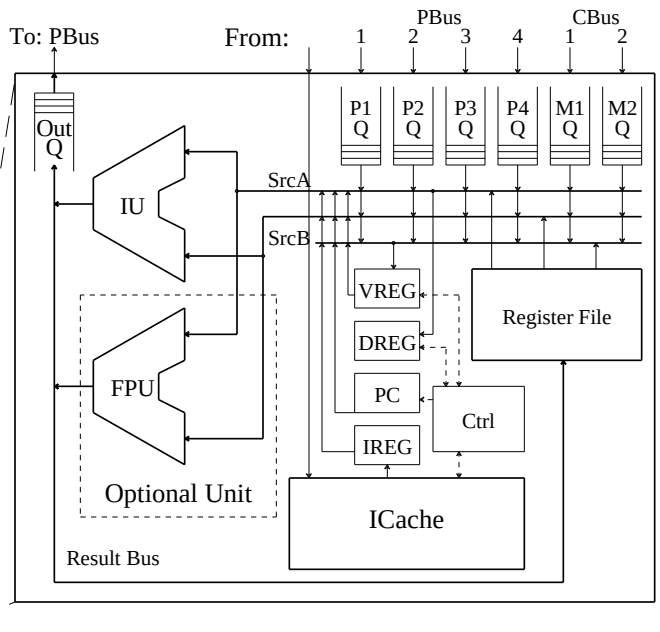


Figure 1: MISC Machine



MISC Processing Element

GPR or *routing* information for a data transfer onto the PE's PBUS. At the instruction issue stage, if a queue is specified as a source input and that queue is currently empty, that instruction is delayed until all required input operands are available.

There are three types of MISC instructions: predicated operations [AKPW83], vector operations, and sentinel operations (which use a sentinel value to terminate iterations of the instruction). Predicated ALU/FPU operations perform all scalar operations as well as allow conditional operations to be specified concisely. A predicate operation uses the third source field to determine whether or not the operation will complete (and thereby change the state of the machine). This allows the issue logic to proceed without interrupt through short segments of conditionally executed code by conditionally completing instead of branching around code.

In the case of control flow operations, the *dest* field is used as a constant to determine the number of *delayed branch* slots [FaPl91] to be filled. The address of the branch is calculated as the sum of the *src1* and *src2* operands, and the *src3* operand specifies the register to be tested.

Vector instructions use the third source operand (*src3*) to specify a vector count. When a vector instruction arrives at the issue stage of the instruction pipeline, the vector register (*VREG*) is cleared and a vector count register (*VCOUNT*) is loaded from *src3*. The scalar version of the vector instruction is then executed and the *VREG* is incremented until the contents of *VREG* are equal to *VCOUNT*. Once the *VREG* equals *VCOUNT* normal instruction pipeline function continues.

In a sentinel instruction, the *src3* field specifies a register whose contents are compared to the sentinel value (assumed to be zero in the initial design). If the contents of the register do not match the sentinel, the scalar version of the instruction is allowed to issue. This pattern of compare and issue is repeated until the comparison produces a match.

5. Scheduling code on a MISC

The MISC compiler makes use of existing optimization techniques, both to simplify its construction and to provide a comparable compilation environment for performance evaluation of the architecture. For those optimization that are unique to MISC, or where existing techniques require modification (e.g. register allocation incorporating queues), care has been taken to maintain the same level of complexity found in current optimizers. The *very portable C compiler* (vpcc) [BeDa91] under development at the University of Virginia is used as the base compiler for MISC.

The code generator translates a dependency graph representation of a program (in *Register Transfer List* **(RTL)** form) to produce parallel machine code. This translation has four phases: code separation, processor load balancing, loop translation and list scheduling. The code separation phase partitions the operations required by the program onto multiple (virtual) processing elements in a manner that subsumes the effects of memory and operational unit latencies. Processor load balancing then repartitions the schedule to evenly distribute the operations on the number of physical processing elements available on the target machine. The list scheduling phase then generates the machine instructions for each of the processing elements. A detailed description of this process follows in the remainder of this section.

5.1. Code separation

The task of the code separator is to partition the task across processing elements, with the goal of minimizing the effects of high memory latency and high functional unit latency for operations like multiply and divide by decoupling the *creation* of the data item from its *use*. Much like initial register allocation strategies, code separation assumes an infinite number of processing elements. The mapping of operations to the physical processing elements available on the target architecture is left to the processor load balancing phase.

The partitioning algorithm processes the dependency graph from the *root* (first operation) to the *leaves* (dependent operations). Information concerning operational latency and register use is examined and operations are assigned to processing elements in a bottom-up or dependent-first order. This leads to a reverse schedule in which the final operations for a block are scheduled first and those instructions that use data items that are not available (due to operational latencies) are scheduled on a different processing element.

Branching is handled slightly differently. Branch instructions are duplicated across all processing elements to ensure that each processing element conforms to the same control flow, and all instructions that calculate branch conditions are assigned to the processing element that has no instructions containing data dependencies; this allows a single processor to *lead* the execution.

5.2. Using leading processes to hide latency

The concept of a *leading* processing element is central to the understanding of code separation. In a MIMD architecture, each of the instruction streams executes independently of the other (ignoring for a moment any data dependencies). If operations are scheduled carefully, we can affect the relative execution of the streams, allowing some of the streams to proceed farther ahead in the computation than others. Staggering the relative entry cycles for the execution of a section of code provides a perfect method for hiding the delay imposed by high latency operations. For example, if the instruction that issues the high latency operation is scheduled on a processor that enters that section of code a sufficient number of cycles before the processor that uses the item, the effects of the latency will be hidden. In such a case it is possible for the *leader* PE will be executing instructions

in a new section of code while *trailing* PEs are still completing previous sections.

To illustrate each phase of the code generation process, a simple example (Lawrence Livermore Loop 3) will be used.

The RTL representation of the intermediate code prior to code scheduling appears in figure 3. Each line of the RTL description either defines a label or describes an operation to be performed in the resulting code. Each RTL line shown throughout this section will contain a comment (delimited by ';') to explain its operation. Virtual register labels (specified as t1, t2, t3, etc.) define intermediate points in the calculation and may or may not map to physical registers or queues. To avoid confusion the actual register mapping has been omitted.

As referred to previously, the code scheduler can use the knowledge that basic blocks are entered by successive processing elements on different cycles to hide the latency of long latency instructions. For example, if an instruction with a completion latency of 5 cycles can be issued by a leading PE with the result destined for a PE trailing by 5 or more cycles, the affects of the operational latency are completely subsumed.

A simple approach the scheduler can take is to assume a fixed latency period between processing element block entries and enforce a strict ordering on PE leadership. This approach can be used to get reasonable performance from the scheduler; however, greater benefit can be found from calculating the expected *stagger* in relative basic block entry/exit cycles. This can be accomplished by augmenting the standard dataflow analysis gathered in the code separation phase with expected completion times for each of the processing elements in the virtual machine.

```
int inner_product() {
        int k, q=0;
        for (k=0; k<1024 ; k++)
                q = q + z[k] * x[k];
}
```

Figure 2: Livermore Loop 3 (Inner Product)

```
[1]   t1 = 0              ; q=0
[2]   t2 = 0              ; k=0
[3]   t3 = 1024           ; set register for test
[4]   t4 = LOC[_z]        ; t4 = base of array z
[5]   t5 = LOC[_x]        ; t5 = base of array x

[6]   L1:
[7]   t6 = (t2>=t3)       ; calculate branch cond
[8]   PC = t6, L2         ; branch if true
[9]   t7 = M[ t4+t2 ]     ; load t7= z[k]
[10]  t8 = M[ t5+t2 ]     ; load t8= x[k]
[11]  t9 = t7 * t8        ; (z[k] * x[k])
[12]  t1 = t1 + t9        ; q = q + (z[k] * x[k])
[13]  t2 = t2 + 1         ; k++
[14]  PC = L1
[15]  L2:
```

Figure 3: RTL Representation of LLL3 prior to Code Separation

It is possible that a constant value for the entry time stagger will be grossly erroneous; this is the case when a loop containing a recurrence is scheduled such that only one processing element is delayed by a recurrence relation while the remaining processing elements proceed to the next iteration (and finally to the loop completion). In this situation, the processing element delayed by the recurrence (and any processing elements dependent on that one) will enter the basic block following the loop termination many cycles later than the lead processing elements. If the code scheduler can recognize when this occurs, a normally less efficient schedule can be created for the basic block succeeding the loop that avoids any use of the processing elements that are lagging far behind. This has the effect of concurrently executing the successor blocks and the loop containing the recurrence. The MISC code scheduler incorporates this augmentation to the dataflow analysis and the benefits of this approach are discussed in the analysis in section 6.

In order to determine which operations should migrate to a new processing element, the scheduler locates all data dependencies that have latencies that can not be hidden with a simple reordering of instructions. From this list of candidate operations, ones that are involved in a dependency circuit are scheduled to a single processing element. This tends to negate the transmission time penalties between processing elements as well as greatly simplifying the deadlock detection scheme in the code generator. Unfortunately, migrating the operation to another processing element may create new hardware dependencies, which may conflict with others. For example, if migrating a multiply operation requires that the destination for the product be a FIFO ordered queue (since queues are used to transfer data between processing elements), a conflict may arise if that queue is already allocated to a previously existing instruction. In these cases, the migration cannot proceed unless that conflict can be resolved in a later optimization step.

This separation tends to identify operations by function: *flow control*, *memory access* or *data manipulation*, with the leader PE performing control flow operations, and trailing PEs supporting memory access and data manipulation. This creates a dependency between the PE that generates the data and a trailing PE which consumes the data, which tends to separate code into memory access and data manipulation functions due to the long latency of memory loads. A graph of these inter-PE dependencies is constructed to aid in the scheduling process, and the scheduler attempts to avoid circuits in the dependency graph to simplify deadlock elimination and to ensure that the ordering of PE execution is maintained (i.e. PE

leader-ship does not transfer).

Code separation provides a model which is capable of balancing PE loads in applications that contain little balance between memory access and data manipulation. It also provides the flexibility to handle code generation for a variable number of processing elements. It should be noted that unlike more coarsely grained parallel computations, the separation of instructions on MISC tends to be between dependent operations.

Applying the code partitioning strategy to the example code in figure 3 results in the specification in figure 4. A second column has been inserted after each line number to indicate which processing element(s) process this RTL line. For instance, the first line ([1]) initializes the variable **q** to zero. Since the only use of **q** is in code allocated to PE3, the initialization of **q** is allocated to PE3. Notice also that branch operations (lines [8] and [14]) are allocated across all active processing elements. The separation of operations occurs in lines [9] through [12]. Since line [12] has a high latency dependency (due to the multiplication) to line [11] they are partitioned to different PEs. Similarly, the memory latency between lines [9] and [11], and lines [10] and [11] require a third PE to be included in the schedule. Lines [9] and [10] can be issued to the same PE because no dependency exists between them; loop control variable calculation also are issued to the lead PE.

5.3. Processor load balancing

The schedule generated by the code separator is incomplete in two areas. First, the partitioning is performed on virtual processors, which may not equal the actual number of physical processors. Second, no consideration has been given to equally distributing operations among the processors. The goal of the load balancing phase is to remedy these deficiencies.

In the example program (figure 4) the code separation phase allocates only three processing elements while the physical target machine contains four processing elements. Therefore, the two independent memory operations (the vector loads for *x* and *z*) are split onto two processing elements. This leads to a schedule that utilizes the full capabilities of the target architecture. The modified schedule is show in figure 5.

```
[1]  [PE3]    t1 = 0              ; q=0
[2]  [PE1]    t2 = 0              ; k=0
[3]  [PE1]    t3 = 1024           ; set register for test
[4]  [PE1]    t4 = LOC[_z]        ; t4 = base of array z
[5]  [PE1]    t5 = LOC[_x]        ; t5 = base of array x

[6]  [PE1-3]  L1:
[7]  [PE1]    t6 = (t2>=t3)       ; PE1 calcs branch cond
[8]  [PE1-3]  PC = t6, L2         ; branch if true
[9]  [PE1]    t7 = M[ t4+t2 ]     ; load t7= z[k]
[10] [PE1]    t8 = M[ t5+t2 ]     ; load t8= x[k]
[11] [PE2]    t9 = t7 * t8        ; (z[k] * x[k])
[12] [PE3]    t1 = t1 + t9        ; q = q + (z[k] * x[k])
[13] [PE1]    t2 = t2 + 1         ; k++
[14] [PE1-3]  PC = L1
[15] [PE1-3]  L2:
```

Figure 4: RTL for LLL3 after Code Separation

5.4. Loop translation

While code separation generally achieves excellent results in code containing no branches, greater benefit can be achieved by devising optimizations that examine the more general nature of control flow. Of particular importance is the optimization of natural loops. Two optimization techniques are applied to loops in this compiler - *branch reduction* and *induction variable calculation*. These two optimizations attempt to eliminate instructions in inner loops, which can lead to significant performance improvements when these loop iterations account for a large portion of the execution time.

5.4.1. Branch reduction

One ramification of having multiple processors cooperating on the execution of a task is that many more branch instructions are required in order to keep the instruction flows synchronized. This duplication of branch instructions in each stream can lead to a significant increase in the total number of instructions required to perform a task.

The MISC architecture provides two mechanisms to reduce the need for branch duplication: VLOOP/SLOOP instructions and predicated execution. The *vector loop* (VLOOP) instruction uses the vector register in conjunction with a *Delay Register* (DREG) to realize a very simple branch hiding (or zero cycle branch) instruction. It is used in cases where the number of times a basic block will execute is known at the initial entry into

```
[1]  [PE4]    t1 = 0              ; q=0
[2]  [PE1]    t2 = 0              ; k=0
[3]  [PE1]    t3 = 1024           ; set register for test
[4]  [PE1]    t4 = LOC[_z]        ; t4 = base of array z
[5]  [PE2]    t5 = LOC[_x]        ; t5 = base of array x

[6]  [PE1-4]  L1:
[7]  [PE1]    t6 = (t2>=t3)       ; PE1 calcs branch cond
[8]  [PE1-4]  PC = t6, L2         ; branch if true
[9]  [PE1]    t7 = M[ t4+t2 ]     ; load t7= z[k]
[10] [PE2]    t8 = M[ t5+t2 ]     ; load t8= x[k]
[11] [PE3]    t9 = t7 * t8        ; (z[k] * x[k])
[12] [PE4]    t1 = t1 + t9        ; q = q + (z[k] * x[k])
[13] [PE1-2]  t2 = t2 + 1         ; k++
[14] [PE1-4]  PC = L1
[15] [PE1-4]  L2:
```

Figure 5: RTL for LLL3 after Load Balancing

the loop (either as a constant or register variable).

The *sentinel loop* (SLOOP) instruction provides branch hiding capability to more generalized (*while*) loops. SLOOP operates in a manner similar to the VLOOP except that a sentinel comparison is made instead of a VREG calculation. When a SLOOP instruction reaches the issue stage of the pipeline, the *src3* operand is tested to see if the first iteration of the loop should be executed. The *src3* operand specifier is then saved to allow for additional sentinel tests to be performed during the last instruction issued each time through the loop body. Loop iteration continues until the sentinel marker is reached.

Conventional wisdom holds that instruction level implementations of higher level semantics seldom lead to performance improvements because of the complexity of implementation and the scarcity of application. While this may be true for a standard sequential processor, the vector and sentinel instructions defined in MISC can be implemented with minimal hardware modification to the issue logic and the existence of multiple instruction streams leads to a greater potential for application of these instructions. Often the application of a complex construct (e.g. a *while* loop) is virtually impossible in a single instruction stream design because of the complexity of evaluating the test condition. The need to both evaluate a complex test condition and perform the control flow operation cannot be reduced to a single instruction. However, in a multiple instruction stream machine such as MISC, a single PE can evaluate the test condition and broadcast the boolean result to all PEs, increasing the number of simple boolean tests evaluated during the execution of these loop semantics.

Hyperblock transformation [MLCH92] also reduces the effects of branching by eliminating branching operations in favor of predicated execution. One problem with hyperblock scheduling is that while it may be useful to eliminate a branch to allow more efficient use of some operational units (e.g. load/store operations), the necessity of translating all instructions effected by the branch into a predicated form can lead to an overall decrease in execution performance. Hyperblock scheduling of MISC code allows for PEs to be transformed in an independent manner. This allows PEs that can benefit from the removal of a branch operation to proceed with the hyperblock transformation while other PEs may more efficiently schedule instructions by retaining that branch. This also allows the application of VLOOP and SLOOP on a greater number of inner loops (containing small segments of conditionally executed code), since loop operations cannot contain branches, but can contain predicated instructions.

During branch reduction, a variant form of hyperblock transformation is used to convert jumps (primarily around short segments of conditionally executed code) into predicated instructions. After this has been accomplished, each inner loop is examined to determine whether the VLOOP or SLOOP operations can be inserted in place of the explicit branching operations. For loops containing no remaining branch operations (other than a single branch exit and the backward branch at the end of the loop), the VLOOP or SLOOP operator is inserted.

In the example of figure 5 it is a simple matter to determine that a VLOOP operation provides the required loop control operation. The effects of applying the loop translation along with induction variable calculation can be seen in figure 6.

5.4.2. Induction variable calculation

An induction variable is a variable whose value is consistently modified (incremented or decremented) by a constant value on each iteration of a loop. These variables are often used to determine the number of iterations. Furthermore, induction variables are often used to index array data items or manipulate memory pointers. Induction variables can be defined in terms of an *induction expression*. While a number of expressions are possible, a useful induction expression is:

$$IV_{(1)}=dee, \qquad IV_{(i+1)}=IV_{(i)} + cee \text{ for } all\ i > 1$$

where **i** is the iteration count (value 1 on the first iteration). The detection of induction variables is a well understood problem. The algorithm used in this compiler is derived from [AhS] (Algorithm 10.9).

Once the control state of the machine has been modified to support loop operations, it is a simple modification to handle the calculation of induction variables used in the loop. The *src1* and *src2* fields of the loop instructions are free to contain the **cee** and **dee** values; **VREG** will maintain the induction value and *src3* will control loop termination as described above.

In the example in figure 5 both array index calculations can be performed by the hardware. This leads to the RTL description after loop translation show in figure 6.

5.5. List scheduling

List scheduling refers to a class of code scheduling algorithms that seek to generate a near optimal schedule for a set of instructions available for immediate execution. A least cost schedule is developed that attempts to schedule all instructions in the shortest time. List scheduling on MISC operates on each of the processing element individually, scheduling to avoid wasted cycles (due to latency). Simple list scheduling is complicated by the necessity to interpret queue register specifications in the RTL and avoid reordering queuing operations. Furthermore, the implementation of loop operations is left to this phase of the code generation. In the example the VLOOP operations for each of the processors can be replaced with the vector version since each loop consist of a single instruction and the default induction calculation is used (or no induction variable is referenced in PE3 and PE4). A flow graph representation of the MISC

```
 [1]  [PE4]    t1 = 0                  ; q=0
 [2]
 [3]  [PE1-4]  t3 = 1024               ; set register for test
 [4]  [PE1]    t4 = LOC[_z]            ; t4 = base of array z
 [5]  [PE2]    t5 = LOC[_x]            ; t5 = base of array x

 [6]
 [7]
 [8]  [PE1-4]  VLOOP 1,0,t3            ; cee = 1, dee = 0
 [9]  [PE1]    t7 = M[ t4+VREG ]       ; load t7=z[k]
[10]  [PE2]    t8 = M[ t5+VREG ]       ; load t8=x[k]
[11]  [PE3]    t9 = t7 * t8            ; (z[k] * x[k])
[12]  [PE4]    t1 = t1 + t9            ; q = q + (z[k]*x[k])
[13]
[14]  [PE1-4]  VLOOP_END
[15]  [PE1-4]  L2:
```

Figure 6: RTL for LLL3 after Loop Translation

object code is shown in figure 7.

6. Analysis

The Lawrence Livermore Loops were selected as the benchmarks, because they are amenable to hand-coding, and are representative of a large class of scientific programs. The first 12 loops were compiled for both the MIPS and MISC architectures. The MIPS code was compiled using the cc compiler with optimization -O2, and the MISC code was generated using the IAGO [TySF93] compilation environment using the techniques discussed in section 5 of this paper. Operational latencies for MISC were set as follows:

Memory Loads	**10**	Branches	2
Integer Add	**2**	Integer Mult	5

In order to compare the performance of MISC to the MIPS processor we examined the total number of cycles required to complete (**Cmplet**) a given loop and the number of cycles required before the execution of instructions after the loop can start (**Next**). The **Next** number is interesting because, since MISC PEs operate independently, one PE can complete it's work on a given loop and begin execution of the code following the loop prior to the official loop termination.

As can be seen in table 1, for a majority of the loops we see a three to four-fold decrease in the cycles required by the MISC machine over the MIPS processor in completion of a loop. This demonstrates that MISC is effectively extracting the parallelism available in the benchmark. In several of these benchmarks there is less of a performance increase (most notably in LLL6 and LLL11); this is due to a recurrence constraint found in the data manipulated by the loop. In these cases, adding more processors will not increase performance regardless of the approach used, since the parallelism is simply not available in the loop.

To provide a comparison with a similarly configured single instruction stream/multiple issue architecture, the loops were also hand compiled for a four-issue VLIW architecture based upon the version found in [RaS92]. This VLIW machine allows four instructions to be issued per clock cycle, and places no limitations on the type of instructions that can be issued. Furthermore, it assumes sufficient resources (e.g. register transfer bandwidth) to sustain a four instructions per cycle execution rate. The "ideal" entry in table 2 is derived by determining the total number of instructions required to complete the program loop, exclusive of branches (which can be removed by software or hardware techniques in any

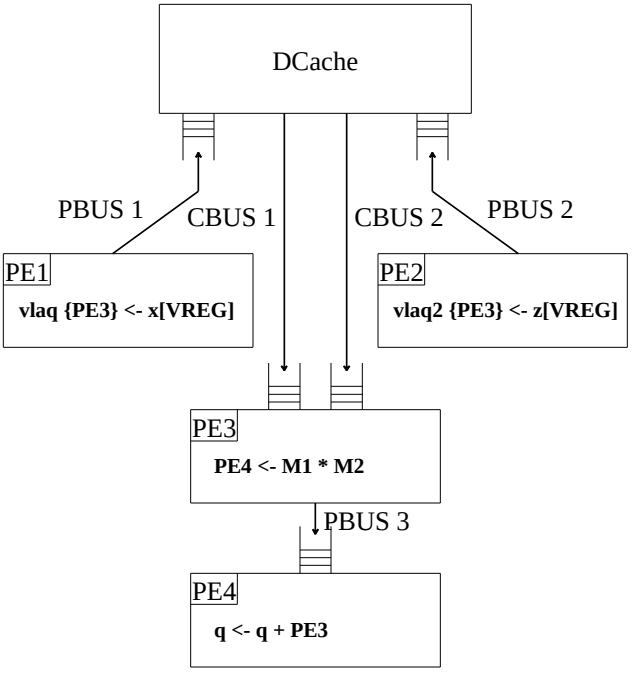


Figure 7: Execution flow for LLL3

Table 1. Comparison of MIPS and MISC cycle counts for Livermore Loops

Bnch	**MIPS**	MISC		Improvement	
		Cmplet	Next	Cmplet	Next
LLL1	5611	1232	1205	4.55	4.65
LLL2	1112	256	201	4.34	5.53
LLL3	6664	2063	1025	3.23	6.50
LLL4	3011	753	385	3.99	7.82
LLL5	6979	1994	977	3.50	7.14
LLL6	7726	4982	0	1.55	∞
LLL7	4338	859	727	5.05	5.97
LLL8	3218	1476	586	2.18	5.49
LLL9	4081	813	609	5.02	6.70
LLL10	3107	1007	506	3.08	6.14
LLL11	3049	2003	0	1.52	∞
LLL12	3759	1013	1002	3.71	3.75

ideal machine). Barring any recurrence relations, the total instruction count is then divided by the issue bandwidth (four in this analysis).

The results in table 2 show that the MISC approach and the VLIW model are capable of extracting about 80% to 99% of the instruction level parallelism available in these loops. However, as mentioned previously, the MISC PEs that finish prior to the overall completion of the loop are available to begin execution of the code following the loop exit. This demonstrates an important point in the performance capabilities of a multiple instruction stream processor; the MISC processor requires only those processing elements necessary to perform the task to be allocated to the loop while the unallocated processing elements can proceed into the following code blocks. In contrast, all functional units in the VLIW processor are locked into the loop (even if they have nothing to do).

A superscalar design might be capable of allocating processing resources across loops, but only with a sufficiently large instruction window and the ability to correctly predict many branches ahead in the instruction stream. The MISC approach of separating instruction streams alleviates these requirements.

To demonstrate the effects of this splitting of resources let us examine two of the loops in more detail. If we look at the execution of LLL6 we notice that only the final processing element is required to perform the majority of the loop calculation. This is due to a tight recurrence relation found in the loop equation. In the VLIW machine all functional units are forced to sit idle in the loop body until the machine (as a whole) completes calculation of the loop. In the MISC approach, the three processing elements not involved in the recurrence calculation are free to continue execution.

If we now assume that LLL11 follows the execution of LLL6, we can determine the different *stagger* rates on exit from LLL6 and reschedule LLL11 to take advantage of the free processing elements. Table 3 shows the result of this rescheduling (done at compile time) and compares it to the VLIW architecture. As seen in the table, the ability to overlap execution of the loops allows the MISC processor to perform both loops in the time required by the VLIW architecture to perform the first alone.

Table 2. Comparison of cycle counts for MISC, VLIW and Ideal Machines

Bnch	**PE1**	**PE2**	**PE3**	**PE4**	**MISC**	**VLIW**	Ideal
LLL1	1205	1215	1221	1232	1232	1236	1000
LLL2	201	201	211	256	256	228	200
LLL3	1025	1025	1035	2063	2063	2070	2048
LLL4	385	395	404	753	753	771	576
LLL5	997	999	1993	4982	4982	4984	4980
LLL6	0	997	1995	4982	4982	4984	4980
LLL7	727	846	736	859	859	863	780
LLL8	586	720	1240	1476	1476	----	950
LLL9	609	707	712	813	813	708	700
LLL10	506	506	1006	1007	1007	1014	750
LLL11	0	999	1000	2003	2003	2004	1998
LLL12	1002	1012	1013	1013	1013	1013	1000

Table 3. Comparison of cycle times for MISC and VLIW executing two loops sequentially

Loop	**PE1**	**PE2**	**PE3**	**PE4**	**MIPS**	VLIW	Improvement	
6	0	997	1995	4982	4982	4984	∞	1.00
11	999	1000	2003	0	2003	2004	∞	1.00
6 - 11	999	1997	3998	4982	4982	6988	6.99	1.40

We expect this final result to be demonstrative of the advantage that the multiple instruction stream attains across basic blocks. Little dataflow analysis is required to achieve this capability.

7. Conclusions/Future work

In this paper we have examined the feasibility of using a MIMD approach to extracting instruction level parallelism. Current MIMD architectures suffer from various deficiencies which prevent their direct application to instruction level parallel tasks. A new architecture has been proposed which alleviates these deficiencies and provides both reduced hardware complexity and simplified software scheduling compared to conventional (single instruction stream) approaches. When supported with a new code scheduling method, this architecture provides performance equivalent to the most powerful VLIW and Superscalar architectures proposed, while maintaining simple hardware and software schemes.

The ability to specify more information in the object language of the MISC machine (by explicitly defining separate instruction streams) simplifies the hardware mechanisms required to support out-of-order execution. This provides a powerful combination of software based control flow optimizations with the dynamic features found in out-of-order execution models in a more cooperative way than that found in existing VLIW and Superscalar designs, by increasing the information content between the two.

We believe these results will prove to be even more significant in non-vectorizable code, because of the latency hiding effects inherent in the decoupled model. We are currently refining the compilation environment, in order to examine a much wider range of benchmark programs.

8. Acknowledgements

This work was supported by the National Science Foundation under Grant MIP-9257259, and by a generous donation from SUN Microsystems.

9. References

[AhS] A. V. Aho, R. Sethi and J. D. Ullman, ‘‘Compilers Principles, Techniques and Tools’’, *Addison-Wesley Publishing*, pp. 644.

[AKPW83] J. R. Allen, K. Kennedy, C. Porterfield and J. Warren, ‘‘Conversion of control dependencies to data dependencies’’, *Proceeding of the 10th ACM Symposium on Principles of Programming Languages*(January 1983), pp. 177-189.

[AuSo92] T. Austin and G. Sohi, ‘‘Dynamic Dependency Analysis of Ordinary Programs’’, *Proceedings of the 19th Annual Symposium on Computer Architecture*, vol. 20, no. 2 (May 19-21, 1992), pp. 342-351.

[BeDa91] M. E. Benitez and J. W. Davidson, ‘‘Code Generation for Streaming: an Access/Execute Mechanism’’, *Proceedings of the Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, Santa Clara, CA (April 8-11, 1991), pp. 132-141.

[BuYP91] M. Butler, T. Yeh and Y. Patt, ‘‘Single Instruction Stream Parallelism is Greater than Two’’, *Proceedings of the Eighteenth Annual International Symposium on Computer Architecture*, Toronto, Canada (May 27-30, 1991), pp. 276-286.

[CGKP87] G. L. Craig, J. R. Goodman, R. H. Katz, A. R. Pleszkun, K. Ramachandran, J. Sayah and J. E. Smith, ‘‘PIPE: A High Performance VLSI Processor Implementation’’, *Journal of VLSI and Computer Systems*, vol. 2 (1987).

[FaPl91] M. Farrens and A. Pleszkun, ‘‘Overview of the PIPE Processor Implementation’’, *Proceedings of the 24th Annual Hawaii International Conference on System Sciences*, Kapaa, Kauai (January 9-11, 1991), pp. 433-443.

[GHLP85] J. R. Goodman, J. T. Hsieh, K. Liou, A. R. Pleszkun, P. B. Schechter and H. C. Young, ‘‘PIPE: a VLSI Decoupled Architecture’’, *Proceedings of the Twelveth Annual International Symposium on Computer Architecture*(June 1985), pp. 20-27.

[JoWa89] N. P. Jouppi and D. W. Wall, ‘‘Available Instruction-Level Parallelism for Superscalar and Superpipelined Machines’’, *Proceedings of the Third International Conference on Architectural Support for Programming Languages and Operating Systems*, Boston, Mass (April 3-6, 1989), pp. 272-282.

[Lam88] M. S. Lam, ‘‘Software Pipelining: An Effective Scheduling Technique for VLIW Machines’’, *Proceedings of the ACM SIGPLAN Notices 1988 Conference on Programming Languages and Implementations*(June 1988), pp. 318-328.

[MLCH92] S. A. Mahlke, D. C. Lin, W. Y. Chen, R. E. Hank and R. A. Bringmann, ‘‘Effective Compiler Support for Predicated Execution Using the Hyperblock’’, *Proceedings of the 25th Annual International Symposium on Microarchitecture*, Portland, Oregon (December 1-4, 1992), pp. 45-54.

[RaS92] B. R. Rau, M. S. Schlansker and P. P. Tirumalai, ‘‘Code Generation Schema for Modulo Scheduled Loops’’, *Proceedings of the 25th Annual International Symposium on Microarchitecture*, Portland, Oregon (December 1-4, 1992), pp. 158-169.

[Site93] R. L. Sites, ‘‘Alpha AXP Architecture’’, *Communications of the ACM*, vol. 36, no. 2 (February, 1993), pp. 33-44.

[Smit82] J. E. Smith, ‘‘Decoupled Access/Execute Computer Architectures’’, *Proceedings of the Ninth Annual International Symposium on Computer Architecture*, Austin, Texas (April 26-29, 1982), pp. 112-119.

[SDVK87] J. E. Smith, G. E. Dermer, B. D. Vanderwarn, S. D. Klinger, C. M. Rozewski, D. L. Fowler, K. R. Scidmore and J. P. Laudon, ‘‘The ZS-1 Central Processor’’, *Proceedings of the Second International Conference on Architectural Support for Programming Languages and Operating Systems*, Palo Alto, California (October 1987), pp. 199-204.

[SCHP92] C. Stephens, B. Cogswell, J. Heinlein, G. Palmer and J. P. Shen, ‘‘Instruction Level Profiling and Evaluation of the IBM RS/6000’’, *Proceedings of the 19th Annual Symposium on Computer Architecture*, vol. 20, no. 2 (May 19-21, 1992), pp. 180-189.

[TjFl70] G. S. Tjaden and M. J. Flynn, ‘‘Detection and Parallel Execution of Parallel Instructions’’, *IEEE Transactions on Computer*(May 1970), pp. 889-895.

[TyFP92] G. Tyson, M. Farrens and A. Pleszkun, ‘‘MISC: A Multiple Instruction Stream Computer’’, *Proceedings of the 25th Annual International Symposium on Microarchitecture*, Portland, Oregon (December 1-4, 1992), pp. 193-196.

[TySF93] G. S. Tyson, R. Shaw and M. Farrens, ‘‘An Interactive Compiler Development System’’, *Tcl/Tk Workshop*(June 10-11, 1993).

[Wall91] D. Wall, ‘‘Limits of Instruction-Level Parallelism’’, *Proceedings of the Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, Santa Clara, CA (April 8-11, 1991), pp. 176-189.

[WeSm87] S. Weiss and J. E. Smith, ‘‘A Study of Scalar Compaction Techniques for Pipelined Supercomputers’’, *Proceedings of the Second International Conference on Architectural Support for Programming Languages and Operating Systems*, Palo Alto, California (October 1987), pp. 105-109.

[Wulf92] W. Wulf, ‘‘Evaluation of the WM Architecture’’, *Proceedings of the 19th Annual Symposium on Computer Architecture*, vol. 20, no. 2 (May 19-21, 1992), pp. 382-390.