EFFICIENT COMPLEX MULTIPLICATION AND FAST FOURIER TRANSFORM (FFT) IMPLEMENTATION ON THE MANARRAY ARCHITECTURE

Abstract

Efficient computation of complex multiplication results and very efficient fast Fourier transforms (FFTs) are provided. A parallel array VLIW digital signal processor is employed along with specialized complex multiplication instructions and communication operations between the processing elements which are overlapped with computation to provide very high performance operation. Successive iterations of a loop of tightly packed VLIWs are used allowing the complex multiplication pipeline hardware to be efficiently used. In addition, efficient techniques for supporting combined multiply accumulate operations are described.
We claim:

1. An apparatus for the efficient processing of complex multiplication computations, the apparatus comprising: at least one controller sequence processor (SP); a memory for storing process control instructions; a first multiply complex numbers instruction stored in the memory and operative to control the PEs to carry out a multiplication operation involving a pair of complex numbers; and hardware for implementing the first multiply complex numbers instruction.

2. The apparatus of claim 1 further comprising a plurality of processing elements (PEs) interconnected with said SP and arranged in an N.times.N array interconnected in a manifold array interconnection network.

3. The apparatus of claim 1 wherein the first multiply complex instruction completes execution in 2 cycles.

4. The apparatus of claim 1 wherein the first multiply complex instruction is tightly pipelineable.

5. The apparatus of claim 1 wherein each complex number is stored as a word, each word comprising a first half word and a second half word, with a real component of each complex number being stored as the first half word and an imaginary component of each complex number being stored as the second half word.

6. The apparatus of claim 1 wherein the first multiply complex instruction includes a plurality of rounding modes, the rounding modes including: rounding toward a nearest integer; rounding toward zero; rounding toward infinity; and rounding toward negative infinity.

7. The apparatus of claim 1 wherein the first multiply complex numbers instruction is one of the following group of instructions: a multiply complex numbers (MPYCX), a multiply complex numbers instruction (MPYCXJ) operative to carry out the multiplication of a pair of complex numbers where an argument is conjugated, a multiply complex numbers instruction (MPYCXD2) operative to carry out the multiplication of a pair of complex numbers with a result divided by two, and a multiply complex numbers instruction (MPYCXJD2) operative to carry out the multiplication of a pair of complex numbers where an argument is conjugated with a result divided by two.

8. The apparatus of claim 1 further comprising a multiply accumulate unit including the memory for storing the first multiply complex numbers instruction.

9. The apparatus of claim 8 wherein the multiply accumulate unit operates in response to a multiply accumulate instruction (MPYA) to extend a multiplication operation with an accumulate operation.

10. The apparatus of claim 8 wherein the multiply accumulate unit operates in response to a sum two product accumulate instruction (SUM2PA) to extend two multiplication operations with an accumulate operation.

11. The apparatus of claim 9 wherein the multiply accumulate unit operates in response to a multiply complex with accumulate instruction (MPYCXA) to carry out the multiplication of a pair of complex numbers with
accumulation of a third complex number.

12. The apparatus of claim 11 wherein the MPYCXA instruction completes execution in 2 cycles.

13. The apparatus of claim 12 wherein the MPYCXA instruction is tightly pipelineable.

14. The apparatus of claim 1 further comprising one or more of the following additional instructions (MPYCXA, MPYCXAD2, MPYCXJA or MPYCXJAD2) stored in the memory to carry out complex multiplication operations pipelined in 2 cycles.

15. A method for the computation of an FFT by a plurality of processing elements (PEs), the method comprising the steps of: loading input data from a memory into each PE in a cyclic manner; calculating a local FFT by each PE; multiplying by the twiddle factors and calculating a FFT by the cluster of PEs; and loading the FFTs into the memory.

16. A method for the computation of a distributed FFT by an N.times.N processing element (PE) array, the method comprising the steps of: loading a complex number x and a corresponding twiddle factor w from a memory into each of the PEs; calculating a first product by the multiplication of the complex numbers x and w; transmitting the first product from each of the PEs to another PE in the N.times.N array; receiving the first product and treating it as a second product in each of the PEs; selectively adding or subtracting the first product and the second product to form a first result; calculating a third product in selected PEs; transmitting the first result or third product in selected PEs to another PE in the N.times.N array; selectively adding or subtracting the received values to form a second result; and storing the second results in the memory.

17. A method for efficient computation by a 2.times.2 processing element (PE) array interconnected in a manifold array interconnection network, the array comprising four PEs (PE0, PE1, PE2 and PE3), the method comprising the steps of: loading a complex number x and a corresponding twiddle factor w from a memory into each of the four PEs, complex number x including subparts x0, x1, x2 and x3, twiddle factor w including subparts w0, w1, w2 and w3; multiplying the complex numbers x and w, such that PE0 multiplies x0 and w0 to produce a product0, PE1 multiplies x1 and w1 to produce a product1, PE2 multiplies x2 and w2 to produce a product2, and PE3 multiplies x3 and w3 to produce a product3; transmitting the product0, the product1, the product2 and the product3, such that PE0 transmits the product0 to PE2, PE1 transmits the product1 to PE3, PE2 transmits the product2 to PE0, and PE3 transmits the product3 to PE1; and performing arithmetic logic operations, such that PE0 adds the product0 and the product2 to produce a sum t0, PE1 adds the product1 and the product3 to produce a sum t2, PE2 subtracts the product2 from the product0 to produce a sum t1, and PE3 subtracts the product3 from the product1 to produce a result which is multiplied by -i to produce a sum t3.

18. The method of claim 17 further comprising the steps of: transmitting the sums t0, t1, t2 and t3, such that PE0 transmits t0 to PE1, PE1 transmits t2 to PE0, PE2 transmits t1 to PE3, and PE3 transmits t3 to PE2; performing the arithmetic logic operations, such that PE0 adds t0 and t2 to produce a y0, PE1 subtracts t2 from t0 to produce a y1, PE2 adds t1 and t3 to produce a y2, and PE3 subtracts t3 from t1 to produce a y3; and storing y0, y1, y2 and y3 in a memory.

19. A special hardware instruction for handling the multiplication with accumulate for two complex numbers from a source register whereby utilizing said instruction and accumulated complex product of two source operands is rounded according to a rounding mode specified in the instruction and loaded into a target register with the complex numbers organized in the source such that a halfword (H1) contains the real component and a halfword (H0) contains the imaginary component.
20. The special hardware instruction of claim 19 wherein the accumulated complex product is divided by two before it is rounded.

21. An apparatus to efficiently fetch instructions including complex multiplication instructions and an accumulate form of multiplication instructions from a memory element and dispatch the fetched instruction to at least one of a plurality of multiply complex and multiply with accumulate execution units to carry out the instruction specified operation, the apparatus comprising: a memory element; means for fetching said instructions from the memory element; a plurality of multiply complex and multiply with accumulate execution units; and means to dispatch the fetched instruction to at least one of said plurality of execution units to carry out the instruction specified operation.

22. The apparatus of claim 21 further comprising: an instruction register to hold a dispatched multiply complex instruction (MPYCX); means to decode the MPYCX instruction and control the execution of the MPYCX instruction; two source registers each holding a complex number as operand inputs to the multiply complex execution hardware; four multiplication units to generate terms of the complex multiplication; four pipeline registers to hold the multiplication results; an add function which adds two of the multiplication results from the pipeline registers for the imaginary component of the result; a subtract function which subtracts two of the multiplication results from the pipeline registers for the real component of the result; a round and select unit to format the real and imaginary results; and a result storage location for saving the final multiply complex result, whereby the apparatus is operative for the efficient processing of multiply complex computations.

23. The apparatus of claim 21 wherein the means for fetching said instructions is a sequence processor (SP) controller.

24. The apparatus of claim 22 wherein the round and select unit provides a shift right as a divide by 2 operation for a multiply complex divide by 2 instruction (MPYCXD2).

25. The apparatus of claim 21 further comprising: an instruction register to hold a dispatched multiply complex instruction (MPYCXJ); means to decode the MPYCXJ instruction and control the execution of the MPYCXJ instruction; two source registers each holding a complex number as operand inputs to the multiply complex execution hardware; four multiplication units to generate terms of the complex multiplication; four pipeline registers to hold the multiplication results; an add function which adds two of the multiplication results from the pipeline registers for the imaginary component of the result; a subtract function which subtracts two of the multiplication results from the pipeline registers for the real component of the result; a round and select unit to format the real and imaginary results; and a result storage location for saving the final multiply complex conjugate result, whereby the apparatus is operative for the efficient processing of multiply complex conjugate computations.

26. The apparatus of claim 25 wherein the round and select unit provides a shift right as a divide by 2 operation for a multiply complex conjugate divide by 2 instruction (MPYCXJD2).

27. The apparatus of claim 21 further comprising: an instruction register to hold the dispatched multiply accumulate instruction (MPYA); means to decode the MPYA instruction and control the execution of the MPYA instruction; two source registers each holding a source operand as inputs to the multiply accumulate execution hardware; at least two multiplication units to generate two products of the multiplication; at least two pipeline registers to hold the multiplication results; at least two accumulate operand inputs to the second pipeline stage accumulate hardware; at least two add functions which each adds the results from the pipeline registers with the third accumulate operand creating two multiply accumulate results; a round and select unit to format the results if required by the MPYA instruction; and a result storage location for saving the final multiply accumulate result,
whereby the apparatus is operative for the efficient processing of multiply accumulate computations.

28. The apparatus of claim 21 further comprising: an instruction register to hold a dispatched multiply accumulate instruction (SUM2PA); means to decode the SUM2PA instruction and control the execution of the SUM2PA instruction; at least two source registers each holding a source operand as inputs to the SUM2PA execution hardware; at least two multiplication units to generate two products of the multiplication; at least two pipeline registers to hold the multiplication results; at least one accumulate operand input to the second pipeline stage accumulate hardware; at least one add function which adds the results from the pipeline registers with the third accumulate operand creating a SUM2PA result; a round and select unit to format the results if required by the SUM2PA instruction; and a result storage location for saving the final result, whereby the apparatus is operative for the efficient processing of sum of 2 products accumulate computations.

29. The apparatus of claim 21 further comprising: an instruction register to hold the dispatched multiply complex accumulate instruction (MPYCXA); means to decode the MPYCXA instruction and control the execution of the MPYCXA instruction; two source registers each holding a complex number as operand inputs to the multiply complex accumulate execution hardware; four multiplication units to generate terms of the complex multiplication; four pipeline registers to hold the multiplication results; at least two accumulate operand inputs to the second pipeline stage accumulate hardware; an add function which adds two of the multiplication results from the pipeline registers and also adds one of the accumulate operand input for the imaginary component of the result; a subtract function which subtracts two of the multiplication results from the pipeline registers and also adds the other accumulate operand input for the real component of the result; a round and select unit to format the real and imaginary results; and a result storage location for saving the final multiply complex accumulate result, whereby the apparatus is operative for the efficient processing of multiply complex accumulate computations.

30. The apparatus of claim 29 wherein the round and select unit provides a shift right as a divide by 2 operation for a multiply complex accumulate divide by 2 instruction (MPYCXAD2).

31. The apparatus of claim 21 further comprising: an instruction register to hold the dispatched multiply complex conjugate accumulate instruction (MPYCXJA); means to decode the MPYCXJA instruction and control the execution of the MPYCXJA instruction; two source registers each holding a complex number as operand inputs to the multiply complex conjugate accumulate execution hardware; four multiplication units to generate terms of the complex multiplication; four pipeline registers to hold the multiplication results; at least two accumulate operand inputs to the second pipeline stage accumulate hardware; an add function which adds two of the multiplication results from the pipeline registers and also adds one of the accumulate operand input for the real component of the result; a subtract function which subtracts two of the multiplication results from the pipeline registers and also adds the other accumulate operand input for the imaginary component of the result; a round and select unit to format the real and imaginary results; and a result storage location for saving the final multiply complex conjugate accumulate result, whereby the apparatus is operative for the efficient processing of multiply complex conjugate accumulate computations.

32. The apparatus of claim 31 wherein the round and select unit provides a shift right as a divide by 2 operation for a multiply complex conjugate accumulate divide by 2 instruction (MPYCXJAD2).

33. The apparatus of claim 21 wherein the complex multiplication instructions and accumulate form of multiplication instructions include MPYCX, MPYCXD2, MPYCXJ, MPYCXJD2, MPYCXA, MPYCXAD2, MPYCXJA, MPYCXJAD2 instructions, and all of these instructions complete execution in 2 cycles.

34. The apparatus of claim 21 wherein the complex multiplication instructions and accumulate form of
multiplication instructions include MPYCX, MPYCXD2, MPYCXJ, MPYCXJD2, MPYCXA, MPYCXAD2, MPYCXJA, MPYCXJAD2 instructions, and all of these instructions are tightly pipelineable.

35. An apparatus for the efficient processing of an FFT, the apparatus comprising: at least one controller sequence processor (SP); a plurality of processing elements (PEs) arranged in an N.times.N array interconnected in a manifold (ManArray) interconnection network; and a memory for storing instructions to be processed by the SP and by the array of PEs.

36. The apparatus of claim 22 wherein the add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction.

37. The apparatus of claim 25 wherein the add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction.

38. The apparatus of claim 29 wherein the add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction.

39. The apparatus of claim 31 wherein the add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction.

Description

[0001] This application claims the benefit of U.S. Provisional Application Serial No. 60/103,712 filed Oct. 9, 1998 which is incorporated by reference in its entirety herein.

FIELD OF THE INVENTION

[0002] The present invention relates generally to improvements to parallel processing, and more particularly to methods and apparatus for efficiently calculating the result of a complex multiplication. Further, the present invention relates to the use of this approach in a very efficient FFT implementation on the manifold array ("ManArray") processing architecture.

BACKGROUND OF THE INVENTION

[0003] The product of two complex numbers x and y is defined to be 
\[ z = x_{R}y_{R} - x_{I}y_{I} + i(x_{R}y_{I} + x_{I}y_{R}) \]
where \( x = x_{R} + ix_{I} \) and \( y = y_{R} + iy_{I} \) and i is an imaginary number, or the square root of negative one, with \( i^2 = -1 \). This complex multiplication of x and y is calculated in a variety of contexts, and it has been recognized that it will be highly advantageous to perform this calculation faster and more efficiently.

SUMMARY OF THE INVENTION

[0004] The present invention defines hardware instructions to calculate the product of two complex numbers encoded as a pair of two fixed-point numbers of 16 bits each in two cycles with single cycle pipeline throughput efficiency. The present invention also defines extending a series of multiply complex instructions with an accumulate operation. These special instructions are then used to calculate the FFT of a vector of numbers efficiently.
A more complete understanding of the present invention, as well as other features and advantages of the invention will be apparent from the following Detailed Description and the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates an exemplary 2×2 ManArray iVLIW processor;

FIG. 2A illustrates a presently preferred multiply complex instruction, MPYCX;

FIG. 2B illustrates the syntax and operation of the MPYCX instruction of FIG. 2A;

FIG. 3A illustrates a presently preferred multiply complex divide by 2 instruction, MPYCXD2;

FIG. 3B illustrates the syntax and operation of the MPYCXD2 instruction of FIG. 3A;

FIG. 4A illustrates a presently preferred multiply complex conjugate instruction, MPYCXJ;

FIG. 4B illustrates the syntax and operation of the MPYCXJ instruction of FIG. 4A;

FIG. 5A illustrates a presently preferred multiply complex conjugate divide by two instruction, MPYCXJD2;

FIG. 5B illustrates the syntax and operation of the MPYCXJD2 instruction of FIG. 5A;

FIG. 6 illustrates hardware aspects of a pipelined multiply complex and its divide by two instruction variant;

FIG. 7 illustrates hardware aspects of a pipelined multiply complex conjugate, and its divide by two instruction variant;

FIG. 8 shows an FFT signal flow graph;

FIGS. 9A-9H illustrate aspects of the implementation of a distributed FFT algorithm on a 2×2 ManArray processor using a VLIW algorithm with MPYCX instructions in a cycle-by-cycle sequence with each step corresponding to operations in the FFT signal flow graph;

FIG. 9I illustrates how multiple iterations may be tightly packed in accordance with the present invention for a distributed FFT of length four;

FIG. 9J illustrates how multiple iterations may be tightly packed in accordance with the present invention for a distributed FFT of length two;

FIGS. 10A and 10B illustrate Kronecker Product examples for use in reference to the mathematical presentation of the presently preferred distributed FFT algorithm;

FIG. 11A illustrates a presently preferred multiply accumulate instruction, MPYA;

FIG. 11B illustrates the syntax and operation of the MPYA instruction of FIG. 11A;

FIG. 12A illustrates a presently preferred sum of 2 products accumulate instruction, SUM2PA;
[0025] FIG. 12B illustrates the syntax and operation of the SUM2PA instruction of FIG. 12A;
[0026] FIG. 13A illustrates a presently preferred multiply complex accumulate instruction, MPYCXA;
[0027] FIG. 13B illustrates the syntax and operation of the MPYCXA instruction of FIG. 13A;
[0028] FIG. 14A illustrates a presently preferred multiply complex accumulate divide by two instruction, MPYCXAD2;
[0029] FIG. 14B illustrates the syntax and operation of the MPYCXAD2 instruction of FIG. 14A;
[0030] FIG. 15A illustrates a presently preferred multiply complex conjugate accumulate instruction, MPYCXJA;
[0031] FIG. 15B illustrates the syntax and operation of the MPYCXJA instruction of FIG. 15A;
[0032] FIG. 16A illustrates a presently preferred multiply complex conjugate accumulate divide by two instruction, MPYCXJAD2;
[0033] FIG. 16B illustrates the syntax and operation of the MPYCXJAD2 instruction of FIG. 16A;
[0034] FIG. 17 illustrates hardware aspects of a pipelined multiply complex accumulate and its divide by two variant; and
[0035] FIG. 18 illustrates hardware aspects of a pipelined multiply complex conjugate accumulate and its divide by two variant.

DETAILED DESCRIPTION


[0037] In a presently preferred embodiment of the present invention, a ManArray 2.times.2 iVLIW single instruction multiple data stream (SIMD) processor 100 shown in FIG. 1 contains a controller sequence processor (SP) combined with processing element-0 (PE0) SP/PE0 101, as described in further detail in U.S. application Ser. No. 09/169,072 entitled "Methods and Apparatus for Dynamically Merging an Array Controller with an Array
Processing Element”. Three additional PEs 151, 153, and 155 are also utilized to demonstrate the implementation of efficient complex multiplication and fast fourier transform (FFT) computations on the ManArray architecture in accordance with the present invention. It is noted that the PEs can be also labeled with their matrix positions as shown in parentheses for PE0 (PE00) 101, PE1 (PE01)151, PE2 (PE10) 153, and PE3 (PE11) 155.

[0038] The SP/PE0 101 contains a fetch controller 103 to allow the fetching of short instruction words (SIWs) from a 32-bit instruction memory 105. The fetch controller 103 provides the typical functions needed in a programmable processor such as a program counter (PC), branch capability, digital signal processing, EP loop operations, support for interrupts, and also provides the instruction memory management control which could include an instruction cache if needed by an application. In addition, the SIW I-Fetch controller 103 dispatches 32-bit SIWs to the other PEs in the system by means of a 32-bit instruction bus 102.

[0039] In this exemplary system, common elements are used throughout to simplify the explanation, though actual implementations are not so limited. For example, the execution units 131 in the combined SP/PE0 101 can be separated into a set of execution units optimized for the control function, e.g. fixed point execution units, and the PE0 as well as the other PEs 151, 153 and 155 can be optimized for a floating point application. For the purposes of this description, it is assumed that the execution units 131 are of the same type in the SP/PE0 and the other PEs. In a similar manner, SP/PE0 and the other PEs use a five instruction slot iVLIW architecture which contains a very long instruction word memory (VIM) memory 109 and an instruction decode and VIM controller function unit 107 which receives instructions as dispatched from the SP/PE0's I-Fetch unit 103 and generates the VIM address-and-control signals 108 required to access the iVLIWs stored in the VIM. These iVLIWs are identified by the letters SLAMD in VIM 109. The loading of the iVLIWs is described in further detail in U.S. patent application Ser. No. 09/187,539 entitled "Methods and Apparatus for Efficient Synchronous MIMD Operations with iVLIW PE-to-PE Communication". Also contained in the SP/PE0 and the other PEs is a common PE configurable register file 127 which is described in further detail in U.S. patent application Ser. No. 09/169,255 entitled "Methods and Apparatus for Dynamic Instruction Controlled Reconfiguration Register File with Extended Precision".

[0040] Due to the combined nature of the SP/PE0, the data memory interface controller 125 must handle the data processing needs of both the SP controller, with SP data in memory 121, and PE0, with PE0 data in memory 123. The SP/PE0 controller 125 also is the source of the data that is sent over the 32-bit broadcast data bus 126. The other PEs 151, 153, and 155 contain common physical data memory units 123', 123", and 123"' though the data stored in them is generally different as required by the local processing done on each PE. The interface to these PE data memories is also a common design in PEs 1, 2, and 3 and indicated by PE local memory and data bus interface logic 157, 157' and 157". Interconnecting the PEs for data transfer communications is the cluster switch 171 more completely described in U.S. patent application Ser. No. 08/885,310 entitled "Manifold Array Processor", U.S. application Ser. No. 09/949,122 entitled "Methods and Apparatus for Manifold Array Processing", and U.S. application Ser. No. 09/169,256 entitled "Methods and Apparatus for ManArray PE-to-PE Switch Control". The interface to a host processor, other peripheral devices, and/or external memory can be done in many ways. The primary mechanism shown for completeness is contained in a direct memory access (DMA) control unit 181 that provides a scalable ManArray data bus 183 that connects to devices and interface units external to the ManArray core. The DMA control unit 181 provides the data flow and bus arbitration mechanisms needed for these external devices to interface to the ManArray core memories via the multiplexed bus interface represented by line 185. A high level view of a ManArray Control Bus (MCB) 191 is also shown.

[0041] All of the above noted patents are assigned to the assignee of the present invention and incorporated herein by reference in their entirety.

[0042] Special Instructions for Complex Multiply
Turning now to specific details of the ManArray processor as adapted by the present invention, the present invention defines the following special hardware instructions that execute in each multiply accumulate unit (MAU), one of the execution units 131 of FIG. 1 and in each PE, to handle the multiplication of complex numbers:

MPYCX instruction 200 (FIG. 2A), for multiplication of complex numbers, where the complex product of two source operands is rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCX instruction format is shown in FIG. 2A. The syntax and operation description 210 is shown in FIG. 2B.

MPYCXD2 instruction 300 (FIG. 3A), for multiplication of complex numbers, with the results divided by 2, FIG. 3, where the complex product of two source operands is divided by two, rounded according to the rounding mode specified in the instruction, and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXD2 instruction format is shown in FIG. 3A. The syntax and operation description 310 is shown in FIG. 3B.

MPYCXJ instruction 400 (FIG. 4A), for multiplication of complex numbers where the second argument is conjugated, where the complex product of the first source operand times the conjugate of the second source operand, is rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXJ instruction format is shown in FIG. 4A. The syntax and operation description 410 is shown in FIG. 4B.

MPYCXJD2 instruction 500 (FIG. 5A), for multiplication of complex numbers where the second argument is conjugated, with the results divided by 2, where the complex product of the first source operand times the conjugate of the second operand, is divided by two, rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXJD2 instruction format is shown in FIG. 5A. The syntax and operation description 510 is shown in FIG. 5B.

All of the above instructions 200, 300, 400 and 500 complete in 2 cycles and are pipelineable. That is, another operation can start executing unit after the first cycle. All complex multiplication instructions return a word containing the real and imaginary part of the complex product in half words H1 and H0 respectively.

To preserve maximum accuracy, and provide flexibility to programmers, four possible rounding modes are defined:

Round toward the nearest integer (referred to as ROUND)

Round toward 0 (truncate or fix, referred to as TRUNC)

Round toward infinity (round up or ceiling, the smallest integer greater than or equal to the argument, referred to as CEIL)

Round toward negative infinity (round down or floor, the largest integer smaller than or equal to the argument, referred to as FLOOR).
Hardware suitable for implementing the multiply complex instructions is shown in FIG. 6 and FIG. 7. These figures illustrate a high level view of the hardware apparatus 600 and 700 appropriate for implementing the functions of these instructions. This hardware capability may be advantageously embedded in the ManArray multiply accumulate unit (MAU), one of the execution units 131 of FIG. 1 and in each PE, along with other hardware capability supporting other MAU instructions. As a pipelined operation, the first execute cycle begins with a read of the source register operands from the compute register file (CRF) shown as registers 603 and 605 in FIG. 6 and as registers 111, 127, 127', 127", and 127"" in FIG. 1. These register values are input to the MAU logic after some operand access delay in halfword data paths as indicated to the appropriate multiplication units 607, 609, 611, and 613 of FIG. 6. The outputs of the multiplication operation units, X.sub.R*Y.sub.R607, X.sub.R*Y. sub.I609, X.sub.I*Y.sub.R611, and X.sub.I*Y.sub.I613, are stored in pipeline registers 615, 617, 619, and 621, respectively. The second execute cycle, which can occur while a new multiply complex instruction is using the first cycle execute facilities, begins with using the stored pipeline register values, in pipeline register 615, 617, 619, and 621, and appropriately adding in adder 625 and subtracting in subtractor 623 as shown in FIG. 6. The add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction. The values generated by the apparatus 600 shown in FIG. 6 contain a maximum precision of calculation which exceeds 16-bits. Consequently, the appropriate bits must be selected and rounded as indicated in the instruction before storing the final results. The selection of the bits and rounding occurs in selection and rounder circuit 627. The two 16-bit rounded results are then stored in the appropriate halfword position of the target register 629 which is located in the compute register file (CRF). The divide by two variant of the multiply complex instruction 300 selects a different set of bits as specified in the instruction through block 627. The hardware 627 shifts each data value right by an additional 1-bit and loads two divided-by-2 rounded and shifted values into each half word position in the target registers 629 in the CRF.

The hardware 700 for the multiply complex conjugate instruction 400 is shown in FIG. 7. The main difference between multiply complex and multiply complex conjugate is in adder 723 and subtractor 725 which swap the addition and subtraction operation as compared with FIG. 6. The results from adder 723 and subtractor 725 still need to be selected and rounded in selection and rounder circuit 727 and the final rounded results stored in the target register 729 in the CRF. The divide by two variant of the multiply complex conjugate instruction 500 selects a different set of bits as specified in the instruction through selection and rounder circuit 727. The hardware of circuit 727 shifts each data value right by an additional 1-bit and loads two divided-by-2 rounded and shifted values into each half word position in the target registers 729 in the CRF.

The FFT Algorithm

The power of indirect VLIW parallelism using the complex multiplication instructions is demonstrated with the following fast Fourier transform (FFT) example. The algorithm of this example is based upon the sparse factorization of a discrete Fourier transform (DFT) matrix. Kronecker-product mathematics is used to demonstrate how a scalable algorithm is created.

The Kronecker product provides a means to express parallelism using mathematical notation. It is known that there is a direct mapping between different tensor product forms and some important architectural features of processors. For example, tensor matrices can be created in parallel form and in vector form. J. Granata, M. Conner, R. Tolimieri, The Tensor Product: A Mathematical Programming Language for FFTs and other Fast DSP Operations, IEEE SP Magazine, January 1992, pp. 40-48. The Kronecker product of two matrices is a block matrix with blocks that are copies of the second argument multiplied by the corresponding element of the first argument. Details of an exemplary calculation of matrix vector products

\[ y = (\text{Im\{circle over \((\cdot)\times\}\}}A)x \]
[0059] are shown in FIG. 10A. The matrix is block diagonal with m copies of A. If vector x was distributed block-wise in m processors, the operation can be done in parallel without any communication between the processors. On the other hand, the following calculation, shown in detail in FIG. 10B,

\[ y = (A \times IM)x \]

[0060] requires that x be distributed physically on m processors for vector parallel computation.

[0061] The two Kronecker products are related via the identity

\[ I \times A = P(A \times I)P^T \]

[0062] where P is a special permutation matrix called stride permutation and P.sup.T is the transpose permutation matrix. The stride permutation defines the required data distribution for a parallel operation, or the communication pattern needed to transform block distribution to cyclic and vice-versa.

[0063] The mathematical description of parallelism and data distributions makes it possible to conceptualize parallel programs, and to manipulate them using linear algebra identities and thus better map them onto target parallel architectures. In addition, Kronecker product notation arises in many different areas of science and engineering. The Kronecker product simplifies the expression of many fast algorithms. For example, different FFT algorithms correspond to different sparse matrix factorizations of the Discrete Fourier Transform (DFT), whose factors involve Kronecker products. Charles F. Van Loan, Computational Frameworks for the Fast Fourier Transform, SIAM, 1992, pp 78-80.

[0064] The following equation shows a Kronecker product expression of the FFT algorithm, based on the Kronecker product factorization of the DFT matrix,

\[ F_n = (F_p I_m) D_p, m (I_p F_m) P_{n,p} \]

[0065] where:

[0066] n is the length of the transform

[0067] p is the number of PEs

[0068] m=n/p

[0069] The equation is operated on from right to left with the P.sub.n,p permutation operation occurring first. The permutation directly maps to a direct memory access (DMA) operation that specifies how the data is to be loaded in the PEs based upon the number of PEs p and length of the transform n.

\[ F_n = (F_p I_m) D_p, m (I_p F_m) P_{n,p} \]

[0070] where P.sub.n,p corresponds to DMA loading data with stride p to local PE memories.

[0071] In the next stage of operation all the PEs execute a local FFT of length m=n/p with local data. No communications between PEs is required.

\[ F_n = (F_p I_m) D_p, (I_p F_m) P_{n,p} \]

[0072] where (I.sub.p \times m)F.sub.m specifies that all PEs execute a local FFT of length m sequentially, with local data.

[0073] In the next stage, all the PEs scale their local data by the twiddle factors and collectively execute m
distributed FFTs of length p. This stage requires inter-PE communications. $4^F_n = (F_p I_m)D_p, m (I_p F_m)P n, p$

[0074] where $(F_{sub.p}(\times .I_{sub.m})D_{sub.p,m}$ specifies that all PEs scale their local data by the twiddle factors and collectively execute multiple FFTs of length p on distributed data. In this final stage of the FFT computation, a relatively large number m of small distributed FFTs of size p must be calculated efficiently. The challenge is to completely overlap the necessary communications with the relatively simple computational requirements of the FFT.

[0075] The sequence of illustrations of FIGS. 9A-9H outlines the ManArray distributed FFT algorithm using the indirect VLIW architecture, the multiply complex instructions, and operating on the 2.times.2 ManArray processor 100 of FIG. 1. The signal flow graph for the small FFT is shown in FIG. 8 and also shown in the right-hand-side of FIGS. 9A-9H. In FIG. 8, the operation for a 4 point FFT is shown where each PE executes the operations shown on a horizontal row. The operations occur in parallel on each vertical time slice of operations as shown in the signal flow graph figures in FIGS. 9A-9H. The VLIW code is displayed in a tabular form in FIGS. 9A-9H that corresponds to the structure of the ManArray architecture and the iVLIW instruction. The columns of the table correspond to the execution units available in the ManArray PE: Load Unit, Arithmetic Logic Unit (ALU), Multiply Accumulate Unit (MAU), Data Select Unit (DSU) and the Store Unit. The rows of the table can be interpreted as time steps representing the execution of different iVLIW lines.

[0076] The technique shown is a software pipeline implemented approach with iVLIWs. In FIGS. 9A-9I, the tables show the basic pipeline for PE3 155. FIG. 9A represents the input of the data X and its corresponding twiddle factor W by loading them from the PEs local memories, using the load indirect (Lii) instruction. FIG. 9B illustrates the complex arguments X and W which are multiplied using the MPYCX instruction 200, and FIG. 9C illustrates the communications operation between PEs, using a processing element exchange (PEXCHG) instruction. Further details of this instruction are found in U.S. application Ser. No. 09/169,256 entitled "Methods and Apparatus for ManArray PE-PE Switch Control" filed Oct. 9, 1998. FIG. 9D illustrates the local and received quantities are added or subtracted (depending upon the processing element, where for PE3 a subtract (sub) instruction is used). FIG. 9E illustrates the result being multiplied by -i on PE3, using the MPYCX instruction. FIG. 9F illustrates another PE-to-PE communications operation where the previous product is exchanged between the PEs, using the PEXCHG instruction. FIG. 9G illustrates the local and received quantities are added or subtracted (depending upon the processing element, where for PE3 a subtract (sub) instruction is used). FIG. 9H illustrates the step where the results are stored to local memory, using a store indirect (sii) instruction.

[0077] The code for PEs 0, 1, and 2 is very similar, the two subtractions in the arithmetic logic unit in steps 9D and 9G are substituted by additions or subtractions in the other PEs as required by the algorithm displayed in the signal flow graphs. To achieve that capability and the distinct MPYCX operation in FIG. 9E shown in these figures, synchronous MIMD capability is required as described in greater detail in U.S. patent application Ser. No. 09/187,539 filed Nov. 6, 1998 and incorporated by reference herein in its entirety. By appropriate packing, a very tight software pipeline can be achieved as shown in FIG. 9I for this FFT example using only two VLIWs.

[0078] In the steady state, as can be seen in FIG. 9I, the Load, ALU, MAU, and DSU units are fully utilized in the two VLIWs while the store unit is used half of the time. This high utilization rate using two VLIWs leads to very high performance. For example, a 256-point complex FFT can be accomplished in 425 cycles on a 2.times.2 ManArray.

[0079] As can be seen in the above example, this implementation accomplishes the following:

http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2...PG01&s1=Pitsianis.IN.&OS=IN/Pitsianis&RS=IN/Pitsianis (13 of 16)10/16/2006 4:32:01 PM
An FFT butterfly of length 4 can be calculated and stored every two cycles, using four PEs.

The communication requirement of the FFT is completely overlapped by the computational requirements of this algorithm.

The communication is along the hypercube connections that are available as a subset of the connections available in the ManArray interconnection network.

The steady state of this algorithm consists of only two VLIW lines (the source code is two VLIW lines long).

All execution units except the Store unit are utilized all the time, which lead us to conclude that this implementation is optimal for this architecture.

**Problem Size Discussion**

The equation:

$$5 F_n = (F_p I_m) D_p, m (I_p F_m) P_n, p$$

where:

- **n** is the length of the transform,
- **p** is the number of PEs, and
- **m=n/p**

is parameterized by the length of the transform *n* and the number of PEs, where **m=n/p** relates to the size of local memory needed by the PEs. For a given power-of-2 number of processing elements and a sufficient amount of available local PE memory, distributed FFTs of size *p* can be calculated on a ManArray processor since only hypercube connections are required. The hypercube of *p* or fewer nodes is a proper subset of the ManArray network. When *p* is a multiple of the number of processing elements, each PE emulates the operation of more than one virtual node. Therefore, any size of FFT problem can be handled using the above equation on any size of ManArray processor.

For direct execution, in other words, no emulation of virtual PEs, on a ManArray of size *p*, we need to provide a distributed FFT algorithm of equal size. For **p**=1, it is the sequential FFT. For **p**=2, the FFT of length 2 is the butterfly:

$$Y_0 = x_0 + w * X_1,$$

$$Y_1 = x_0 - w * X_1$$

where **X0** and **Y0** reside in or must be saved in the local memory of **PE0** and **X1** and **Y1** on **PE1**, respectively. The VLIWs in **PE0** and **PE1** in a 1.times.2 ManArray processor (**p**=2) that are required for the calculation of multiple FFTs of length 2 are shown in FIG. 9J which shows that two FFT results are produced every two cycles using four VLIWs.

**Extending Complex Multiplication**

It is noted that in the two-cycle complex multiplication hardware described in FIGS. 6 and 7, the addition...
and subtraction blocks 623, 625, 723, and 725 operate in the second execution cycle. By including the MPYCX, MPYCXD2, MPYCXJ, and MPYCXJD2 instructions in the ManArray MAU, one of the execution units 131 of FIG. 1, the complex multiplication operations can be extended. The ManArray MAU also supports multiply accumulate operations (MACs) as shown in FIGS. 11A and 12A for use in general digital signal processing (DSP) applications. A multiply accumulate instruction (MPYA) 1100 as shown in FIG. 11A, and a sum two product accumulate instruction (SUM2PA) 1200 as shown in FIG. 12A, are defined as follows.

[0096] In the MPYA instruction 1100 of FIG. 11A, the product of source registers Rx and Ry is added to target register Rt. The word multiply form of this instruction multiplies two 32-bit values producing a 64-bit result which is added to a 64-bit odd/even target register. The dual halfword form of MPYA instruction 1100 multiplies two pairs of 16-bit values producing two 32-bit results: one is added to the odd 32-bit word, the other is added to the even 32-bit word of the odd/even target register pair. Syntax and operation details 1110 are shown in FIG. 11B. In the SUM2PA instruction 1200 of FIG. 12A, the product of the high halfwords of source registers Rx and Ry is added to the product of the low halfwords of Rx and Ry and the result is added to target register Rt and then stored in Rt. Syntax and operation details 1210 are shown in FIG. 12B.

[0097] Both MPYA and SUMP2A generate the accumulate result in the second cycle of the two-cycle pipeline operation. By merging MPYCX, MPYCXD2, MPYCXJ, and MPYCXJD2 instructions with MPYA and SUMP2A instructions, the hardware supports the extension of the complex multiply operations with an accumulate operation. The mathematical operation is defined as: \( Z_t = Z_R + X_R Y_R - X_I Y_I + i(Z_I + X_R Y_I + X_I Y_R) \), where \( X = X_R + iX_I \), \( Y = Y_R + iY_I \) and \( i \) is an imaginary number, or the square root of negative one, with \( i^2 = -1 \). This complex multiply accumulate is calculated in a variety of contexts, and it has been recognized that it will be highly advantageous to perform this calculation faster and more efficiently.

[0098] For this purpose, an MPYCXA instruction 1300 (FIG. 13A), an MPYCXAD2 instruction 1400 (FIG. 14A), an MPYCXJA instruction 1500 (FIG. 15A), and an MPYCXJAD2 instruction 1600 (FIG. 16A) define the special hardware instructions that handle the multiplication with accumulate for complex numbers. The MPYCXA instruction 1300, for multiplication of complex numbers with accumulate is shown in FIG. 13. Utilizing this instruction, the accumulated complex product of two source operands is rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXA instruction format is shown in FIG. 13A. The syntax and operation description 1310 is shown in FIG. 13B.

[0099] The MPYCXAD2 instruction 1400, for multiplication of complex numbers with accumulate, with the results divided by two is shown in FIG. 14A. Utilizing this instruction, the accumulated complex product of two source operands is divided by two, rounded according to the rounding mode specified in the instruction, and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXAD2 instruction format is shown in FIG. 14A. The syntax and operation description 1410 is shown in FIG. 14B.

[0100] The MPYCXJA instruction 1500, for multiplication of complex numbers with accumulate where the second argument is conjugated is shown in FIG. 15A. Utilizing this instruction, the accumulated complex product of the first source operand times the conjugate of the second source operand, is rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXJA instruction format is shown in FIG. 15A. The syntax and operation description 1510
The MPYCXJAD2 instruction 1600, for multiplication of complex numbers with accumulate where the second argument is conjugated, with the results divided by two is shown in FIG. 16A. Utilizing this instruction, the accumulated complex product of the first source operand times the conjugate of the second operand, is divided by two, rounded according to the rounding mode specified in the instruction and loaded into the target register. The complex numbers are organized in the source register such that halfword H1 contains the real component and halfword H0 contains the imaginary component. The MPYCXJAD2 instruction format is shown in FIG. 16A. The syntax and operation description 1610 is shown in FIG. 16B.

All instructions of the above instructions 1100, 1200, 1300, 1400, 1500 and 1600 complete in two cycles and are pipeline-able. That is, another operation can start executing on the execution unit after the first cycle. All complex multiplication instructions 1300, 1400, 1500 and 1600 return a word containing the real and imaginary part of the complex product in half words H1 and H0 respectively.

To preserve maximum accuracy, and provide flexibility to programmers, the same four rounding modes specified previously for MPYCX, MPYCXD2, MPYCXJ, and MPYCXJD2 are used in the extended complex multiplication with accumulate.

Hardware 1700 and 1800 for implementing the multiply complex with accumulate instructions is shown in FIG. 17 and FIG. 18, respectively. These figures illustrate the high level view of the hardware 1700 and 1800 appropriate for these instructions. The important changes to note between FIG. 17 and FIG. 6 and between FIG. 18 and FIG. 7 are in the second stage of the pipeline where the two-input adder blocks 623, 625, 723, and 725 are replaced with three-input adder blocks 1723, 1725, 1823, and 1825. Further, two new half word source operands are used as inputs to the operation. The Rt.H1 1731 (1831) and Rt.H0 1733 (1833) values are properly aligned and selected by multiplexers 1735 (1835) and 1737 (1837) as inputs to the new adders 1723 (1823) and 1725 (1825). For the appropriate alignment, Rt.H1 is shifted right by 1-bit and Rt.H0 is shifted left by 15-bits. The add/subtract, add/sub blocks 1723 (1823) and 1725 (1825), operate on the input data and generate the outputs as shown. The add function and subtract function are selectively controlled functions allowing either addition or subtraction operations as specified by the instruction. The results are rounded and bits 30-15 of both 32-bit results are selected 1727 (1827) and stored in the appropriate half word of the target register 1729 (1829) in the CRF. It is noted that the multiplexers 1735 (1835) and 1737 (1837) select the zero input, indicated by the ground symbol, for the non-accumulate versions of the complex multiplication series of instructions.

While the present invention has been disclosed in the context of various aspects of presently preferred embodiments, it will be recognized that the invention may be suitably applied to other environments consistent with the claims which follow.

* * * * *