This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

MICROPROCESSOR WITH PACKING OPERATION OF COMPOSITE OPERANDS

BACKGROUND OF THE INVENTION

FIELD OF INVENTION

The present invention includes an apparatus and method of performing operations using a single control signal to manipulate multiple data elements. The present invention allows execution of move, pack and unpack operations on packed data types.

2. DESCRIPTION OF RELATED ART

Today, most personal computer systems operate with one instruction to produce one result. Performance increases are achieved by increasing execution speed of instructions and the processor instruction complexity, and by performing multiple instructions in parallel; known as Complex Instruction Set Computer (CISC). Such processors as the Intel 80386™ microprocessor, available from Intel Corp. of Santa Clara, California, belong to the CISC category of processor.

Previous computer system architecture has been optimized to take advantage of the CISC concept. Such systems typically have data buses thirty-two bits wide. However, applications targeted at computer supported cooperation (CSC - the integration of teleconferencing with mixed media data manipulation), 2D/3D graphics, image processing, video compression/decompression, recognition algorithms and audio manipulation increase the need for improved performance. But, increasing the execution speed and complexity of instructions is only one solution.

One common aspect of these applications is that they often manipulate large amounts of data where only a few bits are important. That is, data whose relevant bits are represented in much fewer bits than the size of the data bus. For example, processors execute many operations on eight bit and sixteen bit data (e.g., pixel color components in a video image) but have much wider data busses and registers. Thus,

a processor having a thirty-two bit data bus and registers, and executing one of these algorithms, can waste up to seventy-five percent of its data processing, carrying and storage capacity because only the first eight bits of data are important.

As such, what is desired is a processor that increases performance by more efficiently using the difference between the number of bits required to represent the data to be manipulated and the actual data carrying and storage capacity of the processor.

SUMMARY OF THE INVENTION

A processor having improved data manipulation operations is described.

A processor. The processor includes a first register for storing a first packed data, a decoder, and a functional unit. The decoder has a control signal input. The control signal input is for receiving a first control signal and a second control signal. The first control signal is for indicating a pack operation. The second control signal is for indicating an unpack operation. The functional unit is coupled to the decoder and the register. The functional unit is for performing the pack operation and the unpack operation using the first packed data. The processor also supports a move operation.

Although a great deal of detail has been included in the description and figures, the invention is defined by the scope of the claims. Only limitations found in those claims apply to the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not limitation, in the figures. Like references indicate similar elements.

Figure 1 illustrates an embodiment of the computer system using the methods and apparatus of the present invention.

Figure 2 illustrates an embodiment of the processor of the present invention.

Figure 3 is a flow diagram illustrating the general steps used by the processor to manipulate data in the register file.

Figure 4a illustrates memory data types.

Figure 4b, Figure 4c and Figure 4d illustrate in-register integer data representations.

Figure 5a illustrates packed data types.

Figure 5b, Figure 5c and Figure 5d illustrate in-register packed data representations.

Figure 6a illustrates a control signal format used in the computer system to indicate the use of packed data.

Figure 6b illustrates a second control signal format that can be used in the computer system to indicate the use of packed data or integer data.

Figure 7 illustrates one embodiment of a method followed by a processor when performing a pack operation on packed data.

Figure 8a illustrates a circuit capable of implementing a pack operation on packed byte data.

Figure 8b illustrates a circuit capable of implementing a pack operation on packed word data.

Figure 9 illustrates on embodiment of a method followed by a processor when performing an unpack operation on packed data.

Figure 10 illustrates a circuit capable of implementing an unpack operation on packed data.

DESCRIPTION OF THE PREFERRED EMBODIMENT

OVERVIEW OF ONE EMBODIMENT OF THE PRESENT INVENTION

A processor having move, pack, and unpack operations that operate on multiple data elements is described. In the following description, numerous specific details are set forth such as circuits, etc., in order to provide a thorough understanding of the present invention. In other instances, well-known structures and techniques have not been shown in detail in order not to unnecessarily obscure the present invention.

DEFINITIONS

To provide a foundation for understanding the description of the embodiments of the present invention, the following definitions are provided.

Bit X through Bit Y:
: defines a subfield of binary number. For example, bit six through bit zero of the byte 00111010_2 (shown in base two) represent the subfield 111010_2. The '$_2$' following a binary number indicates base 2. Therefore, 1000_2 equals 8_{10}, while F_{16} equals 15_{10}.

R_X:
: is a register. A register is any device capable of storing and providing data. Further functionality of a register is described below. A register is not necessarily part of the processor's package.

DEST:
: is a data address.

SRC1:
: is a data address.

SRC2:
: is a data address.

Result:
: is the data to be stored in the register addressed by DEST.

Source1:
: is the data stored in the register addressed by SRC1.

Source2:
: is the data stored in the register addressed by SRC2.

COMPUTER SYSTEM

Referring to Figure 1, a computer system upon which an embodiment of the present invention can be implemented is shown as computer system 100. Computer system 100 comprises a bus 101, or other communications hardware and software, for communicating information, and a processor 109 coupled with bus 101 for processing information.

Computer system 100 further comprises a random access memory (RAM) or other dynamic storage device (referred to as main memory 104), coupled to bus 101 for storing information and instructions to be executed by processor 109. Main memory 104 also may be used for storing temporary variables or other intermediate information during execution of instructions by processor 109. Computer system 100 also comprises a read only memory (ROM) 106, and/or other static storage device, coupled to bus 101 for storing static information and instructions for processor 109. Data storage device 107 is coupled to bus 101 for storing information and instructions.

Furthermore, a data storage device 107, such as a magnetic disk or optical disk, and its corresponding disk drive, can be coupled to computer system 100. Computer system 100 can also be coupled via bus 101 to a display device 121 for displaying information to a computer user. Display device 121 can include a frame buffer, specialized graphics rendering devices, a cathode ray tube (CRT), and/or a flat panel display. An alphanumeric input device 122, including alphanumeric and other keys, is typically coupled to bus 101 for communicating information and command selections to processor 109. Another type of user input device is cursor control 123, such as a mouse, a trackball, a pen, a touch screen, or cursor direction keys for communicating direction information and command selections to processor 109, and for controlling cursor movement on display device 121. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), which allows the device to specify positions in a plane. However, this invention should not be limited to input devices with only two degrees of freedom.

Another device which may be coupled to bus 101 is a hard copy device 124 which may be used for printing instructions, data, or other information on a medium such as paper, film, or similar types of media. Additionally, computer system 100 can be coupled to a device for sound recording, and/or playback 125, such as an audio digitizer coupled to a microphone for recording information. Further, the device may include a

speaker which is coupled to a digital to analog (D/A) converter for playing back the digitized sounds.

Also, computer system 100 can be a terminal in a computer network (e.g., a LAN). Computer system 100 would then be a computer subsystem of a computer system including a number of networked devices. Computer system 100 optionally includes video digitizing device 126. Video digitizing device 126 can be used to capture video images that can be transmitted to others on the computer network.

Computer system 100 is useful for supporting computer supported cooperation (CSC - the integration of teleconferencing with mixed media data manipulation), 2D/3D graphics, image processing, video compression/decompression, recognition algorithms and audio manipulation.

PROCESSOR

Figure 2 illustrates a detailed diagram of processor 109. Processor 109 can be implemented on one or more substrates using any of a number of process technologies, such as, BiCMOS, CMOS, and NMOS.

Processor 109 comprises a decoder 202 for decoding control signals and data used by processor 109. Data can then be stored in register file 204 via internal bus 205. As a matter of clarity, the registers of an embodiment should not be limited in meaning to a particular type of circuit. Rather, a register of an embodiment need only be capable of storing and providing data, and performing the functions described herein.

Depending on the type of data, the data may be stored in integer registers 201, registers 209, status registers 208, or instruction pointer register 211. Other registers can be included in the register file 204, for example, floating point registers. In one embodiment, integer registers 201 store thirty-two bit integer data. In one embodiment, registers 209 contains eight registers, R_0 212a through R_7 212h. Each register in registers 209 is sixty-four bits in length. R_1 212a, R_2 212b and R_3 212c are examples of individual registers in registers 209. Thirty-two bits of a

register in registers 209 can be moved into an integer register in integer registers 201. Similarly, an value in an integer register can be moved into thirty-two bits of a register in registers 209.

Status registers 208 indicate the status of processor 109. Instruction pointer register 211 stores the address of the next instruction to be executed. Integer registers 201, registers 209, status registers 208, and instruction pointer register 211 all connect to internal bus 205. Any additional registers would also connect to the internal bus 205.

In another embodiment, some of these registers can be used for two different types of data. For example, registers 209 and integer registers 201 can be combined where each register can store either integer data or packed data. In another embodiment, registers 209 can be used as floating point registers. In this embodiment, packed data can be stored in registers 209 or floating point data. In one embodiment, the combined registers are sixty-four bits in length and integers are represented as sixty-four bits. In this embodiment, in storing packed data and integer data, the registers do not need to differentiate between the two data types.

Functional unit 203 performs the operations carried out by processor 109. Such operations may include shifts, addition, subtraction and multiplication, etc. Functional unit 203 connects to internal bus 205. Cache 206 is an optional element of processor 109 and can be used to cache data and/or control signals from, for example, main memory 104. Cache 206 is connected to decoder 202, and is connected to receive control signal 207.

Figure 3 illustrates the general operation of processor 109. That is, Figure 3 illustrates the steps followed by processor 109 while performing an operation on packed data, performing an operation on unpacked data, or performing some other operation. For example, such operations include a load operation to load a register in register file 204 with data from cache 206, main memory 104, read only memory (ROM) 106, or data storage device 107. In one embodiment of the present invention, processor 109 supports most of the instructions supported by the Intel 80486™, available from Intel Corporation of Santa Clara, California. In

another embodiment of the present invention, processor 109 supports all the operations supported by the Intel 80486™, available from Intel Corporation of Santa Clara, California. In another embodiment of the present invention, processor 109 supports all the operations supported by the Pentium™ processor, the Intel 80486™ processor, the 80386™ processor, the Intel 80286™ processor, and the Intel 8086™ processor, all available from Intel Corporation of Santa Clara, California. In another embodiment of the present invention, processor 109 supports all the operations supported in the IA™ - Intel Architecture, as defined by Intel Corporation of Santa Clara, California (see Microprocessors, Intel Data Books volume 1 and volume 2, 1992 and 1993, available from Intel of Santa Clara, California). Generally, processor 109 can support the present instruction set for the Pentium™ processor, but can also be modified to incorporate future instructions, as well as those described herein. What is important is that general processor 109 can support previously used operations in addition to the operations described herein.

At step 301, the decoder 202 receives a control signal 207 from either the cache 206 or bus 101. Decoder 202 decodes the control signal to determine the operations to be performed.

Decoder 202 accesses the register file 204, or a location in memory, at step 302. Registers in the register file 204, or memory locations in the memory, are accessed depending on the register address specified in the control signal 207. For example, for an operation on packed data, control signal 207 can include SRC1, SRC2 and DEST register addresses. SRC1 is the address of the first source register. SRC2 is the address of the second source register. In some cases, the SRC2 address is optional as not all operations require two source addresses. If the SRC2 address is not required for an operation, then only the SRC1 address is used. DEST is the address of the destination register where the result data is stored. In one embodiment, SRC1 or SRC2 is also used as DEST. SRC1, SRC2 and DEST are described more fully in relation to Figure 6a and Figure 6b. The data stored in the corresponding registers is referred to as Source1, Source2, and Result respectively. Each of these data is sixty-four bits in length.

In another embodiment of the present invention, any one, or all, of SRC1, SRC2 and DEST, can define a memory location in the addressable memory space of processor 109. For example, SRC1 may identify a memory location in main memory 104 while SRC2 identifies a first register in integer registers 201, and DEST identifies a second register in registers 209. For simplicity of the description herein, references are made to the accesses to the register file 204, however, these accesses could be made to memory instead.

In another embodiment of the present invention, the operation code only includes two addresses, SRC1 and SRC2. In this embodiment, the result of the operation is stored in the SRC1 or SRC2 register. That is SRC1 (or SRC2) is used as the DEST. This type of addressing is compatible with previous CISC instructions having only two addresses. This reduces the complexity in the decoder 202. Note, in this embodiment, if the data contained in the SRC1 register is not to be destroyed, then that data must first be copied into another register before the execution of the operation. The copying would require an additional instruction. To simplify the description herein, the three address addressing scheme will be described (i.e. SRC1, SRC2, and DEST). However, it should be remembered that the control signal, in one embodiment, may only include SRC1 and SRC2, and that SRC1 (or SRC2) identifies the destination register.

Where the control signal requires an operation, at step 303, functional unit 203 will be enabled to perform this operation on accessed data from register file 204. Once the operation has been performed in functional unit 203, at step 304, the result is stored back into register file 204 according to requirements of control signal 207.

DATA AND STORAGE FORMATS

Figure 4a illustrates some of the data formats as may be used in the computer system of Figure 1. These data formats are fixed point. Processor 109 can manipulate these data formats. Multimedia algorithms often use these data formats. A byte 401 contains eight bits of

information. A word 402 contains sixteen bits of information, or two bytes. A doubleword 403 contains thirty-two bits of information, or four bytes. Thus, processor 109 executes control signals that may operate on any one of these memory data formats.

In the following description, references to bit, byte, word, and doubleword subfields are made. For example, bit six through bit zero of the byte 00111010_2 (shown in base 2) represent the subfield 111010_2.

Figure 4b through Figure 4d illustrate in-register representations used in one embodiment of the present invention. For example, unsigned byte in-register representation 410 can represent data stored in a register in integer registers 201. In one embodiment, a register, in integer registers 201, is sixty-four bits in length. In another embodiment, a register, in integer registers 201, is thirty-two bits in length. For the simplicity of the description, the following describes sixty-four bit integer registers, however, thirty-two bit integer registers can be used.

Unsigned byte in-register representation 410 illustrates processor 109 storing a byte 401 in integer registers 201, the first eight bits, bit seven through bit zero, in that register are dedicated to the data byte 401. These bits are shown as {b}. To properly represent this byte, the remaining 56 bits must be zero. For an signed byte in-register representation 411, integer registers 201 store the data in the first seven bits, bit six through bit zero, to be data. The seventh bit represents the sign bit, shown as an {s}. The remaining bit sixty-three through bit eight are the continuation of the sign for the byte.

Unsigned word in-register representation 412 is stored in one register of integer registers 201. Bit fifteen through bit zero contain an unsigned word 402. These bits are shown as {w}. To properly represent this word, the remaining bit sixty-three through bit sixteen must be zero. A signed word 402 is stored in bit fourteen through bit zero as shown in the signed word in-register representation 413. The remaining bit sixty-three through bit fifteen is the sign field.

A doubleword 403 can be stored as an unsigned doubleword in-register representation 414 or a signed doubleword in-register representation 415. Bit thirty-one through bit zero of an unsigned

doubleword in-register representation 414 are the data. These bits are shown as {d}. To properly represent this unsigned doubleword, the remaining bit sixty-three through bit thirty-two must be zero. Integer registers 201 stores a signed doubleword in-register representation 415 in its bit thirty through bit zero; the remaining bit sixty-three through bit thirty-one are the sign field.

As indicated by the above Figure 4b through Figure 4d, storage of some data types in a sixty-four bit wide register is an inefficient method of storage. For example, for storage of an unsigned byte in-register representation 410 bit sixty-three through bit eight must be zero, while only bit seven through bit zero may contain non-zero bits. Thus, a processor storing a byte in a sixty-four bit register uses only 12.5% of the register's capacity. Similarly, only the first few bits of operations performed by functional unit 203 will be important.

Figure 5a illustrates the data formats for packed data. Each packed data includes more than one independent data element. Three packed data formats are illustrated; packed byte 501, packed word 502, and packed doubleword 503. Packed byte, in one embodiment of the present invention, is sixty-four bits long containing eight data elements. Each data element is one byte long. Generally, a data element is an individual piece of data that is stored in a single register (or memory location) with other data elements of the same length. In one embodiment of the present invention, the number of data elements stored in a register is sixty-four bits divided by the length in bits of a data element.

Packed word 502 is sixty-four bits long and contains four word 402 data elements. Each word 402 data element contains sixteen bits of information.

Packed doubleword 503 is sixty-four bits long and contains two doubleword 403 data elements. Each doubleword 403 data element contains thirty-two bits of information.

Figure 5b through Figure 5d illustrate the in-register packed data storage representation. Unsigned packed byte in-register representation 510 illustrates the storage of packed byte 501 in one of the registers R_0 212a through R_n 212af. Information for each byte data element is stored

in bit seven through bit zero for byte zero, bit fifteen through bit eight for byte one, bit twenty-three through bit sixteen for byte two, bit thirty-one through bit twenty-four for byte three, bit thirty-nine through bit thirty-two for byte four, bit forty-seven through bit forty for byte five, bit fifty-five through bit forty-eight for byte six and bit sixty-three through bit fifty-six for byte seven. Thus, all available bits are used in the register. This storage arrangement increases the storage efficiency of the processor. As well, with eight data elements accessed, one operation can now be performed on eight data elements simultaneously. Signed packed byte in-register representation 511 is similarly stored in a register in registers 209. Note that only the eighth bit of every byte data element is the necessary sign bit; other bits may or may not be used to indicate sign.

Unsigned packed word in-register representation 512 illustrates how word three through word zero are stored in one register of registers 209. Bit fifteen through bit zero contain the data element information for word zero, bit thirty-one through bit sixteen contain the information for data element word one, bit forty-seven through bit thirty-two contain the information for data element word two and bit sixty-three through bit forty-eight contain the information for data element word three. Signed packed word in-register representation 513 is similar to the unsigned packed word in-register representation 512. Note that only the sixteenth bit of each word data element contains the necessary sign indicator.

Unsigned packed doubleword in-register representation 514 shows how registers 209 store two doubleword data elements. Doubleword zero is stored in bit thirty-one through bit zero of the register. Doubleword one is stored in bit sixty-three through bit thirty-two of the register. Signed packed doubleword in-register representation 515 is similar to unsigned packed doubleword in-register representation 514. Note that the necessary sign bit is the thirty-second bit of the doubleword data element.

As mentioned previously, registers 209 may be used for both packed data and integer data. In this embodiment of the present invention, the individual programming processor 109 may be required to track whether an addressed register, R_1 212a for example, is storing packed data or simple integer/fixed point data. In an alternative embodiment, processor

109 could track the type of data stored in individual registers of registers 209. This alternative embodiment could then generate errors if, for example, a packed addition operation were attempted on simple/fixed point integer data.

CONTROL SIGNAL FORMATS

The following describes one embodiment of control signal formats used by processor 109 to manipulate packed data. In one embodiment of the present invention, control signals are represented as thirty-two bits. Decoder 202 may receive control signal 207 from bus 101. In another embodiment, decoder 202 can also receive such control signals from cache 206.

Figure 6a illustrates a general format for a control signal operating on packed data. Operation field OP 601, bit thirty-one through bit twenty-six, provides information about the operation to be performed by processor 109; for example, packed addition, packed subtraction, etc.. SRC1 602, bit twenty-five through twenty, provides the source register address of a register in registers 209. This source register contains the first packed data, Source1, to be used in the execution of the control signal. Similarly, SRC2 603, bit nineteen through bit fourteen, contains the address of a register in registers 209. This second source register contains the packed data, Source2, to be used during execution of the operation. DEST 605, bit five through bit zero, contains the address of a register in registers 209. This destination register will store the result packed data, Result, of the packed data operation.

Control bits SZ 610, bit twelve and bit thirteen, indicates the length of the data elements in the first and second packed data source registers. If SZ 610 equals 01_2, then the packed data is formatted as packed byte 501. If SZ 610 equals 10_2, then the packed data is formatted as packed word 502. SZ 610 equaling 00_2 or 11_2 is reserved, however, in another embodiment, one of these values could be used to indicate packed doubleword 503.

Control bit T 611, bit eleven, indicates whether the operation is to be carried out with saturate mode. If T 611 equals one, then a saturating operation is performed. If T 611 equals zero, then a nonsaturating operation is performed. Saturating operations will be described later.

Control bit S 612, bit ten, indicates the use of a signed operation. If S 612 equals one, then a signed operation is performed. If S 612 equals zero, then an unsigned operation is performed.

Figure 6b illustrates a second general format for a control signal operating on packed data. This format corresponds with the general integer opcode format described in the "Pentium™ Processor Family User's Manual," available from Intel Corporation, Literature Sales, P.O. Box 7641, Mt. prospect, IL, 60056-7641. Note that OP 601, SZ 610, T 611, and S 612 are all combined into one large field. For some control signals, bits three through five are SRC1 602. In one embodiment, where there is a SRC1 602 address, then bits three through five also correspond to DEST 605. In an alternate embodiment, where there is a SRC2 603 address, then bits zero through two also correspond to DEST 605. For other control signals, like a packed shift immediate operation, bits three through five represent an extension to the opcode field. In one embodiment, this extension allows a programmer to include an immediate value with the control signal, such as a shift count value. In one embodiment, the immediate value follows the control signal. This is described in more detail in the "Pentium™ Processor Family User's Manual," in appendix F, pages F-1 through F-3. Bits zero through two represent SRC2 603. This general format allows register to register, memory to register, register by memory, register by register, register by immediate, register to memory addressing. Also, in one embodiment, this general format can support integer register to register, and register to integer register addressing.

DESCRIPTION OF SATURATE/UNSATURATE

As mentioned previously, T 611 indicates whether operations optionally saturate. Where the result of an operation, with saturate enabled, overflows or underflows the range of the data, the result will be

clamped. Clamping means setting the result to a maximum or minimum value should a result exceed the range's maximum or minimum value. In the case of underflow, saturation clamps the result to the lowest value in the range and in the case of overflow, to the highest value. The allowable range for each data format is shown in Table 1.

Data Format	Minimum Value	Maximum Value
Unsigned Byte	0	255
Signed Byte	-128	127
Unsigned Word	0	65535
Signed Word	-32768	32767
Unsigned Doubleword	0	$2^{64}-1$
Signed Doubleword	-2^{63}	$2^{63}-1$

Table 1

As mentioned above, T 611 indicates whether saturating operations are being performed. Therefore, using the unsigned byte data format, if an operation's result = 258 and saturation was enabled, then the result would be clamped to 255 before being stored into the operation's destination register. Similarly, if an operation's result = -32999 and processor 109 used signed word data format with saturation enabled, then the result would be clamped to -32768 before being stored into the operation's destination register.

DATA MANIPULATION OPERATIONS

In one embodiment of the present invention, the performance of multimedia applications is improved by not only supporting a standard CISC instruction set (unpacked data operations), but by supporting operations on packed data. Such packed data operations can include an addition, a subtraction, a multiplication, a compare, a shift, an AND, and an XOR. However, to take full advantage of these operations, it has been determined that data manipulation operations should be included. Such data manipulation operations can include a move, a pack, and an unpack.

THIS SHEET HAS BEEN REPLACED

Move, pack and unpack facilitate the execution of the other operations by generating packed data in formats that allow for easier use by programmers.

For further background on the other packed operations, see "A Microprocessor Having a Compare Operation," filed on _____, serial number _____, "A Microprocessor Having a Multiply Operation," filed on _____, serial number _____, "A Novel Processor Having Shift Operations," filed on _____, serial number _____, "A Method and Apparatus Using Packed Data in a Processor," filed on December 30, 1993, serial number 08/176,123, and "A Method and Apparatus Using Novel Operations in a Processor," filed on December 30, 1993, serial number 08/175,772, all assigned to the assignee of the present invention.

MOVE OPERATION

The move operation transfers data to or from registers 209. In one embodiment, SRC2 603 is the address containing the source data and DEST 605 is the address where the data is to be transferred. In this embodiment, SRC1 602 would not be used. In another embodiment, SRC1 602 is DEST 605.

For the purposes of the explanation of the move operation, a distinction is drawn between a register and a memory location. Registers are found in register file 204 while memory can be, for example, in cache 206, main memory 104, ROM 106, data storage device 107.

The move operation can move data from memory to registers 209, from registers 209 to memory, and from a register in registers 209 to a second register in registers 209. In one embodiment, packed data is stored in different registers than those used to store integer data. In this embodiment, the move operation can move data from integer registers 201 to registers 209. For example, in processor 109, if packed data is stored in registers 209 and integer data is stored in integer registers 201, then a move instruction can be used to move data from integer registers 201 to registers 209, and vice versa.

This Sheet Has Been Replaced



In one embodiment, when a memory address is indicated for the move, the eight bytes of data at the memory location (the memory location indicating the least significant byte) are loaded to a register in registers 209 or stored from that register. When a register in registers 209 is indicated, the contents of that register are moved to or loaded from a second register in registers 209. If the integer registers 201 are sixty-four bits in length, and an integer register is specified, then the eight bytes of data in that integer register are loaded to a register in registers 209 or stored from that register.

In one embodiment, integers are represented as thirty-two bits. When a move operation is performed from registers 209 to integer registers 201, then only the low thirty-two bits of the packed data are moved to the specified integer register. In one embodiment, the high order thirty-two bits are zeroed. Similarly, only the low thirty-two bits of a register in registers 209 are loaded when a move is executed from integer registers 201 to registers 209. In one embodiment, processor 109 supports a thirty-two bit move operation between a register in registers 209 and memory. In another embodiment, a move of only thirty-two bits is performed on the high order thirty-two bits of packed data.

PACK OPERATION

In one embodiment of the present invention, the SRC1 602 register contains data (Source1), the SRC2 603 register contains the data (Source2), and DEST 605 register will contain the result data (Result) of the operation. That is, parts of Source1 and parts of Source2 will be packed together to generate Result.

In one embodiment, a pack operation converts packed words (or doublewords) into packed bytes (or words) by packing the low order bytes (or words) of the source packed words (or doublewords) into the bytes (or words) of the Result. In one embodiment, the pack operation converts quad packed words into packed doublewords. This operation can be optionally performed with signed data. Further, this operation can be optionally performed with saturate.

Figure 7 illustrates one embodiment of a method of performing a pack operation on packed data. This embodiment can be implemented in the processor 109 of Figure 2.

At step 701, decoder 202 decodes control signal 207 received by processor 109. Thus, decoder 202 decodes: the operation code for the appropriate pack operation; SRC1 602, SRC2 603 and DEST 605 addresses in registers 209; saturate/unsaturate, signed/unsigned, and length of the data elements in the packed data. As mentioned previously, SRC1 602 (or SRC2 603) can be used as DEST 605.

At step 702, via internal bus 205, decoder 202 accesses registers 209 in register file 204 given the SRC1 602 and SRC2 603 addresses. Registers 209 provides functional unit 203 with the packed data stored in the SRC1 602 register (Source1), and the packed data stored in SRC2 603 register (Source2). That is, registers 209 communicate the packed data to functional unit 203 via internal bus 205.

At step 703, decoder 202 enables functional unit 203 to perform the appropriate pack operation. Decoder 202 further communicates, via internal bus 205, saturate and the size of the data elements in Source1 and Source2. Saturate is optionally used to maximize the value of the data in the result data element. If the value of the data elements in Source1 or Source2 are greater than or less than the range of values that the data elements of Result can represent, then the corresponding result data element is set to its highest or lowest value. For example, if signed values in the word data elements of Source1 and Source2 are smaller than 0x80 (or 0x8000 for doublewords), then the result byte (or word) data elements are clamped to 0x80 (or 0x8000 for doublewords). If signed values in word data elements of Source1 and Source 2 are greater than 0x7F (or 0x7FFF for doublewords) , then the result byte (or word) data elements are clamped to 0x7F (or 0x7FFF).

At step 710, the size of the data element determines which step is to be executed next. If the size of the data elements is sixteen bits (packed word 502 data), then functional unit 203 performs step 712. However, if the size of the data elements in the packed data is thirty-two bits (packed doubleword 503 data), then functional unit 203 performs step 714.

Assuming the size of the source data elements is sixteen bits, then step 712 is executed. In step 712, the following is performed. Source1 bits seven through zero are Result bits seven through zero. Source1 bits twenty-three through sixteen are Result bits fifteen through eight. Source1 bits thirty-nine through thirty-two are Result bits twenty-three through sixteen. Source1 bits sixty-three through fifty-six are Result bits thirty-one through twenty-four. Source2 bits seven through zero are Result bits thirty-nine through thirty-two. Source2 bits twenty-three through sixteen are Result bits forty-seven through forty. Source2 bits thirty-nine through thirty-two are Result bits fifty-five through forty-eight. Source2 bits sixty-three through fifty-six are Result bits thirty-one through twenty-four. If saturate is set, then the high order bits of each word are tested to determine whether the Result data element should be clamped.

Assuming the size of the source data elements is thirty-two bits, then step 714 is executed. In step 714, the following is performed.Source1 bits fifteen through zero are Result bits fifteen through zero. Source1 bits forty-seven through thirty-two are Result bits thirty-one through sixteen. Source2 bits fifteen through zero are Result bits forty-seven through thirty-two. Source2 bits forty-seven through thirty-two are Result bits sixty-three through forty-eight. If saturate is set, then the high order bits of each doubleword are tested to determine whether the Result data element should be clamped.

In one embodiment, the packing of step 712 is performed simultaneously. However, in another embodiment, this packing is performed serially. In another embodiment, some of the packing is performed simultaneously and some is performed serially. This discussion also applies to the packing of step 714.

At step 720, the Result is stored in the DEST 605 register.

Table 2 illustrates the in-register representation of a pack unsigned word operation with no saturation. The first row of bits is the packed data representation of Source1. The second row of bits is the data representation of Source2. The third row of bits is the packed data representation of the Result. The number below each data element bit is

the data element number. For example, Source1 data element three is 10000000_2.

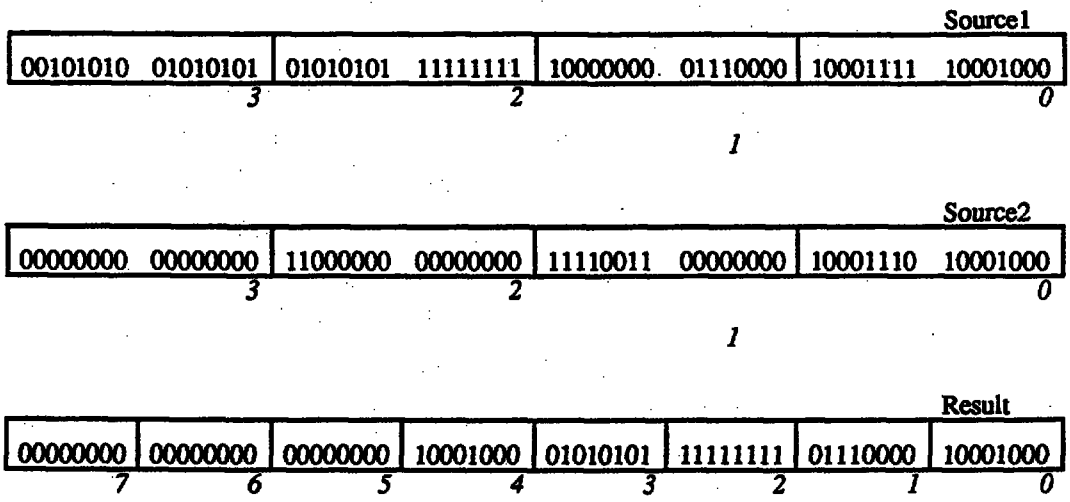


Table 2

Table 3 illustrates the in-register representation of pack signed doubleword operation with saturation.

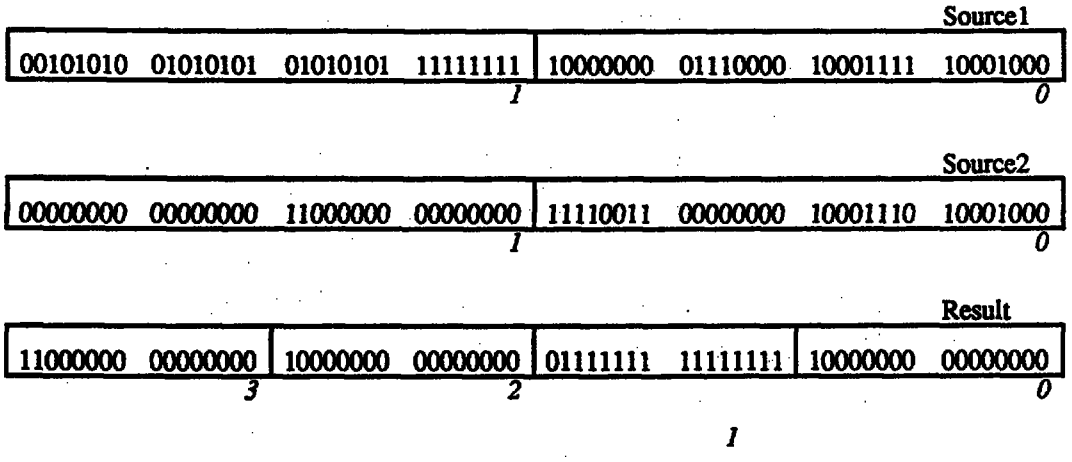


Table 3

PACK CIRCUITS

In one embodiment of the present invention, to achieve efficient execution of the pack operation parallelism is used. Figure 8a and 8b

illustrate one embodiment of a circuit that can perform a pack operation on packed data. The circuit can optionally perform the pack operation with saturation.

The circuit of Figure 8a and 8b includes an operation control circuit 800, a result register 852, a result register 853, eight sixteen bit to eight bit test saturate circuits, and four thirty-two bit to sixteen bit test saturate circuits.

Operation control 800 receives information from the decoder 202 to enable a pack operation. Operation control 800 uses the saturate value to enable the saturation tests for each of the test saturate circuits. If the size of the source packed data is word packed data 503, then output enable 831 is set by operation control 800. This enables the output of output register 852. If the size of the source packed data is doubleword packed data 504, then output enable 832 is set by operation control 800. This enables the output of output register 853.

Each test saturate circuit can selectively test for saturation. If a test for saturation is disabled, then each test saturate circuit merely passes the low order bits through to a corresponding position in a result register. If a test for saturate is enabled, then each test saturate circuit tests the high order bits to determine if the result should be clamped.

Test saturate 810 through test saturate 817 have sixteen bit inputs and eight bit outputs. The eight bit outputs are the lower eight bits of the inputs, or optionally, are a clamped value (0x80, 0x7F, or 0xFF). Test saturate 810 receives Source1 bits fifteen through zero and outputs bits seven through zero for result register 852. Test saturate 811 receives Source1 bits thirty-one through sixteen and outputs bits fifteen through eight for result register 852. Test saturate 812 receives Source1 bits forty-seven through thirty-two and outputs bits twenty-three through sixteen for result register 852. Test saturate 813 receives Source1 bits sixty-three through forty-eight and outputs bits thirty-one through twenty-four for result register 852. Test saturate 814 receives Source2 bits fifteen through zero and outputs bits thirty-nine through thirty-two for result register 852. Test saturate 815 receives Source2 bits thirty-one through sixteen and outputs bits forty-seven through forty for result register 852.

Test saturate 816 receives Source2 bits forty-seven through thirty-two and outputs bits fifty-five through forty-eight for result register 852. Test saturate 817 receives Source2 bits sixty-three through forty-eight and outputs bits sixty-three through fifty-six for result register 852.

Test saturate 820 through test saturate 823 have thirty-two bit inputs and sixteen bit outputs. The sixteen bit outputs are the lower sixteen bits of the inputs, or optionally, are a clamped value (0x8000, 0x7FFF, or 0xFFFF). Test saturate 820 receives Source1 bits thirty-one through zero and outputs bits fifteen through zero for result register 853. Test saturate 821 receives Source1 bits sixty-three through thirty-two and outputs bits thirty-one through sixteen for result register 853. Test saturate 822 receives Source2 bits thirty-one through zero and outputs bits forty-seven through thirty-two for result register 853. Test saturate 823 receives Source2 bits sixty-three through thirty-two and outputs bits sixty-three though forty-eight of result register 853.

For example, in Table 4, a pack word unsigned with no saturate is performed. Operation control 800 will enable result register 852 to output result[63:0] 860.

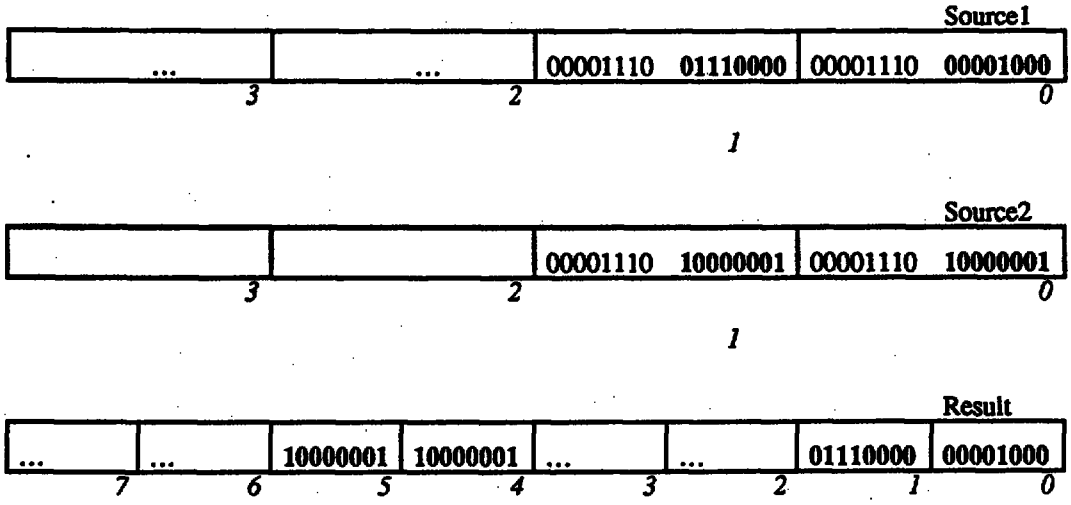


Table 4

However, if a pack doubleword unsigned with no saturate is performed, operation control 800 will enable result register 853 to output result[63:0] 860. Table 5 illustrates this result.

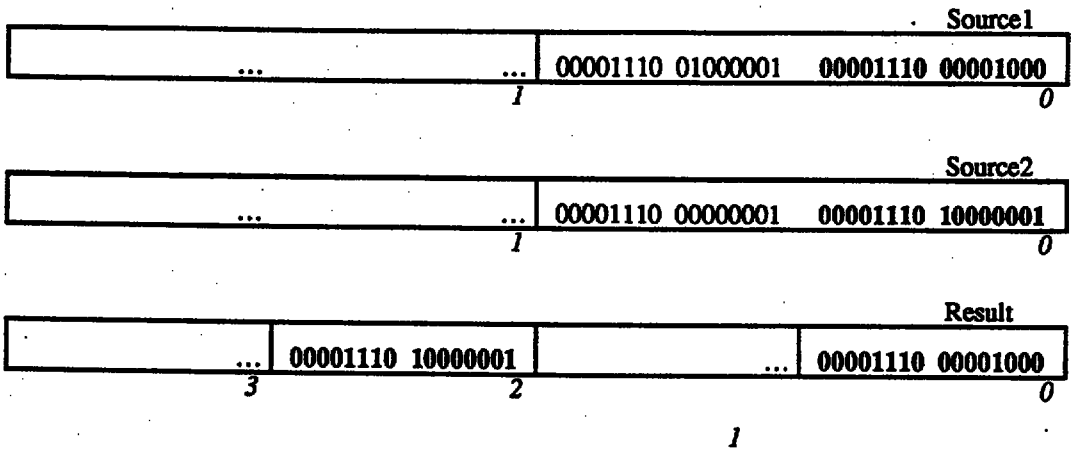


Table 5

UNPACK OPERATION

In one embodiment, an unpack operation interleaves the low order packed bytes, words or doublewords of two source packed data to generate result packed bytes, words, or doublewords.

Figure 9 illustrates one embodiment of a method of performing an unpack operation on packed data. This embodiment can be implemented in the processor 109 of Figure 2.

Step 701 and step 702 are executed first. At step 903, decoder 202 enables functional unit 203 to perform the unpack operation. Decoder 202 communicates, via internal bus 205, the size of the data elements in Source1 and Source2.

At step 910, the size of the data element determines which step is to be executed next. If the size of the data elements is eight bits (packed byte 501 data), then functional unit 203 performs step 712. However, if the size of the data elements in the packed data is sixteen bits (packed word 502 data), then functional unit 203 performs step 714. However, if

the size of the data elements in the packed data is thirty-two bits (packed doubled word 503 data), then functional unit 203 performs step 716.

Assuming the size of the source data elements is eight bits, then step 712 is executed. In step 712, the following is performed. Source1 bits seven through zero are Result bits seven through zero. Source2 bits seven through zero are Result bits fifteen through eight. Source1 bits fifteen through eight are Result bits twenty-three through sixteen. Source2 bits fifteen through eight are Result bits thirty-one through twenty-four. Source1 bits twenty-three through sixteen are Result bits thirty-nine through thirty-two. Source2 bits twenty-three through sixteen are Result bits forty-seven through forty. Source1 bits thirty-one through twenty-four are Result bits fifty-five through forty-eight. Source2 bits thirty-one through twenty-four are Result bits sixty-three through fifty-six.

Assuming the size of the source data elements is sixteen bits, then step 714 is executed. In step 714, the following is performed. Source1 bits fifteen through zero are Result bits fifteen through zero. Source2 bits fifteen through zero are Result bits thirty-one through sixteen. Source1 bits thirty-one through sixteen are Result bits forty-seven through thirty-two. Source2 bits thirty-one through sixteen are Result bits sixty-three through forty-eight.

Assuming the size of the source data elements is thirty-two bits, then step 716 is executed. In step 716, the following is performed. Source1 bits thirty-one through zero are Result bits thirty-one through zero. Source2 bits thirty-one through zero are Result bits sixty-three through thirty-two.

In one embodiment, the unpacking of step 712 is performed simultaneously. However, in another embodiment, this unpacking is performed serially. In another embodiment, some of the unpacking is performed simultaneously and some is performed serially. This discussion also applies to the unpacking of step 714 and step 716.

At step 720, the Result is stored in the DEST 605 register.

Table 6 illustrates the in-register representation of an unpack byte operation.

Source1

00101010	01010101	01010101	11111111	10000000	01110000	10001111	10001000
7	6	5	4	3	2	1	0

Source2

00000000	00000000	11000000	00000000	11110011	00000000	10001110	10001000
7	6	5	4	3	2	1	0

Result

11110011	10000000	00000000	01110000	10001110	10001111	10001000	10001000
7	6	5	4	3	2	1	0

Table 6

Table 7 illustrates the in-register representation of an unpack word operation.

Source1

00101010 01010101	01010101 11111111	10000000 01110000	10001111 10001000
3	2	1	0

Source2

00000000 00000000	11000000 00000000	11110011 00000000	10001110 10001000
3	2	1	0

Result

11110011 00000000	10000000 01110000	10001110 10001000	10001111 10001000
3	2	1	0

Table 7

Table 8 illustrates the in-register representation of an unpack doubleword operation.

Source1

00101010	01010101	01010101	11111111	10000000	01110000	10001111	10001000
			1				*0*

Source2

00000000	00000000	11000000	00000000	11110011	00000000	10001110	10001000
			1				*0*

Result

11110011	00000000	10001110	10001000	10000000	01110000	10001111	10001000
			1				*0*

Table 8

UNPACK CIRCUITS

In one embodiment of the present invention, to achieve efficient execution of the unpack operation parallelism is used. Figure 10 illustrates one embodiment of a circuit that can perform an unpack operation on packed data.

The circuit of Figure 10 includes the operation control circuit 800, a result register 1052, a result register 1053, and a result register 1054.

Operation control 800 receives information from the decoder 202 to enable an unpack operation. If the size of the source packed data is byte packed data 502, then output enable 1032 is set by operation control 800. This enables the output of result register 1052. If the size of the source packed data is word packed data 503, then output enable 1033 is set by operation control 800. This enables the output of output register 1053. If the size of the source packed data is doubleword packed data 504, then output enable 1034 is set by operation control 800. This enables the output of output result register 1054.

Result register 1052 has the following inputs. Source1 bits seven through zero are bits seven through zero for result register 1052. Source2 bits seven through zero are bits fifteen through eight for result register 1052. Source1 bits fifteen through eight are bits twenty-three through sixteen for result register 1052. Source 2 bits fifteen through eight are bits thirty-one through twenty-four for result register 1052. Source1 bits

twenty-three through sixteen are bits thirty-nine through thirty-two for result register 1052. Source2 bits twenty-three through sixteen are bits forty-seven through forty for result register 1052. Source1 bits thirty-one through twenty-four are bits fifty-five through forty-eight for result register 1052. Source2 bits thirty-one through twenty-four are bits sixty-three through fifty-six for result register 1052.

Result register 1053 has the following inputs. Source1 bits fifteen through zero are bits fifteen through zero for result register 1053. Source2 bits fifteen through zero are bits thirty-one through sixteen for result register 1053. Source1 bits thirty-one through sixteen are bits forty-seven through thirty-two for result register 1053. Source2 bits thirty-one through sixteen are bits sixty-three though forty-eight of result register 853.

Result register 1054 has the following inputs. Source1 bits thirty-one through zero are bits thirty-one through zero for result register 1054. Source2 bits thirty-one through zero are bits sixty-three through thirty-two of result register 1054.

For example, in Table 9, an unpack word operation is performed. Operation control 800 will enable result register 1053 to output result[63:0] 860.

Source1

3	2	1	0
...	...	00001110 01110000	00001110 00001000

Source2

3	2	1	0
...	...	00001110 00000001	00001110 10000001

Result

3	2	1	0
00001110 00000001	00001110 01110000	00001110 10000001	00001110 00001000

Table 9

However, if an unpack doubleword is performed, operation control 800 will enable result register 1054 to output result[63:0] 860. Table 10 illustrates this result.

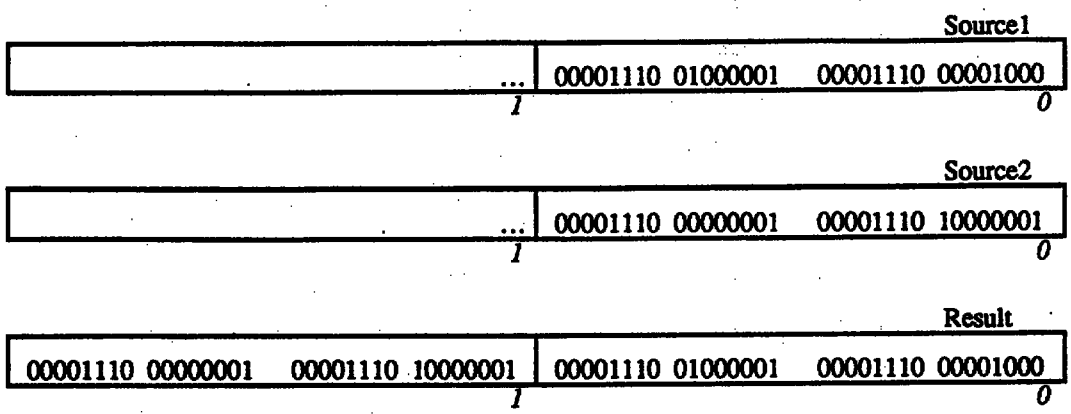


Table 10

Therefore, the move, pack and unpack operations can manipulate multiple data elements. In prior art processors, to perform these types of manipulations, multiple separate operations would be needed to perform a single packed move, pack or unpack operation. The data lines for the packed data operations, in one embodiment, all carry relevant data. This leads to a higher performance computer system.

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

THE CLAIMS

What is claimed is:

1. A processor including:

a first register for storing a first packed data;

a decoder having a control signal input, said control signal input for receiving a first control signal, and a second control signal, said first control signal for indicating a pack operation, said second control signal for indicating an unpack operation;

a functional unit being coupled to said decoder and said first register, said functional unit for performing said pack operation and said unpack operation using said first packed data.

2. The processor of claim 1 wherein said first packed data includes a plurality of data elements, each data element of said plurality of data elements has a size, and said first control signal further includes an indicator corresponding to said size.

3. The processor of claim 2 wherein said size is one of packed byte, packed word, and packed doubleword.

4. The processor of claim 2 where said first packed data is sixty-four bits.

5. The processor of claim 1 wherein said first control signal includes a source address.

* C L M S *

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED

6. The processor of claim 1 wherein said first control signal includes a sign indicator, said sign indicator for determining whether said pack operation is to be performed signed or unsigned.

7. The processor of claim 1 wherein said control signal input is further for receiving a third control signal, said third control signal for indicating a move operation, wherein said processor further includes a second register, said second register for storing integer data, and wherein said move operation causes said processor to move data between said first register and said second register.

8. The processor of claim 1 wherein said processor includes a register file, said register file includes a register, and wherein said first control signal and said second control signal include a second location and said second location corresponds to said register, wherein said functional unit is further for generating a result from an execution of an operation, and wherein said result is stored at said second location.

9. The processor of claim 1 wherein said processor first control signal includes a second location and said second location corresponds to a memory location, wherein said functional unit is further for generating a result from an execution of an operation, and wherein said result is stored at said second location.

10. A method of manipulating packed data in a processor, said processor having a decoder, a functional unit, a first register and a second register, said decoder being coupled to said functional unit, said first register and said second register, said method comprising the steps of:

said decoder decoding a control signal;



This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED

if said decoder determines that said control signal indicates a pack operation, then said processor performing the following steps,

accessing a first packed data stored in said first register,

accessing a second packed data stored in said second register,

packing said first packed data with said second packed data to generate a result packed data,

storing said result packed data in said first register;

if said decoder determines that said control signal indicates an unpack operation, then said processor performing the following steps,

accessing said first packed data stored in said first register,

accessing said second packed data stored in said second register,

unpacking said first packed data and said second packed data to generate said result packed data,

storing said result packed data in said first register; and

if said decoder determines that said control signal indicates a move operation, then said processor performing the following steps,

accessing said second packed data stored in said second register;

storing a portion of said second packed data in a third register, said third register for storing integer data.

11. The method of claim 10 wherein said packing is performed without saturation and wherein said packing includes accessing the low order bits of each data element in said first packed data and said second packed data, and wherein said low order bits are included in said result packed data.



* C L M S *

THIS SHEET HAS BEEN REPLACED

12. The method of claim 10 wherein said first packed data includes a first plurality of data elements, each data element of said first plurality of data elements being represented by a first predetermined number of bits, said control signal includes a size indicator, said size indicator indicating said first predetermined number of bits, and wherein said result packed data includes a second plurality of data elements, said second plurality of data elements having twice as many data elements as said first plurality of data elements, and each data element of said second plurality of data elements being represented by a second predetermined number, and said second predetermined number of bits being half as much as said first predetermined number of bits.

13. The method of claim 10 wherein said first packed data includes a first set of data elements, and wherein said second packed data includes a second set of data elements, and wherein said unpacking includes interleaving low order data elements of said first set of data elements with low order data elements of said second set of data elements.

14. The method of claim 10 wherein said first register is sixty-four bits long, and said first packed data includes four packed word data elements.

This Sheet Has Been Replaced



