PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁴	•		(11) International Publication Number	: WO 88/ 04806
G06F 9/38		A1	(43) International Publication Date:	30 June 1988 (30.06.88)
(21) International Application Number:	PCT/US	887/034	22 (81) Designated States: AU, JP, K	IR.

Published

(22) International Filing Date: 23 December 1987 (23.12.87)

945,486 (31) Priority Application Number:

(32) Priority Date: 23 December 1986 (23.12.86)

(33) Priority Country:

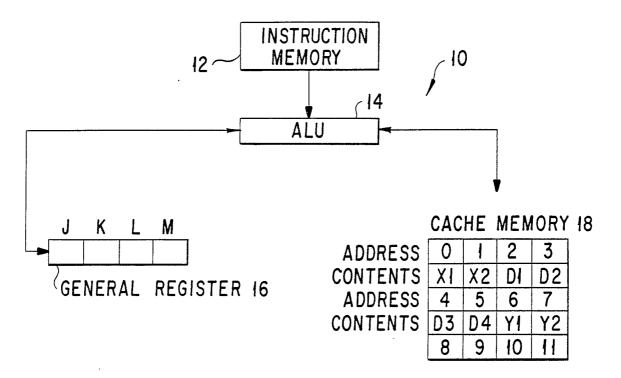
(71) Applicant: MIPS COMPUTER SYSTEMS, INC. [US/ US]; 930 Arques Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors: HANSEN, Craig, C; 14 Forest Glen, Mountain View, CA 94043 (US). RIORDAN, Thomas, J.; 148 Greenoaks Drive, Atherton, CA 94025 (US).

(74) Agents: HUETTNER, Richard, A. et al.; Kenyon & Kenyon, One Broadway, New York, NY 10004 (US).

With international search report.

(54) Title: RISC COMPUTER WITH UNALIGNED REFERENCE HANDLING AND METHOD FOR THE SAME



(57) Abstract

In a RISC device a set of four instructions from Instruction Memory (12) are provided which allow either the loading of an unligned reference into General Register (16) or the storage of an unaligned reference into Cache Memory (18). The instructions are overlapped to reduce the overall execution time of the device. An ALU circuit (14) is also provided for executing the instruction set.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

ΑT	Austria	FR	France	ML	Mali
ΑU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		

WO 88/04806 PCT/US87/03422

-1-

RISC COMPUTER WITH UNALIGNED REFERENCE HANDLING AND METHOD FOR THE SAME

5 BACKGROUND OF THE INVENTION

This invention pertains to a computer with a reduced instruction set capable of handling unaligned references, and more particularly, the reading and writing of data having fractional word length, as well as a method for handling the same.

A new development in computer architecture has been the introduction of so called RISC (Reduced Instruction Set 15 Computer) devices, in which each instruction is ideally performed in a single operational cycle. Such devices are advantageous over computers having standard architecture and instruction sets in that they are capable of much higher data processing speeds due to 20 their ability to perform frequent operations in shorter periods of times. Frequently, computers and similar data processors must be able to handle data having fractional word length. For example, although many computers are designed to handle words two or four bytes 25 in length (i.e., words of 16 or 32 bits each), certain peripheral devices and applications generate or accept data of only one or two bytes. This is often the case with data processing programs and products. One result of this type of data is that it produces an unaligned

reference. Namely, for a machine capable of handling four-byte words (32 bit devices), if incoming data is located sequentially as two bytes of data followed by four bytes of data, the four bytes of data cannot be 5 retrieved or stored in a single cycle because it would overlap a word boundary within the memory. This effect is even more problematical if a word overlaps a page boundary within the memory because, if a virtual memory system is used, only a portion of the word may actually 10 reside in addressable memory. Therefore, prior art RISC devices either do not accept data in this form, in which case special procedures must be followed to ensure that all data is aligned in word boundaries, or very involved programming is required which uses up at least two 15 consecutive instruction cycles. One way to ensure for example that all data is aligned in word boundaries would be to add extra bits to data of shorter length, usually known as bit stuffing. Whether bit stuffing is used or the programming is altered, it is obvious that 20 unaligned references seriously degrade the performance of prior art RISC devices.

Also, it should be noted that data is organized in modern computers in either of two formats or in some 25 combination of those formats. The formats are known as "big endian," in which the high order bit, byte, or other unit of information is located in the lower numbered unit address, and "little endian," in which the high order unit of information is located in the higher numbered unit address. Thus, in a true big endian 30 computer architecture, bits of data are thought of as being lined up from left to right, the lowest numbered and most significant bit being on the left. When this string of bits is divided into, for example, 8-bit bytes, 16-bit halfwords, and/or 32-bit words, the lowest numbered and most significant byte, halfword, or word continues to be located on the left. In a true little

WO 88/04806 PCT/US87/03422

-3-

endian architecture, the scheme is exactly the opposite. Bits, bytes, halfwords, and words are numbered right to left, the least significant bit, byte, halfword, or word being located on the right.

5

SUMMARY OF THE INVENTION

In view of the above, it is an objective of the present invention to provide a means and method of handling unaligned references within a RISC device.

Another objective is to provide a RISC device which is capable of loading or storing an unaligned reference in a reduced number of instruction cycles, thereby

15 maintaining a high processing speed for the device.

A further another objective is to provide a method and means of handling unaligned references which can be easily implemented, without any major changes in the 20 hardware or the operating system.

Other objectives and advantages of the system shall become apparent from the following description of the invention.

25

Briefly, a RISC device for handling unaligned references includes an instruction set which has four instructions: two instructions (Load Word Left and Load Word Right) for loading an n byte unaligned reference from a memory into an n byte general register; and two instructions (Store Word Left and Store Word Right) for storing an unaligned reference from the general register into the memory. The two instructions are used sequentially in a manner which allows the corresponding instruction

35 sequences to overlap. Therefore, the total time required to store or load an unaligned reference is much shorter than the time required to execute two

independent instructions.

The device includes several latches through which data is propagated during the above-mentioned instructions and a multiplexer register used to assemble the different sections of an unaligned reference.

BRIEF DESCRIPTION OF THE DRAWINGS

- 10 The present invention is illustrated by way of example and not limitation in the figures of the accompanying drawings, in which like references indicate similar elements, and in which:
- 15 Figure 1 shows in diagramatic form elements of an emdodiment of the present invention;

Figure 2 shows the general register after a Load Word Left instruction;

20

Figure 3 shows the general register after a Load Word Right instruction;

Figure 4 shows successive operational intervals for Load 25 Word Left and Load Word Right instructions;

Figure 5 shows the general register and the cache memory before the STORE instructions;

30 Figure 6 shows the cache memory after the unaligned reference has been stored; and

Figure 7 shows, in block diagram form, a circuit arrangement used for executing the instruction set.

35

Embodiments of the invention shall be described in connection with a 32-bit device, i.e., a RISC device in which four-byte words with eight bit bytes are handled. However, it should be understood that the means and method for handling unaligned references described herein is equally applicable to devices that handle longer or shorter words or bytes.

Further, although this description is with respect to 10 big endian addressing, it is equally applicable to little endian addressing.

With reference to the drawings, Figure 1 shows a RISC device 10 comprising an instruction memory 12 (which is 15 comprised of random access memory ("RAM"), read-only memory ("ROM"), or an instruction cache memory) which holds the instructions which make up the operation system of the device, an arithmetic logic unit ("ALU") 14, a general register 16, and a cache memory 18. The 20 · general register 16 is four bytes wide, and cells are identified in Fig. 1 as cells J, K, L, and M, respectively. Similarly, cache memory 18 is organized to hold data in rows, with each row of four bytes being addressable simultaneously. Each row therefore can be 25 identified by the first cell of the row. Thus, the cache memory is made up of rows 0, 4, 8, etc. For example, cache memory may contain a two byte data group X1, X2; a four byte data group D1, D2, D3 and D4; and another two byte group Y1 and Y2. As can be seen from 30 Figure 1, because the first group (X1, X2) is only two bytes long, the full or one word long data group D1-D4 overlaps the boundary between rows 0 and 4 of the cache memory. As a result, if a normal load instruction is used such as LOAD WORD 0 to load the contents of memory 35 row 0 into general register 16, only the first two bytes D1 and D2 are obtained. Special provisions must be made to save these bytes and then LOAD WORD 4 to obtain the

10

-6-

remaining bytes D3 and D4. This is accomplished in the present invention by using two special instructions named Load Word Left and Load Word Right hereinafter called LWL and LWR, respectively. Each of these instructions is followed by two arguments. The two instructions and their arguments are defined more specifically below:

TABLE 1 LOAD INSTRUCTIONS

Arguments Function Instruction loads the left portion R, Byte Address Load Word Left of the register R with 15 data beginning at the specified memory byte address and proceeding rightward to the memory word boundary. 20 loads the right portion Load Word Right R, Byte Address of the register R with data beginning at the memory word boundary and 25 proceeding rightward to the specified memory byte address.

30 As shown below, at the end of the fourth interval, the data bytes removed from the cache memory are saved in the general register in such a manner that they are not erased by the next load operation (LWR). This allows the bytes obtained by LWL and LWR instructions to be combined properly.

Thus, in order to load word D1-D4 from the cache memory

into the general register 16, first the following instruction is used: LWL R,2. This instruction loads bytes Dl and D2 into cells J and K, respectively, as shown in Figure 2. Thereafter, the instruction LWR R,5 is used which loads bytes D3 and D4 into cells L and M, respectively, as shown in Figure 3, thereby completing the loading of the word. In general, for a big endian device and a memory having rows four bytes wide, if the Byte Address for the LWL instruction is X, the Byte Address for the corresponding LWR instruction is X + 3.

Advantageously, the two instructions described above may be executed in an overlapping matter. Thus, the following five step sequence may be required to perform one of the load operations described above:

- Fetch instruction from the ROM (step "I");
- Read Register File (step "R");
- Compute address (step "A");
- 20 4. Access Cache Memory (step "M"); and
 - 5. Write into Register File (step "W").

These steps are taken by the ALU 14 and may be overlapped as shown in Figure 4 as follows. The first instruction -- LWL R,2 -- may start during interval 1 and end in interval 5 with each of the intervals being used for one of the steps I, R, A, M, and W as defined above. However, the second instruction -- LWR R,5 -- can start during interval number 2 as shown in Figure 4.

30 Because the device does not have to wait for the completion of the second instruction before the completion of the first, the overall speed of operation of the device is increased. Thus, the total time required to load the unaligned reference word requires only six intervals, only one interval more than the number of intervals required for a single instruction.

The pair of LOAD instructions LWL and LWR can be executed in either order, however; either LWL or LWR can come first. Furthermore, the LOAD instructions still work when they are not adjacent.

5

The above-described procedure is readily extendable to the storage of an unaligned reference. In Figure 5, general register 16 contains a four byte word El, E2, E3 and E4 which is to be stored in the same order in 10 positions Pl-P4. In order to perform this operation, the device uses the instructions Store Word Left ("SWL") and Store Word Right ("SWR"), each having two arguments. The two STORE instructions are defined in table 2 below:

15

Instruction

TABLE 2 STORE INSTRUCTIONS

Arguments

	Store	Word	Left	R,	Byte	Address		stores data from
20								the left portion
								of the register R
								into the
								specified memory
			•					byte address and
25						-	-	proceeding
					٠.			rightward to the
								memory word
						•	٠.	boundary.
								•
30	Store	Word	Right	R,	Byte	Address		stores data from
							-	the right portion
						-	-	of the register R
								into the memory
							•	byte just after
35					•			the memory word
								boundary, and
					_		÷	proceeding

PCT/US87/03422 WO 88/04806

-9-

rightward to the specified memory byte address.

5 In general, for a big endian device and a memory having rows four bytes wide, if the Byte Address for the SWL instruction is X, then the Byte Address for the corresponding SWR instruction is X + 3.

10 At the end of the first STORE instruction, bytes El and E2 are stored at addresses 2 and 3, respectively, and at the end of the second store instruction, bytes E3 and E4 are stored at addresses 4 and 5, respectively, as shown in Figure 6.

15 Like the LOAD instructions, the STORE instructions can be executed in either order; either SWL or SWR can come first. Furthermore, the STORE instructions still work when they are not adjacent.

20

A circuit for executing the four instructions is shown in block diagram form in Figure 7. This circuit may be implemented directly, or by using software. The circuit comprises a shift/merge unit 20 which receives an input 25 from cache memory 18 and generates an output which is fed (in parallel) to a latch 22. The latch 22 in turn feeds a general register 16 to be designated by the argument R in the appropriate instruction. The contents of general register 16 are propagated during each 30 operational interval though a latch 24, shift unit 26, and latch 28. Latch 28 can feed the cache memory 18. There is also a first feedback path from the output of latch 28 to a first input of bypass multiplexer unit 30. The multiplexer unit 30 has a second input connected to 35 the output of latch 22 which therefore forms a second feedback path. The output of multiplexer unit 30 is also fed to shift/merge unit 20. During the STORE

instructions, the multiplexer 30, shift/merge unit 20, and latch 22 are not in operation. During the LOAD instructions, shift unit 26 merely feeds through the data from latch 24 to latch 28 without any appreciable time delay. One of the purposes of latch 24 and latch 28 is to match the delay of the circuit path containing those latches with the number of steps making up an instruction. If the number of steps making up an instruction were increased or decreased, the number of latches in the circuit would change accordingly. The circuit of Fig. 7 operates as follows.

A LWL instruction is received during interval 1 (see Figure 4). Then in interval 4, the four bytes from the 15 row containing the address defined in the argument Byte Address are shifted to the left by the shift/merge unit 20 and merged with what had been the contents of the general register 16 two intervals earlier. (The contents of general register 16 having been fed through 20 latch 24, shift unit 26, latch 28, and bypass multiplexer 30). The results of this operation are stored in latch 22 at the end of interval 4. Thus, if row 0 is read from the cache memory 18, latch 22 will contain the bytes Dl, D2, Y, and Z, wherein Y and Z were 25 the earlier contents of general register 16 memory cells L and M. Earlier, during interval 2, instruction LWR R,5 is also received. In interval 5, the contents of latch 22 are fed to general register 16. At the same time, the LWR instruction causes the contents of the row 30 4 to be read into shift/merge unit 20. This time these bytes are shifted right until the end of the word boundary. Because the two instructions refer to the same general register and are adjacent, multiplexer 30 is now set to feed the contents of latch 22 to 35 shift/merge unit 20. Thus, during interval 5, the bytes D1, D2, D3, and D4 are assembled within the shift/merge unit 20 and fed to latch 22. During interval 6 these

WO 88/04806 PCT/US87/03422

-11-

bytes are fed to register 16.

The STORE instructions are executed as follows. The unaligned reference word is fed from the general register 16 (identified as register R) to latch 24. During the first STORE instruction -- SWL R,2 -- the word fed from latch 24 is shifted in shift unit 26 to the right by two bytes so that bytes El and E2 are in the right hand position. The contents of the shift unit 26 are then fed to latch 28, which then sends the same to the address 2 of the cache memory. More particularly, for SWL R,2, El and E2 are stored at locations Pl and P2, respectively, without disturbing the contents at memory address 0 and 1 (Figs. 5 and 6).

15

The unaligned reference word is again fed from general register 16 to latch 24. In response to the SWR R,5, the contents of latch 24 are shifted to the left so that bytes E3 and E4 are on the left side of the shift unit 26, and are then fed to row 4 by latch 28. More particularly, during SWR, bytes E3 and E4 are stored in locations P3 and P4 without disturbing the contents at addresses 6 and 7.

25 In devices in which error correction coding ("ECC") is used, a read modify write cycle is performed so that a new ECC Code is calculated after each STORE instruction.

As with the LOAD instructions, the STORE instructions

30 SWR and SWL are overlapped to reduce the overall time required to complete the instructions. Thus, the two instructions required to store the unaligned reference require only six intervals, only one interval more than the number of intervals required for a single

35 instruction. It should be appreciated that since each row of the cache memory is handled separately on an individual basis, the fact that a reference may overlap

a page boundary within the memory has no effect on the device.

It should also be noted that the pair of STORE

instructions can be executed in either order; either SWL or SWR can come first. Correspondingly, the pair of LOAD instructions can also be executed in either order; either LWL or LWR can come first. Further, the LOAD instructions still work when they are not adjacent, and the same is true with respect to the STORE instructions.

The above set of instructions are suitable for a big endian device, i.e., a device in which the leftmost bit is the most significant bit. However, the same 15 arrangement and procedure may used for a little endian device, i.e., a device wherein the leftmost bit of a byte is the least significant bit. The only change that needs to be made is to increment the address value of the arguments to the LWL and SWL instructions by 3 20 rather than to increment the arguments to the LWR and SWR instructions (as is done in the big endian device). Alternatively, a generic set of instructions could be used by changing "left" and "right" in the above instructions to "lower address" and "higher address," 25 wherein the "lower address" instructions would operate as "left" on a little endian machine and "right" on a big endian machine, and the "higher address" instructions would operate as "right" on a little endian machine and "left" on a big endian machine. This set of 30 instructions could also be used for devices which can handle both big endian and little endian data (i.e., dual byte order devices).

In the foregoing specification, the invention has been described with reference to specific exemplary embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto

without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

What is claimed is:

- 1. A method of loading an unaligned reference from a memory into a general register of a computer, said reference straddling a word boundary so that one portion of the reference is on one row and a second portion of the reference is on a next consecutive row of said memory comprising the steps of:
 - a. retrieving said first portion during an nth instruction;
 - b. retrieving said second portion during an (n+k)th instruction; and
 - c. merging said first and second portions, wherein k and n are positive integers.
- 2. A method of storing an unaligned reference into a memory of a computer, said unaligned reference straddling a word boundary, whereby said boundary partitions said word into first and second portions, comprising the steps of:
 - a. storing said first portion at a section at one row of said memory within an nth instruction;
 and
 - b. storing said second portion at a section of a next consecutive row of said memory during an (n+k)th instruction, wherein k and n are positive integers.
- 3. The method of claim 1 wherein said nth and (n+k)th instructions are overlapped.
- 4. The method of claim 2 wherein said nth and (n+k)th instructions are overlapped.
- 5. In a reduced instruction set computer, a device for retrieving an unaligned reference from a memory comprising:

- a general register;
- b. means for retrieving a first portion of said unaligned reference in response to an nth instruction and a second portion of said unaligned reference in response to an (n+k)th instruction; and
- c. combining means for combining said first and second portions, wherein k and n are positive integers.
- 6. In a reduced instruction set computer, a device for storing an unaligned reference into a memory comprising:

shifting means for shifting said unaligned reference in a first direction in response to an nth instruction and in a second direction in response to (n+k)th instruction, said means generating sequentially a first and second portion; and

means for storing said first and second portions sequentially into said memory, wherein \boldsymbol{k} and \boldsymbol{n} are positive integers.

7. In a reduced instruction set computer, a device for loading a first unaligned reference having first and second portions, said first portion being stored into a first section of a memory and said second portion being stored into a second section of said memory, and for storing a second unaligned reference into said first and said second sections, comprising:

a shift/merge unit having first and second inputs and being provided to shift first data bytes received from said first input, said first input being coupled to said memory unit to receive said first and second portions sequentially, and merge said first data bytes with second data bytes from

said second input;

a first latch means for storing said first and second data bytes, said latch having an output coupled to said second input;

a general register coupled to said first latch means and provided for holding selectively one of said first or second unaligned references;

a second latch means coupled to said register for storing said second unaligned reference;

shifting means for shifting said second unaligned references; and

output means for storing said second unaligned reference after shifting by said shifting means into said memory.

- 8. The device of claim 7 wherein said shift/merge unit shifts bytes received from said memory in a first direction in response to a first load instruction, and in a second direction in response to a second load instruction.
- 9. The device of claim 7 wherein said shifting means shifts bytes received from said second latch means in a first direction in response to a first store instruction, and in a second direction in response to a second store instruction.
- 10. The device of claim 7 further comprising a bypass multiplexer for selectively coupling to said second input one of the outputs of said first and second latching means.
- 11. The device of claim 8 wherein said first and second load instructions are at least partially overlapped.
- 12. The device of claim 9 wherein said first and second

store instructions are at least partially overlapped.

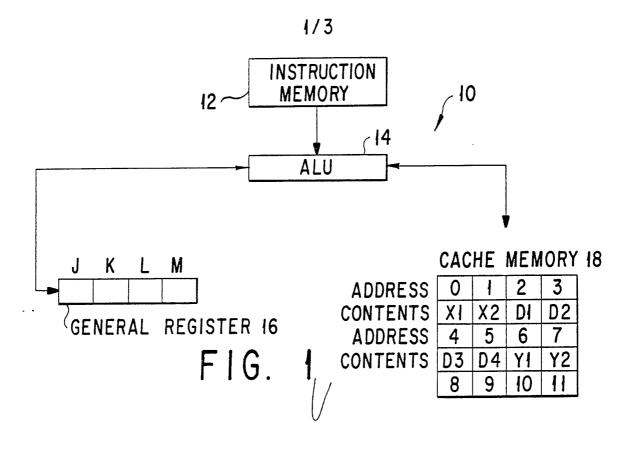


FIG. 2

FIG. 3

INSTR,	1	2	3	4	5	6
LWLR, 2	I	R	Α	М	W	
LWRR, 5		I	R	Α	М	W

FIG. 4

SUBSTITUTE SHEET

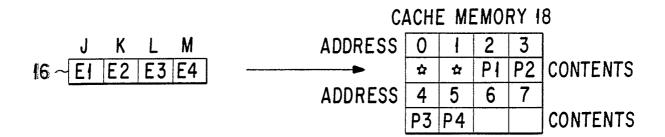
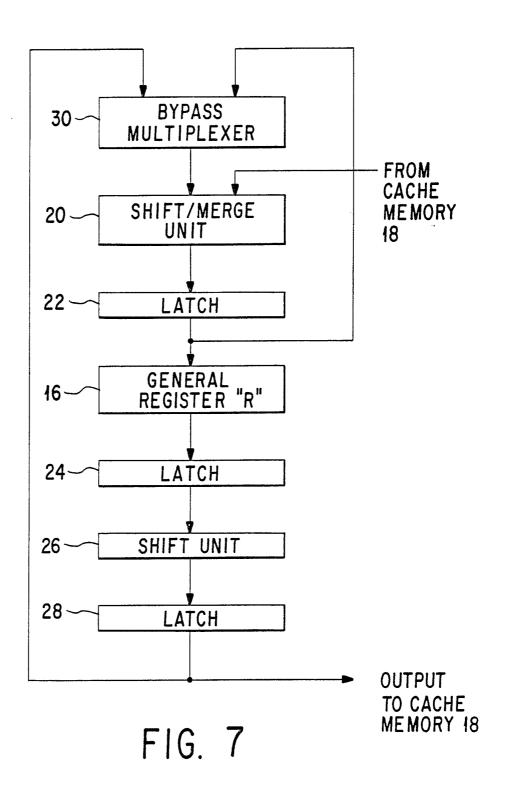


FIG. 5

CACHE MEMORY 18

0	1	2	3
☆	☆	EI	E2
4	5	6	7
E3	E4	☆	☆

FIG. 6



INTERNATIONAL SEARCH REPORT

I. CLA	SSIFICATIO	N OF SUBJECT MATTER (if several	classification symbols apply, indicate ail) 3	/US87/03422
Accord	ing to Interna	tional Patent Classification (IPC) or to bot	National Classification and IDC	
IPC	(4): G	06F 9/38	The state of the s	
_U.S	. CL.	364/200		
	DS SEARCE			
		Minimum Doc	umentation Searched 4	
Classifica	tion System		Classification Symbols	
U.S	•	364/200, 900		
		•		
		Documentation Searched of	her than Minimum Documentation lents are included in the Fields Searched 6	
	·	To the Extent that Joen Docum	ents are included in the rields Searched	
		ONSIDERED TO BE RELEVANT !	• •	
Category •	Citato	on of Document, 16 with indication, where	appropriate, of the relevant passages 17	Relevant to Claim No. 18
Y	US, A	, 4,291,370 (CHARLE	S) 22 September 1981	1_12
		See the entire do	Cument	1-12
Y	US, A	, 4,447,878 (KINNTE) 8 May 1984 See the	1 10
		entire document.	, o may 1904 bee the	1-12
		- Cocament.		
A	US, A	, 4,569,016 (HAO) 4	February 1006 car	
		the entire docume	rebruary 1986 See	•
		o to de la de de diffe		
A	US, A	, 4,339,795 (BRERET	ON) 13 July 1982 See	
		the entire docume:	nt	
A	US, A	, 4,276,596 (FLYNN)	30 June 1981 See	
ļ	1	the entire documen	nt.	
1				
A	US, A	4,271,480 (VINOT)	2 June 1981 See the	
		entire document.	The type of the	
j				
A	US, A,	4,258,419 (BLAHUT)	24 March 1981 See	
		the entire documen	it.	
A	US, A,	4,240,144 (KINDELI) 16 December 1980	
		See the entire doc	ument.	
ļ				
Special	categories o	cited documents: 15	"T" letter description authorise about	
"A" docu	ment defining	the general state of the art which is not	"T" later document published after the or priority date and not in conflict	with the application but
cons	ideled to be	of particular relevance out published on or after the international	cited to understand the principle invention	
nung	date		"X" document of particular relevance cannot be considered novel or	the claimed invention
which	h is cited to	may throw doubts on priority claim(s) or establish the publication date of another	involve an inventive step	
citati	on or other s	pecial reason (as specified)	"Y" document of particular relevance cannot be considered to involve an	inventive step when the
other	ment referring	to an oral disclosure, use, exhibition or	document is combined with one or ments, such combination being ob-	more other such docu- rious to a person skilled
"P" docum	ment published	ed prior to the international filing date but	in the art.	· i
		my sale seemed	"A" document member of the same par	ent ramity
	ICATION			
em or the	Actual Comp	etion of the International Search 3	Date of Mailing of this International Sear	ch Report *
l6 FEI	BRUARY	1988	2 5 HAR 1988	
	Searching A		Signature of Authorized Officer **	
		-	Committee of Additional Princer of	
SA/US	3		I me u // home	

III. DOCU			SIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET	
A				Relevant to Claim No
A	03,	Α,	4,090,237 (DIMMICK) 16 May 1978 See the entire document.	
A	US,	Α,	3,976,979 (PARKINSON) 24 August 1976 See the entire document.	
			in the second of	
			i i	
		-		
			<u>.</u>	
				•
ı				
;				
:				
			-	
•				
;				
<u>.</u>	•			
E Vide Bissor &				