PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : H04L 12/56	A1	(11) International Publication Number: WO 91/08633 (43) International Publication Date: 13 June 1991 (13.06.91)

(21) International Application Number: PCT/EP90/02010

(22) International Filing Date: 27 November 1990 (27.11.90)

(30) Priority data:
68059 A/89 30 November 1989 (30.11.89) IT

(71) Applicant *(for all designated States except US)*: ITALTEL SOCIETÀ ITALIANA TELECOMUNICAZIONI S.P.A. [IT/IT]; Piazzale Zavattari, 12, I-20149 Milan (IT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : BOSTICA, Bruno [IT/IT]; Via dei Pioppi, 12, I-10025 Pino (IT). DANIELE, Antonella [IT/IT]; Corso Peschiera, 179, I-10141 Torino (IT). VERCELLONE, Vinicio [IT/IT]; Via Berino, 42/3, I-10078 Venaria (IT).

(74) Agent: GIUSTINI, Delio; Italtel Società Italiana Telecomunicazioni s.p.a. , P.O. Box 10, I-20019 Settimo Milanese (IT).

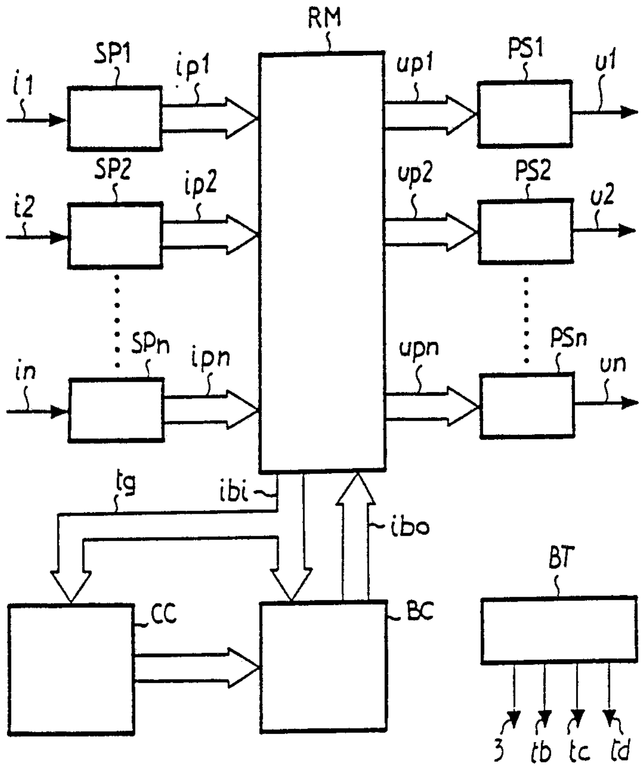
(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: BASIC ELEMENT FOR THE CONNECTION NETWORK OF A FAST PACKET SWITCHING NODE

(57) Abstract

Basic element for the interconnection network of a fast packet switching node, where a synchronization is made at bit input stream level, the cell beginning is identified and a stream conversion from the serial form to a word parallel form is performed. Cells are thus transformed in a completely parallel form and in the same form they are cyclically discharged in the subsequent cell time in a memory (BC), where cells are written and read in a shared way on the basis of instructions given by a control unit (CC), thus performing the switching function. The control unit is essentially based on the use of a content-addressed associative memory, where a fraction of the routing header and a code indicating the time sequence on which the cells arrive are stored. Memory outgoing cells are reconverted from a completely parallel form to a form having the length of one word and therefore in a completely serial form at a bitrate equal to the input one.



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.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CI	Côte d'Ivoire	LK	Sri Lanka	TD	Chad
CM	Cameroon	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark	MG	Madagascar		
ES	Spain				

"Basic element for the connection network of a fast packet switching node"

This invention refers to telecommunication systems employing digital signals for the transmission of speech, video and data signals, and in particular it refers a basic element for the connection network of a fast packet switching node.

The fast packet switching techniques, called ATM from the first letters of the wording in the English language "Asynchronous Transfer Mode", is going to take on an ever growing importance in the integrated switching of digital streams, belonging to the services for speech signal transmission, video and data signals, with different bandwith requirements and differentiated traffic characteristics. The network foreseeing this kind of service integration, with even more wide bandwith, is called B-ISDN (Broadband Integrated Service Digital Network). This technique meets better than others the requirements of the above mentioned services using an integrated switching structure, open to possible future services with not yet defined characteristics. The resources offered by the switching system are not strictly dedicated to a single call for its all length of time, as in the circuit switching systems, but are used only on demand, when the need arises to transfer information.

As known, this technique foresees that information relevant to the various services is organised in continguous units with a fixed length of approximately 400 bits, called cells. These are composed by an information field and a routing field, called header, carrying the information necessary to the route selection through the connection network and other service information.

Cells are received by line interfaces placed at the input of a switching node, essentially consisting of a control unit and of a structure performing the real switching function. The control unit performs all high level functions related to the call processing, to the configuration of the connection network and to the control of other services. Among these functions, a fundamental is the path finding. This path is decided at the call setup phase and is common to all the cells belonging to the same call. The choise is determined call by call by routing bounds throughout the geographical network and by the bandwith allocation state within the interconnection network.

The structure performing the cells switching operates by converting the header, which validity is just link by link, and the routing of cells of the same call towards the appropriate output through the connection network.

The connection network, which has the function to obtain the space switching of the cells from an input port to an output port, must be able to deliver large traffic volumes, in the range of some hundred Gbit/s, with a low cell loss probability, and low blocking probability. Furthermore, the connection network must show a minimum crossing time and has to be open to further modular growth.

Some connection networks are know at present, based on multistage structures almost non blocking, which employ unblocking switching elements of NN capacity, where N is higher or equal to 8.

Each one of these elements controls the space switching of the cells belonging to the same call, which are sent following a path unique per each input-output pair. It works in a self-routing way, since a portion of the header of the cell, called TAG, describes the route of the cell itself through the connection network and in particular the output port of each element, where the cell has to be delivered.

Since it can occur that two or more cells, arrived at different inputs, want to access to the same output port at the same time, it is necessary to foresee an intermediate storage function for the cells which cannot be immediately transferred. One cell can therefore be sent at once to the subsequent stage, while the remaining ones stand-by waiting for the availability of the output port. The known switching elements essentially differ in the way the intermediate storage of cells in conflict is performed.

According to a first method, cells are held in intermediate storages before being sent to the output through a space switching network. The storage memory is usually organized according to a FIFO discipline, in order to prevent inversions in the cells order; however this method has a drawback; in fact if the first cell which entered the memory cannot be switched due to an output conflict, it blocks all the cells arrived later, even if these are adressed towards available outputs. This can be overcome, as described in "Considerations on the structure of an ATM switch in the frame work of a hybrid BB ISDN concept", by Karl Anton Lutz, presented at IEEE COMSOC International Workshop, November 22-24, 1987, Osaka, Japan, using an access algorithm to the memories, not merely FIFO, but this requires a higher complexity of memory control units.

According to another method, described in the paper titled "The Knockout switch: a simple, modular architecture for high-performance packet switching", by Y. S. Yeh and others, published in section B10.2.1 of the proceedings of 1987 ISS, March 15-20,87, Phoenix, Arizona, USA, cells are switched towards the desired output through a crosspoint-type network performed through some buses, which operates at a speed higher than the network speed and sufficient to enable, in the worst case, to receive a cell from each input by an output. In particular, the speed increases in proportion to the number of inputs and outputs; that can originate increasing difficulties in the realization of the connection network.

A third method, described in the paper titled "Prelude: an asynchronous time-division switched network", by J. Coudreuse and others, published at section 22.2.1 of the proceeding of the 1987 ICC conference of June 8,87, Seattle, USA, foresees that all incoming cells are entered in a common memory in the switching element and that cells are drawn from the same, through an adequate control algorithm, already switched to be sent to the appropriate output. The storage is thus considered as an area to which each output port can have free access; this storage is therefore completely shared by all output ports. On the contrary, each input port is able to enter only to a dedicated area, so whenever this area fills up, the subsequent cells arriving to that input cannot fill up other free storage areas, assigned to other inputs. Therefore the capacity of the common storage cannot be completely employed.

Moreover, due to the way cells are stored, the capacity of each memory element must be equal to the number of 8-bit bytes of the cell, which heavily decreases the system flexibility in view of possible format modifications of the cells to be treated.

Finally it must be highlighted that, with equal performances from the loss probability point of view and on equal traffic conditions, the storage schemes realized in the first two solutions require a storage capacity globally higher than the one necessary in the third solution, since storage is not shared in any way, neither at input nor at output. The structures proposed in the article "A shared buffer memory switch for an ATM exchange", by Hiroshi Kuwahara and others, published at sect. 4.4.1 of the proceedings of ICC89 conference, June 11-14, 89, Boston, USA and in "Switching ATM in a Broadband ISDN", by A. J. Wiley, published on page 115 of the proceedings of Network 89 conference, Birmingham, Great Britain, can also be considered, in which cells storage is such to enable the access to a same storage area by all input and output streams, with a consequent save in the required total storage capacity. The realization of the shared access in the storage area is also such to entirely free the number of the inputs and outputs of the elements from the cell length. Access is controlled by a control unit employing a second storage area, where pointer linked lists to the data memory are realized. However, these solutions require in the second storage area an operational speed at least double than that required in the data memory.

The basic element for the connection network of a fast packet switching node, subject of the present invention, can obviate to these disadvantages, since it employs a technique for cell storage useful to minimize the amount of circuitry required for its implementation. The realization of the shared access to the storage area furthermore does not require element operation speeds higher than those defined by the speed of data flows, being also completely independent from the number of inputs and outputs of the element from the cell length.

The particular object of this invention is a basic element for the connection network of a fast packet switching node, as described in claim 1.

These and other characteristics of the present invention shall better be clarified by the following description of a preferred form of realization of the same, given as an example but not limited to, and by the attached drawings, where:

- Fig. 1, is a general block diagram of the basic element;

- Fig. 2, is a block diagram of the block marked SP1 in Fig.1;

- Fig. 3 is a block diagram of the block marked RM in Fig.1;

- Fig. 4, is a block diagram of the bloc marked RMA in Fig.3;

- Fig. 5 is a block diagram of the block marked CC in Fig.1.

The functional block diagram of the basic element for the connection network is shown in Fig.1.

Through input ports i1, i2,...in and output ports u1, u2,...un, serial information streams transit at a bitrate in the range of 150Mbit/s, made of contiguous cells having fixed length, formed by a number m of 8-bit bytes equal approx to 50. As already said, cells are made of an information fields and of a header containing the VCI (Virtual Call Identifier), which identifies the code of the call to which the cell belongs, for that node, and service information. The field containing the virtual call identifier VCI, is treated in line interfaces of the switching node located at the input of the interconnection network. In particular, it is used to address a memory, which supplies the new VCI that has to be associated to the cell for the connection between two adjacent nodes and also, it supplies a field whose bits are used for the routing of the cell itself through the elements of the interconnection network. The new VCI and the routing field are written in the abovementioned memory on the basis of information received at the moment of the call setup by the node controllers. At each stage of the network,

consisting of elements making the object of the present invention duly interconnected, a fraction of this new routing field is used and shifted to the field left, which shall be used in the subsequent stages.
A code is inserted at the beginning of the cell and it shall be used by the elements of the interconnection network to detect the cell start.
Each serial input stream, though it is isochronous with the other streams, has in general a different phase at bit and cell level due to the different length of interconnections among the different stages of the network. It is therefore required to introduce, for each input connection, a block restoring the correct phase relations. Inside these blocks, called in the picture SP1, SP2,...SPn, asynchronization at bit level of the input stream with the element clock signal, distributed by a time basis BT on wire 3, is performed, the cell start is detected and a stream conversion from the serial form to an 8-bit parallel form, supplied at the output on connections ip1, ip2,....,ipn, is performed.
The time basis BT sends on wire 3 a clock signal having a period equal to the bit time, on wire tb a clock signal with period equal to one half of the 8-bit byte time, on wire tc a clock signal with period equal to two cell times and on wire td a signal having a period equal to two cell times but with two different phases, one having a length equal to n cycles of 8-bit byte and the other one equal to the remaining 2m-n 8-bit byte cycles.
Connections ip1, ip2,..., ipn, access to a block RM in which cells are transformed in a completely parallel form and in this form are cyclically discharged in the subsequent cell time on the connection ibi, consisting of a number of wires equal to the cell bit number, towards a block BC. This last block is made of a memory in which

cells are written and read in a shared way on the basis of instructions given by a control unit CC, thus performing the switching function.

The control unit CC is essentially based on the use of a content-addressed memory, of the CAM type (Content Addressable Memory). In this memory a fraction of the routing header, present on the group of wires tg forming part of the ibi connection and relevant to the stage of the connection network to which the element considered belongs, is stored. A code indicating the time sequence of cells arrival, relating to the output specified by the abovementioned header fraction is also stored.

Using this information the control unit CC controls the selection of the appropriate cell when reading, inside the shared memory BC. When writing, cells addresses are identified starting from a bit associated to each memory location, indicating its state.

Cells outgoing the BC memory newly enter through the ibo connection the RM block which in this case reconverts them from a completely parallel form to a form having 8-bit length, made available on connections up1, up2,....,upn.

Blocks PS1, PS2, ..., PSn carry out the conversion of these streams in a completely serial form at a bitrate equal to the input one and supply them on wires u1,u2, ..., un. These are realized with shift registers, parallel loaded with a parallel 8 bit wide bus and read in a serial way at the speed determined by a clock signal supplied by the switching element time basis.

Details of one of input blocks is given in Fig.2, e.g. the one indicated by SP1. The serial stream at the element input on wire i1 is aligned by the block SB with the element internal clock signal, having bit frequency, present on wire 3. The block SB shows a structure which can be carried out according to known diagrams, e.g.

according to the diagram shown in Fig. 6 of the article titled "Technology aspects for System 12 Broadband ISDN", by Dietrich Boettle and others, published on page 1242 of IEEE Journal on selected areas in communications, October 1987.

The output stream on wire 1 access to an RSC block, where the cell start signal is detected and a corresponding signal is generated, sent on wire ta1, with a sunchronization function for the time basis BTI, which supplies at its output on wire tb1 a 8-bit byte time for the conversion of the serial stream present on wire 1 in the parallel form on eight bit, supplied on connection ip1 by a shift register SPB. The block RSC is a finite state machine which detects the cell start, triggering an appropriate synchronism code written at the beginning of the cell itself, as previously said. Both the signal on wire ta1, and the signal on wire tb1 are used for cells writing in block RM (fig.1), as it shall be described hereafter.

Fig. 3 shows the block RM, essentially consisting of two memory planes RMA and RMB. In one of the two memory planes, RMA for instance, cells are stored 8-bit byte after 8-bit byte, arriving, not necessarily in phase, at n inputs ip1,..., ipn in a cell time. Cells stored shall be discharged in the subsequent cell phases through a connection mai, a multiplexer MRI and the connection ibi towards the shared buffer marked BC in Fig.1. The multiplexer MRI is controlled by the signal coming from the element time basis on wire tc.

At the same time, the cells received by the same memory plan RMA in the previous cell phases from the block BC of Fig. 1, through the connection ibo, a demultiplexer MRO and a connection mao, are discharged towards the outputs up1,... upn through the connections a1, ..., an, and multiplexers MU1,..., MUn. All these multiplexers are placed in such a way as to receive the outputs of the digit plan RMA from the same signal on wire tc, obtaining time aligned cells at the module output.

Always in the same time phase, the n cells stored in the previous cell time are discharged in sequence in a completely parallel form from the other memory plan RMB towards the connection ibi through a connection mbi and the multiplexer MRI. This operation is made in n cycles, subsequent and clocked by the clock signal supplied on wire tb by the element time basis BT (Fig. 1).

It must be noted that the length of time in which the n cells are completely discharged towards the connection ibi is equal to n*2*tb. As n is generally lower than m (for instance, n=8, 16 or 32; m= 50) and the discharging operation of cells towards BC (fig.1) takes place in the second phase of time td, having a length of n times of 8-bit byte, an interval is left which can vary from 0 to m-n 8-bit byte times, depending on the moment of the cell arrival. This interval is used in the sequence charging plan RMB to compensate the dispersion of delays of input cell starts versus the cell time reference present on wire tc.

At the same time, from the ibo connection the switched cells coming from the memory BC (Fig. 1) are loaded in a completely parallel form through the demultiplexer MRO and a connection mbo in plan RMB in n subsequent time phases generated by the clock signal tb.

Fig. 4 shows one part of a digit plan. e.g. RMA. It is made of an 8 bit (i=1,...,n; j=1,...,m) location matrix BMij, arranged in n lines and m columns, whose locations are represented placed at crossings of the first two lines with the first two columns. Each line contains the cell which has to be written in the cell storage BC or coming from

this last; each column contains 8-bit bytes coming from the n inputs or which have to be sent to the n outputs.
In one of the two memory plans, RMA in this case, cells arriving at the n inputs ip1, ip2, ..., ipn are stored in a cell time. The storage takes place under the control of signals on wires ta1, ta2,...,tan, and tb1, tb2, ...tbn, supplying for each input the 8-bit byte sunchronism and the cell clocking, respectively, necessary to the logics SC1, SC2,..., SCn for the memory plans routing to control the correct access to the 8-bit byte column which must be written in the considered phase, through wires wc11, wc21, ..., wcnm, belonging to connections wc1, wc2,..., wcn.
The reading of the cells previously stored on the plan considered takes place under the control of an appropriate clocking logic LC linked to the internal clock signal tb and tc; it routes in writing the homonym 8-bit bytes, that is those occupying the same position inside the cell, belonging to the cells contained in the different lines of the matrix through the signals on wires rc1, rc2, ..., rcm of the connection rc.
Concerning the access to the common storage, it is controlled by two addressing logics LR and LS, which supply a reading and writing address of the matrix lines in the two appropriate phases on connections wr and rr through wires wr1, wr2,...,wrn and rr1, rr2, ..., rrn. In this way at each 8-bit byte time tb the reading of the content of one line to dump towards the cell memory and the writing of the cell memory content in the same line can be performed. Blocks SC1,...,SCn, LC, LR and LS are essentially made of counters and related decoders.
Reading and writing control signals of locations BMij are obtained through the ports OR GR11, GR12,...,GRnm and GW11, GW12, ..., GWnm to which the inputs, the row and column

reading signals and row and column writing signals are sent respectively.
The data input port of the generic location BMij receives the information on 8 bit coming from the connection ipi or from the connection maoj through a multiplexer MXij (MX11, MX12,...,MXnm); likewise on the output side MBij supplies the data or towards the connections a1,..., an or towards the connections mai1,..., main, through a demultiplexer DMij (DM11, DM12,...,DMnm). Multiplexers MXij and demultiplexer DMij are controlled by the signal on wire td, outgoing the time basis BT (Fig.1).
The control unit CC is shown on the block diagram of fig. 5, where to better clarify it is also shown the shared memory BC. The purpose of CC is to take care of the selection of the locations of the shared memory BC, both for the writing phase and for the reading phase of cells.
The operational requirements this unit has to meet are:

- the identification of locations which in each phase become free and therefore which can be used for the storage of cells arriving from block RM (Fig.1), inside the shared memory BC: this is the writing operation;
- the sorting of BC cells to be sent to block RM (Fig.1), in order to observe their order of arrival, that is performing a control of the FIFO type (first-in first-out) of cells stored for each output: this is the reading operation.

The control unit CC, as already said, includes an associative content-addressed (CAM) memory MC. This memory has a number of locations q, equal to the number of the BC memory less one, one of the BC locations being destined to store the empty cell configuration. Each MC location is in fact strictly associated to the corresponding one in BC, in such a way that the memory word contained in MC can be conveniently seen as an extension of the corresponding word of the BC memory. The employ of a content-addressed memory, in this configuration, enables to reduce to a number of words equal to BC ones the quantity of memory required to address its cells. At the same time it allows to supply a sorting of BC words already completely decoded.

The control unit CC includes also a writing sorting logic LSS, having q inputs and as many outputs, corresponding to the number of words present in MC. It makes a filtering action on logic signals present on the connection of the bv input, consisting of wires bv1, bv2,...,bvq, which consists in transferring a single active signal among those found simultaneously active at its inputs, to the corresponding outputs is1, is2,...,isq. The sorting strategy can be considered arbitrary since it is not important for the operation of the control mechanism which shall be described overleaf. For instance, the sorting could maintain active the output corresponding to the input with a lower index. The block LSS is also equipped with a further control input, brought to the block on the wire ab belonging to the connection tg, which has the function to inhibit outputs activation, which in this case are all placed at the same zero logic value. An easy implementation of this logic function, even if not the best one from the speed point of view, is a daisy chain structure.

The MC memory is also connected through a connection bd to a block SS, with the function to supply the data to write to MC from time to time in the location sorted by the logic LSS. The block SS is composed by a register bank BRS, made of n registers, as many as the element inputs or outputs are, addressed by the wire group tg, by an incrementer block INCS and by a register RS.

The data, which is stored in the register RS to be written in MC, is made of two bit fields put close together: one field which is directly received by the wire group tg of the connection ibi and one field coming from the register bank BRS.
The first field identifies the output, among the n outputs of the switching element, to which the cell is destined. This field is also used to address the content of the corresponding register inside the bank BRS through a simple decoding logic DES. The content of this register, forming the second field of the register RS, supplies the information of the time sequence belonging to the cell under writing. During the switching element initialization phase, the content of all BRS registers is setted equal to zero. Furtherly, when each register is addressed by tg through DES decoding, its output content is transferred both to register RS, and to the incrementer INCS to be incremented by one unit. The value so updated is thus written again in the same origin register. Infact whenever a cell destined to a given output is received, its sequence number, destined to assure the coherence of the order in which the cells have been queued in the shared memory BC during the subsequent reading, as well as the univoque result of the associative search, must be modulo q incremented.
The block marked SL is very similar to the block SS now described. It has the function to supply to memory MC, through the connection br, the data which must be searched from time to time inside the memory itself. This data corresponds to the presence in the shared memory BC of one cell destined to a determined output and having a definite location in terms of arrival order. Even this block SL is made of a bank of n registers BRL, addressed this time by

a counter CL through a decoding DEL, by an incrementer INCL and a register RL.

The two fields making the content of the register RL have just the function to specify to the sort mechanism the characteristics of the cell already noted above, that is: the output to which the cell is destined and its time order. The first field, identifying the output, is generated by the counter CL, which increments according to the rithm determined by the clock signal present on wire tb, already examined. The counting supplied by CL, besides being stored in the register RL, is also used to address the register bank BRL through a decoding DEL. BRL register when addressed supplies the second field to the register RL, containing the information of time order relevant to that output of the switching element. It must be noted that for each output of the element the information on time order according to which cells are written and read is very important, since the order according to which cells are to be sent to the outputs must observe the order they are received, as previously said.

Even for block SL, the second field RL is sent to the incrementer INCL to be written again in the same register, increased by one unit or possibly unchanged, depending on the state of the signal on wire ht. This signal carries the information on the search result inside MC of the word presented by the register RL. If the search was successfull, that is if the word is present in MC, the signal on wire ht increases in INCL the content of the register of addresses BRL; if unsuccessful the content of the register itself remains unchanged.

The control unit CC includes also a block RV which is made of a rank of 1 bit registers SR1, SR2,...., SRq, in the same number as words contained in MC, that is q. It also contains the blocks A1, A2,...,Aq having the function to

generate the reading sorting for BC and a block M which generates the abovementioned ht signal and the denied signal of this one on the wire cv.
The block RV maintains, for each one of the words present in MC, through the state of SR1,...,SRq, an information on the present validity of the content of the subject word. This information defines in fact if the word to which it refers has already been used and therefore it can be written again with a new word, or if it has a still valid content. In this second case the BC corresponding word contains a cell which is still waiting for its turn to be sent towards the required output.
Finally the control unit CC includes a block RI, made both of an address register of (q+1) bit which, maintaining steady the sorting processed by the control unit during the previous cycle, enables to create a superimposition among the access cycles to memory MC and to memory BC, and of multiplexers necessary to alternate on this register the inputs supplied by the block LSS on wires is1, is2,...,isq and those coming from blocks A1, A2,...Aq of RV.

The content addressed memory MC is able to operate as a conventional memory during the writing phase. The outputs of the LSS logic, select the location where the writing of the word supplied by the register RS has to be made, putting at the same time at the logic state one the corresponding register in the rank SR1, ...,SRq of RV, thus indicating that the location itself was occupied by a valid word. In the same way LSS outputs are stored in the register contained in RI, where they shall be required to address the location of the BC memory where the cell coming from RM shall be stored (Fig.1).
During the associative search phase, the content supplied by the register RL is on the contrary compared in parallel

to all the words present in MC. The comparison result, relative to each MC location, is made available at output by MC on the group of 2q wires directed towards RV block. In fact this group of wires is made of as many wire pairs as are the MC location, one wire for the comparison result and the other one (is1, is2,...,isq) for the same sorting signal sent by LSS to MC. Each pair of wires is connected to the inputs of the corresponding register of the rank SR1,...,SRq contained in RV, whose state is newly resetted any time the comparison result indicates that the corresponding data, contained in MC, matches the one presented on the connection br. For this reason the outputs of each register Sr1,...,SRq and the corresponding signal generated by MC have access to the inputs of a block in the rank A1, A2,...,Aq, inside which the signal carrying the comparison result is conditioned by the information of validity of the relative MC word, represented by the register state of the rank SR1,...,SRq, producing the final outputs of RV, which shall be stored in the register contained in RI.

The block M, by processing the state of all these final outputs, detects if one of them is active, thus generating the signal on wire ht, used to drive the incrementer INCL in SL.

Finally, while on one hand the signals coming out from registers SR1,...,SRq of RV identify all the free MC locations and which can therefore be used to wire new data, through the connection bv and the logic LSS, on the other hand, the same signals have also the function to define which one of the locations of MC which had a positive result in the associative search, contains also a currently valid word. Only in this case the positive result is transformed by the rank of blocks A1,...,Aq of

RV in the sorting of the corresponding location of memory BC, through the register RI.

In summary, the operation sequence of the control unit CC is in short the following. A time phase, corresponding to a 8-bit byte time, is divided in two parts. In the first phase, the cell to store in the BC memory is present on the connection ibi. If the cell is not empty, which condition is indicated by the signal on the wire ab, the field identifying the output to which the cell is destined is present on the group of wires tg belonging to the above mentioned connection. The signals on tg contribute both directely and indirectly through the register bank BRS, to form the content of register RS, which is then written in MC. This data, written in the MC memory, is in fact the information necessary to the subsequent recovery of the corresponding cell which shall be stored in the BC memory in the time phase immediately subsequent. If on the contrary the signal on wire ab informs that it is an empty cell, LSS will not generate the writing address and consequently the writing shall not occur neither in MC, nor in BC.

In the second phase, there is the possible real writing of the cell, still present on the connection ibi, in the BC memory. To this purpose the address which the LSS logic, through the register contained in RI and the corresponding multiplxers generated in the previous phase, during which it needed to the writing in MC, is used again.

In the meantime in this same phase the control prepares the address of the BC location whose content, in the subsequent reading cycle, shall be transferred to RM through the output connection ibo. During this phase, infact, the MC memory is the object of the associative search of the word supplied by the register RL of block SL. In case this search is unsuccessful, this means that in BC no cell for the specified output is present; consequently the block M generates the ht signal, which function has already been examined, and its complement cv which is stored in the approrpiate register of RI. This signal become therefore the sorting signal ma0, addressing the BC location which contains the empty cell code to transmit to the output. The cycle is again repeated with a phase in which a writing operation in MC takes place, while in BC memory it takes place the reading of the cell addressed from the active sorting line among the ones, ma0, ma1,...,maq outgoing RI. The shared memory BC is a random access memory (RAM) with separated input and output connections, ibi and ibo respectively, both consisting of a number of wires equal to the number of bits of one cell (m*8 bit). This memory is made of (q+1) locations; the value of q is defined by statistical evalutations and is in the range of 150; the location addressed by the ma0 wire contains an empty cell code. At each cell time all the information stored in one of the two plans RMA or RMB (Fig. 3) of the RM block are transferred in the BC memory with a clocking equal to the 8-bit byte time and in the first half of this time, on the basis of decoded addresses supplied by the control unit CC on lines ma0,...,maq. In the second half of each 8-bit byte time, the switched cells are read by the cell BC memory, which are made available on the ibo connection. In case no cells are directed to a particular output, the control unit CC sends the empty cell code to the connection through activation of the decoding wire ma0 of the first cell of the BC memory, in which this code has been charged during the element initialization phase.

It is clear that the above description has been given as an example but not limited to it. Variants and modifications are possible within the claims protection field.

Claims

1. Basic element for the connection network of a fast packet switching node, with cells of fixed length and formed by an information field and by a header, expressed with a number m of words, including:

- blocks (SP1, SP2...,SPn) to which input ports (i1, i2,...,in) the serial information streams have access, made of contiguous cells, and where the synchronization at bit level of input streams, the detection of cells beginning and the conversion from the serial form to the word parallel form;
- a time basis (BT);
- a memory block (RM) in which incoming cells in the word parallel form are transformed in a completely parallel form and vice versa for outgoing cells;
- a shared memory (BC) in which cells are written and read in a shared way, performing the switching function;
- a control unit (CC), controlling the reading and writing operations in said memory (BC)
- block (PS1, ...,PSn) performing the conversion of streams outgoing the memory block (RM) in serial streams (u1, u2,...,un) at a bitrate equal to the incoming one;

and characterized by the fact that said memory block (RM) is essentially made of two memory plans (RMA, RMB) in one of which are stored word after word in a cell time the cells arriving, on different instants, from said blocks (SP1,SP2,...,SPn) and at the same time, the cells which the same memory plan received from the memory (BC) are discharged towards the outputs (up1,...,upn) through the multiplexer (MU1,...,MUn) controlled by a single signal, obtaining at the output time aligned cells, while always in the same time phases, the cells stored in the previous cell time are discharged in a completely parallel form from the other memory plan towards the shared memory (BC) through the multiplexer (MRI) and the switched cells coming from the same memory (BC) are charged in a completely parallel form through the demultiplexer (MRO), being these operation alternatively performed in the two memory plans.

2. Basic element as in claim 1, characterized by the fact that each one of said digit plans (RMA, RMB) includes:

- one matrix of locations (BMij: i=1,...,n j=1, ...,m) for said words, organized in row (n) and columns (m), where each row contains the cell which must be written in the memory (BC) or which come from the same and each column contains the words coming from the inputs (ip1, ip2,...,ipn) or the words which have to be sent at outputs (a1,...,an);
- some column routing logics (SC1, SC2,..., SCn) to control (wc11, wc21,...,wcnm) the writing in each single column, under the control of synchronization signals for words (ta1,...,tan) and for cells (tb1,...,tbn) released by said blocks (SP1, SP2,...,SPn), regardless of the delays dispersion of cell start signals at inputs (ip1,...,ipn);
- a clocking logic (LC) synchronized with signals (tb, tc) released by said time basis (BT) for the routing in reading (rc1, rc2,...,rcm) of homonym words;
- two routing logics (LR, LS) of the lines of matrix under reading (rr1, rr2, ..., rrn) and writing (wr1, wr2,...,wrn), synchronized by a signal (tb) released by said time basis (BT);
- ports OR (GR11, GR12,...,GRnm; GW11, GW12,...,GWnm) for enabling said locations (BMij) to line and column reading operations and line and column writing ones, on the basis of signals released by said routing logics (LR, LS, SC1, SC2,..., SCn) and clocking logics (LC);

- some multiplexers (MX11, MX12,..., MXnm) through which the words coming from the inputs (ip1, ip2,...,ipn) or from said memory (BC) are sent at the input of said locations (BMij);
- some multiplexers (DM11, DM12,...,DMnm) through which the words outcoming said locations (BMij) are sent either to the outputs (a1,...,an) or to said memory (BC).

3. Basic element as in claim 1, characterized by the fact that said control unit (CC), includes a content-addressed memory (MC) of associative type, supplied with a number of locations (q), equal to the one of said memory (BC) less one, where a fraction of the cell heading is stored, relevant to the routing and a code indicating the time sequence of the arrival of the cells, relating to the specified output of the abovementioned fraction, and with the use of this information controls the sorting during reading of the cell which is appropriate from time to time, inside the memory (BC), while in writing the cells addresses are identified starting from a bit associated to each memory location, which indicates the busy state.

4. Basic element as in claims 1 or 3, characterized by the fact that said control unit (CC) includes also:
- a sorting logic as for writing (LSS), having a number (q) of inputs (bv1,...,bvq) and of outputs (is1,...,isq), equal to the number of words present in said memory of the content-addressed type (MC), which transfers to the corresponding outputs a single active signal among those active at the same time at its inputs according to an appropriate criterion, and outputs can be inhibited through a control wire (ab);
- a first block (SS), which has the function to supply through a first connection (bd) to said memory (MC) the data to write from time to time in the location selected by the logic (LSS);

- a second block (SL), which has the function to supply the data which must be from time to time searched inside the memory, through a second connection (br) to said memory (MC);
- a third block (RV), which keeps for each one of the words present in the memory (MC) an information on the present validity of the content of the word itself;
- a fourth block (RI), which enables to alternate at the output the signals supplied by the logic (LSS) with those coming from the third block (RV), maintaining them steady for the access cycle to the shared memory (BC).

5.- Basic element as in claim 4, characterized by the fact that said first block (SS) includes:
- a register bank (BRS), consisting of a number of registers equal to the one of the element inputs or outputs, addressed through a decoding (DES) from said fraction of the cell heading (tg);
- an incrementer block (INCS) receiving the content of a bank register (BRS) and re-writes it incremented in the same register;
- a register (RS), in which are written the fraction of the cell heading (tg) and the content of the bank register (BRS) forming the data to write in a location of said memory (MC) through said first connection (bd).

6. Basic element as in claim 4, characterized by the fact that said second block (SL) includes:
- a counter (CL) generating a binary digit which increases at the rithm of a clock signal (tb);
- a register bank (BRL), consisting of a number of registers equal to the element inputs or outputs number, cyclically addressed through said binary digit through a decoding (DEL);
- an incrementer block (INCL) receiving the content of one bank register (BRL) and which writes it again

incremented in the same register, when enabled by an appropriate signal (ht);

- one register (RL), in which the binary digit generated by said counter (CL) and the content of a bank register (BRL) forming the data to find in a location of said memory (MC) through said second connection (br), are written.

7. Basic element as in claim 4, characterized by the fact that said third block (RV) includes:

- a rank of 1 bit registers (SR1, SR2,...,SRq), in a number (q) equal to that of the locations of said memory (MC) and having an input connected to the corresponding memory (MC) output, on which the comparison result is available and the other input (is1, is2,...,isq) for the same sorting signal generated by said logic (LSS), the outputs of the registers of the rank being connected to the inputs (bv1, bv2,...,bvq) of said logic (LSS);

- a rank of blocks (A1, A2,...,Aq), in a number (q) equal to the one of the locations of said memory (MC) and having an input connected to the corresponding memory (MC) output, on which the comparison result is available, and the other input for a conditioning signal drawn at the output of the corresponding register of said rank of registers;

- a block (M) which generates at the output a signal (ht) and its complement (cv) if one of the signals outgoing said block rank (A1, A2,...,Aq) is active and sends them to the second block (SL) and to the fourth block (RI), respectively.

8. Basic element as in claim 1, characterized by the fact that said blocks (SP1, SP2,...,SPn) to which input ports (i1, i2,...,in) the serial information streams have access, including:

- a fifth block (SB) for the alignment of serial streams with a clock signal having bit frequency generated by said time basis (BT);
- a sixth block (RSC), consisting of a finite state machine, in which it is detected the beginning of the cell outgoing from said fifth block (SB) and generated by a corresponding signal (ta1);
- a time basis (BT1), synchronized by the signal (ta1) generated by the sixth block and suitable to generate a word clocking (tb1);
- a shift register (SPB) suitable to perform the serial conversion of the stream supplied by the fifth block in a word parallel stream, under the control of the clocking signal supplied by said time basis (BTI).

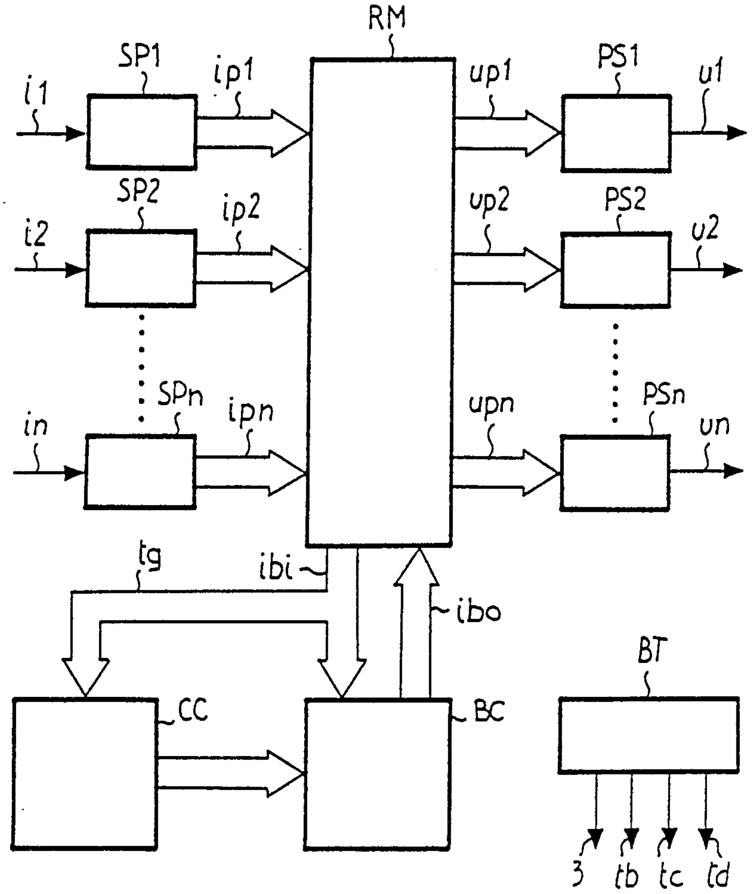

RM
SP1
ip1
i1
SP2
ip2
i2
SPn
ipn
in
up1
PS1
u1
up2
PS2
u2
upn
PSn
un
tg
ibi
ibo
CC
BC
BT
3
tb
tc
td

FIG. 1

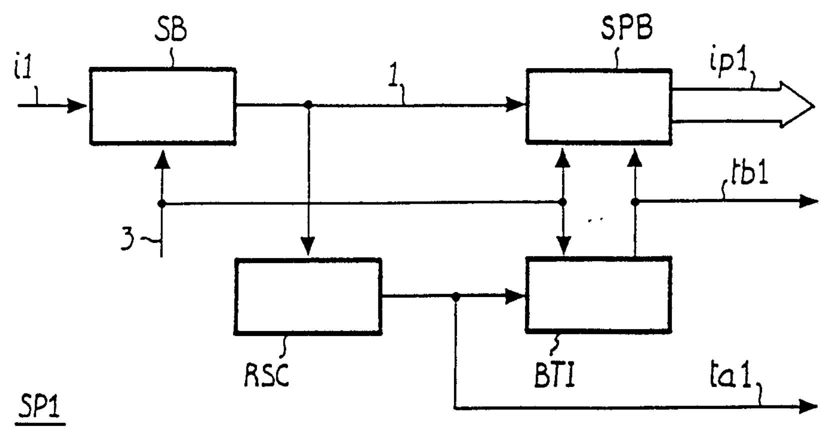

SB
SPB
i1
1
ip1
tb1
3
RSC
BTI
ta1
SP1

FIG. 2

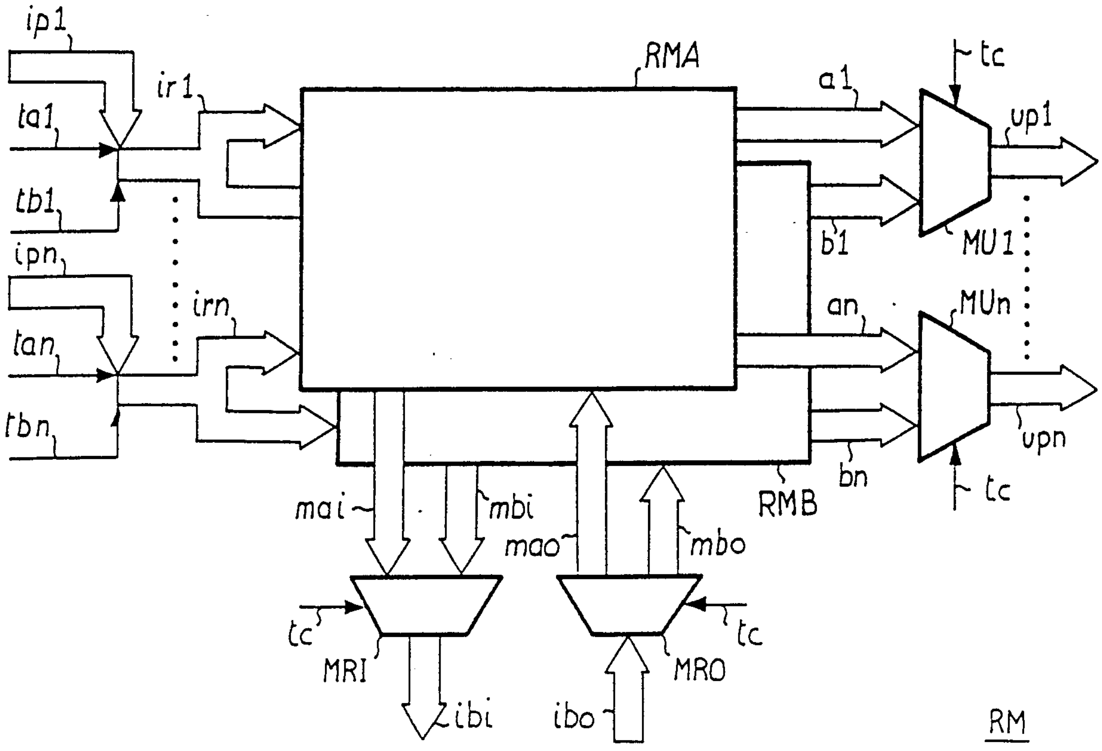

ip1
ir1
RMA
a1
tc
up1
ta1
tb1
b1
MU1
ipn
irn
an
MUn
tan
tbn
upn
mai
mbi
mao
mbo
RMB
bn
tc
tc
MRI
MRO
tc
ibi
ibo
RM

FIG. 3

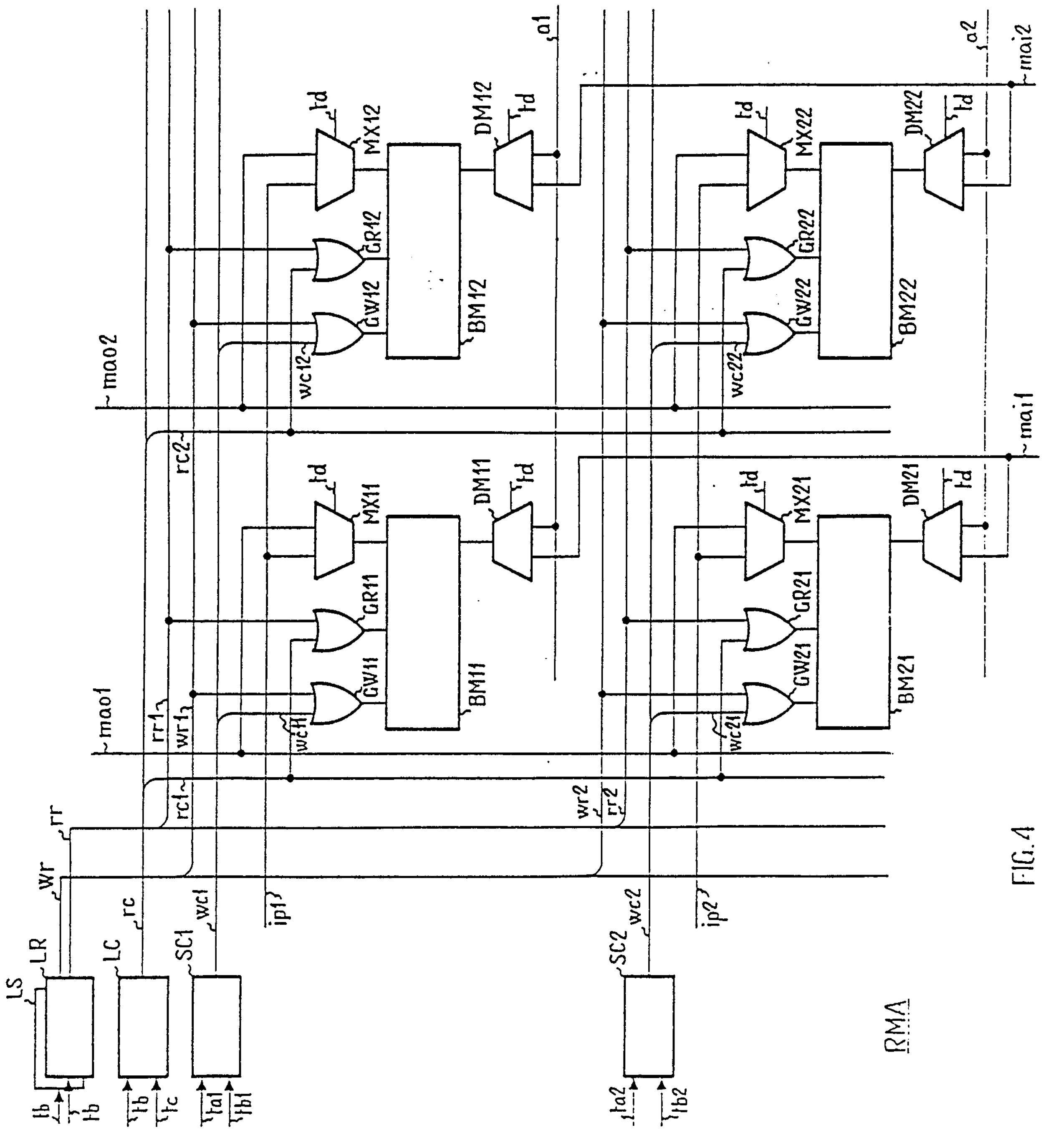

LS
LR
LC
SC1
SC2
wr
rr
rc
wc1
wc2
ip1
ip2
rr1
wr1
rc1
rc2
wr2
rr2
mao1
mao2
mai1
mai2
a1
a2
wc11
wc12
wc21
wc22
GW11
GR11
MX11
BM11
DM11
GW12
GR12
MX12
BM12
DM12
GW21
GR21
MX21
BM21
DM21
GW22
GR22
MX22
BM22
DM22
td
tb
tc
ta1
tb1
ta2
tb2
RMA

FIG. 4

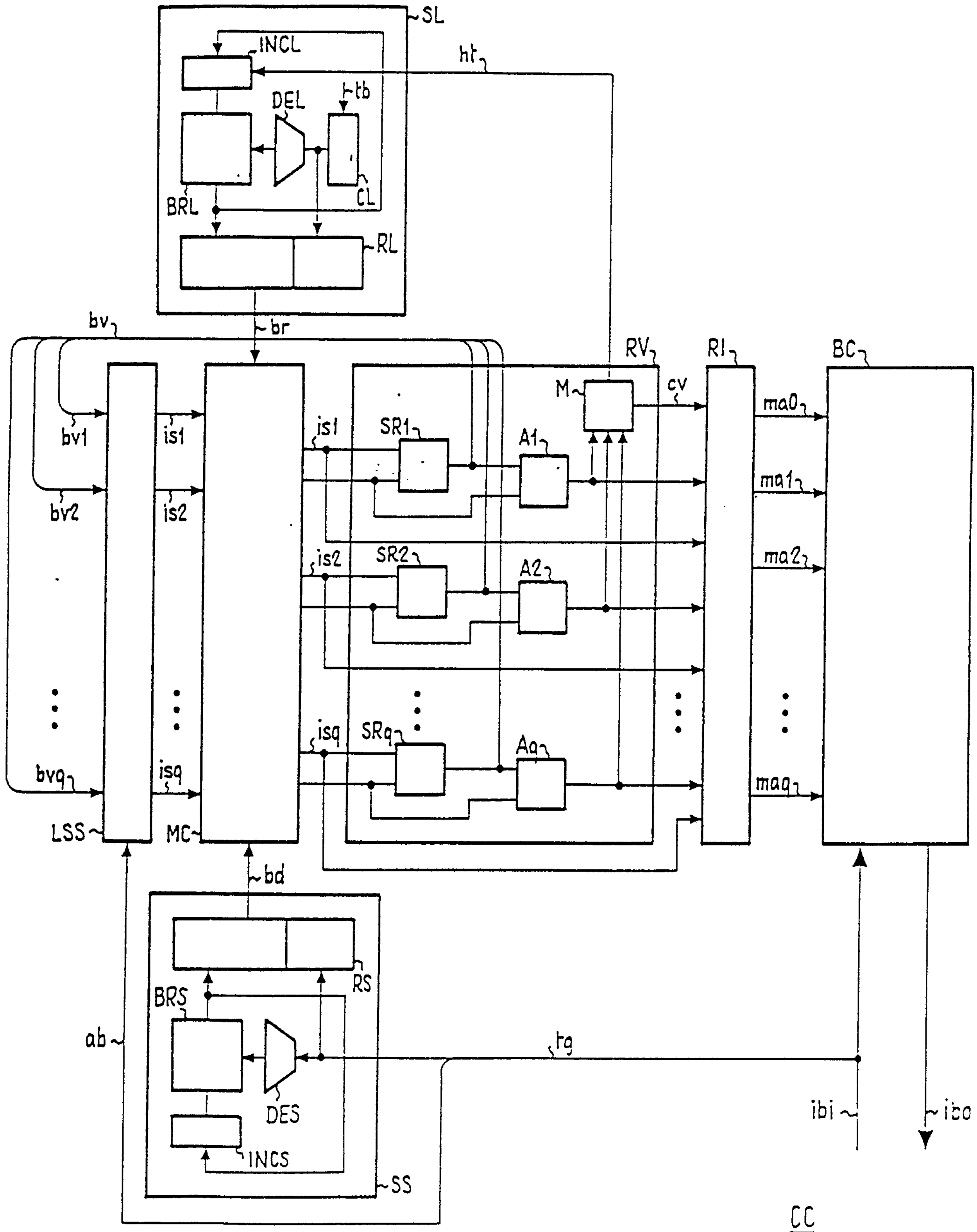

SL
INCL
ht
DEL
tb
BRL
CL
RL
bv
br
RV
RI
BC
M
cv
ma0
bv1
is1
is1
SR1
A1
ma1
bv2
is2
is2
SR2
A2
ma2
isq
SRq
Aq
bvq
isq
maq
LSS
MC
bd
RS
BRS
ab
tg
DES
INCS
ibi
ibo
SS
CC

FIG. 5

INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 90/02010

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC^5: H 04 L 12/56

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC^5	H 04 L

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [*]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
A	EP, A, 0338558 (NEC) 25 October 1989 see column 10, line 14 - column 11, line 9; figures 6,7 --	1,2
A	FR, A, 2504760 (WESTERN ELECTRIC) 29 October 1982 see page 12, line 20 - page 16, line 25; figure 11 --	1-3
A	IEEE International Conference on Communications, Seattle, Washington, 7-10 June 1987, vol. 1, IEEE, J.-P. Coudreuse et al.: "Prelude: an asynchronous time-division switched network", pages 769-773 see the whole article cited in the application ----------	1

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
20th February 1991	25. 03. 91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	N. Toribio Nuria TORIBIO

Form PCT/ISA/210 (second sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. EP 9002010

SA 41823

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 18/03/91
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0338558	25-10-89	JP-A- 1270431	27-10-89
		JP-A- 2067045	07-03-90
FR-A- 2504760	29-10-82	US-A- 4382295	03-05-83
		AU-B- 546537	05-09-85
		AU-A- 8293382	28-10-82
		BE-A- 892928	16-08-82
		CA-A- 1170379	03-07-84
		CH-A- 636487	31-05-83
		DE-A,C 3214189	11-11-82
		GB-A,B 2098831	24-11-82
		JP-A- 57181294	08-11-82
		NL-A- 8201678	16-11-82
		SE-B- 445869	21-07-86
		SE-A- 8202233	24-10-82

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82