

## Description

The present invention relates generally to packet switching networks employed in multi-processor and parallel computer systems, and digital communications systems, and the like, and more particularly to packet switching nodes which employ a multiport memory and control logic coupled thereto which is employed to store and forward data packets to a plurality of output ports and which reduces output port contention among data packets arriving at the same input port.

One developing area of computer technology involves the design and development of large-scale, multi-processor-based distributed and parallel computer systems. Typical of these classes of computer systems and architectural approaches are the single instruction stream, multipledata stream (SIMD) computer architecture and the multiple instruction stream, multiple data stream (MIMD) computer architecture.

A SIMD computer typically comprises a control unit, N processors, N memory modules and an interconnection network. The control unit broadcasts instructions to all of the processors, and all active processors execute the same instruction at the same time. Each active processor executes the instruction on data in its own associated memory module. The interconnection network provides a communications facility for the processors and memory modules.

A MIMD computer typically comprises N processors and N memories, and each processor can execute an independent instruction stream. Each of the processors may communicate to any other processor. Similar interconnection networks may be employed in the MIMD computer.

Various interconnection networks may be employed to interconnect processors and memories employed in either type of computer system. These interconnection networks include delta networks, omega networks, indirect binary n-cube networks, flip networks, cube networks and banyan networks, for example.

The above-cited networks are discussed in some detail in the following publications: "LSI implementation of modular interconnection networks for MIMD machines," *1980 Int'l Conf. Parallel Processing*, Aug. 1980, pp. 161-162; "Analysis and simulation of buffered delta networks," *IEEE Trans. Computers*, Vol. C-30, pp. 273-282, April 1981; "Processor-memory interconnections for multiprocessors," *6th Annual Int'l. Symp. Computer Architecture*, April 1979, pp. 168-177; "Design and implementation of the banyan interconnection network in TRAC," *AFIPS 1980 Nat'l. Computer Conf.* June 1980, pp. 643-653; "The multistage cube: a versatile interconnection network," *Computer*, Vol. 14, pp. 65-76, Dec. 1981; "The hybrid cube network," *Distributed Data Acquisition Computing and Control Symp.* Dec. 1980, pp. 11-22; and "Performance and implementation of 4x4 switching nodes in an interconnection network for PASM,"

*1981 Int'l Conf. on Parallel Processing*, Aug. 1981, pp. 229-233.

Several types of data switching techniques may be employed to transfer data in SIMD and MIMD computers, and the like, including packet switching, message switching, time-division circuit switching or space-division circuit switching. Packet switching involves sending one or more words of data at a time through the system.

Conventional packet switching interconnection networks and nodes have a well-known problem involving the speed of transmission of information through the network. Conventional designs have typically employed a single queue coupled to each input port of the network to store and forward signal packets to all the output ports. In the single queue system, a contention problem occurs due to the fact that a signal packet destined for output port 2, for example, is blocked from exiting through that port because a single packet destined for output port 1 is physically ahead of it in the queue and has not yet exited through port 1. This contention problem causes unnecessary transport delays and thus reduces system throughput.

One packet switching node which attempts to overcome this problem is described in a presently copending patent application entitled "Packet switched multiple queue  $N \times M$  switch node," invented by R. J. McMillen and assigned to the assignee of the present invention. This invention utilizes a plurality of queue sets individually connected to each of the input ports in order to sort the signal packets according to output port destination. However, although this node improves the communications capability of computer systems employing it, it requires a large buffering capacity. This implies the use of a large number of storage transistors in the logic circuitry.

According to the present invention there is provided a packet switching node which processes applied data packets containing routing tag signals indicative of the output port destination to which said data packets are to be applied, said packet switching node comprising a plurality of input ports; a plurality of output ports; a multiport memory coupled between said input and output ports, said memory having a predetermined number of memory locations available for storage of said data packets applied to each of said input ports; and control logic coupled to said input and output ports and said multiport memory for controlling the storage of data packets in said memory and the routing thereof to said output ports in accordance with said routing tag signals, wherein the improvement is in said control logic characterized by:

a status register coupled to said input ports and said memory for storing said routing tag signals and status signals indicative of the full/empty status of each of the memory locations in said memory;

priority encoder logic coupled to said input ports, said memory and said status register for

monitoring the full/empty status of said memory locations, writing said routing tag signals into registers of said status register and writing data packets corresponding to said routing tag signals into selected memory locations;

an arbitration logic coupled to said status register, said memory and said output ports for reading said data packets from said memory locations and applying them to the output ports identified in the corresponding routing tag signals based upon a predetermined priority arbitration scheme;

buffer arbitration logic which arbitrates among routing tags entering the same input port that contend for output from the same output port and randomly selects a particular routing tag therefrom;

port arbitration logic coupled to said buffer arbitration logic which arbitrates among all routing tags transmitted by said buffer arbitration logic which contend for the same output port and selects one routing tag therefrom for output from that particular output port;

address generation logic of said arbitration logic which generates read addresses that control the output of data packets through the appropriate output port for data packets corresponding to those routing tag signals selected by said port arbitration logic; and

status update logic which generates signals which reset the full/empty status of the registers corresponding to the data packets applied to any of the output ports;

According to another aspect of the present invention there is provided a method of processing applied data packets containing routing tag signals indicative of the output port destinations to which said data packets are to be applied in order to route said data packets to the appropriate output ports, said method comprising the steps of:

storing said data packets in predetermined memory locations;

sorting said data packets based upon the output port destination thereof contained in said routing tag signals;

arbitrating among those data packets which have entered the same input port and which are destined for the same output port;

arbitrating among those data packets that have entered different input ports and which are destined for the same output port; and

routing the data packets selected during the arbitration process from said memory locations to said output ports identified in said routing tag signals.

The various objects and features of the present invention may be more readily understood with reference to the following detailed description taken in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Fig. 1 illustrates a generalized embodiment of an  $N \times M$  multiport memory packet switching

node in accordance with the principles of the present invention;

Fig. 2 illustrates a block diagram of a four input, four output implementation of the packet switching node of Fig. 1;

Figs. 3a and 3b show a detailed block diagram of control logic for use in the packet switching node of Fig. 2;

Figs. 4a and 4b show a design of status, buffer arbitration and priority encoder logic for use in the control logic of Fig. 3;

Figs. 5a and 5b show a design of port arbitration and latch logic for use in the control logic of Fig. 3;

Figs. 6a and 6b show a design of status update (clear) logic used in the control logic of Fig. 3;

Fig. 7 shows a design of a demultiplexer for use in the buffer arbitration logic of Fig. 4;

Fig. 8 shows a design of a three-to-one multiplexer for use in the buffer arbitration logic of Fig. 4;

Fig. 9 shows a design of a modulo 3 counter for use in the buffer arbitration logic of Fig. 4;

Fig. 10 shows a design of a three input arbitration circuit for use in the buffer arbitration circuit of Fig. 4;

Fig. 11 shows a design of a four-to-twelve demultiplexer circuit for use in the status update logic of Fig. 6;

Fig. 12 shows a design of a multiport memory for use in the switch node of Fig. 2;

Figs. 13a and 13b show a design of a  $12 \times 4 \times 3$  bit array for use in the memory of Fig. 12;

Figs. 14a and 14b show a design of a  $4 \times 3$  memory array for use in the bit array of Fig. 13;

Fig. 15 shows a design of write address select decode logic for use in the memory of Fig. 12;

Fig. 16 shows the design of a demultiplexer for use in the logic circuit of Fig. 15;

Fig. 17 shows a design of read address row and column select decode logic for use in the memory of Fig. 12;

Fig. 18 shows the design of a decoder for use in the logic circuit of Fig. 17; and

Fig. 19 shows a timing diagram for the switch node of Fig. 2.

Referring to Fig. 1, there is shown a generalized embodiment of an  $N \times M$  multiport memory packet switching node 20 in accordance with the principles of the present invention. The node 20 comprises a plurality of input ports 21 and a plurality of output ports 22. A multiport memory 23 is coupled between the input and output ports 21, 22. The memory 23 has a predetermined number of memory locations available for storage of data packets applied to each of the input ports 21. Control logic 24 is coupled to the input and output ports 21, 22 and the multiport memory 23 which controls the storage of data packets in the memory 23. The control logic 24 also controls the routing of the data packets to the output ports 22 in accordance with routing tag signals contained in the data packets.

The transmission protocol for receiving data packets by way of the input ports 21 is executed by a plurality of receiving handshake signal lines 25 coupled to the control logic 24. Similarly the transmission protocol for sending data packets from the output ports 22 is executed by a plurality of sending handshake signal lines 26 also coupled to the control logic 24.

The control logic 24 comprises tag status logic register 30 coupled to the input ports 21, the receive handshake signal lines 25, and the memory 23, which stores both the routing tag signals and status signals indicative of the full/empty status of each of the memory locations in the memory 23. The control logic 24 also includes priority encoder logic 31 coupled to the input ports 21, the memory 23 and the tag status register logic 30. The priority encoder logic 31 is adapted to monitor the full/empty status of the memory locations, write the routing tag signals into registers of the status register 30 and write data packets corresponding to the routing tag signals into selected memory locations.

The control logic 24 further comprises arbitration logic 32 coupled to the tag status register logic 30 and latch logic 36. The arbitration logic 32 is designed to examine routing tags stored in the status register logic 30 and resolve conflicting requests based upon a predetermined priority arbitration scheme. In the arbitration process, read addresses are generated and stored in the latch logic 36 that determine the output of data packets through the appropriate output port 22 such that they correspond to the routing tag signals selected by the arbitration logic 32.

The arbitration logic 32 includes buffer arbitration logic 33 which is designed to arbitrate among routing tags entering the same input port 21 that contend for output from the same output port 22. The arbitration logic 32 randomly selects a particular routing tag therefrom. In addition, port arbitration logic 34 is coupled to the buffer arbitration logic 33 and to the latch logic 36 which is adapted to arbitrate among all routing tags transmitted by the buffer arbitration logic 33 which contend for the same output port 22 and selects one routing tag and its associated packet therefrom for output from that particular output port 22.

Status update logic 35 is coupled to the latch logic 36 and the sending handshake signal lines 26. The status update logic 35 is designed to generate signals which reset the full/empty status of the registers corresponding to the data packets applied to and transmitted from any of the output ports 22.

In operation, the present invention may be employed as an  $N \times M$  switch node that accepts data packets at any of  $N$  input ports 21 and routes each to any of  $M$  output ports 22. The output port 22 selected is determined by the routing tag signal in the data packet. The control logic 24 is designed so that the data packets are effectively sorted according to their desired output port destination. The arbitration logic 32 randomly, in

a statistical sense, chooses among data packets that are directed to the same output port 22. The algorithm implemented by the arbitration logic 32 is designed so that data packets will not wait indefinitely to be routed from the switch node 20.

The control logic 24 performs handshaking with other switch nodes or processors, or the like, to which the node 20 is connected. The control logic 24 generates addresses into the memory 23 and arbitrates among data packets. All incoming requests to the switch node 20 are accompanied by the routing tag signals. These routing tag signals are a predetermined number of bits in the data packet. These bits, along with a full/empty status bit, are stored in the tag status register logic 30. The status register logic 30 contains  $N \times P$  cells, where  $N$  is the number of input ports 21 and  $P$  is the number of storage locations allowed per input port 21. Each cell corresponds to a memory location and stores both the status and tag bits.

The priority encoder logic 31 selects the first available cell allotted for a given input port 21 when a request is received. As long as the status bits are not all full, a "not full," or grant signal is sent to the sending node or processor which is interpreted to mean proceed. The requesting packet is then written into its assigned location in the memory 23, the tag bits are written into the tag status register logic 30 and the corresponding full/empty bit is set to "full".

Arbitration among packets takes place in two steps. First, all the tag bits in full cells are decoded and grouped according to the desired output port 22. Then the buffer arbitration logic 33 randomly chooses among packets that entered the same input port 21 that also are destined for the same output port 22. In the second step, the port arbitration logic 34 randomly chooses between any packets presented by the buffer arbitration logic 33 that are destined for the same output port 22. As long as data packets that entered the same input port 21 are destined for different output ports 22, all are allowed to participate in the second step of arbitration. Thus, none are blocked due to their position in the memory 23. A packet can only be blocked due to competition with another packet for the same output port 22.

The results of the arbitration process are latched in the latch logic 36 and requests are generated for those output ports 22 desired by the data packets. Appropriate read addresses are generated for each output port 22 that is to be used. For those requests that are granted, the appropriate full/empty status bits are reset to "empty." If any request is not granted, the packet remains in the memory 23 and repeats the arbitration process during the next cycle.

Fig. 2 illustrates a block diagram of the currently preferred embodiment of Fig. 1, in which  $N = M = 4$ , and which comprises a four by four multiport memory switch node. Referring to Fig. 3, a top-level detailed block diagram of the control logic 24 which may be employed in the packet switching node of Fig. 2 is shown. Fig. 3 shows the

interconnection of a plurality of logic blocks which incorporate the status update logic 35, tag status register logic 30, priority encoder logic 31, buffer arbitration logic 33, port arbitration logic 34 and latch logic 36 which comprise the various logic portions of the control logic 24 utilized in the present invention.

The logic circuit shown in Fig. 3 is partitioned so that those portions of the tag status register logic 30, priority encoder logic 31 and buffer arbitration logic 33 associated with one input port 25 are grouped together to form four identical logic modules which comprise four input port logic sections 70a-d. Similarly, those portions of the port arbitration logic 34 and latch logic 36 associated with each output port 26 are grouped together to form four identical logic modules comprising four output port logic sections 71a-d.

Other required signal lines are shown including clock and reset signal lines. Detailed schematics of the various logic portions shown in Fig. 3 are illustrated in Figs. 4 through 11 hereinbelow. There is a one to one correspondance between the signal line identifiers in Fig. 3 and the signal line identifiers used in the Figs. 4 through 11. The balance of the Figures will not be discussed in explicit detail except to the extent necessary to explain their operation and function in the overall switch node 20.

Fig. 4 shows the detailed logic design of one input port section 70. The status register logic 30 comprises nine flip-flops 72a-i which are coupled to tag decoders or demultiplexers 73a-c and the priority encoder logic 31 which employs a plurality of nand gates 78a-c. Decoded tag signals are coupled from the demultiplexers 73 to the buffer arbitration logic 33 which comprises buffer arbitration logic circuits 74a-d. Intermediate status register logic comprising registers 75a-h is used to ensure correct timing of the signals. A modulo 3 counter 76 is employed to determine the current priority used by the buffer arbitration logic circuits 74a-d and is enabled by counter enable logic 77. The priority encoder logic 31 is coupled through the intermediate status register 75f-h to tag status register logic clock enable logic 79. Full/empty flip-flops 72g-i utilized in the tag status register logic 30 are cleared by status reset logic 80.

The logic circuit of Fig. 4 operates as follows. The input port logic section 70a shown in Fig. 4 stores a 2 bit tag and its associated request signal into one of the three locations in the tag status register logic 30 comprising flip-flops 72a-i. The particular storage location is determined by the priority encoder logic 31 comprising nand gates 78 and the clock enable logic 79. The priority encoder logic 31 produces a 2 bit address indicating the first available location and a signal to indicate if there are none. These three signals are also provided to the multiport memory 23 and form the write address for the corresponding input port 21.

The tags that have been stored in the tag status register logic 30 are each decoded by the demulti-

plexers 73. At most, one signal output from each demultiplexer 73 will be active and will correspond to the desired output port 22. The signal wires from the demultiplexers 73 are physically grouped by corresponding output port 22 and are input to four buffer arbitration logic circuits 74, comprising the buffer arbitration logic 33, with each circuit being associated with a different output port 22. The priority used by each buffer arbitration logic circuit 74 is provided by a modulo 3 counter 76. Each buffer arbitration logic circuit 74 outputs a two bit address of the selected tag and corresponding packet, along with a signal that indicates if any tag desires that output port 22.

The modulo 3 counter 76 is incremented, and hence the priority changes, according to a signal derived from counter enable logic 77. The modulo 3 counter 76 is enabled only if the location with highest priority is empty, as indicated by the tag status register logic full/empty flip-flops 72g-i, or if a clear signal is produced by the status update logic 35 indicating that the packet in the location with highest priority has been transmitted. The clear signals from the status update logic 35 are combined with a master reset signal and a clock signal in status reset logic 80.

Referring to Fig. 5, there is shown the detailed logic design of an output port logic section 71. The four to one address multiplexers 93a-b and gates 90 which comprise the port arbitration logic 34 are coupled to the latch logic 36 comprising flip-flops 94a-h. The output of the port arbitration logic 34 is also directly coupled to the address multiplexers 93a-b. A modulo 4 counter 91 is coupled to the port arbitration logic 34 and determines the current priority. Counter enable logic 92 controls whether or not the modulo 4 counter 91 is incremented.

In operation, each output port logic section 71 receives one address and request signal from each input port logic section 70, as shown in Fig. 3. The request signals are processed by the port arbitration logic 34 according to the priority indicated by the modulo 4 counter 91. The 2 bit address of the selected input port 21 is output to the latch logic 36 along with a request signal indicating if there is any request. This 2 bit input port address also allows the 4 to 1 address multiplexers 93 to select the corresponding 2 bit status register address from the selected input port 21 which is also output to the latch logic 36. The 4 bits of address form a read address which is coupled to the corresponding output port 22 in the multiport memory 23 to access the packet associated with the selected tag bits.

The modulo four counter 91 is incremented, changing the priority used by the port arbitration logic 34, according to signals from the counter enable logic 92. The counter enable logic 92 produces an increment signal whenever the input port 21 with highest priority has no request for the associated output port 22 or it has a request which is honored and the packet is transmitted. As long as the FULL signal is TRUE, indicating

that an output device or switch node attached to the associated output port 22 is "not full," the selected request will be sent. Otherwise the request must wait until the output device is no longer full.

Referring to Fig. 6, there is shown a detailed logic design of the status update logic 35. It is comprised of four identical read address decoders 97a-d which are coupled to clear signal merge logic 98. In operation, each read address decoder 97 enables a single signal, which corresponds to one of 12 status flip-flops in the tag status register logic 30 if there is a request associated with the read address at the input and the "not full" signal is true. The clear signals from all the read address decoders 97 are merged into a single set of signals which are connected to the status reset logic 80 and the counter enable logic 77 in Fig. 4. There are at most four clear signals active at one time.

Referring to Fig. 7, a design of the demultiplexer 73 used in the input port logic 70 of Fig. 4 is shown. Fig. 8 shows a design of the three to one multiplexer 77 used in the input port logic 70 of Fig. 4. Fig. 9 shows a design of the modulo 3 counter 76 used in the input port logic 70 of Fig. 4. Fig. 10 shows a design of the 3 input arbitration logic 74 used in the input port logic 70 of Fig. 4. Fig. 11 shows a design of a four-to-twelve demultiplexer circuit 97 for use in the status update logic 35 of Fig. 6.

Referring to Fig. 12, a detailed logic design of the multiport memory 23 is illustrated. The memory 23 is comprised of write address select decode logic 40 and read address row and column select decode logic 41 which are coupled to a  $12 \times 4 \times 3$  bit array 42. The multiport memory 23 is designed to accept up to four write addresses and 12 bits which comprise a data packet associated with each write address at an input port 21 and up to four read addresses and produce the four corresponding 12 bit data packets at an output port 22 during each clock cycle. Each write address is decoded in the write address logic 40 so that one of three memory cell select lines is enabled. Each read address is decoded in the read address logic 41 so that one of four row select and one of three column select lines is enabled. The memory cell at the intersection of the row and column select lines presents its contents at the corresponding output port 22. Detailed schematics of the various logic blocks shown in Fig. 12 are shown in Figs. 13 through 18 hereinbelow.

Fig. 13 shows the organization of the  $12 \times 4 \times 3$  bit array 42. It is comprised of twelve identical copies of a  $4 \times 3$  bit memory array 50. The write, row, and column select lines are all connected to each  $4 \times 3$  bit memory array 50. Fig. 14 shows the detailed logic design of the  $4 \times 3$  memory array 50. Flip-flops 60a-l are used to store one bit of information each. Each input port 21 has an exclusive set of three memory locations into which it can write, which corresponds to three flip-flops that are in the same row, represented by

60a-c, 60d-f, etc. There are four identical copies of memory read flip-flop select logic 61a-d, one for each output port 22, which outputs the value of the bit in the flip-flop 60 corresponding to the intersection of the enabled row and column select lines.

Fig. 15 shows the detailed design of the write address select decode logic 40 which is comprised of four identical 2 to 3 demultiplexers 51. The detailed logic design of the 2 to 3 demultiplexer 51 is shown in Fig. 16. Fig. 17 shows the detailed design of the read address row and column select decode logic 41 which is comprised of four identical row and column decoders 52. The detailed logic design of a row and column decoder 52 is shown in Fig. 18.

Fig. 19 shows a timing diagram for the multiport memory packet switching node 20 of Fig. 2.

The system clock, which is externally generated, is comprised of two non-overlapping phases,  $\phi 1$  and  $\phi 2$ .

The rising edge of  $\phi 1$  initiates the generation of all read and write addresses. Any data packet and its corresponding tag and request signal must be stable prior to the rising edge of  $\phi 2$ . Upon the rising edge of  $\phi 2$ , new data packets are written into the multiport memory 23 and the tag status register logic 30 is updated. Conflicts among tags are resolved at this time. The buffer and port priorities are also changed if enabled at this time.

Thus, there has been described a new and improved packet switching node which may be employed as a switch node in a multi-processor or parallel computer application. The packet switching node employs a multiport memory to store and sort data packets based upon the output port destinations thereof. The packet switching node eliminates the problem of contention between data packets arriving at an input port of the node whose destinations are different output ports. The present invention also provides for a packet switching node which has improved performance and higher throughput.

It is to be understood that the abovedescribed embodiments are merely illustrative of some of the many specific embodiments which represent applications of the principles of the present invention. Clearly, numerous and varied other arrangements may be readily devised by those skilled in the art without departing from the spirit and scope of the invention. For example, the present invention is not limited to uses related only to interconnection devices for computers. It also has application in the telecommunications field and may be adapted for use with telecommunications equipment, and the like. In particular, interconnection networks employing the present invention may be used to connect telephone systems that communicate both data and voice information by way of data packets.

#### Claims

1. A packet switching node (20) which processes applied data packets containing routing

that an output device or switch node attached to the associated output port 22 is "not full," the selected request will be sent. Otherwise the request must wait until the output device is no longer full.

Referring to Fig. 6, there is shown a detailed logic design of the status update logic 35. It is comprised of four identical read address decoders 97a-d which are coupled to clear signal merge logic 98. In operation, each read address decoder 97 enables a single signal, which corresponds to one of 12 status flip-flops in the tag status register logic 30 if there is a request associated with the read address at the input and the "not full" signal is true. The clear signals from all the read address decoders 97 are merged into a single set of signals which are connected to the status reset logic 80 and the counter enable logic 77 in Fig. 4. There are at most four clear signals active at one time.

Referring to Fig. 7, a design of the demultiplexer 73 used in the input port logic 70 of Fig. 4 is shown. Fig. 8 shows a design of the three to one multiplexer 77 used in the input port logic 70 of Fig. 4. Fig. 9 shows a design of the modulo 3 counter 76 used in the input port logic 70 of Fig. 4. Fig. 10 shows a design of the 3 input arbitration logic 74 used in the input port logic 70 of Fig. 4. Fig. 11 shows a design of a four-to-twelve demultiplexer circuit 97 for use in the status update logic 35 of Fig. 6.

Referring to Fig. 12, a detailed logic design of the multiport memory 23 is illustrated. The memory 23 is comprised of write address select decode logic 40 and read address row and column select decode logic 41 which are coupled to a  $12 \times 4 \times 3$  bit array 42. The multiport memory 23 is designed to accept up to four write addresses and 12 bits which comprise a data packet associated with each write address at an input port 21 and up to four read addresses and produce the four corresponding 12 bit data packets at an output port 22 during each clock cycle. Each write address is decoded in the write address logic 40 so that one of three memory cell select lines is enabled. Each read address is decoded in the read address logic 41 so that one of four row select and one of three column select lines is enabled. The memory cell at the intersection of the row and column select lines presents its contents at the corresponding output port 22. Detailed schematics of the various logic blocks shown in Fig. 12 are shown in Figs. 13 through 18 hereinbelow.

Fig. 13 shows the organization of the  $12 \times 4 \times 3$  bit array 42. It is comprised of twelve identical copies of a  $4 \times 3$  bit memory array 50. The write, row, and column select lines are all connected to each  $4 \times 3$  bit memory array 50. Fig. 14 shows the detailed logic design of the  $4 \times 3$  memory array 50. Flip-flops 60a-l are used to store one bit of information each. Each input port 21 has an exclusive set of three memory locations into which it can write, which corresponds to three flip-flops that are in the same row, represented by

60a-c, 60d-f, etc. There are four identical copies of memory read flip-flop select logic 61a-d, one for each output port 22, which outputs the value of the bit in the flip-flop 60 corresponding to the intersection of the enabled row and column select lines.

Fig. 15 shows the detailed design of the write address select decode logic 40 which is comprised of four identical 2 to 3 demultiplexers 51. The detailed logic design of the 2 to 3 demultiplexer 51 is shown in Fig. 16. Fig. 17 shows the detailed design of the read address row and column select decode logic 41 which is comprised of four identical row and column decoders 52. The detailed logic design of a row and column decoder 52 is shown in Fig. 18.

Fig. 19 shows a timing diagram for the multiport memory packet switching node 20 of Fig. 2.

The system clock, which is externally generated, is comprised of two non-overlapping phases,  $\phi 1$  and  $\phi 2$ .

The rising edge of  $\phi 1$  initiates the generation of all read and write addresses. Any data packet and its corresponding tag and request signal must be stable prior to the rising edge of  $\phi 2$ . Upon the rising edge of  $\phi 2$ , new data packets are written into the multiport memory 23 and the tag status register logic 30 is updated. Conflicts among tags are resolved at this time. The buffer and port priorities are also changed if enabled at this time.

Thus, there has been described a new and improved packet switching node which may be employed as a switch node in a multi-processor or parallel computer application. The packet switching node employs a multiport memory to store and sort data packets based upon the output port destinations thereof. The packet switching node eliminates the problem of contention between data packets arriving at an input port of the node whose destinations are different output ports. The present invention also provides for a packet switching node which has improved performance and higher throughput.

It is to be understood that the abovedescribed embodiments are merely illustrative of some of the many specific embodiments which represent applications of the principles of the present invention. Clearly, numerous and varied other arrangements may be readily devised by those skilled in the art without departing from the spirit and scope of the invention. For example, the present invention is not limited to uses related only to interconnection devices for computers. It also has application in the telecommunications field and may be adapted for use with telecommunications equipment, and the like. In particular, interconnection networks employing the present invention may be used to connect telephone systems that communicate both data and voice information by way of data packets.

#### Claims

1. A packet switching node (20) which processes applied data packets containing routing



tag signals indicative of the output port destination to which said data packets are to be applied, said packet switching node (20) comprising a plurality of input ports (21); a plurality of output ports (22); a multiport memory (23) coupled between said input (21) and output (22) ports, said memory (23) having a predetermined number of memory locations available for storage of said data packets applied to each of said input ports (21); and control logic (24) coupled to said input (21) and output (22) ports and said multiport memory (23) for controlling the storage of data packets in said memory (23) and the routing thereof to said output ports (22) in accordance with said routing tag signals, wherein the improvement is in said control logic (24) characterized by:

a status register (30) coupled to said input ports (21) and said memory (23) for storing said routing tag signals and status signals indicative of the full/empty status of each of the memory locations in said memory (23);

priority encoder logic (31) coupled to said input ports (21), said memory (23) and said status register (30) for monitoring the full/empty status of said memory locations, writing said routing tag signals into registers of said status register and writing data packets corresponding to said routing tag signals into selected memory locations;

an arbitration logic (32) coupled to said status register (30), said memory (23) and said output ports (22) for reading said data packets from said memory locations and applying them to the output ports (22) identified in the corresponding routing tag signals based upon a predetermined priority arbitration scheme;

buffer arbitration logic (33) which arbitrates among routing tags entering the same input port (21) that contend for output from the same output port (22) and randomly selects a particular routing tag therefrom;

port arbitration logic (34) coupled to said buffer arbitration logic (33) which arbitrates among all routing tags transmitted by said buffer arbitration logic (33) which contend for the same output port (22) and selects one routing tag therefrom for output from that particular output port (22);

address generation logic (36) of said arbitration logic (32) which generates read addresses that control the output of data packets through the appropriate output port (22) for data packets corresponding to those routing tag signals selected by said port arbitration logic (34); and

status update logic (35) which generates signals which reset the full/empty status of the registers corresponding to the data packets applied to any of the output ports (22).

2. A method of processing applied data packets containing routing tag signals indicative of the output port destinations to which said data packets are to be applied in order to route said data packets to the appropriate output ports, said method comprising the steps of:

storing said data packets in predetermined memory locations;

sorting said data packets based upon the output

port destination thereof contained in said routing tag signals;

arbitrating among those data packets which have entered the same input port and which are destined for the same output port;

arbitrating among those data packets that have entered different input ports and which are destined for the same output port; and

routing the data packets selected during the arbitration process from said memory locations to said output ports identified in said routing tag signals;

## Patentansprüche

1. Paketschaltungsknoten (20), der angelegte Datenpakete verarbeitet, die Leitmarkensignale enthalten, welche die Ausgangsanschlußbestimmung anzeigen, an die die Datenpakete angelegt werden sollen, wobei der Paketschaltungsknoten (20) eine Vielzahl von Eingangsanschlüssen (21); eine Vielzahl von Ausgangsanschlüssen (22); einen zwischen den Eingangs- (21) und Ausgangsanschlüssen (22) verbundenen Mehrfachanschlußspeicher (23), wobei der Speicher (23) eine vorbestimmte Anzahl von Speicherplätzen, die zum Speichern der an jeden der Eingangsanschlüsse (21) angelegten Datenpakete verfügbar sind, aufweist; und eine mit den Eingangs- (21) und Ausgangsanschlüssen (22) und dem Mehrfachspeicher (23) verbundene Steuerlogik (24) zum Steuern der Speicherung der Datenpakete in dem Speicher (23) und deren Leiten an die Ausgangsanschlüsse (22) entsprechend der Leitmarkensignale aufweist, wobei die Verbesserung in der Steuerlogik (24) besteht und gekennzeichnet ist durch:

ein mit den Eingangsanschlüssen (21) und dem Speicher (23) verbundenes Statusregister (30) zum Speichern der Leitmarkensignale und von Statussignalen, die den vollen/leeren Status von jedem der Speicherplätze in dem Speicher (23) anzeigt;

eine mit den Eingangsanschlüssen (21), dem Speicher (23) und dem Statusregister (30) verbundene Prioritätsdecoderlogik (31) zum Überwachen des vollen/leeren Status der Speicherplätze, zum Schreiben der Leitmarkensignale in Register der Statusregister und zum Schreiben von Datenpaketen entsprechend der Leitmarkensignale in ausgewählte Speicherplätze;

eine mit dem Statusregister (30), dem Speicher (23) und den Ausgangsanschlüssen (22) verbundene Entscheidungslogik (32) zum Lesen der Datenpakete von den Speicherplätzen und Anlegen dieser an die in den entsprechenden Leitmarkensignalen aufgrund eines vorbestimmten Prioritätsentscheidungsschemas identifizierten Ausgangsanschlüsse (22);

eine Pufferentscheidungslogik (33), die unter den denselben Eingangsanschluß (21) betretenden Leitmarken, die für die Ausgabe von demselben Ausgangsanschluß (22) konkurrieren, entscheidet und wahlfrei davon eine bestimmte Leitmarke auswählt;

eine mit der Pufferentscheidungslogik (33) ver-



bundene Anschlußentscheidungslogik (34), die unter sämtlichen von der Pufferentscheidungslogik (33) übertragenen Leitmarken, die für denselben Ausgangsanschluß (22) konkurrieren, entscheidet und eine Leitmarke davon für die Ausgabe von diesem bestimmten Ausgangsanschluß (22) auswählt;

eine Adreßerzeugungslogik (36) der Entscheidungslogik (32), die Leseadressen erzeugt, die die Ausgabe der Datenpakete über den geeigneten Ausgangsanschluß (22) für Datenpakete entsprechend denjenigen Leitmarkensignalen, die von der Anschlußentscheidungslogik (34) ausgewählt sind, steuert; und

eine Statuserneuerungslogik (35), die Signale erzeugt, die den vollen/leeren Status der Register entsprechend den anirgendeinen der Ausgangsanschlüsse (22) angelegten Datenpakete zurücksetzen.

2. Verfahren zum Verarbeiten von angelegten Datenpaketen, die Leitmarkensignale enthalten, die die Ausgangsanschlußbestimmungen anzeigen, an die die Datenpakete angelegt werden sollen, um die Datenpakete an die geeigneten Ausgangsanschlüsse zu leiten, wobei das Verfahren die Schritte aufweist:

Speichern der Datenpakete in vorbestimmte Speicherplätze;

Sortieren der Datenpakete aufgrund der Ausgangsanschlußbestimmung, die in den Leitmarkensignalen enthalten sind;

Entscheiden unter denjenigen Datenpaketen, die denselben Eingangsanschluß betreten haben und für denselben Ausgangsanschluß bestimmt sind;

Entscheiden unter denjenigen Datenpaketen, die unterschiedliche Eingangsanschlüsse betreten haben und für denselben Ausgangsanschluß bestimmt sind; und Leiten der während des Entscheidungsprozesses ausgewählten Datenpakete von den Speicherplätzen auf die in den Leitmarkensignalen identifizierten Ausgangsanschlüsse;

## Revendications

1. Noeud de commutation de paquets (20) qui traite des paquets de données appliqués contenant des signaux formant étiquettes d'acheminement indicateurs de l'accès de sortie de destination auquel les paquets de données doivent être appliqués, ce noeud de commutation de paquets (20) comprenant plusieurs accès d'entrée (21), plusieurs accès de sortie (22), et une mémoire à accès multiples (23) couplée entre les accès d'entrée (21) et les accès de sortie (22), la mémoire (23) ayant un nombre prédéterminé d'emplacements de mémoire disponibles pour le stockage des paquets de données appliqués à chacun des accès d'entrée (21), et une logique de commande (24) reliée aux accès d'entrée (21), aux accès de sortie (22) et à la mémoire à accès multiples (23), et destinée à commander le stockage de paquets de données dans la mémoire (23) et l'acheminement de ces paquets de données jusqu'aux accès de sortie (22) conformément aux signaux formant

étiquettes d'acheminement, dans lequel le perfectionnement consiste en ce que la logique de commande (24) est caractérisée par:

un registre d'état (30) relié aux accès d'entrée (21) et à la mémoire (23) pour stocker les signaux formant étiquettes d'acheminement et des signaux d'état indicateurs de l'état plein/vide de chacun des emplacements de mémoire dans la mémoire (23);

une logique de codage des priorités (31) reliée aux accès d'entrée (21), à la mémoire (23) et au registre d'état (30), destinée à surveiller l'état plein/vide des emplacements de mémoire, à écrire les signaux formant étiquettes d'acheminement dans des registres du registre d'état (30), et à écrire des paquets de données correspondant aux signaux formant étiquettes d'acheminement dans des emplacements de mémoire sélectionnés;

une logique d'arbitrage (32) reliée au registre d'état (30), à la mémoire (23) et aux accès de sortie (22), destinée à lire les paquets de données dans lesdits emplacements de mémoire et à les appliquer aux accès de sortie (22) identifiés dans les signaux formant étiquettes d'acheminement correspondants sur la base d'un plan prédéterminé d'arbitrage entre les priorités;

une logique d'arbitrage tampon (33) qui arbitre entre les étiquettes d'acheminement entrant par le même accès d'entrée (21) et qui sont en conflit pour sortir par le même accès de sortie (22), et sélectionne au hasard une étiquette d'acheminement particulière parmi celles-ci;

une logique d'arbitrage entre les accès (34) reliée à la logique d'arbitrage tampon (33), qui arbitre entre toutes les étiquettes d'acheminement transmises par cette logique d'arbitrage tampon (33) qui sont en conflit pour le même accès de sortie (22), et sélectionne une étiquette d'acheminement parmi celles-ci pour sortie par cet accès de sortie (22);

une logique d'élaboration d'adresses (36) de la logique d'arbitrage (32) qui élabore des adresses de lecture commandant la sortie de paquets de données par l'accès de sortie (22) approprié pour des paquets de données correspondant aux signaux formant étiquettes d'acheminement sélectionnés par la logique d'arbitrage entre les accès (34); et

une logique de mise à jour d'état (35) qui génère des signaux remettant à l'état initial l'état plein/vide des registres correspondant aux paquets de données appliqués à n'importe lequel des accès de sortie (22).

2. Procédé de traitement de paquets de données appliqués contenant des signaux formant étiquettes d'acheminement indicateurs des accès de sortie de destination auxquels les paquets de données doivent être appliqués afin d'acheminer ces paquets de données jusqu'aux accès de sortie appropriés, ce procédé comprenant les étapes consistant à:

stocker les paquets de données dans des emplacements de mémoire prédéterminés;

trier les paquets de données d'après l'accès de



sortie de destination de ceux-ci contenu dans les signaux formant étiquettes d'acheminement;

arbitrer entre les paquets de données qui sont entrés par le même accès et qui sont destinés au même accès de sortie;

arbitrer entre les paquets de données qui

sont entrés par différents accès d'entrée et qui sont destinés au même accès de sortie; et

acheminer les paquets de données sélectionnés pendant le processus d'arbitrage, desdits emplacements de mémoire jusqu'aux accès de sortie identifiés dans les signaux formant étiquettes d'acheminement.

10

15

20

25

30

35

40

45

50

55

60

65

9



\* C L N S \*