



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

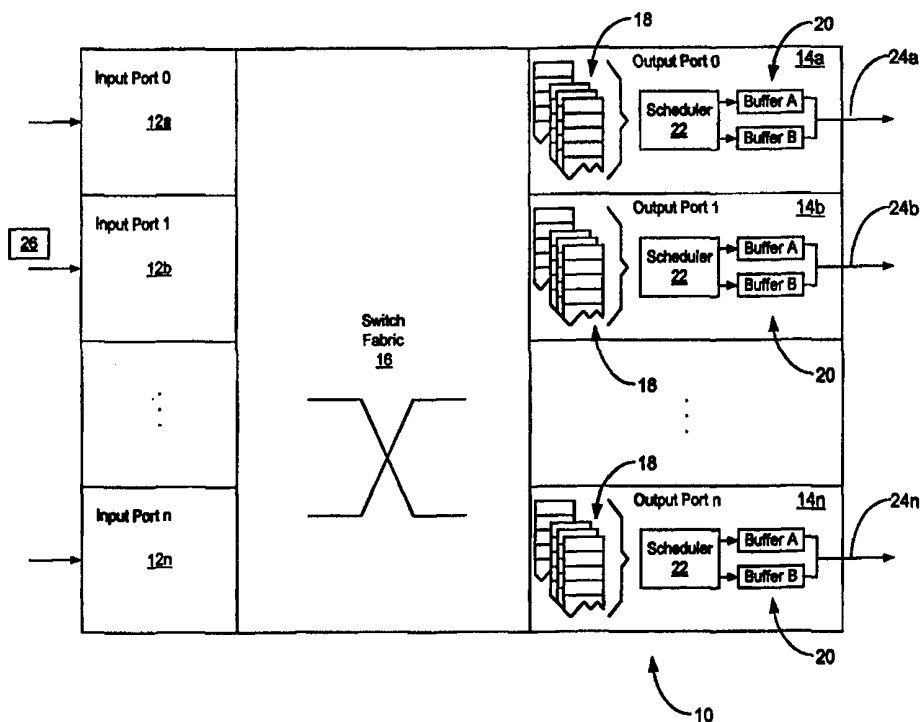
(51) International Patent Classification ⁶ : H04L 12/28, 12/56	A1	(11) International Publication Number: WO 99/66674 (43) International Publication Date: 23 December 1999 (23.12.99)
(21) International Application Number: PCT/US99/13681 (22) International Filing Date: 16 June 1999 (16.06.99) (30) Priority Data: 09/100,584 19 June 1998 (19.06.98) US (71) Applicant: ASCEND COMMUNICATIONS, INC. [US/US]; 1 Robbins Road, Westford, MA 01886 (US). (72) Inventors: HEBB, Andrew, T.; 62 Lakeside Avenue, Hudson, MA 01749 (US). GOSS, Gregory, S.; Unit 2, 1267 Pawtucket Boulevard, Lowell, MA 01854 (US). (74) Agents: SCHURGIN, Stanley, M. et al.; Weingarten, Schurgin, Gagnebin & Hayes LLP, Ten Post Office Square, Boston, MA 02109 (US).		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>With amended claims.</i>

(54) Title: LOGICAL MULTICASTING METHOD AND APPARATUS

(57) Abstract

A method and apparatus for performing logical multicasting within a network switch (10) are disclosed so as to permit a downstream network device to perform cell forwarding in a manner which resembles spatial multicasting. Each input cell (26) received at an input port (12a—12n) is checked to determine if it is a logical multicast cell. A logical multicast cell has a connection identifier associated with the cell within the switch (10). The connection identifier is employed to initiate a series of lookup operations to generate unique VPI/VCI addresses which are transmitted out of an output port (14a—14n) of the switch as copies or leaves of the multicast cell.

Successive look-ups of outgoing VPI/VCI addresses for a logical multicast cell are performed using a chain bit to indicate when all leaves of the cell have been transmitted. Leaves of a logical multicast cell are transmitted to the downstream network device from one output cell buffer of a plurality of output cell buffers (20) associated with an output port (14) of the switch (10) so as to minimize the likelihood of blocking and so as to minimize cell delay variation. In one embodiment, the output cell buffers (20) having cells to transmit are afforded transmit opportunities in a round robin sequence by a scheduler (22) associated with the output port (14).



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.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

LOGICAL MULTICASTING METHOD AND APPARATUS

CROSS REFERENCE TO RELATED APPLICATIONS

NOT APPLICABLE

**STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR
DEVELOPMENT**

NOT APPLICABLE

BACKGROUND OF THE INVENTION

The present invention relates to computer networking and more particularly to a method and apparatus for performing logical multicasting in a computer network.

Telecommunications networks deliver data by routing data units, such as Asynchronous Transfer Mode ("ATM") cells, from source to destination through switches. Each switch includes input/output ("I/O") ports through which the ATM cells are received and transmitted, respectively. The switch determines the appropriate output port to be employed for transmission of the cell based upon an address in the received cell header, such as the VPI/VCI address within the cell header of a received ATM cell.

Depending on the destination indicated in the cell header, it may be desirable to transmit the cell from multiple ports on the switch. Cells which enter a switch through one port and exit through one port are referred to as unicast cells. Cells which enter the switch through a single port and exit through multiple ports are known as multicast cells. While most existing switches are configured to handle unicast cells, not all switches are capable of multicast operation.

Two types of multicasting are known in the art; namely spatial multicasting and logical multicasting. Spatial multicasting involves the receipt of a cell at one input port of a network switch and the forwarding of copies of the cell out of plural output ports of the network switch.

Each copy of the cell which is forwarded out of an output

port in a spatial multicasting operation is provided with a unique VPI/VCI address. Logical multicasting involves the receipt of a cell at an input port of a network switch and the forwarding of plural copies of the cell out of a single output port of the network switch. Each of the leaves of the logical multicast cell is transmitted from the output port with a unique VPI/VCI address.

While many switches are capable of spatial multicast operation (transmission of single copies of a cell through multiple output ports), not all multicast switches are capable of logical multicast operation (transmission of multiple copies of a cell through a single output port). Logical multicast switches are known, but are more complex and costly than typical spatial-only multicast switches.

One technique for performing logical multicast of cells within a network switch is disclosed in U.S. Patent Application No. 08/710,673 filed September 18, 1996 and assigned to the predecessor in interest of the present application. The technique therein disclosed provides a first copy of the cell to the desired output port. A second copy of the cell is provided to a dedicated loop back output port which causes the cell to be looped back to an input port dedicated to loop back. The cell provided to the loop-back input port is subsequently directed to the desired output port such that the first and second copies of the cell are serially transmitted from the single desired output port. The cell may be repeatedly provided to the loop-back port to permit the cell to be transmitted in a serial fashion from the desired output as many times as desired. This technique disadvantageously requires the dedication of one input port and one output port to the logical multicasting function.

Accordingly, it would be desirable to be able to handle logical multicast operation without specifically allocating input or output ports for the support of multicast operation. Additionally, it would be useful to

provide multicast type capability in switches which do not support multicast operation by design.

BRIEF SUMMARY OF THE INVENTION

5 A method and apparatus for performing logical multicasting in a first network device are disclosed so as to permit a downstream network device to forward cells in a manner which resembles spatial multicasting. A multicast cell which is received at a network switch is converted
10 into a plurality of cells, each having a unique VPI/VCI address. The cells are then transmitted out of a single physical port to a downstream network node. The downstream node can then switch the cells out of different physical output ports based upon the received incoming VPI/VCI
15 addresses transmitted by the upstream node.

 More specifically, an internal cell header is prepended to an ATM multicast cell within the input port prior to forwarding the cell to the output port from which the logical multicast is to occur. The internal cell
20 header includes a multicast identifier (MID) which is converted to a local connection identifier (CID) via a table look up within the respective output port. The local connection identifier serves as an offset value. The multicast cell is stored in one of a plurality of queues
25 within the respective output port along with the internal cell header.

 The multicast cell is retrieved from the respective queue based upon a scheduling algorithm which may employ strict priority scheduling algorithm, a weighted priority
30 scheduling algorithm or any other suitable scheduling technique. When the multicast cell is retrieved from the respective queue, it is placed in one of a plurality of Output Cell Buffers associated with the respective output port. Each Output Cell Buffer stores a single cell for
35 transmission over the communications link.

 At call setup, a plurality of tables are configured which are employed in the presently disclosed logical

multicasting technique. A first table contains VPI/VCI addresses and a chaining bit in each record entry and is indexed by the local connection identifier. A second table contains the next offset value to be employed and is likewise indexed by the local connection identifier offset. A third table contains a number of entries for cell statistics and is also indexed by the local connection identifier offset.

Logical multicasting proceeds in the following manner.

The outgoing VPI/VCI address to be employed in the first outgoing cell of a series of outgoing logical multicast cells is read from the first table using the local connection identifier (CID) in the internal cell header as an offset into the table. The VPI/VCI address from the indexed location is inserted into the cell header for the outgoing cell within the applicable Output Cell Buffer and the cell is transmitted out of the respective output port.

The first table also includes a chaining bit for each indexed record. If the chaining bit is set, such indicates that the cell is to be logically multicast and a further leaf of the logical multicast cell exists for transmission.

Thus, if the chaining bit is set, the cell is maintained within the respective Output Cell Buffer for the duration of the logical multicast sequence.

In the event a further cell is queued for transmission, a second cell is read from one of the queues in accordance with the specified scheduling algorithm and the cell is stored in a second one of the plurality of Output Cell Buffers. The local connection identifier within the internal cell header is employed as an index into the first table to locate the outgoing VPI/VCI address for the additional cell.

When a chaining bit is set in the respective record, an additional lookup is performed in the second table using the same index value as was employed to locate the VPI/VCI address for the prior cell in the logical multicast sequence. The lookup yields a new offset which replaces

the local connection identifier within the internal cell header. The new offset is employed for the next VPI/VCI lookup in the logical multicast sequence. In the foregoing manner, a linked list or chain of offsets may be generated to locate the outgoing VPI/VCI addresses for the cells to be logically multicast from the respective port. The chain terminates when the chaining bit is not set. At this point, the respective Output Cell Buffer is available to receive another cell specified by the Scheduler.

Cells are transmitted from the plurality of Output Cell Buffers in a round robin sequence so long as cells are present for transmission in the respective buffers. In the event that a buffer is empty, the cell within the next Output Cell Buffer in the round robin sequence is transmitted from the respective output port. As a consequence of the toggling of transmit opportunities among the Output Cell Buffers containing a cell to transmit, cell delay variation is lowered for cells within the plurality of cell queues.

The logically multicast cells are received at a downstream network switch or node and are switched out the port at the downstream node associated with the VPI/VCI address specified within the respective cells. In the foregoing manner, even if the downstream node does not support spatial multicast, it will appear to be spatially multicasting cells by virtue of the logical multicasting operation performed at the upstream switch.

The presently disclosed logical multicasting technique reduces the possibility that a low priority logical multicast cell with many leaves will block a high priority cell, avoids the need to dedicate an input and output port to logical multicast support, and minimizes cell delay variation for cells within the cell queues.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

The invention will be more fully understood by reference to the following Detailed Description of the Invention in conjunction with the drawing, of which:

Fig. 1 is a block diagram of a network switch operative to perform logical multicasting in accordance with the present invention;

Fig. 2 illustrates exemplary data structures employed to perform logical multicasting in the network switch of Fig. 1;

Fig. 3 is a more detailed block diagram of one of the output ports of the network switch of Fig. 1 employed to achieve logical multicasting of ATM cells;

Fig. 4a is a flow diagram depicting the method of storing cells in output queues within the output ports of the network switch of Fig. 1; and

Fig. 4b is a flow diagram depicting the method of performing logical multicasting in accordance with the present invention.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, a method and apparatus are disclosed for logically multicasting cells from an Asynchronous Transfer Mode (ATM) switch to permit spatial multicast of cells in a downstream node which does not support spatial multicast. A network switch 10 operative in accordance with the present invention is depicted in Fig. 1 and includes a plurality of input ports 12 designated herein as input ports 12a through 12n, respectively, a plurality of output ports 14 designated herein as output ports 14a through 14n, respectively, and a switch fabric 16 for selectively forwarding cells received at one of the input ports 12 of the network switch 10 to one or more of the output ports 14 of the switch. Each output port 14a through 14n includes a plurality of output queues 18 for storing cells in advance of dispatch from the

respective output port 14. Additionally, each output port 14a through 14n includes a plurality of Output Cell Buffers which are depicted as Buffer A and Buffer B in Fig. 1. While only two Output Cell Buffers 20 are illustrated, a
5 greater number of Output Cell Buffers may be employed within each output port. A Scheduler 22 is operative to identify cells which are queued on the output queues and to select cells for dispatch to one of the Output Cell Buffers in accordance with techniques which are discussed
10 subsequently in greater detail. Cells are transmitted from the Output Cell Buffers onto an associated communication link 24. The communications links are designated herein as links 24a through 24n.

Upon receipt of a cell 26 at input port 12b of the
15 switch 10, VPI/VCI address of the received cell is employed within the input port in a lookup operation to identify a connection identifier which is used within the switch to specify the manner of cell processing. More specifically, if the received cell is a multicast cell, a table lookup is
20 performed using the VPI/VCI address as an index to locate a multicast identifier (MID). The MID is stored within an internal cell header which is prepended to the received ATM cell. If the received cell 26 is a unicast cell a table lookup is performed using the VPI/VCI address as an index
25 to locate a connection identifier (CID) which is stored within the internal cell header.

The received cell in conjunction with the prepended cell header is switched via the switch fabric 16 to one or more of the output ports 14 of the switch 10. Upon receipt
30 at the respective output port, a table lookup is performed using the MID or the CID, as applicable, as an index into a look-up table, to generate a local connection identifier.

The local connection identifier is stored within the internal cell header in place of the MID or CID. The local
35 connection identifier stored within the internal cell header is employed within the output port as an offset in various table lookup operations which are subsequently

described in greater detail. The internal cell header and the associated cell are stored within one of the plurality of queues 18 within the respective output port 14 of the switch 10.

5 Referring to Fig. 2, the data structures employed to achieve logical multicasting in accordance with the presently disclosed method are depicted. In a preferred embodiment, two tables associated with the output port are employed to accomplish logical multicasting; namely, a
10 VPI/VCI lookup table 40 and a Next Offset Table 42. The VPI/VCI lookup table is indexed by the local connection identifier 44 within the internal cell header 46. The VPI/VCI lookup table includes a VPI/VCI field 48 which specifies the outgoing VPI/VCI address to be employed for
15 the outgoing cell and a Chain Bit field 50 which is employed in the logical multicasting of cells. When the chain bit is set, an additional lookup is performed in the Next Offset table 42 using the same CID as was employed for the last lookup within the VPI/VCI lookup table 40 for the
20 respective cell. The result of the lookup within the Next Offset table 42 is a new local connection identifier or offset which replaces the one carried within the internal cell header 46. The new offset is employed for the next lookup operation within the VPI/VCI lookup table 40. Thus,
25 a linked list of offsets is provided which allows for a series of outgoing VPI/VCI addresses to be located within the lookup table 40 and inserted in the Output Cell Buffer 20 for transmission over the applicable communication link 24.

30 A statistics table 52 is also provided in which cell statistics may be stored. In a preferred embodiment, the statistics table 52 includes a first field 54 which contains a count of User cells having the Cell Loss Priority (CLP) bit=0, a second field 56 which contains a
35 count of User cells having the CLP bit=1, a third field 58 which contains a count of Operation, Administration and Maintenance (OAM) cells having the CLP bit=1 and a fourth

field 60 which contains a count of OAM cells having the CLP bit=0.

Fig. 3 depicts in greater detail, the logic associated with each output port 14 employed to support logical multicast operation. It is noted that while the VPI/VCI lookup table, the Next Offset table and the Statistics Table are illustrated in Fig. 3, such tables may be shared across a plurality of ports which would typically be disposed on a single Input/Output Module in the network switch 10. As depicted in Fig. 3, a plurality of queues 100a through 100d are provided for storing ATM cells for dispatch from the output port. The various queues may correspond to various levels of priority or levels of service. For example, queue 100a may receive cells which are designated as having the highest priority, queue 100b may receive cells which are designated as having a priority lower than the cells stored in queue 100a, queue 100c may receive cells which are designated as having a priority lower than the cells stored in queue 100b and queue 100d may receive cells which are designated as having a priority lower than the cells stored in queue 100c. Accordingly, the cells stored within queue 100a comprise the highest priority cells and the cells stored within queue 100d comprise the lowest priority cells. ATM cells are switched to the respective output ports and stored within the queues in a conventional manner noting that an internal cell header containing an offset value is stored in the queue in conjunction with the respective cell. The outputs of the queues 100a through 100b are coupled to Output Cell Buffer A 102 and Output Cell Buffer B 104 via a bus (Queue_DAT) 106. Transfer of a cell from one of the queues to one of the Output Cell Buffers proceeds under the control of the Scheduler and Table Lookup Control Logic 106 in the manner described below.

The Scheduler employs a predetermined algorithm to select the next cell from one of the plurality of queues for dispatch to a free Output Cell Buffer 102 or 104. For

example, the queues may be organized as First In First Out queues (FIFOs) and the Scheduler may employ strict priority scheduling in which the next cell dispatched to a free Output Cell Buffer always comprises the first cell in the highest priority queue. Using the strict priority scheduling algorithm, it should be appreciated that there is a possibility that ongoing service of the higher priority queues will preclude the lower priority queues from being serviced.

Alternatively, a weighted priority scheduling algorithm may be employed to assure that all queues, irrespective of priority, are serviced. In one example of a weighted queue, assuming all queues have cells to transmit, the Scheduler 108 would cause the queue 100b to forward a cell to a free Output Cell Buffer 102 or 104 after two cells were forwarded from queue 100a. Additionally, the Scheduler 108 would cause the queue 100c to forward a cell to a free Output Cell Buffer 102 or 104 after two cells were forwarded from queue 100b and the Scheduler 108 would cause the queue 100d to forward a cell to a free Output Cell Buffer 102 or 104 after two cells were forwarded from queue 100c. It should be appreciated that the weightings employed within the Scheduler 108 may be varied for different applications and need not be equal. Applying a weighted priority scheduling algorithm allows the lower priority queues to obtain at least some of the scheduled bandwidth.

It should also be appreciated that any other suitable scheduling technique may be employed for dispatching cells from the plurality of queues to an available one of the plurality of Output Cell Buffers.

The Scheduler and Table Lookup Control Logic controls the Scheduling and dispatch of cells to an available Output Cell Buffer and ascertains the outgoing VPI/VCI address for each outgoing cell. The Scheduler 108 has an address output bus Q_Add 110 which is coupled to the address inputs of the plurality of queues 100a - 100d. Additionally, the

Scheduler 108 drives queue output enable lines Q0_OE 112, Q1_OE 114, Q2_OE 116 and Q3_OE 118 which are coupled to the respective queue FIFOs. Assertion of the applicable output enable line causes the internal cell header and the cell information to be driven from the addressed location within the respective queue onto the Queue_DAT bus 106. The Scheduler 108 also drives VPI/VCI address lines which are coupled to the Output Cell Buffers 102 and 104 and additionally drives Output Cell Buffer load signals LOB_A 122 and LOB_B 124 to cause the respective Output Cell Buffer to be loaded with data. The outputs of the Output Cell Buffers 102 and 106 are coupled to the communication link 128. A transmit controller 126 causes cells to be transmitted out of the Output Cell Buffers in an alternating or round robin sequence so long as cells are present in both buffers for transmission onto the communications link 128. Thus, if cells are present in both Output Buffers 102 and 104, the transmit controller toggles transmit opportunities between the respective Output Cell Buffers. If only one cell buffer has cells for transmission over the link, that cell buffer is afforded all transmit opportunities by the transmit controller until once again both cell buffers contain cells for transmission. If more than two Output Cell Buffers are employed, transmit opportunities are afforded to the Output Buffers having cells to transmit in a round robin sequence.

The Scheduler and Table Control Lookup logic 108 also has access to the VPI/VCI lookup table 40 the Next Offset table 42 and the Statistics table 52.

The VPI/VCI Lookup table 40 and the Next Offset Table 42 are configured when a connection is established or torn down and may be modified from time to time if, for example, a leaf of logical multicast transmission is pruned for any reason. Such configurations are performed as switch 10 administrative functions. Since chaining of leaves is permitted within the VPI/VCI Lookup Table, contiguous memory locations are not required. Accordingly, the

management and updating of the VPI/VCI Lookup Table and the Next Offset table is simplified.

Logical multicast operation in accordance with the presently disclosed technique will be more fully understood by reference to the illustrative example described below and with reference to Figs. 1-3. Assume a cell 26 is received at an input port 12b of the switch 10 and the switch 10 has been configured to logically multicast the cell out a single output port because the downstream switch coupled to that port does not support spatial multicast. The VPI/VCI address of the cell 26 is used as an index into a lookup table within the input port 12b in a conventional manner to obtain a multicast connection identifier (MID) which is inserted within an internal cell header. The internal cell header is prepended to the cell and the cell, along with the the internal cell header is switched through the switch fabric 16 for receipt by the applicable output port. Upon receipt at the output port, the multicast connection identifier within the internal cell header is used as an index into a lookup table to obtain a local connection identifier which is stored within the internal cell header. The cell 26 along with the prepended cell header is then stored on one of the plurality of queues 100a - 100d based upon the priority of the cell. While, in the present embodiment, cells are queued based upon cell priority, any other queuing criteria may be employed without departing from the inventive concepts disclosed herein. Additionally, the logical multicast functionality herein disclosed are located after any spatial multicast mechanisms (not shown) if such are employed in the switch.

In this manner the switch 10 may perform any mix of spatial and logic multicast operations.

When the Scheduler 108 determines that the cell 26, which comprises the logical multicast cell is the next cell for dispatch, based upon the scheduling algorithm employed, and an Output Cell Buffer is free, the cell 26 within the respective queue is addressed by the Scheduler 108 over

Q_Add lines 110 and the queue output is enabled to drive the Queue_DAT data bus 106. The Scheduler reads the local connection identifier 44 information from the internal cell header to obtain an offset into the VPI/VCI Lookup Table 40. In the present example, referring to Fig. 2, the offset 1 as used as an index into the Lookup Table 40 to obtain the outgoing VPI/VCI address A for the cell 26. Assuming that Output Cell Buffer A 102 is available, the Scheduler 108 asserts the load output buffer A signal LOB_A to cause the cell to be loaded into Output Cell Buffer A with the proper outgoing VPI/VCI address. Applicable cell statistics are written to the Statistics table 52 at the same offset. The Chain bit 50 at the offset 1 is next tested to determine whether such bit is set. If the chain bit 50 is set, such indicates that the cell is a logical multicast cell with additional leaves that need to be transmitted and the cell is retained in the applicable Output Cell Buffer for the remainder of the logical multicast sequence. Since the chain bit 50 is set in the record entry within the VPI/VCI Lookup Table corresponding to offset entry 1, the cell 26 is kept in Output Cell Buffer A. If the Chain bit is set, the Table Lookup Control Logic performs a further look-up within the Next Offset Table 42 using the current offset value as the index. Using the current offset value of "1" as an index into the Next Offset Table 42, the next offset of "2" is obtained. This offset value is then employed as an index into the VPI/VCI Lookup Table 40 to obtain the next outgoing VPI/VCI address within the logical multicast sequence. In the present example, the VPI/VCI address B is retrieved from the VPI/VCI Lookup Table 40, and the VPI/VCI address B is loaded into the Output Cell Buffer 102 for transmission after the transmission of the cell bearing the VPI/VCI address A. The payload for the cell remains the same. Applicable cell statistics are written to the Statistics table 52 at the same offset. The Chain bit 50 at offset 2 within the VPI/VCI Lookup Table is tested and,

in the present example, it is determined that the chain bit is set indicating that a further leaf exists for transmission. Accordingly, the offset 2 is employed as an index into the Next Offset Table 42 to identify the offset to be employed in the next VPI/VCI lookup. In the present example, using the offset 2 as the index into the Next Offset Table, the offset 4 is retrieved. This offset is employed as the index into the VPI/VCI Lookup Table 40 to identify the outgoing VPI/VCI address C for the next cell to be transmitted in the logical multicast sequence of cells. After the transmission of the leaf bearing the VPI/VCI address B, the VPI/VCI address C is loaded into the Output Cell Buffer 102 for transmission out of the respective output port. Applicable cell statistics are written to the Statistics table 52 at the same offset. The Chain Bit 50 at offset 4 is next tested and it is determined that the chain bit 50 at such offset is not set.

Since the Chain bit 50 is not set, the Output Cell Buffer 102 is freed after transmission of the cell bearing the VPI/VCI address C which comprises the last leaf in the logical multicast sequence. Once the Output Cell Buffer is freed, another cell may be dispatched from the Scheduler to that Output Cell Buffer.

It should be noted that the number of leaves that can be supported per circuit for logical multicast is equal to the number of connections supported within the switch. In a preferred embodiment of the switch, sixteen thousand connections may be accommodated at any given time.

It should further be noted that cells within the queue may comprise logical multicast or unicast cells. Unicast cells are handled as a degenerate case of the logical multicast case in which the first test of the Chain bit 50 within the VPI/VCI Lookup Table 40 results in a determination that the chain bit is not set. In such event, the Output Cell Buffer for the respective unicast cell is freed after the transmission of the single unicast cell from the Output Cell Buffer.

The utilization of multiple output buffers which are provided transmit opportunities in a round robin sequence provides significant performance advantages in a network switch supporting logical multicast operation. It will be appreciated that a low priority logical multicast cell having numerous leaves may be dispatched to the Output Cell Buffer. If a single Output Cell Buffer is provided, all cells, including higher priority cells would be blocked pending completion of the logical multicast operation.

Blocking introduces undesirable cell delay variations. The provision of a plurality of Output Cell Buffers which are afforded transmit opportunities in a round robin sequence allows logical multicast operation to be provided in a non-blocking manner and minimizes the introduction of such undesirable cell delay variations.

The methods employed in accordance with the present invention are further illustrated by reference to the flow charts of Figs. 4a and 4b.

Fig. 4a depicts the process by which cells are received, switched to the applicable output port and queued for dispatch to an Output Cell Buffer. Referring to Fig. 4a, a cell is received at an input port of the switch as depicted in step 200. Following receipt of the cell, a determination is made whether the cell is a logical multicast cell or a unicast cell as depicted in decision step 202. If the cell is a logical multicast cell, a Multicast Connection identifier (MID) is assigned to the cell based upon the incoming VPI/VCI address and the MID is prepended to the cell in an internal cell header as illustrated in step 204. The cell is then switched to the applicable output port as shown in step 206. The cell may be switched to plural output ports if the cell is a spatial multicast cell. If it is determined in step 202 that the cell is not a multicast cell, but rather a unicast cell, a Unicast Connection Identifier (CID) is assigned based upon the incoming VPI/VCI address and is the CID prepended to the cell within an internal cell header as

depicted in step 208. The unicast cell along with the cell header is next switched to the applicable output port as depicted in step 206. The multicast identifier (MID) or the (CID) is next converted to a local connection identifier at the I/O module containing the respective output port and the local connection identifier is stored within the internal cell header. The cell along with the prepended cell header is stored within the appropriate queue for the respective output port as illustrated in step 210. Dequeuing of cells proceeds as illustrated in Fig. 4b.

Referring to Fig. 4b the Output Cell Buffers are tested to determine whether one of the plurality of cell buffers is free as illustrated in decision step 250. If none of the Output Cell Buffers are free, the Scheduler defers until one of the buffers is free. If it is determined by the Scheduler that one of the cell buffers is available to receive a cell, the Scheduler retrieves a cell from one of the queues in accordance with a predetermined scheduling algorithm and stores the selected cell within the available Output Cell Buffer as depicted in step 252. The outgoing VPI/VCI for the respective cell is next retrieved from a lookup table using the local connection identifier within the internal cell header for the respective cell as an index. The retrieved VPI/VCI address is stored in the applicable Output Cell Buffer as shown in step 254. As previously discussed, Output Cell Buffers are afforded transmit opportunities in a round robin sequence.

Thus, as illustrated in inquiry step 256, a determination is made whether the respective Output Cell Buffer may transmit its cell out of the output port onto the communication link. If the respective Output Cell Buffer cannot transmit, it defers until afforded a transmit opportunity by the transit controller 126. When afforded a transmit opportunity, the cell is transmitted from the Output Cell Buffer onto the associated communications link as depicted in step 258. The chain bit at the offset

specified by the local connection identifier is next tested as illustrated in decision step 260 to determine whether the cell comprises a logical multicast cell having further outgoing leaves. If inquiry step 260 results in a determination that the chain bit is not set, the respective Output Cell Buffer is freed and made available to receive another cell following the transmission of the current cell within the buffer as depicted in step 262 and flow passes to inquiry step 250. If the chain bit is set, indicating that the cell comprises a logical multicast cell with a further leaf for transmission, the cell payload remains within the applicable Output Cell Buffer and the present local connection identifier is employed as an index into the Next Offset Table to retrieve a new local connection identifier to be employed for further look-ups as shown in step 264. The new connection identifier is then employed as an index into the VPI/VCI Lookup Table to retrieve the VPI/VCI address for the next outgoing leaf in the logical multicast sequence. Flow proceeds to inquiry step 256 and the process continues.

It should be appreciated that the cells dispatched to the Output Cell Buffers may comprise logical multicast or unicast cells and thus, at any given time either type of cell may be resident in either Output Cell Buffer. Thus, look-ups of outgoing VPI/VCI addresses for the cells contained within the Output Cell Buffers may be interleaved.

While the embodiment depicted in Fig. 3 includes four FIFO queues 100a - 100d, it should be appreciated that any number of queues may be employed and any suitable scheduling technique may be used for dispatching cells from the respective queues to the free Output Cell Buffers. For example, one queue may be provided for each connection being serviced within the switch 10. Additionally, while two Output Cell Buffers are employed in the embodiment depicted in Figs. 1 and 3, a greater number of Output Cell Buffers may be employed to further minimize the possibility

of blocking and further reduce cell delay variations that may be introduced as a consequence of logical multicast support. Transmit opportunities are granted to the Output Cell Buffers by the Transmit Controller in a round robin sequence.

In the above described embodiments, the VPI/VCI lookup table and the Next Offset table are described as two separate tables. It should be appreciated that such tables may comprise a single table indexed by the local connection identifier.

The above described embodiments are intended to be exemplary. It will be appreciated that variations to and modification of the above described methods and apparatus for performing logical multicasting of cells from a network switch are possible and will be apparent to those of ordinary skill in the art without departing from the presently disclosed inventive concepts. Accordingly, the invention is not to be viewed as limited to the exemplary embodiments but rather, solely by the scope and spirit of the appended claims.

CLAIMS

1. A method for performing logical multicasting of a cell from an output port of a network device comprising the steps of:

5 in a first storing step, storing a logical multicast cell in a first output cell buffer of a plurality of output cell buffers associated with said output port, wherein each one of said plurality of output cell buffers is in electrical communication with a single communications link;

10 in a first transmitting step, transmitting a plurality of cells corresponding to leaves of said logical multicast cell including a last cell of said plurality of cells corresponding to a last leaf onto said communications link from said first output cell buffer;

15 in a second storing step, storing at least one other cell in a second output cell buffer of said plurality of output cell buffers; and

20 in a second transmitting step, transmitting said at least one other cell over said communications link from said second output cell buffer between a pair of leaves of said logical multicast cell transmitted from said first output cell buffer.

25 2. The method of claim 1 further comprising the step of providing an indication that the respective output cell buffer of said plurality of output cell buffers is available for receipt of an additional cell following the transmission of one of a unicast cell and the last leaf of said logical multicast cell from the respective output cell
30 buffer.

3. The method of claim 2 further comprising the step of:

35 prior to said first transmitting step, storing said logical multicast cell in one of a plurality of output queues associated with said output port according to a predetermined storage criteria and forwarding said logical multicast cell to said first output cell buffer from the

respective output queue in accordance with a predetermined scheduling methodology when the respective output cell buffer is identified as being available to receive a cell.

5 4. The method of claim 3 wherein said predetermined scheduling methodology comprises the steps of:

 associating priorities with each of the plurality of output queues; and

 dispatching cells from the respective output queues to
10 selected ones of said plurality of output cell buffers based upon a strict priority methodology.

5. The method of claim 3 wherein said predetermined scheduling methodology comprises the steps of:

15 associating priorities with each of the plurality of output queues; and

 dispatching cells from the respective output queues to
 selected ones of said plurality of output cell buffers
 based upon a weighted priority methodology.

20 6. The method of claim 3 further comprising the steps of:

 storing an offset value within the respective queue in association with said logical multicast cell;

 performing an address lookup operation utilizing said
25 offset value as an index to locate an address to be transmitted with the respective leaf of said logical multicast cell and storing said address within said first output cell buffer;

 performing a chaining information lookup operation
30 utilizing said offset value as an index to locate chaining information associated with said offset value, said chaining information comprising an indication of whether any further leaves of the logical multicast cell are to be transmitted from the respective output cell buffer;

35 testing said chaining information to determine whether a further leaf associated with said multicast cell is to be transmitted from the respective output cell buffer;

in the event said testing step indicates that a further leaf is to be transmitted, performing a next offset value lookup operation utilizing said offset value as an index to locate a new offset value and repeating said performing steps and said testing step using said new offset value as said offset value;

in the event said testing step indicates that no further leaves are to be transmitted, identifying said first output cell buffer as available to receive another cell following the transmission of the last leaf associated with the respective logical multicast cell.

7. The method of claim 6 further comprising the step of storing statistics information in at least one statistic table indexed by said offset value.

8. The method of claim 6 wherein said address comprises a Virtual Path Identifier/Virtual Connection Identifier address for transmission with the corresponding leaf of said logical multicast cell.

9. The method of claim 1 wherein said second transmitting step further comprises the step of transmitting cells from each output cell buffer having a cell stored in the respective buffer for transmission in addition to said cell stored in said first and second output cell buffers in a predetermined transmission sequence so as to interpose the transmission of cells stored in said output cell buffers other than said first output cell buffers between the transmission of leaves of said multicast cell from said first output cell buffer.

10. The method of claim 9 wherein said predetermined transmission sequence comprises a round robin sequence.

11. Apparatus for logical multicasting a cell having a predetermined number of leaves from an output port in a

network device comprising:

5 a plurality of output cell buffers associated with said output port, each one of said plurality of output cell buffers being coupled to a single communication link associated with said output port, each of said plurality of output cell buffers for storing one of a unicast cell and a logical multicast cell and wherein said logical multicast cell has a plurality of leaves corresponding in number to the number of copies of the logical multicast cell to be transmitted;

10 a cell transmission controller in electrical communication with each of said plurality of output cell buffers, said controller operative to control the transmission of said unicast cell and said leaves of said logical multicast cell from the respective ones of said plurality of output cell buffers containing cells available for transmission;

15 said cell transmission controller being operative to cause the transmission of leaves of a multicast cell stored in a first one of said plurality of output cell buffers and at least one other cell stored in at least one other one of said plurality of output cell buffers in a predetermined sequence such that said at least one other cell is transmitted between a pair of said leaves of said logical multicast cell.

20 12. The apparatus of claim 11 wherein said predetermined sequence comprises a round robin sequence among said plurality of output cell buffers.

25 13. The apparatus of claim 12 wherein said predetermined sequence comprises a round robin sequence among said plurality of output cell buffers containing cells available for transmission.

30 14. The apparatus of claim 11 further comprising:

35 a plurality of output queues, each one of said

plurality of output queues for storing cells received by said computer network device and an offset value associated with the respective cell, wherein said received cells comprise said logical multicast cells and said unicast cells;

a scheduler operative to selectively forward said cells from said queues to selected ones of said output cell buffers upon selection of the respective cell for dispatch by said scheduler and upon an indication that the respective cell buffer is available to receive the respective cell.

15. The apparatus of claim 14 further comprising:

at least one lookup table indexed by said offset value, said at least one lookup table containing an address field containing an address, a chaining field containing an indication of whether the address within the address field at the corresponding offset value corresponds to the last address associated with the last leaf of a multicast cell and a next offset value field containing an indication of the next offset value to be employed in the event the contents of said chaining field indicate another leaf of said multicast cell is to be transmitted; and

a multicast controller operative to iteratively:

access an address contained within said address field in response to said offset value and to store said address within the respective output cell buffer containing the respective multicast cell in advance of the transmission of that cell over the associated communication link;

test the contents of the chaining field indexed by the respective offset value to determine if any further leaves of said multicast cell remain to be transmitted;

access the next offset value to be employed as said offset value in the next subsequent lookup operation employing said at least one lookup table in

the event the testing of said chaining field indicates that a further leaf of said multicast cell remains to be transmitted.

5 16. The apparatus of claim 15 wherein said at least one lookup table further comprises at least one field indexed by said offset value for storing statistics information.

10 17. The apparatus of claim 15 wherein said address comprises a virtual path identifier/virtual connection identifier address address.

15 18. The apparatus of claim 14 wherein each of said plurality of output queues has an associated priority and said scheduler is operative to selectively forward a cell from a selected one of the plurality of queues to a selected one of said output cell buffers applying a strict priority scheduling technique.

20 19. The apparatus of claim 14 wherein each of said plurality of output queues has an associated priority and said scheduler is operative to selectively forward a cell from a selected one of the plurality of queues to a selected one of said output cell buffers applying a
25 weighted priority scheduling methodology in which cells are forwarded from at least one of said plurality of output queues more frequently than from at least one other one of said plurality of output queues.

AMENDED CLAIMS

[received by the International Bureau on 10 November 1999 (10.11.99);
original claims 1-2,6,8-9,11 and 15 amended;
remaining claims unchanged (7 pages)]

1. A method for performing logical multicasting of a
cell having a predetermined number of leaves from an
5 output port of a network device comprising the steps of:

in a first storing step, storing said logical
multicast cell in a first output cell buffer of a
plurality of output cell buffers associated with said
output port, wherein each one of said plurality of output
10 cell buffers is in electrical communication with a single
communications link;

in a first transmitting step, transmitting in a
predetermined sequence a predetermined number of leaf
cells corresponding to said leaves of said logical
15 multicast cell onto said communications link from said
first output cell buffer, said leaf cells including a last
leaf cell corresponding to a last leaf of said leaves;

in a second storing step, storing at least one other
cell in a second output cell buffer of said plurality of
20 output cell buffers; and

in a second transmitting step, transmitting said at
least one other cell over said communications link from
said second output cell buffer between two of said leaf
cells transmitted from said first output cell buffer in
25 said predetermined sequence.

2. The method of claim 1 further comprising the step of
providing an indication that the respective output cell
buffer of said plurality of output cell buffers is
30 available for receipt of an additional cell following the
transmission of one of a unicast cell and the last leaf

cell of said logical multicast cell from the respective output cell buffer.

3. The method of claim 2 further comprising the step of:

5 prior to said first transmitting step, storing said logical multicast cell in one of a plurality of output queues associated with said output port according to a predetermined storage criteria and forwarding said logical multicast cell to said first output cell buffer from the
10 respective output queue in accordance with a predetermined scheduling methodology when the respective output cell buffer is identified as being available to receive a cell.

4. The method of claim 3 wherein said predetermined
15 scheduling methodology comprises the steps of:

associating priorities with each of the plurality of output queues; and

dispatching cells from the respective output queues to selected ones of said plurality of output cell buffers based
20 upon a strict priority methodology.

5. The method of claim 3 wherein said predetermined scheduling methodology comprises the steps of:

25 associating priorities with each of the plurality of output queues; and

dispatching cells from the respective output queues to selected ones of said plurality of output cell buffers based upon a weighted priority methodology.

30 6. The method of claim 3 further comprising the steps of:

storing an offset value within the respective queue in association with said logical multicast cell;

performing an address lookup operation utilizing said offset value as an index to locate an address to be transmitted with a corresponding leaf cell of said logical multicast cell and storing said address within said first output cell buffer;

performing a chaining information lookup operation utilizing said offset value as an index to locate chaining information associated with said offset value, said chaining information comprising an indication of whether any further leaf cells are to be transmitted from the respective output cell buffer;

testing said chaining information to determine whether said chaining information indicates that a further leaf cell is to be transmitted from the respective output cell buffer; and

in the event said chaining information indicates that a further leaf cell remains to be transmitted, performing a next offset value lookup operation utilizing said offset value as an index to locate a new offset value and repeating said performing steps and said testing step using said new offset value as said offset value; or

in the event said chaining information does not indicate that a further leaf cell remains to be transmitted, identifying said first output cell buffer as available to receive another cell following the transmission of said last leaf cell.

7. The method of claim 6 further comprising the step of storing statistics information in at least one statistic table indexed by said offset value.

5 8. The method of claim 6 wherein said address comprises a Virtual Path Identifier/Virtual Connection Identifier address for transmission with the corresponding leaf cell of said logical multicast cell.

10 9. The method of claim 1 wherein said second transmitting step further comprises the step of transmitting additional cells from each output cell buffer storing said additional cells for transmission, in addition to said logical multicast cell stored in said
15 first output cell buffer and said at least one other cell stored in said second output cell buffer, in a predetermined transmission sequence so as to interpose the transmission of said additional cells between the transmission of said leaf cells of said multicast cell
20 from said first output cell buffer.

10. The method of claim 9 wherein said predetermined transmission sequence comprises a round robin sequence.

25 11. Apparatus for logical multicasting of a cell having a predetermined number of leaves from an output port in a network device comprising:

a plurality of output cell buffers associated with said output port, each one of said plurality of output
30 cell buffers being coupled to a single communication link associated with said output port, a first of said

plurality of output cell buffers for storing said logical multicast cell and a second of said plurality of output cell buffers for storing a unicast cell;

5 a cell transmission controller in electrical communication with each of said plurality of output cell buffers, said controller operative to control the transmission of a predetermined number of leaf cells corresponding to said leaves of said logical multicast cell in a predetermined sequence from said first output
10 cell buffer and said unicast cell from said second output cell buffer;

said cell transmission controller being operative to cause said transmission of said unicast cell between transmissions of two of said leaf cells in said
15 predetermined transmission sequence.

12. The apparatus of claim 11 wherein said predetermined sequence comprises a round robin sequence among said plurality of output cell buffers.

20

13. The apparatus of claim 12 wherein said predetermined sequence comprises a round robin sequence among said plurality of output cell buffers containing cells available for transmission.

25

14. The apparatus of claim 11 further comprising:

a plurality of output queues, each one of said plurality of output queues for storing cells received by said computer network device and an offset value associated
30 with the respective cell, wherein said received cells

comprise said logical multicast cells and said unicast cells;

a scheduler operative to selectively forward said cells from said queues to selected ones of said output cell buffers upon selection of the respective cell for dispatch by said scheduler and upon an indication that the respective cell buffer is available to receive the respective cell.

15. The apparatus of claim 14 further comprising:

10 at least one lookup table indexed by said offset value, said at least one lookup table containing an address field containing an address, a chaining field containing an indication of whether the address within the address field at the corresponding offset value
15 corresponds to the last address associated with the last leaf cell of a multicast cell and a next offset value field containing an indication of the next offset value to be employed in the event the contents of said chaining field indicate another leaf cell is to be transmitted; and

20 a multicast controller operative to iteratively:

access an address contained within said address field in response to said offset value and to store said address within the respective output cell buffer containing the respective leaf cell in advance of the
25 transmission of that leaf cell over the associated communication link;

test the contents of the chaining field indexed by the respective offset value to determine if any further leaf cells remain to be transmitted; and

30 access the next offset value to be employed as said offset value in the next subsequent lookup operation

employing said at least one lookup table in the event the testing of said chaining field indicates that a further leaf cell remains to be transmitted.

5 16. The apparatus of claim 15 wherein said at least one lookup table further comprises at least one field indexed by said offset value for storing statistics information.

10 17. The apparatus of claim 15 wherein said address comprises a virtual path identifier/virtual connection identifier address address.

15 18. The apparatus of claim 14 wherein each of said plurality of output queues has an associated priority and said scheduler is operative to selectively forward a cell from a selected one of the plurality of queues to a selected one of said output cell buffers applying a strict priority scheduling technique.

20 19. The apparatus of claim 14 wherein each of said plurality of output queues has an associated priority and said scheduler is operative to selectively forward a cell from a selected one of the plurality of queues to a selected one of said output cell buffers applying a weighted priority
25 scheduling methodology in which cells are forwarded from at least one of said plurality of output queues more frequently than from at least one other one of said plurality of output queues.

1/4

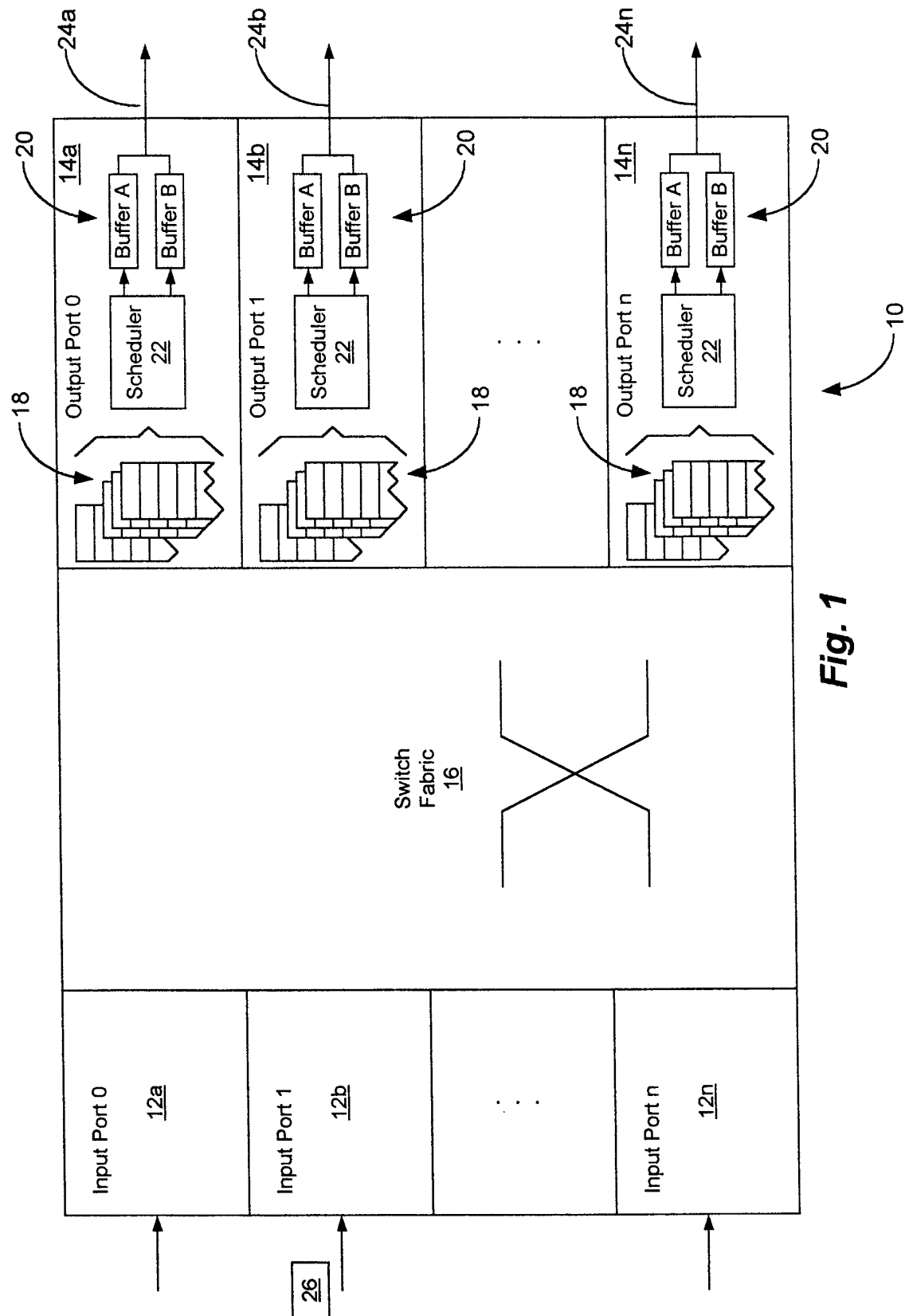


Fig. 1

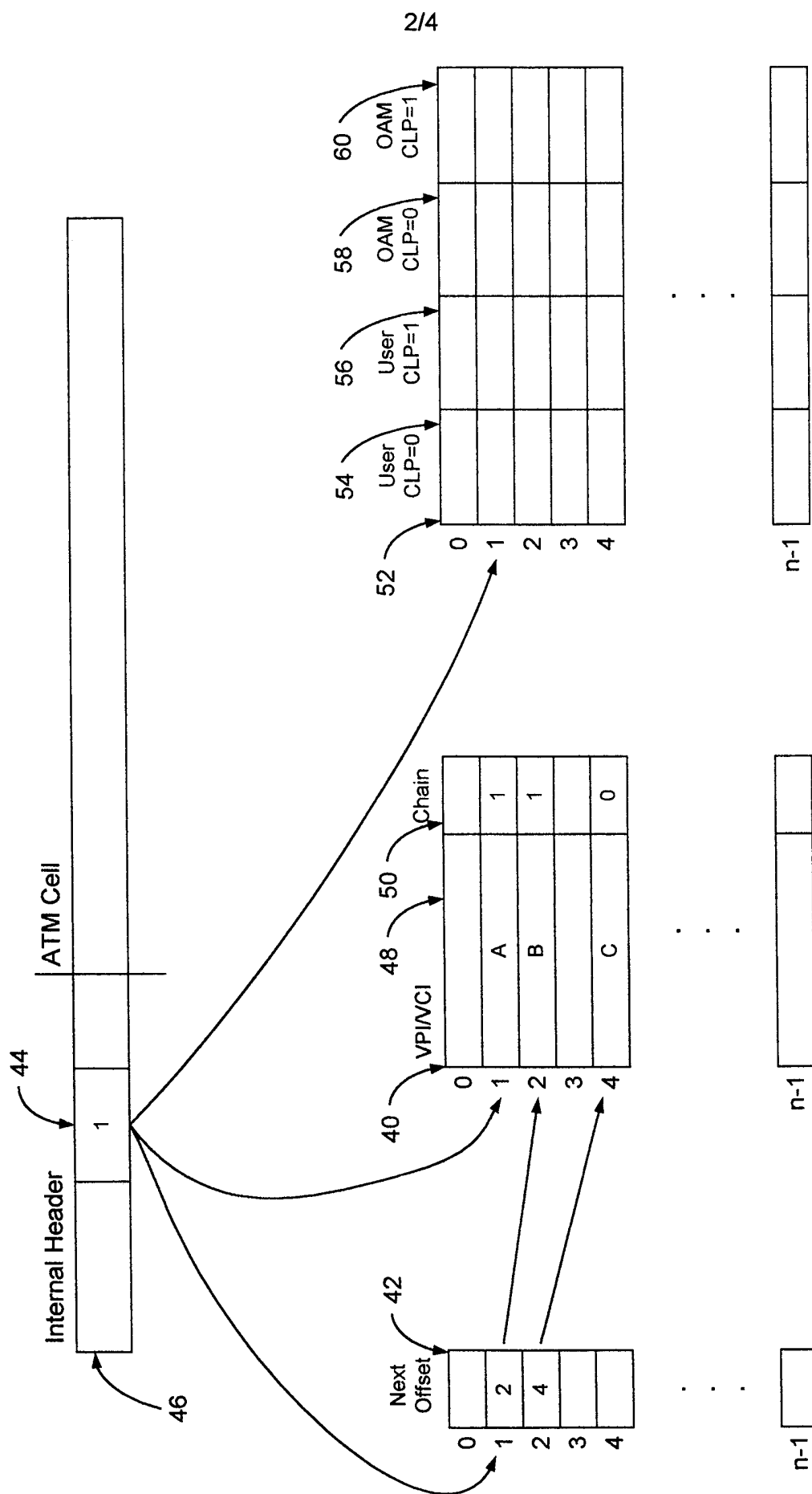
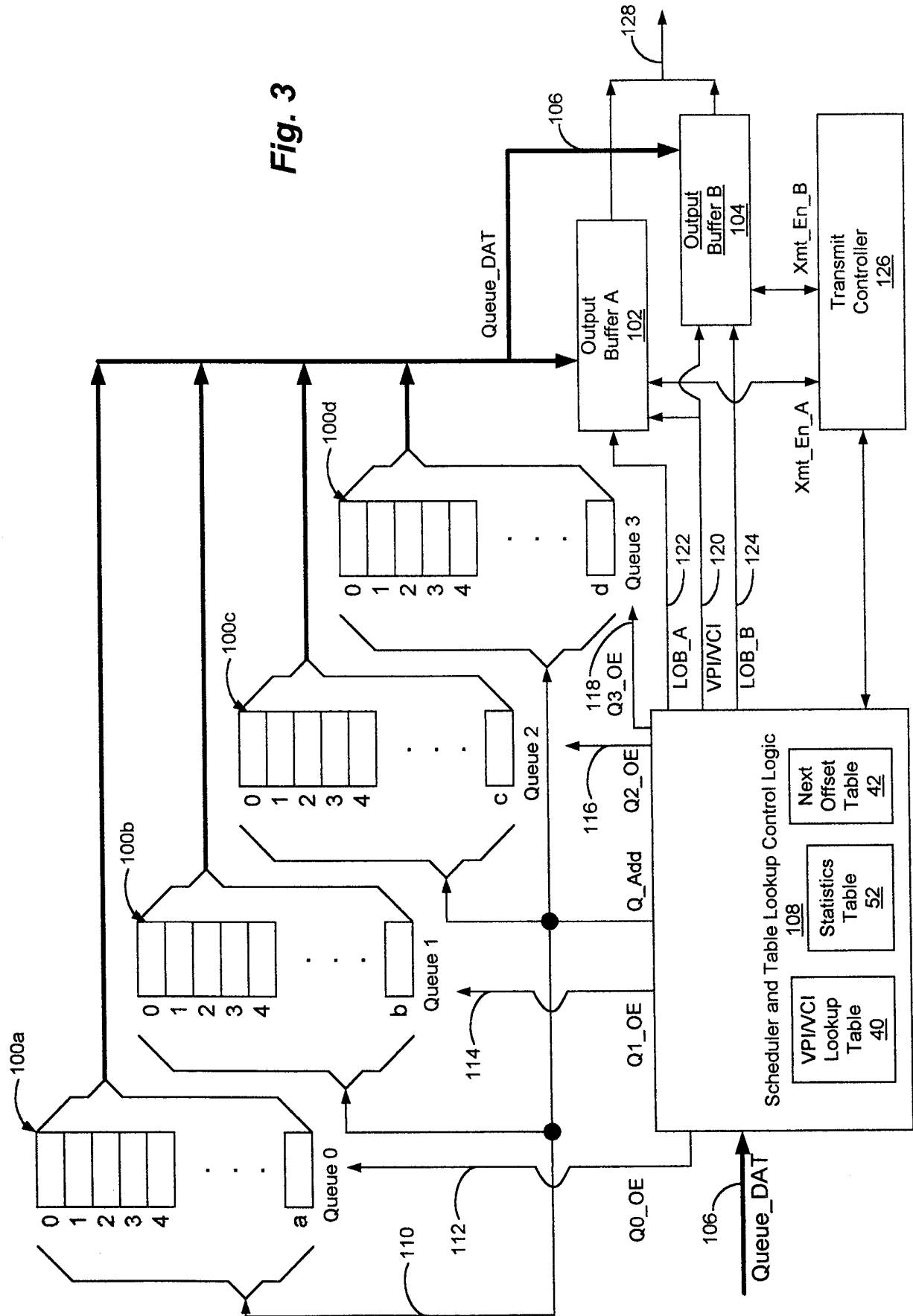
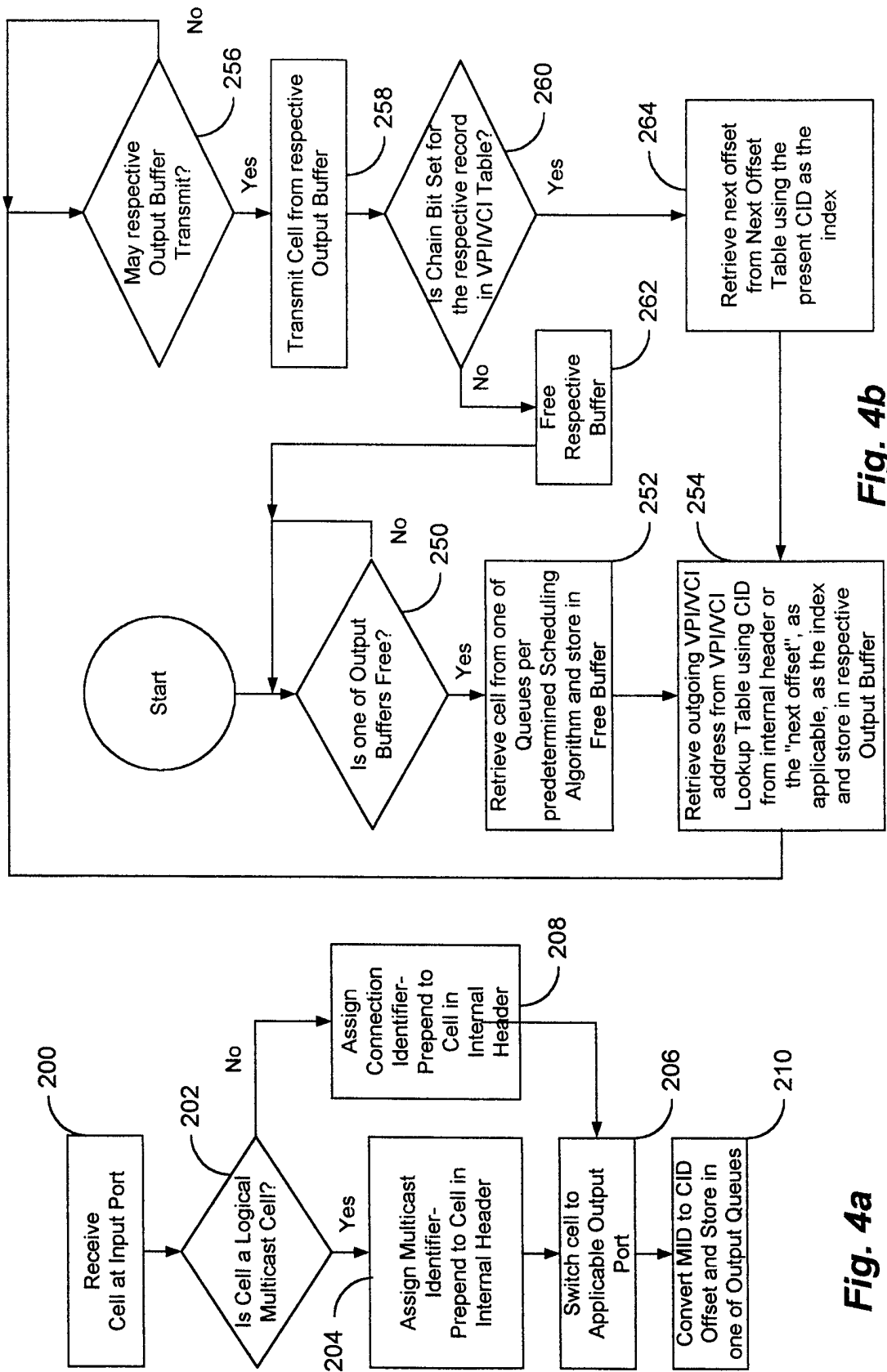


Fig. 2

Fig. 3





INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/13681

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : H04L 12/28, 12/56

US CL : 370/417

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/412, 417, 418, 428, 429, 395, 398, 399, 229, 230, 235, 237, 389, 392

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,577,035 A (HAYTER ET AL) 19 NOVEMBER 1996, see fig. 1.	1-19
A,P	US 5,859,835 A (VARMA ET AL) 12 JANUARY 1999, see entire reference.	1-19
A	US 5,583,861 A (HOLDEN) 10 DECEMBER 1996, see fig. 4.	1-19
A	US 5,689,506 A (CHIUSSI ET AL) 18 NOVEMBER 1997, see entire reference.	1-19



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents	*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
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*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
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)	*Z* document member of the same patent family
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	

Date of the actual completion of the international search

02 AUGUST 1999

Date of mailing of the international search report

10 SEP 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

AJIT PATEL

Telephone No. (703) 305-3900