

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 March 2003 (06.03.2003)

PCT

(10) International Publication Number
WO 03/019882 A1

(51) International Patent Classification⁷: H04L 12/56, H04Q 11/04, G06F 13/00

(21) International Application Number: PCT/US02/26743

(22) International Filing Date: 21 August 2002 (21.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/935,444 22 August 2001 (22.08.2001) US

(71) Applicant: PLURIS, INC. [US/US]; 10455 Bandley Drive, Cupertino, CA 95014 (US).

(72) Inventors: TUCK, Russell, R., III; 1136 S. Blaney Avenue, San Jose, CA 95129 (US). AGARWAL, Puneet; 50 Chumasero Drive, 10M, San Francisco, CA 94132 (US).

(74) Agent: BOYS, Donald, R.; P.O. Box 187, Aromas, CA 95004 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/019882 A1

(54) Title: VIRTUAL EGRESS PACKET CLASSIFICATION AT INGRESS

(57) Abstract: In a network packet router having one or more ingress and egress ports, a method is implemented at an ingress port for egress pass/drop determination for packets, comprising the steps of (a) noting header combinations and values, and egress port destination for incoming packets; (b) comparing the header combinations and values with rule sets associated with the header combinations and values including egress port identities, and (c) returning a determination of pass or drop for the packet.

Virtual Egress Packet Classification at Ingress

by inventor(s)

Russ Tuck and Puneet Agarwal

5

Field of the Invention

The present invention is in the field of routing digital data over a data-packet-network (DPN) and pertains more particularly to methods and apparatus for
10 classifying packets to determine routing requirements.

Background of the Invention

15 In the art of routing digital data through data-packet networks, research and development of methods for more efficient handling of data packets continues. Generally speaking, a data packet is a digitized and organized block of binary data that is a “packaged” portion of a specific communication or data transfer from a source location to an ultimate destination on a network. A data packet typically has
20 one or more headers, and a data body. The packet headers are used for data routing through the network. During routing from a source location to a routed destination, data packets may be processed at one or more stops or routing points along the way. These hops, as they are often termed, are between data routers and, in some cases, server nodes distributed through the network. Data packets are, for example,
25 commonly routed over the Internet network and commonly include Internet Protocol (IP) headers and Transmission Control Protocol (TCP) headers. It is well-known in the routing art that there are typically several hops for a packet between a source and a destination. It is also well-known that there are industry-accepted and applied procedures and protocols in routing, so that the many parts of the Internet (for
30 example) will continue to operate seamlessly, even though hardware and software from many different sources and companies is used. Further, the procedures by which

- 2 -

routers function involve many internal processes and messages between routers. For example, to operate successfully every router needs to keep track of its own position in the overall network, the position and functioning characteristics of its nearest neighbors, the output (egress) ports by which it must transmit packets previously 5 received at input (ingress) ports to be sure the packets follow a best route to the final destination, and so on.

A router typically has more than one ingress port and more than one egress port. The ports are often organized so each specific port functions for both ingress and egress. For descriptive purposes, however, it is quite useful to treat ingress and 10 egress ports as separate entities, because they are logically separate and are often implemented as separate entities. A packet received at any ingress port is pre-processed at that port by, for example, checking the header information for type, source and destination, port numbers, and so forth, and determining which of potentially many rules and processes apply, and then processing the packet by 15 applying the determined procedures. Some packets may be data packets for such as a video stream or a Web page, for example, which may be processed by re-transmitting them at whatever egress port is determined to be coupled to the next node to which they should go on the way to the final destination. Other packets may be determined to be queries from a neighboring router, which may be diverted to a CPU for a 20 subsequent answer to be prepared and sent back to the neighbor. There are many possibilities, and every packet must be processed so a determination can be made, and the correct action taken.

One of the functions in the art that routers practice on packets has to do with a quick decision to pass a packet on to further processing, or to simply drop the packet. 25 This function in the state-of-the-art takes place at both ingress and egress ports. There are a variety of reasons that an incoming packet received at an ingress port might be dropped. For example, packets of an unknown type, which may be malicious, should not be processed. As another simple example, packets that have a destination that cannot be reached according to current routing tables can be dropped, because

- 3 -

they cannot go anywhere anyway. There are a number of other possibilities well-known in the art.

There are also a variety of reasons why packets at an egress port might need to be dropped, even though they have been accepted and processed in the router, and 5 have arrived at an egress port. There may be, for example, requests from certain organizations for receiving only packets meeting certain criteria. There are a number of other reasons for dropping packets at output ports well-known in the art. In many cases the reasons for pass and drop are port specific.

A typical way that pass/drop determination is made at either ingress or egress 10 is by matching certain fields from a packet with fields in an Access Control List (ACL). The ACL is simply a lookup mechanism for matching the fields and returning a bit determined by the match (or mismatch) to pass (further process or transmit) or drop the packet. The implementation of such ACLs is described further in this specification in the section entitled "Description of the Preferred Embodiments" 15 below.

Because the reasons for dropping at ingress or at egress ports may differ, it makes good sense in the art to implement ingress pass/drop determinations separately from egress pass/drop determinations. Still, doing the determination to pass or drop requires the mechanisms to accomplish the function. That is, to do an ACL lookup 20 and return a bit every ingress port and egress port must have the hardware and/or software to do so. Typically hardware implementation is preferred for speed.

Because of the additional complexity, latency and expense of doing pass/drop determination at both ingress and egress, some manufacturers may prefer to implement the ACL pass/drop function at only ingress ports, and to forego 25 implementing the function at egress. Until the present invention this would mean not having the function at all at the egress ports.

What is clearly needed is a method and apparatus that enables a router or server to accomplish both ingress and egress pass/drop functionality at ingress ports, obviating the need for the mechanisms to accomplish the function at egress ports.

Summary of the Invention

In a preferred embodiment of the invention, in a network packet router having
5 one or more ingress and egress ports, a system implemented at an ingress port for
egress pass/drop determination for packets is provided, comprising a rule set
comprising header combinations and values for which a pass/drop decision may be
applied, the set including an egress port identity for each header combination and
value set, and a mechanism noting pertinent headers of a specific incoming packet,
10 and an egress port to which the packet is to be sent, comparing the headers with rules
in the rule set, and returning a determination of a rule to be applied.

In some embodiments of the invention the rule set association takes the form
of a lookup table, and the rule returned may determine the pass or drop for the packet.
Certain preferred embodiments are capable of implementation in the Internet.

15 In some cases there is a second lookup table for accomplishing ingress rule
determinations without egress port numbers. Also in some embodiments there is a
content addressable memory (CAM) used for header combination matches, and
pass/drop is determined by the address location of any match. In some embodiments
of the system there are two sections in the CAM, and pass or drop is determined by
20 the section of the CAM wherein a content match is found. In some other
embodiments the CAM is divided into more than two sections, and a result or action
is associated with each section.

In another aspect of the invention, in a network packet router, an ingress port
is provided, comprising an interface for receiving packets, a first mechanism for
25 noting header combinations and values, and egress ports for transmission, for
individual ones of the received packets, and a second mechanism comparing the
headers and values with a rule set comprising header combinations and values and
egress ports, and returning a rule for the packet.

In some embodiments the rule returned is a pass or drop determination for the
30 packet. Also in some embodiments the rule set association is in the form of a lookup

- 5 -

table. Further the network may be the Internet. In some embodiments there may be a second rule set association for accomplishing ingress rule determinations without egress port numbers. Also in some embodiments a content addressable memory (CAM) is used for header combination matches, and an associated rule is determined 5 by the address location of any match. The CAM may be divided into two sections, and a pass or drop determination is determined by the section wherein the address lies for the content match. In some embodiments the CAM is divided into more than two sections, and a result or action is associated with each section.

In yet another aspect a network packet router is provided, comprising one or 10 more ingress ports, and one or more egress ports. The router is characterized in that individual ones of the ingress ports comprise a first mechanism for noting header field combinations and values, and egress ports for transmission, for individual ones of the received packets, and a second mechanism comparing the headers with rules associated with the egress ports and field values, and returning a rule determination 15 for the packet.

In some embodiments of the router the rule returned determines pass or drop for the associated packet, and in some embodiments the rule set association takes the form of a lookup table. The network may be the Internet. In some cases there may be a second rule set association for accomplishing ingress pass/drop determinations 20 without egress port numbers. Also in some cases a content addressable memory (CAM) is used for header combination matches, and a rule is determined by the address location of any match. The CAM may be divided into two sections, and a pass or drop determination is determined by the section wherein the address lies for the content match. Further, in some cases the CAM is divided into more than two 25 sections, and a result or action is associated with each section.

In yet another aspect of the invention, in a network packet router having one or more ingress and egress ports, a method implemented at an ingress port for determining both ingress and egress rules for packets is provided, comprising the steps of (a) noting header combinations and values, and egress port destination for 30 incoming packets; (b) comparing the header combinations and values and egress port

- 6 -

destinations with a rule set associated with the header combinations and values including egress port destinations; and (c) returning a rule for the packet.

In preferred embodiments, in step (c), the rule returned indicates pass or drop for the associated packet. Also, in some preferred embodiments the rule set 5 associations take the form of a lookup table. The network may be the Internet.

In some embodiments of the method there is a second lookup in a separate rule set association for accomplishing ingress rule application determinations without egress port numbers.

In some embodiments a content addressable memory (CAM) is consulted for 10 rule set matches, and a rule is determined by the address location of any match. The CAM may be divided into two sections, and pass or drop is then determined by the section wherein the address is found for a content match. In some cases the CAM is divided into more than two sections, and each section is associated with a different result or action.

15 In embodiments of the invention described in enabling detail below, for the first time a system and method is provided wherein egress port pass/drop determinations may be made at ingress.

20

Brief Description of the Drawing Figures

Fig. 1 is a simplified diagram representing an Access Control List (ACL) lookup in the prior art.

25 Fig. 2 is a simplified diagram representing an Access Control List in an embodiment of the present invention.

Fig. 3 is a diagram representing a content-addressable memory in an embodiment of the present invention.

Description of the Preferred Embodiments

Fig. 1 is a simplified diagram representing an Access Control List (ACL) lookup in the prior art. The usual process is, for each packet received, to match 5 certain field values from the packet with a list of field values in the table. A value or values associated with the matching entry is returned, and the packet is processed accordingly. A pass or drop indication is normally part of the returned value. While multiple uses may be made of the result at ingress, including for example class of service, at egress it is common to return only a pass or drop indication. In Fig. 1, a, 10 for pass or drop purposes, a 1 or a 0 is returned for the set of values, and the packet is passed on for further processing or dropped accordingly. There are a variety of ways the ACL may be implemented, with preference typically given to hardware as a faster solution. Further, essentially the same mechanism typically is used at both ingress and egress ports.

15 In Fig. 1 the column in the ACL for field values shows four fields represented by capital letters A-D. This is exemplary only, as the fields and the patterns can vary according to, for example, packet type, and according to purpose for the rules to be applied. In this example the letters A, B C, D indicate particular fields, and not the values for those fields. The values in the fields are indicated as lower-case letters in 20 the field columns. As indicated in Fig. 1, an asterisk is a wild-card, and any value will do. The particular fields and the values can be anything within reason, dictated by the kinds of fields available in different sorts of packets, and the rules that one may wish to apply. IP packets, as a simple example, have fields for source address, source port, destination address and destination port, and these may well be the fields that would 25 need to be checked in many instances to apply a rule. For other packets and other needs, different fields and values may be used.

As was discussed briefly in the Background section above, a cost saving may be had for the router by implementing pass/drop functionality at only ingress ports, with the tradeoff of a loss in egress function. Theoretically the same cost saving may 30 be had for implementing only at egress, but then all packets have to be processed,

- 8 -

even though a certain percentage will be dropped, which is quite inefficient. A further advantage for doing all pass/drop functions at ingress is that no packets will then be processed in the router that will be dropped at egress.

Fig. 2 illustrates a new ACL according to an embodiment of the present invention. In this lookup, implemented only at ingress ports, the same field-matching function is performed as shown in Fig. 1, and an additional column is added for the egress port to which an incoming packet will be sent. This requires that a forwarding lookup for egress port be made prior to the ACL lookup. The forwarding lookup is a consultation of forwarding tables to determine which egress port is the proper port for sending a packet along toward its destination. A match for both the field values and the egress port now is required to return a pass/drop determination. This allows all functionality to be implemented at the ingress ports, and there is no loss of functionality at egress ports.

In a preferred embodiment there is some leeway in the order of operations. There are some good reasons for doing the ingress and egress pass/drop lookups separately, although they could be combined in some embodiments.

For example, because the rules are often independent, combining the rules in a single table requires multiplying them to generate all the combinations – which wastes a lot of table space and might make the rules not fit in the table.

In a preferred embodiment therefore, the ingress and egress lookups, both performed at the ingress ports, are performed separately. The ingress lookup in this case may be done either before or after the forwarding lookup, but the egress pass/drop lookup must be done after the forwarding lookup, because the egress port must be known.

In the case of the egress pass/drop lookup, a specific header match may be a pass for some egress ports, and a drop for others. In Fig. 2 it is assumed that there are three egress ports. The figure is therefore drawn showing a specific header set with three port indications, showing that a pass/drop entry (line) is implemented in a preferred embodiment for each field match of interest for each egress port.

- 9 -

With an ACL according to the present invention, all pass/drop functionality occurs at ingress, and the mechanisms at egress are no longer necessary, without any loss in functionality.

In a preferred embodiment of the present invention a mechanism is provided 5 for ACL lookups incorporating a content-addressable memory. A content-addressable memory is known in the art as a memory for which a lookup is done by using the content one wishes to match, in this case the field values for an incoming packet, rather than addresses.

Fig. 3 is a simplified diagram of a block 301 of content-addressable memory 10 (CAM) divided into two sections 303 and 305. The lines of memory are the header values and egress port numbers for which a pass or drop determination needs be made. In this example the memory is entered with the header values and port numbers, and pass or drop is a matter of in which section of memory, 303 or 305, the match is made. In alternative embodiments the CAM may be divided into more than two 15 sections allowing more than just pass or drop determination to be made.

The inventor is aware that there are a variety of options in implementation. One might, for example, implement the pass/drop lookups so that only pass (or drop) 20 combinations of header values are implemented in the lookup table, and the opposite is assumed if no match is found. There are also a variety of alterations that may be made in the embodiments of the invention thus far described without departing from the spirit and scope of the invention, so the invention should be accorded the scope of the claims which follow:

- 10 -

What is claimed is:

1. In a network packet router having one or more ingress and egress ports, a system implemented at an ingress port for egress pass/drop determination for packets,
5 comprising:

a rule set comprising header combinations and values for which a pass/drop decision may be applied, the set including an egress port identity for each header combination and value set; and

a mechanism noting pertinent headers of a specific incoming packet, and an
10 egress port to which the packet is to be sent, comparing the headers with rules in the rule set, and returning a determination of a rule to be applied.

2. The system of claim 1 wherein the rule set association takes the form of a lookup table.

15

3. The system of claim 1 wherein the rule returned determines pass or drop for the packet.

4. The system of claim 1 wherein the network is the Internet.

20

5. The system of claim 1 further comprising a second lookup table for accomplishing ingress rule determinations without egress port numbers.

25

6. The system of claim 1 wherein a content addressable memory (CAM) is used for header combination matches, and pass/drop is determined by the address location of any match.

7. The system of claim 1 comprising two sections in the CAM, and wherein pass or drop is determined by the section of the CAM wherein a content match is found.

30

- 11 -

8. The system of claim 6 wherein the CAM is divided into more than two sections, and a result or action is associated with each section.
9. In a network packet router, an ingress port, comprising:
 - 5 an interface for receiving packets;
 - a first mechanism for noting header combinations and values, and egress ports for transmission, for individual ones of the received packets; and
 - 10 a second mechanism comparing the headers and values with a rule set comprising header combinations and values and egress ports, and returning a rule for the packet.
10. The ingress port of claim 9 wherein the rule returned is a pass or drop determination for the packet.
- 15 11. The ingress port of claim 9 wherein the rule set association is in the form of a lookup table.
12. The ingress port of claim 9 wherein the network is the Internet.
- 20 13. The ingress port of claim 9 further comprising a second rule set association for accomplishing ingress rule determinations without egress port numbers.
14. The ingress port of claim 9 wherein a content addressable memory (CAM) is used for header combination matches, and an associated rule is determined by the address 25 location of any match.
15. The ingress port of claim 14 wherein the CAM is divided into two sections, and a pass or drop determination is determined by the section wherein the address lies for the content match.

- 12 -

16. The ingress port of claim 14 wherein the CAM is divided into more than two sections, and a result or action is associated with each section.

17. A network packet router comprising:

5 one or more ingress ports; and
 one or more egress ports;
 characterized in that individual ones of the ingress ports comprise a first
mechanism for noting header field combinations and values, and egress ports for
transmission, for individual ones of the received packets, and a second mechanism
10 comparing the headers with rules associated with the egress ports and field values, and
 returning a rule determination for the packet.

18. The router of claim 17 wherein the rule returned determines pass or drop for the
associated packet.

15
19. The router of claim 17 wherein the rule set association takes the form of a lookup
table.

20. The router of claim 17 wherein the network is the Internet.

20
21. The router of claim 17 further comprising a second rule set association for
accomplishing ingress pass/drop determinations without egress port numbers.

22. The router of claim 17 wherein a content addressable memory (CAM) is used for
25 header combination matches, and a rule is determined by the address location of any
match.

23. The packet router of claim 22 wherein the CAM is divided into two sections, and
a pass or drop determination is determined by the section wherein the address lies for
30 the content match.

- 13 -

24. The router of claim 22 wherein the CAM is divided into more than two sections, and a result or action is associated with each section.

5 25. In a network packet router having one or more ingress and egress ports, a method implemented at an ingress port for determining both ingress and egress rules for packets, comprising the steps of:

- (a) noting header combinations and values, and egress port destination for incoming packets;
- 10 (b) comparing the header combinations and values and egress port destinations with a rule set associated with the header combinations and values including egress port destinations; and
- (c) returning a rule for the packet.

15 26. The method of claim 25 wherein, in step (c), the rule returned indicates pass or drop for the associated packet.

27. The method of claim 25 wherein the rule set associations take the form of a lookup table.

20 28. The method of claim 25 wherein the network is the Internet.

29. The method of claim 25 further comprising a second lookup in a separate rule set association for accomplishing ingress rule application determinations without egress

25 port numbers.

30. The method of claim 25 wherein, in step (b), a content addressable memory (CAM) is consulted for rule set matches, and a rule is determined by the address location of any match.

- 14 -

31. The method of claim 30 wherein the CAM is divided into two sections, and pass or drop is determined by the section wherein the address is found for a content match.
32. The method of claim 30 wherein the CAM is divided into more than two sections, and each section is associated with a different result or action.

1/3

Fields

Pass/Drop

A	B	C	D	
e	f	g	h	0
e	f	j	h	1
		⋮	⋮	
m	f	*	h	0
e	*	g	k	1
1	m	n	o	1

* = wild card, any value will do

1 = Pass
0 = Drop

Fig. 1 (Prior Art)

2/3

Fields

Values

A	B	C	D	(1)	(2)
1	m	n	o	(1)	0
1	m	p	o	(2)	1
1	m	n	q	(3)	0
m	n	*	p	(1)	0
*	n	l	q	(2)	1
q	r	s	t	(3)	1

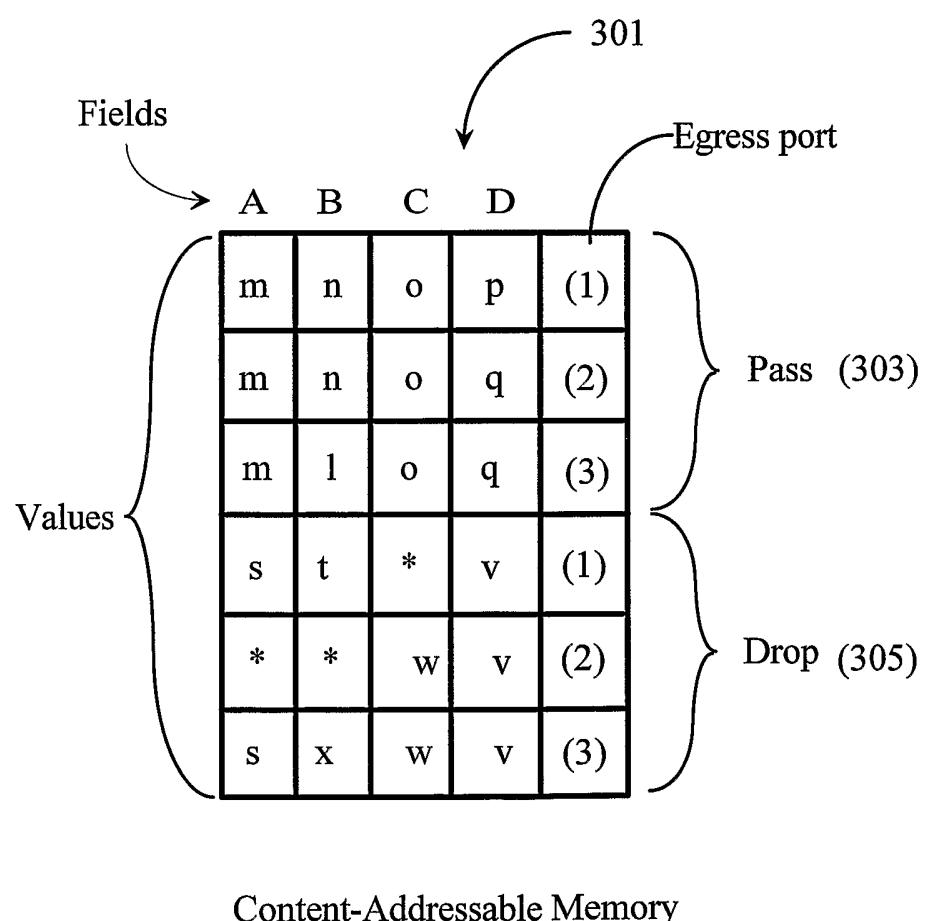
* = wild-card, any value will do

1 = Pass

0 = Drop

Fig. 2

3/3



* wild card - any value is acceptable

Fig. 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/26748

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04L 12/56; H04Q 11/04; G06F 19/00
US CL :709/238; 370/392, 389

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. : 709/238; 370/392, 389

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

NONE

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

EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,721,819 A (GALLES ET AL.) 24 February 1998, an abstract, cols. 1-2 and 26-27.	1-32
Y	US 5,996,021 A (CIVANLAR ET AL.) 30 November 1999, an abstract, cols. 5, 10-11.	1-32
Y	US 6,185,635 B1 (O'LOUGHLIN ET AL.) 06 February 2001, an abstract, cols. 5-7, 42-43.	1-32
Y	US 6,260,072 B1 (RODRIGUEZ-MORAL) 10 July 2001, an abstract, cols. 4-5.	1-32
Y, P	US 6,385,209 B1(SKIRMONT et al.) 07 May 2002, an abstract, Fig. 5, cols. 1-3.	1-32

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)	"&"	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 OCTOBER 2002

Date of mailing of the international search report

13 DEC 2002

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20251

Faxsimile No. (703) 305-3230

Authorized officer

KRISMA LIM

Telephone No. (703)-305-9672

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/26748

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	US 6,363,056 B1 (BEIGI et al.) 26 March 2002, an abstract, cols. 2-4.	1-32
A	US 6,055,561 A (FELDMAN et al.) 25 April 2000, an abstract, cols. 2-6.	1-32