

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 April 2008 (17.04.2008)

PCT

(10) International Publication Number
WO 2008/043374 A1

(51) International Patent Classification:

H04L 12/24 (2006.01) **H04L 12/56** (2006.01)
H04L 12/46 (2006.01)

(21) International Application Number:

PCT/EP2006/009730

(22) International Filing Date: 9 October 2006 (09.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): **TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)**
[SE/SE]; SE-164 83 Stockholm (SE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **FIORONE, Raoul** [IT/IT]; Viale Villa Gavotti 117/1, I-16155 Genova (IT). **MARTINOTTI, Riccardo** [IT/IT]; C. So Mazzini 10/4, I-17100 Savona (IT). **CORTI, Andrea** [IT/IT]; Via Degli Olenadri 10/2, I-17019 Varazze (sv) (IT).

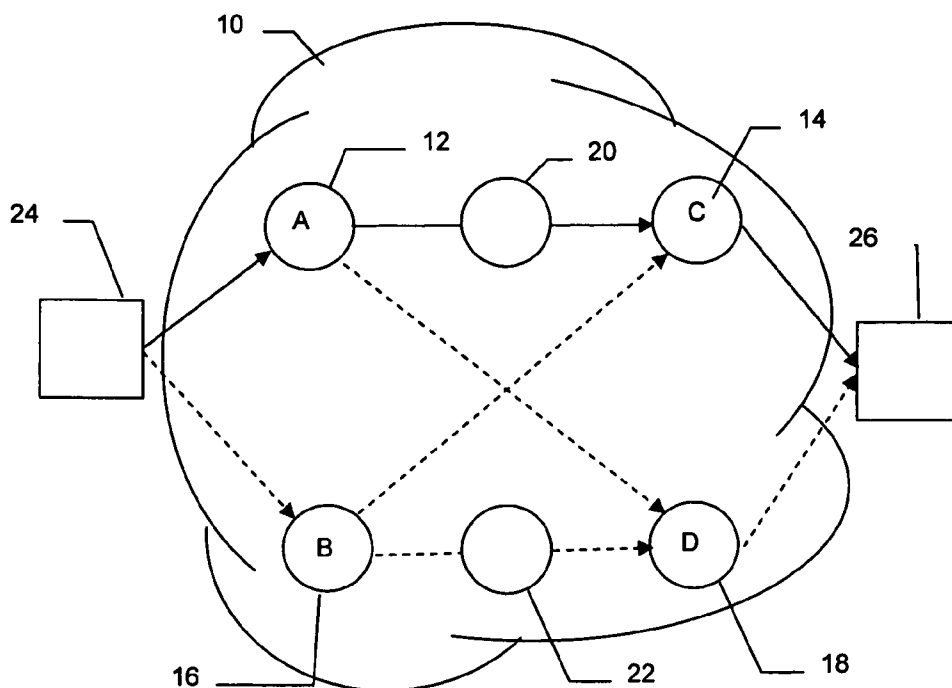
(74) Agent: **COLLIER, Ian, Terry**; Intellectual Property Group, Ericsson Limited, Maplewood, Crockford Lane, Chineham Business Park, Basingstoke, Hampshire RG24 8YB (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: RESILIENCY SCHEMES IN COMMUNICATIONS NETWORKS



(57) Abstract: A connection oriented communications network 10 comprises a plurality of nodes including a plurality of edge nodes 12, 14, 16, 18. The network is arranged to define a primary tunnel connecting a primary one 12, 14 of the edge nodes to another edge node 14, 12 and a secondary tunnel connecting a secondary one 16, 18 of the edge nodes to another edge node 12, 14, 16, 18. The network is arranged to enable switching of traffic from the primary tunnel to the secondary tunnel in the event of detection of a failure in the primary tunnel.

WO 2008/043374 A1



Published:

— *with international search report*

RESILIENCY SCHEMES IN COMMUNICATIONS NETWORKS

Field of the Invention

The present invention relates to network resiliency schemes in connection
5 oriented networks. It has particular application in multi-protocol label switched
(MPLS), and connection oriented Ethernet networks.

Background of the Invention

Network resiliency schemes are generally arranged to compensate for failures in
10 a network by detecting failure of one of the nodes, or inter-node connections, in the
network, and re-routing traffic to bypass the failure. Networks generally have a number
of edge nodes at which traffic can enter and leave the network, and a number of
intermediate nodes through which traffic can pass to travel from any one edge node to
any other. Customer equipment that is arranged to communicate over the network will
15 generally communicate with one or more edge nodes. In the simplest type of scheme,
any piece of customer equipment can only communicate with one edge node. Therefore
any re-routing carried out by the resiliency scheme cannot bypass the ingress node at
which traffic enters the network, or the egress node at which it leaves the network. In
some systems, dual parenting schemes are known in which the customer equipment can
20 communicate with more than one ingress or egress node. This has the advantage that, if
one of the edge nodes suffers a failure, then the customer equipment can still
communicate over the network. However, no such dual parenting scheme is currently
available for MPLS or other connection oriented networks. This means that resiliency
schemes for MPLS networks rely on path diversity between the same ingress and egress
25 nodes, and are therefore limited in the degree of resilience they can provide.

Summary of the Invention

Accordingly, the invention seeks to preferably mitigate, alleviate or eliminate
one or more of the disadvantages mentioned above singly or in any combination.

30

According to a first aspect of the present invention there is provided a
connection oriented communications network comprising a plurality of nodes including

a plurality of edge nodes, wherein the network is arranged to define a primary tunnel connecting together a primary pair of the edge nodes, and a secondary tunnel connecting together a secondary pair of the edge nodes, which is different from the primary pair, and wherein the network is arranged to enable switching of traffic from the primary tunnel to the secondary tunnel in the event of detection of a failure affecting the primary tunnel.

The first and second pair of nodes may have a common node, but each have at least one node that is not common to both pairs.

10

According to a second aspect of the present invention there is provided a node for a connection oriented network having a plurality of edge nodes, the node being arranged to identify a primary tunnel connecting together a primary pair of the edge nodes, and a secondary tunnel connecting together a secondary pair of the edge nodes, which is different from the primary pair, and wherein the node is arranged to enable switching of traffic from the primary tunnel to the secondary tunnel in the event of detection of a failure in the primary tunnel.

Further features of the present invention are as claimed in the dependent claims.

20

The present invention beneficially allows for network resiliency in an MPLS network, or other connection oriented network with node diversity.

Preferred embodiments of the present invention will now be described by way of example only with reference to the accompanying drawings.

25

Brief description of the drawings

FIG. 1 is a diagram illustrating an MPLS network according to one embodiment of the present invention,

30

FIG. 2 is a state diagram illustrating changes in the operation of the network of Figure 1,

FIG. 3 is a diagram illustrating an MPLS network according to a further
5 embodiment of the present invention; and

FIG. 4 is a diagram illustrating an MPLS network according to a further embodiment of the present invention.

10 Description of preferred embodiments of the invention

Referring to Figure 1, a multiprotocol label switching (MPLS) network 10 according to an embodiment of the invention comprises a number of edge nodes 12, 14, 16, 18 and a number of intermediate nodes 20, 22. Various devices, in this case referred to as items 24, 26 of customer equipment, are arranged to communicate over the network 10. Only
15 two customer equipment units 24, 26 are shown for simplicity although it will be appreciated that in practice the number will generally be much larger.

As is conventional each of the nodes 12, 14, 16, 18, 20, 22 includes one or more processors and associated memory and is arranged to perform a number of functions
20 relating to the operation of the network, and these functions are grouped, from an organisational point of view, into planes. The management plane provides overall management of the network and includes a network management system (NMS) that provides an interface for a user to provide manual control of the network. The control plane uses a number of protocols and allows for communication between the nodes and
25 configuration of the forwarding plane. The forwarding plane performs the forwarding operations that transmit data across the network. The control plane protocols, which can include, for example, OSPF-TE and RSVP-TE, are also arranged to provide communication between the edge nodes 12, 14, 16, 18 and optionally the customer equipment 24, 26. One of the main functions of the control plane protocols is to
30 determine the route that any communication will take across the network.

In order for the two customer equipment (CE) units 24, 26 to be able to communicate with each other, the control plane protocol of the network identifies one primary edge

node 12 for one of the customer equipment units 24, and one primary edge node 14 for the other 26. For communication from the first unit 24 to the second 26, the first primary edge node 12 is defined as the primary ingress node and the second primary edge node 14 is defined as the primary egress node. The control plane protocols also
5 define a number of communications tunnels each of which provides communication from one of the edge nodes 12, 14, 16, 18 to another. Each of the tunnels is a route between two nodes along which traffic can be transmitted in the network 10 such that the traffic, which is in the form of packets in an MPLS network, is opaque to intermediate nodes. In an MPLS network this can be achieved because the intermediate
10 nodes only read the label on each packet and not the data within the packet. The ingress node 12 that directs the traffic across the network identifies the egress node 14 to which it needs to send the traffic, and then attaches labels to the information which identify which tunnel it is to be transmitted through. Each tunnel is defined by a number of parameters, which include the two edge nodes that the tunnel connects. Other defining
15 parameters of the tunnels can include any one or more of the following, all of which are included in this embodiment: the characteristics of the traffic it can carry, the quality of service (QoS) scheme it operates, the service class associated to the MPLS tunnel, its maximum and mean bandwidth, its maximum burst size for example. It is therefore possible to have more than one tunnel between any pair of edge nodes, the two tunnels
20 having different characteristics or parameters. Each tunnel may include only a single label switch path (LSP), or it may include a plurality of LSPs the most appropriate of which will be selected at any given time depending on the characteristics and performances required by the client service. It will be appreciated that the customer traffic transmitted in the tunnel can be encapsulated into pseudo-wires arranged to
25 support Ethernet services as well as any type of service admitted in the network.

During communication over the network, a number of checks are carried out by the nodes as defined by the control plane protocols to identify failures in the network. Many methods of detecting failures are well known and will not be described in detail. A
30 management information base (MIB), which is related to and assists the control plane, is arranged to store data relating to what faults are detected in the network. Also all of the nodes of the network are arranged to communicate with each other to check for, and communicate information about, failures in the network, which will affect which of the tunnels is available for use at any time. This form of checking is carried out when the

CE unit 24 first transmits a communication for transmission across the network and will be described in more detail below. This allows selection of which of the tunnels is used depending on what faults are present.

- 5 For any type of communication between the CE unit 24 and the CE unit 26 a primary tunnel is defined which has one edge node 12 as its ingress node, and one edge node 14 as its egress node. In this embodiment the sending CE unit 24 is linked to, and can communicate with, a secondary ingress node 16, as well as the primary ingress node 12, and the receiving CE unit 26 is linked to, and can communicate with, a secondary egress node 18, as well as the primary egress node 14. The primary tunnel AC is therefore defined from the primary ingress node 12 to the primary egress node 14. Three secondary tunnels are also defined: one AD from the primary ingress node 12 to the secondary egress node 18, one BC from the secondary ingress node 16 to the primary egress node 14, and one BD from the secondary ingress node 16 to the secondary egress node 18. Therefore each of the tunnels is between a different pair of nodes, having one of its nodes in common with one other tunnel, another of its nodes in common with another tunnel, and no nodes in common with the remaining tunnel.
- 20 When the sending CE unit 24 wants to start communicating with the CE unit 26 it is arranged to direct the communication traffic to the primary ingress node 12. This traffic includes an identification of the customer equipment unit 26 that the communication is to be transmitted to, as well as information regarding the characteristics of the communication. On receiving this communication, the primary ingress node is arranged to start transmitting over the network to the primary egress node, and to trigger a number of communications between the edge nodes to check for any faults that might result in a need for a tunnel other than the primary tunnel to be used for the communication. It will be appreciated that these communications between the edge nodes can take a number of formats, but it is desirable that they are sufficient to identify which of the tunnels are available for use and which are not.
- 30

In this embodiment the primary ingress node 12, which is now the active ingress node, sends the traffic from the CE unit 24 to the primary egress node 14. The active egress node, on receiving the traffic, sends it on to the receiving CE unit 26 and monitors the

traffic. Suitable instruments, using either standard messages or a proprietary protocol, are used to check whether or not the LSP tunnel is successfully transporting the traffic from the active ingress node to the active egress node. This enables the active ingress node 12 to determine if there is a fault in the primary tunnel AC. Provided no fault is
5 detected on the primary tunnel, the traffic continues to be transmitted via the primary tunnel.

Also, when the communication begins, the primary and secondary edge nodes are arranged to exchange a series of messages so that each of them can determine which is
10 the current active tunnel, which are the secondary or standby tunnels, and whether there are any faults in the primary or any of the secondary tunnels. It will be appreciated that these messages can take a number of forms. In this embodiment, the state message to verify LSP integrity is also sent to the standby ingress node 16. Proper signalling messages are also exchanged between the primary ingress node 12 and the secondary
15 ingress node 16 to identify their respective status as ingress nodes. This enables the active ingress node 12 to check whether the secondary ingress node 16 has experienced a fault or not. The active ingress node 12 also sends a check signal to the secondary egress node 18 to enable it to identify its status as secondary egress node. The active ingress node 12 is enabled to check whether the standby egress node 18 is available for
20 use if necessary, by means of keep-alive signalling messages, which are periodically exchanged between the two nodes. The active ingress node 12 is also arranged to send an acknowledgement signal back to the sending CE unit 24 to confirm that it is receiving the traffic and sending it over the network to the receiving CE unit 26. All of these signals and checks can be repeated during the communication to check continually
25 on the status of each of the tunnels. These signals and checks can be exchanged among primary and standby edge nodes enabling them to check the status of the network.

For as long as the primary tunnel AC continues to function and provide the communication required, the communication will continue over it. However, if a fault is
30 detected in either of the primary edge nodes 12, 16, or indeed in any intermediate nodes 20 forming part of the primary tunnel AC that cannot be bypassed within the primary tunnel, then the network is arranged to switch to one of the standby tunnels so that communication can continue. The most appropriate tunnel will be selected depending on which of the nodes and LSPs are functioning correctly. Each of the nodes 12, 14, 16,

18, 20, 22 is arranged to monitor for and identify faults during the communication, and if any of them identifies a fault that requires, or might require, a change of active tunnel, then it is arranged to communicate this to each of the other edge nodes. In this way all of the edge nodes can maintain an accurate record of which is the currently active
5 tunnel, and whether there are any faults in any of the edge nodes 12, 14, 16, 18 or intermediate nodes 20, 22 that might affect the choice of active tunnel.

Referring to Figure 2, the status of the network can be described in terms of the active tunnel and how that will change depending on the detection of various faults. Where the
10 ingress nodes 12 and 16 are indicated as A and B and the egress nodes 14, 18 are indicated as C and D, the state of the network can be indicated by the pair of active nodes, such as AC if the primary tunnel is active. Each state therefore corresponds to a different active tunnel between a different pair of nodes. In Figure 2 the four states are indicated, the arrows between states indicate the possible transitions between states, and
15 the labels on each of the arrows indicate the node as a single letter, or the LSP between two nodes by the appropriate two letters, in which a fault will cause the transition to occur. When the communication starts, the system first moves to the state AC because AC is the primary or default tunnel. However, if a fault is then detected in egress node C, or anywhere in the LSP AC between nodes A and C, the network switches to state
20 AD, with the communication being sent to the standby egress node D. Similarly if a fault is detected in the ingress node A, then the network switches to state BD, with traffic being communicated via tunnel BD. It will be appreciated that there is a degree of flexibility as to which state the network moves to on detection of any particular fault. Apart from the need to avoid all nodes and LSPs with known faults, the choice may
25 further depend on other factors such as the condition of the traffic on the network as a whole.

If the active ingress node 12 receives a message indicating a fault affecting the primary tunnel, but is still functioning itself, then it can re-direct the traffic to the appropriate
30 secondary tunnel, and the new active egress node can transmit the traffic on to the receiving CE 26. Therefore neither of the CE units 24, 26 needs to take an active part in the process, except that the receiving CE obviously needs to be able to receive data from the secondary egress node. If a fault occurs which results in a need to switch to the secondary ingress node 16, the communication is transported in the secondary tunnel.

The CE unit 24 needs to be able to transmit data to the secondary ingress node 16 as well as to the primary ingress node 12. Optionally, external means can be used to communicate with the sending CE unit 24, instructing it to switch to sending the traffic to the secondary ingress node 16. This means can be a proprietary message set or
5 proprietary extensions to existing protocols (e.g. Spanning Tree), and is conceived to shorten the time, which is necessary for the CE unit to become aware of the necessity to transmit traffic to the secondary ingress node (without waiting for the learning process to take place). The CE unit 24 therefore needs to be arranged to respond to instructions from either of the ingress nodes 12, 16 to switch from directing traffic to one of them to
10 sending traffic to the other. However, in this embodiment, the sending CE unit 24 does not take an active part in determining which ingress node it should communicate with. It is however not required that the sending CE unit 24 is aware of any fault occurring in the MPLS network.

15 Referring to Figure 3 in a second embodiment of the invention, the basic setup is the same as in the first embodiment, and corresponding parts are indicated by the same reference numerals increased by 100. However in this case there is only one ingress node 124 which can communicate with either the primary 114 or secondary 118 egress node via respective primary and secondary tunnels 150, 152.

20

Referring to Figure 4, in a third embodiment, a service provider network 210 includes two MPLS networks 211, 213 run by different network operators. One CE unit 224 communicates with an edge node 212 of the first network 211, and a second CE unit 226 communicates with an edge node 256 of the second network 213. A first tunnel is
25 defined between these two edge nodes 212, 256 via primary edge nodes 214, 252 in the first and second networks. A second tunnel is defined via secondary edge nodes 218, 254 in the first and second networks. Here all of the edge nodes 212, 214, 218, 252, 254, 256 are arranged to communicate with one another, in a similar manner to those of the first embodiment of Figure 1, to identify any faults affecting the primary tunnel
30 which would require transfer of the traffic to the secondary tunnel, and to keep each other updated regarding the status of each of the nodes and each of the tunnels as in the previous embodiments.

CLAIMS

1. A connection oriented communications network comprising a plurality of nodes including a plurality of edge nodes, wherein the network is arranged to define a primary
5 tunnel connecting together a primary pair of the edge nodes, and a secondary tunnel connecting together a secondary pair of the edge nodes, which is different from the primary pair, and wherein the network is arranged to enable switching of traffic from the primary tunnel to the secondary tunnel in the event of detection of a failure affecting the primary tunnel.
- 10 2. A network according to claim 1 wherein the primary and secondary tunnels have different egress nodes.
3. A network according to claim 1 or claim 2 wherein the primary and secondary
15 tunnels have different ingress nodes.
4. A network according to any foregoing claim wherein at least two of the edge nodes are arranged to communicate with each other by means of failure identifying communications to identify a relevant failure requiring a switch to the secondary tunnel.
- 20 5. A network according to claim 4 wherein the failure identifying communications can be between the ingress node of the primary tunnel and the egress node of the primary tunnel.
- 25 6. A network according to claim 4 or claim 5 wherein the failure identifying communications can be between a node of the primary tunnel and a node of the secondary tunnel.
7. A network according to any of claims 4 to 6 wherein the failure identifying
30 communications can be between the ingress node of the primary tunnel and the ingress node of the secondary tunnel.

8. A network according to any of claims 4 to 7 wherein the failure identifying communications can be between the egress node of the primary tunnel and the egress node of the secondary tunnel.

5 9. A network according to any foregoing claim including a plurality of ingress nodes and a plurality of egress nodes, and wherein a tunnel is defined between each of the ingress nodes and each of the egress nodes.

10 10. A node for a connection oriented network having a plurality of edge nodes, the node being arranged to identify a primary tunnel connecting together a primary pair of the edge nodes, and a secondary tunnel connecting together a secondary pair of the edge nodes, which is different from the primary pair, and wherein the node is arranged to enable switching of traffic from the primary tunnel to the secondary tunnel in the event of detection of a failure in the primary tunnel.

15

11. A node according to claim 10 which is an edge node arranged to communicate with other edge nodes in the network to identify a failure in the network requiring switching of the traffic to the secondary tunnel.

Fig. 1

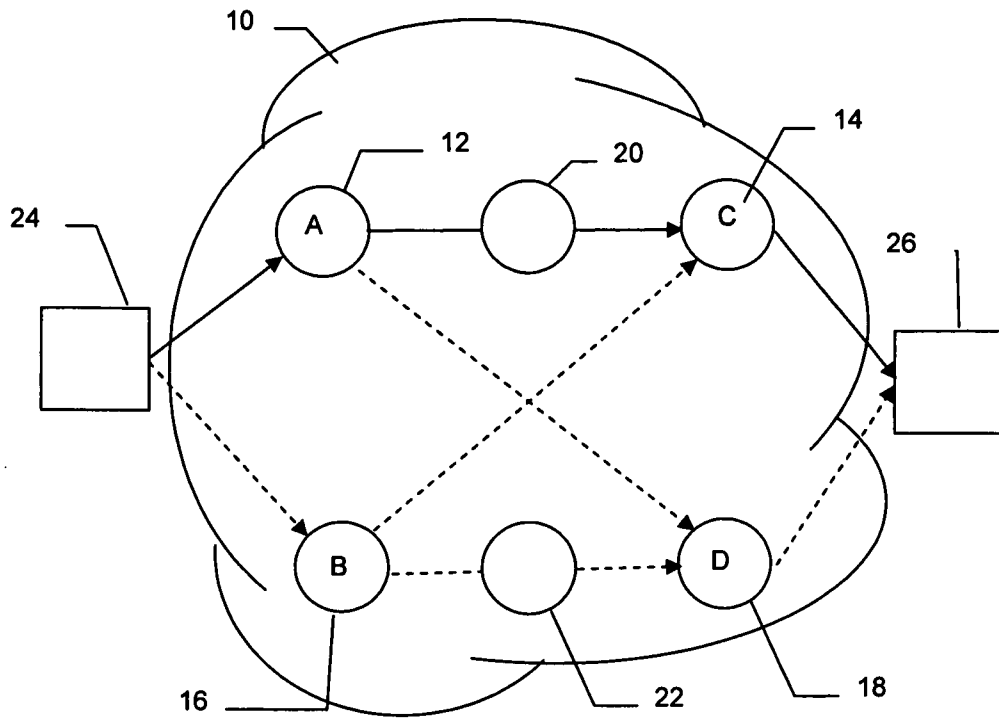


Fig. 2

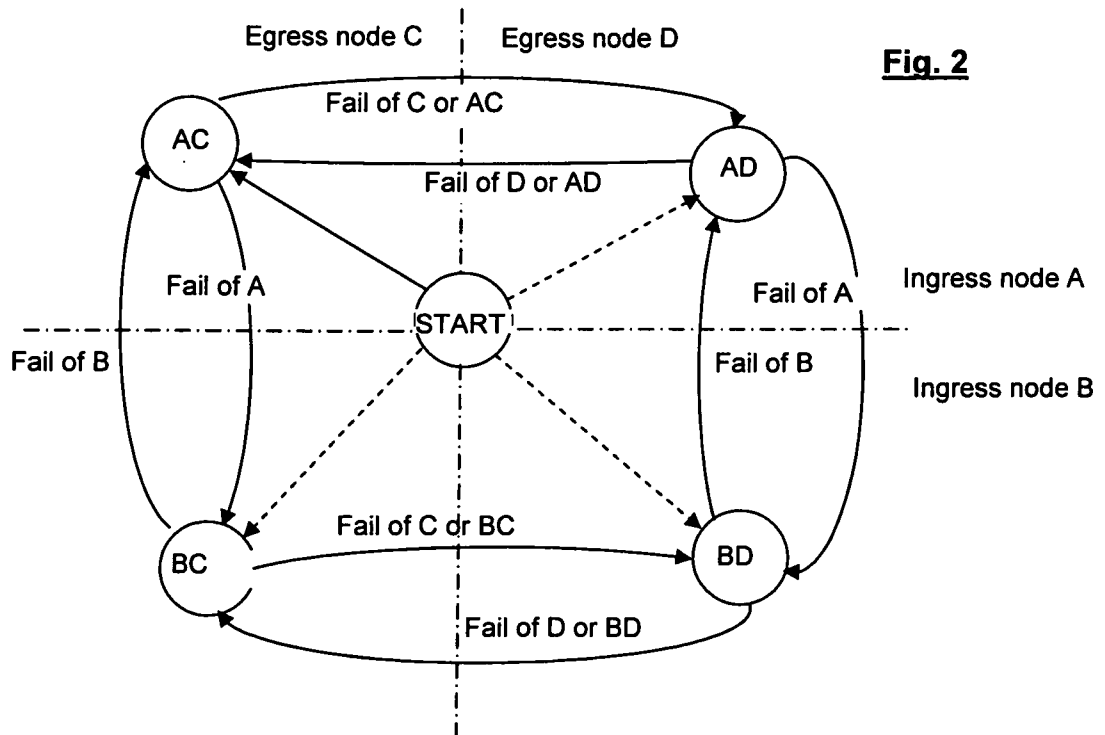


Fig. 3

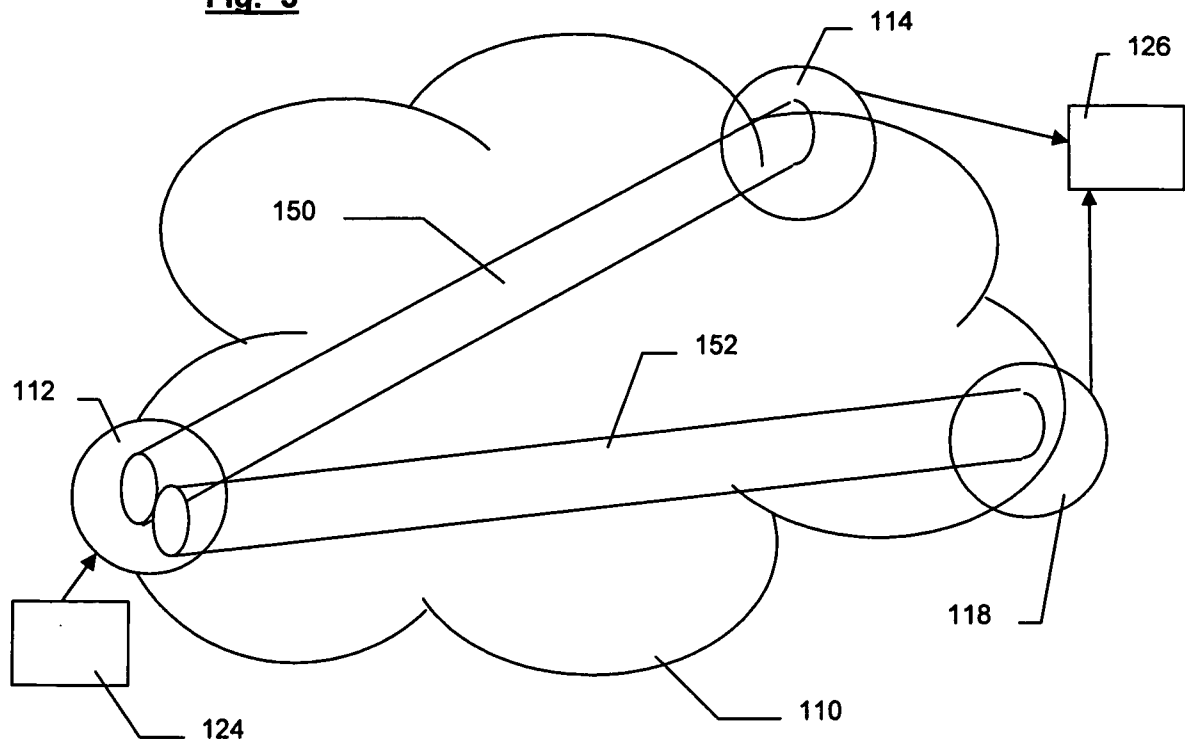
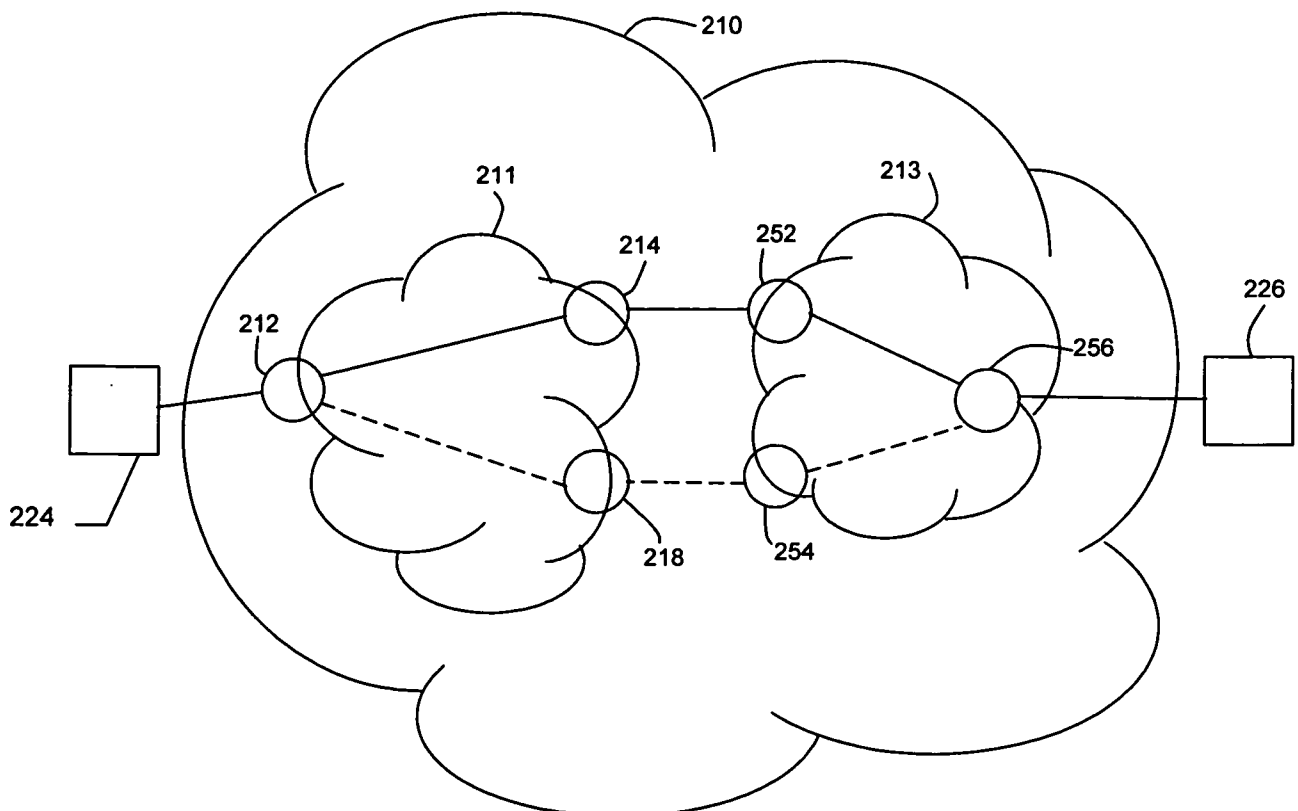


Fig. 4



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/009730

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04L12/24 H04L12/46 H04L12/56

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)
H04L

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 practical, search terms used)

EPO-Internal, WPI Data, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHARMA V ET AL: "Framework for Multi-Protocol Label Switching (MPLS)-based Recovery" IETF STANDARD, INTERNET ENGINEERING TASK FORCE, IETF, CH, February 2003 (2003-02), XP015009252 ISSN: 0000-0003 abstract page 1, line 1 - page 6, line 20 page 14, line 1 - page 18, line 20 page 21, line 20 - page 24, line 15 page 27, line 20 - page 28, line 10 ----- -/--	1,2,4-6, 9-11
Y		8

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *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.
- *G* document member of the same patent family

Date of the actual completion of the international search

15 August 2007

Date of mailing of the international search report

24/08/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

AURA MARCOS, F

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2006/009730

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2006/221813 A1 (SCUDDER JOHN G [US] ET AL) 5 October 2006 (2006-10-05) abstract paragraph [0030] - paragraph [0049] paragraph [0063] - paragraph [0065] figures 1-3,8	8
A	----- "RSVP-TE Extensions in support of End-to-End GMPLS-based Recovery" INTERNET DRAFT:CCAMP WORKING GROUP, [Online] February 2004 (2004-02), XP002446707 Retrieved from the Internet: URL: http://tools.ietf.org/html/draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03 [retrieved on 2007-06-07] abstract page 9, line 1 - line 15 page 16, line 20 - page 21, line 25	1-11
A	----- P. PAPADIMITRIOU: "Re: Shared mesh restoration-e2e recovery signaling draft"[Online] 26 July 2006 (2006-07-26), XP002446708 Newsgroup (http://psg.com) Retrieved from the Internet: URL: http://psg.com/lists/ccamp/ccamp.2006/msg00574.html [retrieved on 2007-06-07] the whole document	1-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2006/009730

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2006221813 A1	05-10-2006	WO 2006107875 A2	12-10-2006