

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 February 2009 (12.02.2009)

PCT

(10) International Publication Number
WO 2009/019257 A2

(51) International Patent Classification:
H04L 12/437 (2006.01) *H04L 12/24* (2006.01)
H04L 12/46 (2006.01)

(21) International Application Number:
PCT/EP2008/060245

(22) International Filing Date: 4 August 2008 (04.08.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07015526.2 7 August 2007 (07.08.2007) EP

(71) Applicant (for all designated States except US): **NOKIA SIEMENS NETWORKS OY** [FI/FI]; Karaportti 3, FIN-02610 Espoo (FI).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **RAMALHO RIBEIRO DOS SANTOS, José Miguel** [PT/PT]; Rua José Cardoso Pires n°1, 8°B., P-1750-430 Lisboa (PT).

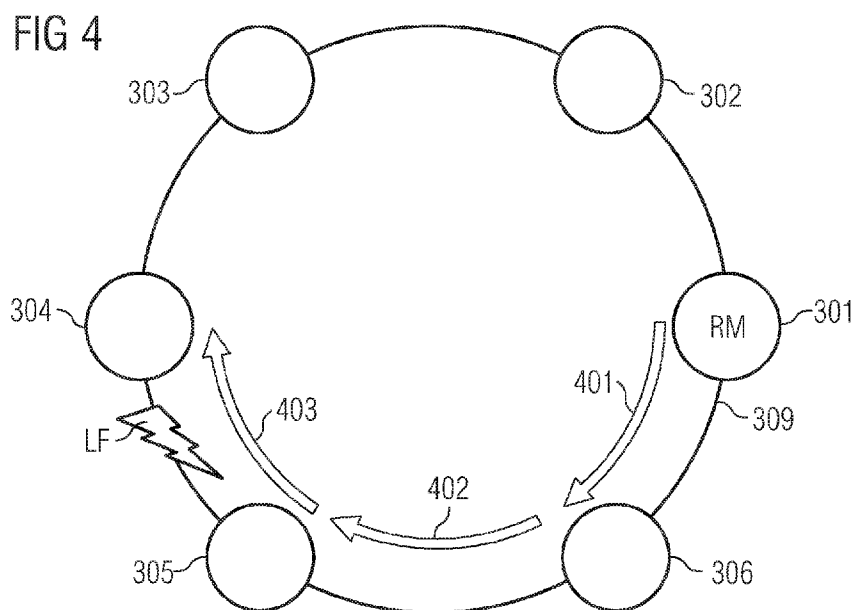
(74) Common Representative: **NOKIA SIEMENS NETWORKS OY**; COO RTP IPR / Patent Administration, 80240 Munich (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, 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, ST, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

(54) Title: METHOD TO BE RUN IN AND DEVICE OF A NETWORK AS WELL AS COMMUNICATION SYSTEM COMPRISING SUCH DEVICE



(57) Abstract: A method and a device are provided to be run in a network (or in particular on a network component of such network). The network comprises several network elements that are connected via a ring, wherein one network element is a ring master comprising a primary port and a secondary port. The approach comprises the steps of (i) a failure along one direction of the ring is detected by the ring master; and (ii) the ring master sends a first message via its port that indicates the direction of the failure.

WO 2009/019257 A2

Description

Method to be run in and device of a network as well as
communication system comprising such device

5

The invention relates to a method to be run in and a device of a network as well as to a communication system comprising such device.

10 An Ethernet Ring Protection (ERP) mechanism and protocol are disclosed in, e.g., EP 1 062 787 B1. In addition, there exists another ring protection mechanism called Ethernet Automatic Protection Switching (EAPS) as described in, e.g., IETF RRC3619.

15

Such ring protection mechanisms comprise a ring master RM (also referred to as a redundancy manager) to coordinate ring protection activities.

20 Protection in this sense means in particular that a link-layer loop in a physical Ethernet is avoided. The ring master is equipped to prevent the ring from forming such Ethernet loops.

25 When the ring master is notified that the ring is healthy (e.g., via test packets that are sent by the ring master via both of its ports), i.e. all ring nodes (network elements) and links (segments or arcs) are operational, the ring master breaks the link-layer loop by blocking traffic reception and
30 transmission at one of its ring ports (the ring master's secondary port).

All traffic is blocked at that secondary port except for Ethernet ring protection control traffic, e.g., test packets.

35 Preferably, such control traffic is sent via a control virtual LAN (VLAN).

From a link-layer's perspective, blocking traffic at the ring master's secondary port transforms the ring's topology into a chain of nodes (network elements). This is necessary in typical layer 2 (L2) networks (see also document IEEE 802.1
5 for further explanation). The ring master blocking its secondary port resulting in a topology of a chain of network elements is considered a normal operational state of the Ethernet Ring Protection mechanism.

10 **Fig.1** shows such an ERP structure. The ring comprises network elements or nodes 101 to 106, wherein the node 101 is a Ring Master RM (also referred to as redundancy manager) with a primary port P and a secondary port S. As stated before, in normal operation, the Ring Master blocks its secondary port S
15 resulting in the nodes 101 to 106 building a chain topology for the user traffic.

Link or Port Failure:

20 When a failure emerges in the ring, e.g., a link failure of a ring segment, the Ring Master unblocks its secondary port S thereby reestablishing communication between all ring nodes.

The failure can be directly detected by the Ring Master
25 itself if the failure occurs at one of its ports.
Alternatively, the Ring Master can be notified by another network element of the ring about a failure detected at one of the network element's ports. In such case, the affected network element sends a Link Down message to the Ring Master.
30 The Ring Master subsequently unblocks its secondary port S (see **Fig.2**).

Failure Recovery:

35 When a network element of the ring detects that a failure recovered, it sends a notification to the Ring Master indicating that the link or port is operative again. This can be achieved by the network element sending a Link Up message

to the Ring Master. The network element will switch over to a pre-forwarding state blocking all traffic except test packets (health-check messages conveyed via the VLAN). In this pre-forwarding state the network element waits for a message from
5 the Ring Master to switch over to normal operation (or forwarding state) again.

The Ring Master blocks the secondary port S again and sends the message to the network element to get back to normal
10 operation. The Ring Master allows the network element to migrate from its pre-forwarding state to normal operation (forwarding state) only after the Ring Master blocked its secondary port S. This avoids configuration of a link-layer loop.

15 Preferably, the Ring Master assesses the operational state of the whole ring by frequently sending health-check packets via both of its ring interfaces, i.e. via its primary port P and its secondary port S. These health-check packets (also
20 referred to as test packets) may be conveyed via a control VLAN. If the ring is operational, the Ring Master receives its test packets sent via the respective other interface. If the test packets are not received, the ring may be broken and protection recovery actions should be initiated.

25 It is disadvantageous that according to a non fatal error test packets may be lost in just one direction. This leads to a partially defective traffic, i.e. a portion of the traffic may get lost whereas another portion may still arrive at its
30 destination.

The **object** to be solved is to avoid the disadvantages stated before and to provide an approach that handles a link that is
35 defective in only one direction, but operative in the other direction.

This problem is solved according to the features of the independent claims. Further embodiments result from the depending claims.

- 5 In order to overcome this problem a method is provided to be run in a network (or in particular on a network component of such network). The network comprises several network elements that are connected via a ring, wherein one network element is a ring master comprising a primary port and a secondary port.
- 10 The method comprises the steps:
- a failure along one direction of the ring is detected by the ring master;
 - the ring master sends a first message via its port that indicates the direction of the failure.

15 In particular, the ring master allows the failure in the ring to be detected, said failure happens only in one direction along the path, but not in the other direction. The ring master acts on such detection by sending the first message

20 via the port (its primary port or secondary port) that points towards the failure or indicates the direction of the failure.

In an embodiment, the failure is detected by the ring master

25 via an at least one second message that is not arriving at (at least) one of its port.

Due to the fact that only one second message does not reach the port, but the other does reach the opposite port, the

30 ring master determines that the communication path is corrupt in one direction only.

In another embodiment, the at least one second message is a test message and/or a health check message. Such at least one

35 second message may be conveyed as a data packet.

In a further embodiment, the at least one second message is sent by the ring master. Also, the at least one second

message may be sent by the ring master via its primary port and via its secondary port.

Advantageously, such at least one second message sent via one
5 of the ring master's ports travels through the ring network and arrives at the ring master's opposite port. Receiving, e.g., the at least one second message at the ring master's secondary port, the ring master may recognize that it has been generated by itself and sent via its primary port.
10 Hence, the communication path from the primary port to the secondary port (through the ring network) is operative. The same mechanism may be applied in the opposite direction from the secondary port to the primary port thereby indicating whether or not this particular direction through the ring
15 network is defective.

In yet another embodiment, the at least one second message is sent via a control virtual local area network (control VLAN).

20 Such control VLAN may be operative in both directions through the ring and is advantageously administered by the ring master. This, however, is different from user traffic that - in normal operation, i.e. without a failure in the ring network - may not be conveyed via the secondary port of the
25 ring master (which, as stated above, is blocked in normal operation).

In a next embodiment, the first message is sent by the ring master via the port that did not receive the at least one
30 second message.

Hence, in the case of an unidirectional failure within the ring network, the at least one second message may not arrive at one of the ring master's ports. This port is chosen for
35 dispatching the first message into the ring network.

The first message may in particular be a message that advises a recipient to block an interface upon certain conditions.

In an embodiment, the method comprises the following steps:

- a network element receiving the first message forwards the first message if the at least one second message does not arrive at both of its ports;
- a network element receiving the first message enters a predefined state if the at least one second message arrives at both of its ports.

Hence, receiving the first message, the network element determines whether or not the at least one second message (e.g., test message) arrives at both of its ports. If this is the case, this particular network element is not the last one that received the at least one second message at both of its ports. Therefore, the first message is forwarded to the next network element along the direction as specified by the ring master.

Thus, one network element after the other runs this mechanism and forwards the first message along the given direction (from the ring master via its port that did not receive the at least one second message, e.g., test packet) until the network element is found that received the second message at both of its ports. This network element hence is the last network element that received the at least one second message at both sides (ports). The link failure is between this network element identified and the one that at last dispatched the first message.

Accordingly, this network element enters the predefined state.

It is an embodiment that in the predefined state the network element blocks the port at which the first message has been received.

Preferably, the predefined state may be a pre-forwarding state.

In yet another embodiment, the network element that enters the predefined state sends a third message to the ring master. This third message may be an acknowledge message to the first message.

In an embodiment, the third message is sent via the path without the failure.

Furthermore, upon receipt of the third message, the ring master may unblock its secondary port.

Hence, the ring master becomes aware that according to its first message sent a network element entered a pre-forwarding state and sends via the third message an acknowledge to this first message. Pursuant to receiving the third message the ring master knows that a port of a network element has been blocked and thus the ring master may unblock its secondary port to keep up the traffic flow within the ring network.

In another embodiment, the method comprises the following steps:

- the ring master detects the at least second message at its secondary port and at its primary port;
- the ring master blocks its secondary port;
- the ring master sends a forth message to the network element that has entered the predefined state.

This allows the ring master to switch to normal operation mode, i.e. to block its secondary port, once the ring master becomes aware of the at least one second messages arriving (again) at both of its ports. Hence, the one directional link failure is corrected, the ring master gets back to normal operation and sends the forth message toward the network element that entered the predefined (e.g., pre-forwarding) state.

As another embodiment, the network element that is in the predefined state unblocks its port upon reception of the forth message. In particular, the network element that is in the predefined state may enter a forwarding state.

5

In yet a further embodiment, the network element that switched back to the forwarding state (subsequently) informs the ring master of its state change by sending a respective message to the ring master.

10

Thus, the ring master may already have blocked its secondary port and informed the network element in the predefined state to switch to normal operation, i.e. to unblock its previously blocked port and switch to the forwarding state. After the affected network element acted accordingly, the ring master is informed that this network element is back to normal operation. Hence, the ring network is fully operative in normal state (again) as it was prior to any failure detected by the ring master.

20

The problem as stated supra is also solved by a device comprising a processor unit that is arranged and/or equipped such that the method as described herein can be run and/or executed on said processor.

25

The device may in particular be a communication device, e.g., a network element or a ring master.

The problem as stated supra is further solved by a communication system comprising the device as described herein.

30

Embodiments of the invention are shown and illustrated in the following figures:

35

Fig.3 shows an ring network with an unidirectional link failure that is detected by a ring master;

Fig.4 shows the ring network of Fig.3, wherein "block interface"-messages initiated by the ring master are sent towards the network element of the ring network that is adjacent to the unidirectional link failure;

5

Fig.5 shows the ring network of Fig.4, wherein the network element that entered a pre-forwarding state sends a "block interface acknowledge"-message to the ring master via the healthy path;

10

Fig.6 shows the ring network of Fig.5, wherein the link failure is corrected and the test messages are no longer interrupted;

15

Fig.7 shows the ring network of Fig.6, wherein due to the healthy ring the ring master blocks its secondary port and sends "unblock interface"-messages to the network element that still is in the pre-forwarding state;

20

Fig.8 shows the ring network of Fig.7, wherein the network element that has been in the pre-forwarding state switches to a forwarding state and sends "unblock interface acknowledge"-messages to the ring master;

25

Fig.9 shows a float diagram comprising steps to be run and/or executed on a ring master of a ring network.

The approach described herein preferably is based upon a link failure in one direction only, i.e. a communication channel can transfer data from A to B, but not vice versa.

The example described hereinafter uses in particular the following mapping to the messages as described supra:

35

- a) The first message as described corresponds to a block interface message (BI) sent by a ring master RM

indicating that a network element has to block a particular port.

- 5 b) The at least one second message as described corresponds to a test message (or health check message). Test messages are preferably sent by the ring master RM via its primary port and via its secondary port. The test messages are (in normal operation without a failure in the ring network) received by the respective other port of the ring master thereby informing the ring master that the ring is operational (in both directions). The test messages are preferably sent via a control VLAN of the ring network.
- 10
- 15 c) The third message is an acknowledge message to the first message, preferably a block interface acknowledge message (BI_ack).
- 20 d) The forth message is a message sent by the ring master telling a particular network element of the ring network to unblock its port and switch to normal operation (i.e. a forwarding state). This forth message is further referred to as unblock interface message (UI).
- 25 e) In addition, there is a further message sent by the network element that received the unblock interface message and entered the forwarding state thereby sending an unblock interface acknowledge message (UI_ack) to the ring master.

30

The block interface message, the block interface acknowledge message, the unblock interface message and the unblock interface acknowledge message are preferably conveyed via an Ethernet slow protocol.

35

Fig.3 to Fig.8 show an example as how the approach presented herein may operate. These figures show a ring network comprising network elements (or nodes) 301 to 306, wherein

the network element 301 is the ring master RM (also referred to as a redundancy manager). As far as ring protection is concerned, the ring network may run an Ethernet Ring Protection (ERP) protocol or an Ethernet Automatic Protection Switching (EAPS) protocol.

In **Fig.3**, a link failure LF occurs at a connection between the network element 304 and the network element 305, wherein the failure only affects one direction of the connection, i.e. traffic may pass from the network element 305 to the network element 304, but not vice versa. The ring master 301 on a regular basis dispatches test messages via both of its ports 309 and 310.

Due to the link failure LF, test messages 307 sent via port 309 are interrupted, but test messages 308 sent via port 310 get back to port 309 of the ring master 301 thereby indicating that there is a unidirectional link failure somewhere within the ring network.

Hence, as shown in **Fig.4**, the ring master 301 sends the block interface message 401 via its port 310 that did not receive the test message. The block interface message 401 is received at the network element 306.

Next, the network element 306 evaluates whether test messages have been received at both of its ports. According to this example and as can be seen in Fig.3, the network element 306 only received test messages 308 at one of its ports, the test messages 307 cannot reach this network element 306 due to the unidirectional link failure LF.

As test messages did not arrive at both ports of network element 306, this network element 306 forwards the block interface message 402 to the network element 305.

With regard to the network element 305 only one of its ports received test messages (i.e. test messages 308 sent via port

310 of the ring master 301). Thus, the network element 305 forwards the block interface message 403 to the network element 304.

5 As the link failure LF is unidirectional from the network element 304 to the network element 305 only, the block interface message 403 arrives at the network element 304.

10 The network element 304 is the last one in the test message path 307 that received test messages at both of its ports. Hence, the network element 304 enters a pre-forwarding state and blocks (see reference 502 in **Fig.5**) its port pointing towards the link failure, the direction receiving the block interface message. It is to be noted that the blocked
15 interface 502 still may receive test messages via, e.g., a control VLAN.

In addition, the network element 304 sends a block interface acknowledge message 501 to the ring master 301 via the path
20 without link failure.

Upon reception of the block interface acknowledge message 501, the ring master 301 unblocks its secondary port thereby restoring connectivity of the ring network.

25 **Fig.6** shows that the link failure LF has been corrected, test messages 601 and 602 arrive at both ports of the ring master 301. The ring master 301 blocks its secondary port to avoid a loop within the ring network.

30 However, due to the link failure LF, the network element 304 still is in the pre-forwarding state and the its port 502 still is blocked.

35 As shown in **Fig.7**, the ring master 301 sends - via both of its ports 309 and 310 - unblock interface messages 701 and 702 to the network element 304.

Upon receipt of the unblock interface messages 701 and 702 at both of its ports, the network element 304 unblocks its interface 502 and switches from the pre-forwarding state to a forwarding state. In addition, the network element 304 sends
5 unblocking interface acknowledge messages 801 and 802 via both of its ports to the ring master 301 (as shown in **Fig.8**).

It is to be noted that each message may be sent more than once, in particular in order to avoid any errors due to lost
10 data packets.

It is further to be noted that the link failure may happen on a shared link. In such case, just one ring master may detect the failure and may provide the measures as described herein.
15

It is an advantage of this approach that unidirectional link failures within a network may be detected and correctly removed from layer 2 paths. Temporary loops within the ring network are avoided.
20

Fig.9 shows a flow diagram comprising steps to be run and/or executed on a ring master of a ring network. In a step 901 the ring master detects an unidirectional link failure and sends in a step 902 a block interface message via its port
25 that did not receive test messages due to the unidirectional link failure. Upon reception of a block interface acknowledge message the ring master in a step 903 unblocks its secondary port to keep up traffic flow throughout the ring network. Upon detection of test packets arriving at both of its ports
30 (again), the ring master in a step 904 blocks the secondary port and (in a step 905) sends unblock interface messages to the network element that is in a pre-forwarding state.

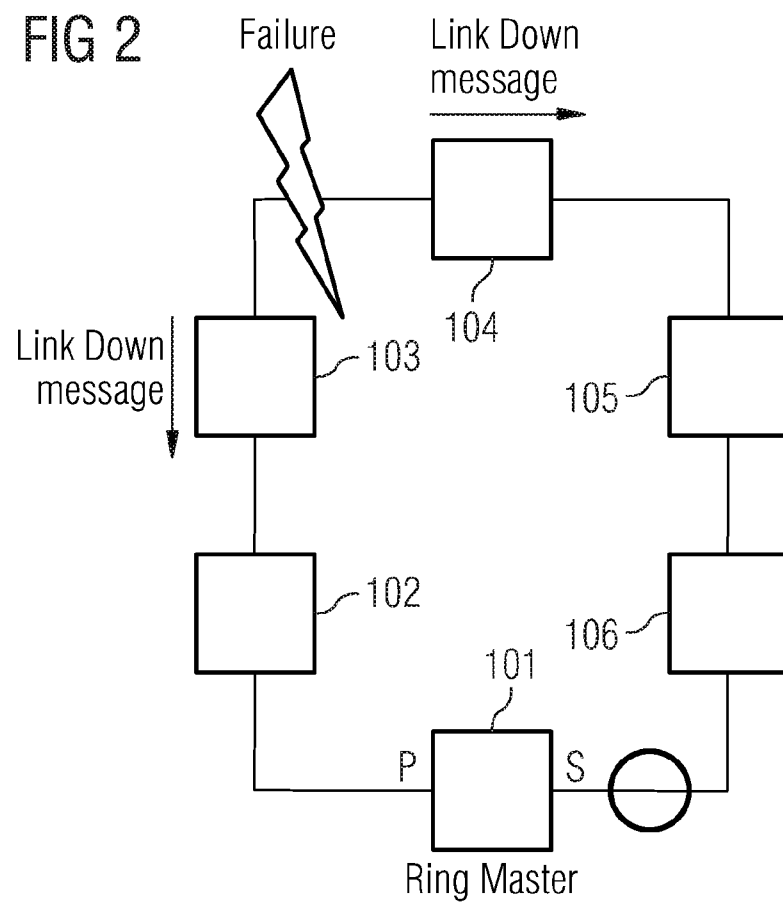
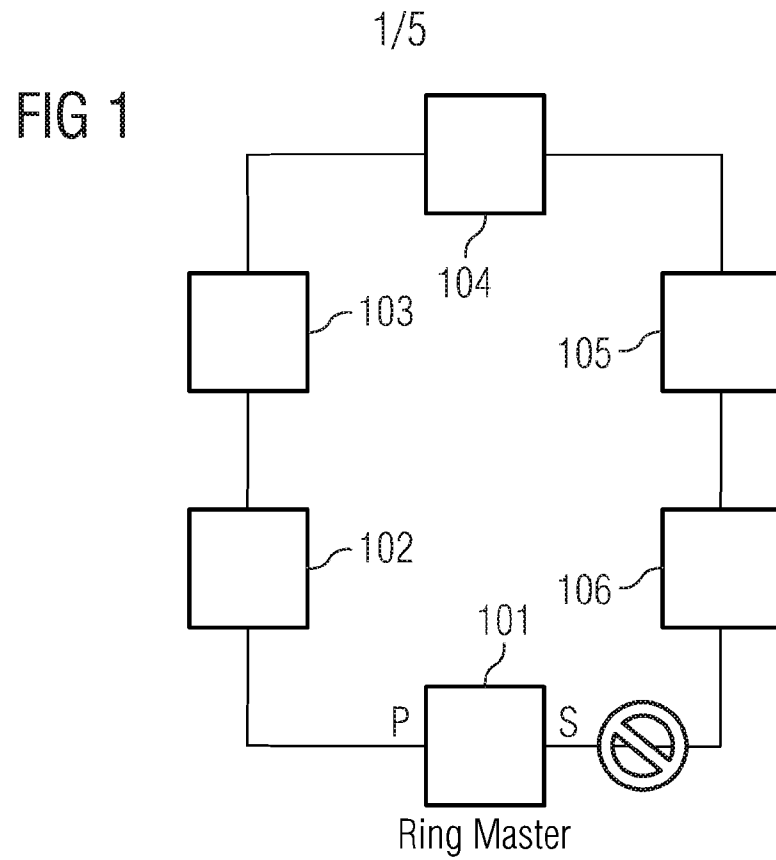
Claims:

1. A method to be run in a network
 - wherein the network comprises several network
 - 5 elements that are connected via a ring;
 - wherein one network element of the network elements is a ring master comprising a primary port and a secondary port;
 - comprising the steps:
 - 10 – a failure along one direction of the ring is detected by the ring master;
 - the ring master sends a first message via its port that indicates the direction of the failure.
- 15 2. The method according to claim 1, wherein the failure is detected by the ring master via an at least one second message not arriving at one of its ports.
3. The method according to claim 2, wherein the at least
- 20 one second message is a test message and/or a health check message.
4. The method according to any of claims 2 or 3, wherein the at least one second message is sent by the ring
- 25 master.
5. The method according to any of claims 2 to 4, wherein the at least one second message is sent by the ring master via its primary port and via its secondary port.
- 30 6. The method according to any of claims 2 to 5, wherein the at least one second message is sent via a control virtual local area network (VLAN).
- 35 7. The method according to any of claims 2 to 6, wherein the first message is sent by the ring master via the port that did not receive the at least one second message.

8. The method according to any of claims 2 to 7 comprising the steps:
- a network element receiving the first message
 - 5 forwards the first message if the at least one second message does not arrive at both of its ports;
 - a network element receiving the first message enters a predefined state if the at least one second message arrives at both of its ports.
- 10
9. The method according to claim 8, wherein in the predefined state the network element blocks the port that received the first message.
- 15
10. The method according to any of claims 8 or 9, wherein the predefined state is a pre-forwarding state.
11. The method according to any of claims 8 to 10, wherein the network element that enters the predefined state
- 20 sends a third message to the ring master.
12. The method according to claim 11, wherein the third message is an acknowledge message to the first message.
- 25
13. The method according to any of claims 11 or 12, wherein the third message is sent via the path without the failure.
14. The method according to any of claims 11 to 13, wherein
- 30 upon receipt of the third message the ring master unblocks its secondary port.
15. The method according to any of claims 2 to 14 comprising the following steps:
- 35 – the ring master detects the at least second message at its secondary port and at its primary port;
 - the ring master blocks its secondary port;

- the ring master sends a forth message to the network element that has entered the predefined state.

- 5 16. The method according to claim 15, wherein the network element that is in the predefined state upon reception of the forth message unblocks its port.
- 10 17. The method according to claim 16, wherein the network element enters a forwarding state.
- 15 18. The method according to claim 17, wherein the network element informs the ring master that it entered the forwarding state.
- 20 19. A device comprising a processor unit that is arranged such that the method according of any of the preceding claims is executable on said processor.
- 25 20. The device according to claim 19, wherein said device is a communication device, in particular a network element or a ring master.
21. Communication system comprising the device according to any of claims 19 to 20.



2/5

FIG 3

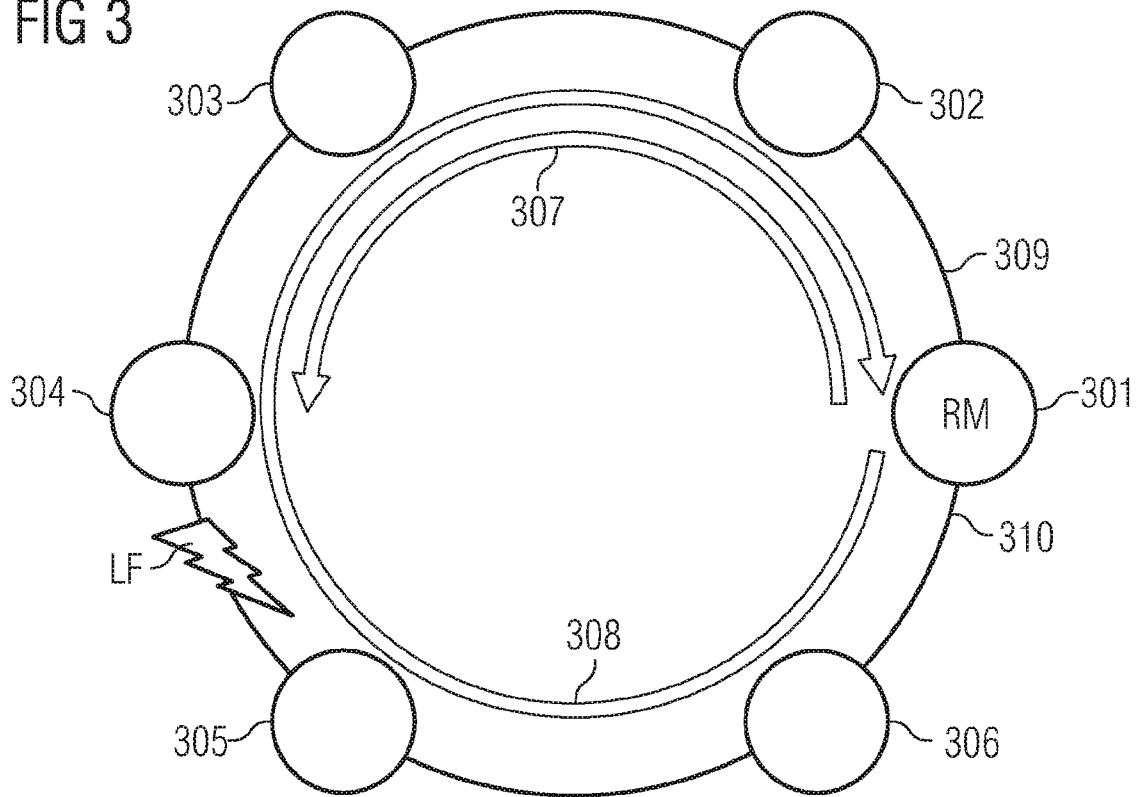
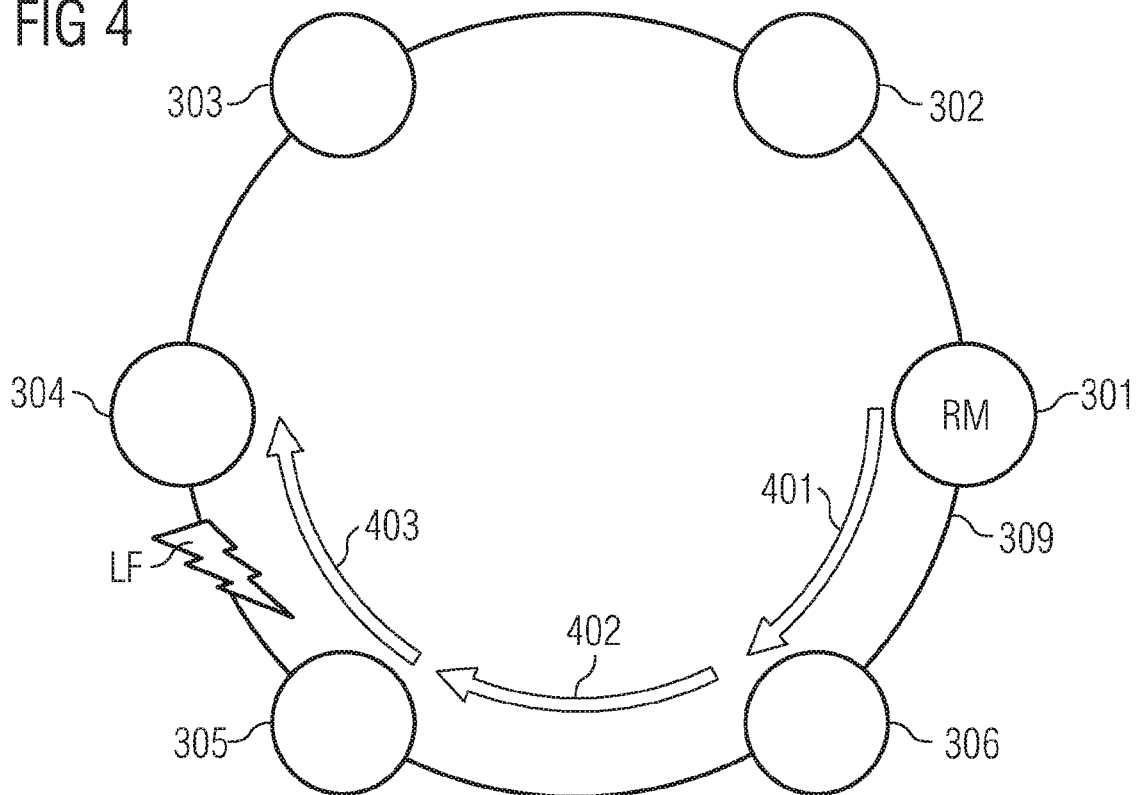


FIG 4



3/5

FIG 5

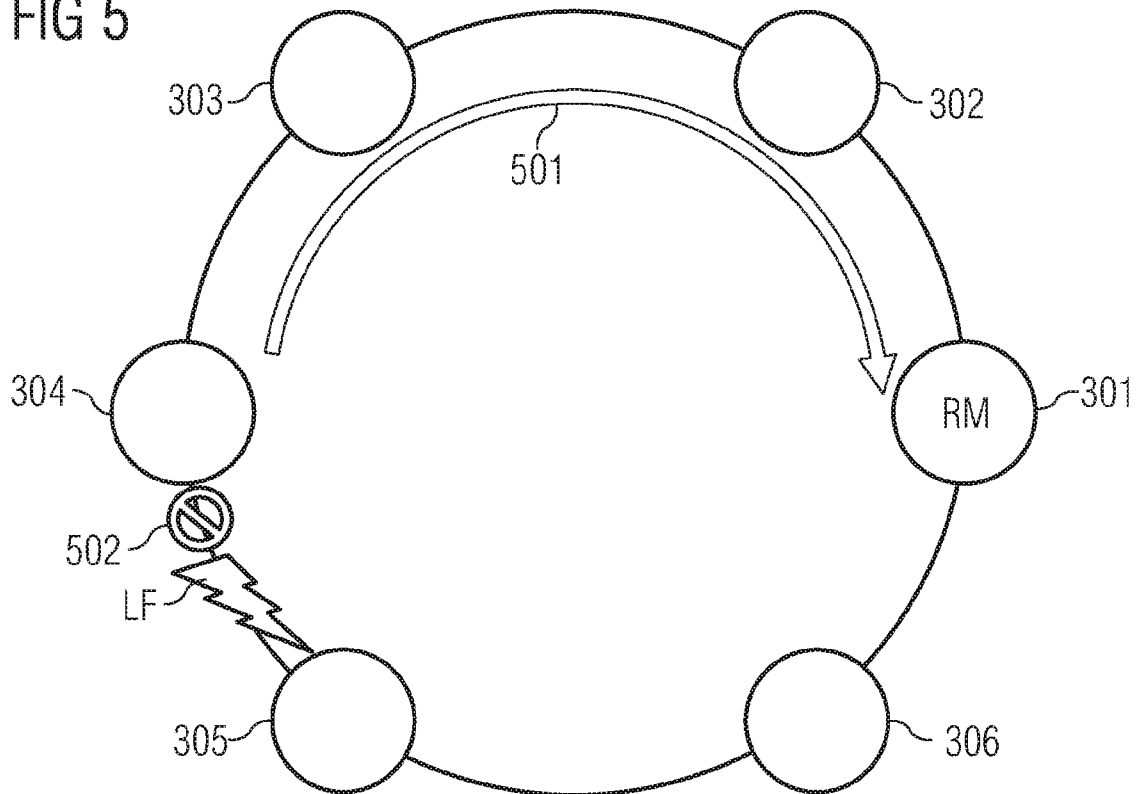
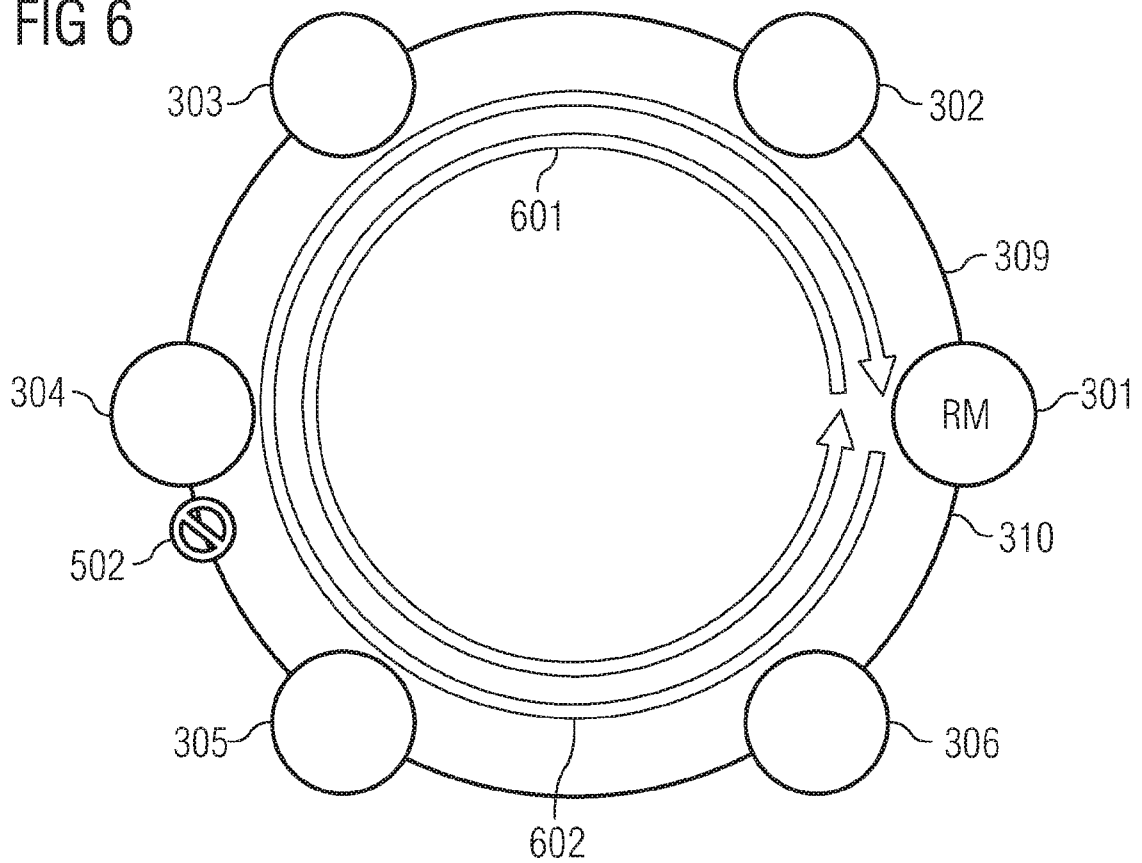


FIG 6



4/5

FIG 7

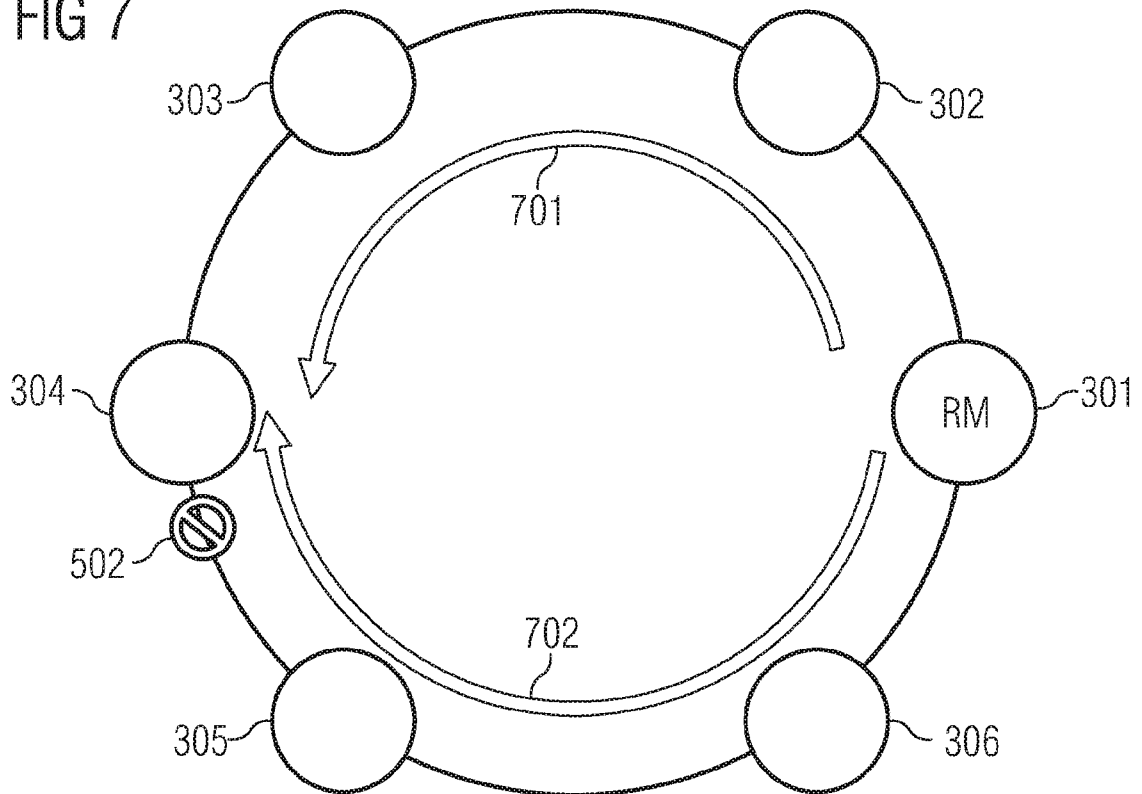
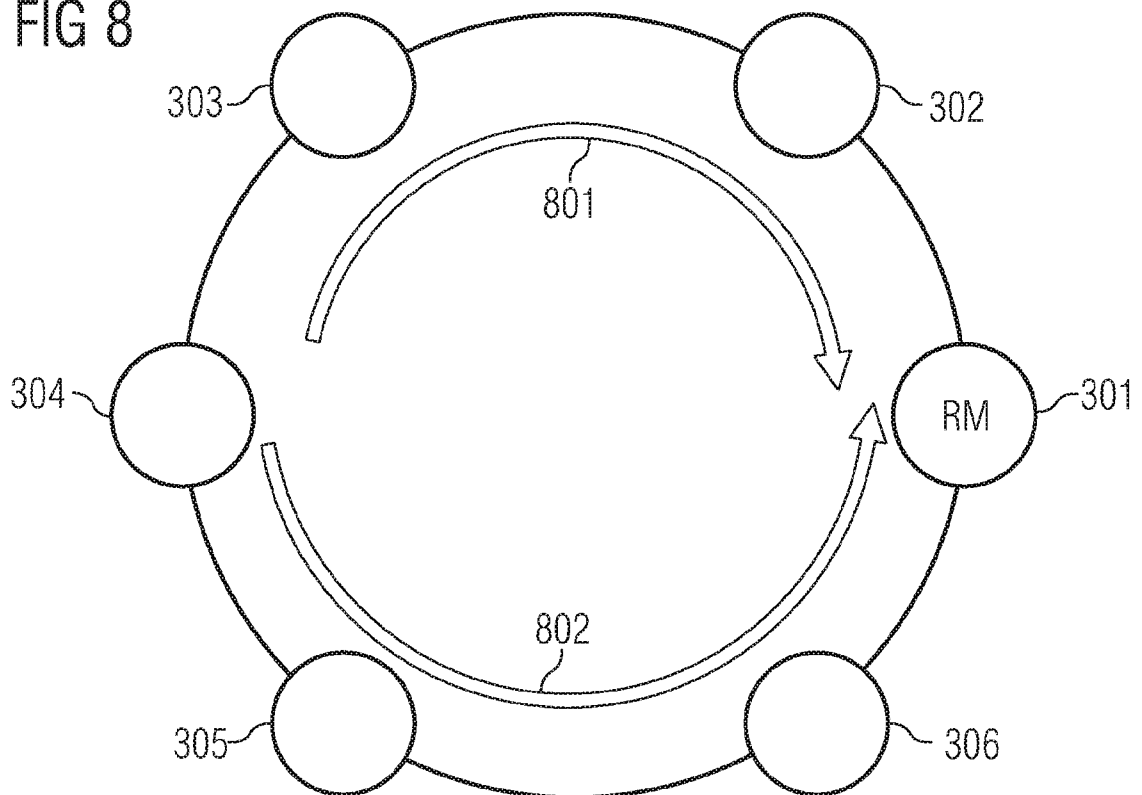


FIG 8



5/5

FIG 9

