10. Topology discovery

10.1 Scope

This section describes the RPR topology discovery protocol, which implements a reliable and accurate means for all RPR stations on a ring to discover the initial topology of the stations on the ring and any changes to that topology. The protocol is intended to scale up to hundreds of stations, to cause insignificant overhead for ring traffic, and to cause insignificant overhead on software and ASICs.

The services and features provided are:

- Determine/validate connectivity and ordering of stations on the ring
- Ensure all stations on the ring will converge to a uniform and current image of the topology normally within 1 ringlet circulation time
- Immediate reaction to changes
- Tolerant of message loss
- Operate without any master station on the ring
- Operate independently of and in the absence of any management systems
- Usable with all supported topologies: ring, bus (broken ring), and isolated station
- Support dynamic addition and removal of stations to/from the ring
- Detect mis-cabling between stations

> ***Editors' Notes (jl):*** *To be removed prior to final publication.*
>
> Detecting mis-cabling (above) assumes static local assignment of ring IDs (as opposed to dynamic discovery). How the assignments are made is a local, non-standardized decision. This might need to be discussed in the OAM&P clause.

- Provide means of sharing additional information between stations
- Cause insignificant overhead

The RPR topology discovery protocol is used to discover the physical link configuration between stations. It is not within the scope of the RPR topology discovery protocol to determine the dynamic link status information, i.e. which ringlet links are up or down, ring segment failures, etc. The discovered topology is used by other protocols such as the RPR protection protocol and the RPR congestion avoidance protocol.

10.2 Algorithm overview

The RPR topology discovery protocol provides each station on the ring with knowledge of the number and arrangement of other stations on the ring. This collection of information is referred to as the **topology image**. Each station maintains its own local copy of the topology image for the entire ring. Initially, the station's topology image contains information only about itself.

Ring topology discovery is initiated as needed and periodically. No station acts as a master for the topology image or for the protocol. All ringlet segments that can be discovered are included. A fully connected ring is not needed for the protocol.

In addition to station identifiers and physical connectivity relationships, the topology discovery protocol is also used to propagate additional station information, for use in other parts of this standard.

The messages sent as part of the RPR topology discovery protocol are indicated in the RPR frame header as control frames.

10.2.1 At initialization

At station initialization, the local topology image is initialized to contain only the local station and no links, and the addresses of the neighboring stations are initialized to all 0’s. The station starts the topology algorithm by broadcasting a Topology_Status message on all ringlets. Then it continually listens for Topology_Status messages broadcast on the ring, and broadcasts Topology_Status messages periodically and whenever there is a local topology change.

10.2.2 At addition of a station

When a station is inserted into an existing ring, it initilizes itself as described above. Its neighbors will detect the station as a new neighbor by receiving its Topology_Status message with a TTL set to Max_Ring_Size (indicating the message has traveled exactly one hop) and with the SA set to a value other than previously record (if any). The neighbors will respond by sending an immediate Topology_Status message indicating the new neighbor information to the rest of the ring, and providing topology information to the new station. All other stations will detect the new station by receiving its Topology_Status message with its neighbor MAC addresses set to all 0’s, and each will send an immediate Topology_Status message, providing topology information to the new station.

10.2.3 At removal of a station

When a station is removed from a ring, the stations on either side do not change their topology image until a new neighbor is detected. The new neighbor station could be a new station placed in the same location as the old station, or the former neighbor on the other side of the removed station.

10.2.4 Topology discovery message

At first bring up, at any point that a station detects a change in local status, at any point that a station detects a new station on the ring, and periodically, a station broadcasts a Topology_Status message to all stations on the ring. The Topology_Status message contains all the information about the local station, including its links to its neighbors. When a station receives a Topology_Status message, it updates its local topology image.

The Topology_Status message contains the ringlet ID on which it is sent. When a station receives a Topology_Status message with a TTL set to Max_Ring_Size (indicating the message has traveled exactly one hop), it verifies that the ringlets of both stations are connected with the same ringlet ID value (as indicated in the received message and by local configuration).

The Topology_Status message, like all RPR frames, contains the source MAC address of the station from which it is sent. When a station receives a Topology_Status message with a TTL set to Max_Ring_Size (indicating the message has traveled exactly one hop), it verifies that it knows who its neighbor is.

10.2.5 Checking topology consistency

It can be easily determined when an image is complete and consistent by examining the image contents. When the contents of the local topology image show station information for each station described in the link information of another station, then the image is complete. For each ringlet, when the contents of the local topology image show that all stations on that ringlet are connected to each other in a logical ring, bus (broken ring), or isolated station, then the topology image is consistent.

A canonical form for the topology image allows all the stations to eventually arrive at the same image for the topology.

10.2.6 Topology database and distance determination

The distance to another station can be calculated either through examining the topology image or by using the difference of the received TTL from Max_Ring_Size (255) + 1. For each direction, either the topology image must be complete up to the station whose distance is desired, or a Topology_Status message must have been received from the station from the desired direction.

> ***Editors' Notes (jl):*** *To be removed prior to final publication.*
>
> The following should be moved into the OAMP and/or MIB clauses.

For each ringlet, the distance to each station is used as the index into the topology and status database. The values of each entry are: the MAC address of the station, the MAC address of the station right from the station, the MAC address of the station left from the station, the protection status of the link immediately after the station, and the fair rate of the link immediately after the station. An example table for one ringlet is shown in Table 1.

Table 1—Topology and status database example

values/ distance	local MAC	right neighbor's MAC	left neighbor's MAC	next link's availability	next link's fair rate
0	00-10-A4-97-A8-DE	00-10-A4-97-A8-EF	00-10-A4-97-A8-BD	IDLE	3
1	00-10-A4-97-A8-EF	00-10-A4-97-A8-AC	00-10-A4-97-A8-DE	IDLE	4
2	00-10-A4-97-A8-AC	00-10-A4-97-A8-BD	00-10-A4-97-A8-EF	IDLE	3
3	00-10-A4-97-A8-BD	00-10-A4-97-A8-DE	00-10-A4-97-A8-AC	IDLE	3

10.2.7 Isolated station

At any point that a station loses connection with both of its neighbors, as detected by both links having a protection state other than IDLE, it is considered isolated. An isolated station shall remove all entries from its local topology image, and reset its neighbor MAC addresses to all 0's.

10.2.8 Determination and validation of ringlet ID

Each station determines which interface is associated with which ringlet and assigns the corresponding ringlet ID either through fixed mapping between hardware locations or through configuration. Each topology control message is sent separately on each ringlet, identifying the ringlet on which it is being sent. Any Topology_Status message with a TTL set to Max_Ring_Size received on a ringlet different from the ringlet on which it is identified as being sent shall cause the link to be declared non-operational and trigger a misconfiguration alarm.

10.3 Topology discovery process

10.3.1 Topology discovery process description

1) On startup
— Trigger
  At start of state machine.
— Action
  a) Clear all entries from the local topology image.
  b) Set the neighbor MAC addresses to all 0's.

c) Broadcast a Topology_Status message on each ringlet.
d) Start the Topology_Status_Timer.

2) Station status received
— Trigger
A Topology_Status message is received.
— Action
a) If the TTL=Max_Ring_Size and ringlet_id in the message does not match the ID of the ringlet on which the message was received, then raise a mis-configuration alarm and declare the link as non-operational.
b) If ringlet_id in the message matches ID of ringlet on which message was received, then continue.
c) Replace the remote station information in the local topology image for the ringlet on which the message was received.
d) If the TTL=Max_Ring_Size and the SA is different from the stored neighbor address, then change the local neighbor MAC address to that of the SA, and broadcast a Topology_Status message on each ringlet.
e) Reset the Topology_Status_Timer.

3) Status timer pop
— Trigger
The Topology_Status_Timer expires.
— Action
a) Send a Topology_Status on each ringlet.
b) Start the Topology_Status_Timer with the next backoff value.

4) Isolated
— Trigger
Both links enter protection states other than IDLE.
— Action
a) Clear the local topology image.
b) Set the neighbor MAC addresses to all 0’s.

10.4 Topology discovery message format

10.4.1 Topology_Status

Topology_Status messages report changes in neighbor identity. They are sent as MAC Control messages with a control opcode of Topology_Status. They are sent as broadcast frames on all ringlets, with TTL of Max_Ring_Size (255), removed by the source station, and with the source MAC set to the actual MAC of the sending station.

The information field of the message is as follows:

Bytes 0..1	station_capabilities: unsigned 16-bit integer
Bytes 2..7	right_station_address: IEEE-48 MAC address
Bytes 8..13	left_station_address: IEEE-48 MAC address
Bytes 14..17	right_A_bandwidth: unsigned 32-bit integer
Bytes 18..21	left_A_bandwidth: unsigned 32-bit integer
Bytes 22..25	right_B_bandwidth: unsigned 32-bit integer

Bytes 26..29	left_B_bandwidth: unsigned 32-bit integer

Table 2—Topology_Status message format

Parameters (see Table 2 **above for codings):**

station_capabilities: The station_capabilities parameter contains the capabilities of the sending station. The contents and format of this field are encoded as shown below in Table 3.

Table 3—station_capabilities format

Bit	Value
0	Wrap protection capable (1)
1	Jumbo frame receive capable (1)
2-8	Reserved
9-15	Weight

right_station_address: The right_station_address parameter carries the MAC address of the station right to the sending station. If the station's MAC address is unknown, it shall be all 0's.

left_station_address: The left_station_address parameters carries the MAC address of the station left to the sending station. If the station's MAC address is unknown, it shall be all 0's.

right_A_bandwidth: The right_A_bandwidth parameter provides the total provisioned Class A bandwidth on the right link from the sending station. The units and acceptable values are TBD.

left_A_bandwidth: The left_A_bandwidth parameter provides the total provisioned Class A bandwidth on the left link from the sending station. The units and acceptable values are TBD.

right_B_bandwidth: The right_B_bandwidth parameter provides the total provisioned Class B bandwidth on the right link from the sending station. The units and acceptable values are TBD.

left_B_bandwidth: The left_B_bandwidth parameter provides the total provisioned Class B bandwidth on the left link from the sending station. The units and acceptable values are TBD.

Editors' Notes (jl): *To be removed prior to final publication.*

The bandwidth fields are tentatively here until a definitive request from RAH is made to keep or drop them. If kept, then the units in which bandwidth is provisioned and the valid values need to be established.

10.4.1.1 When generated

The Topology_Status message is broadcast on the initial start of the RPR topology discovery, at any point that a station detects a change in local status, at any point that a station detects a new station on the ring, and periodically. The period between broadcasts of the Topology_Status message starts at a value of 500 ms for the first message sent after a changed message, and increasing by factors of 2 up to a maximum of 10 s.

> ***Editors' Notes (jl):*** *To be removed prior to final publication.*
>
> The values of 500 ms and 10 s are arbitrary, and should be finalized after a decision has been made on whether to combine this algorithm with protection and/or fairness. Much smaller values (e.g. 2 ms and 1000 ms) would be needed if 2 protocols were combined.

10.4.1.2 Effect of receipt

The receipt of this message on the same ringlet as which it was sent (as indicated in the ringlet_id field) from any station causes the MAC Control sublayer to update its current local topology image.

The receipt of this message on the same ringlet as which it was sent (as indicated in the ringlet_id field) from a neighbor station (as determined by TTL=Max_Ring_Size) causes the MAC Control sublayer to validate and (if needed) update the identity of its neighbor.

The receipt of this message on a ringlet other than that on which is was sent (as indicated in the ringlet_id field) from a neighbor station (as determined by TTL=Max_Ring_Size) causes the MAC Control sublayer to discard the message, place the link upon which the message was received into a non-operational state, and generate a miscabling error.

> ***Editors' Notes (jl):*** *To be removed prior to final publication.*
>
> Exactly what the non-operational state is needs to be chosen by the OAM&P team. Is it ForcedSwitch (implying that the link is physically working but we're making a concious decision not to use it anyway), SignalFail (implying that the link is effectively not working), or some new state to be defined by OAM&P or Protection?

A station in wrapped protection state shall not wrap a Topology_Status message, and shall strip it after receiving it.

> ***Editors' Notes (jl):*** *To be removed prior to final publication.*
>
> In order to indicate not to wrap these messages, either the type field needs another code point for non-wrappable control packets, or a separate wrap capable bit should be added to the RPR packet header. To allow for future possibilities, it would be cleaner to define a bit separate from packet type.