P802.17/D0.1, January 2002

Proposed Draft Standard for Resilient Packet Ring Access Method & Physical Layer Specifications

(Transit path and fairness behavior)

Draft 0.1

Comments on this proposal can be directed to the contributing editors:

Italo Busi
Alcatel
Via Trento, 30
20059 Vimercate (MI)
Italy
Phone: +39 039 686 7054
FAX: +39 039 686 3590
Email: italo.busi@alcatel.it

NOTE — The editors recognize that the some text of this document has been taken from the draft Darwin. The purpose of this draft is to propose some changes to the draft Darwin in order to specify a behavior instead of possible implementations.

Table of contents

Abbreviations

MAC	Medium Access Control
PHY	Physical Interface

References

[B1] IEEE 802.3 – 2000 Edition

Carrier sense multiple access with collision detection (CSMA/CD) MAC and physical layer specification.

6. Media Access Control data path

6.1 Transit buffer

To be able to detect when to transmit and receive packets from the ring, RPR MAC makes use of a transit buffer as shown in Figure 6.1.

The structure of the transit buffer is an implementation option and out of the scope of the IEEE 802.17 Standard. There are different implementations of the transit buffers. Some examples are shown in Annex K.

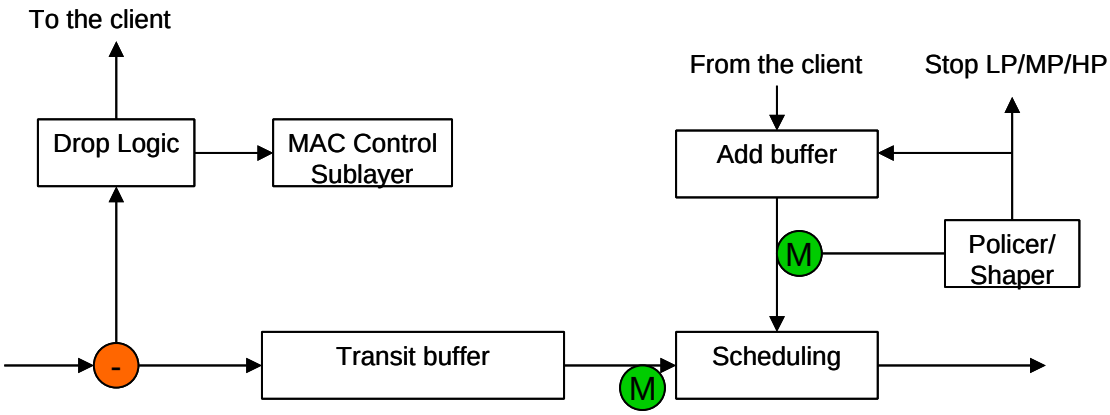

Figure 6.1 – MAC Data Path

6.2 Transmit and forwarding operation

A RPR MAC can transmit data packets from six possible flows:

1) High priority transit frame
2) Medium priority transit frame
3) Low priority transit frame
4) High priority frame from the client
5) Medium priority frame from the client
6) Low priority frame from the client

Note that Medium priority traffic is assigned a Committed Access Rate (CAR). Traffic within the CAR is treated as if it is high priority traffic while it is being accepted to the ring. Traffic above the CAR will be treated as low priority, and will be referred to as “excess” MP traffic, or eMP.

The RPR MAC will decide which traffic to send based on a scheduling algorithm. The specification of such an algorithm is implementation specific and out of the scope of the IEEE 802.17 Standard. Some examples are shown in Annex K.

The scheduling algorithm shall ensure:

1) The commitments on the HP and cMP transit and add traffic.
2) A fair access between the LP and eMP transit and add traffic. The IEEE 802.17 Standard defines a per-station weighted fairness allocation.
3) The LP and eMP add traffic should not exceed the allowed_rate parameter defined by the fairness protocol as specified in section 12.

6.3 Receive operation

Receive Packets entering a node are subject to the Header Error Check (HEC) test. If this test fails, the frame is stripped from the ring.

If the HEC test is passed, frames are then subject to the Destination Address (DA) match. If the DA matches, control frames are passed to the MAC Control Sublayer unit, while user data frames are passed to the client. If a DA matched packet is also a unicast, then the packet will be stripped.

If a packet does not DA match or is a multicast and the packet does not Source Address (SA) match, then the packet is passed to the Transit Buffer (TB) for forwarding to the next node if the packet passes Time To Live tests.

6.4 Transit operation

A series of decisions based on the type of packet, source and destination addresses are made on the MAC incoming packets. Packets can either be control or data packets. The rules for reception and stripping are given below as well as in the flow chart in Figure 6.2.

The flowchart is TBD

Figure 6.2 – RPR Receive Flowchart

1) Received packets will be discarded if there is a HEC error.
2) Conditionally decrement TTL on receipt of a packet, discard if it gets to zero; do not forward. The conditions to decrement TTL are as follows: always decrement unless the ring is in the wrap state (anywhere) and the ring id in the packet an in the MAC do not match.
3) Strip unicast packets at the destination station. Control frames are passed to the MAC Control Sublayer while user data frames are passed to the MAC client.
4) Copy multicast control frames to the MAC Control Sublayer or multicast user data frames to the MAC client.
5) Do not process packets other than for TTL and forwarding if ring identifier bit is not matched for the direction in which they are received unless the node is wrapped.
6) Packets to be sent to the MAC client due to destination address match may be optionally discarded at the MAC if there is an FCS error.
7) Transit packets may be optionally discarded at the MAC if there is an FCS error.
8) Packets with source address and ring identifier bit match should be stripped. If the node is wrapped and source address matches then the packet should be stripped.

6.4.2 Transit operation in a Bridge (Promiscuous Mode)

When the RPR MAC is part of a bridge, all data packets that do not DA match are copied to the Bridge Relay Entity and forwarded to the transit buffer.

Optional behaviors to improve bridging performance include the use of a MAC Filtering Database to hold the DA and SA of stations that are located behind the bridge: In this case the DA and SA of the packet can be checked to determine if the packet is to be dropped or stripped. If the addresses are not found in the database then the same rules as promiscuous mode apply.

6.5 Circulating packet detection (stripping)

Packets continue to circulate when transmitted packets fail to get stripped. Unicast packets are normally stripped by the destination station or by the source station if the destination station has failed. Multicast packets are only stripped by the source station. If both the source and destination stations drop out of the ring while a unicast packet is in flight, or if the source node drops out while its multicast packet is in flight, the packet will rotate around the ring continuously.

The solution to this problem is to have a TTL or Time To Live field in each packet that is set to the number of nodes in the ring. As each node forwards the packet, it decrements the TTL. If the TTL reaches zero it is stripped off of the ring. In order to allow 256 nodes on a wrapped ring, the TTL is not decremented when the packet is on the opposite ring and the ring is still wrapped. Once the ring unwraps, TTL decrements are per-formed on all packets. This catches the case where the packet is stuck on the wrong ring.

The ring identifier bit is used to qualify all stripping and receive decisions. This is necessary to handle the case where packets are being wrapped by some node in the ring. The sending node may see its packet on the reverse ring prior to reaching its destination so must not source strip it.

A potential optimization would be to allow ring identifier bit independent destination stripping of unicast packets. One problem with this is that packets may be delivered out of order during a transition to a wrap condition. For this reason, the ring identifier bit should always be used as a qualifier for all strip and receive decisions.

6.6 Wrapping of data

Normally, transmitted data is sent on the same ring to the downstream neighbor. However, if a node is in the wrapped state, transmitted data is sent on the opposite ring to the upstream neighbor. Packets of type 0x3 are marked for steering only, and when they reach a wrap point they are stripped.

6.7 Pass-thru mode

An optional mode of operation is pass-thru mode. In pass-thru mode, a node transparently forwards data. The node does not source or sink packets. It may optionally decrement the TTL and adjust the HEC but does no other modifications to the packets that it forwards. The node does not source any control packets (e.g. topology discovery or protection switch protocol) and basically looks like a signal regenerator with delay (caused by packets that happened to be in the transit buffer when the transition to pass-thru mode occurred). A node can enter pass-thru mode because of an operator command or due to a error condition such as a software crash.

The justification for continuing with the TTL decremented operation is to prevent a packet from being delivered twice if the node that sourced the packet is the node that goes into pass-thru. This could cause packets to be stripped early when topology discovery has determined that the ring contains fewer stations and adjusts the TTL value down in magnitude.

NOTE — We can use this mode also during auto-configuration to ensure that all the nodes on the ring support the same options. This will require this node to exchange topology data. Alternatively we can define another operation mode.

11. Media Access Control

11.1 Overview

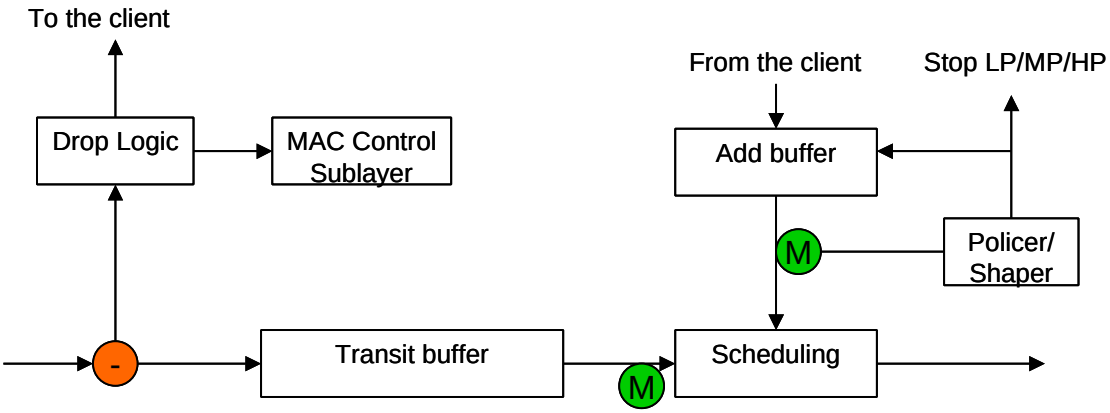

Figure 11.1 – MAC Reference Model

11.2 Transmit and forwarding operation

This section is TBD

11.3 Dynamic traffic shaping

This section is TBD

11.4 Pre-provision bandwidth for high priority traffic

This section is TBD

11.5 RPR ring access operation

This section is TBD

12. MAC fairness

12.1 Overview

This section is TBD

12.2 Congestion detection

Congestion detection is strictly dependent on the actual queuing and the scheduling implementations.

Thus congestion detection mechanisms are implementation specific and out of the scope of IEEE 802.17 Standard. Some examples are shown in Annex K.

12.3 RPR fairness packet format

This section is TBD.

Annex K Implementation Guidelines

(Informative)

This section is TBD.

NOTE — Some text can be grabbed from the Darwin proposal to show the Darwin implementations as possible example of IEEE 802.17 standard compliant implementations.