Information technology—
Telecommunications and information exchange between systems
LAN/MAN specific requirements—

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

Sponsored by the
LAN/MAN Standards Committee
of the IEEE Computer Society

Abstract: This document is a standard for medium-access and physical layer components that meet the functional requirements of a Resilient Packet Ring system as defined by the IEEE 802.17 Working Group. Detailed logical, electrical, and signal processing specifications are presented that enable the production of interoperable equipment.

Keywords: metropolitan area network standards, resilient packet rings, ring networks

Introduction

(This introduction is not part of IEEE P802.17, Resilient Packet Ring Access Method & Physical Layer Specifications

At the time this standard was completed, the working group had the following membership:

TABLE OF CONTENTS

INTRODUCTION **3**

1 OVERVIEW **1**

1.1 SCOPE 1
1.2 PURPOSE 2
1.3 TERMINOLOGY 2
1.4 IEEE ARCHITECTURAL CONFORMANCE 3
1.4.1 Layer definitions *3*

2 NORMATIVE REFERENCES **3**

3 DEFINITIONS **3**

4 ABBREVIATIONS AND ACRONYMS **3**

5 RPR CONCEPTS AND REFERENCE MODEL **4**

5.1 THIRD-PARTY OPERATION 4
5.2 SERVICES 4
5.3 NETWORK PROPERTIES 4
5.3.1 Network scale *4*
5.3.2 Topologies *5*
5.3.3 Shared medium *6*
5.3.4 Packet-based Operation *6*
5.3.5 Resilience *7*
5.3.6 Destination Removal *7*
5.4 BANDWIDTH MANAGEMENT FEATURES 7
5.4.1 Congestion Avoidance *7*
5.4.2 Fairness algorithm *8*
5.4.3 Jitter and Delay Considerations *9*
5.5 PHYSICAL LAYER INDEPENDENCE 10

1 Overview

1.1 Scope

This standard is intended to provide interoperability between implementations of equipment providing communication with a variety of service classes in the metropolitan domain. As is the case with most IEEE 802 standards, it is based on a shared medium serving all the nodes on the network.

Shared media are well suited to bursty and unpredictable demands, but they require a coordination mechanism, known as Medium Access Control, or MAC, to provide equitable access to the medium and prevent interference between users. In local and metropolitan networks, the distances are such that usage information can propagate across the network within a relatively short time interval. This makes it possible to achieve both high link utilization and responsiveness well within the delay requirements of various real-time applications. This distinguishes these networks from wide-area networks, where longer distances make the same goals much more difficult to achieve.

For the purposes of this document, a “system” constitutes: an 802.17 MAC implementation, in which at least two stations communicate via a dual or multipath ring, the interfaces to external networks, the services transported by the MAC protocol layers, and the mapping of these layers onto a variety of physical layers through convergence procedures..

The 802.17 Resilient Packet Ring standard is part of a family of standards for local and metropolitan area networks. The following diagram illustrates the relationship of 802.17 protocols to other 802 standards, and to the OSI reference model. (The numbers in the figure refer to IEEE standard numbers.) Note that the 802.17 standard does not specify its own physical layer. Rather, it makes use of existing physical layers standardized by other bodies.

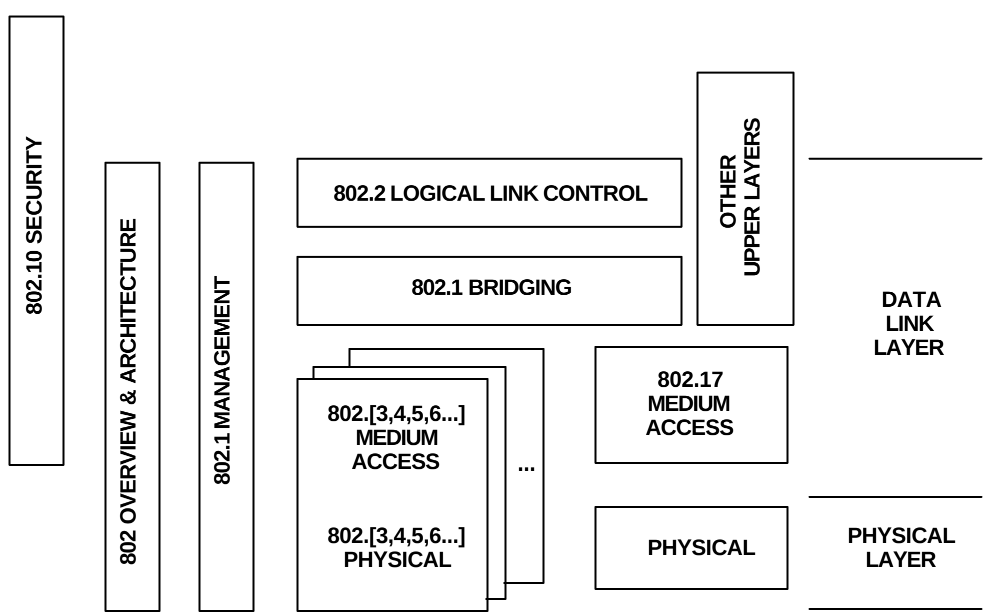

Figure 1-1. IEEE 802 standards family

This family of standards deals with the Physical and Data Link layers as defined by the International Organization for Standardization (ISO) Open Systems Interconnection Basic Reference Model (ISO 7498: 1984). The access standards define several types of medium access technologies and associated physical media, each appropriate for particular applications or system objectives. Other types are under investigation.

The standards that define the technologies noted in the above diagram are as follows:

IEEE Std 802: Overview and Architecture. This standard provides an overview to the family of IEEE 802 Standards. This document forms part of the 802.1 scope of work.

ANSI/IEEE Std 802.1B [ISO/IEC 15802-2]: LAN/MAN Management. Defines an Open Systems Interconnection (OSI) management-compatible architecture, environment for performing remote management.

ANSI/IEEE Std 802.1D [ISO/IEC 10038]: MAC Bridging. Specifies an architecture and protocol for the interconnection of IEEE 802 LANs below the MAC service boundary.

ANSI/IEEE Std 802.1E [ISO/IEC 15802-4]: System Load Protocol. Specifies a set of services and protocols for those aspects of management concerned with the loading of systems on IEEE 802 LANs.

ANSI/IEEE Std 802.2 [ISO/IEC 8802-2]: Logical Link Control

ANSI/IEEE Std 802.3 [ISO/IEC 8802-3]: CSMA/CD Access Method and Physical Layer Specifications

ANSI/IEEE Std 802.4 [ISO/IEC 8802-4]: Token Bus Access Method and Physical Layer Specifications

IEEE Std 802.10: Interoperable LAN/MAN Security, Secure Data Exchange (SDE)

1.2 Purpose

This standard provides detailed architectural specifications for a Resilient Packet Ring metropolitan network. It describes the method to access the shared medium and the implementation interfaces to higher layers and to a variety of physical layers based on other standards.

1.3 Terminology

Throughout this document, the words that are used to define the significance of particular requirements are capitalized. These words are:

"MUST" or “SHALL” These words or the adjective "REQUIRED" means that the item is an absolute requirement for any implementation conforming to this standard.

"MUST NOT" This phrase means that the item is an absolute prohibition.

"SHOULD" This word or the adjective "RECOMMENDED" means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighed before choosing a different course.

"SHOULD NOT" This phrase means that there may exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label.

"MAY" This word or the adjective "OPTIONAL" means that this item is optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product, for example; another vendor may omit the same item.

1.4 IEEE Architectural Conformance

1.4.1 Layer definitions

The protocol layering used in this standard, and its relationship to the OSI Reference Model, is shown in Figure 1-2.

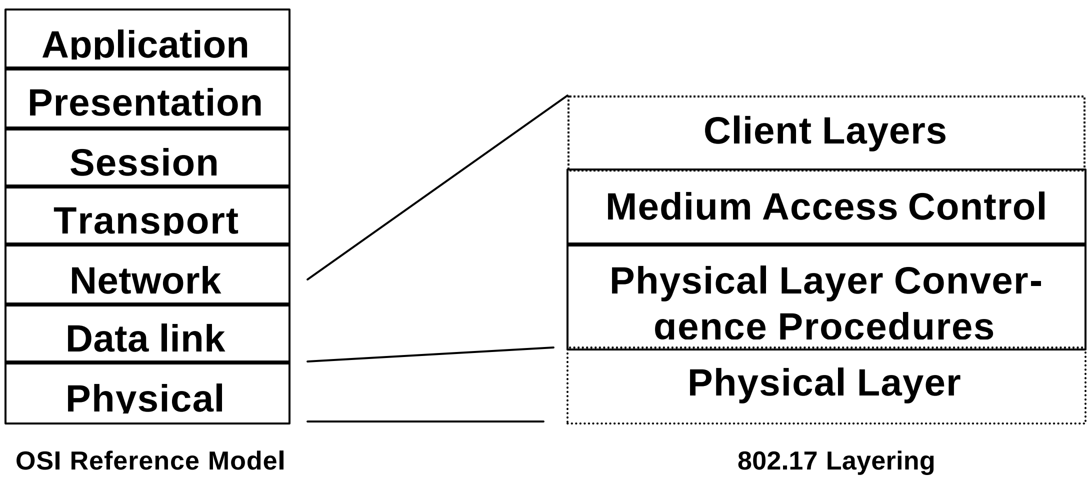

Figure 1-2. Layer structure in relation to the OSI Reference Model; the three sublayers with solid borders are covered in this standard.

2 Normative references

3 Definitions

4 Abbreviations and acronyms

5 RPR Concepts and Reference Model

5.1 Third-Party Operation

The RPR standard deals with metropolitan ring networks. Such networks are generally operated by third-party service providers, although in some instances the operating environment may be university and industrial campuses where there is no third party. Both types of operation are supported.

Within the IEEE 802 family, most of the standards, such as the various versions of Ethernet, apply to local area networks, owned and operated by a single organization. In other cases, such as the 802.16 Broadband Wireless Access standard, a third-party operator is assumed. In that case, the nature of the wireless medium makes it necessary makes it necessary to provide a high level of authentication and encryption, in order to assure access only by authorized users and to guarantee privacy of the transmissions.

In the case of RPR, the need for such facilities is less, given that the need for physical access to wireline facilities provides security otherwise absent in wireless systems.

However, it is important to third-party service providers to be able to provide guarantees of service through service-level agreements (SLAs). The level of service provided to one user, in terms of throughput, delay, and other characteristics, should not be impacted by the actions of other users.

Thus this standard provides for a variety of classes of service and provides mechanisms to prevent users from degrading the service seen by others.

5.2 Services

The applications that are expected to use an RPR network encompass the complete range of networking applications, including the familiar set of voice, video, and data. The resilient, or fail-safe, ring is ideal for critical applications that require high availability; since there are two paths between any two points on the ring, failure of a link need not prevent applications from running.

Typical uses of the network would include Internet access; in this case most traffic would exit from the RPR ring into a wide-area network or WAN for delivery to distant locations. Other uses would be for virtual private networks (VPNs) where bandwidth guarantees might be established to provide the equivalent of a hard-wired private line but with the cost savings associated with the ability of the network to recycle unused bandwidth to other users, or to allow a customer to burst at higher rates on an occasional basis.

5.3 Network Properties

5.3.1 Network scale

RPR technology is optimized for the needs of metropolitan networks, in other words on a scale larger than a LAN but smaller than a typical WAN. Circumferences ranging up to several hundred kilometers are likely to be the most common. However, this does not preclude its use in other situations, for example as a building or campus backbone. It may also be used with ring circumferences ranging into thousands of kilometers with relatively little loss of responsiveness, should the fail-safe or bandwidth management properties of the network make it attractive to do so.

The data rates that can be accommodated by the RPR design cover a wide range. The protocols are designed to operate over a variety of physical layers, including SONET/SDH, Gigabit Ethernet (IEEE 802.3ab), and DWDM fiber. As higher-speed physical layers become available, it is expected that RPR will be able to work over them as well.

5.3.2 Topologies

This standard is intended for network configurations that have a ring topology, where there is a well-defined ring structure that offers two possible paths from any source to any destination. This is in contract to a fortuitous loop in the cabling configuration, which is generally avoided in networking practice except as backup.

5.3.2.1 Dual Ring

The basic ring configuration as used in RPR systems is a dual ring, arranged so that transmission on one ring (sometimes referred to as a ringlet) goes one way around, and the other direction on the other ring. [Other media than fiber can be used, but for simplicity we will refer to the physical medium as fiber.]

The nodes are able to send data on either ringlet, in other words clockwise or counter-clockwise. Generally the shorter of the two possible paths to a given destination is used, based on the node discovery scheme, but this is not required, due for instance to congestion or malfunction on one of the links. The structure of an individual node can be thought of as a dual add-drop multiplexer, as shown for a dual ring in Figure 5-1. The inputs to the switch include:

Receive from upstream
Local transmission queues

Outputs are:

Packets for this node
Transmit downstream

Note that multicast and broadcast packets are copied to the local receiver and transmitted downstream.

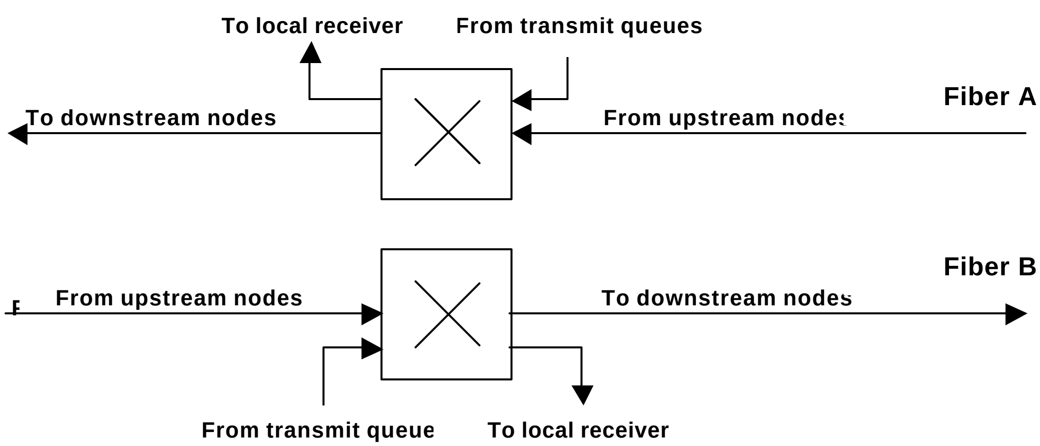

Figure 5-1. Switching functions at a dual-ring node.

The overall structure of the system is as shown in Figure 5-2.

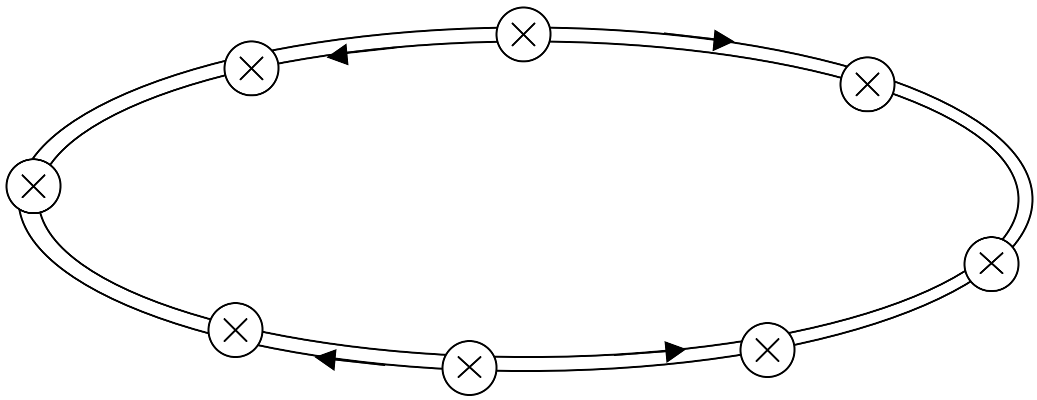

Figure 5-2. Dual-ring MAN; individual nodes are as shown in Figure 5-1.

The use of a dual ring has substantial advantages in terms of the capacity of the system. Since there are two paths to each destination, it is possible to send traffic over the shorter path rather than the longer one; hence the longest path taken by any packet is half the ring circumference. If the destinations are randomly distributed along the ring, then the average path length is half of this, or a quarter of the circumference. Given that packets are stripped at the destination node (see 5.3.6) extensive spatial re-use is possible, with the total capacity of the dual ring approaching 8 times that of a single fiber.

5.3.2.2 Multi-ring

It is also possible to have a ring structure with more than two physical paths, which may be additional links or possibly different wavelengths multiplexed within the same fiber. (Indeed, with wavelength-division multiplexing, it is possible to have a dual or mult-ring system implemented on a single strand of fiber.)

In the multi-ring case, there exist two or more choices of path to a given destination. Assuming all of them are operational, it is up to the node to choose which one to use. (This choice is beyond the current scope of this standard.) Once chosen, the MAC protocol as applied to that path is used.

5.3.3 Shared medium

An RPR ring acts as a broadcast medium, in which a single transmission is capable of reaching all stations on the network. This means that the large number of mechanisms that have come into use with broadcast networks such as Ethernet will still work with RPR. Examples of these include the ARP protocol, the spanning tree protocol (IEEE 802.1D), and Layer 3 protocols in general.

Even under protection-switching conditions, where the rings operate as a bidirectional bus, the broadcast property still holds. Protection switching is an internal function of the RPR system and does not impact higher layers.

5.3.4 Packet-based Operation

The basic unit of data on the RPR ring is a packet, consistent with current networking practice. This follows the precedent of the IEEE 802.5 Token-Passing Ring and the FDDI ring. It does differ, however, from the 802.6 Distributed Queue Dual Bus ring, which was based on cells, and SONET/SDH rings, which operate on a TDM basis with 8-bit granularity.

5.3.5 Resilience

The redundancy provided by two or more paths to the same destination provides resiliency in the face of fiber or equipment failure. This standard provides resiliency without the need to assign half of the fiber links to standby status. During normal operation, all paths carry traffic.

On the occurrence of network faults, due either to fiber breaks or equipment malfunction, the RPR system can continue to function. While throughput may be degraded, all functioning nodes continue to operate on the ring.

5.3.5.1 Source Steering

Given that the ring provides two routes to any destination, if one is not operational, the other can be used. The source node simply must be aware of the existence of the failure; for the duration of the outage it can send all data over the operational path.

5.3.6 Destination Removal

RPR achieves a high degree of link utilization by having the destination node remove packets from the ring. This is called spatial reuse because it provides empty space on the ring which that node or another node down the fiber can use to send additional data. This is in contrast to the 802.5 Token-passing Ring or FDDI, in which packets which have already been received continue on the ring until they are removed their original source node.

There is a provision for dealing with failed or non-existent destinations. A time-to-live or hop count field in the header makes it possible to detect packets circulating excessively and to remove them.

5.3.6.1 Multicast Issues

An exception to destination stripping is made for multicast and broadcast packets. Since there is no single destination to remove the packet, stripping is done only by the original sender. If the ring fails, the packet must be sent in both direction with the TTL set to expire after one trip and not allowing duplicate packets to exist when the station is re-stored. When a station is inserted on the ring it initially sends in both directions until it finds out the location of the destination nodes.

5.4 Bandwidth Management Features

One of the key features of RPR is a sophisticated set of bandwidth management features. These features operate in a fully distributed manner: there is no master node on the ring, but all features are carried out by the collective operation of the algorithms in each node.

These algorithms make it possible to support service-level agreements (SLAs) between the service operator and the customers, providing guarantees of levels of service while at the same time permitting unused bandwidth to be used by nodes that need it. This permits a very high degree of link utilization, without impeding the quality of service needed by applications running over the network.

Since the RPR network will usually be operated by third parties, setting of the parameters that govern the bandwidth management algorithms is under the control of the service operator.

5.4.1 Congestion Avoidance

As a shared medium, each Resilient Packet Ring carries data from many sources to many destinations. The use of such a shared medium requires a Medium Access Control, or MAC, to govern transmissions and meet the needs of all nodes with minimal delay and negligible packet loss. At the same time, it is necessary to comply with service level agreements between the system operator and the customer, ensuring levels of throughput and delay that are not impacted by the presence or absence of traffic from others.

RPR achieves this by focusing traffic control on the node where the traffic enters the ring. Once a packet is transmitted onto the ring, it proceeds directly to its destination without intermediate queueing. This minimizes delay through the system and enables RPR to scale well as additional nodes are added to the ring.

The decision of which packet to forward onto the ring when an opportunity occurs does not affect interoperability of the nodes and may be left to the implementer.

5.4.1.1 In-transit vs. Transmitted Traffic

At each node, the primary activity is the removal of packets addressed to that node (multicast and broadcast excepted, as noted in 5.3.6.1) and the insertion of packets originating at that node.

Packets already in transit through the node receive priority over packets waiting to be transmitted. Pre-emption is not done; if the node is in process of transmitting a packet when another one comes in from upstream, the first packet is allowed to finish before the transit packet is sent.

This operational model of the ring may be thought of as similar to a traffic roundabout, where vehicles already in the roundabout have right-of-way over vehicles attempting to enter. The result is that traffic within the circle moves smoothly, while queues are established as necessary on the feeder roads.

5.4.1.2 Per-destination Queuing

One phenomenon which can occur in the automotive case is head-of-line blocking. If one output road is heavily congested, it can cause traffic to back up even for vehicles destined for other exits. However, the use of multiple queues can solve this problem in the electronic case.

In RPR, transmission of packets onto the ring is governed by the downstream congestion. Each node broadcasts the congestion status of its downstream link; hence all nodes are aware of the congestion situation of all downstream links. Each node maintains separate queues for all destinations on the ring. With destination-based queuing, the node is able to choose packets for transmission that do not pass over congested links. This eliminates head-of-line blocking. Consider a 4-node ring set up as

A ® B ® C ® D ® A

If link CD is congested, A defers sending packets to D but can still send freely to B and C.

5.4.2 Fairness algorithm

The mechanism for dealing with congestion in the presence of SLAs is based on each node broadcasting the amount of unused bandwidth it has available. Typically, RPR customers enter into SLAs that guarantee them a certain amount of bandwidth, and provide various priority, delay, and jitter parameters.

However, a customer is not likely to use the committed amount of bandwidth all the time. RPR makes it possible to make the unused bandwidth available to other users, along with any excess over and above the service agreements. Each node broadcasts an Available Bandwidth Factor periodically; this information is used by other nodes to tailor their transmissions to the available capacity of the system. In short, bandwidth unused by its "owner" is used by others but it can be repossessed on short notice. The net result is to provide very high link utilization in the face of widely fluctuating demand.

The Rate Control Factor (RCF) is computed as follows:

Assume that there are multiple flows, indexed by ***i***, and that the capacity of the outgoing link from the node is ***C***. The service operator assigns each flow a weight w and a nominal rate ***r***. Then the allowable flow rate ***f*** is the sum of the committed rate plus a share of the allocated but unused bandwidth, weighted by the ratio of ***w*** for this flow to the sum of the weights of all the active flows.

$$f_i = r_i + w_i \frac{(C - \sum_{active} r_i)}{\sum_{active} w_i}$$

The congestion state of the outgoing link is summarized by the available bandwidth factor ***RCF***, given by

$$RCF = \frac{(C - \sum_{active} r_i)}{\sum_{active} w_i}$$

In order to manage congestion around the ring, each node broadcasts its ***RCF*** for each outgoing path. To determine the bandwidth that a node can utilize in sending to a given destination on the ring, it computes

$$min\ (RCF_{\boldsymbol{n}})$$

where ***n*** ranges over all links between the source and the destination. Thus, in the example given in 5.4.1.2, if A wants to send to D, it finds the minimum of the available bandwidth on links AB, BC, and CD. Link CD is the congested one, and therefore it has the lowest ***RCF*** value, and this is the one used by A in determining its transmission rate to D.

Of course this transmitted load contributes to the load on links AB and BC as well. In the case of transmission to C, the lower of the RCF values of AB and BC is used, with allowance for the currently allocated traffic through these links to D.

Use of RCF determines the total amount that a node is allowed to send to each destination, but within this amount, the node may allocate transmission opportunities as it likes to the packets that it has queued up. Fairness algorithms may be applied as needed; standardization is not required since the choice of which packet to send does not affect ring operation as seen by other nodes. Control messages are an exception; transmission of control information may be specified by the standard as appropriate for each type of message.

5.4.3 Jitter and Delay Considerations

Any shared-medium system is subject to some degree of delay and delay variance, known as jitter. RPR minimizes these effects by queuing traffic at the entrance node. Once traffic is in the ring, it has priority over queued traffic at all intermediate points. This minimizes end-to-end delay in most situations and avoids having transmitted traffic subject to queuing algorithms at nodes along the way.

Packets for which transmission has started are allowed to go to completion. This means that there is a delay seen by a transit packet, if the node is currently transmitting. The average delay at each node is half the transmission time of the average-size packet. For a 1000-byte packet, this transmission time is under a microsecond for a 10-gigabit line speed. For slower links like Gigabit Ethernet, the time is proportionately longer, but it is clear that overall delay in even a large network will rarely exceed one millisecond. Since the propagation time of signals in fiber is 5 microseconds per kilometer, a network with a circumference of 500 km will in any case introduce propagation delays up to 1.25 milliseconds (half the fiber length) during normal operation, and longer during ring reconfiguration periods.

The jitter introduced by allowing packets to complete their transmission is due to random variations in the amounts of remaining data that each node must send before it is able to forward the transit packet. It can be removed by use of a de-jitter buffer at the receiver.

5.5 Physical Layer Independence

The RPR design allows for wide latitude in the choice of physical layers. Various fiber-based layers, such as Gigabit Ethernet and SDH/Sonet may be used. This standard includes Physical Layer Convergence Procedures (PLCP) to specify how RPR packets are carried over the physical layer, and how control information is passed between the layers. It is expected that new PLCP standards may be generated in the future as new physical layers come into use.