

一种面向运营级以太网的自适应前向纠错方法

戴锦友^{1,2} 余少华²

¹(华中科技大学计算机科学与技术学院 武汉 430074)

²(武汉邮电科学研究院 武汉 430074)

(djy96@fhn.com.cn)

An Adaptive Forward Error Correction Method for Carrier Ethernet

Dai Jinyou^{1,2} and Yu Shaohua²

¹(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074)

²(Wuhan Research Institute of Post and Telecommunication, Wuhan 430074)

Abstract Huge change has happened with physical channel of the present Ethernet while compared with that of traditional Ethernet. Carrier Ethernet brings forward higher request to Ethernet physical channel. Under the application environment of carrier Ethernet, it is of necessity to improve the transport performance of the Ethernet physical channel by using a special method. Based on the frame structure defined in IEEE Std 802.3, a kind of frame structure is presented, which comprises FEC (forward error correction) contents and an FEC method is also brought forward. The method includes an error correction function as well as a mechanism of detecting the performance of the physical channel and adaptively adjusting the FEC configuration, so the cost brought by the FEC method can be cut short to the minimum while the error correction performance is guaranteed. An Ethernet interface using the FEC method can communicate with an ordinary non FEC Ethernet interface, so the backward compatibility can be reached. Detailed analysis for the performance of the FEC method is given too, and the analysis result proves that the FEC method is feasible. A kind of network device that carries out the method is designed and a test topology is set up to test and verify the method. The test data show that the method has the expected performance.

Key words carrier Ethernet; optical channel; forward error correction; adaptive; MAC

摘要 相对于传统以太网,现有的以太网物理信道已发生巨大变化。运营级以太网则对以太网信道提出了更高的要求,在运营级以太网应用环境下,采用适当机制提高以太网信道可靠性是必要的。在IEEE802.3标准帧结构的基础上,提出了一种包含前向纠错(FEC)机制的帧结构,并阐述了基于该帧结构的FEC方法。该方法除实现纠错功能外,还具有检测信道性能并自适应调节FEC配置的机制,使得可以在保障纠错性能的基础上尽可能减少FEC带来的开销。并且使用该方法的以太网接口能够与普通以太网接口互通互联以实现后向兼容。对该方法的性能进行了透彻的分析,证明该方法可行。设计了网络设备实现该FEC方法,并搭建测试拓扑对该方法进行评测,测试结果表明,该方法可以实现预期功能。

关键词 运营级以太网;光信道;前向纠错;自适应;MAC

中图法分类号 TP302

以太网早已成为一种应用广泛的技术,它在接入层面占有 90% 以上的应用份额,在城域网中也占有越来越重要的地位.运营级以太网的盛行更是将以太网技术的应用提高到一个全新的层面.同时运营级以太网也提出了更高的要求.与传统以太网相比,运营级以太网具有下列 5 个特性:提供可靠的业务保护模式、可扩展性、对 TDM (time division multiplexing) 的支持、高 QoS (quality of service)、OAM (operation, administration, management) 能力^[1].要具备这 5 个特性,作为数据传送基础的以太网物理层和链路层必须是可靠的.

另一方面,相对于传统以太网,现有的以太网物理信道已发生巨大变化,很多情况下以太网在长距离光纤上传送,传送距离已从 100 多米扩展到几十公里甚至更远.以太网的传送速度从最初的 10 Mbps 扩展到 100 Mbps, 1 Gbps 以及 10 Gbps,将来甚至可到 100 Gbps.这要求传送以太网的光信道具备更好的性能.但现有的光信道具有多样性,多数情况下并不满足必需的性能要求.各种噪声源、色散引起的码间干扰以及定位抖动产生的误码都会造成信道发生错误.

SDH (synchronous digital hierarchy)、WDM (wave-division multiplexing) 和 ATM (asynchronous transport mode) 等基于光信道的主流传送技术已应用信道误码纠错的机制,相关标准已有描述^[2-3],也有很深入的研究^[4-5],同样在光信道上传送的以太网也必须解决信道误码带来的难题.而且,作为 Internet 基础技术的以太网,其上承载的语音和视频等业务对可靠性有要求,虽然有在应用层面保证其可靠性的相关研究^[6],但基于不可靠信道的应用纠错机制重复性高、开销大,并非最佳方案.

现有以太网标准包括基本以太网标准^[7]、千兆以太网标准^[8]以及 10Gbps 以太网标准^[9]都未涉及信道纠错的内容,更没有考虑到面向数据包的纠错技术和方法,当前也基本没有具有 FEC (forward error correction) 机制的以太网产品在实际工程中应用.

本文基于 IEEE802.3 提出了一种包含 FEC 内容的帧结构,并基于该帧结构提出了一种面向运营级以太网,应用在 MAC 和物理层之间的基于以太网数据包的 FEC 方法.

该方法通过特定机制,得到描述物理信道性能的定量数据,根据性能数据确定是否采用 FEC 机制以及 FEC 纠错机制的配置.在两端都具有同类

FEC 机制的情况下,可以利用 FEC 机制纠正信道产生的误码,从而改善物理信道的传送性能.

本方法还可检测与之相连的远端是否具有相同的 FEC 机制,从而确保与不具有 FEC 机制的普通以太网接口能够互通互联,实现后向兼容.

1 以太网物理信道性能分析

设一个以太网数据包包含的数据位为 N ,第 k ($k=1,2,\dots,N$) 位的出错概率为 p_k ,则该数据包正确传送的概率 P_c 可以用式(1)表示:

$$P_c = \prod_{k=1}^N (1 - p_k). \quad (1)$$

一般来说,信道的噪声分布是随机的,各数据位出错的概率可以视为相同,也就是说 $p_1 = p_2 = \dots = p_N =$ 信道误码率.如果数据包内有错误,通常会被丢弃,因此错包率和丢包率是相同的.

以 L 表示以太网的包长(以字节数表示), E 表示以太网物理信道的误码率,以 D 表示实际的传送丢包率, D 将随着 L, E 的变化而变化,将 D 与 L, E 的关系表示成函数: $D = f(L, E)$.结合前面的分析和式(1)则可得到式(2):

$$D = (1 - (1 - E)^{8L}) \times 100\%. \quad (2)$$

由式(2)可以看出,信道丢包率随着信道误码率的递增而递增,也随着以太网数据包长度的递增而递增.当信道误码率在 1.0×10^{-3} 量级时,长度大于 512 B 的数据包的丢包率接近 100%.

2 FEC 方法描述

FEC 方法将以太网数据包作为纠错单位,尽最大可能保证以太网数据包传送的正确性,同时,根据以太网物理信道的当前性能情况调整 FEC 码长度,尽量减少 FEC 机制带来的开销,力争以较小的开销获得最大程度的性能改善.

该方法 FEC 码的生成采用 $R-S(n, k)$ 编码体系,其中 k 是信息长度(字节数), n 是加上纠错码的信息总长度^[10].

由理论分析可知,采用 $R-S(n, k)$ 编码,Hamming 距离为

$$d_{\min} = n - k + 1. \quad (3)$$

则可以纠正的错误数至多为 t 个错误(每个错误指的是一个字节错误,即使是该字节内 8 位全错,也算是一个错误):

$$t = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor = \left\lfloor \frac{n - k}{2} \right\rfloor. \quad (4)$$

编码时,将原始以太网数据包从逻辑上划分成长度为 32B 的数据节,每个数据节使用编码器独立生成相应的纠错码。

前述的 k 值等于一个数据节的长度,为 32, n 的取值则有 $k+2, k+4, k+8, k+16, k+24, k+32$ 等 6 个选项,可以根据远端端口反馈的以太网物理信道的实际性能情况来调整 n 的取值,目的是在达到预期性能目标的前提下,尽可能减少 FEC 方法带来的附加开销.并且以太网数据包中应有 n 的取值信息的显式表示, n 的取值可以由系统自动选取,也可以手工强制设定。

本方法的设计原则可以归纳为两个:吞吐量优先和正确率优先.吞吐量优先的原则是在物理信道带宽和误包率确定的情况下,选择适当的 FEC 机制,使得在单位时间内能在物理信道上正确传送的以太网数据包个数尽可能多.正确率优先的原则是,在物理信道带宽和误包率确定的情况下,选择适当的 FEC 机制,使得能在物理信道上正确传送的以太网数据包与传送的数据包总数的比值尽可能大。

图 1 是 FEC 方法的功能结构, FEC 子层位于以太网接口的接入访问控制子层(media access control, MAC)和物理层(PHY)之间:

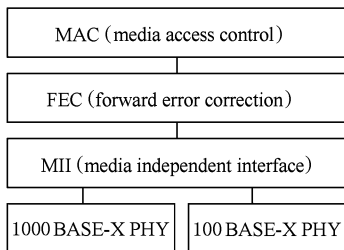


Fig. 1 Function structure of the Ethernet interface with the FEC mechanism.

图 1 具有 FEC 机制的以太网接口功能结构

该 FEC 方法主要有以下两方面的功能:

1) 发送数据包时,发送实体在发送侧从 MAC 层接收以太网数据包.如果远端以太网口(与本端以太网接口直接连接)反馈的信息表明远端能够完成 FEC 功能,则根据手工配置信息或物理信道的性能情况,按照以太网数据包的内容生成纠错码,并将纠错码与数据包组合在一起形成一个新的数据包,再将数据包发送到本端 PHY 层.如果远端以太网口(与本端以太网相连接)反馈的信息表明它不具备 FEC 机制处理能力,则跳过纠错逻辑子层,直接将数据包传递到本端 PHY 层,以实现后向兼容,即与普通以太网接口能够互通互联。

2) 接收数据包时,接收实体在接收侧从 PHY 层接收包含冗余码的以太网数据包.如果数据包是带 FEC 机制的数据包,则剥掉纠错码,检查数据包是否包含有错误,并利用纠错码纠正数据包包含的错误,再将正确的数据包送交本端 MAC 层.如果数据包是不带 FEC 机制的普通数据包,则直接将数据包送交 MAC 层。

如果远端以太网口具有相同的 FEC 机制,则从本端以太网端口接收数据包时,应有错误包计数和总包计数的功能,以便估算物理信道的当前性能.物理信道的性能可用误包率来表示.以 P_e 表示误包率, S_e 和 S 分别表示收到的错误数据包数和收到的总数据包数,则

$$P_e = (s_e/s) \times 100\%. \quad (5)$$

同时,需要从收到的含 FEC 信息的数据包中提取数据包中表示物理信道当前性能的信息,以用于确定新的 n 值。

值得强调的是,首先需确定与本端口互联的远端以太网口的类型,换句话说,需要知道远端以太网口是普通以太网口,还是与本端具有相同 FEC 机制的以太网口.如果判断远端是普通以太网端口,则将本端设置成普通以太网工作模式,以便本端口可以与普通以太网口互通互联,实现后向兼容.判断的方法如下:①如果在本端发送之前,已收到远端发送的数据包,则对照第 3 节的帧结构确定数据包是否携带相应的 FEC 信息,如携带则远端与本端具有相同的 FEC 机制,否则,远端是普通以太网接口.②如果在本端发送之前未能确定远端端口的性质,则在发送数据之前,发送一个净信息为 64B 全 1(包括地址域)的 FEC 数据包,远端端口如与本端具有相同性质,则回应一相同的数据帧,否则不作操作。

3 本方法采用的帧结构

本方法采用的帧结构如图 2 所示。

图 2 第 1 行对应原 IEEE802.3 协议帧,该协议帧的帧结构说明可参考文献[6],该协议帧可看做 FEC 机制采用的协议帧的净载荷。

图 2 的第 2 行表示 FEC 码组,每一列表示每个数据节的 FEC 码,长度是 $n-k$ 字节,而列数 I 的取值由式(6)确定.以 L 表示 IEEE802.3 帧结构的帧长度,则

$$I = \left\lceil \frac{L}{32} \right\rceil. \quad (6)$$

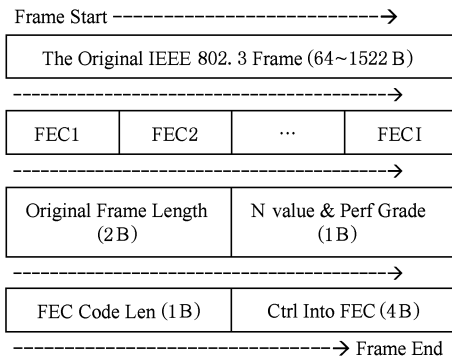


Fig. 2 Frame structure of the FEC method.

图2 FEC方法采用的帧结构

但是,最后一个数据节很可能不足 32 B,需作特别处理,这里采用假填充法,意思是编码解码时仍然把最后一个数据节当成 32 B 处理,不足部分以全 0 填充,以此生成本数据节的纠错码。但填充部分并不作为以太网数据包的一部分,也不会从以太网端口发送出去。

“Original Frame Length”字段的值等于 L 。用于确定数据包中纯信息域(即 FEC 处理前的 IEEE802.3 帧)长度。

n 值由“N value & Perf Grade”字段的高半字节 H_h 表示, H_h 取值为 0 时,表示数据包不包含 FEC 码,但具有控制信息,以传送信道性能信息。

信道性能由“N value & Perf Grade”字段的低半字节 H_l 表示,表示从远端以太网口到本地以太网口的信道单向性能,由本地以太网口检测远端以太网口发送的数据包,通过计算得到,用来表示误包率的等级。

“FEC Code Len”字段:其值等于原 IEEE 802.3 帧划分成的数据节数,也就是前述的 I 。设置这样一个字段结合“Original Frame Length”字段可以更准确地判断数据包的类型(是普通以太网数据包还是 FEC 数据包),也方便计算整个数据包的长度等信息。

“Ctrl Into FEC”字段的长度为 4B,其目的为控制信息纠错。控制信息包含“Original Frame Length”,“N value & Perf Grade”,“FEC Code Len”等等。

帧结构里包含的控制信息非常重要,因此更需保证其传送正确性,因此控制信息字段需要独立的 FEC 机制确保可靠性。

4 FEC 性能分析

由于信道的噪声分布是随机的,各数据位出错

概率可视为独立同分布,也就是说各数据位出错概率等于物理信道的误码率 E 。那么每个数据节出现 j 位错误的概率 $P(j)$ 符合二项分布,即

$$P(j) = C_{(8n)}^j E^j (1-E)^{(8n-j)}, \quad (7)$$

其中, n 是数据节包含的字节数, $8n$ 则是数据节包含的位数。

采用 $R-S(n, k)$ 编码体系,当每个数据节的误字节数小于等于 t (见式(4)),则该数据节可以有效纠错。以 j 表示一个数据节出错的位数,以 m 表示数据节出错的字节数。当 $j \leq t$ 时,则 $m \leq t$ 一定成立,那么该数据节一定可以纠错。当 $j > 8 \times t$ 时,则 $m \leq t$ 一定不成立,那么该数据节的错误不能够修正。当 $t < j \leq 8 \times t$ 时,如果 j 个错误都在 t 个以内的字节上,则能够纠错,反之则不能够修正错误。

由式(7)可知:一个数据节出现 $j \leq t$ 个错误位的概率 $P(j \leq t)$ 为

$$P(j \leq t) = \sum_{j=0}^t C_{8n}^j E^j (1-E)^{(8n-j)}. \quad (8)$$

一个数据节出现 $t < j \leq 8 \times t$ 个错误位的概率 $P(t < j \leq 8 \times t)$ 为

$$P(t < j < 8t) = \sum_{j=t+1}^{8t-1} C_{8n}^j E^j (1-E)^{(8n-j)}. \quad (9)$$

当发生 $t < j \leq 8 \times t$ 时, j 个错误都在 t 个以内的字节上的概率 $P(m < t)$ 为

$$P(m \leq t) = \frac{C_n^j C_{8t}^j}{C_{8n}^j}. \quad (10)$$

由式(8)~(10)可以得到,每个数据节能够纠错的概率 P_{corr} 为

$$P_{\text{corr}} = \sum_{j=0}^t C_{8n}^j E^j (1-E)^{(8n-j)} + \sum_{j=t+1}^{8t-1} C_n^j C_{8t}^j E^j (1-E)^{(8n-j)}. \quad (11)$$

由于数据包包含 I 个数据节,则数据包能正确传送的概率 P_{fc} 为

$$P_{fc} = (P_{\text{corr}})^I. \quad (12)$$

根据式(11)(12)可以得到图 3,它表示 FEC 纠错后的包正确传送率 P_{fc} 、实际物理信道丢包率 P_l 以及数据节总长度 n (包含净信息和纠错码)的关系。

图 3 中, X, Y 和 Z 轴分别表示裸物理信道(无 FEC 机制)丢包率 P_l 、数据节总长度 n 和采用 FEC 机制后的包正确传送率 P_{fc} 。

显然采用 FEC 机制后,信道的传送性能有相当大的提高。

FEC 虽然最大程度地满足基于正确率优先的

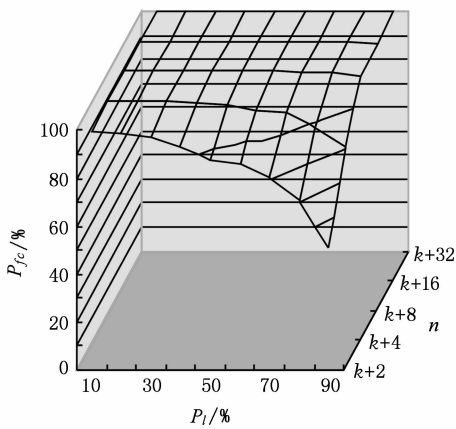


Fig. 3 Channel performance after using the FEC method.

图3 采用FEC机制后的信道性能

设计目标,但由于FEC码需要占用信道带宽,那么基于通过率优先的设计目标情况下,则该FEC不一定最优。

这里分析FEC带来的开销情况,以 P_i 表示有效信息率,它可以定义为原以太网信息帧所占物理信道比特数与对应的FEC信息帧所占信道比特数的百分比。以太网帧间隔和前导码为20 B或160 b,那么结合式(6)和图(2)的帧结构可以得到:

$$P_i = \frac{L + 20}{\left\lceil \frac{L}{32} \right\rceil \times n + 20 + 8} \times 100\% \quad (13)$$

由于 $64 \leq L \leq 1518^{[7]}$, $2 + k \leq n \leq 32 + k$, $k = 32$,根据式(13)可得到 $38.6\% \leq P_i \leq 92.5\%$,其中 P_i 最小值发生在 $L = 65$, n 取最大值时,而 P_i 最大值发生在 $L = 1504$, n 取最小值时。可见FEC带来的开销的影响不可忽视。

但FEC的信息长度可以根据检测到的物理信道性能情况自适应设置,信道性能良好时 n 取较小值,带来的开销就可以很小。另一方面,运营级以太网应用中无论是业务接入层,还是承载层,常常有大量剩余带宽可利用,所以应用意义上带来一定开销并无问题。再则,运营级以太网需要对TDM和OAM完善支持,可靠性是首要的前提,否则传送大量的错包是没有意义的。因此采用合适的FEC机制仍然具有很高的应用价值。

5 实现和测试

根据上述方法设计了千兆运营级以太网设备来验证该FEC方法,该设备的接口都是千兆光口。在PHY和网络处理器之间用FPGA(field-

programmable gate array)实现上述的FEC机制。测试网络由6台该运营级以太网设备(依次以 F_1, F_2, \dots, F_6 表示)组成,用光纤连接成链形结构,每段光纤使用光衰减器以产生一定程度的误码。测试仪的一个千兆光口作为输入与 F_1 的一个千兆光口连接, $F_2 \sim F_6$ 分别有一个千兆光口作为输出与测试仪的对应光口连接。

本次测试的测试条件一共有5类(5个测试集):FEC配置信息、物理信道误码率、信道路径长度、以太网数据包长和测试流量。

由上述的测试条件组合可得到几千个测试用例,实际测试选取了部分特征测试用例进行验证,但选取测试用例时已确保每个测试集的所有成员都被引用到。

图4表示的是FEC配置为 $k+16$ 、信道路径为 $F_1 \rightarrow F_2$ 、测试流量为500 MB时的测试结果。 X 轴表示的是 F_1 和 F_2 之间的光信道的误码率, Y 轴表示的是包正确传送率。显然,在采用FEC机制的情况下,传送性能有明显的改观。

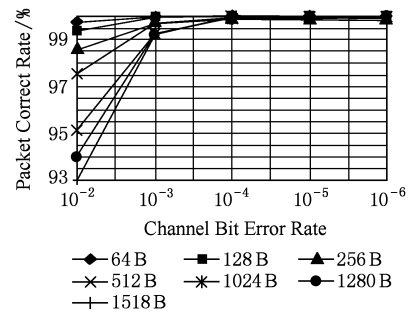


Fig. 4 Test result of the FEC method.

图4 对FEC方法的测试结果

测试结果还表明,在无FEC机制的情况下,传送性能随信道路径的长度增加而迅速下降。在采用适当FEC机制的情况下,传送性能则基本与信道路径的长度无关。

6 结论

运营级以太网对物理信道的可靠性提出了更高的要求,但也不能通过改变物理介质来满足要求,那样,成本的急剧上升将使以太网丧失其固有的优势。

可以通过FEC方法改善物理信道性能来满足要求。使用FEC方法,现有以太网物理信道可以充分利用,而可靠性将大大提高,而且由于采用自适应机制,带来的额外开销也较小。

实验表明,以以太网数据包为单位的自适应

FEC 方法对信道性能的改善是非常明显的, 应用可行性得到确认.

参 考 文 献

- [1] Stuart E, Haidar C, William B, *et al.* Carrier Ethernet: A reality check [C]//Optical Fiber Communication and the National Fiber Optic Engineers Conference. Anaheim: Institute of Electrical and Electronics Engineers (IEEE), 2007: 1-6
- [2] International Telecommunication Union, ITU-T G. 707/Y. 1322. Network node interface for the synchronous digital hierarchy (SDH) [S]. Geneva: ITU Telecommunication Standardization Sector, 2000
- [3] International Telecommunication Union, ITU-T G. 709/Y. 1331. Interfaces for the optical transport network (OTN) [S]. Geneva: ITU Telecommunication Standardization Sector, 2001
- [4] Howard K, Nanda R, Imran H. Performance improvement in high capacity, ultra-long distance, WDM systems using forward error correction codes [C]//Proc of Optical Fiber Communication Conference. Baltimore: Institute of Electrical and Electronics Engineers (IEEE), 2000: 274-276
- [5] Li L, Li L. Performance analysis of forward error correction schemes for wireless ATM systems in a Rician fading channel [J]. Journal of Chinese Institute of Communications, 1999, 20(S1): 293-298 (in Chinese)
(李立忠, 李乐民. ATM 在无线衰落信道上的前向纠错技术及其性能分析[J]. 通信学报, 1999, 20(S1): 293-298)
- [6] Shan Y, Chai Q. Forward error control schemes and utility in real-time networked multimedia applications on the Internet [J]. Journal of Computer Research and Development, 2001, 38(7): 816-820 (in Chinese)

(单玉峰, 柴乔林. 基于 Internet 的实时多媒体应用中前向纠错控制策略研究[J]. 计算机研究与发展, 2001, 38(7): 816-820)

- [7] IEEE Computer Society. IEEE Std 802. 3. Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications [S]. New York: IEEE Standards Board, 2003
- [8] IEEE Computer Society. IEEE Std 802. 3z. Media access control (MAC) parameters, physical layer and management parameters for Gb/s operation [S]. New York: IEEE Standards Board, 1998
- [9] IEEE Computer Society. IEEE Std 802. 3ae. Media access control (MAC) parameters, physical layers, and management parameters for 10 Gb/s operation [S]. New York: IEEE Standards Board, 2001
- [10] Suman M, Daniel I, Andrei I. Instruction set extensions for Reed-Solomon encoding and decoding [C] //The 16th Int Conf on Application-Specific Systems. Samos, Greece: Institute of Electrical and Electronics Engineers (IEEE), 2005: 364-369



Dai Jinyou, born in 1968. Ph. D. candidate. His current research interests include multi-service Ethernet and IP network.

戴锦友, 1968 年生, 博士研究生, 高级工程师, 主要研究方向为多业务以太网和 IP 互连网络。



Yu Shaohua, born in 1962. Ph. D., professor, and Ph. D. supervisor. His main research interests include IP network and MAN.

余少华, 1962 年生, 博士, 教授级高工, 博士生导师, 主要研究方向为 IP 互连网络和城域网。

Research Background

Ethernet is one of the widest-applied network technologies, and carrier Ethernet brings forward higher request for reliability of the Ethernet. On the other hand, Ethernet's port speed and transport distance have been extended hundreds of times during the past several decades. The present physical channel can't guarantee error-free transporting the Ethernet data packet. So, just like the mainstream transport technology such as SDH, WDM and ATM, it is of necessity for Ethernet to adopt the forward error correction method. This work is partially supported by the National Natural Science Foundation of China under grant Nos. 2003AA121110 and 2005AA121410, and the Electronic Foundation Project "Metro Area Network Multi-Service Ring(MSR) Technology and the Method for Bit-Based Ethernet over SDH(EOS)".