電腦網路 ch2.4

2.4 Ethernet (IEEE 802.3)

2.4以太網（IEEE 802.3）

Originally proposed by Bob Metcalfe in 1973,

最初由Bob Metcalfe於1973年提出，

Ethernet was once one of the competitors of the LAN technology,

以太網曾經是LAN技術的競爭對手之一，

and has been the winner. In over 30 years,

並一直是贏家。在30多年來，

Ethernet has been reinvented many times to accommodate new demands,

以太網已經進行了多次改造，以適應新的需求，

resulting in the large IEEE 802.3 Standard,

產生了大型的IEEE 802.3標準，

and the evolution continues well into the future.

進化將持續到未來。

We introduce readers to appreciate the evolution and philosophy of Ethernet,

我們向讀者介紹了以太網的發展和原理，

and also brief the hot topics currently under development.

並簡要介紹當前正在開發的熱門話題。

2.4.1 Ethernet Evolution: A Big Picture

2.4.1以太網演進：概況

As the title of the standard, "Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specification," suggests,

作為該標準的標題，"帶有衝突檢測的載波偵聽多路訪問（CSMA / CD）訪問方法和物理層規範"建議：

Ethernet is most distinguished from other LAN technologies,

以太網與其他局域網技術最有區別，

such as Token Bus and Token Ring, by its medium access method. A lab at Xerox gave birth to the technology in 1973,

例如令牌總線和令牌環，通過其媒體訪問方法。施樂公司的一個實驗室於1973年誕生了該技術，

which was later standardized by DEC, Intel and Xerox in 1981,

後來由DEC，Intel和Xerox於1981年標準化，

known as the DIX Ethernet. Although this standard bore little resemblance to the original design at Xerox,

被稱為DIX以太網。儘管此標準與Xerox的原始設計幾乎沒有相似之處，

the essence of CSMA/CD was preserved. In 1983,

CSMA / CD的本質得以保留。 1983年

the IEEE 802.3 Working Group approved a standard based on the DIX Ethernet with only insignificant changes.

IEEE 802.3工作組批准了基於DIX以太網的標準，僅做了很小的改動。

This standard becomes the well known IEEE 802.3 Standard. Since Xerox relinquished the trademark name "Ethernet",

該標準成為眾所周知的IEEE 802.3標準。由於Xerox放棄了商標名稱" Ethernet"，

distinction between the two terms no longer exists when we refer to the Ethernet and the IEEE 802.3 Standard. In fact, the IEEE 802.3 Working Group has been leading the Ethernet development as of its first version of the standard.

當我們引用以太網和IEEE 802.3標準時，這兩個術語之間不再存在區別。實際上，從標準的第一個版本開始，IEEE 802.3工作組就一直領導著以太網的發展。

Figure 2.10 illustrates the milestones in the Ethernet standards. It has experienced several significant revisions during the past 30 years. We list the major trends below.

圖2.10說明了以太網標準中的里程碑。在過去的30年中，它經歷了幾次重大修訂。我們在下面列出了主要趨勢。

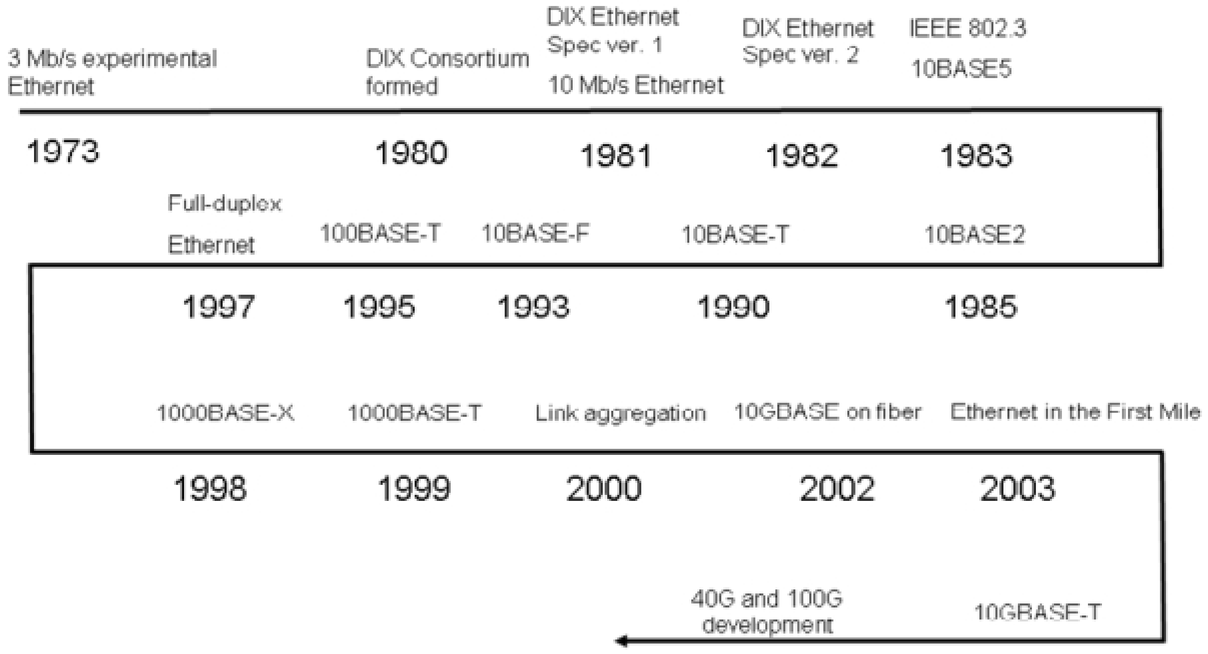


Figure 2.10 Milestones in the Ethernet Standards.

From low to high speed: Starting from a prototype running at 3 Mbps,

從低速到高速：從以3 Mbps運行的原型開始，

Ethernet has grown up to 10 Gbps – a boost of more than 3000 times in speed.

以太網已增長到10 Gbps –速度提高了3000倍以上。

An ongoing work (IEEE 802.3ba) to further boost the data rate up to 40 Gbps and 100 Gbps has started. An astonishing development as it is, the technology still remains cheap, making it widely accepted around the world. Ethernet has been built-in on almost every motherboard on desktop computers and laptops. We would be sure that Ethernet will be ubiquitous for wired connectivity.

正在進行的一項工作（IEEE 802.3ba）將數據速率進一步提高到40 Gbps和100 Gbps。儘管它是一個驚人的發展，但它的價格仍然很便宜，使其在世界範圍內被廣泛接受。以太網已經內置在台式機和筆記本電腦的幾乎所有主板上。我們確信以太網將在有線連接中無處不在。

From shared to dedicated media: The original Ethernet runs on a bus topology of coaxial cables.

從共享媒體到專用媒體：原始以太網運行在同軸電纜的總線拓撲上。

Multiple stations share the bus with the CSMA/CD MAC algorithm where collisions are normal. As of the development of 10BASE-T,

在衝突正常的情況下，多個站使用CSMA / CD MAC算法共享總線。隨著10BASE-T的發展，

dedicated media between two devices becomes the majority. Dedicated media are necessary to the later development of full-duplex Ethernet. Full-duplex allows both stations to transmit over the dedicated media simultaneously,

兩台設備之間的專用媒體成為主流。專用媒體對於全雙工以太網的後續開發是必需的。全雙工允許兩個站同時通過專用媒體進行傳輸，

which in effect doubles the bandwidth!

實際上，帶寬增加了一倍！

Form LAN to MAN and WAN: Ethernet was well known as a LAN technology.

從LAN到MAN和WAN：以太網是眾所周知的LAN技術。

Two factors help the technology move toward the MAN and WAN markets.

有兩個因素有助於該技術向MAN和WAN市場發展。

The first is the cost. Ethernet has low cost in implementation for its simplicity.

首先是成本。以太網由於其簡單性而實現成本較低。

It takes less pain and money in interoperability if the MAN and WAN are also Ethernet.

如果城域網和廣域網也是以太網，則可減少互操作性的痛苦和金錢。

The second comes from full duplex, which eliminates the need of CSMA/CD and thus lifts the distance restriction due to it.

第二個來自全雙工，這消除了對CSMA / CD的需求，因此取消了對CSMA / CD的距離限制。

The data can be transmitted as far as a physical link can reach.

數據可以傳輸到物理鏈路可以到達的範圍。

Richer medium: The term “ether” was once thought of as the medium to propagate electromagnetic waves through space.

更豐富的介質：“醚”一詞曾經被認為是通過空間傳播電磁波的介質。

Although Ethernet never uses ether to transmit data,

儘管以太網從不使用以太網來傳輸數據，

it does carry messages onto a variety of media: coaxial cables, twisted pairs and optical fibers.

它確實將消息傳遞到各種介質上：同軸電纜，雙絞線和光纖。

“Ethernet is Multimedia!” -- The amusing words by Rich Seifert in his book Gigabit Ethernet best depict the scenario.

“以太網就是多媒體！” -Rich Seifert在他的書《千兆以太網》中有趣的話最能說明這種情況。

Table 2.6 lists all the 802.3 family members in terms of speed and media.

表2.6從速度和媒體方面列出了所有802.3系列成員。

Table 2.6 The 802.3 family

medium / speed	Coaxial cable	Twisted pairs	Fiber
under 10 Mb/s		1BASE5 (1987) 2BASE-TL (2003)	
10 Mb/s	10BASE5 (1983) 10BASE2 (1985) 10BROAD36 (1985)	10BASE-T (1990) 10PASS-TS (2003)	10BASE-FL (1993) 10BASE-FP (1993) 10BASE-FB (1993)
100 Mb/s		100BASE-TX (1995) 100BASE-T4 (1995) 100BASE-T2 (1997)	100BASE-FX (1995) 100BASE-LX/BX10 (2003)
1 Gb/s		1000BASE-CX (1998) 1000BASE-T (1999)	1000BASE-SX (1998) 1000BASE-LX (1998) 1000BASE-LX/BX10 (2003) 1000BASE-PX10/20 (2003)
10 Gb/s		10GBASE-T (2006)	10GBASE-R (2002) 10GBASE-W (2002) 10GBASE-X (2002)

Not all members are commercially successful. For example,

並非所有成員在商業上都成功。例如，

100BASE-T2 has never been a commercial product.

100BASE-T2從來都不是商業產品。

In contrast, some are so successful that almost everybody can find a Network Interface Card (NIC) of 10BASE-T or 100BASE-TX behind a computer on a LAN.

相比之下，有些產品是如此成功，以至於幾乎每個人都可以在LAN上的計算機後面找到10BASE-T或100BASE-TX的網絡接口卡（NIC）。

Most new motherboards for desktop computers come with an Ethernet interface of 100BASE-TX or 1000BASE-T nowadays.

如今，大多數台式機新主板都帶有100BASE-TX或1000BASE-T以太網接口。

The number in the parentheses indicates the approval year of the specification by the IEEE.
括號中的數字表示IEEE批准該規範的年份。

The Ethernet Nomenclature
以太網命名法
Ethernet is rich in its physical specification, as presented in Table 2.6. The notation follows the format {1/10/100/1000/10G}{BASE/BROAD/PASS}[-]phy.
以太網具有豐富的物理規格，如表2.6所示。該表示法遵循{1/10/100/1000 / 10G} {BASE / BROAD / PASS} [-] phy格式。
The first item is the speed. The second item depends on whether the signaling is baseband or broadband. Almost all Ethernet signaling is baseband, except the old 10BROAD36 and 10PASS-TS.
第一項是速度。第二項取決於信令是基帶還是寬帶。除了舊的10BROAD36和10PASS-TS，幾乎所有的以太網信令都是基帶。
The third item is the maximum length in units of 100m in the beginning.
第三項是開頭的最大長度，以100m為單位。
No dash was between the second and the third item. The convention had later been changed to indicate the physical specifications,
第二和第三項之間沒有破折號。後來對約定進行了更改以指示物理規格，
such as medium type and signal encoding, and a dash is located between the second and the third item.
（例如媒體類型和信號編碼），並且在第二個和第三個項目之間放置一個破折號。

2.4.2 The Ethernet MAC
2.4.2以太網MAC
Ethernet Framing, Addressing, and Error control The 802.3 MAC sublayer is the medium-independent part of the Ethernet.
以太網成幀，尋址和錯誤控制802.3 MAC子層是以太網的與介質無關的部分。
Along with the Logical Link Control (LLC) sublayer specified in IEEE 802.2, they compose the data-link layer in the OSI layer model. The functions associated with the MAC sublayer include data encapsulation and media access control,
它們與IEEE 802.2中指定的邏輯鏈路控制（LLC）子層一起構成OSI層模型中的數據鏈路層。與MAC子層相關的功能包括數據封裝和媒體訪問控制，
and those for the LLC sublayer are intended to be common interfaces for Ethernet,
和LLC子層的接口旨在用作以太網的通用接口，
Token Ring, WLAN and so on. Linux also implements the latter part in functions like bridge configuration, since the configuration frames are specified in the LLC format (See Section 2.7).
令牌環，WLAN等。 Linux還通過網橋配置等功能實現了後一部分，因為配置框架以LLC格式指定（請參見2.7節）。
Figure 2.11 presents the untagged4 Ethernet frame.
圖2.11展示了未標記的4以太網幀。
Through the frame format, we first introduce framing, addressing and error control,
通過幀格式，我們首先介紹成幀，尋址和錯誤控制，
and leave issues of medium access control and flow control later.
並在以後保留介質訪問控制和流控制的問題。

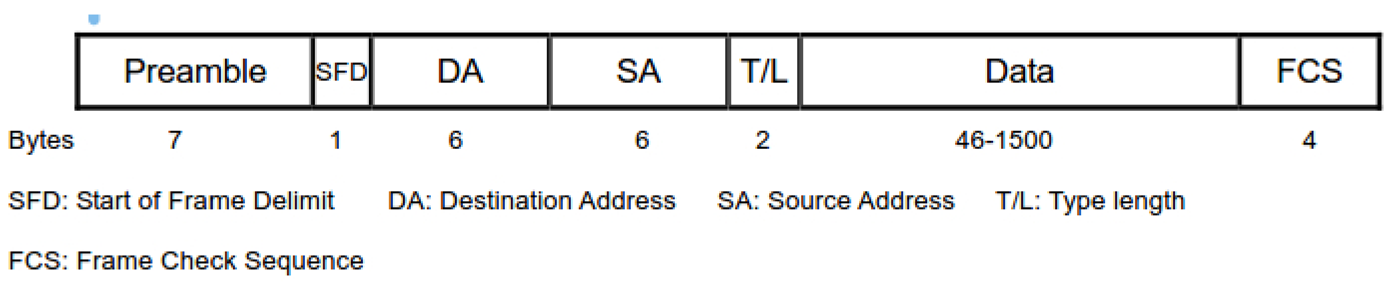


Figure 2.11 Ethernet frame format.

Preamble: This field synchronizes the physical signal timing on the receiver side.

前導：此字段同步接收器側的物理信號時序。

Its value is fixed at 1010....10 in the transmission order5, totally 56 bits long.

它的值在傳輸順序5中固定為1010....10，共56位長。

Note that the frame boundary may be marked by special physical encoding or the presence (absence) of signal, depending on the PHY.

注意，取決於PHY，可以通過特殊的物理編碼或信號的存在（不存在）來標記幀邊界。

For example, 100BASE-X Ethernet converts the first byte of the Preamble, /1010/1010/, into two special code groups /J/K/ of the value /11000/10001/ using 4B/5B encoding.

例如，100BASE-X以太網使用4B / 5B編碼將前同步碼的首個字節/ 1010/1010 /轉換為兩個特殊的代碼組/ J / K /，其值是/ 11000/10001 /。

The 4B/5B encoding converts 1010 (in the transmission order) to 01011 for normal data without ambiguity. Similarly,

對於正常數據，4B / 5B編碼將1010（按傳輸順序）轉換為01011，沒有歧義。同樣，

100BASE-X appends two special code groups /T/R/ of the value /01101/10001/ to mark a frame end.

100BASE-X附加兩個特殊代碼組/ T / R /，其值為/ 01101/10001 /，以標記幀結束。

SFD: This field indicates the start of the frame with the value 10101011 in the transmission order.

SFD：此字段指示傳輸順序中值為10101011的幀的開始。

Historically, the DIX Ethernet Standard specified an 8-byte preamble with exactly the same value as the first two fields in an 802.3 frame,

從歷史上看，DIX以太網標準指定了一個8字節的前同步碼，其值與802.3幀中的前兩個字段完全相同，

but they are different only in nomenclature.

但是它們僅在術語上有所不同。

DA: This field is the 48-bit destination MAC address in the format introduced in Subsection 2.1.2.

DA：此字段是第2.1.2節介紹的格式的48位目標MAC地址。

SA: This field is the 48-bit source MAC address.

SA：此字段是48位源MAC地址。

Type/Length: This field has two meanings for historical reasons.

類型/長度：出於歷史原因，此字段具有兩個含義。

The DIX Standard specified the field to be a code of payload protocol type, say

DIX標準將該字段指定為有效載荷協議類型的代碼，例如

IP, while the IEEE 802.3 Standard specified the field to be the length of the data field6 and left the protocol type to the LLC sublayer.

IP，而IEEE 802.3標準將字段指定為數據字段的長度6，並將協議類型留給LLC子層。

The 802.3 Standard later (in 1997)

後來的802.3標準（1997年）

approved the type field, resulting in the dual roles of this field today.

批准了type字段，從而導致了該字段的雙重作用。

The way to distinguish is simple. Because the data field is never larger than 1500 bytes,

區分的方法很簡單。由於數據字段永遠不會大於1500個字節，

a value less than or equal to1500 means a length field. A value larger than or equal

小於或等於1500的值表示長度字段。大於或等於的值

to 1536 (=0x600) means a type field. Although the purposes are different,

到1536（= 0x600）表示類型字段。儘管目的不同，

they can co-exist due to the distinction. The values in between are intentionally not defined.

由於兩者的不同，它們可以共存。中間的值有意未定義。

Most frames use it as the type field because the dominating network layer protocol, IP, uses it as the type field.
大多數幀使用它作為類型字段，因為主要的網絡層協議IP使用它作為類型字段。

Data: This field carries the data varying from 46 to 1500 bytes.
數據：此字段承載的數據范圍從46到1500字節不等。
FCS: This field carries a 32-bit CRC code as a frame check sequence.
FCS：此字段攜帶32位CRC碼作為幀校驗序列。
If the receiver finds an incorrect frame,
如果接收方發現幀不正確，
it silently discards the frame. The transmitter knows nothing about whether the frame is discarded.
它無聲地丟棄了幀。發射機對是否丟棄幀一無所知。
The responsibility of a retransmission is left to upper-layer protocols, such as TCP.
重傳的責任留給了上層協議，例如TCP。
This approach is quite efficient because the transmitter does not need to wait an acknowledgement for the next transmission.
這種方法非常有效，因為發射機不需要等待下一次傳輸的確認。
The error is not a big problem here because the bit error rate is assumed to be very low in the Ethernet physical layer.
這裡的錯誤不是大問題，因為在以太網物理層中假定誤碼率非常低。
The frame size is variable. We often exclude the first two fields and say a minimum Ethernet frame has 64 (=6+6+2+46+4) bytes and a maximum Ethernet frame has 1518 (=6+6+2+1500+4) bytes.
幀大小是可變的。我們經常排除前兩個字段，並說最小的以太網幀有64（= 6 + 6 + 2 + 46 + 4）個字節，最大的以太網幀有1518（= 6 + 6 + 2 + 1500 + 4）個字節。
One may think the maximum length is not long enough so that the header overhead is larger, compared with Token Ring or FDDI. We shall analyze the Ethernet efficiency in Section 2.6.
人們可能會認為最大長度不夠長，因此與令牌環或FDDI相比，報頭開銷更大。我們將在2.6節中分析以太網效率。

Medium Access Control: Transmission and Reception Flow
媒體訪問控制：發送和接收流程
We now show how a frame is transmitted and received, and you shall see
現在我們展示如何發送和接收幀，您將看到
how CSMA/CD works in great details. Figure 2.12 shows what role the MAC
CSMA / CD的工作原理非常詳細。圖2.12顯示了MAC的作用
sublayer plays during the frame transmission and reception.
子層在幀發送和接收期間播放。

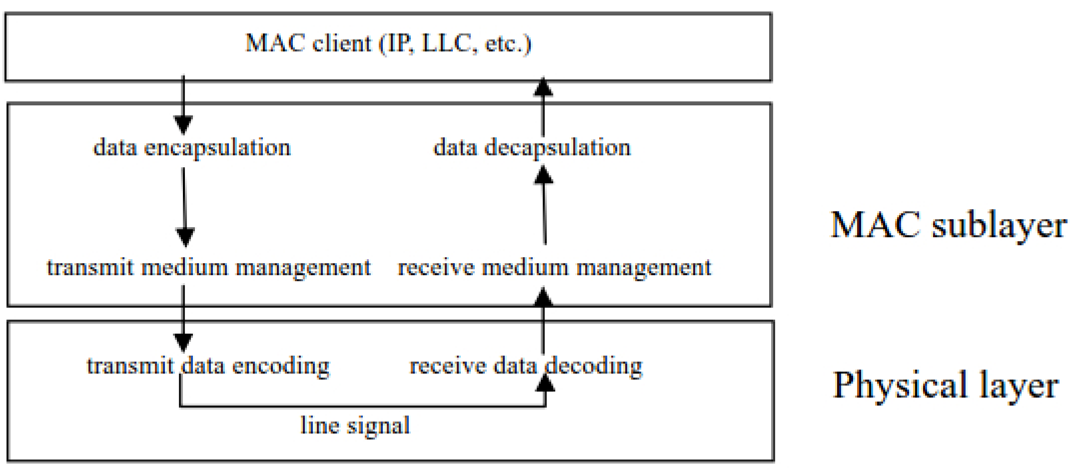


Figure 2.12 Frame transmission and reception.

CSMA/CD works in a simple way as its name implies.
顧名思義，CSMA / CD以一種簡單的方式工作。
With a frame to transmit, CSMA/CD senses the cable first.
有了要傳輸的幀，CSMA / CD首先檢測電纜。
If a carrier signal is sensed, i.e. the cable is busy, it continues sensing till idle, waits a small gap, and then transmits.
如果感測到載波信號，即電纜正忙，它將繼續感測到空閒，等待一小段間隙然後傳輸。
If a collision is detected during transmission, it jams the cable, aborts the transmission, and waits a random back-off time to retry.
如果在傳輸過程中檢測到衝突，則它將阻塞電纜，中止傳輸，並等待隨機的退避時間以重試。
Figure 2.13 presents the transmission flow. We list the exact procedure below. Note that on full-duplex links carrier sense and collision detection disappear effectively.
圖2.13給出了傳輸流程。我們在下面列出了確切的過程。請注意，在全雙工鏈路上，載波偵聽和衝突檢測會有效消失。

1. The MAC client (IP, LLC, etc.) asks for frame transmission.
1. MAC客戶端（IP，LLC等）請求幀傳輸。
2. The MAC sublayer prepends and appends MAC information (Preamble, SFD, DA, SA, type and FCS) to the data from the MAC client.
2. MAC子層將MAC信息（Preamble，SFD，DA，SA，類型和FCS）前置並附加到來自MAC客戶端的數據中。
3. In the half-duplex mode, the CSMA/CD method senses carrier to determine whether the transmission channel is busy. If so, the transmission is deferred until the channel is clear.
3.在半雙工模式下，CSMA / CD方法感測載波以確定傳輸信道是否繁忙。如果是這樣，則將傳輸推遲到清除信道為止。
4. Wait for a period of time called inter-frame gap (IFG). The time length is 96 bit times for all Ethernet types. The bit time is the duration of one bit transmission and thus the reciprocal of the bit rate. The IFG allows time for the receiver to do processing such as interrupts and pointer adjustment for incoming frames.
4.等待一段時間，稱為幀間間隔（IFG）。所有以太網類型的時間長度均為96位時間。比特時間是一位傳輸的持續時間，因此是比特率的倒數。 IFG允許接收器有時間進行處理，例如中斷和對傳入幀進行指針調整。
5. Start to transmit the frame.
5.開始傳輸幀。
6. In the half-duplex mode, the transmitter should keep monitoring if there is a collision during transmission. The monitoring method depends on the attached medium. Multiple transmissions on a coaxial cable result in higher absolute voltage levels than normal. For twisted pairs, a collision is asserted by perceiving received signal on the receive pair while transmitting.
6.在半雙工模式下，發送器應繼續監視發送過程中是否存在衝突。監視方法取決於所連接的介質。同軸電纜上的多次傳輸導致絕對電壓水平高於正常水平。對於雙絞線，通過在發送時感知接收對上的接收信號來聲明衝突。
7. In case of no collision during transmission, the frame is transmitted until done.
7.在傳輸過程中沒有衝突的情況下，將傳輸幀直到完成。

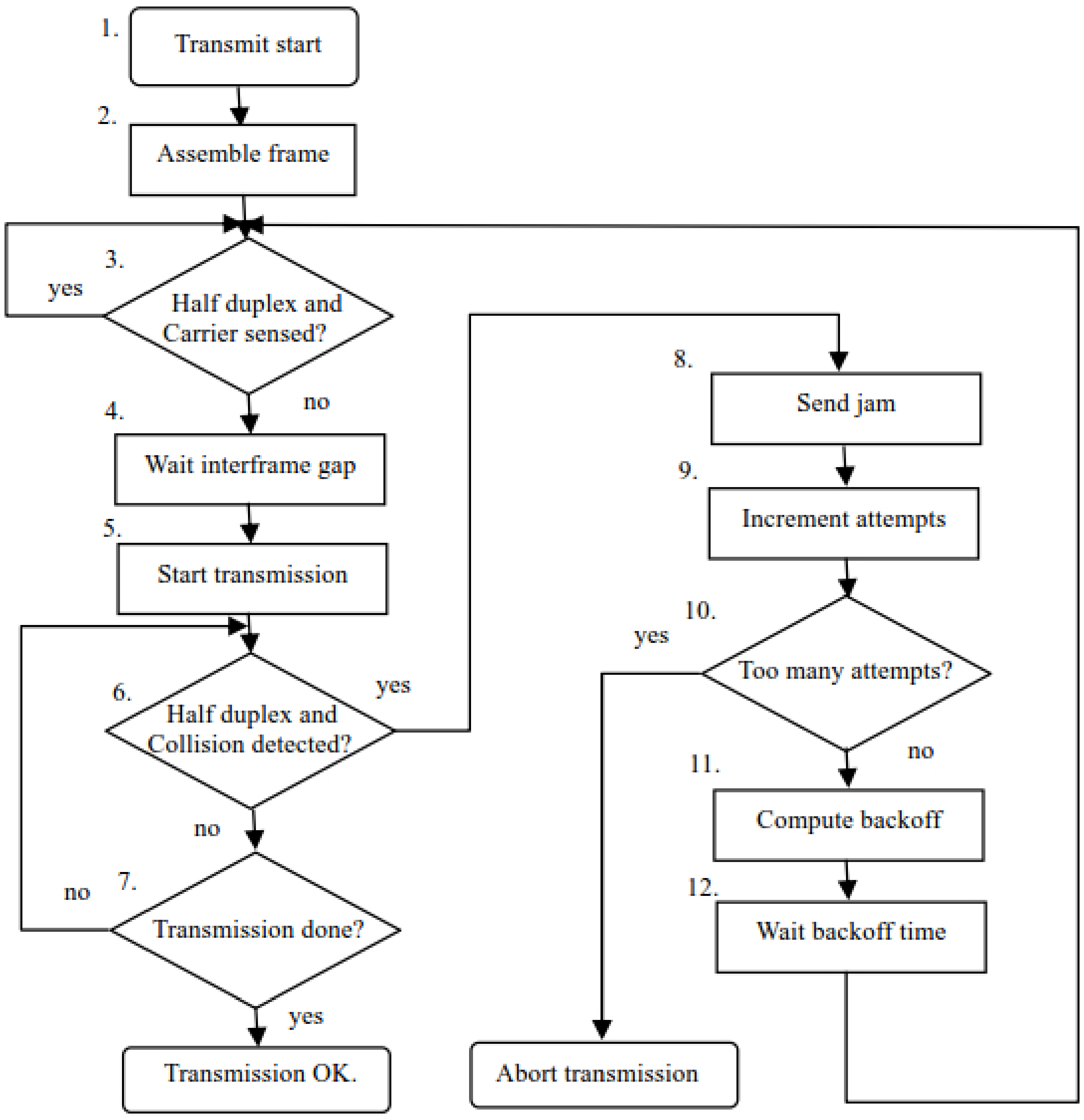


Figure 2.13 Frame transmission flow.

If a collision is detected in the half duplex mode, follow steps 8-12 to go on.

如果在半雙工模式下檢測到衝突，請按照步驟8-12繼續進行。

8. The transmitter transmits a 32-bit long jam signal to ensure the collision is long enough that all involved stations are aware of it. The pattern of the jam signal is unspecified. Common implementations are keeping transmitting 32 more data bits or using the circuit that generates the preamble to transmit alternating 1's and 0's.

8.發送器發送一個32位長的阻塞信號，以確保衝突足夠長，以使所有參與的站都知道它。卡紙信號的模式未指定。常見的實現方式是繼續發送32個以上的數據位，或使用生成前同步碼的電路交替發送1和0。

9. Abort the current transmission and attempt to schedule another transmission!

9.中止當前傳輸並嘗試安排另一個傳輸！

10. The maximum number of attempts to retransmit is 16. If still not able to transmit, abort the frame.

10.重新發送的最大嘗試次數為16。如果仍然無法發送，請中止該幀。

11. On an attempt to retransmit, a back-off time in units of slots is chosen randomly from the range of 0 to 2k-1, where k = min(n, 10) and n is the number of attempts. The range grows exponentially, so the algorithm is referred to as truncated binary exponential back-off. The duration of the slot time is 512 bit times for 10/100 Mb/s Ethernet and 4096 bit times for 1 Gb/s Ethernet. We shall talk about the reason when we discuss Gigabit Ethernet in Subsection 2.4.3.

11.在嘗試重傳時，從0到2k-1的範圍內隨機選擇以時隙為單位的退避時間，其中k = min（n，10），n為嘗試次數。範圍呈指數增長，因此該算法稱為截斷二進制指數補償。對於10/100 Mb / s以太網，時隙時間的持續時間為512位時間，對於1 Gb / s以太網，時隙時間的持續時間為4096位時間。我們將在2.4.3小節中討論千兆以太網的原因。

12. Wait for the back-off time and attempt to retransmit

12.等待退避時間，然後嘗試重新傳輸

Receiving a frame is much easier, where a sequence of checks is done on frame length (too short or too long), destination MAC address, FCS, and octet boundary before passing it to the MAC client. Figure 2.14 illustrates the reception flow. We list the procedure below.

接收幀要容易得多，在對幀長度（太短或太長），目標MAC地址，FCS和八位字節邊界進行一系列檢查之前，先將其傳遞給MAC客戶端。圖2.14說明了接收流程。我們在下面列出了過程。

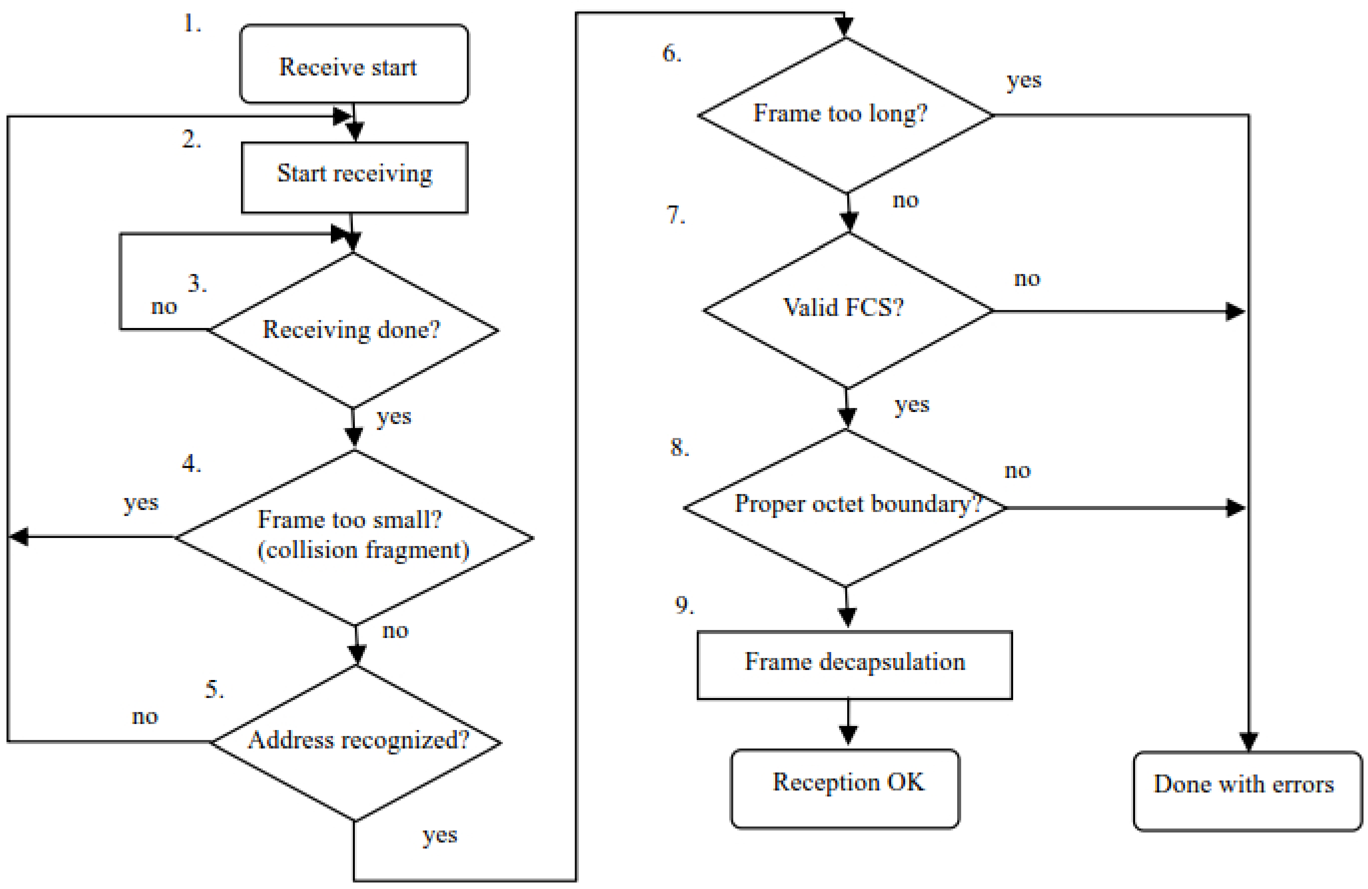


Figure 2.14 Frame reception flow.

1. The arrival of a frame is detected by the physical layer of the receiver.

1.幀的到達由接收器的物理層檢測。

2. The receiver decodes the received signal and passes the data, except the preamble and SFD, up to the MAC sublayer.

2.接收器解碼接收到的信號，並將除前導碼和SFD之外的數據傳遞到MAC子層。

3. The receiving process goes on as long as the received signal continues. When the signal ceases, the incoming frame is truncated to an octet boundary.

3.只要接收信號繼續，接收過程就繼續進行。當信號停止時，傳入的幀被截斷到八位字節邊界。

4. If the frame is too short (shorter than 512 bits), it is thought of as a collision fragment and dropped.

4.如果幀太短（短於512位），則將其視為衝突片段並丟棄。

5. If the destination address is not for the receiver, the frame is dropped.

5.如果目標地址不是用於接收方的，則丟棄該幀。

6. If the frame is too long, it is dropped and the error is recorded for management statistics.

6.如果幀太長，則會丟棄該幀，並記錄錯誤以進行管理統計。

7. If the frame has an incorrect FCS, it is dropped and the error is recorded.

7.如果框架的FCS不正確，則將其丟棄並記錄錯誤。

8. If the frame size is not an integer number of octets, it is dropped and the error is recorded.

8.如果幀大小不是八位位組的整數，則將其丟棄並記錄錯誤。

9. If everything is OK, the frame is de-capsulated and the fields are passed up to the MAC client.

9.如果一切正常，則將幀解封裝，並將字段傳遞給MAC客戶端。

Can Collision Cause Bad Performance?

碰撞會導致性能下降嗎？

The term collision sounds terrible! However, collision is part of the normal arbitration mechanism of CSMA/CD, instead of system malfunction. Collision can cause a garbled frame,

碰撞一詞聽起來很可怕！但是，衝突是CSMA / CD正常仲裁機制的一部分，而不是系統故障。碰撞會導致框架亂碼，

but it is not so bad if with collision detection.

但是如果有碰撞檢測，還算不錯。

A transmission can be stopped if a collision is detected. Before further analysis of wasted bit times due to a collision, we first answer a critical question:

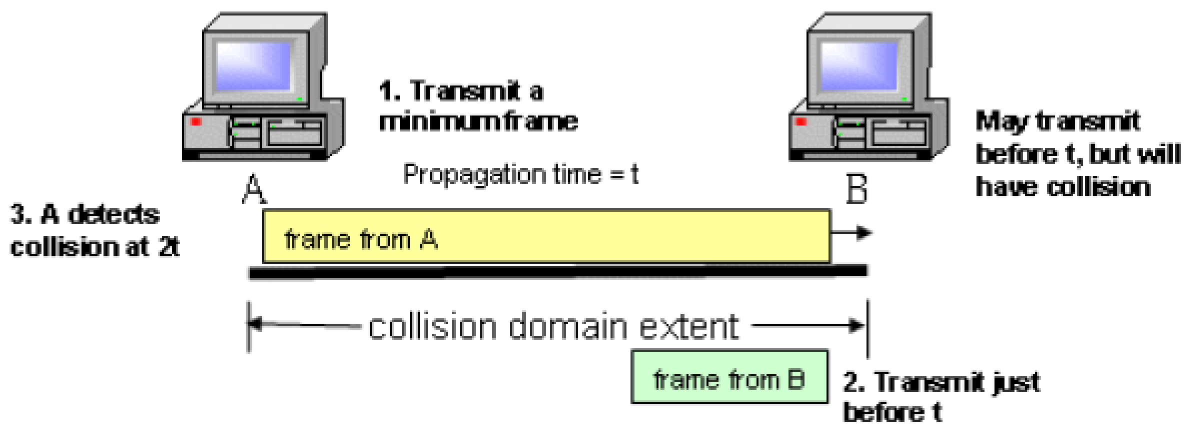
如果檢測到衝突，則可以停止傳輸。在進一步分析由於碰撞而浪費的比特時間之前，我們首先回答一個關鍵問題：

Where can a collision occur?

哪裡會發生碰撞？

We model the frame transmission in Figure 2.15.

我們在圖2.15中對幀傳輸進行建模。



Principle: round-trip time 2t < time to transmit a minimum frame

Figure 2.15 Collision detection with propagation delay.

Suppose station A transmits a minimum frame of 64 bytes,

假設站點A傳輸的最小幀為64字節，

and the propagation time before the frame arrives at station B is t. Even with carrier sense,

幀到達站B之前的傳播時間為t。即使有載體意識

station B is likely to transmit anytime before t. Further suppose station B transmits right at time t, which results in a collision. The collision takes another t to propagate back to station A.

B站可能在t之前的任何時間發送。進一步假設站B在時間t正確發送，這導致衝突。衝突需要再花費t才能傳播回站點A。

If station A finishes transmitting the minimum frame before the round-trip time, 2t, expires, it has no chance to invoke collision detection and is unable to schedule a retransmission, and thus the frame is lost.

如果站點A在往返時間2t到期之前完成了最小幀的發送，則它沒有機會調用衝突檢測並且無法安排重傳，因此幀丟失。

For CSMA/CD to function normally, the round-trip time should be less than the time to transmit a minimum frame, meaning the CSMA/CD mechanism limits the extent between two stations in a collision domain.

為了使CSMA / CD正常運行，往返時間應小於傳輸最小幀的時間，這意味著CSMA / CD機制會限制衝突域中兩個站點之間的範圍。

This limitation complicates the half-duplex Gigabit Ethernet design. We shall talk more about this issue when we

此限制使半雙工千兆位以太網設計複雜化。我們將在我們討論更多有關此問題的信息

introduce Gigabit Ethernet in Subsection 2.4.3. Because the minimum size is 64 bytes, it also means a collision must occur during the first 64 bytes of a frame under the distance limitation. If more than 64 bytes have been transmitted, the chance of collision has been ruled out due to carrier sense by other stations.

在2.4.3小節中介紹千兆以太網。由於最小大小為64個字節，因此這也意味著在距離限制下，幀的前64個字節期間必須髮生衝突。如果傳輸的字節數超過64個，則由於其他站的載波偵聽，已經排除了衝突的機會。

If we take the 32-bit jam into consideration, the actual number of bits in a frame that have been transmitted plus the jam cannot exceed 511 bits, as described in step 4 of the frame reception flow, as 512 bits (= 64 bytes) are the minimum length of a normal frame.

如果考慮到32位卡紙，則幀接收流程的第4步中所述，已發送的幀中的實際位數加卡紙不能超過511位（= 64字節）是正常幀的最小長度。

If exceeded, the receiver will think of these bits as a frame, rather than a collision fragment.

如果超出，則接收器會將這些位視為幀，而不是衝突片段。

Therefore, the maximum number of wasted bit times is 511 + 64 (from the preamble) + 96 (from the IFG) = 671.

因此，浪費的最大比特時間為511 + 64（來自前同步碼）+ 96（來自IFG）= 671。

This is only a small portion for a large frame. In addition, we must emphasize it is the worst case. For most collisions, they are detected during the preamble because the distance between two transmitting stations is not that far. In this case,

對於大框架來說，這只是一小部分。另外，我們必須強調這是最壞的情況。對於大多數衝突，在前導期間會檢測到它們，因為兩個發射站之間的距離並不那麼遠。在這種情況下，

the number of wasted bit times is only 64 (from the preamble) + 32 (from the jam) + 96 (from the IFG) = 192.

浪費的比特時間僅為64（從前同步碼開始）+ 32（從阻塞開始）+ 96（從IFG開始）= 192。

Maximum Frame Rate

How many frames can a transmitter (receiver) transmits (receives) in a second? This is an interesting question, especially when you design or analyze a packet processing device, say a switch, to know how many frames a second your device may need to process.

Frame transmission begins with a 7-byte Preamble and a 1-byte SFD. To reach the maximum number of frames per second, all frames should be of minimum size, i.e. 64 bytes. Do not forget the IFG of 12 bytes (= 96 bits) between two successive frame transmissions. Totally, a frame transmission occupies (7+1+64+12) x 8 = 672 bit times. In a 100 Mb/s system, the maximum number of frames per second is therefore 100 x 10^6 / 672 = 148,800. This value is referred to as the *maximum frame rate* for the 100 Mb/s link. If a switch has 48 interface ports, the aggregated maximum frame rate would be 148,800 x 48 = 7,140,400, i.e. over

7 millions!

Maximum Frame Rate

最大幀率

How many frames can a transmitter (receiver) transmits (receives) in a second?

發射器（接收器）在一秒鐘內可以發射（接收）多少幀？

This is an interesting question, especially when you design or analyze a packet processing device,

這是一個有趣的問題，尤其是在設計或分析數據包處理設備時，

say a switch, to know how many frames a second your device may need to process.

說一下交換機，以了解您的設備每秒可能需要處理多少幀。

Frame transmission begins with a 7-byte Preamble and a 1-byte SFD. To reach the maximum number of frames per second,

幀傳輸以7字節的前同步碼和1字節的SFD開始。為了達到每秒最大幀數，

all frames should be of minimum size, i.e. 64 bytes.

所有幀都應為最小大小，即64個字節。

Do not forget the IFG of 12 bytes (= 96 bits) between two successive frame transmissions.

不要忘記兩次連續幀傳輸之間的12個字節（= 96位）的IFG。

Totally, a frame transmission occupies (7+1+64+12) x 8 = 672 bit times. In a 100 Mb/s system,

總共，幀傳輸佔用（7 + 1 + 64 + 12）x 8 = 672位時間。在100 Mb / s的系統中，

the maximum number of frames per second is therefore 100 x 106 / 672 = 148,800.

因此，每秒最大幀數為100 x 106/672 = 148,800。

This value is referred to as the maximum frame rate for the 100 Mb/s link.

此值稱為100 Mb / s鏈路的最大幀速率。

If a switch has 48 interface ports,

如果交換機具有48個接口端口，

the aggregated maximum frame rate would be 148,800 x 48 = 7,140,400, i.e. over 7 millions!

總的最大幀速率將為148,800 x 48 = 7,140,400，即超過700萬！

Full-duplex MAC

全雙工MAC

Early Ethernet uses coaxial cables as the transmission media and forms a bus topology to connect stations. Twisted pairs have replaced most of them due to easier management.

早期的以太網使用同軸電纜作為傳輸介質，並形成總線拓撲來連接站點。由於易於管理，雙絞線已取代了大多數雙絞線。

A twisted pair cable connects a station and a concentration device, such as a hub or switch, and this star topology becomes dominant.

雙絞線電纜將工作站和集中設備（例如集線器或交換機）連接起來，這種星形拓撲變得占主導地位。

For popular 10BASE-T and 100BASE-TX, a wire pair in a twisted pair cable is dedicated to either transmitting or receiving 8 . A collision is thus defined by perceiving received signal on the receive pair while transmitting on the transmit pair.

對於流行的10BASE-T和100BASE-TX，雙絞線電纜中的線對專用於發送或接收8。因此，通過在接收對上接收到的信號同時在發射對上進行發送來定義衝突。

However, this is still inefficient. Since the medium is dedicated, why does it need “arbitration”?

但是，這仍然是無效的。由於媒體是專用的，為什麼需要“仲裁”？

In 1997, the IEEE 802.3x Task Force added full-duplex operation in Ethernet,

1997年，IEEE 802.3x工作隊在以太網中添加了全雙工操作，

i.e., transmission and reception can proceed at the same time.

即，發送和接收可以同時進行。

No carrier sense or collision detection is in the full duplex mode because they are not needed.

全雙工模式下沒有載波偵聽或衝突檢測，因為不需要它們。

There is no “multiple access” on a dedicated medium. Therefore, CS, MA,

專用介質上沒有“多重訪問”。因此，CS，MA，

and CD are all gone! Interestingly,

和CD都消失了！有趣的是

this is quite a significant change to Ethernet since Ethernet was known for its CSMA/CD.

由於以太網以其CSMA / CD而著稱，因此這對以太網是一個相當大的變化。

Three conditions should be satisfied to run full-duplex Ethernet:

運行全雙工以太網應滿足三個條件：

1. The transmission medium must be capable of transmitting and receiving on both ends without interference.

1.傳輸介質必須能夠在兩端進行傳輸和接收而不會受到干擾。

2. The transmission medium should be dedicated for exactly two stations, forming a point-to-point link.

2.傳輸介質應專用於恰好兩個站，形成點對點鏈接。

3. Both stations should be able to be configured in the full-duplex mode.

3.兩個工作站都應能夠配置為全雙工模式。

The IEEE 802.3 Standard explicitly rules out the possibility of running the full-duplex mode on a repeater hub. The bandwidth in the hub is shared, IEEE 802.3標準明確排除了在中繼器集線器上運行全雙工模式的可能性。集線器中的帶寬是共享的，

not dedicated. Three typical scenarios of full-duplex transmission are the station-to-station link,

不敬業。全雙工傳輸的三種典型情況是站到站鏈路，

the station-to-switch link, and the switch-to-switch link.

站點到交換機的鏈接以及交換機到交換機的鏈接。

These links need to be dedicated.

這些鏈接需要專用。

Full-duplex Ethernet in effect doubles the bandwidth between two stations.

實際上，全雙工以太網使兩個站點之間的帶寬增加了一倍。

It also lifts the distance limitation because of CSMA/CD.

由於CSMA / CD，它還解除了距離限制。

This is very important for high-speed and wide-area, as we shall discuss it in Subsection 2.4.3. Nowadays,

這對於高速和廣域非常重要，正如我們將在2.4.3小節中討論的那樣。如今，

virtually all Ethernet interfaces support full duplex.

實際上，所有以太網接口都支持全雙工。

Both interfaces can perform auto-negotiation to determine whether both parties support full duplex.

兩個接口都可以執行自動協商，以確定雙方是否都支持全雙工。

If so, both will operate in full duplex for higher efficiency.

如果是這樣，則兩者都將以全雙工運行以提高效率。

Ethernet Flow Control

以太網流量控制

Flow control in the Ethernet depends on the duplex mode.

以太網中的流控制取決於雙工模式。

In the half-duplex mode,

在半雙工模式下，

if the receiver cannot afford more incoming frames,

如果接收者負擔不起更多的傳入幀，

it could transmit a carrier,

它可以傳送一個載波，

say a series of 1010....10, on the shared medium until it can afford more frames. The transmitter will sense the carrier and defer its subsequent transmission.

在共享媒體上說一系列1010....10，直到它負擔得起更多幀為止。發射機將感知載波並推遲其後續傳輸。

This technique is called false carrier. Alternatively, the receiver can force a collision whenever any frame transmission is detected.

此技術稱為偽載波。可替代地，每當檢測到任何幀傳輸時，接收器都可以強制衝突。

This forces the transmitter to back off and reschedule its transmission.

這迫使發送器退回並重新安排其發送時間。

This technique is referred to as force collision. Both techniques are collectively called back pressure.

該技術稱為力碰撞。兩種技術統稱為背壓。

However, back pressure is void in full duplex mode because CSMA/CD is gone.

但是，由於CSMA / CD已消失，所以在全雙工模式下背壓無效。

IEEE 802.3 specifies a PAUSE frame to do flow control in the full duplex mode.

IEEE 802.3指定暫停幀以全雙工模式進行流控制。

The receiver explicitly sends a PAUSE frame to ask for a stop.

接收者顯式發送一個PAUSE幀以請求停止。

Upon receiving the PAUSE frame, the transmitter stops transmitting immediately.

接收到暫停幀後，發送器立即停止發送。

The PAUSE frame carries a field, pause_time, to tell the transmitter how long it should stop.

PAUSE幀帶有一個pause_time字段，以告知發送器應停止多長時間。

Since the pause time is not easy to estimate in advance, in practice,

由於暫停時間很難預先估算，因此在實踐中，

pause_time is set to be the maximum to stop the transmission and another PAUSE frame with pause_time = 0 is sent to the transmitter to tell it to resume when the receiver can accept more frames.
將pause_time設置為停止傳輸的最大值，然後將另一個pause_time = 0的PAUSE幀發送到發送器，以告訴它在接收器可以接受更多幀時恢復。

Flow control is optional in Ethernet. It can be enabled by the user or through auto-negotiation.
流量控制在以太網中是可選的。它可以由用戶啟用，也可以通過自動協商啟用。
IEEE 802.3 Standard provides an optional sublayer between MAC and LLC, namely MAC Control sublayer,
IEEE 802.3標准在MAC和LLC之間提供了一個可選的子層，即MAC控制子層，
which defines MAC Control frames to provide real-time manipulation of MAC sublayer operation.
它定義了MAC控制幀以提供MAC子層操作的實時操作。
The PAUSE frame is a kind of MAC Control frames.
暫停幀是一種MAC控制幀。

Open Source Implementation 2.5: CSMA/CD
開源實施2.5：CSMA / CD
CSMA/CD is part of the Ethernet MAC and most of the Ethernet MAC is implemented in hardware.
CSMA / CD是以太網MAC的一部分，大多數以太網MAC是在硬件中實現的。
An open source Ethernet example is available from OPENCORE (www.opencores.org),
OPENCORE（www.opencores.org）提供了一個開源以太網示例，
which consists of a synthesizable Verilog code.
由可綜合的Verilog代碼組成。
By synthesizable we mean the Verilog code is complete enough to be compiled, through a series of tools, into a circuit. It provides the implementation of the layer-2 protocol according to the IEEE specifications for the 10 Mbps and 100 Mbps Ethernet.
可綜合性是指Verilog代碼足夠完整，可以通過一系列工具編譯成電路。它根據IEEE規範為10 Mbps和100 Mbps以太網提供了第2層協議的實現。
Note that this open source implementation is the only hardware example in this text. All others are software.
請注意，此開源實現是本文中唯一的硬件示例。其他所有軟件。

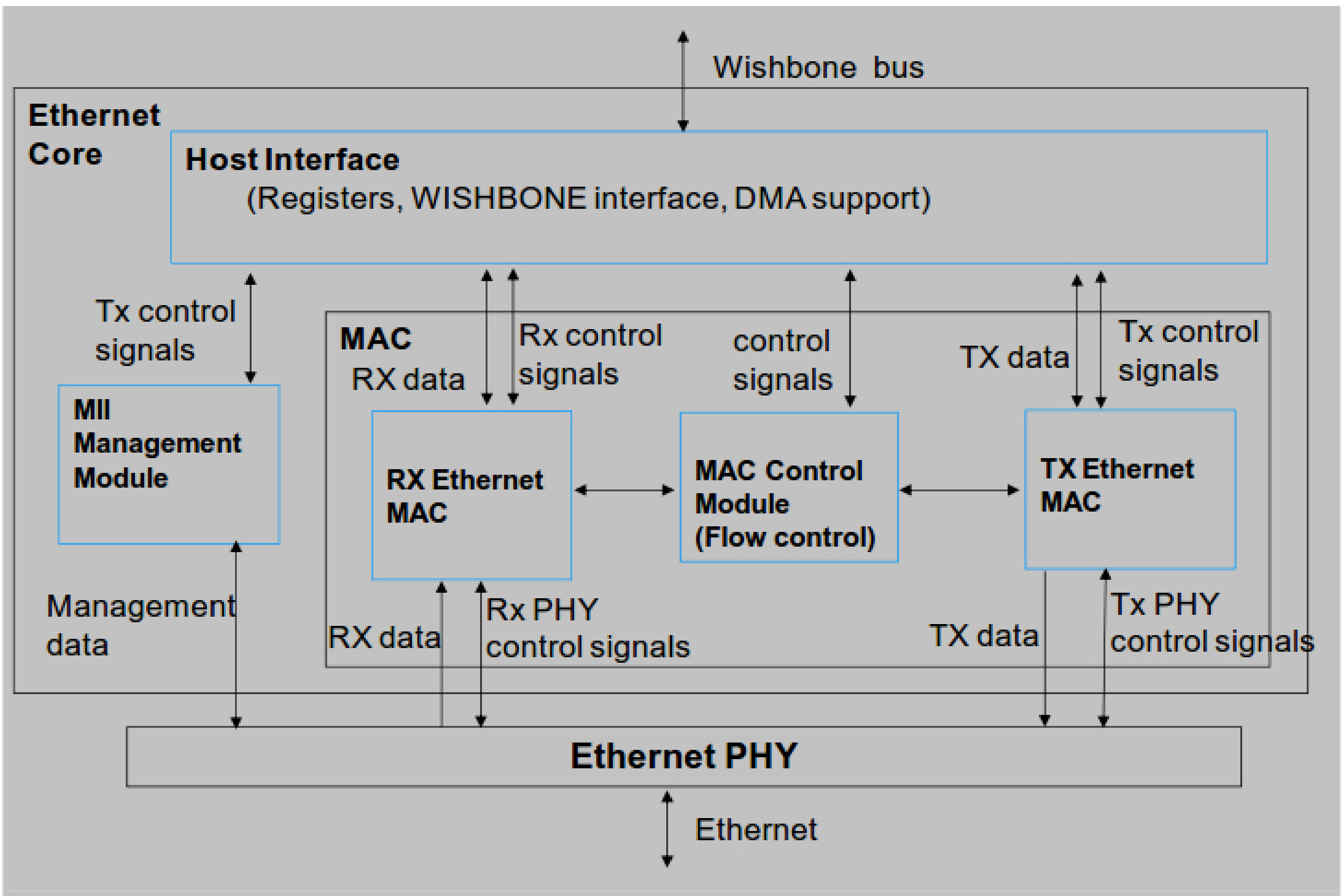


Figure 2.16 Architecture of Ethernet MAC core.

Figure 2.16 Architecture of Ethernet MAC core.
圖2.16以太網MAC內核的架構。
Figure 2.16 illustrates the architecture of OPENCORE Ethernet Core,
圖2.16說明了OPENCORE以太網核心的體系結構，
which mainly consists of host interface, transmit (TX) module, receive (RX) module,
主要由主機接口，發送（TX）模塊，接收（RX）模塊，
MAC control module,
MAC控制模塊
and Media Independent Interface (MII) management module.
和媒體獨立接口（MII）管理模塊。
They are described below.
如下所述。

1. The TX and RX modules enable all transmit and receive functionalities.
1. TX和RX模塊啟用所有發送和接收功能。
These modules handle preamble generation and removal. Both modules incorporate the CRC generators for error detection. In addition, the TX
這些模塊處理前導碼的生成和刪除。這兩個模塊都集成了用於錯誤檢測的CRC發生器。此外，TX
module has a random time generation used in the back-off process,
模塊在退避過程中使用隨機時間生成，
and monitors the CarrierSense and Collision signals to exercise the main body of CSMA/CD.
並監視CarrierSense和Collision信號，以鍛煉CSMA / CD的主體。
2. The MAC control module provides full duplex flow control,
2. MAC控制模塊提供全雙工流控制，
which transfers the PAUSE control frames between the communicating stations.
在通信站之間傳送暫停控制幀。
Therefore, the MAC Control Module has control frame detection and generation, interfaces to TX and RX MAC, PAUSE timer and Slot timer.
因此，MAC控制模塊具有控制幀檢測和生成，與TX和RX MAC，PAUSE計時器和Slot計時器的接口。
3. The MII management module implements the standard of IEEE 802.3 MII,
3. MII管理模塊執行IEEE 802.3 MII的標準，
which provides the interconnections between the Ethernet PHY and MAC layers.
它提供了以太網PHY和MAC層之間的互連。
Through the MII interface, the processor can force Ethernet PHY to run at 10 Mbps or 100 Mbps,
通過MII接口，處理器可以強制以太網PHY以10 Mbps或100 Mbps的速度運行，
and configure it to perform at full or half duplex mode.
並將其配置為以全雙工或半雙工模式執行。
The MII management module has the sub-modules for operation controller, shift registers, output control module and clock generator.
MII管理模塊具有用於操作控制器，移位寄存器，輸出控制模塊和時鐘發生器的子模塊。

4. The host interface is a WISHBONE (WB) bus connecting the Ethernet MAC to the processor and external memory.
4.主機接口是WISHBONE（WB）總線，將以太網MAC連接到處理器和外部存儲器。
The WB is an interconnection specification of OPENCORE projects.
WB是OPENCORE項目的互連規範。
Only DMA transfers are supported so far for data transferring.
到目前為止，僅支持DMA傳輸進行數據傳輸。
The host interface also has status and register modules.
主機接口還具有狀態和寄存器模塊。
The status module records the statuses written to the related buffer descriptors.
狀態模塊記錄寫入相關緩衝區描述符的狀態。
The register module is used for Ethernet MAC operations,
寄存器模塊用於以太網MAC操作，

and it includes configuration registers, DMA operation,
它包括配置寄存器，DMA操作，
and transmit and receive status.
並發送和接收狀態。

State Machines: TX and RX
狀態機：TX和RX
In the TX and RX modules, there are TX and RX state machines, respectively,
在TX和RX模塊中，分別有TX和RX狀態機，
to control their behaviors. Figure 2.17 present the state machines of both.
控制他們的行為。圖2.17展示了兩者的狀態機。
We only describe the behaviors of the TX state machine here.
我們在這裡僅描述TX狀態機的行為。
The RX state machine works similarly. When the WB Interface requests a transmission,
RX狀態機的工作方式與此類似。當WB接口請求傳輸時，
it forces the TX state machine to go to the preamble state,
它強制TX狀態機進入前導狀態，
which informs the Ethernet PHY chip to start a transmission.
通知以太網PHY芯片開始傳輸。
After the preamble 0x5555555 and Start Frame Delimiter 0xd are sent,
發送前同步碼0x5555555和起始幀定界符0xd之後，
the TX state machine goes to the data0 and data1 states to transmit the Least Significant Byte (LSB) nibbles of the data byte,
TX狀態機進入data0和data1狀態以傳輸數據字節的最低有效字節（LSB）半字節，
and then informs the Wishbone Interface to provide next data byte until the end of the
然後通知Wishbone接口提供下一個數據字節，直到
packet.
包。
When there is just one byte left to be sent, it performs the following operations.
當只剩下一個字節要發送時，它將執行以下操作。

If the data length is greater or equal to the minimum frame length and CRC is enabled,
如果數據長度大於或等於最小幀長度，並且啟用了CRC，
then TX state machine enters the FCS state to calculate the 32-bit CRC value from the data, then to the defer state, then to the IPG state,
然後TX狀態機進入FCS狀態以根據數據計算32位CRC值，然後進入延期狀態，然後進入IPG狀態，
and to the idle state again.
並再次進入空閒狀態。
However if CRC is disabled the TX state machine goes to the defer state directly,
但是，如果禁用了CRC，則TX狀態機將直接進入延遲狀態，
then to the IPG state and to the idle state again.
然後進入IPG狀態並再次進入空閒狀態。

z If the data length is smaller than the minimum frame length and padding is enabled,
如果數據長度小於最小幀長度，並且啟用了填充，
then the TX state machine goes to the pad state that data is padded with zeros until the minimum frame length is achieved.
然後TX狀態機進入填充狀態，在數據中填充零，直到達到最小幀長度。
Then the TX state machine goes to the FCS state,
然後TX狀態機進入FCS狀態，
defer state, IPG state and idle state sequentially. However the pad state is skipped when padding is disabled.
依次推遲狀態，IPG狀態和空閒狀態。但是，禁用填充時會跳過填充狀態。

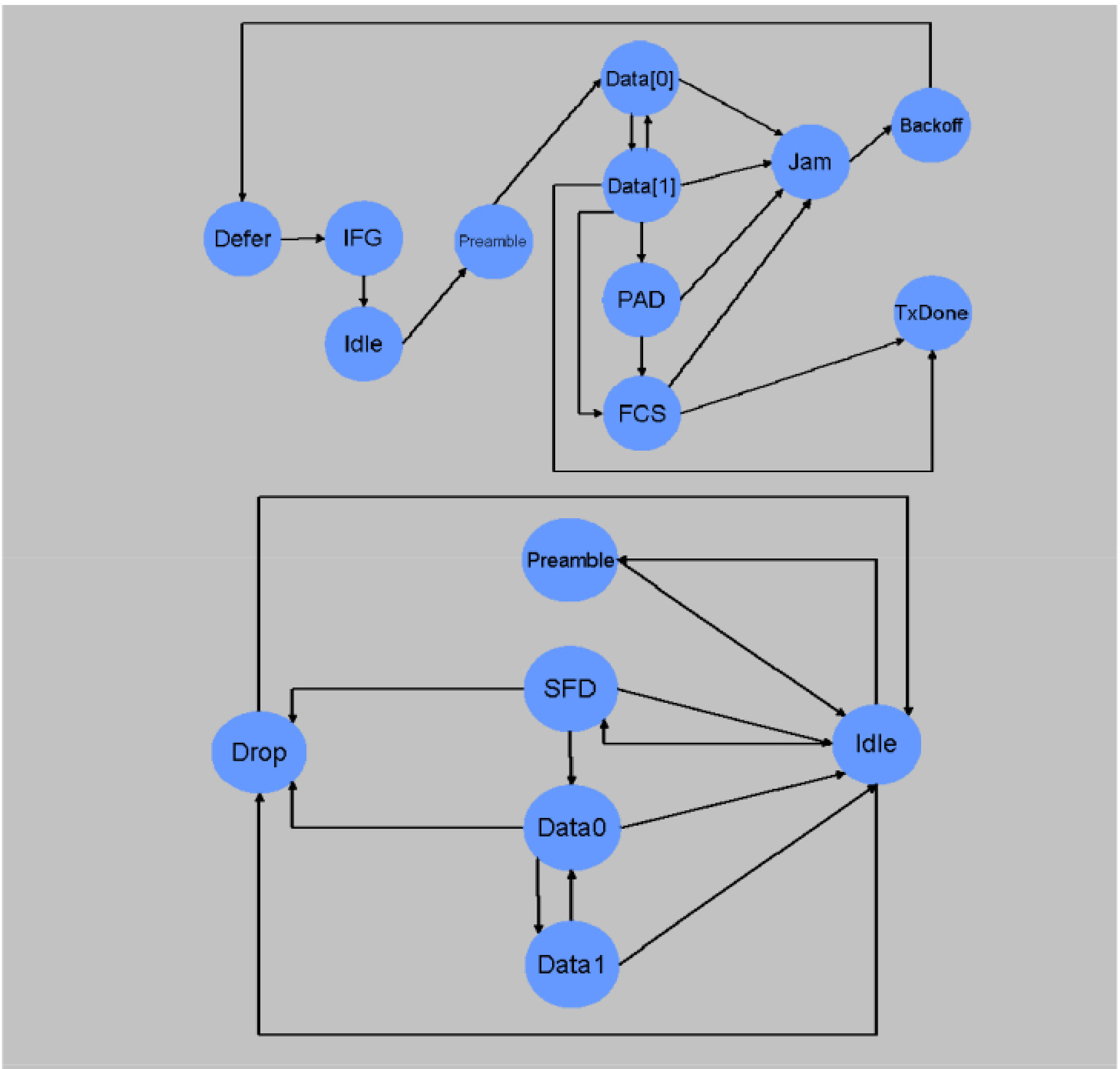


Fig. 2.17 The TX (upper) and RX (lower) state machines.

Programming CSMA/CD Signals

編程CSMA / CD信號

Figure 2.18 is a segment of Verilog code that programs the key CSMA/CD signals.

圖2.18是一段Verilog代碼，可對關鍵的CSMA / CD信號進行編程。

An output signal is an arithmetic combination of various input signals and updated once in every clock cycle.

All output signals are updated in parallel,

輸出信號是各種輸入信號的算術組合，並在每個時鐘週期更新一次。所有輸出信號並行更新，

which is the key difference with the sequentially executed software code.

這是與順序執行的軟件代碼的關鍵區別。

The symbols “~”, “&”, “|” and “=” denote the operations of “not”, “and”, “or” and “assign”, respectively.

符號“～”，“&”，“ |” “ =”分別表示“ not”，“ and”，“ or”和“ assign”的運算。

When a station transmits in the half-duplex mode, it does carrier sense by observing the activity on the PHY media (from the CarrierSense variable that uses “true” or “false” to denote the presence or absence of the carrier).

當一個站以半雙工模式發送時，它通過觀察PHY介質上的活動來進行載波偵聽（來自使用“ true”或“ false”表示載波存在或不存在的CarrierSense變量）。

If no carrier is present, any station can start transmitting (See equation 1).

如果不存在載波，則任何站都可以開始發送（請參見公式1）。

A collision results from simultaneous transmission of two or more stations (denoted by the Collision variable).

同時傳輸兩個或更多站（由Collision變量表示）會導致衝突。

All stations stop transmitting and back off for a random time (the Startbackoff is set to “true”).

所有電台都停止發送並退回一段隨機的時間（Startbackoff設置為“ true”）。

```
assign StartDefer = StateIdle & ~TxStartFrm & CarrierSense
```



```
                | StateBackOff & (TxUnderRun | RandomEqByteCnt) ;
assign StartData[1] = ~Collision & StateData[0] & ~TxUnderRun &
                 ~MaxFrame;
assign StartJam = (Collision | UnderRun) & ((StatePreamble & NibCntEq15)
                | (StateData[1:0]) | StatePAD | StateFCS);
assign StartBackoff = StateJam & ~RandomEq0 & ColWindow & ~RetryMax
                & NibCntEq7 & ~NoBckof;
```

Figure 2.18 CSMA/CD Signals

Since the TX module starts the back-off process after a collision has been detected, it waits for some duration derived from a pseudo random as shown in Figure 2.19. It applies the "binary exponential" algorithm to generate a random back-off time within the predefined restriction. An element `x[i]` in the `x` array are random bits of either 0 or 1, and the Random array can be viewed as the binary representation of the random value (totally 10 bits, as the range of the random number is from 0 to 2^k-1, where $k = min(n, 10)$ and *n* is the number of re-trials.) When `RetryCnt` is larger than `i`, `Random[i]` may be set to 1 if `x[i] = 1`; otherwise `Random[i]` is set to 0 by assigning bit 0 (denoted by 1'b0) to it. In other words, one more high-bit in the random values is likely to be set to 1, which means the range of the random values exponentially grows. After the random value is derived, it will be latched into the `RandomLatched` variable if the transmission channel is jammed (judged from the `StateJam` and `StateJam_q` variables), e.g., due to collision; otherwise, the random value is reset to 0.

Figure 2.18 CSMA/CD Signals

圖2.18 CSMA / CD信號

Since the TX module starts the back-off process after a collision has been

由於TX模塊在發生衝突後開始退避過程

detected, it waits for some duration derived from a pseudo random as shown in

檢測到後，它會等待一段時間（從偽隨機數中得出），如

Figure 2.19.

圖2.19。

It applies the "binary exponential" algorithm to generate a random back-off time within the predefined restriction.

它應用"二進制指數"算法在預定義限制內生成隨機退避時間。

An element x[i] in the x array are random bits of either 0 or 1, and the Random array can be viewed as the binary representation of the random value (totally 10 bits, as the range of the random number is from 0 to 2k-1,

x數組中的元素x [i]是0或1的隨機位，並且Random數組可以看作是隨機值的二進製表示形式（總共10位，因為隨機數的範圍是從0到0）。 2k-1，

where k = min(n, 10) and n is the number of re-trials.)
其中k = min（n，10），n是重試次數。）
When RetryCnt is larger than i, Random[i] may be set to 1 if x[i] = 1;
當RetryCnt大於i時，如果x [i] = 1，則Random [i]可以設置為1；否則，Random [i]可以設置為1。
otherwise Random[i] is set to 0 by assigning bit 0 (denoted by 1'b0) to it.
否則，通過為它分配位0（用1'b0表示）將Random [i]設置為0。
In other words, one more high-bit in the random values is likely to be set to 1, which means the range of the random values exponentially grows.
換句話說，隨機值中的另一個高位可能會被設置為1，這意味著隨機值的範圍呈指數增長。
After the random value is derived, it will be latched into the RandomLatched variable if the transmission channel is jammed (judged from the StateJam and StateJam_q variables), e.g., due to collision; otherwise, the random value is reset to 0.
導出隨機值後，例如，由於衝突導致傳輸通道被阻塞（由StateJam和StateJam_q變量判斷）時，它將被鎖存到RandomLatched變量中。否則，將隨機值重置為0。

```
assign Feedback = ~(x[2] ^ x[9]);
assign Random [0] = x[0];
assign Random [1] = (RetryCnt > 1) ? x[1] : 1'b0;
assign Random [2] = (RetryCnt > 2) ? x[2] : 1'b0;
assign Random [3] = (RetryCnt > 3) ? x[3] : 1'b0;
assign Random [4] = (RetryCnt > 4) ? x[4] : 1'b0;
assign Random [5] = (RetryCnt > 5) ? x[5] : 1'b0;
assign Random [6] = (RetryCnt > 6) ? x[6] : 1'b0;
assign Random [7] = (RetryCnt > 7) ? x[7] : 1'b0;
assign Random [8] = (RetryCnt > 8) ? x[8] : 1'b0;
assign Random [9] = (RetryCnt > 9) ? x[9] : 1'b0;
always @ (posedge MTxClk or posedge Reset)
begin
  if(Reset)
    RandomLatched <= 10'h000;
  else
    begin
      if(StateJam & StateJam_q)
        RandomLatched <= Random;
    end
end
```

Figure 2.19 Jam Random Generator

Gigabit Ethernet
千兆以太網
The stipulation of Gigabit Ethernet was originally divided into two Task Forces: 802.3z and 803.3ab.
千兆以太網的規定最初分為兩個任務組：802.3z和803.3ab。
A later Task force for Ethernet in the First Mile (EFM) also specified three new PHYs running at the gigabit rate. For clearance,
後來的第一英里以太網專責小組（EFM）也指定了三個以千兆位速率運行的新PHY。為了通關
we leave the latter part to the introduction of EFM.
我們將後一部分留給EFM來介紹。
Table 2.7 lists the specifications only in 802.3z and 803.3ab.
表2.7僅列出了802.3z和803.3ab中的規格。
A difficulty in Gigabit Ethernet design is the distance restriction from CSMA/CD.

千兆以太網設計的一個困難是與CSMA / CD的距離限制。

For 10 Mb/s and 100 Mb/s Ethernet,

對於10 Mb / s和100 Mb / s以太網，

this would not be a problem.

這不會有問題。

The distance is about 200 m for copper connection in 100 Mb/s Ethernet,

對於100 Mb / s以太網中的銅線連接，距離約為200 m，

which is enough for normal configurations.

這對於常規配置已經足夠。

The distance is even longer for 10 Mb/s Ethernet,

對於10 Mb / s以太網，距離甚至更長，

but Gigabit Ethernet is ten times faster to transmit a frame than 100 Mb/s Ethernet,

但是千兆位以太網傳輸幀的速度比100 Mb / s以太網快十倍，

making the distance restriction ten times shorter.

使距離限制縮短十倍。

A restriction of about 20 m is unacceptable for many network deployments.

對於許多網絡部署，約20 m的限制是不可接受的。

Table 2.7 Physical specifications of Gigabit Ethernet

Task Forces	Specification name	Description
IEEE 802.3z (1998)	1000BASE-CX	25 m 2-pair Shielded Twisted Pairs (STP) with 8B/10B encoding
	1000BASE-SX	Multi-mode fiber of short-wave laser with 8B/10B encoding
	1000BASE-LX	Multi- or single-mode fiber of long-wave laser with 8B/10B encoding
IEEE 802.3ab (1999)	1000BASE-T	100 m 4-pair Category 5 (or better) Unshielded Twisted Pairs (UTP) with 8B1Q4.

The IEEE 802.3 Standard appends a series of extension bits after a frame. The extension bits can be any non-data symbols in the physical layer.

IEEE 802.3標准在幀後附加了一系列擴展位。擴展位可以是物理層中的任何非數據符號。

This technique,

這種技術

called carrier extension,

稱為運營商擴展，

in effect extends the frame length without changing the minimum frame size. Their extension bits are 4096 bits.

實際上，可以在不更改最小幀大小的情況下擴展幀長度。它們的擴展位是4096位。

They are for CSMA/CD purpose only,

它們僅用於CSMA / CD，

and will be discarded silently by the receiver.

並將被接收方靜默丟棄。

Although carrier extension addresses the problem,

儘管運營商擴展解決了該問題，

the data throughput is low because the transmission channel is mostly occupied by the extension bits if the frames are short.

數據吞吐量很低，因為如果幀較短，則傳輸通道主要由擴展位佔用。

The solution is to allow the transmitter to transmit the next frame,

解決方案是允許發射機發送下一幀，

if any,

如果有的話

without extension bits,

沒有擴展位，

by filling the IFG with a carrier.

通過在IFG中填充載體。

Because the IFG between two successive frames is occupied with a carrier,

由於兩個連續幀之間的IFG被一個載波佔用，

the transmission channel is not relinquished by the transmitter.

發射器不會放棄傳輸通道。

The transmitter can transmit more frames following the first frame,

發送器可以在第一個幀之後發送更多幀，

as long as it has,

只要有

up to a limit.

達到極限。

This technique is called frame bursting. Figure 2.20 depicts the scenario.

這種技術稱為幀突發。圖2.20描述了這種情況。

The maximum length in the bursting is 65,536 bits.

突發中的最大長度為65,536位。

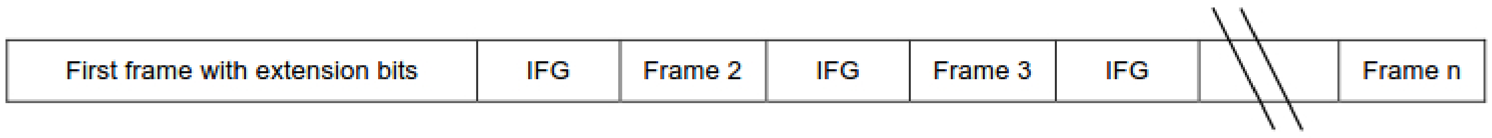


Figure 2.20 Illustration of frame bursting.

Both carrier extension and frame bursting complicates the MAC design.

載波擴展和幀突發都使MAC設計複雜化。

The throughput is not good despite it works.

儘管工作正常，但吞吐量卻不好。

In contrast,

相反，

full duplex Ethernet does not need CSMA/CD at all,

全雙工以太網完全不需要CSMA / CD，

making these solutions unnecessary.

使這些解決方案變得不必要。

Its implementation is simpler and the throughput is much higher.

它的實現更簡單，吞吐量更高。

Why do we bother to implement half-duplex Gigabit Ethernet if it is unnecessary?

如果不必要的話，為什麼還要打擾半雙工千兆以太網呢？

With the advance of ASIC technology,

隨著ASIC技術的進步，

switched networks are no longer much more expensive than shared networks.

交換網絡不再比共享網絡昂貴得多。

For the deployment of Gigabit Ethernet,

為了部署千兆以太網，

it is the performance rather than the cost anymore that is of concern.

關注的是性能而不是成本。

The market has proved the failure of half-duplex Gigabit Ethernet.

市場已經證明了半雙工千兆以太網的失敗。

Only full duplex Gigabit Ethernet products exist on the market.

市場上僅存在全雙工千兆以太網產品。

10 Gigabit Ethernet

10千兆以太網

Just like Moore's Law stating the power of microprocessors doubles every 18 months,

就像摩爾定律指出微處理器的能力每18個月翻一番，

the speed of Ethernet has also grown exponentially.

以太網的速度也呈指數增長。

The 10 GigabitEthernet Standard, developed by the IEEE 802.3ae Task Force came out in 2002.

由IEEE 802.3ae工作組開發的10 GigabitEthernet標準於2002年發布。

It was later extended to operate on twisted pairs in 2006, namely 10GBASE-T.

後來在2006年擴展為可在雙絞線上使用，即10GBASE-T。

The 10 Gigabit Ethernet bears the following features:

10 Gb以太網具有以下功能：

Full duplex only: The IEEE 802.3 people learned a lesson from the development of Gigabit Ethernet. Only full duplex mode is in the 10 Gigabit Ethernet.

僅限全雙工：IEEE 802.3人員從千兆以太網的發展中學到了教訓。 10千兆位以太網中僅全雙工模式。

Half duplex mode is not even considered.

甚至不考慮半雙工模式。

Compatibility with past standard: The frame format and the MAC operations remain unchanged, making the interoperability with existing products rather easy.

與過去標準的兼容性：幀格式和MAC操作保持不變，從而使與現有產品的互操作性變得非常容易。

Move toward the WAN market: Since Gigabit Ethernet has moved toward the MAN market,

邁向WAN市場：由於千兆以太網已邁向MAN市場，

10 Gigabit Ethernet will go further into the WAN market.

10 Gb以太網將進一步進入WAN市場。

On one hand,

一方面，

the longest distance in the new standard is 40 km.

新標準中最長的距離是40公里。

On the other hand,

另一方面，

a WAN PHY is defined to interface with OC-192 (OC: Optical Carrier) in the Synchronous Optical Networking (SONET) infrastructure,

WAN PHY被定義為與同步光網絡（SONET）基礎架構中的OC-192（OC：光載波）接口，

which operates at a rate very close to 10 Gigabit.

它的運行速度非常接近10 Gb。

The IEEE 802.3ae comes with an optional WAN PHY besides the LAN PHY.

除了LAN PHY之外，IEEE 802.3ae還帶有可選的WAN PHY。

Both PHYs have the same transmission media,

兩個PHY具有相同的傳輸媒體，

and thus the same transmission distance. The difference is that the WAN PHY has a WAN Interface Sublayer (WIS) in the Physical Coding Sublayer (PCS).

因此，相同的傳輸距離。區別在於WAN PHY在物理編碼子層（PCS）中具有WAN接口子層（WIS）。

The WIS is a framer that maps an Ethernet frame into a SONET payload.

WIS是將以太網幀映射到SONET有效負載的成幀器。

This makes attaching Ethernet to OC-192 devices easy.

這使得將以太網輕鬆連接到OC-192設備變得容易。

Table 2.8 lists the physical specifications in IEEE802.ae. The character 'W' in the code names denotes a WAN PHY,

表2.8列出了IEEE802.ae中的物理規格。代碼名稱中的字符“ W”表示WAN PHY，

which can be directly connected to an OC-192 interface. The others are for LAN only.

可以直接連接到OC-192接口。其他僅用於局域網。

Every physical specification except 10GBASE-LX4 uses a complex 64B/66B block coding. 10GBASE-LX4 uses 8B/10B blocking coding,

除10GBASE-LX4以外的所有物理規範都使用複雜的64B / 66B塊編碼。 10GBASE-LX4使用8B / 10B塊編碼，

and relies on four wave-length division multiplexing (WDM) channels to achieve 10 Gb/s.

並依靠四個波分複用（WDM）通道來達到10 Gb / s。

Besides the first batch of 10 Gigabit specifications in IEEE 802.3ae,

除了IEEE 802.3ae中的第一批10 Gigabit規範外，

later specifications such as 10GBASE-CX4 and 10GBASE-T allow even copper wires to transmit at 10 Gb/s.

後來的規格（例如10GBASE-CX4和10GBASE-T）甚至允許銅線以10 Gb / s的速率傳輸。

An extension to Ethernet Passive Optical Network (EPON) running at 10 Gb/s is also under development in 2008.

還在2008年開發以10 Gb / s速度運行的以太網無源光網絡（EPON）的擴展。

Table 2.8 Physical specifications in the IEEE 802.3ae.

Code name	Wave length	Transmission distance (m)
10GBASE-LX4	1310 nm	300
10GBASE-SR	850 nm	300
10GBASE-LR	1310 nm	10,000
10GBASE-ER	1550 nm	10,000
10GBASE-SW	850 nm	300
10GBASE-LW	1310 nm	10,000
10GBASE-EW	1550 nm	40,000

Ethernet in the First Mile

第一英里以太網

We have Ethernet dominating the wired LAN,

我們以以太網主導著有線局域網，

and are seeing it to dominate in the WAN.

並看到它在WAN中占主導地位。

But how about the interface between LAN and WAN?

但是，LAN和WAN之間的接口又如何呢？

Given broad bandwidth both on the LAN and WAN,

鑑於LAN和WAN都具有寬帶寬，

you might still need to access Internet at home through ADSL,

您可能仍然需要通過ADSL在家訪問Internet，

cable modems and so on.

電纜調製解調器等。

The segment of the subscriber access network,

訂戶訪問網絡的網段，

also called the first mile or last mile,

也稱為第一英里或最後一英里，

between LAN and WAN may become a bottleneck.

LAN和WAN之間的連接可能會成為瓶頸。

The protocol conversion due to different technologies in LAN,

由於LAN中的技術不同，導致協議轉換，

first mile and WAN is an overhead.

第一英里，而WAN是開銷。

With the popularity of subscriber access network,

隨著用戶訪問網絡的普及，

the potential market becomes highly noticeable. An effort in the IEEE 802.3ah Ethernet in the First Mile (EFM) Task Force defined a standard for this market.

潛在市場變得非常引人注目。第一英里（EFM）工作組的IEEE 802.3ah以太網的努力為該市場定義了一個標準。
If Ethernet could be everywhere in the wired networks,
如果以太網可以在有線網絡中無處不在，
no protocol conversion would be needed,
無需協議轉換，
which also reduces the overall cost.
這也降低了總體成本。
All in all,
總而言之，
the standard is expected to provide a cheap and fast technology in the potentially broad first mile market.
該標準有望在潛在的廣泛的"第一英里"市場中提供一種廉價且快速的技術。
Ethernet is poised to be ubiquitous,
以太網勢必無處不在，
and the goals of the standard include the following:
該標準的目標包括：

New topologies: The requirements for the subscriber access network include point to point on fiber, point to multipoint on fiber, and point to point on copper. The standard meets these requirements.
新拓撲：用戶接入網絡的要求包括光纖上的點對點，光纖上的點對多點以及銅線上的點對點。該標準滿足這些要求。

New PHYs:
新的PHY：
Table 2.9 summarizes the PHYs in IEEE 802.3ah,
表2.9總結了IEEE 802.3ah中的PHY，
including the following specifications: Point-to-Point optics The PHYs are single-mode fibers from one point to the other.
包括以下規範：點對點光學器件PHY是從一個點到另一個點的單模光纖。
They include 100BASE-LX10,
其中包括100BASE-LX10，
100BASE-BX10,
100BASE-BX10，
1000BASE-LX10,
1000BASE-LX10，
1000BASE-BX10,
1000BASE-BX10，
where LX denotes a pair of fibers and BX denotes a single fiber.
其中LX表示一對光纖，BX表示單個光纖。
Here 10 means the transmission distance is 10 km,
這裡10表示傳輸距離是10 km，
which is longer than the longest distance of 5 km in IEEE 802.3z Gigabit Ethernet.
它比IEEE 802.3z千兆以太網中最長的5 km距離更長。
Point-to-Multipoint optics: The topology serves multiple premises from a single point.
點對多點光學器件：拓撲從單個點為多個前提提供服務。
In the branch is a passive optical splitter that is not powered,
分支中是一個無源分光器，它沒有供電，
so the topology is also called Passive Optical Network (PON).
因此該拓撲也稱為無源光網絡（PON）。
The PHYs include 1000BASE-PX10 and 1000BASE-PX20.
PHY包括1000BASE-PX10和1000BASE-PX20。
The former can transmit 10 km,
前者可以傳輸10公里，
while the latter can transmit up 20 km.
而後者可以傳輸20公里。
Another effort to push Ethernet PON up to 10 Gb/s is ongoing in IEEE 802.3av.

在IEEE 802.3av中，正在進行將以太網PON提升至10 Gb / s的另一項努力。

Point-to-Point copper: The PHYs are for non-loaded voice grade copper cables.

點對點銅纜：PHY用於非負載語音級銅纜。

The PHYs include 2BASE-TL and 10PASS-TS.

PHY包括2BASE-TL和10PASS-TS。

The former is at least 2 Mb/s up to 2700 m over SHDSL,

前者在SHDSL上達2700 m時至少為2 Mb / s，

and the latter is at least 10 Mb/s up to 750 m over VDSL.

後者在VDSL上達到至少750 m時至少為10 Mb / s。

They are more economical solutions if the optical fibers are unavailable.

如果沒有光纖，它們是更經濟的解決方案。

Table 2.9 Physical specifications in the IEEE 802.3ah.

Code name	Description
100BASE-LX10	100 Mb/s on a pair of optical fibers up to 10 km
100BASE-BX10	100 Mb/s on a optical fiber up to 10 km
1000BASE-LX10	1000 Mb/s on a pair of optical fibers up to 10 km
1000BASE-BX10	1000 Mb/s on a optical fiber up to 10 km
1000BASE-PX10	1000 Mb/s on passive optical network up to 10 km
1000BASE-PX20	1000 Mb/s on passive optical network up to 20 km
2BASE-TL	At least 2 Mb/s over SHDSL up to 2700 m
10PASS-TS	At least 10 Mb/s over VDSL up to 750 m

Far-end Operations, Administration, and Maintenance (OAM):

遠端運作，管理和維護（OAM）：

Reliability is critical in the subscriber access network.

可靠性在用戶訪問網絡中至關重要。

For easy OAM,

為了輕鬆進行OAM，

the standard defines new methods of remote failure indication,

該標准定義了遠程故障指示的新方法，

remote loopback and link monitoring.

遠程環回和鏈接監視。