FOURTH EDITION

Computer Networks

Virtual LAN

ANDREW S. TANENBAUM



CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 1 of 11
INFOEXPRESS, INC. - EX2566

Library of Congress Cataloging-in-Publication Data

Tanenbaum, Andrew S.
Computer networks / Andrew S. Tanenbaum.--4th ed.
p. cm.
Includes bibliographical references.
ISBN 0-13-066102-3
1. Computer networks. I. Title.
TK5105.5 .T36 2002
004.6--dc21 2002029263

Editorial/production supervision: *Patti Guerrieri*
Cover design director: *Jerry Votta*
Cover designer: *Anthony Gemmellaro*
Cover design: *Andrew S. Tanenbaum*
Art director: *Gail Cocker-Bogusz*
Interior Design: *Andrew S. Tanenbaum*
Interior graphics: *Hadel Studio*
Typesetting: *Andrew S. Tanenbaum*
Manufacturing buyer: *Maura Zaldivar*
Executive editor: *Mary Franz*
Editorial assistant: *Noreen Regina*
Marketing manager: *Dan DePasquale*

PH PTR
© 2003, 1996 Pearson Education, Inc.
Publishing as Prentice Hall PTR
Upper Saddle River, New Jersey 07458

Prentice Hall books are widely used by corporations and government agencies for training, marketing, and resale.

For information regarding corporate and government bulk discounts please contact:
Corporate and Government Sales (800) 382-3419 or corpsales@pearsontechgroup.com

All products or services mentioned in this book are the trademarks or service marks of their respective companies or organizations.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-066102-3

Pearson Education LTD.
Pearson Education Australia PTY, Limited
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd.
Pearson Education Canada, Ltd.
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education — Japan
Pearson Education Malaysia, Pte. Ltd.

a consequence, when host *A* in Fig. 4-47(b) wants to send a frame to host *B*, the bridge gets the frame but just discards it. In contrast, in Fig. 4-47(c), the switch must actively forward the frame from *A* to *B* because there is no other way for the frame to get there. Since each switch port usually goes to a single computer, switches must have space for many more line cards than do bridges intended to connect only LANs. Each line card provides buffer space for frames arriving on its ports. Since each port is its own collision domain, switches never lose frames to collisions. However, if frames come in faster than they can be retransmitted, the switch may run out of buffer space and have to start discarding frames.

To alleviate this problem slightly, modern switches start forwarding frames as soon as the destination header field has come in, but before the rest of the frame has arrived (provided the output line is available, of course). These switches do not use store-and-forward switching. Sometimes they are referred to as **cut-through switches**. Usually, cut-through is handled entirely in hardware, whereas bridges traditionally contained an actual CPU that did store-and-forward switching in software. But since all modern bridges and switches contain special integrated circuits for switching, the difference between a switch and bridge is more a marketing issue than a technical one.

So far we have seen repeaters and hubs, which are quite similar, as well as bridges and switches, which are also very similar to each other. Now we move up to routers, which are different from all of the above. When a packet comes into a router, the frame header and trailer are stripped off and the packet located in the frame's payload field (shaded in Fig. 4-46) is passed to the routing software. This software uses the packet header to choose an output line. For an IP packet, the packet header will contain a 32-bit (IPv4) or 128-bit (IPv6) address, but not a 48-bit 802 address. The routing software does not see the frame addresses and does not even know whether the packet came in on a LAN or a point-to-point line. We will study routers and routing in Chap. 5.

Up another layer we find transport gateways. These connect two computers that use different connection-oriented transport protocols. For example, suppose a computer using the connection-oriented TCP/IP protocol needs to talk to a computer using the connection-oriented ATM transport protocol. The transport gateway can copy the packets from one connection to the other, reformatting them as need be.

Finally, application gateways understand the format and contents of the data and translate messages from one format to another. An e-mail gateway could translate Internet messages into SMS messages for mobile phones, for example.

4.7.6 Virtual LANs

In the early days of local area networking, thick yellow cables snaked through the cable ducts of many office buildings. Every computer they passed was plugged in. Often there were many cables, which were connected to a central back-

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 3 of 11
INFOEXPRESS, INC. - EX2566

bone (as in Fig. 4-39) or to a central hub. No thought was given to which computer belonged on which LAN. All the people in adjacent offices were put on the same LAN whether they belonged together or not. Geography trumped logic.

With the advent of 10Base-T and hubs in the 1990s, all that changed. Buildings were rewired (at considerable expense) to rip out all the yellow garden hoses and install twisted pairs from every office to central wiring closets at the end of each corridor or in a central machine room, as illustrated in Fig. 4-48. If the Vice President in Charge of Wiring was a visionary, category 5 twisted pairs were installed; if he was a bean counter, the existing (category 3) telephone wiring was used (only to be replaced a few years later when fast Ethernet emerged).

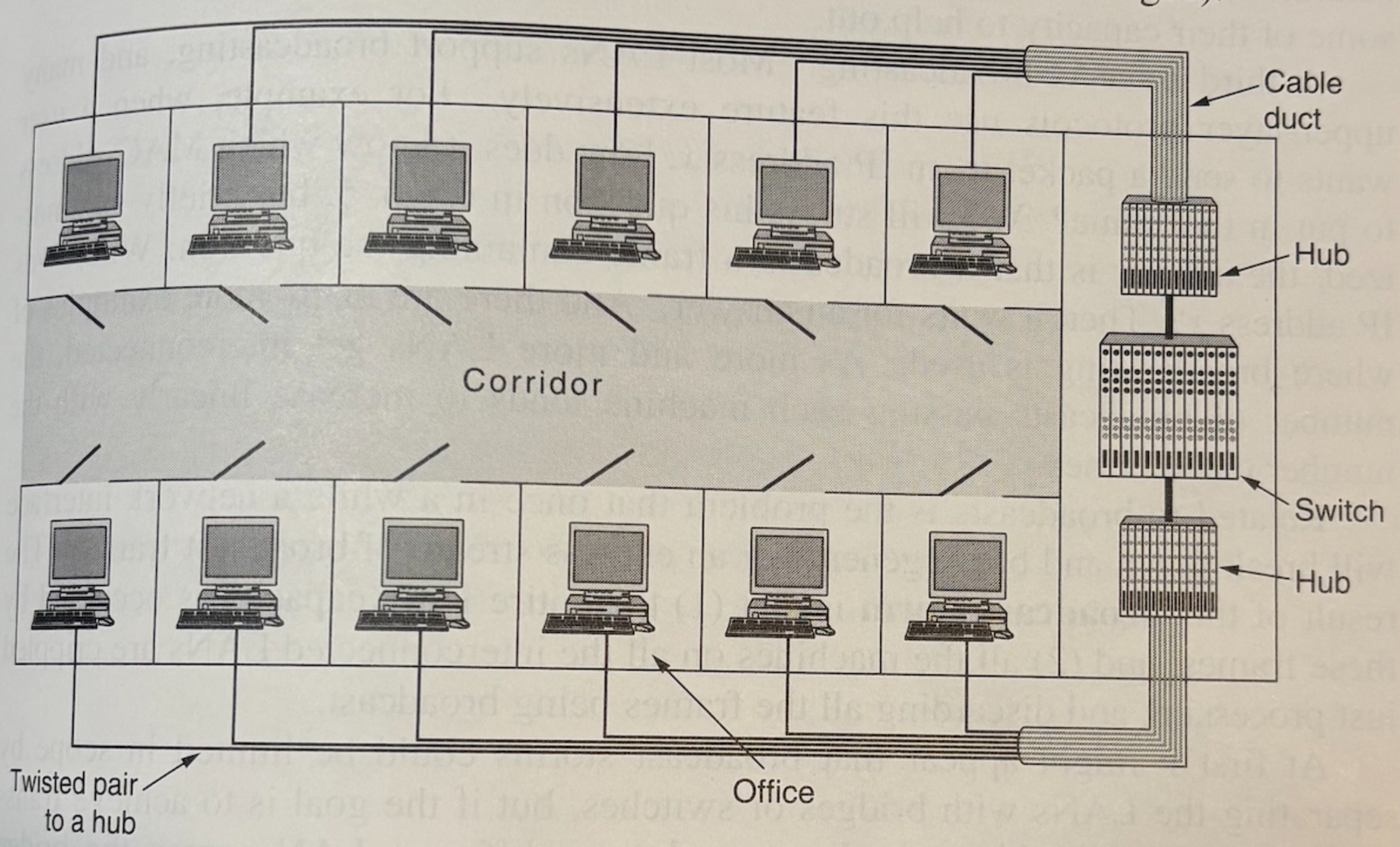


Figure 4-48. A building with centralized wiring using hubs and a switch.

With hubbed (and later, switched) Ethernet, it was often possible to configure LANs logically rather than physically. If a company wants k LANs, it buys k hubs. By carefully choosing which connectors to plug into which hubs, the occupants of a LAN can be chosen in a way that makes organizational sense, without too much regard to geography. Of course, if two people in the same department work in different buildings, they are probably going to be on different hubs and thus different LANs. Nevertheless, the situation is a lot better than having LAN membership entirely based on geography.

Does it matter who is on which LAN? After all, in virtually all organizations, all the LANs are interconnected. In short, yes, it often matters. Network administrators like to group users on LANs to reflect the organizational structure rather than the physical layout of the building for a variety of reasons. One issue is security. Any network interface can be put in promiscuous mode, copying all the

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 4 of 11
INFOEXPRESS, INC. - EX2566

traffic that comes down the pipe. Many departments, such as research, patents, and accounting, have information that they do not want passed outside their department. In such a situation, putting all the people in a department on a single LAN and not letting any of that traffic off the LAN makes sense. Management does not like hearing that such an arrangement is impossible unless all the people in each department are located in adjacent offices with no interlopers.

A second issue is load. Some LANs are more heavily used than others and it may be desirable to separate them at times. For example, if the folks in research are running all kinds of nifty experiments that sometimes get out of hand and saturate their LAN, the folks in accounting may not be enthusiastic about donating some of their capacity to help out.

A third issue is broadcasting. Most LANs support broadcasting, and many upper-layer protocols use this feature extensively. For example, when a user wants to send a packet to an IP address x, how does it know which MAC address to put in the frame? We will study this question in Chap. 5, but briefly summarized, the answer is that it broadcasts a frame containing the question: Who owns IP address x? Then it waits for an answer. And there are many more examples of where broadcasting is used. As more and more LANs get interconnected, the number of broadcasts passing each machine tends to increase linearly with the number of machines.

Related to broadcasts is the problem that once in a while a network interface will break down and begin generating an endless stream of broadcast frames. The result of this **broadcast storm** is that (1) the entire LAN capacity is occupied by these frames, and (2) all the machines on all the interconnected LANs are crippled just processing and discarding all the frames being broadcast.

At first it might appear that broadcast storms could be limited in scope by separating the LANs with bridges or switches, but if the goal is to achieve transparency (i.e., a machine can be moved to a different LAN across the bridge without anyone noticing it), then bridges have to forward broadcast frames.

Having seen why companies might want multiple LANs with restricted scope, let us get back to the problem of decoupling the logical topology from the physical topology. Suppose that a user gets shifted within the company from one department to another without changing offices or changes offices without changing departments. With hubbed wiring, moving the user to the correct LAN means having the network administrator walk down to the wiring closet and pull the connector for the user's machine from one hub and put it into a new hub.

In many companies, organizational changes occur all the time, meaning that system administrators spend a lot of time pulling out plugs and pushing them back in somewhere else. Also, in some cases, the change cannot be made at all because the twisted pair from the user's machine is too far from the correct hub (e.g., in the wrong building).

In response to user requests for more flexibility, network vendors began working on a way to rewire buildings entirely in software. The resulting concept is

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 5 of 11
INFOEXPRESS, INC. - EX2566

called a **VLAN (Virtual LAN)** and has even been standardized by the 802 committee. It is now being deployed in many organizations. Let us now take a look at it. For additional information about VLANs, see (Breyer and Riley, 1999; and Seifert, 2000).

VLANs are based on specially-designed VLAN-aware switches, although they may also have some hubs on the periphery, as in Fig. 4-48. To set up a VLAN-based network, the network administrator decides how many VLANs there will be, which computers will be on which VLAN, and what the VLANs will be called. Often the VLANs are (informally) named by colors, since it is then possible to print color diagrams showing the physical layout of the machines, with the members of the red LAN in red, members of the green LAN in green, and so on. In this way, both the physical and logical layouts are visible in a single view.

As an example, consider the four LANs of Fig. 4-49(a), in which eight of the machines belong to the G (gray) VLAN and seven of them belong to the W (white) VLAN. The four physical LANs are connected by two bridges, *B1* and *B2*. If centralized twisted pair wiring is used, there might also be four hubs (not shown), but logically a multidrop cable and a hub are the same thing. Drawing it this way just makes the figure a little less cluttered. Also, the term "bridge" tends to be used nowadays mostly when there are multiple machines on each port, as in this figure, but otherwise, "bridge" and "switch" are essentially interchangeable. Fig. 4-49(b) shows the same machines and same VLANs using switches with a single computer on each port.

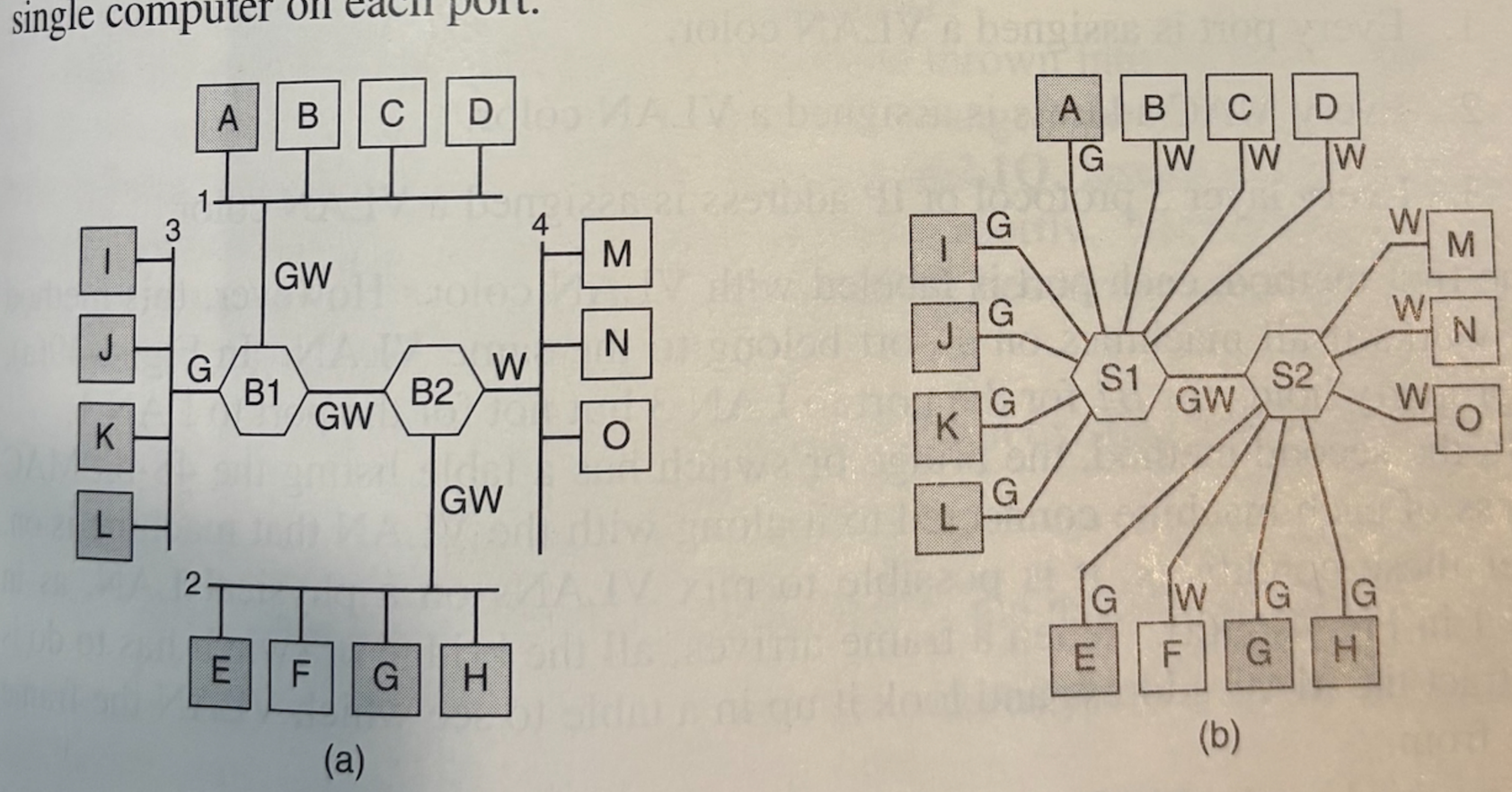


Figure 4-49. (a) Four physical LANs organized into two VLANs, gray and white, by two bridges. (b) The same 15 machines organized into two VLANs by switches.

To make the VLANs function correctly, configuration tables have to be set up in the bridges or switches. These tables tell which VLANs are accessible via which ports (lines). When a frame comes in from, say, the gray VLAN, it must be

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 6 of 11
INFOEXPRESS, INC. - EX2566

forwarded on all the ports marked G. This holds for ordinary (i.e., unicast) traffic as well as for multicast and broadcast traffic.

Note that a port may be labeled with multiple VLAN colors. We see this most clearly in Fig. 4-49(a). Suppose that machine *A* broadcasts a frame. Bridge *B1* receives the frame and sees that it came from a machine on the gray VLAN, so it forwards it on all ports labeled G (except the incoming port). Since *B1* has only two other ports and both of them are labeled G, the frame is sent to both of them.

At *B2* the story is different. Here the bridge knows that there are no gray machines on LAN 4, so the frame is not forwarded there. It goes only to LAN 2. If one of the users on LAN 4 should change departments and be moved to the gray VLAN, then the tables inside *B2* have to be updated to relabel that port as GW instead of W. If machine *F* goes gray, then the port to LAN 2 has to be changed to G instead of GW.

Now let us imagine that all the machines on both LAN 2 and LAN 4 become gray. Then not only do *B2*'s ports to LAN 2 and LAN 4 get marked G, but *B1*'s port to *B2* also has to change from GW to G since white frames arriving at *B1* from LANs 1 and 3 no longer have to be forwarded to *B2*. In Fig. 4-49(b) the same situation holds, only here all the ports that go to a single machine are labeled with a single color because only one VLAN is out there.

So far we have assumed that bridges and switches somehow know what color an incoming frame is. How do they know this? Three methods are in use, as follows:

1. Every port is assigned a VLAN color.
2. Every MAC address is assigned a VLAN color.
3. Every layer 3 protocol or IP address is assigned a VLAN color.

In the first method, each port is labeled with VLAN color. However, this method only works if all machines on a port belong to the same VLAN. In Fig. 4-49(a), this property holds for *B1* for the port to LAN 3 but not for the port to LAN 1.

In the second method, the bridge or switch has a table listing the 48-bit MAC address of each machine connected to it along with the VLAN that machine is on. Under these conditions, it is possible to mix VLANs on a physical LAN, as in LAN 1 in Fig. 4-49(a). When a frame arrives, all the bridge or switch has to do is to extract the MAC address and look it up in a table to see which VLAN the frame came from.

The third method is for the bridge or switch to examine the payload field of the frame, for example, to classify all IP machines as belonging to one VLAN and all AppleTalk machines as belonging to another. For the former, the IP address can also be used to identify the machine. This strategy is most useful when many machines are notebook computers that can be docked in any one of several places. Since each docking station has its own MAC address, just knowing which docking station was used does not say anything about which VLAN the notebook is on.

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 7 of 11
INFOEXPRESS, INC. - EX2566

The only problem with this approach is that it violates the most fundamental rule of networking: independence of the layers. It is none of the data link layer's business what is in the payload field. It should not be examining the payload and certainly not be making decisions based on the contents. A consequence of using this approach is that a change to the layer 3 protocol (for example, an upgrade from IPv4 to IPv6) suddenly causes the switches to fail. Unfortunately, switches that work this way are on the market.

Of course, there is nothing wrong with routing based on IP addresses—nearly all of Chap. 5 is devoted to IP routing—but mixing the layers is looking for trouble. A switch vendor might pooh-pooh this argument saying that its switches understand both IPv4 and IPv6, so everything is fine. But what happens when IPv7 happens? The vendor would probably say: Buy new switches, is that so bad?

The IEEE 802.1Q Standard

Some more thought on this subject reveals that what actually matters is the VLAN of the frame itself, not the VLAN of the sending machine. If there were some way to identify the VLAN in the frame header, then the need to inspect the payload would vanish. For a new LAN, such as 802.11 or 802.16, it would have been easy enough to just add a VLAN field in the header. In fact, the *Connection Identifier* field in 802.16 is somewhat similar in spirit to a VLAN identifier. But what to do about Ethernet, which is the dominant LAN, and does not have any spare fields lying around for the VLAN identifier?

The IEEE 802 committee had this problem thrown into its lap in 1995. After much discussion, it did the unthinkable and changed the Ethernet header. The new format was published in IEEE standard **802.1Q**, issued in 1998. The new format contains a VLAN tag; we will examine it shortly. Not surprisingly, changing something as well established as the Ethernet header is not entirely trivial. A few questions that come to mind are:

1. Need we throw out several hundred million existing Ethernet cards?
2. If not, who generates the new fields?
3. What happens to frames that are already the maximum size?

Of course, the 802 committee was (only too painfully) aware of these problems and had to come up with solutions, which it did.

The key to the solution is to realize that the VLAN fields are only actually used by the bridges and switches and not by the user machines. Thus in Fig. 4-49, it is not really essential that they are present on the lines going out to the end stations as long as they are on the line between the bridges or switches. Thus, to use VLANs, the bridges or switches have to be VLAN aware, but that was already a requirement. Now we are only introducing the additional requirement that they are 802.1Q aware, which new ones already are.

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 8 of 11
INFOEXPRESS, INC. - EX2566

As to throwing out all existing Ethernet cards, the answer is no. Remember that the 802.3 committee could not even get people to change the *Type* field into a *Length* field. You can imagine the reaction to an announcement that all existing Ethernet cards had to be thrown out. However, as new Ethernet cards come on the market, the hope is that they will be 802.1Q compliant and correctly fill in the VLAN fields.

So if the originator does not generate the VLAN fields, who does? The answer is that the first VLAN-aware bridge or switch to touch a frame adds them and the last one down the road removes them. But how does it know which frame belongs to which VLAN? Well, the first bridge or switch could assign a VLAN number to a port, look at the MAC address, or (heaven forbid) examine the payload. Until Ethernet cards are all 802.1Q compliant, we are kind of back where we started. The real hope here is that all gigabit Ethernet cards will be 802.1Q compliant from the start and that as people upgrade to gigabit Ethernet, 802.1Q will be introduced automatically. As to the problem of frames longer than 1518 bytes, 802.1Q just raised the limit to 1522 bytes.

During the transition process, many installations will have some legacy machines (typically classic or fast Ethernet) that are not VLAN aware and others (typically gigabit Ethernet) that are. This situation is illustrated in Fig. 4-50, where the shaded symbols are VLAN aware and the empty ones are not. For simplicity, we assume that all the switches are VLAN aware. If this is not the case, the first VLAN-aware switch can add the tags based on MAC or IP addresses.

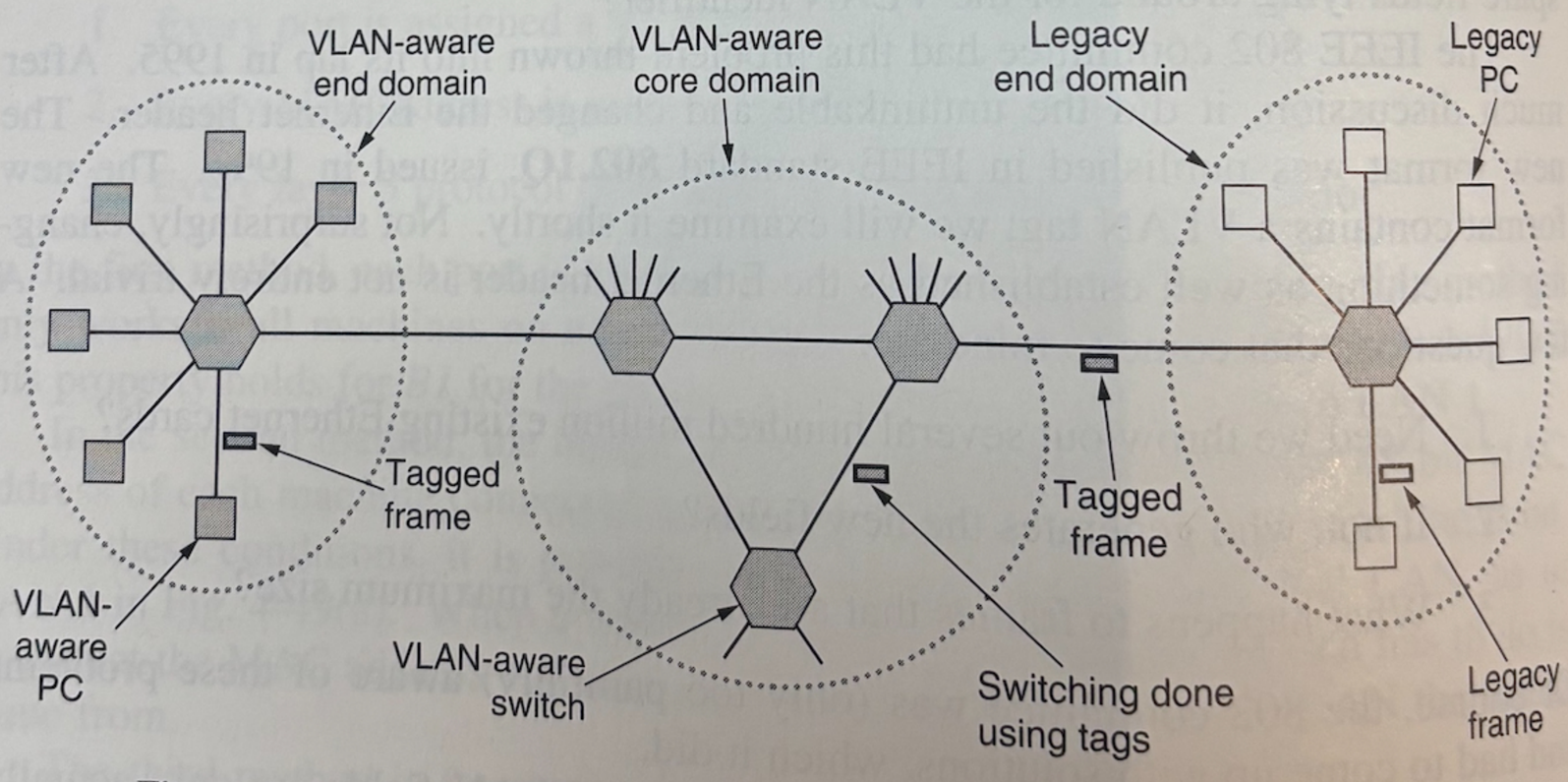


Figure 4-50. Transition from legacy Ethernet to VLAN-aware Ethernet. The shaded symbols are VLAN aware. The empty ones are not.

In this figure, VLAN-aware Ethernet cards generate tagged (i.e., 802.1Q) frames directly, and further switching uses these tags. To do this switching, the switches have to know which VLANs are reachable on each port, just as before. Knowing that a frame belongs to the gray VLAN does not help much until the

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 9 of 11
INFOEXPRESS, INC. - EX2566

switch knows which ports connect to machines on the gray VLAN. Thus, the switch needs a table indexed by VLAN telling which ports to use and whether they are VLAN aware or legacy.

When a legacy PC sends a frame to a VLAN-aware switch, the switch builds a new tagged frame based on its knowledge of the sender's VLAN (using the port, MAC address, or IP address). From that point on, it no longer matters that the sender was a legacy machine. Similarly, a switch that needs to deliver a tagged frame to a legacy machine has to reformat the frame in the legacy format before delivering it.

Now let us take a look at the 802.1Q frame format. It is shown in Fig. 4-51. The only change is the addition of a pair of 2-byte fields. The first one is the *VLAN protocol ID*. It always has the value 0x8100. Since this number is greater than 1500, all Ethernet cards interpret it as a type rather than a length. What a legacy card does with such a frame is moot since such frames are not supposed to be sent to legacy cards.

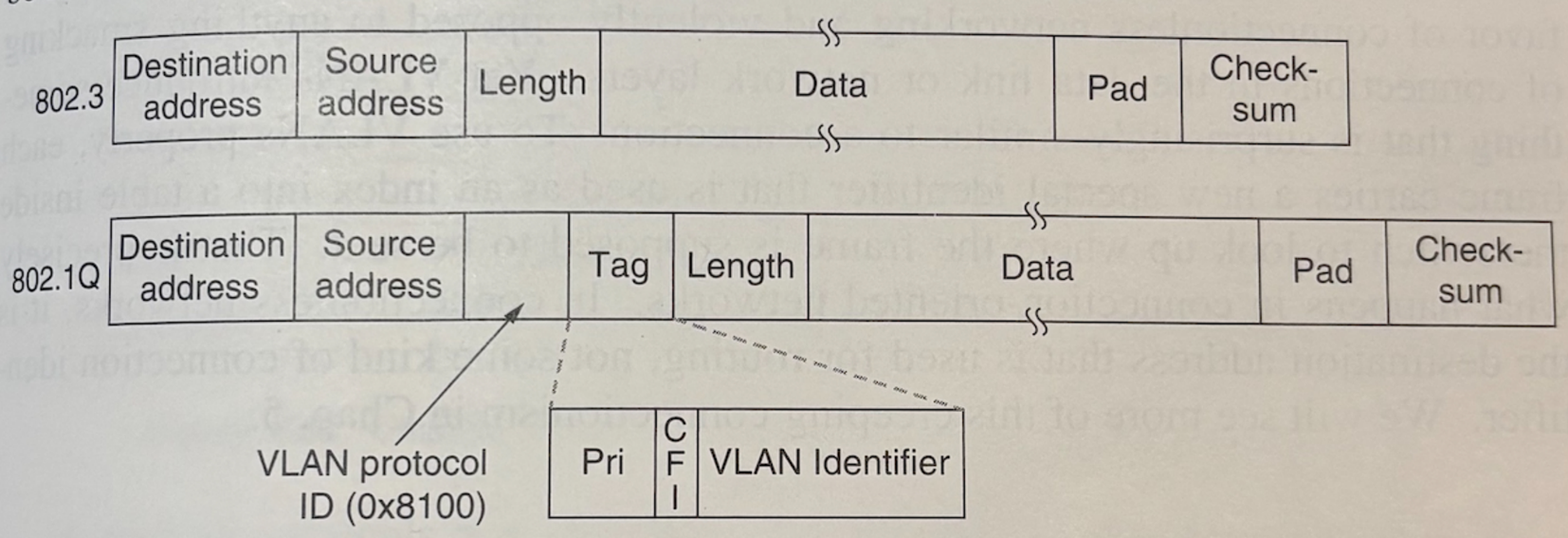


Figure 4-51. The 802.3 (legacy) and 802.1Q Ethernet frame formats.

The second 2-byte field contains three subfields. The main one is the *VLAN identifier*, occupying the low-order 12 bits. This is what the whole thing is about—which VLAN does the frame belong to? The 3-bit *Priority* field has nothing to do with VLANs at all, but since changing the Ethernet header is a once-in-a-decade event taking three years and featuring a hundred people, why not put in some other good things while you are at it? This field makes it possible to distinguish hard real-time traffic from soft real-time traffic from time-insensitive traffic in order to provide better quality of service over Ethernet. It is needed for voice over Ethernet (although in all fairness, IP has had a similar field for a quarter of a century and nobody ever used it).

The last bit, *CFI* (*Canonical Format Indicator*) should have been called the *CEI* (*Corporate Ego Indicator*). It was originally intended to indicate little-endian MAC addresses versus big-endian MAC addresses, but that use got lost in other controversies. Its presence now indicates that the payload contains a freeze-dried 802.5 frame that is hoping to find another 802.5 LAN at the destination while being carried by Ethernet in between. This whole arrangement, of course, has

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 10 of 11
INFOEXPRESS, INC. - EX2566

nothing whatsoever to do with VLANs. But standards' committee politics is not unlike regular politics: if you vote for my bit, I will vote for your bit.

As we mentioned above, when a tagged frame arrives at a VLAN-aware switch, the switch uses the VLAN ID as an index into a table to find out which ports to send it on. But where does the table come from? If it is manually constructed, we are back to square zero: manual configuration of bridges. The beauty of the transparent bridge is that it is plug-and-play and does not require any manual configuration. It would be a terrible shame to lose that property. Fortunately, VLAN-aware bridges can also autoconfigure themselves based on observing the tags that come by. If a frame tagged as VLAN 4 comes in on port 3, then apparently some machine on port 3 is on VLAN 4. The 802.1Q standard explains how to build the tables dynamically, mostly by referencing appropriate portions of Perlman's algorithm standardized in 802.1D.

Before leaving the subject of VLAN routing, it is worth making one last observation. Many people in the Internet and Ethernet worlds are fanatically in favor of connectionless networking and violently opposed to anything smacking of connections in the data link or network layers. Yet VLANs introduce something that is surprisingly similar to a connection. To use VLANs properly, each frame carries a new special identifier that is used as an index into a table inside the switch to look up where the frame is supposed to be sent. That is precisely what happens in connection-oriented networks. In connectionless networks, it is the destination address that is used for routing, not some kind of connection identifier. We will see more of this creeping connectionism in Chap. 5.

4.8 SUMMARY

Some networks have a single channel that is used for all communication. In these networks, the key design issue is the allocation of this channel among the competing stations wishing to use it. Numerous channel allocation algorithms have been devised. A summary of some of the more important channel allocation methods is given in Fig. 4-52.

The simplest allocation schemes are FDM and TDM. These are efficient when the number of stations is small and fixed and the traffic is continuous. Both are widely used under these circumstances, for example, for dividing up the bandwidth on telephone trunks.

When the number of stations is large and variable or the traffic is fairly bursty, FDM and TDM are poor choices. The ALOHA protocol, with and without slotting, has been proposed as an alternative. ALOHA and its many variants and derivatives have been widely discussed, analyzed, and used in real systems.

When the state of the channel can be sensed, stations can avoid starting a transmission while another station is transmitting. This technique, carrier sensing, has led to a variety of protocols that can be used on LANs and MANs.

CISCO SYSTEMS, INC. et, al. v. INFOEXPRESS, INC. [IPR2024-00679]
Page 11 of 11
INFOEXPRESS, INC. - EX2566