October 28, 1997

To All IEEE P802.3 CSMA/CD Working Group Voters, Observers, and Liaisons:

This is the Working Group **Recirculation** ballot draft, D2.0, of IEEE P**802.3aa Maintenance Revision #5** (100BASE-T.) This document has been revised since D1.1 to reflect the resolution of comments received during the Working Group ballot which closed July 3, 1997.

The cover page of the document contains additional information concerning its status. IEEE Standards Committee participants may reproduce this document for purposes of IEEE standardization activities. Please see additional copyright information on the front page of the document.

The major difference between D 2.0 and D 1.1 is the modification of changes to the partition state diagram for port X so as to ensure continued support for legacy implementations.

The **opening date of the ballot is October 28th, 1997**. The **closing date of the ballot is November 13th, 1997**, which provides for the required 10-day recirculation ballot cycle plus required mail time for distribution. Please return your ballots and comments via email (preferred) or fax as soon as possible. We wish to deal with all issues at the Montreal meeting. We will also accept ballots in person at the Montreal meeting.

During a recirculation it is not necessary for you to respond unless you wish to change your vote from that which was recorded during the initial Working Group Ballot. For example, if you voted "Approve" or "Abstain" during the initial ballot, and you see no reason to change your vote, then it is not necessary for you to respond. However, if you wish to submit additional comments, or if you wish to change your ballot from what has already been recorded, then you must respond to this recirculation ballot within the allotted time. The ballot will close promptly on schedule.

The ballot draft D2.0 of IEEE P802.3aa Maintenance Revision #5 (100BASE-T) consists of 15 pages, including the cover sheet and defines three proposed changes to 802.3u and is attached.

On behalf of the IEEE P802.3, we ask that you please give this document your most careful review and consideration. We look forward to receiving your ballots and comments on IEEE P802.3aa.

Geoffrey O. Thompson
Chair, IEEE P802.3 CSMA/CD Working Group
Bay Networks, Inc.
4401 Great America Parkway
PO Box 58185, MS SC01-05
Santa Clara, CA 95052-8185
email: geoff_thompson@baynetworks.com
phone: +1 408 495 1339
fax: +1 408 988 5525

Colin K. Mick.
Chair, IEEE P802.3aa
The Mick Group
2130 Hanover St.
Palo Alto, CA 94306
email: ckm@mickgroup.com
phone: +1 650-856-3666
fax : +1 650-494-3737

802.3 WORKING GROUP
RECIRCULATION BALLOT

IEEE802.3 WG Ballot 1st WG Recirculation Ballot, 10/28/97
Subject: P802.3aa Maintenance #5 (100BASE-T)
(Revision to CSMA/CD, 100 Mb/s Operation)

BALLOT RESPONSE:
RESPONSE IS DUE BY E-MAIL/FAX **NO LATER THAN November 13**, 1997

_____ 802.3 VOTER

_____ APPROVE

_____ APPROVE WITH COMMENTS (Comments are non-binding)

_____ DO NOT APPROVE(Must attach specific comments for remedy)

_____ ABSTAIN

RETURN BALLOT (with comments via internet e-mail or FAX TO:
Colin Mick
Chairman, 100BASE-T Maint.
The Mick Group
2130 Hanover
Palo Alto, CA 94306 USA

Phone: +1 650 856 3636
FAX: +1 650 494 3737
Internet E-Mail: ckm@ix.netcom.com

RETURN COPY & COMMENTS BY E-MAIL OR FAX TO:
Geoff Thompson M/S SC01-05
Chairman, IEEE 802.3
Bay Networks, Inc.
4401 Great America Parkway
Post Office Box 58185
Santa Clara, CA 95052-8185 USA

Phone: +1 408 495 1339
FAX : +1 408 988 5525
Internet E-Mail: thompson@baynetworks.com

Signature: _____

Date: _____

Print Name: _____

Draft revision to

ANSI/IEEE Std 802.3, 1996 Edition

Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method & Physical Layer Specifications:

Maintenance Revision #5 (100BASE-T)

Sponsor: LAN MAN Standards Committee
of the
IEEE Computer Society

This is Draft 2.0 with items approved for balloting by IEEE 802.3 and approved by a Working Group Ballot which closed July 3, 1997. This draft includes changes approved at the 802.3 Plenary meeting in Lahaina, HI the week of July 7-10, 1997 in response to comments received during the initial Working Group balloting. This draft is for a 10-day recirculation ballot and is scheduled to close in time for comment resolution at the 802.3 Plenary meeting in Montreal, Canada the week of November 10-13, 1997. This document expires November 15, 1997.

P802.3aa, 802.3 MAINTENANCE BALLOT #5 (100BASE-T)

CHANGE 1: Changes to Partition State Diagram for Port X (Figure 27-8)
Clause 27: Repeater for 100Mb/s baseband networks
Proposed by: Lloyd Oliver
Approved for ballot 11/12/96

PROPOSED CHANGES

In Figure 27-8 Repeater Partition State Diagram for Port X:

1) Change transition:

COLLISION COUNT IDLE to WATCH FOR COLLISION

Update to read:

(scarrier_present(x) = true) + (command(x) ≠ quiet)

This ensures that the COLLISION COUNT IDLE state is exited for both receive activity (scarrier_present(x) = true) and transmit activity (command(x) ≠ quiet). The term (command(x) ≠quiet) has to be ORed in to ensure transmits also cause an exit from the COLLISION COUNT IDLE state.

2) Change transition:

WATCH FOR COLLISION to COLLISION COUNT IDLE

Update to read:

(scarrier_present(x) = false) * (command(x) = quiet)

This ensures that the collision counter is neither incremented nor cleared if both transmit and receive activity have ceased before the no_collision_timer has completed. The term (command(x) = quiet) has to replace the term (command(x) ≠ collision) to ensure that if the WATCH FOR COLLISION state is entered due to a transmit, it remains there until that transmit ~~is~~ is completed.

3) Change Transition:

WATCH FOR COLLISION to COLLISION COUNT INCREMENT

Updateto read:

(command(x) = collision) * (scarrier_present(x) = true)

This change ensures that the counter is only incremented when a collision is happening on port X. The term (scarrier_present(x) = true) is ANDed to qualify the fact that a collision is occurring and the fact that port x is receiving and hence taking part in a collision.

4) Change transition:

WATCH FOR COLLISION to CLEAR COUNTER

Update to read:

no_collision_timer_Done * (command(x) ≠ collision) * ((scarrier_present(x) = true) + (command(x) = copy))

This change ensures that either a transmit or a receive for duration greater than no_collision_timer will reset the collision counter.

5) Change transition:

PARTITION HOLD to PARTITION COLLISION WATCH

Update to read:

(command(x) ≠ quiet) + (scarrier_present(x) = true)

This change ensures a receive or a transmit will start the no_collision_timer and un-partition the port once the timer is done.

6) Change transition:

PARTITION COLLISION WATCH to PARTITION WAIT

Update to read:

(command(x) ≠ quiet) * (scarrier_present(x) true)

This change ensures that if a receive is occurring while the port is transmitting (Command(x) = copy or collision) the port remains partitioned.

7) Change transition:

PARTITION COLLISION WATCH to WAIT TO RESTORE PORT

Update to read:

no_collison_timer_Done * (((scarrier_present(x) = false) * (command(x) = copy)) + ((scarrier_present(x) = true) * (command(x) = quiet)))

This change ensures that ~~a receive~~ receipt of a good packet as well ~~transmitting~~ as transmission of a packet without contention restores the port to full operation.

8) Change Text in sub-clause 27.3.1.6 (Second paragraph):

Change "The count shall be incremented on each transmission that suffers a collision and shall be reset on a successful transmission." to read "The count shall be incremented on each collision and shall be reset on a transmit or receive ~~without incurring a collision.~~ event which continues beyond the duration of no collision_timer."

9) Change Text in subclause 27.7.4.8 (PICS items PA3 & PA4).

Change PA3 comment "Count incremented on each transmission that suffers a collision" to "Count incremented on a collision".

Change PA4 comment "Count reset on successful collision" to "Count reset on a transmit or receive ~~without incurring a collision~~ event which continues beyond the duration of no_collision_timer".

Rationale for revision:

1. Cater for "capture" effect on heavily loaded network where one node could be streaming lots of packets under burst mode with a second node experiencing short term unfairness once it lost the first collision leading it to see a few max. collisions which can lead to partition of the node as only successful receives (from the repeaters viewpoint) currently clear the collision count. Newer protocols being investigated are looking at windows of up to 128kbytes which may highlight this behavior more.

2.To fix the inconsistencies between the text describing the partition function (27.3.1.6) and the state diagram of figure 27-8.

3.To modify the Partition state machine to be the same as the 10Mb/s partition algorithm of chapter 9 providing symmetrical operation (Partition and Un-Partition for the same reasons.)

Impact on existing networks:

Reduces or eliminates the potential to partition a 100Mb/s network port unless a real fault has occurred rather than normal collisions. This proposal is a "superset" of the existing state diagram and would inter-operate with repeaters implementing existing state diagram.

CHANGE 2: Change to Repeater Core State Diagram (Figure 27-2)
Clause 27: Repeater for 100Mb/s baseband networks
Proposed by: Lloyd Oliver
Approved for ballot 11/12/96
PROPOSED CHANGE

In Figure 27-2 Repeater Core State Diagram, change exit term out of ACTIVE state.

FROM: (activity(ALL) = 0) * (all_data_sent = true)

TO: (activity(N) = 0) * (all_data_sent = true)

Rationale for revision:

Fixes anomaly of repeater operation during collision storms at high traffic loading. During these periods, inter-packet gaps seen from the perspective of the Repeater Core State Diagram regarding multiple ports, can become zero or, in fact, overlap. The proposed change allows the core to enable sourcing packet activity ~~form~~ from the remaining port active and reassign N based on real activity rather than going blind until all network activity stops.

Current operation isolates port N such that a station (or repeater) attached to this port is isolated from carrier activity until all carrier activity is halted. This effects stations connected (in any way; i.e., through a repeater or direct) to port N from properly deferring to network traffic.

Impact on existing networks:

Reduces or eliminates late collision events recorded by management tools during periods of peak traffic loading.

CHANGE 3: Change Repeater Partition State Diagram (Figure 27-8)
Clause 27: Repeater for 100Mb/s baseband networks
Proposed by: Lloyd Oliver
Approved for ballot 11/12/96
PROPOSED CHANGE

In Figure 27-8 Partition State Diagram, change right exit term out of COLLISION COUNT INCREMENT state to PARTITION WAIT state;

FROM: CC(X) ≥ CCLimit

TO: CC(X) ≥ CCLimit + jabber_timer_done

(Editor's note: "_" in D 1.0 changed to "≥" in D2.0 for both statements above)

Rationale for revision:

Consider a case where loop back plugs are present on multiple ports of a repeater. Further, that a station on yet another port initiates a packet transmission. The loop back plugs cause a collision and the station backs off. But, the event lingers. The ports with loop back plugs cause the Repeater Core State Diagram to remain in the JAM state since "activity(ALL) > 1". Which, of course, holds all of the Transmit state machines in the COLLISION state, sending Jam. The loop back continues. This situation is maintained until the Receive Timer State Diagram intervenes when the "jabber_timer_done" interrupts the looped back carrier by forcing the Receive State Diagram into SILENT and the Transmit State Diagram (for that port) into QUIET. But, only briefly. The Receive State Diagram would reset to NO INPUT state when carrier is interrupted. However, due to the tolerancing of the Jabber timers, other ports with loop back plugs are still holding the Repeater

Core in JAM state. This cycling repeats ad infinitum. Even the Partition State diagram can't help because it is stuck in the COLLISION COUNT INCREMENT state waiting for the core to issue a quiet command. The change listed above cures this situation by isolating the offending ports with loop back plugs attached.

Impact on existing networks:

Clears net for working traffic sooner when loop back plugs are present in system.

CHANGE 4: A series of editorial changes to Clause 22.

Proposed by: Bob Grow during 802.3aa Working Group ballot

Editorial changes only, published for record

PROPOSED CHANGES:

Change "registers 2 through 7" in paragraph 3 of 22.2.4 to read "registers 2 through 10".

Change "1.15:11" to read "1.15:9" in paragraph 1 of 22.2.4.1.3

Change "1.15:11" to read "1.15:9" in paragraph 2 of 22.2.4.1.3

Change "1.15:11" to read "1.15:9" in paragraph 1 of 22.2.4.1.8

Change "1.15:11" to read "1.15:9" in paragraph 3 of 22.2.4.1.8

Change "4,5,6, and 7" to read "4,5,6,7,and 8" in 22.2.4.2.10 (two occurrences)

Change "six registers" to read "nine registers" in paragraph 2 of 22.2.4.3.

Change the subclause references in 22.7.3.4, MF39 through MF51 to reflect subclause numbering changes caused by the insertion of two new subsections (22.2.4.2.6 and 22.2.4.2.7) and the renumbering of current subsections 22.2.4.2.6 through 22.2.4.2.13 as per changes made by 802.3x and 802.3y.

Rationale for revision:

This is a series of editorial corrections to Clause 22 to harmonize the text with changes made by 802.3y. These changes were identified by a reviewer to the 802.3aa ballot and are posted here as an editorial record. These corrections were inserted during the publication of 802.3x and 802.3y.

Impact on existing networks:

None

CHANGED TEXT

NOTE: The editing instructions contained in this supplement define how to merge the material contained herein into IEEE Std 802.3u-1995.

The editing instructions are shown in ***bold italic***. Three editing instructions are used: change, insert, and replace. Change is used to make small corrections in existing text or tables. The editing instruction specifies the location of the change and describes what is being changed either by using ~~strikethrough~~ (to remove old material) and underscore (to add new material). Insert adds new material without disturbing the existing material. Insertions may require renumbering. If so, renumbering instructions are given in the editing instruction. Replace is used to make large changes in existing text, subclauses, tables, or pages by removing existing material and replacing it with new material. When modifications are made to paragraphs of existing text, deletions are shown in ~~strikethrough type~~ and additions are underscored. Editorial notes will not be carried over into future editions.

Change 27.3.1.6 as follows (change bars, underline and strikethrough are against base document):-

27.3.1.6 Partition functional requirements

In large multisegment networks it may be desirable that the repeater set protect the network from some fault conditions that would disrupt network communications. A potentially likely cause of this condition could be due to a cable fault.

Each repeater PMA interface shall contain a self-interrupt capability as described in figure 27-8 to prevent a faulty segment's carrier activity from reaching the repeater unit and hence propagating through the network. The repeater PMA interface shall count ~~consecutive~~ collisions. The count shall be incremented on each transmission that suffers a collision ~~and shall be reset on a successful transmission~~. The count shall be reset on a carrier event of duration in excess of no_collision_timer (see 27.3.2.1.4) without incurring a collision. If this count reaches ~~exceeds~~ the value CCLimit (see 27.3.2.1.1) the Partition condition shall be detected.

Upon detection of Partition, the port shall perform the following:

a) Inhibit sending further input messages to the repeater unit.
b) Continue to output messages from the repeater unit.
c) Continue to monitor activity on that PMA interface.

The repeater shall reset the Partition function when one of the following conditions is met:

a) On power-up reset.
b) The repeater has transmitted ~~detected activity~~ on the port for a duration in excess of ~~more than the number of bits specified for~~ no_collision_timer (see 27.3.2.1.4) without incurring a collision.

NOTE: It is possible that under some network conditions the partition state machine will partition a port due to normal network collisions rather than a fault condition. It is also possible that some double fault conditions will remain undetected. To reduce the likelihood of these events occurring the following optional measures, as described in figure 27-8, are recommended.
(1) The collision count is additionally reset when the repeater has transmitted on the port for a duration in excess of no_collision_timer without detecting a collision.
(2) The Partition function is additionally reset when the repeater has received activity on the port for a duration in excess of no_collision_timer without detecting a collision.
(3) The Partition condition is additionally detected due to a carrier event of duration in excess of jabber_timer (see 27.3.1.7) in which a collision has occurred.

Insert the following after "partition(X)" in 27.3.2.1.2:-

part_opt(X)

Implementation option. Either value may be chosen for repeater implementation (see 27.3.1.6).

Values: true; port supports the recommended optional measures in the partition state machine.
false; port does not support the recommended optional measures in the partition state machine.

Change the "jabber_timer" in 27.3.2.1.4 as follows (change bars, underline and strikethrough are against base document):-

jabber_timer

Timer for length of carrier which must be present before the Jabber state ~~is entered~~ (27.3.1.7), and optionally during a collision the Partition state (27.3.1.6), is entered. The timer is done when it reaches 40 000 – 75 000 BT.

Insert the following addition PICS entry to the end of 27.7.3 as follows:-

27.7.3 Major capabilities/options

*OPF	Partition function supports the recommended optional measures as described	27.3.1.6	O		

Change 27.7.4.8 as follows (change bars, underline and strikethrough are against base document):-

27.7.4.8 Partition function

Item	Feature	Subclause	Status	Support	Value/Comment
PA1	Partition function implementation	27.3.1.6	M		Self-interrupt of data reception
PA2	~~Consecutive~~ Collision count for entry into partition state	27.3.1.6	M		~~Consecutive~~ Collision count greater than or equal to ~~in excess of~~ CCLimit
PA3	~~Consecutive~~ Collision counter incrementing	27.3.1.6	M		Count incremented on a collision ~~on each transmission that suffers a collision~~.
~~PA4~~ PA4 PA5	~~Consecutive~~ Collision counter reset	27.3.1.6	~~M~~ M OPF:M		~~Count reset on successful collision~~. Count reset on receive activity in excess of no_collision_timer without collision Count reset on transmission in excess of no_collision_timer without collision
PA~~5~~6	Messages sent to repeater unit in Partition state	27.3.1.6	M		Inhibited sending messages to repeater unit
PA~~6~~7	Messages sent from repeater unit in Partition state	27.3.1.6	M		Continue sending output messages
PA~~7~~8	Monitoring activity on PMA interface in Partition state	27.3.1.6	M		Continue monitoring activity at PMA interface
~~PA8~~ PA9 PA10	Reset of Partition state	27.3.1.6	~~M~~ M OPF:M		~~Power-up reset or Detecting activity for greater than duration no_collision_timer without a collision~~ Power-up reset or transmission in excess of no_collision_timer without collision Receive activity in excess of no_collision_timer without collision
PA11	Excessive carrier entry into Partition state	27.3.1.6	OPF:M		Carrier duration in excess of jabber_timer in which a collision occurs

Replace Figure 27-2 as follows:-

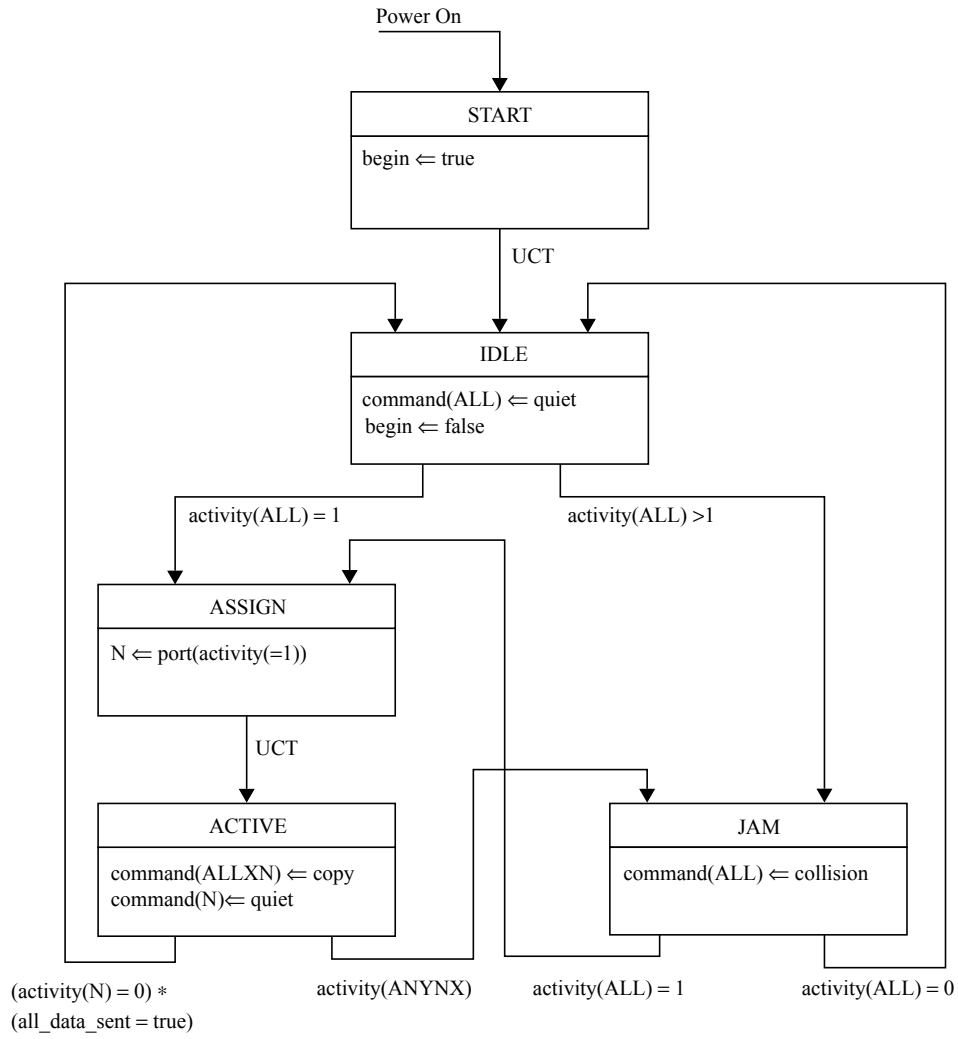

Figure 27-2—Repeater core state diagram (AS REVISED BY THESE CHANGES)

Replace Figure 27-8 as follows:-

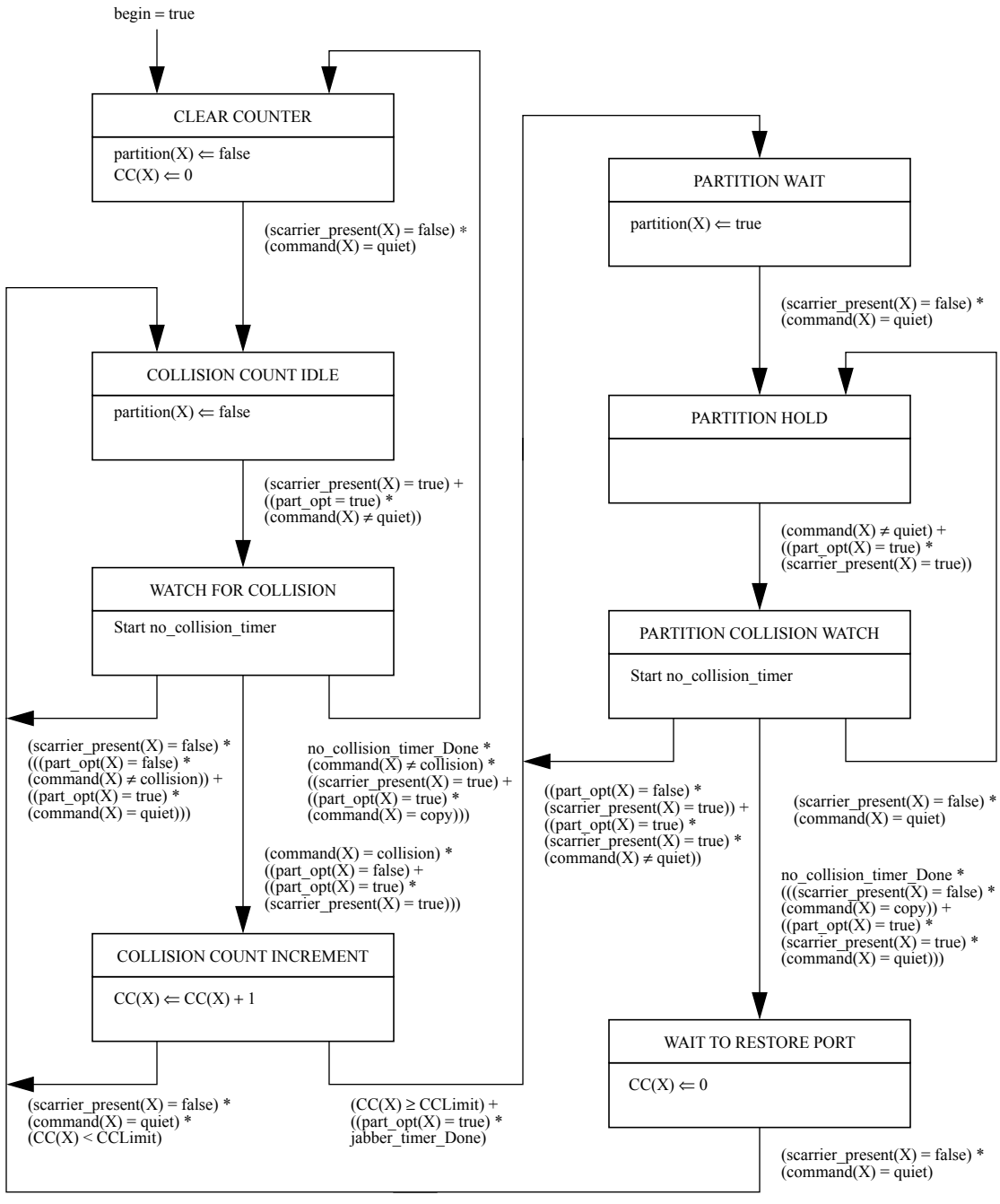

Figure 27-8—Partition state diagram for port X (AS REVISED BY THESE CHANGES)

Old Figures

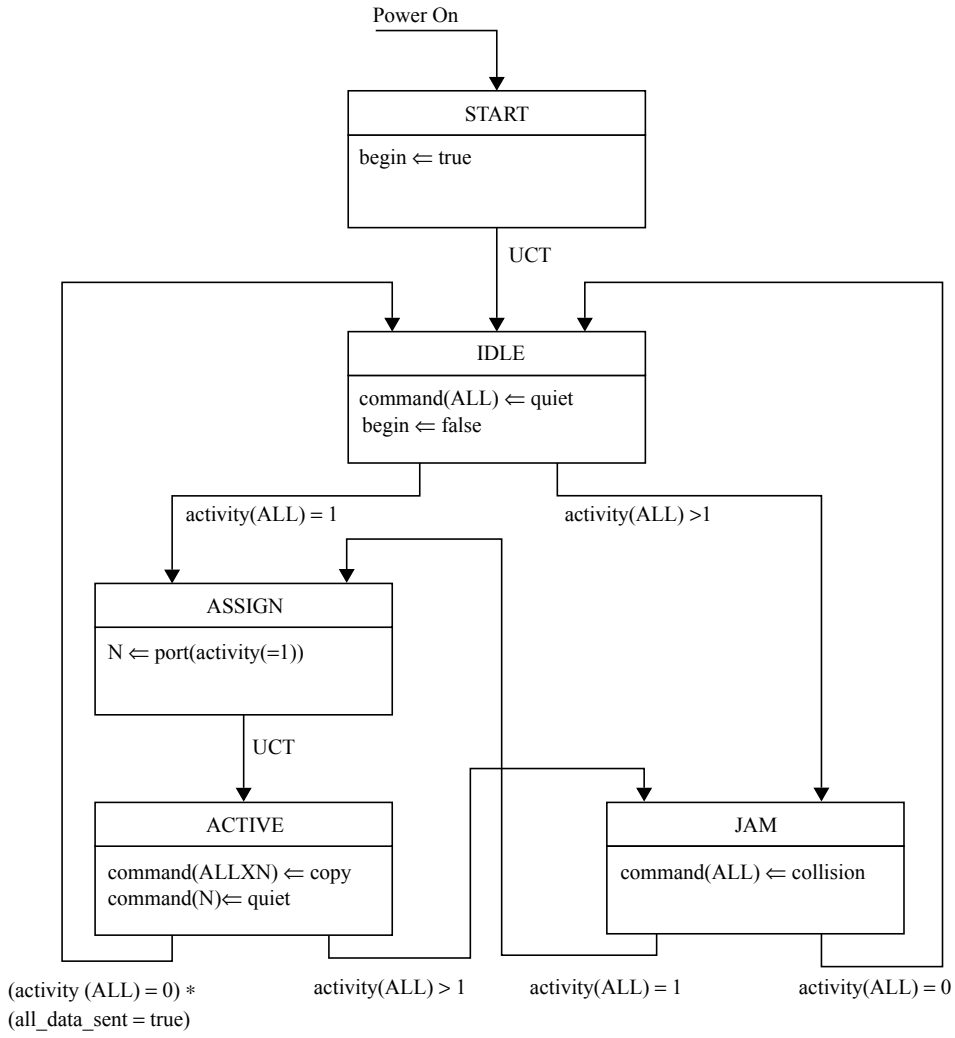

Figure 27-2—Repeater core state diagram (AS CURRENTLY PUBLISHED)

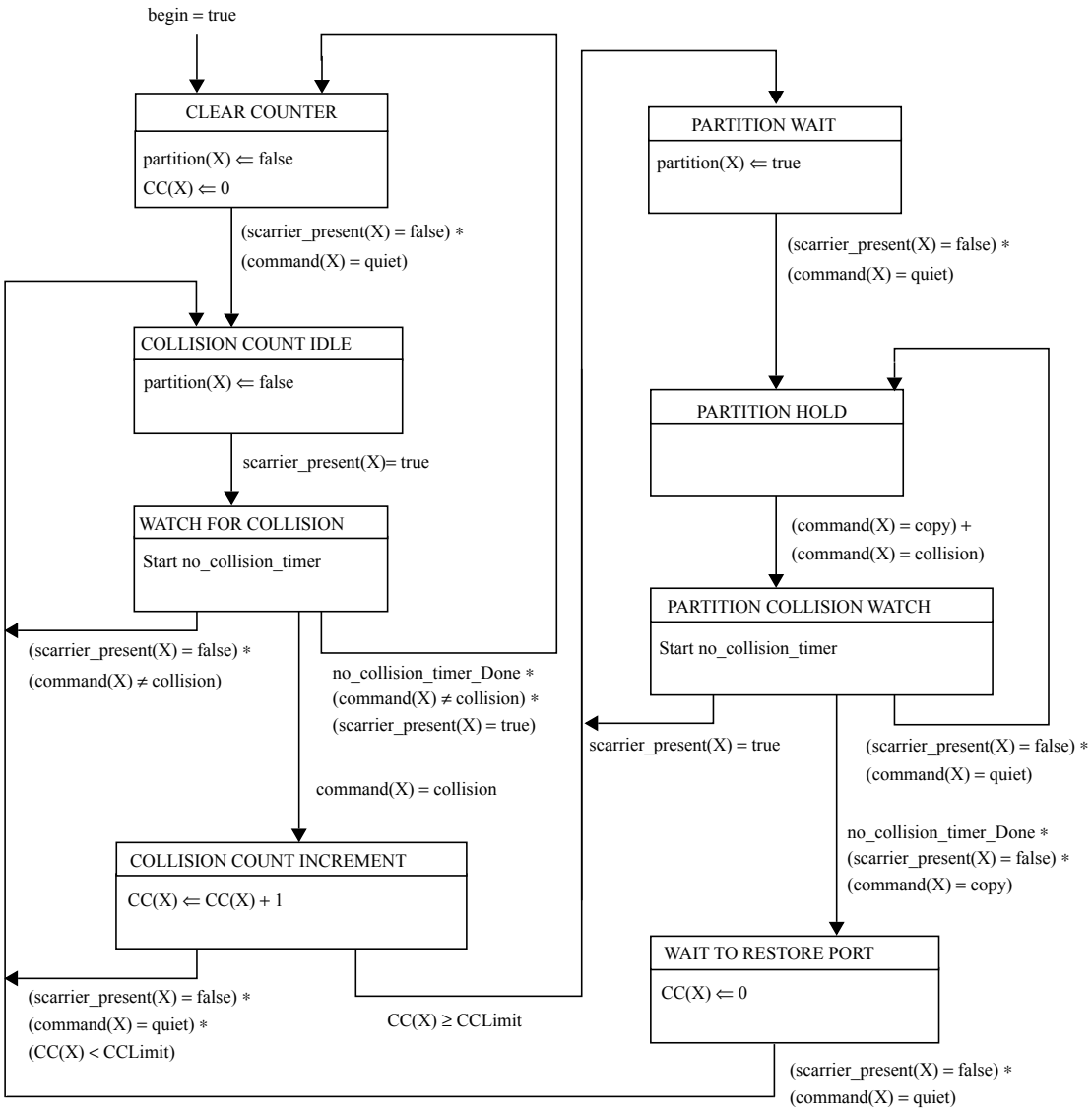

Figure 27-8—Partition state diagram for port X (AS CURRENTLY PUBLISHED)

Change "registers 2 through 7" in paragraph 3 of 22.2.4 to read "registers 2 through 10".

The basic register set consists of two registers referred to as the Control Register (register 0) and the Status Register (register 1). The status and control functions defined here are considered basic and fundamental to 100 Mb/s PHYs. All PHYs that provide an MII shall incorporate the basic register set. Registers 2 through ~~7~~ 10 are part of the extended register set.

Change "1.15:11" to read "1.15:9" in paragraph 1 of 22.2.4.1.3
Change "1.15:11" to read "1.15:9" in paragraph 2 of 22.2.4.1.3

Link speed can be selected via either the Auto-Negotiation process, or manual speed selection. Manual speed selection is allowed when Auto-Negotiation is disabled by clearing bit 0.12 to zero. When Auto-Negotiation is disabled, setting bit 0.13 to a logic one configures the PHY for 100 Mb/s operation, and clearing bit 0.13 to a logic zero configures the PHY for 10 Mb/s operation. When Auto-Negotiation is enabled, bit 0.13 can be read or written, but the state of bit 0.13 has no effect on the link configuration, and it is not necessary for bit 0.13 to reflect the operating speed of the link when it is read. If a PHY reports via bits 1.15:~~11~~ 9 that it is able to operate at only one speed, the value of bit 0.13 shall correspond to the speed at which the PHY can operate, and any attempt to change the setting of the bit shall be ignored.

The default value of bit 0.13 is one, unless the PHY reports via bits 1.15:~~11~~ 9 that it is able to operate only at 10 Mb/s, in which case the default value of bit 0.13 is zero.

Change "1.15:11" to read "1.15:9" in paragraph 1 of 22.2.4.1.8
Change "1.15:11" to read "1.15:9" in paragraph 3 of 22.2.4.1.8

The duplex mode can be selected via either the Auto-Negotiation process, or manual duplex selection. Manual duplex selection is allowed when Auto-Negotiation is disabled by clearing bit 0.12 to zero. When Auto-Negotiation is disabled, setting bit 0.8 to a logic one configures the PHY for full-duplex operation, and clearing bit 0.8 to a logic zero configures the PHY for half-duplex operation. When Auto-Negotiation is enabled, bit 0.8 can be read or written, but the state of bit 0.8 has no effect on the link configuration. If a PHY reports via bits 1.15:~~11~~ 9 that it is able to operate in only one duplex mode, the value of bit 0.8 shall correspond to the mode in which the PHY can operate, and any attempt to change the setting of bit 0.8 shall be ignored.

When a PHY is placed in the loopback mode of operation via bit 0.14, the behavior of the PHY shall not be affected by the state of bit 0.8.

The default value of bit 0.8 is zero, unless a PHY reports via bits 1.15:~~11~~ 9 that it is able to operate only in full-duplex mode, in which case the default value of bit 0.8 is one.

Change "4,5,6, and 7" to read "4,5,6,7,and 8" in 22.2.4.2.10 (two occurrences)

When read as a logic one, bit 1.5 indicates that the Auto-Negotiation process has been completed, and that the contents of registers 4, 5, 6,~~ and 7~~ 7, and 8 are valid. When read as a logic zero, bit 1.5 indicates that the Auto-Negotiation process has not been completed, and that the contents of registers4, 5, 6,~~ and 7~~ 7, and 8 are meaningless. A PHY shall return a value of zero in bit 1.5 if Auto-Negotiation is disabled by clearing bit 0.12. A PHY shall also return a value of zero in bit 1.5 if it lacks the ability to perform Auto-Negotiation.

Change "six registers" to read "nine registers in paragraph 1 of 22.2.4.3.

In addition to the basic register set defined in 22.2.4.1 and 22.2.4.2, PHYs may provide an extended set of capabilities that may be accessed and controlled via the MII management interface.~~ Six~~ Nine registers have been defined within the extended address space for the purpose of providing a PHY-specific identifier to layer management, and to provide control and monitoring for the Auto-Negotiation process.

Change the subclause references in 22.7.3.4, MF39 through MF51 to reflect subclause numbering changes caused by the insertion of two new subsections (22.2.4.2.6 and 22.2.4.2.7) and the renumbering of current subsections 22.2.4.2.6 through 22.2.4.2.13 as per changes made by 802.3x and 802.3y.

MF39	Reserved bits ignored when read	22.2.4.2.~~6~~ 8	M		
MF40	PHY returns 0 in reserved bits	22.2.4.2.~~6~~ 8	M		
MF41	PHY returns 0 if Auto-Negotiation disabled	22.2.4.2.~~8~~ 10	M		Yes (1.5 = 0 when 0.12 = 0)
MF42	PHY returns 0 if it lacks ability to perform Auto-Negotiation	22.2.4.2.~~8~~ 10	M		Yes (1.5 = 0 when 1.3 = 0)
MF43	Remote fault has latching function	22.2.4.2.~~9~~ 11	M		Yes (once set will remain set until cleared)
MF44	Remote fault cleared on read	22.2.4.2.9 11	M		Yes
MF45	Remote fault cleared on reset	22.2.4.2.~~9~~ 11	M		Yes (when 0.15 = 1)
MF46	PHY without remote fault returns value of zero	22.2.4.2.~~9~~ 11	M		Yes (1.4 always 0)
MF47	Link status has latching function	22.2.4.2.~~11~~ 13	M		Yes (once cleared by link failure will remain cleared until read by MII)
MF48	Jabber detect has latching function	22.2.4.2.~~12~~ 14	M		Yes (once set will remain set until cleared)
MF49	Jabber detect cleared on read	22.2.4.2.~~12~~ 14	M		
MF50	Jabber detect cleared on reset	22.2.4.2.~~12~~ 14	M		
MF51	100BASE-T4 and 100BASE-X PHYs return 0 for jabber detect	22.2.4.2.~~12~~ 14	M		Yes (1.1 always = 0 for 100BASE-T4 and 100BASE-TX)