Modified Clause 73 Auto-Negotiation Detailed Proposal

IEEE 802.3bp - Plenary Meeting - March 2014

William Lo, Marvell

MARVELL®

Agenda

▶ **Brief Recap on Theory of Operation**

▶ **State Machine Changes**
  - **3 Existing State Machine + New one**

▶ **Timer and Variable Changes**

▶ **Possible Optimizations**

▶ **Electricals**

Half Duplex Differential Manchester Encoded Auto-Negotiation

▶ **Use modified 802.3ap Clause 73 Auto-Negotiation for Backplane**

▶ **Differential Manchester Encoding instead of linkpulses**
  - **But not continuously transmitted (half duplex)**
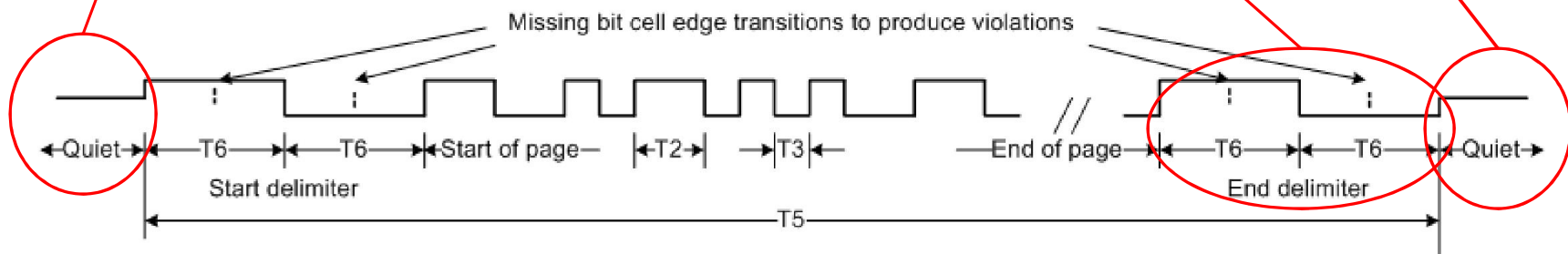 - **DC balanced, Partially Randomized**

MARVELL®
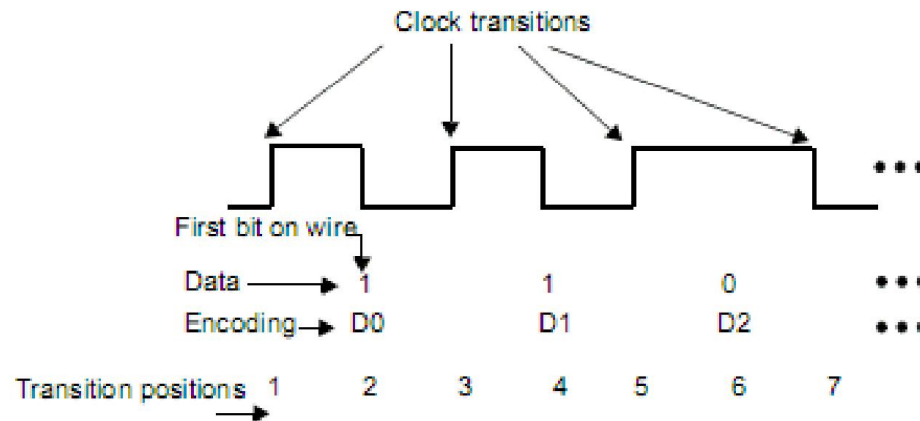
Auto-Negotiation DME Page

▶ **DME Page consists of 48 data bits + 1 random bit + delimiters**

▶ **What's New:**

  ▪ **Not continuously transmitted – DME pages separated by quiet time**
  ▪ **Extra delimiter added at end of DME page to delineate from quite time**

▶ **Proposed values**

#	Parameter	Min	Typ	Max	Units
T1	Transmit position spacing (period)	39	40	41	ns
T2	Clock transition to clock transition	76	80	84	ns
T3	Clock transition to data transition (data = 1)	36	40	44	ns
T4a	+1 to -1 or -1 to +1 transitions in a DME page	52	-	101	-
T4b	0 to +/1 or +/-1 to 0 transitions in a DME page	2	2	2	-
T5	DME page width	4446	4560	4674	ns
T6	DME Manchester violation delimiter width	156	160	164	ns

4

Auto-Negotiation DME Page

▶ **The 49 data bit is a random bit is generated according to Clause 73: $X^7 + X^3 + 1$ or $X^7 + X^6 + 1$**

▶ **The generator is advanced once per DME page**

▶ **The DME page is polarity insensitive and the random bit has the effect of inverting the DME page polarity.**

▶ **The polarity of the next DME page shall be such that Start delimiter is the same polarity as the End delimiter of the current DME page**

5

MARVELL®

Half Duplex Operation

▶ **PHYs take turns sending and receiving DME pages**

▶ **Blind period to avoid seeing echo from self**

▶ **After page received Wait (Silent) period needed since link partner may be in Blind period**

▶ **Wait period slightly longer than Blind Period**

▶ **Examples below with max cable length and zero cable length**
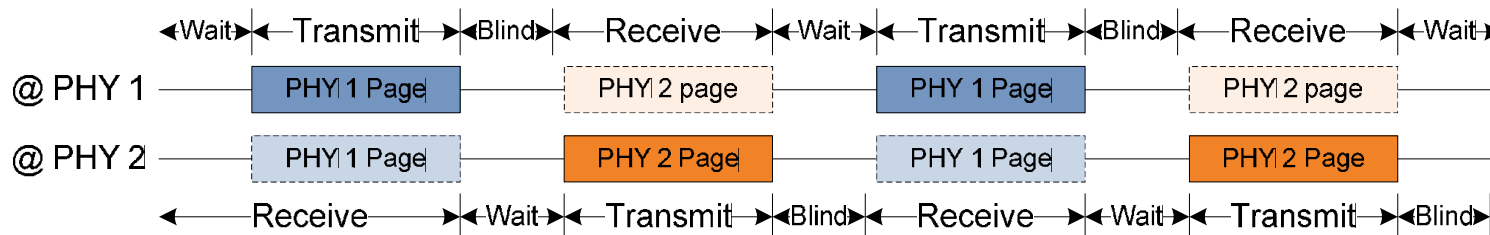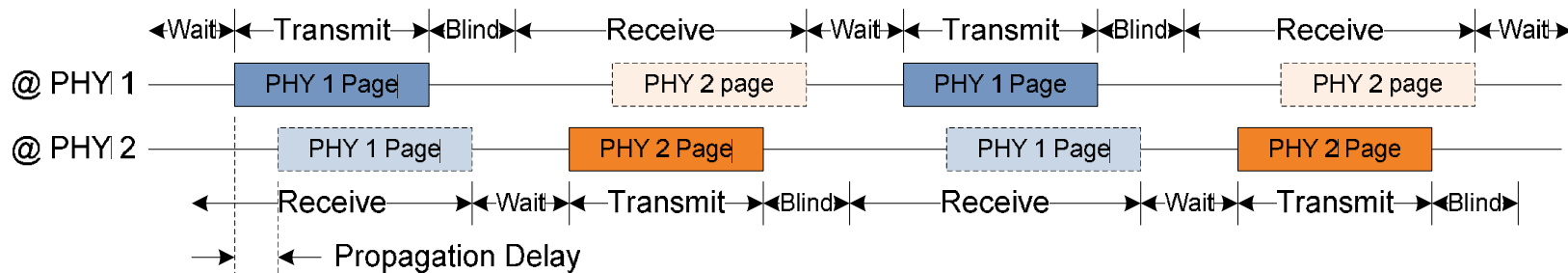
Signal Separation on Startup

▶ **Both transmit overlap and no page seen – receive times out**

  ▪ **Timeout value should be longer than DME_page_width + Wait (Silent) period**

▶ **PHYs will randomly decide whether to delay zero or more additional Wait period while still keeping receiver on**

  ▪ **Assuming 16 possible delay settings 1 in 1 trillion chance of going beyond 10 tries**

▶ **PHY that waits more should receive page and should sync up system and no more random waits are needed**

▶ **At startup PHY should always listen for at least one Wait period prior to transmitting first page**
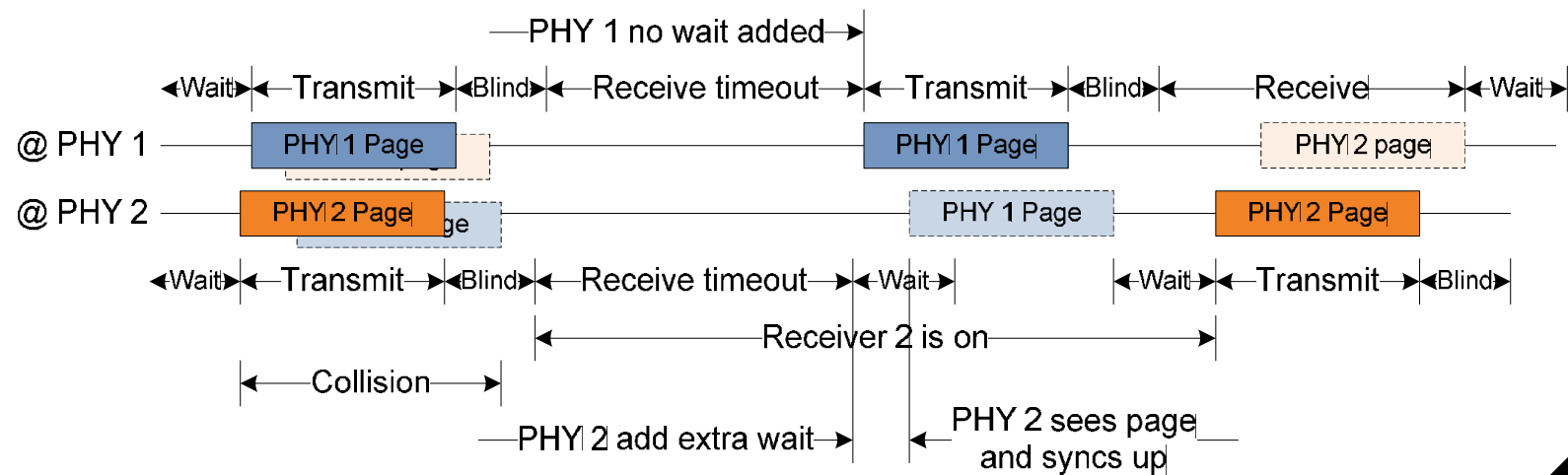
Clause 73 Arbitration State Machine Changes

▶ **Remove states related to parallel detection since there is no legacy PHY without Auto-Negotiation to support**

▶ **SCAN_FOR_CARRIER does not apply since there is no legacy PHY generating signals**

▶ **Otherwise no changes**

Clause 73 Arbitration State Machine Changes

PARALLEL DETECTION FAULT
mr_parallel_detection_fault ⇐ true
link_control_[al] ⇐ DISABLE

UCT

LINK STATUS CHECK
Start autoneg_wait_timer
transmit_disable ⇐ true

single_link_ready=false

link_status_[KX]=OK +
link_status_[KX4]=OK

single_link_ready=true *
autoneg_wait_timer_done

States removed since
no parallel detect is
supported or needed
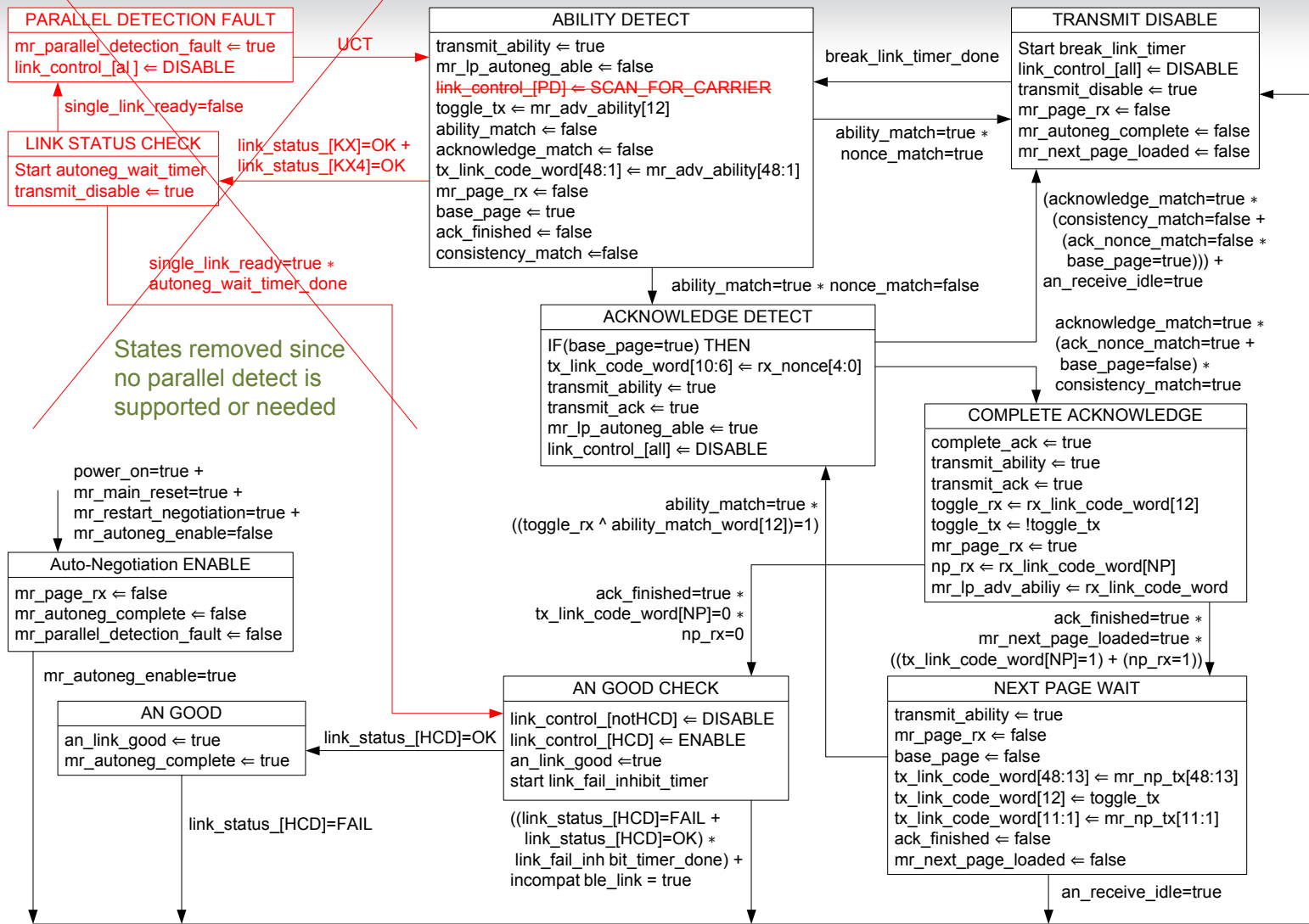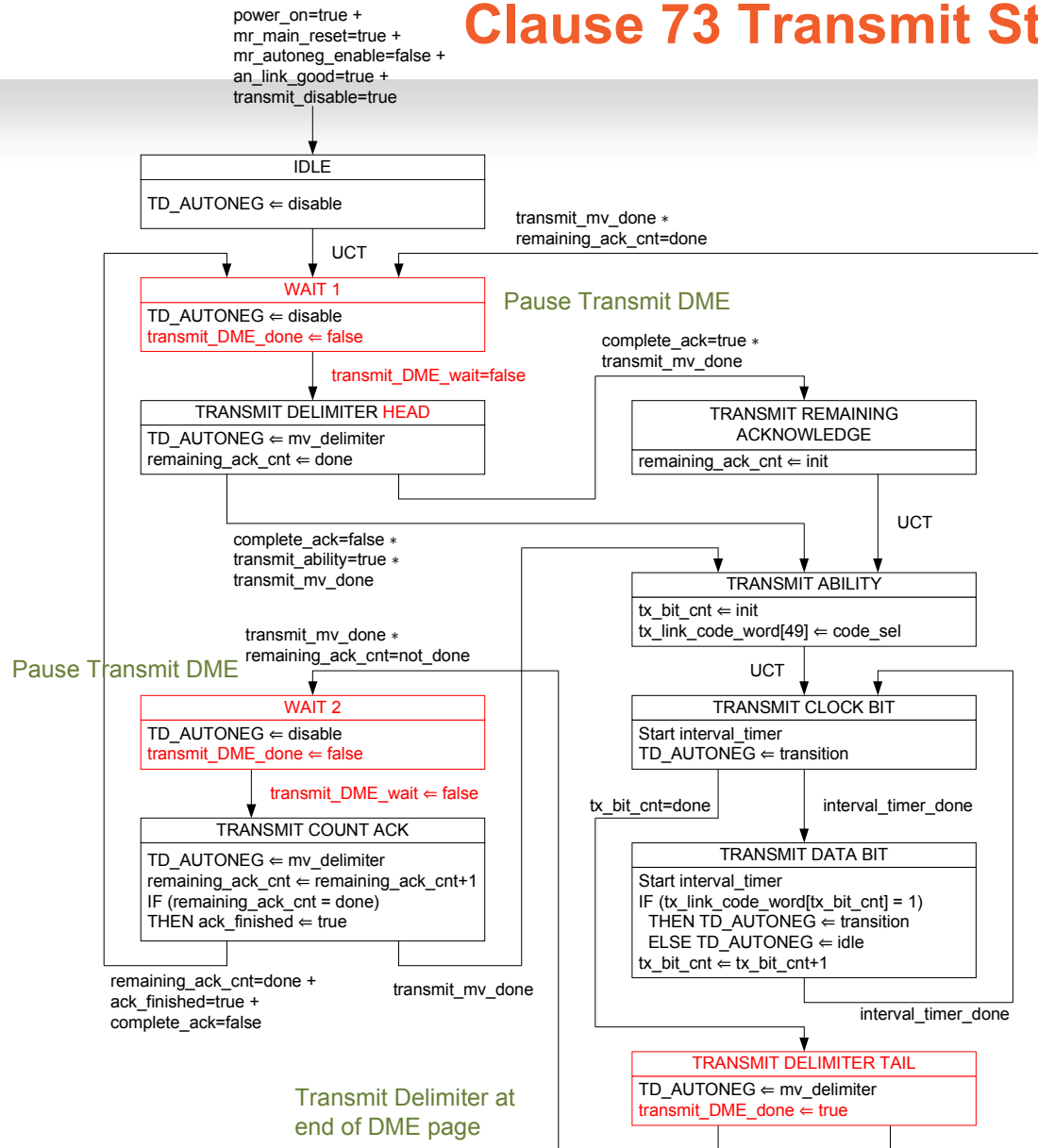
ABILITY DETECT
transmit_ability ⇐ true
mr_lp_autoneg_able ⇐ false
link_control_[PD] ⇐ SCAN_FOR_CARRIER
toggle_tx ⇐ mr_adv_ability[12]
ability_match ⇐ false
acknowledge_match ⇐ false
tx_link_code_word[48:1] ⇐ mr_adv_ability[48:1]
mr_page_rx ⇐ false
base_page ⇐ true
ack_finished ⇐ false
consistency_match ⇐false

break_link_timer_done

ability_match=true *
nonce_match=true

TRANSMIT DISABLE
Start break_link_timer
link_control_[all] ⇐ DISABLE
transmit_disable ⇐ true
mr_page_rx ⇐ false
mr_autoneg_complete ⇐ false
mr_next_page_loaded ⇐ false

(acknowledge_match=true *
(consistency_match=false +
(ack_nonce_match=false *
base_page=true))) +
an_receive_idle=true

ability_match=true * nonce_match=false

ACKNOWLEDGE DETECT
IF(base_page=true) THEN
tx_link_code_word[10:6] ⇐ rx_nonce[4:0]
transmit_ability ⇐ true
transmit_ack ⇐ true
mr_lp_autoneg_able ⇐ true
link_control_[all] ⇐ DISABLE

acknowledge_match=true *
(ack_nonce_match=true +
base_page=false) *
consistency_match=true

COMPLETE ACKNOWLEDGE
complete_ack ⇐ true
transmit_ability ⇐ true
transmit_ack ⇐ true
toggle_rx ⇐ rx_link_code_word[12]
toggle_tx ⇐ !toggle_tx
mr_page_rx ⇐ true
np_rx ⇐ rx_link_code_word[NP]
mr_lp_adv_abiliy ⇐ rx_link_code_word

power_on=true +
mr_main_reset=true +
mr_restart_negotiation=true +
mr_autoneg_enable=false

Auto-Negotiation ENABLE
mr_page_rx ⇐ false
mr_autoneg_complete ⇐ false
mr_parallel_detection_fault ⇐ false

mr_autoneg_enable=true

ability_match=true *
((toggle_rx ^ ability_match_word[12])=1)

ack_finished=true *
tx_link_code_word[NP]=0 *
np_rx=0

ack_finished=true *
mr_next_page_loaded=true *
((tx_link_code_word[NP]=1) + (np_rx=1))

AN GOOD
an_link_good ⇐ true
mr_autoneg_complete ⇐ true

link_status_[HCD]=OK

AN GOOD CHECK
link_control_[notHCD] ⇐ DISABLE
link_control_[HCD] ⇐ ENABLE
an_link_good ⇐true
start link_fail_inhibit_timer

link_status_[HCD]=FAIL

((link_status_[HCD]=FAIL +
link_status_[HCD]=OK) *
link_fail_inh bit_timer_done) +
incompat ble_link = true

NEXT PAGE WAIT
transmit_ability ⇐ true
mr_page_rx ⇐ false
base_page ⇐ false
tx_link_code_word[48:13] ⇐ mr_np_tx[48:13]
tx_link_code_word[12] ⇐ toggle_tx
tx_link_code_word[11:1] ⇐ mr_np_tx[11:1]
ack_finished ⇐ false
mr_next_page_loaded ⇐ false

an_receive_idle=true

Clause 73 Transmit State Machine Changes

```
power_on=true +
mr_main_reset=true +
mr_autoneg_enable=false +
an_link_good=true +
transmit_disable=true
```

IDLE
TD_AUTONEG ⇐ disable

UCT

transmit_mv_done *
remaining_ack_cnt=done

WAIT 1
TD_AUTONEG ⇐ disable
transmit_DME_done ⇐ false

Pause Transmit DME

transmit_DME_wait=false

TRANSMIT DELIMITER HEAD
TD_AUTONEG ⇐ mv_delimiter
remaining_ack_cnt ⇐ done

complete_ack=true *
transmit_mv_done

TRANSMIT REMAINING ACKNOWLEDGE
remaining_ack_cnt ⇐ init

complete_ack=false *
transmit_ability=true *
transmit_mv_done

UCT

transmit_mv_done *
remaining_ack_cnt=not_done

TRANSMIT ABILITY
tx_bit_cnt ⇐ init
tx_link_code_word[49] ⇐ code_sel

Pause Transmit DME

UCT

WAIT 2
TD_AUTONEG ⇐ disable
transmit_DME_done ⇐ false

transmit_DME_wait ⇐ false

TRANSMIT CLOCK BIT
Start interval_timer
TD_AUTONEG ⇐ transition

tx_bit_cnt=done

interval_timer_done

TRANSMIT COUNT ACK
TD_AUTONEG ⇐ mv_delimiter
remaining_ack_cnt ⇐ remaining_ack_cnt+1
IF (remaining_ack_cnt = done)
THEN ack_finished ⇐ true

TRANSMIT DATA BIT
Start interval_timer
IF (tx_link_code_word[tx_bit_cnt] = 1)
  THEN TD_AUTONEG ⇐ transition
  ELSE TD_AUTONEG ⇐ idle
tx_bit_cnt ⇐ tx_bit_cnt+1

remaining_ack_cnt=done +
ack_finished=true +
complete_ack=false

transmit_mv_done

interval_timer_done

Transmit Delimiter at
end of DME page

TRANSMIT DELIMITER TAIL
TD_AUTONEG ⇐ mv_delimiter
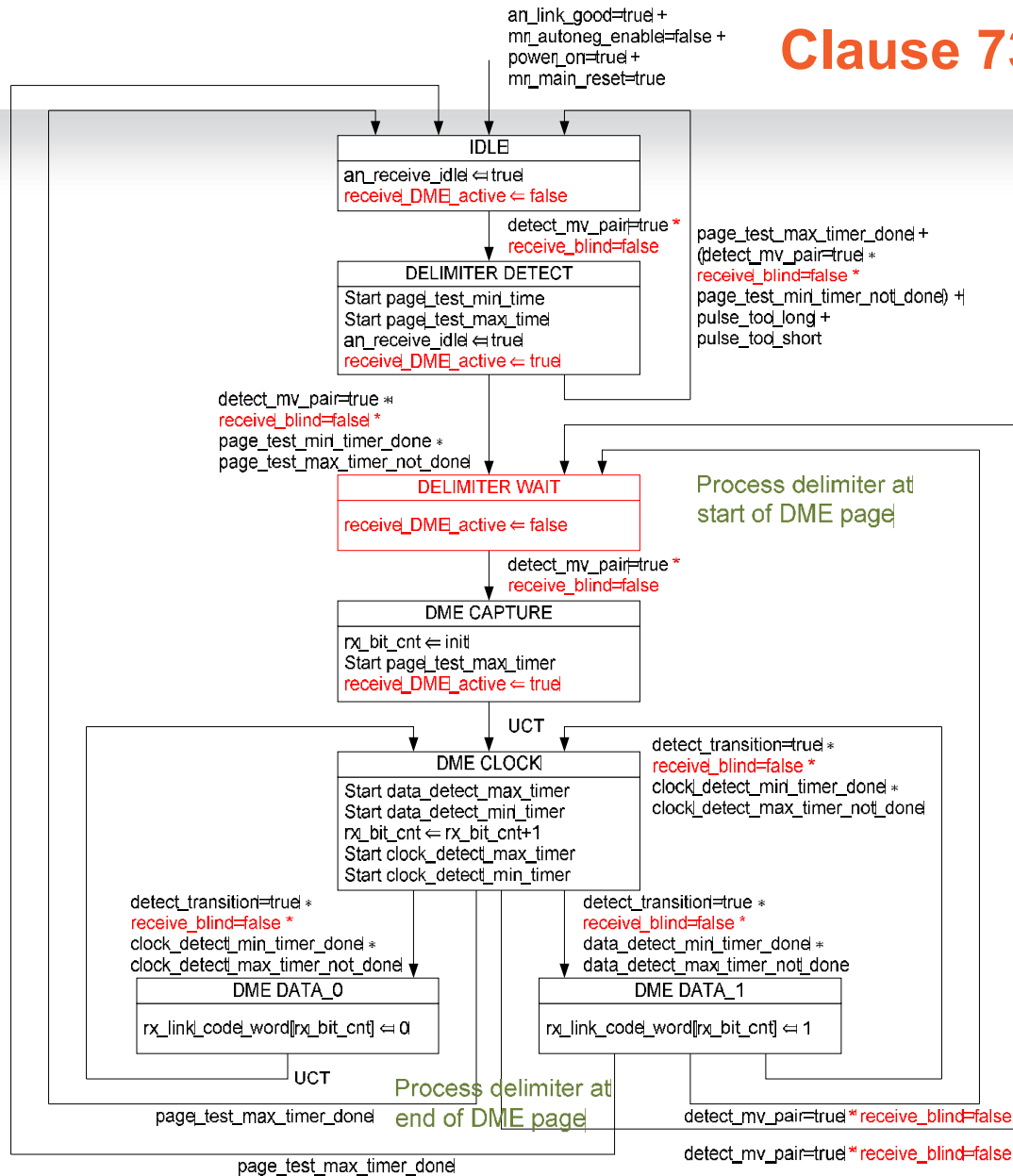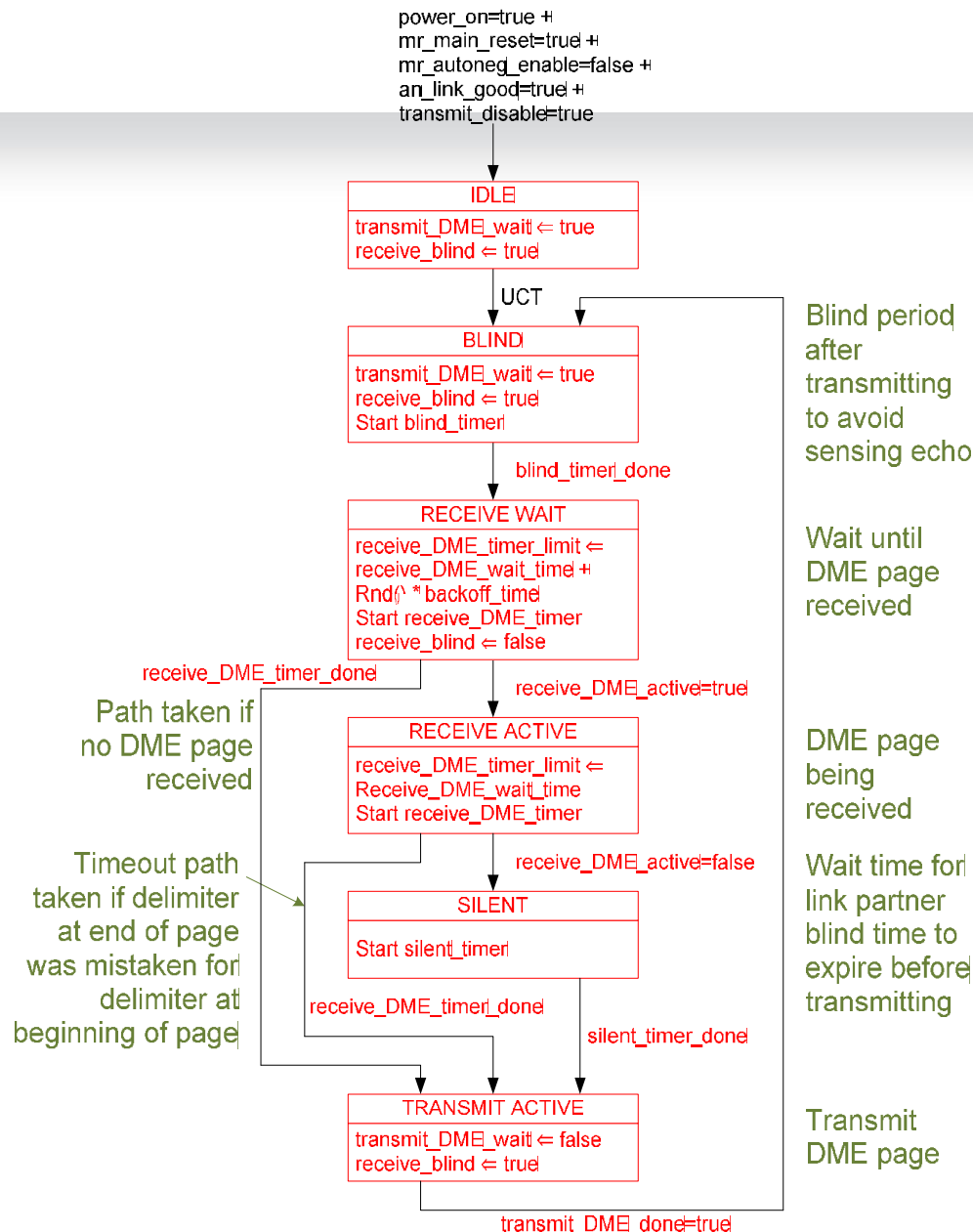transmit_DME_done ⇐ true

- ▸ **WAIT1 and WAIT2 states pauses the next DME page transmission until Half-Duplex state machine gives go ahead to continue via transmit_DME_wait**

- ▸ **transmit_DME_done lets the Half-Duplex state machine know if a DME page is being transmitted or not**

- ▸ **TRANSMIT DELIMITER TAIL state adds the end delimiter to the DME page**

Clause 73 Receive State Machine Changes

State machine diagram:

an_link_good=true +
mr_autoneg_enable=false +
power_on=true +
mr_main_reset=true

IDLE
an_receive_idle ⇐ true
receive_DME_active ⇐ false

detect_mv_pair=true *
receive_blind=false

DELIMITER DETECT
Start page_test_min_time
Start page_test_max_time
an_receive_idle ⇐ true
receive_DME_active ⇐ true

page_test_max_timer_done +
(detect_mv_pair=true *
receive_blind=false *
page_test_min_timer_not_done) +
pulse_too_long +
pulse_too_short

detect_mv_pair=true *
receive_blind=false *
page_test_min_timer_done *
page_test_max_timer_not_done

DELIMITER WAIT
receive_DME_active ⇐ false

Process delimiter at start of DME page

detect_mv_pair=true *
receive_blind=false

DME CAPTURE
rx_bit_cnt ⇐ init
Start page_test_max_timer
receive_DME_active ⇐ true

UCT

DME CLOCK
Start data_detect_max_timer
Start data_detect_min_timer
rx_bit_cnt ⇐ rx_bit_cnt+1
Start clock_detect_max_timer
Start clock_detect_min_timer

detect_transition=true *
receive_blind=false *
clock_detect_min_timer_done *
clock_detect_max_timer_not_done

detect_transition=true *
receive_blind=false *
clock_detect_min_timer_done *
clock_detect_max_timer_not_done

detect_transition=true *
receive_blind=false *
data_detect_min_timer_done *
data_detect_max_timer_not_done

DME DATA_0
rx_link_code_word[rx_bit_cnt] ⇐ 0

DME DATA_1
rx_link_code_word[rx_bit_cnt] ⇐ 1

UCT

Process delimiter at end of DME page

page_test_max_timer_done

detect_mv_pair=true * receive_blind=false

detect_mv_pair=true * receive_blind=false

page_test_max_timer_done

▸ **DELIMITER WAIT state processes delimiter at the end of the DME page**

▸ **receive_DME_active lets the Half-Duplex state machine know if a DME page is being received or not**

▸ **The Half-Duplex state machine blinds the Receive State Machine with receive_blind**

Half-Duplex State Machine

power_on=true +
mr_main_reset=true +
mr_autoneg_enable=false +
an_link_good=true +
transmit_disable=true

IDLE
transmit_DME_wait ⇐ true
receive_blind ⇐ true

UCT

BLIND
transmit_DME_wait ⇐ true
receive_blind ⇐ true
Start blind_timer

blind_timer_done

RECEIVE WAIT
receive_DME_timer_limit ⇐
receive_DME_wait_time +
Rnd() * backoff_time
Start receive_DME_timer
receive_blind ⇐ false

receive_DME_timer_done

Path taken if
no DME page
received

receive_DME_active=true

RECEIVE ACTIVE
receive_DME_timer_limit ⇐
Receive_DME_wait_time
Start receive_DME_timer

receive_DME_active=false

Timeout path
taken if delimiter
at end of page
was mistaken for
delimiter at
beginning of page

SILENT
Start silent_timer

receive_DME_timer_done

silent_timer_done

TRANSMIT ACTIVE
transmit_DME_wait ⇐ false
receive_blind ⇐ true

transmit_DME_done=true

Blind period
after
transmitting
to avoid
sensing echo

Wait until
DME page
received

DME page
being
received

Wait time for
link partner
blind time to
expire before
transmitting

Transmit
DME page

▶ **State Machine cycles through Blind, Receive, Silent (Wait), and Transmit**

▶ **Interacts with Transmit and Receive SM via 4 variables**
- **receive_blind**
- **receive_DME_active**
- **transmit_DME_wait**
- **transmit_DME_done**

▶ **RECEIVE WAIT to TRANSMIT ACTIVE has random backoff component if no DME page received**

▶ **RECEIVE ACTIVE to TRANSMIT ACTIVE to recover if delimiter at end of page is mistaken for delimiter at start of page**

IEEE 802.3bp RTPGE – March 2014 Plenary Meeting

12

MARVELL®

Timers – Proposed Values

Parameter	Min	Typ	Max	Units
autoneg_wait_timer		deleted		
break_link_timer		TBD		
clock_detect_min_timer	68		76	ns
clock_detect_max_timer	84		92	ns
data_detect_min_timer	28		36	ns
data_detect_max_timer	44		52	ns
interval_timer	39	40	41	ns
link_fail_inhibit_timer		TBD		
page_test_min_timer	4366		4406	ns
page_test_max_timer	4714		4754	ns
blind_timer	2000		2040	ns
silent_timer	2120		2200	ns
receive_DME_wait_time	6794		6874	ns
backoff_time	2120		2200	ns
receive_DME_timer	receive_DME_wait_time + random(0 to 15) x backoff_time			

▶ **New timers in red**

Timer Notes

▶ **autoneg_wait_timer**
- **Deleted since no parallel detect**

▶ **break_link_timer – Recommend 100us**
- **Need further study if PCS, PMA can detect link drop this quickly**

▶ **linkfail_inhibit_timer**
- **Need further study to determine how fast PCS, PMA can link up**
- **Provision should be made to add more margin for future PHYs**

▶ **blind_timer**
- **Amount of time needed for worst case round trip echo propagation**
- **2000 ns is good for 200 meter – more than sufficient for future proofing**
- **Note that echos may propagate for more than 1 round trip**

▶ **silent_timer**
- **Wait time should be slightly longer than blind_timer to guarantee that link partner is not blinded when DUT starts transmitting**

Timer Notes

▶ **receive_DME_wait_time**

- **Longer than sum of DME_page_width + silent_timer**
- **Allows sufficient time to see link partner's DME page**

▶ **backoff_time**

- **Same as silent_time**

▶ **interval_timer**

- **N x 25MHz friendly**

New Variables

Variables	
transmit_DME_wait	true = pause DME page transmission false = continue DME page transmission
receive_blind	true = ignore received DME transitions false = accept received DME transitions
receive_DME_timer_limit	Timer duration for receive_DME_timer
receive_DME_active	true = DME page reception in progress false = DME page reception completed
transmit_DME_done	true = DME page transmission completed false = DME page transmission in progress

▶ **Rnd()**

▪ **Random number generator outputting integer from 0 to 15 inclusive**

MARVELL®

DME Page Formatting

▶ **S[4:0] and C[2:0] – Same as 802.3ap**

- **S[4:0] set to 00001**
- **C[2:0] used for pause**

▶ **E[4:0] and T[4:0] randomized – Same as 802.3ap**

- **Also use this for Master/Slave Random Seed**

▶ **A[24:0] – Technology Ability Field**

- **This task force and future task force to define**

▶ **F[1:0] – Propose change from FEC to Master/Slave**

- **F[0] – Master Slave Config**
- **F[1] – Master Slave Manual Config Enable**

D 0	D 1	D 2	D 3	D 4	D 5	D 6	D 7	D 8	D 9	D 10	D 11	D 12	D 13	D 14	D 15
S 0	S 1	S 2	S 3	S 4	E 0	E 1	E 2	E 3	E 4	C 0	C 1	C 2	RF	Ack	NP

F

D 16	D 17	D 18	D 19	D 20	D 21	D 22	D 23	D 24	D 25	D 26	D 27	D 28	D 45	D 46	D 47
T 0	T 1	T 2	T 3	T 4	A 0	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 24	F 0	F 1

Master Slave Random Seed

▶ **If both PHYs prefer to be master or both prefer to be slave the master / slave random seed kicks in**

▶ **If the randomly generated T[4:0] field of both PHYs the same the arbitration state machine forces restart (nonce_match)**

  ▪ **Enhancement possible to guarantee nonce_match is always false**

▶ **Hence T[4:0] of both PHYs will never be the same upon successful exchange**

  ▪ **PHY with higher T[4:0] value will be the master**

```
ABILITY DETECT                                    TRANSMIT DISABLE

                        break_link_timer_done
                   ◄────────────────────────

                   ────────────────────────►
                        ability_match=true *
                        nonce_match=true
```
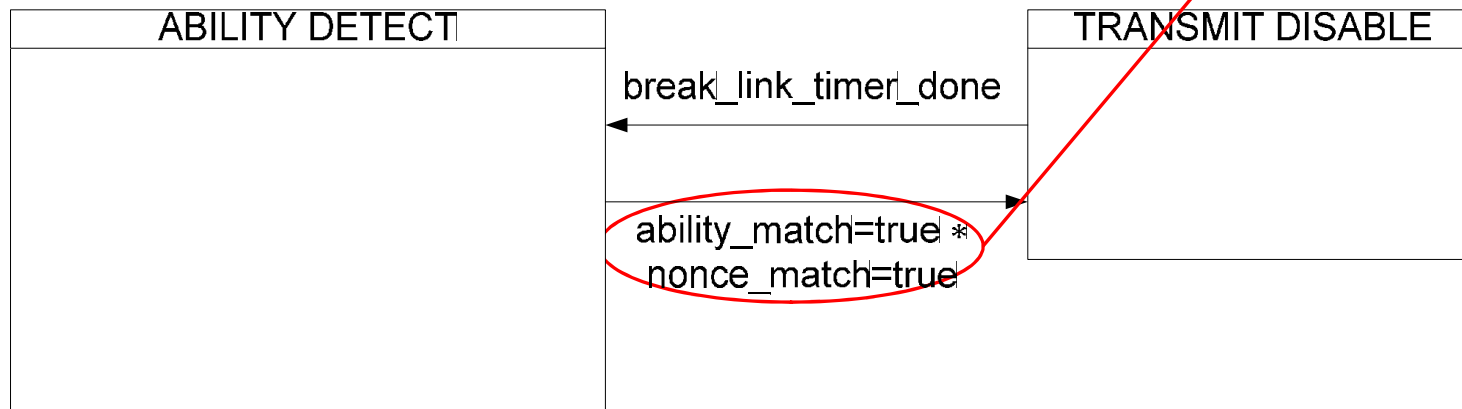
Optional Enhancement for State Machine Optimization

▶ **remaining_ack_cnt**

- **Currently need to send 6 to 8 pages to complete exchange**
- **Since half duplex operation forces both PHYs to be in sync we can set this to a lower number (recommend 3).**

▶ **nonce_match**

- **Currently forces the state machine to recycle to TRANSMIT DISABLE state and wait another break_link_timer**
- **Instead of waiting to see 3 link partner pages, change the transmitted T[4:0] on the fly while in the ABILITY MATCH state if the T[4:0] of the first page received matches.**
- **Since the system is half-duplex, the link partner probably never received the DUT page and does not know the DUT's T[4:0] value so changing on the fly does not hurt.**
- **To prevent the new T[4:0] from being the same as the previous one, the proposal is to invert the previous T[0] and randomly generate a new T[4:1].**

Electrical Interface

▶ **Use PAM3 Transmitter as is**

- **Electrical levels – TBD by task force**
- **Rise and fall times – TBD by task force**
- **DME – use PAM3 +1, -1 levels.**
- **Quiet – either PAM3 0 level, or turn off**

▶ **Two comparators needed for receiver**

- **One for PAM3 +1 detection, other for PAM3 -1 detection**
- **Propose +50mV and -50mV thresholds**
- **Need to bypass any on chip receive filtering**
- **Comparators oversampled to reconstruct DME pulses**

MARVELL®

THANK YOU

BACKUP

Startup Times With Proposed Values

▶ **break_link_timer – proposed 100 us**

▶ **Startup retry resolution per attempt - worst case wait**
  - **DME page + Blind period + receive time out + Wait period = 4.56 us + 2.0 us + 6.8 us + 15 * 2.16 us = 45.76 us**

▶ **Once in sync each pair of pages takes**
  - **2 x (Wait period + DME page + propagation) = 2 x (2.16 us + 4.56 us + 1.0 us) = 15.44 us**

▶ **Base page exchange**
  - **3 in Ability Detect, 3 in Acknowledge Detect, 6 in Complete Acknowledge = (3 + 3 + 6) * 15.44us = 185.28 us**

▶ **DSP linkup time**
  - **RTPGE Group to define based on modulation scheme and PCS**

▶ **Total time not counting DSP link up time**
  - **No retry – 100 us + 185.28 us < 300 us**
  - **10 retry – 100 us + 10 x 45.76 us + 185.28 us < 750 us**

MARVELL®