802.3ap Auto-Negotiation Proposal
with Clause 28 State Machines

Presentation to

IEEE 802.3ap Task Force

Sep 2004 Interim Meeting

Contributors and Supporters

Contributors/Supporters

Ilango Ganga, Luke Chang, Brad Booth	Intel
Andre Szczepanek	Texas Instruments
Thomas Kirkegaard Jørgensen	Vitesse
Pat Thaler	Agilent

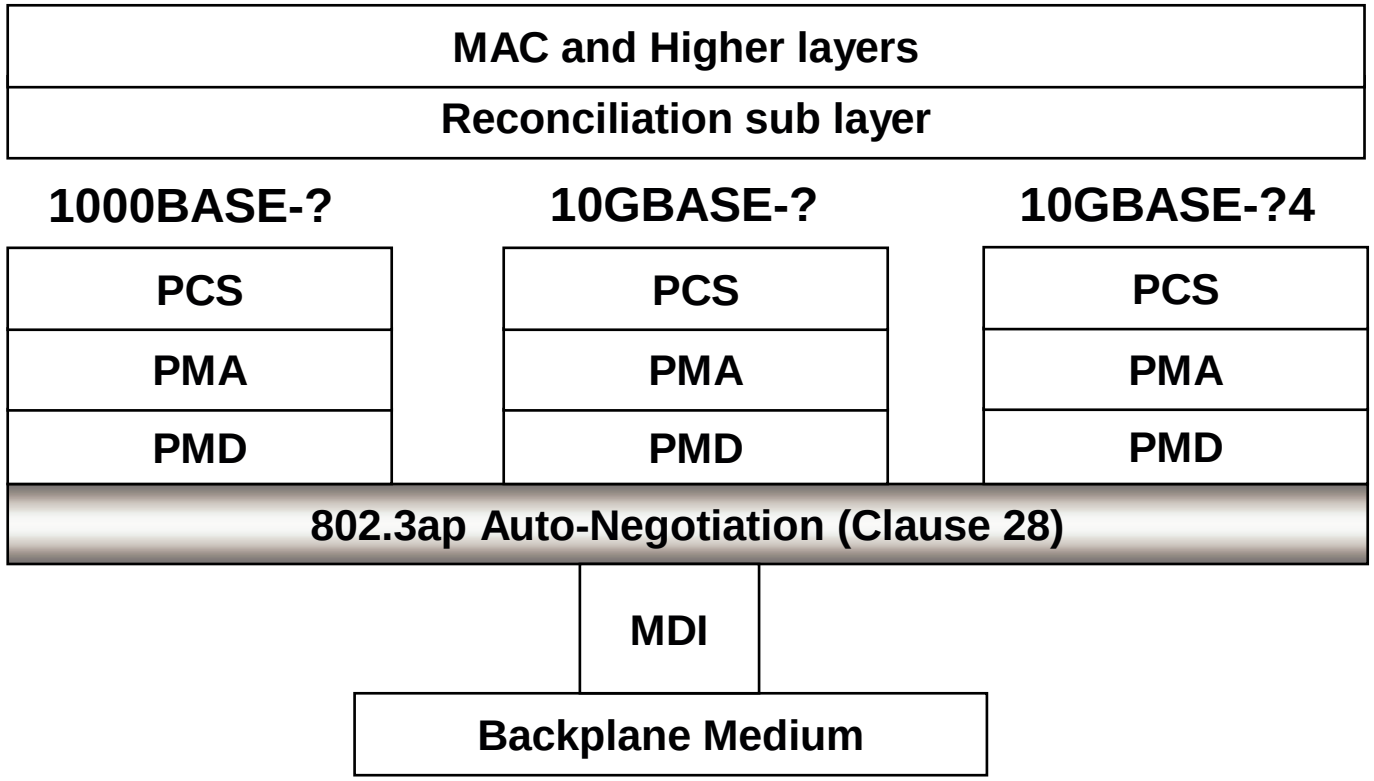
Supporters

Nitish Amin, Glen Koziuk	Vitesse
Rob Brink, Max Olsen	Agere
Rick Rabinovich	Spirent Comm
Petre Popescu	Quake Tech
Joel Goergen	Force10
Harmeet Bhugra	IDT
Dimitry Taich , Amir Bar-Niv,	Mysticom

Objective

- Auto-Neg to exchange information between peer devices to automatically configure to take maximum advantage of their abilities
- New base page and selector fields for 802.3 Backplane Ethernet
  - Higher efficiency for exchanging pages
- Use Proposed SERDES compatible FLP
- Use Clause 28 state machines for exchanging pages and Parallel Detection
- Define parallel detection mechanism to support backplane links that do not support 802.3ap Auto-Negotiation and be backward compatible with legacy 1000BASE-X and XAUI backplane links

High Level Layer Model



Layer Model

Clause 28 background

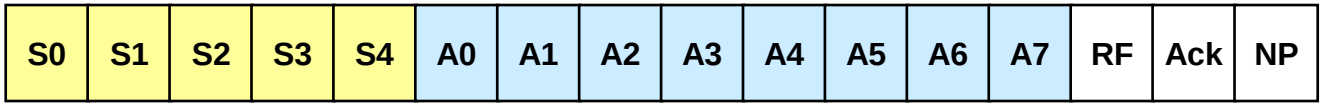
- Established for copper technologies and can Negotiate Multiple speeds
- Clause 28 state machines verified over years and have been made robust
- 7 years of AN Interoperability testing shows that most issues are associated with Software/Drivers that control the AN process (Device Management)
  - The protocol and state machine itself is robust
  - Most problems show up when system does not properly use the information provided by the AN protocol[2]
  - Initial interoperability tests focused on accuracy of AN State Machines but later new tests were added to focus on Device Management to address system level issues (Reference: [2] UNH- IOL Clause 28 AN management system test suite

Advantages of using Clause 28

- Auto-Negotiation function located below the PMD layer (Fig 28-2) [4]
  - Ideal position of sub layer within PHY
  - Ability to negotiate multiple speeds
- Ability to work with different signaling
  - NRZ and non NRZ signaling (MLS)
- Does not depend on PCS encoding
- Self Clocking
- The state machine has been field proven over multiple PHY generations

Clause 28 Base Page

Clause 28 Base Page definition



Next Page

Acknowledge

Remote Fault

Table 28A–1—Selector Field value mappings

S4	S3	S2	S1	S0	Selector description
0	0	0	0	0	Reserved for future Auto-Negotiation development
0	0	0	0	1	IEEE Std 802.3®
0	0	0	1	0	IEEE Std 802.9® ISLAN-16T
0	0	0	1	1	IEEE Std 802.5®
1	1	1	1	1	Reserved for future Auto-Negotiation development[a]

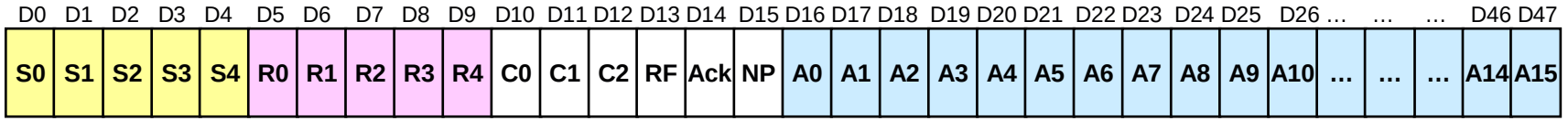
[a]For up-to-date information on the allocation of Auto-Negotiation Selector fi
http://www.ieee802.org/3/selectors/selectors.html

Table 28B–1—Technology Ability Field bit assignments

Bit	Technology	Minimum cabling requirement
A0	10BASE-T	Two-pair category 3
A1	10BASE-T full duplex	Two-pair category 3
A2	100BASE-TX	Two-pair category 5
A3	100BASE-TX full duplex	Two-pair category 5
A4	100BASE-T4	Four-pair category 3
A5	PAUSE operation for full duplex links	Not applicable
A6	Asymmetric PAUSE operation for full duplex Links	Not applicable
A7	Reserved for future technology	

Proposed 802.3ap Base Page

Proposed 48-bit Base Page definition (Base Link Code Word)



Proposed selector field for 802.3 Backplane Ethernet

S4	S3	S2	S1	S0	Selector description
0	0	0	0	0	Reserved for future Auto-Negotiation development
0	0	0	0	1	IEEE Std 802.3
0	0	0	1	0	IEEE Std 802.9 ISLAN-16T
0	0	0	1	1	IEEE Std 802.5
0	0	1	0	0	IEEE Std 1394
0	0	1	0	1	IEEE 802.3 Backplane Ethernet
1	x	x	x	x	Reserved for future Auto-Negotiation development*

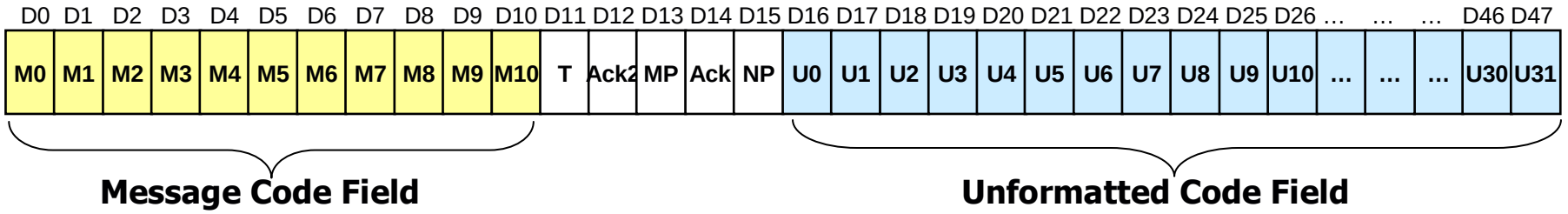
* Annex 28A specifies this as 11111. Correction to Base spec has been submitted to 802.3

Proposed 802.3ap Base page bit assignments

Bit	Base Page bit assignments
A0	802.3ap 1Gb/s 1-Lane (1000BASE-?)
A1	802.3ap 10Gb/s 4-Lane (10GBASE-?4)
A2	802.3ap 10Gb/s 1-Lane (10GBASE-?)
A3:9	Reserved for Future Technology
A10:22	Reserved / TBD (other 802.3ap parameters)
R0:R4	Reserved, (Could be sub selector in future)
C0	Pause
C1:2	Reserved
RF	Remote Fault
Ack	Acknowledge
NP	Next Page

Next Page Assignment

- Continue to have Next Pages for future proofing and to negotiate additional parameters if any
  - Goal is to complete most negotiations in 1 next page
  - Align with Page Extension proposal in 802.3an [6]
- Next Page is also 48-bit size (NP=0 indicates last page)
- The message page and Unformatted pages are merged into a single 48-bit next page for higher efficiency
- Next Page encodings are shown below
  - 11 bit Message code field and 32 bit Unformatted code field



Functional reference diagram

- Use Clause 28 functional reference diagram

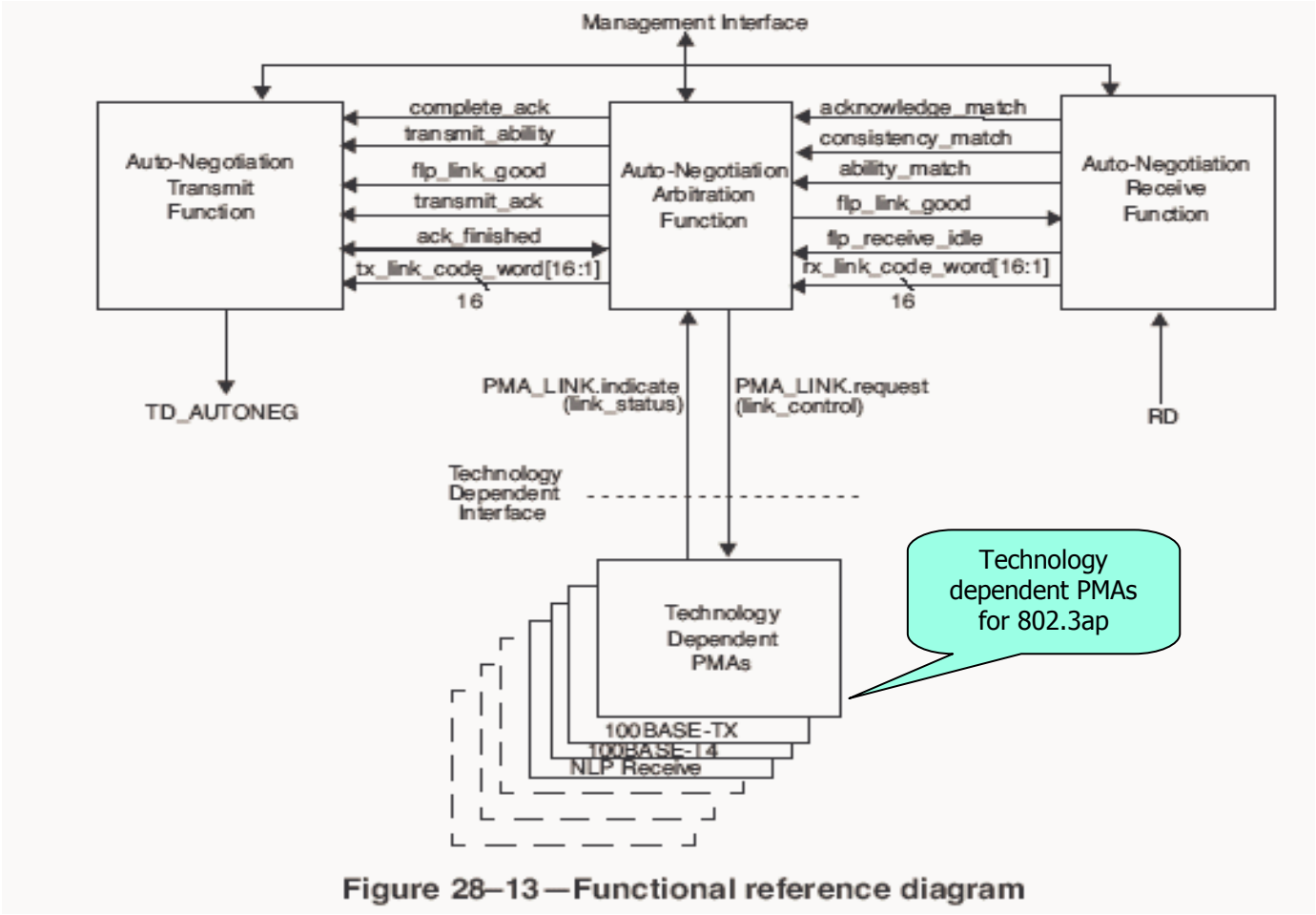


Figure 28–13—Functional reference diagram

Transmit & Receive state machines

- Use Clause 28.3 State diagrams and transmit/receive/arbitration state machine definitions for exchanging base and next pages [4]
- The State Diagram Counters shall be increased to accommodate 48 bit Link Code Words
  - Example: increase tx_bit_cnt from integer (1 – 17) to (1 – 49)
  - Align with 802.3an proposal on increasing page size [6]
- The state diagram timers in clause 28.3.2 re-defined with the goal of reducing page transfer time to below 1ms
  - Timer variables are lowered for higher efficiency and faster link convergence for Backplane Ethernet
  - Goal: maximum 1 sec (or less) for link convergence

Timing

- Lower Timer Values for high efficiency and fast link convergence
- Significantly faster page transfer time
  - 48-bit SSP Burst less than 40µs instead of 6ms
  - 80µs SSP Burst to Burst instead of 16ms

Table 28-1 – FLP Burst timing summary (Redefined)

#	Technology	Typical	Units
T1	Clk/Data Pulse width	104	ns
T2	Clock Pulse to Clock Pulse	800	ns
T3	Clock Pulse to Data Pulse	400	ns
T4	Pulses in a Burst	49 to 97	#
T5	Burst Width	38.4	µs
T6	FLP Burst to FLP Burst	80	µs

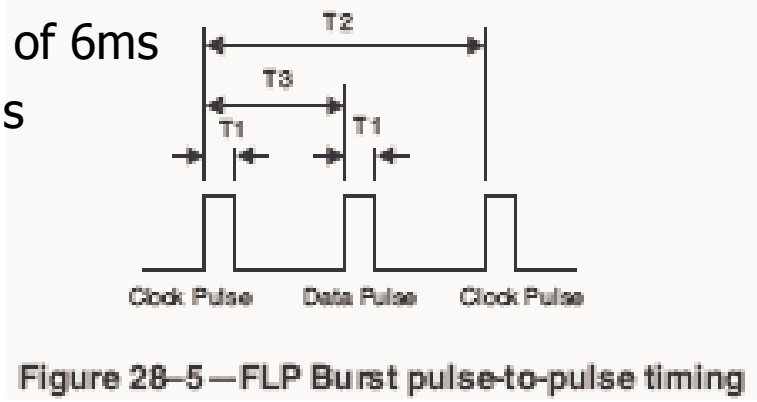


Figure 28–5—FLP Burst pulse-to-pulse timing

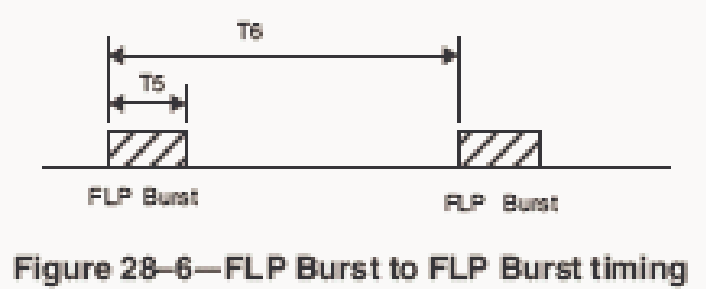


Figure 28–6—FLP Burst to FLP Burst timing

Priority Resolution

- Local and Link partner devices may have multiple abilities in common
- Hence Priority resolution function to predefine hierarchy of supported technologies
- Local Device and Link Partner negotiates to the Highest common denominator (HCD)
- Relative priorities defined below (Listed from highest to lowest)

Priority	Technology	Capability
1	10GBASE-?	802.3ap 10Gb/s 1-Lane
2	10GBASE-?4	802.3ap 10Gb/s 4-Lane
3	1000BASE-?	802.3ap 1Gb/s 1-Lane

Management Data Interface

- Use Clause 45 MDIO interface and Register space [4]
  - Clause 22 Management frame format
  - Clause 45.4 Electrical interface definition
  - Clause 45.2.1 PMA/PMD register set definition
- Use Clause 45 interface (ST=00) to access Clause 45 registers (Recommended)
  - Access Clause 22 registers if present in a legacy device through Clause 45 management interface
  - Since there is no need to access control bits relevant to 10/100/1000 twisted pair operation it is proposed to use clause 45 interface to access clause 45 register space
    - Other options considered: Using Clause 22 mechanism to access Clause 45 register set, similar to 802.3ah EFM [5]

Management Registers

- Use Clause 45.2.1 PMA/PMD register set [4][5]
  - Define additional bits and registers for 802.3ap
  - Lot of space available for future proofing
- Define 1.0.5:2 speed selector bits for 1G speeds: 0x00=10Gb/s; 0x01 taken by EFM; 0x02=1Gb/s
  - Add 1Gb/s speed ability bit to Register 1.4.3 to indicate 1G speed (Table 45-5)
- Bit 1.5.0 to indicate clause 22 registers implemented and accessed through clause 45 interface
- Bit 1.5.1 indicates PMA/PMD device present in package
  - Define 1.7.4:3 (two bits) to indicate different 802.3ap PMA/PMD types
  - Add additional bits to the 10G PMA/PMD extended abilities register (Register 1.11) to indicate 802.3ap PMD types (Bits:1.11:3:2); Bit 1.11.0 taken by 10GBASE-CX4 and 1.11.1 possibly by 10GBASE-T
- Continue to use Registers 1.9 and 1.10 for Transmit/Receive functions
- Define 1.110 to 1.1xx new registers for 802.3ap specific functions

SSP = Symbol Sequence Pulse

- Clause 14 Link pulses are Baseband 100ns pulses
- SSPs are ~100ns Amplitude Modulated carrier pulses
  - The carrier consists of a sequence of symbols at one of the 802.3ap data rates
- SSP envelopes are detected using a differential signal detector
- The modulation and detection of SSPs can be asynchronous to the carrier symbol rate
  - Asynchronously modulate the output of the existing PCS/SERDES datapath
    - Benign symbol sequences chosen for legacy compatability
  - AN can communicate even if link partners share no common data-rate
- This is a an Out-Of-Band signalling technique already used by SATA & PCI-Express
  - Practicality of 100ns bursts already demonstrated by SATA
  - Already being implemented in SERDES
- Transmitter shall transmit SSPs at its lowest capable rate
- Replace clause 14 Link pulses with SSPs in clause 28 protocol

1.25G Signaling

- To transmit an SSP, the transmitter shall transmit between 96-112ns of D21.5 symbols (1010101010).
  - D21.5 is chosen because it's DC balanced and disparity neutral.
  - At this symbol rate, the time to transmit 13 8b10b symbols is 104ns.
- An SSP can start and end Asynchronously to the symbol boundaries.
  - Allows asynchronous modulation of the analog Transmitter (as in SATA)
- Within an SSP all normal 1.25G data transmit electrical specifications hold
  - Allows existing datapath to be used as source of SSP
  - Ensures a non SSP-capable 1.25G Rx'er will see a benign symbol stream

3.125G Signaling

- SSPs are sent on Lane 0 only
- To transmit an SSP, the transmitter shall transmit between 96-112ns of D21.5 symbols (1010101010).
  - D21.5 is chosen because it's DC balanced and disparity neutral.
  - At this symbol rate, the time to transmit 32 8b10b symbols is 104ns.
- An SSP can start and end Asynchronously to the symbol boundaries.
  - Allows asynchronous modulation of the analog Transmitter (as in SATA)
- Within an SSP all normal 3.125G data transmit electrical specifications hold
  - Allows existing datapath to be used as source of SSP
  - Ensures a non SSP-capable 3.125G Rx'er will see a benign symbol stream

10G Signaling

- To transmit an SSP, the transmitter shall transmit between 96-112ns of appropriate symbols.
  - The Symbol sequence will depend on the Signaling and coding schemes chosen for 10G operation.
    - It will be DC balanced and of sufficient energy to meet the detection threshold levels defined for 1 & 2.5Gbps operation.
    - At this symbol rate, the time to transmit 16 64b66b symbols is 104ns.
- An SSP can start and end Asynchronously to the symbol boundaries.
  - Allows asynchronous modulation of the analog Transmitter (as in SATA)
- Within an SSP all normal 10G data transmit electrical specifications hold
  - Allows existing datapath to be used as source of SSP

SSP detection and validation

- Use a differential signal detector to detect these SSPs irrespective of Tx symbol rate
  - Define threshold range : 65-175mV
    - All signals detected above this threshold are potential SSPs
    - All signals below this threshold are regarded as "IDLE"
- Validate SSP width
  - Detector shall reject all SSPs outside of range 48-176ns
    - 6-22 symbols @ 1.25Gbd
    - 15-55 symbols @ 3.125Gbd
  - Detector shall accept all SSPs inside of range 96-112ns
    - 12-14 symbols @ 1.25Gbd
    - 30-35 symbols @ 3.125Gbd

SERDES requirements

- SERDES Transmitter must support a driven electrical idle state
  - In Electrical Idle Vdiff = 0, with normal Vcm maintained
- Receiver must implement a true signal detect
  - A signal detection threshold in the range of 65-175mV would be sufficient and compatible with PCI-Express & SATA
  - Detection latency should be no more than 50ns to meet detection thresholds
- Multi-purpose SERDES with support for PCI-Express or SATA will already have these features

Parallel Detect

- Re-use Clause 28 Parallel Detect function
- Ensures that a valid link is always brought up - also when the link partner does not support Auto-Negotiation
- Ensures backward compatibility with legacy 1000BASE-X and XAUI ports

Functional Description …1

- The Local Device detects a Link Partner that supports Auto-Negotiation by SSP Burst detection. The Parallel Detect function allows detection of Link Partners that support 1000BASE-X, 10GBASE-X4 (XAUI) and 10GBASE-Serial, but do not support Auto-Negotiation.

- Prior to detection of SSP Bursts, the Receive Switch shall direct MDI receive activity to the 1000BASE-X, 10GBASE-X4 and 10GBASE-Serial PMAs, if present. If at least one of the 1000BASE-X, 10GBASE-X4 or 10GBASE-Serial PMA establishes link_status=READY, the LINK STATUS CHECK state is entered and the autoneg_wait_timer is started.

Functional Description …2

- If exactly one link_status=READY indication is present when the autoneg_wait_timer expires, then Auto-Negotiation shall set link_control=ENABLE for the PMA indicating link_status=READY. If a PMA is enabled, the Arbitration function shall set link_control=DISABLE to all other PMAs and indicate that Auto-Negotiation has completed.

- On transition to the SSP LINK GOOD CHECK state from the LINK STATUS CHECK state the Parallel Detection function shall set the bit in the link partner ability register (Register 5) corresponding to the technology detected by the Parallel Detection function.

Updates to state machine – 28-16

- No new states needed
- Only update to variables

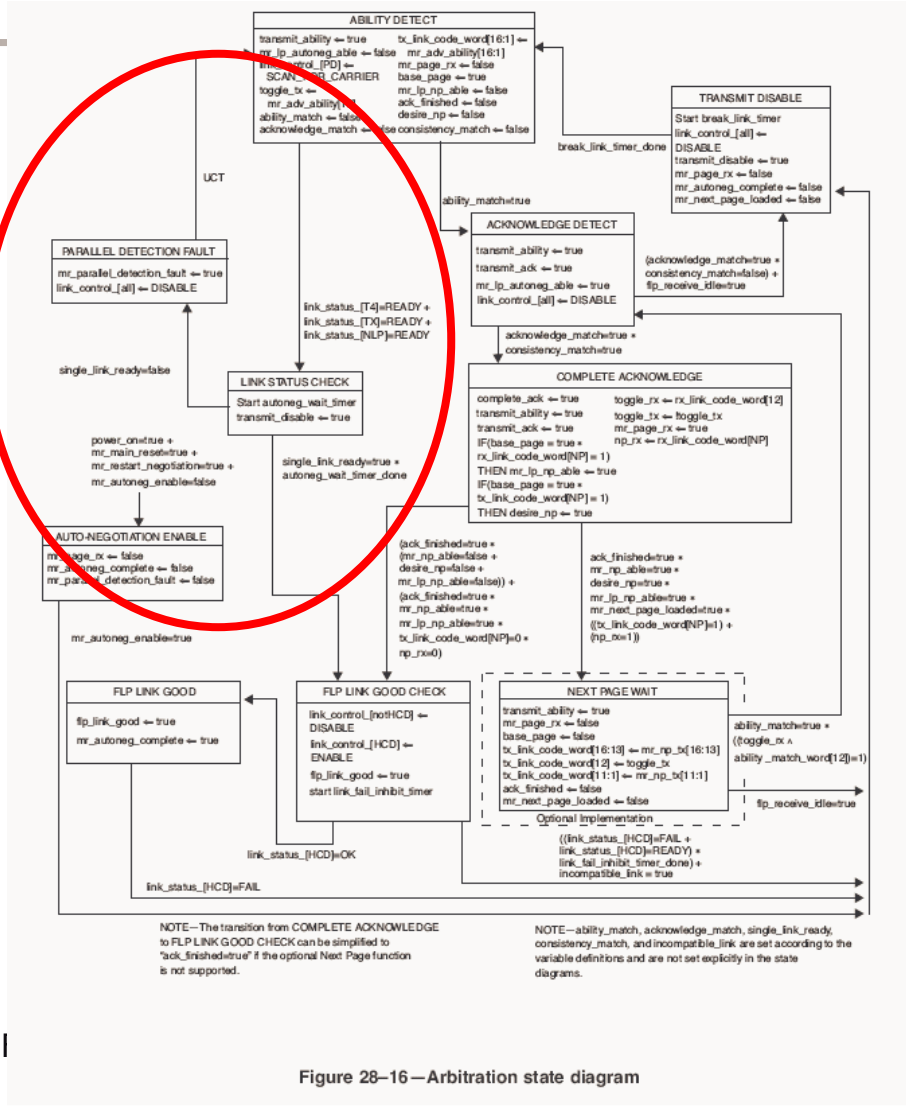
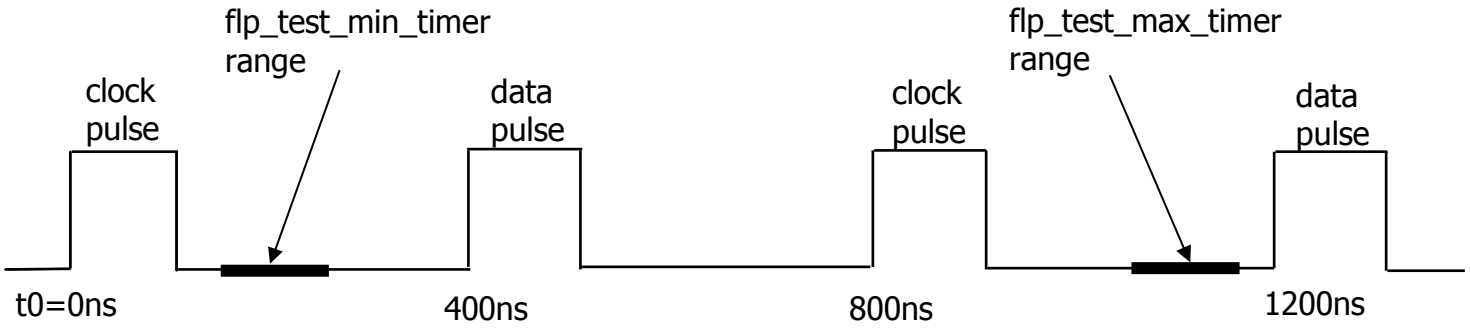


Figure 28–16—Arbitration state diagram

How to detect SSP bursts versus legacy

- Signal Detect constantly at a high level indicates
  - If Auto-Negotiation has completed: a valid link is present
  - If Auto-Negotiation has not completed: a legacy link is present
- Signal Detect pulsing indicates
  - That SSP Bursts are present (if the pulses adhere to SSP timing spec)
  - That an unstable link is present (if the pulses do not adhere to the SSP timing specification



Proposed Figure 28-8 – SSP detect timers (flp_test_min/max_timers)

link_status=ready

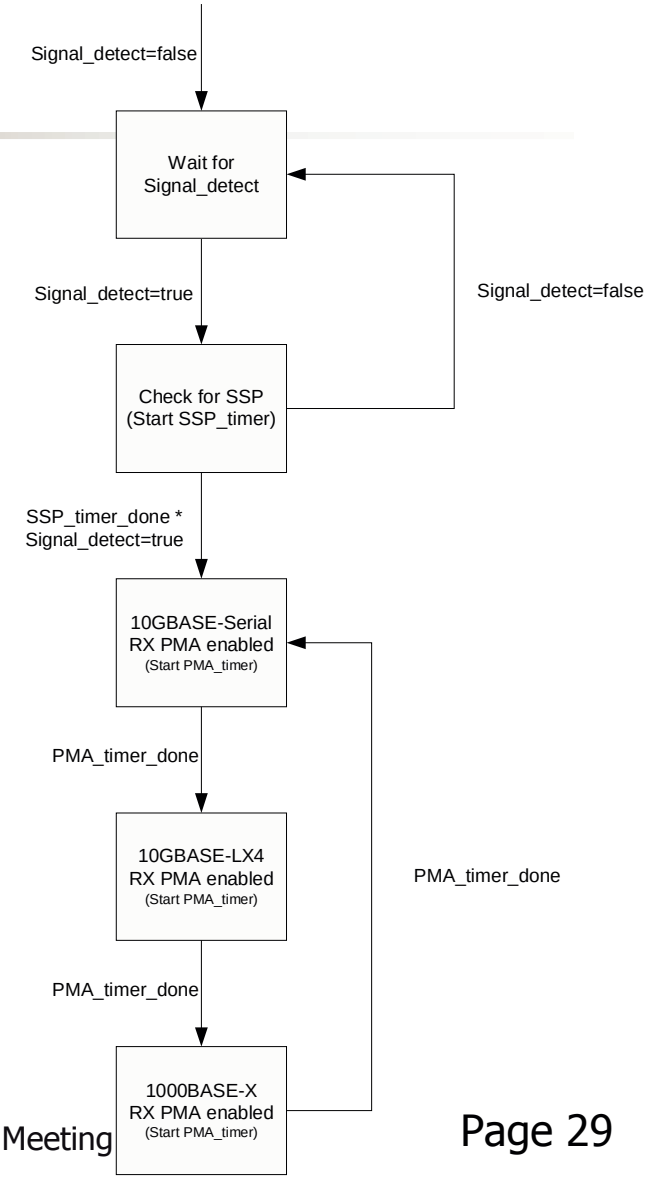
- 1000BASE-X
  - RX synchronization state machine in sync_acquired state (valid commas received)
  - sync_status=OK
- 10GBASE-X4 (XAUI)
  - Sync acquired on all four lanes
  - All lanes aligned
  - align_status=OK
- 10GBASE-Serial
  - TBD

Clause 28 Parallel Detection Issues

- Clause 28 Parallel Detection requires all the PMA functionality to be active simultaneously – no shared resources!
  - In theory yes, but specific implementations can avoid this, see next slide
  - This is not part of the proposal, but is for information only

Parallel Detection with shared PMA functions

- 2 x PMA_timer < autoneg_wait_timer (currently 500-1000ms)
- SSP timer is ~ 300ns-1000ns
- PMA_timer should be long enough to allow PMA to get link_status=ready
- Link_status for each PMA should be saved until Signal_detect is false or Auto Negotiation has completed



Optional Enhancements

- It is allowed to support Clause 37 Auto-Negotiation after the parallel detect function has detected a valid 1000BASE-X link

Summary

- Auto-Neg for Plug-N-Play link configuration
- Use Clause 28 state machines for exchanging pages and Parallel detection [4]
  - Significantly fast page transfers by lowering timer values
  - High efficiency and Fast Link Convergence
- New 48-bit pages and selector fields for 802.3 Backplane Ethernet
- Use Proposed SERDES compatible FLP(SSP) [1]
- Support Parallel Detection for devices that do not support 802.3ap Auto-Negotiation and with legacy 1000Base-X and XAUI backplane links

References

- [1] SERDES compatible FLP AN proposal for 802.3ap, July 04 Plenary
  - http://ieee802.org/3/ap/public/jul04/szczepanek_02_0704.pdf
- [2] UNH-IOL Clause 28 Auto-Negotiation Management System Test Suite, UNH IOL Ethernet Interoperability Testing presentation
  - http://ftp.iol.unh.edu/fec/anegSystem.pdf
  - http://www.nbl.org.tw/nbl_old/nbl-iol-workshop/07_Ethernet_Overview.pdf
- [3] Earlier Auto-Neg presentations to Backplane Ethernet SG/TF
  - http://www.ieee802.org/3/bladesg/public/mar04/chang_01_0304.pdf
  - http://www.ieee802.org/3/ap/public/may04/ganga_01_0504.pdf
  - http://www.ieee802.org/3/ap/public/jul04/ganga_01_0704.pdf
- [4] IEEE Std 802.3-2002, IEEE Std 802.3ae-2002
- [5] IEEE Std 802.3ak-2004, P802.3ah approved draft D3.3 (to be published as IEEE Std 802.3ah-2004)
- [6] Improving Autoneg Efficiency proposal to 802.3an/802.3ap TF by Pat Thaler
  - http://www.ieee802.org/3/an/public/sep04/thaler_1_0904.pdf
  - http://www.ieee802.org/3/ap/public/sep04/thaler_01_0904.pdf