Specification of optical PHY type auto-negotiation (OAN)

Matt Brown, Alphawave Semi
Kent Lusted, Intel
Ali Ghiasi, Ghiasi Quantum

Contributors

- <contributors>

Supporters

- <supporters>

Introduction

- Contribution brown_3dj_01a_2311 proposed use of the electrical auto-negotiation specifications as a starting point for optical automatic link configuration.
  - Some slides are repeated in this presentation for context.
- This contribution takes that proposal a step further as follows providing specific details for Optical Auto-Negotiation (OAN).

Scope

- This contributions limits specification as follows…
- OAN for IMDD PHY types for SMF fiber, either parallel fiber or duplex fiber.
  - This could be extended in the future to support coherent PHY types or PHY types using MMF fiber.
- OAN between PHYs using the same lane types, either
  - Between PHYs supporting parallel fiber
  - Between PHYs supporting duplex fiber
  - Single-lane PHYs such as 200GBASE-DR1/FR1 fall into either category

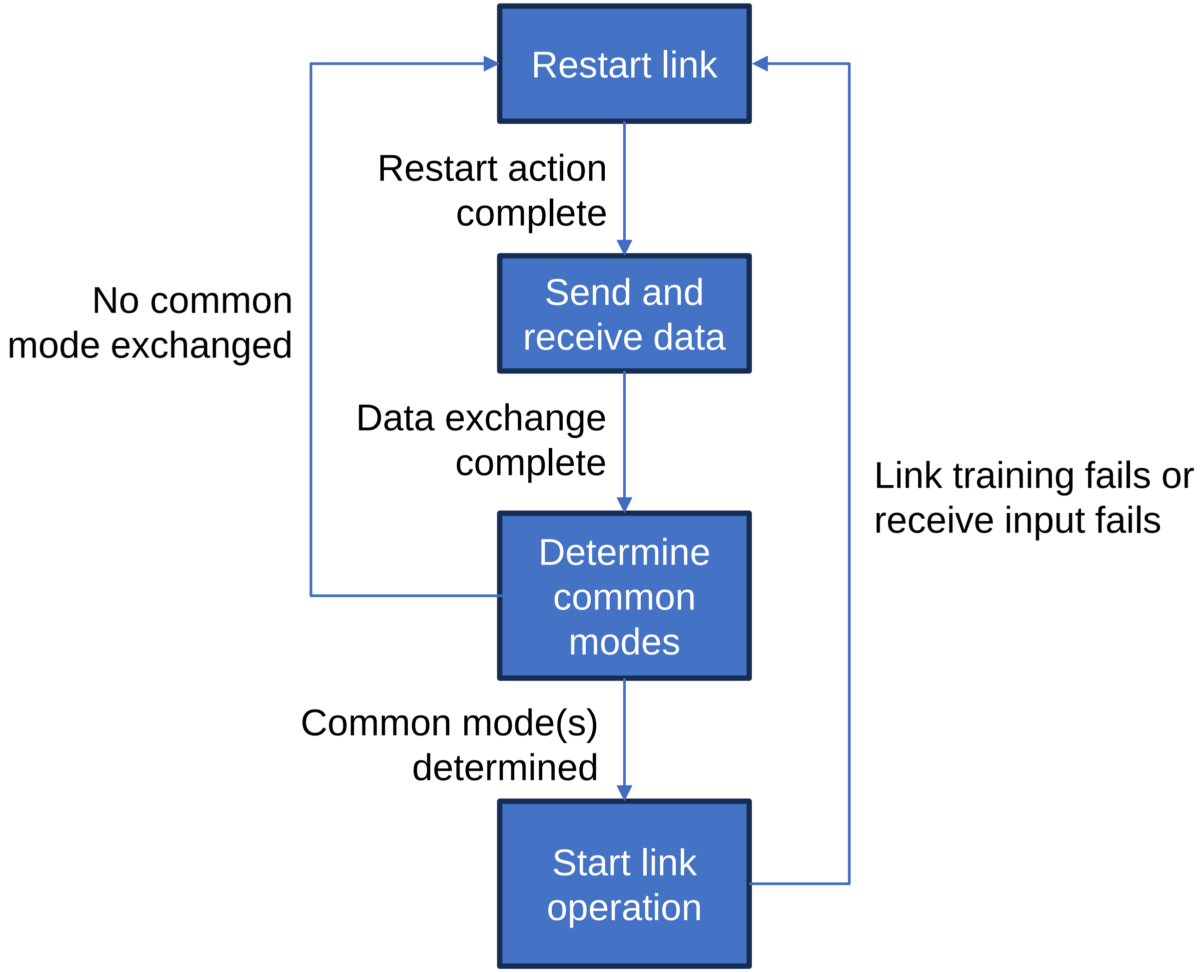
Purpose (redux)

- Establish common configuration (or lack thereof) for link partners at each end of an optical fiber.
- For example, determine common PHY type
    - Both ends use the same PMD type, e.g., 800GBASE-DR4
- Other selectable parameters could be addressed as well.

Useful features (redux)

- Means to exchange information between link partners
  - e.g., signaling and data structure
- Means to initiate automatic configuration
- Means to select which technologies to permit
- Means to select a common mode of operation
- Means to transition from automatic configuration to data mode
- State machines to coordinate the above features

High level state diagram (redux)



Reuse of Clause 73 (electrical) Auto-Negotiation

- Include new optical auto-negotiation (OAN) sublayer below the PMD.
- Specify whether OAN is mandatory to implement and optional to use or optional to implement/use.
- Imply that OAN is co-resident with the PMD on the module, i.e., not on the host.
- Use signaling as defined in Clause 73, except…
  - edge spacing at 75.3 ps (53.125 GBd / 4 = 13.28125 GBd) rather than 3.2 ns (312.5 MBd)
    - Same as training frame control channel proposed in ghiasi_3dj_01_2311; 1/8 PAM4 symbol rate
  - Specify OMA requirements rather than electrical peak to peak swing as well as other optical parameters as necessary, inclusive of any anticipated PMD specifications
- Signaling over a specific lane TBD for multi-lane PHYs, other lanes disabled.
- Use page structure (base and next pages) and delineation based on those defined in Clause 73.
- Specify new allocation of PHY types and capabilities to base page.
- Specify new PHY prioritization table.
- Specify selection criteria for other capabilities, e.g., FEC, interleaving.
  - At this time no such optional capabilities have been adopted.
- Use state machines defined in Clause 73, modified as appropriate.

Optical Auto-Negotiation (OAN) sublayer

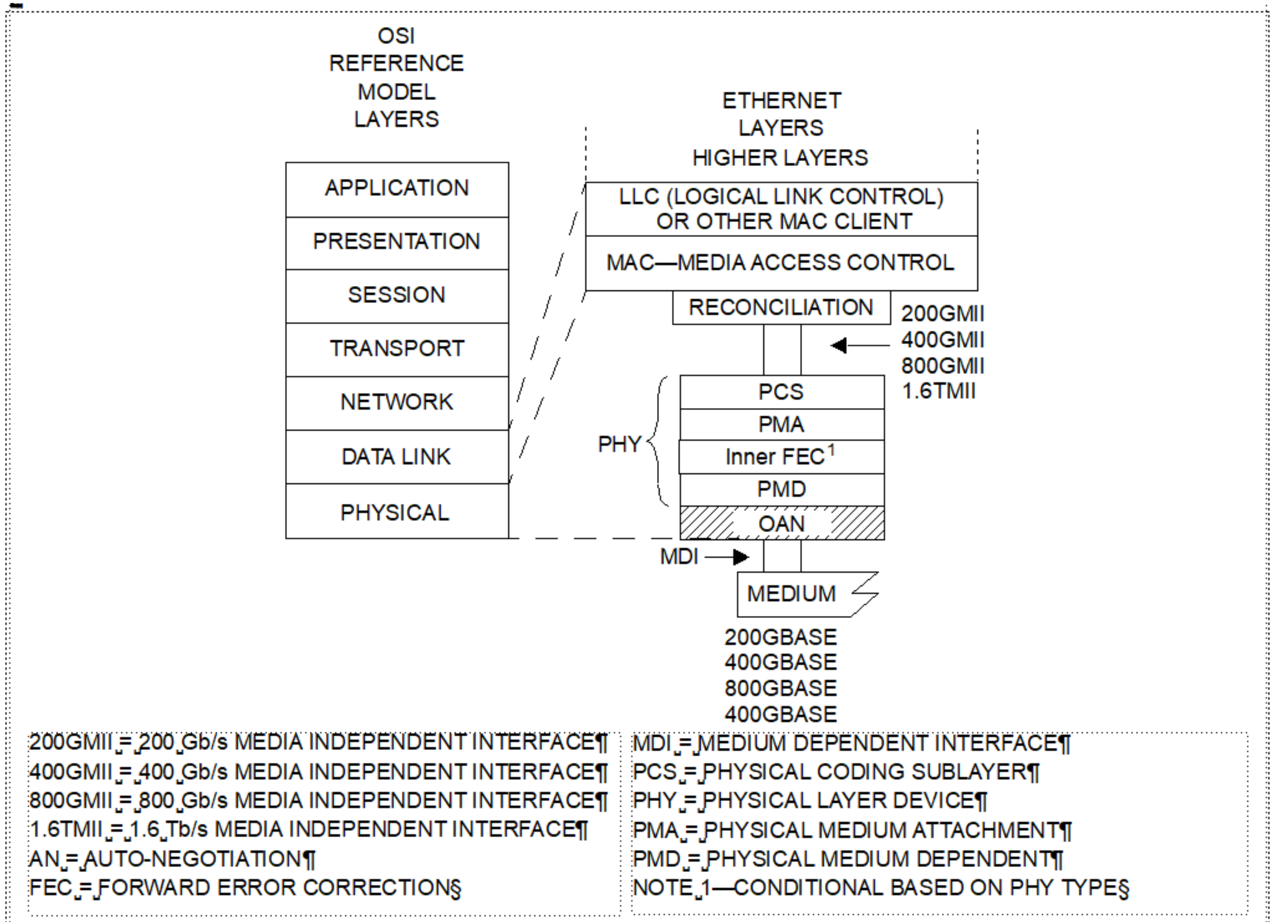


Figure 500–1—OAN function relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 Ethernet model§

- OAN specified in a standalone clause or annex.
  - In the figure to the left, it is assumed that it is specified in Clause 500.
  - Based upon, but not the same as Clause 73.
- OAN sublayer is positioned below the PMD.
- Implementation of OAN is optional.
- Use of OAN is optional.
- The PHY is configured only once AN has determined a common technology (highest common denominator or HCD).

Functions and signaling

- OAN has three functions
  - OAN transmit – send advertised information to the link partner
  - OAN receive – receive advertised information from the link partner
  - OAN arbitration – determines common advertised technology
- For multiple-lane PHY types, AN signal is sent on
  - Lane 0 for parallel fiber PHY types
  - Lane TBD (@ particular wavelength) for duplex fiber PHY types

Clause 73 – link codeword— encoding (redux)

73.6 Link codeword encoding

The base link codeword (Base Page) transmitted within a DME page shall convey the encoding shown in Figure 73–6. The Auto-Negotiation function supports additional pages using the Next Page function. Encoding for the link codeword(s) used in the Next Page exchange are defined in 73.7.7. In a DME page, D0 shall be the first bit transmitted.

D[4:0] contains the Selector Field. D[9:5] contains the Echoed Nonce field. D[12:10] contains capability bits to advertise capabilities not related to the PHY. C[1:0] is used to advertise pause capability. The remaining capability bit C[2] is reserved. D[15:13] contains the RF, Ack, and NP bits. These bits shall function as specified in 28.2.1.2. D[20:16] contains the Transmitted Nonce field. D[43:21] contains the Technology Ability Field. D[47:44] contains FEC capability (see 73.6.5).

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

D 16	D 17	D 18	D 19	D 20	D 21	D 22	D 23	D 24	D 25	D 26	D 43	D 44	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 22	F 2	F 3	F 0	F 1

Figure 73–6—Link codeword Base Page

73.6.1 Selector Field

Selector Field (S[4:0]) is a five-bit wide field, encoding 32 possible messages. Selector Field encoding definitions are shown in Annex 28A. Combinations not specified are reserved for future use. Reserved combinations of the Selector Field shall not be transmitted.

The Selector Field for IEEE Std 802.3 is shown in Table 73–3.

Table 73–3—Selector Field Encoding

S4	S3	S2	S1	S0	Selector description
0	0	0	0	1	IEEE Std 802.3

73.6.2 Echoed Nonce Field

Echoed Nonce Field (E[4:0]) is a 5-bit wide field containing the nonce received from the link partner. When Acknowledge is set to logical zero, the bits in this field shall contain logical zeros. When Acknowledge is set to logical one, the bits in this field shall contain the value received in the Transmitted Nonce Field from the link partner.

73.6.3 Transmitted Nonce Field

Transmitted Nonce Field (T[4:0]) is a 5-bit wide field containing a random or pseudo-random number. A new value shall be generated for each entry to the Ability Detect state. The method of generating the nonce is left to the implementer. The transmitted nonce should have a uniform distribution in the range from 0 to $2^5 - 1$. The method used to generate the value should be designed to minimize correlation to the values generated by other devices.

Link codeword fields

- Define the link codeword fields based on Clause 73 as follows:
  - The base page provides 48 bits, denoted D0 to D47 or D[0:47] of information to be exchanged.
    - Selector field S[0:4] from D[0:4]
    - Echoed nonce field E[0:4] from D[5:9]
    - Pause capability field C[0:2] from D[10:12]
    - RF/ACK/NP fields from D(13:15)
    - Transmitted nonce field T[0:4] from D[16:20]
    - Technology advertisement A[0:26] from D[21:47]
  - Define D[0:20] as defined in Clause 73.
  - Define A[0:26] as in Clause 73 with technologies mapped on next slide.

Link codeword advertisement field A[0:27]

- Allocate A[0:27] to technologies as shown in the table to the right.
- A value of 1 indicates that the PHY is advertising the corresponding technology.
- It is expected that if advertising a technology, the implementation can support that PHY type.
- All technologies supported by the implementation need not be advertised.
  - e.g., A particular PHY type can be forced by advertising only that technology.

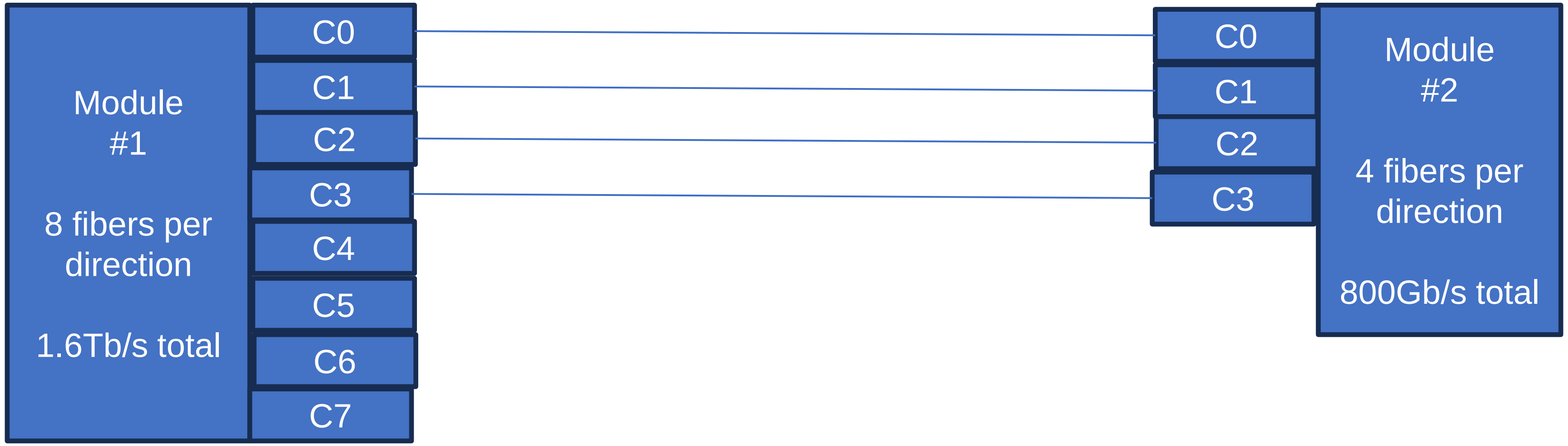
Bit	Technology	Capability
A0	1.6TBASE-DR8-2	1.6 Tb/s, 2 km, parallel
A1	1.6TBASE-DR8	1.6 Tb/s, 500 m, parallel
A2	800GBASE-LR4	800 Gb/s, 10 km, duplex
A3	800GBASE-FR4	800 Gb/s, 2 km, duplex
A4	800GBASE-DR4-2	800 Gb/s, 2 km, parallel
A5	800GBASE-FR4-500m	800 Gb/s, 500 m, duplex
A6	800GBASE-DR4	800 Gb/s, 500 m
A7	400GBASE-DR2-2	400 Gb/s, 2 km, parallel
A8	400GBASE-DR2	400 Gb/s, 500 m
A9	200GBASE-FR1	200 Gb/s, 2 km, parallel
A10	200GBASE-DR1	200 Gb/s, 500 m
A11	Reserved	
…		
A27	Reserved	

Priority resolution table

- If the two devices, one at each end of a fiber, advertise multiple common technologies, then a mechanism to reconcile is required.
- A prioritization table is defined to determine which of a set of common technologies to select.
- A priority table, to the right, is proposed as a means to determine the highest priority common technology.
- The proposed order of priority based upon the following criteria, in order of prioritization:
  - Ethernet rate: higher rate = higher priority
  - Reach: longer reach = higher priority
  - Data rate per lane: higher data rate = higher priority
  - Number of fibers (inverse): fewer fibers = higher priority
    - e.g., FR4 higher priority than DR4-2

Priority	Technology	Capability
1 highest	1.6TBASE-DR8-2	1.6 Tb/s, 2 km, parallel
2	1.6TBASE-DR8	1.6 Tb/s, 500 m, parallel
3	800GBASE-LR4	800 Gb/s, 10 km, duplex
4	800GBASE-FR4	800 Gb/s, 2 km, duplex
5	800GBASE-DR4-2	800 Gb/s, 2 km, parallel
6	800GBASE-FR4-500m	800 Gb/s, 500 m, duplex
7	800GBASE-DR4	800 Gb/s, 500 m
8	400GBASE-DR2-2	400 Gb/s, 2 km, parallel
9	400GBASE-DR2	400 Gb/s, 500 m
10	200GBASE-FR1	200 Gb/s, 2 km, parallel
11 lowest	200GBASE-DR1	200 Gb/s, 500 m

Example of technology resolution



Scenario	Module #1 advertises on C0	Module #2 advertises on C0	Common advertised capabilities	HCD selected based on priority table
1	1.6TBASE-DR8 (A1) 800GBASE-DR4 (A6) 400GBASE-DR2 (A8) 200GBASE-DR1 (A10)	400GBASE-DR2 (A8) 200GBASE-DR1 (A10)	400GBASE-DR2 (priority = 9) 200GBASE-DR1 (priority = 11)	400GBASE-DR2 (select highest rate)
2	800GBASE-DR4-2 (A4)	800GBASE-DR4-2 (A4) 800GBASE-DR4 (A6)	800GBASE-DR4-2	800GBASE-DR4-2 (module #1 forces FECi)
3	800GBASE-DR4 (A6)	800GBASE-DR4-2 (A4) 800GBASE-DR4 (A6)	800GBASE-DR4	800GBASE-DR4 (module #1 forces FECo)

Use of “next pages”

- “next pages” provide extra bits to expand beyond the base page.
- Any number of “next pages” may be used.
- These can be used for various reasons:
  - Not enough bits on base page; use for forward compatibility.
  - Use for vendor specific purpose outside the scope of IEEE.
- For OAN, include “next pages” as specified in 73.7.7 for forward compatibility and vendor specific use.

OAN Signal Characteristics

- For OAN, the signal will be defined by optical rather than electrical characteristics.
- The table below is proposed as a starting point.
- The range of OMA and wavelength should be specified to be inclusive of the ranges natural for the expected PMD types.
- Are there other optical parameters that must be constrained?

Parameter	Value	Unit
Transmit OMA	TBD to TBD	dBm
Receive OMA	TBD to TBD	dBm
Wavelength	TBD to TBD	nm

Clause 73 – AN signaling (redux)

Table 73–1—DME electrical characteristics

Parameter	Value	Units
Transmit differential peak-to-peak output voltage	600 to 1200	mV
Receive differential peak-to-peak input voltage	200 to 1200	mV

Table 73–2— DME page timing summary

	Parameter	Min.	Typ.	Max.	Units
T1	Transition position spacing (period)	3.2 −0.01%	3.2	3.2 +0.01%	ns
T2	Clock transition to clock transition	6.2	6.4	6.6	ns
T3	Clock transition to data transition (data = 1)	3.0	3.2	3.4	ns
T4	Transitions in a DME page	51	—	100	—
T5	DME page width	338.8	339.2	339.6	ns
T6	DME Manchester violation delimiter width	12.6	12.8	13.0	ns

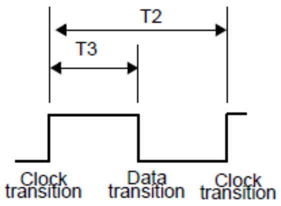


Figure 73–4—DME page transition timing

Given 3.2 ns width (T3) this is equivalent to NRZ signaling rate 312.5 MBd

The encoding of data using DME bits in an DME page is illustrated in Figure 73–3.

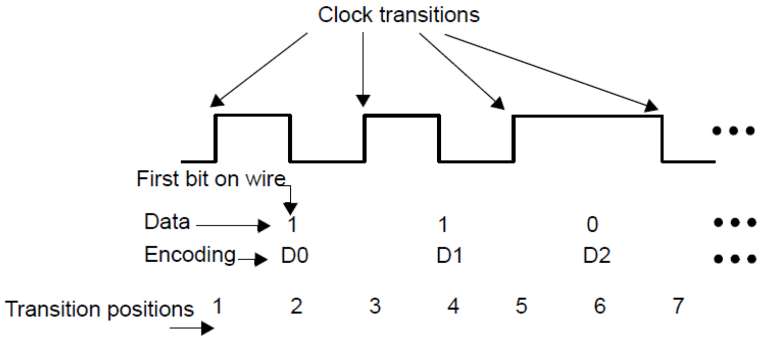


Figure 73–3—Data bit encoding within DME pages

73.5.3.1 Manchester violation delimiter

A violation is signaled as shown in Figure 73–5.

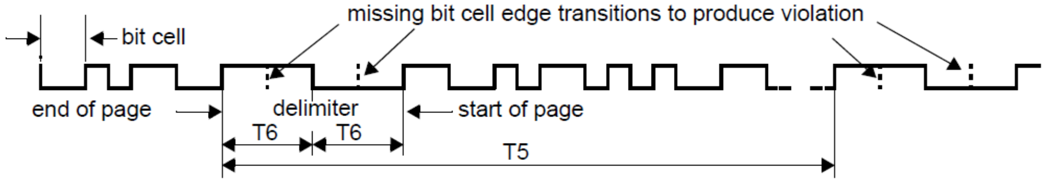


Figure 73–5—Manchester violation

OAN Timing Characteristics

- For OAN, the signal will be defined by optical rather than electrical characteristics.
- The table below is proposed as a starting point.
- Typical transition distance based on unit interval for 53.125 GBd / 4 = 13.28125 GBd and with minimum and maximum values scaled as in Table 73-2.

	Parameter	Min.	Typ.	Max.	Units
T1	Transition position spacing (period)	75.2941 – 0.01%	75.2941	75.2941 + 0.01%	ps
T2	Clock transition to clock transition	145.8823	150.5882	155.2941	ps
T3	Clock transition to data transition (data = 1)	70.5882	75.2941	80	ps
T4	Transitions in a DME page	51	—	100	—
T5	DME page width	7482.3529	7981.1765	8480	ps
T6	DME Manchester violation delimiter width	282.3529	301.1765	320	ps

Clause 73 AN – State diagrams – transmitter and receiver (redux)

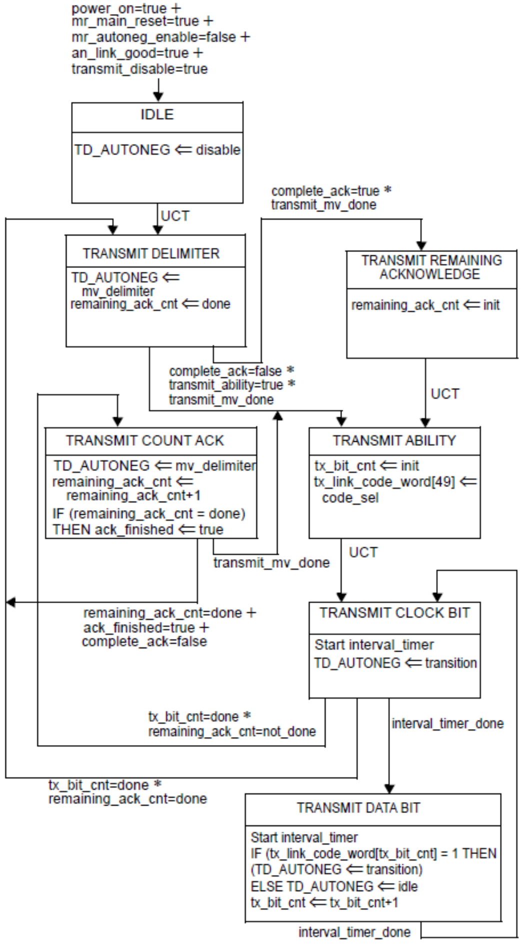


Figure 73–9—Transmit state diagram

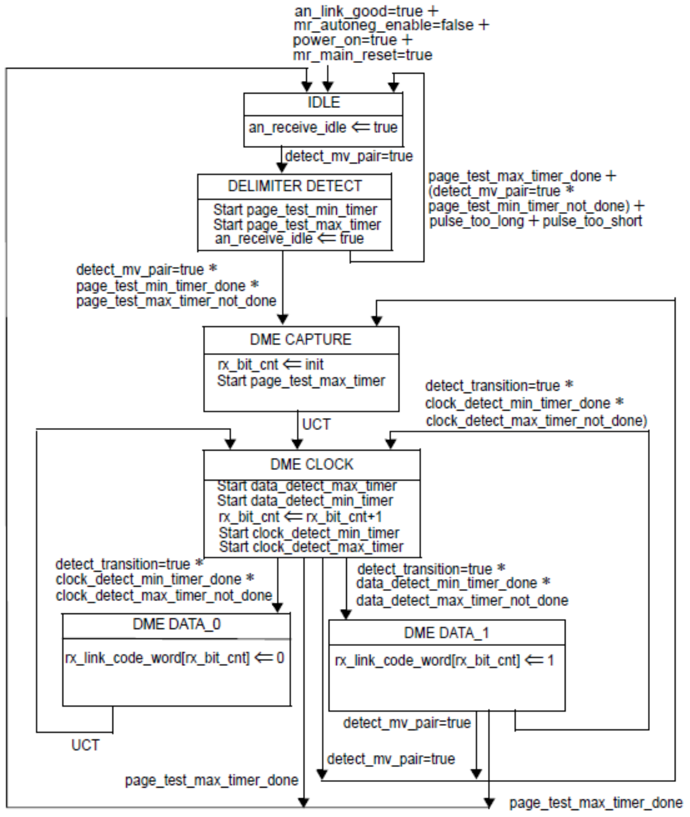


Figure 73–10—Receive state diagram

Clause 73 AN – arbitration (redux)

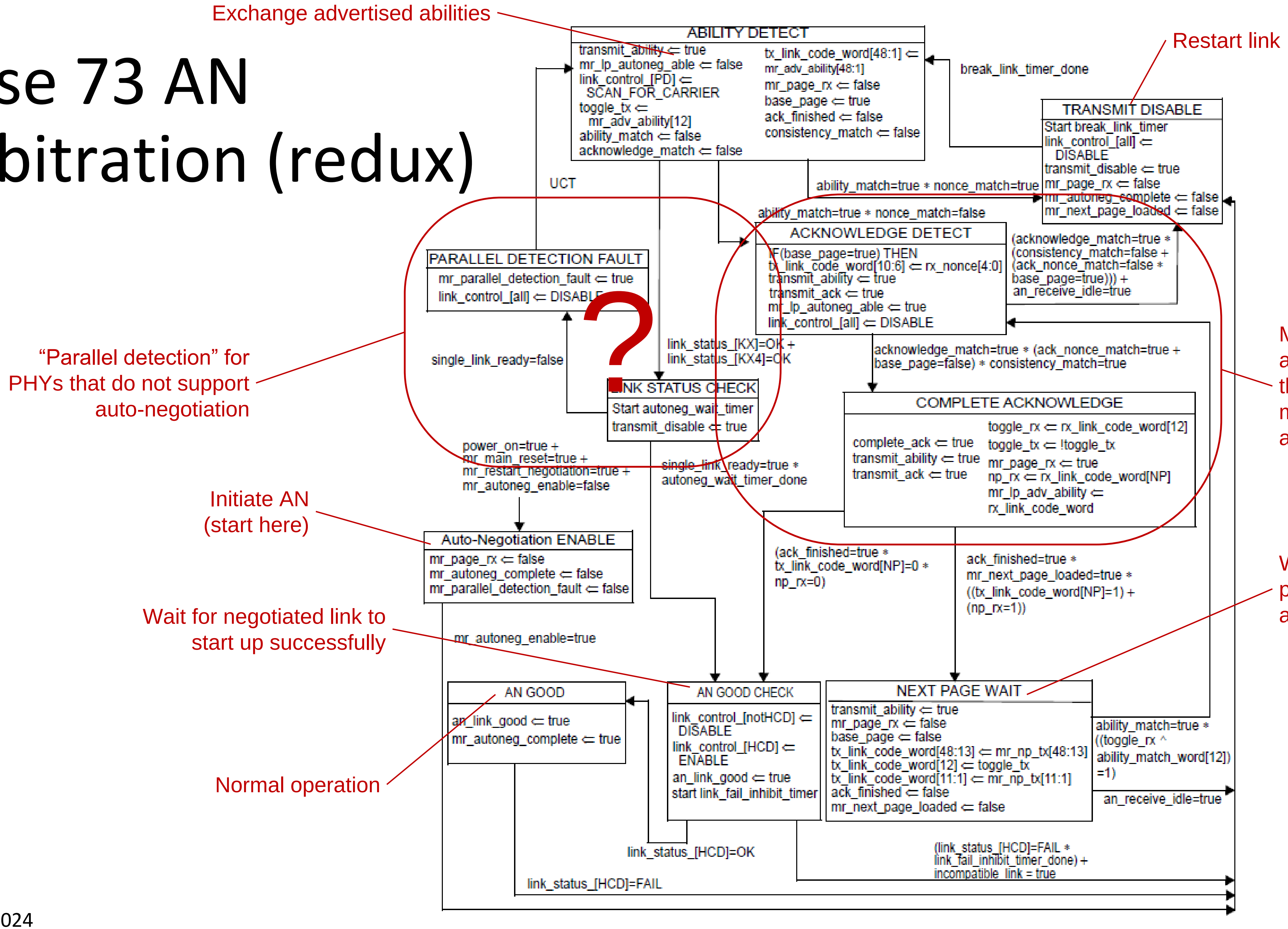


Figure 73–11—Arbitration state diagram

State Diagrams

- The transmit and receive state diagrams manage the pattern generation and detection
  - The Clause state diagrams should be equally applicable to OAN.
  - For OAN use the state diagrams in Figure 73-9 and Figure 73-10 without modification.
- The arbitration state diagram manages AN as a whole including initializing and restarting the link, interpretation of the pages.
  - The Clause 73 state diagram is equally applicable, except that the parallel detect function might not be relevant.
  - For OAN use the state diagrams in Figure 73-11 except
    - Exclude the parallel detect function for now. Can add parallel detect later if we want AN to work with legacy 100 Gb/s per PMDs or with 200 Gb/s per lane PMDs without OAN implemented.
    - link_fail_inhibit_timer value is TBD.

Summary

- The presentation has provided details for specifics of auto-negotiation for optical PHYs (OAN).
- Consider this as a starting point for a baseline.

Thanks