Gigabit Ethernet Serial Link Codes

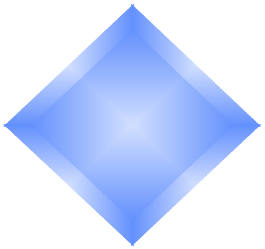
Proposal for serial link codes and receiver/transmitter states based on the PCS protocol requirements.

Contents

IEEE 802.3z Gigabit Ethernet Task Force
September 9-11, 1996 Interim Meeting
Coeur d'Alene Inn and Conference Center
Coeur d'Alene, Idaho
Richard Taborek
amdahl *Corp.*

The following companies have indicated their support for the concepts outlined in this proposal (in alphabetical order):

3Com, Amdahl, Cisco, Compaq, DEC, Granite, H-P, Madge, Packet Engines, Sun, VLSI Logic and Xaqti



High-Level Assumptions

- Codings are based on 8B10B data codes (Dx.y) and special control characters (Kx.y) as defined in ANSI X3.230 FC-PH (Fibre Channel) Clause 11.
- Codings should take into careful consideration the error detection properties of code words selected.
- Codings are required for all primitives specified for the link startup sequence.
- Codings should provide sufficient coding space for all automatic capability detection parameters and shall be similar in nature to the base capability register defined in IEEE 802.3u clause 28.
- Automatic capability detection codings should allow for future expansion of the associated protocol.
- Codings are required for a link Idle indication.
- Codings are required for IEEE 802.3 packets and bursts of packets during a carrier event.
- Codings are required for packet start and end delimiters.
- Coding is required for a bad packet information indication.
- Coding is required for an indication which allows the carrier to be extended.
- Coding is required for an indication which ensures that packet or carrier extension always concludes on an odd-numbered character.

Nomenclature

- One or more data codes and special characters are grouped into transmission words.
- Special transmission words called ordered sets are defined.
- Packet delimiter ordered sets are used to mark packet boundaries.
- Signal ordered sets are used to signal events.
- Sequence ordered sets are repeated ordered sets used to signal states, carry information and operate in environments with relatively (to frames) high error rates.
- The link is not considered usable until Link_Startup is complete.
- Synchronization is achieved when the receiver identifies the same transmission word boundary on the received bit stream as that established by the transmitter at the other end of the link.
- The detection of invalid transmission words is an indication that a receiver is out of synchronization.
- The clock recovery system in a receiver shall have sufficient hysteresis to prevent a single transmission error from causing it to go out of synchronization.
- A Loss_of_Synchronization procedure defines the method by which the receiver changes from the Synchronization_Acquired state to the Loss_of_Synchronization state.
- The first transmission character of an ordered set transmitted over an operational link is transmitted in an even-numbered character position. Subsequent characters continuously alternate as odd and even-numbered characters.

8B10B Transmission Code

- 8B10B transmission code provides the following functions:
  - Improves transmission characteristics
  - Enables bit-level clock recovery
  - Improves error detection
  - Separates data symbols from control symbols
  - Derives bit and word synchronization
- 8-bit data bytes are encoded as 10-bit Data Characters.
- 12 Special Characters are defined for special signaling.
- One or more Data and/or Special Transmission Characters may be grouped into Transmission Words.
- Special Transmission Words called Ordered Sets are defined.
- Ordered Sets are flexible building blocks which may be used for in-band or out-of-band protocol functions.

Ordered Set Usage

- Various coding objectives suggest the the specification of multiple length ordered sets.
  - A single character ordered set is required to ensure that the packet or carrier extension always ends on an odd-numbered character.
  - Multiple character ordered sets are required for sequences which include a comma and communicate a small amount of information. Two-character ordered sets provide 16/20 bit alignment.
  - The current requirements for automatic capability detection call for the communication of greater than 8 bits of information which requires either multiple 2 character sequences or the use of larger ordered sets. The use of 4 character ordered sets provides ample coding space and provides information within a single ordered set.
- Packets and sequences shall start only on even-numbered characters for consistency and error robustness.
  - This includes intermediate packets within a burst of packets.

Special Character Usage

- The K28.5 special character is chosen as the first character of all sequences for the following reasons:
  - Bits abcdeif make up a comma. A comma is a singular bit pattern which in the absence of transmission errors cannot appear in any other location of a transmission character and cannot be generated across the boundaries of any two adjacent transmission characters.
  - The comma can be used to easily find and verify character and word boundaries of the received bit stream.
  - Bits ghj of the encoded character present the maximum number of transitions, simplifying receiver acquisition of bit synchronization.
- Special characters other than K28.5 are specified for single character ordered sets for alignment purposes or to provide timely function recognition.
  - Other special characters are chosen from the list of 12 available special characters.
  - The list is reduced to 9 when special characters containing a comma are excluded to prevent ordered set alignment on odd-numbered characters.

Defined Ordered Sets
Idle Signal and Sequence

- The I signal or sequence is used to signal Idle.
  - I effects the Link Up primitive which indicates that the link is available to send packets.
  - The first I following a packet or Link_Configuration sequence shall restore the current positive or negative running disparity to a negative value. One I signal ordered set is required for this purpose.
  - All subsequent I sequence ordered sets shall insure negative running disparity.
  - I codes should have a high transition density to keep the receiver in optimum sync during the high-frequency Idle sequence.
  - Two I codes, I1 and I2, are defined consisting of a K28.5 special character followed by a data character.
  - Distinct carrier events should be separated by Idles. When a receiver sees any Idle, it drops carrier.

Defined Ordered Sets
Start and End of Packet Delimiters

- The S delimiter is used to signal Start_Of_Packet (SOP).
  - A single character ordered set is desired in order to quickly assert the CSMA/CD carrier indication.
  - S follows Idle for the first packet during a carrier event.
  - S follows R for subsequent packets during a carrier event.

- The T delimiter is used to signal End_Of_Packet (EOP).
  - A single character ordered set is desired in order to assist in quick de-assertion of the carrier indication.

Defined Ordered Sets
Align/Extend Sequence

- The R signal is used to force even-numbered character alignment for the first Idle following a carrier event and to extend the carrier following a packet when necessary.
  - A single character ordered set is required.
  - If T is transmitted as an odd-numbered character, and the carrier is not extended, T shall be followed by two successive R signals.
    - Two R's rather than none are required because of the possibility of running disparity error propagation to characters beyond those during which the error occurred.
  - If T is transmitted as an even-numbered character, and the carrier is not extended, exactly one R shall be transmitted after the T.
  - Additional R's shall be transmitted following a T to extend the carrier as required. The last R to extend the carrier must be transmitted in an odd-numbered character position.
- R shall also be used to separate the packets within a burst.

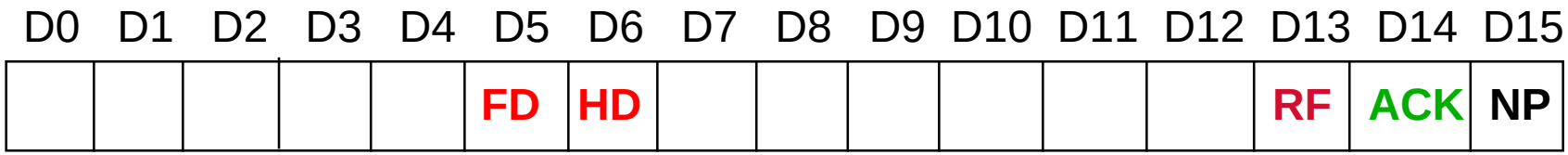
Defined Ordered Sets
Invalid Packet Delimiter/Signal

- The H delimiter and signal is used to replace individual invalid data and delimiters characters and mark those entities as invalid.
  - A single character ordered set is required.
  - An inter-station unit, such as a repeater, may substitute H for invalid data and delimiters upon detection of invalid characters.
  - H shall **not** be used to replace R.
    - The last H to extend the carrier must be transmitted in an odd-numbered character position.

Defined Ordered Sets
Link_Not_Available & Link_Configuration

- The F sequence is used to indicate Link_Not_Available.
  - F is signaled continuously while the associated receiver attempts to acquire synchronization.
- The C Sequence is used to indicate Link_Configuration.
  - C conveys a single parameter to the other end of the link.
    - The parameter is a 16-bit Config_Register that includes bits sufficient to specify the capabilities of a PHY as well as an ACK bit.
    - Config_Register format is styled after the Link Code Word (LCW) defined in IEEE 802.3u clause 28.
  - Config_Register data bits are coded as the third and forth characters of the Link_Configuration ordered set.
  - The order of transmission of Config_Register data is encoded bits d0:d7 followed by encoded bits d8:d15.
    - The transmitted bit streams bears little resemblance to order of the Config_Register.

Config_Register



Config_Register bit usage:

- D5/FD: Full duplex capable
- D6/HD: Half duplex capable
- D13/RF: Remote Fault
- D14/ACK: Acknowledge
- D15/NP: Next Page (Escape)

Listing of Ordered Sets

Code	Function	Encoding	Beg. RD	End RD
F	LINK_NOT_AVAILABLE	K28.5 D21.5	?	flip
C	LINK_CONFIGURATION	K28.5 D10.5 config_reg	?	?
I1	Idle/Flip Disparity	K28.5 D5.6	+	-
I2	Idle/Disparity OK	K28.5 D16.2	-	-
S	SOP	K27.7	?	same
T	EOP1	K29.7	?	same
R	EOP2	K23.7	?	same
H	EOPinvalid	K30.7	?	same

Coding Distance of Characters Used in Ordered Sets

Char	D21.5	D16.2	D10.5
D16.2	7		
D10.5	6	7	
D5.6	4	4	6

- Data characters are chosen for high transition density, proper disparity control, and sufficient coding distance.

Link Information Example
Part 1/2

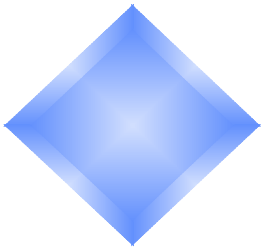
802.3 info	8B10B codes
LINK_NOT_AVAILABLE	F
LINK_NOT_AVAILABLE	F
~	~
LINK_NOT_AVAILABLE	F
LINK_CONFIGURATION	C
LINK_CONFIGURATION	C
~	~
LINK_CONFIGURATION	C
Idle/Disparity OK	I2
Idle/Disparity OK	I2
~	~
Idle/Disparity OK	I2
0 PREAMBLE	S (single short packet)
1 PREAMBLE	Data = 10101010
2 PREAMBLE	Data = 10101010
3 PREAMBLE	Data = 10101010
4 PREAMBLE	Data = 10101010
5 PREAMBLE	Data = 10101010
6 PREAMBLE	Data = 10101010
7 SFD	Data = 10101011
8 DA	Data = DA...
~	~
LLC DATA	Data = LLC DATA...
~	~
FCS1	Data = FCS1 octet...

802.3 info	8B10B codes
FCS2	Data = FCS2 octet...
FCS3	Data = FCS3 octet...
FCS4	Data = FCS4 octet...
EOP1	T
EOP2	R
EOP2	R
~	~
EOP2	R (odd-num character)
Idle/Flip Disparity	I1
Idle/Disparity OK	I2
~	~
Idle/Disparity OK	I2
LINK_NOT_AVAILABLE	F (error or reconfig)
LINK_NOT_AVAILABLE	F
~	~
LINK_NOT_AVAILABLE	F
LINK_CONFIGURATION	C
LINK_CONFIGURATION	C
~	~
LINK_CONFIGURATION	C
Idle/Flip Disparity	I1
Idle/Disparity OK	I2
~	~
Idle/Disparity OK	I2
~	~

Link Information Example
Part 2/2

802.3 info	8B10B codes
Idle/Disparity OK	I2
0 PREAMBLE	S (1st packet of burst)
1 PREAMBLE	Data = 10101010
2 PREAMBLE	Data = 10101010
3 PREAMBLE	Data = 10101010
4 PREAMBLE	Data = 10101010
5 PREAMBLE	Data = 10101010
6 PREAMBLE	Data = 10101010
7 SFD	Data = 10101011
8 DA	Data = DA...
~	~
LLC DATA	Data = LLC DATA...
~	~
FCS1	Data = FCS1 octet...
FCS2	Data = FCS2 octet...
FCS3	Data = FCS3 octet...
FCS4	Data = FCS4 octet...
EOP1	T (even char/no extension)
EOP2	R
0 PREAMBLE	S (2nd packet of burst)
1 PREAMBLE	Data = 10101010
2 PREAMBLE	Data = 10101010
3 PREAMBLE	Data = 10101010
4 PREAMBLE	Data = 10101010
5 PREAMBLE	Data = 10101010

802.3 info	8B10B codes
6 PREAMBLE	Data = 10101010
7 SFD	Data = 10101011
8 DA	Data = DA...
~	~
LLC DATA	Data = LLC DATA...
~	~
FCS1	Data = FCS1 octet...
FCS2	Data = FCS2 octet...
FCS3	Data = FCS3 octet...
FCS4	Data = FCS4 octet...
EOP1	T (even char/no extension)
EOP2	R
Idle/Flip Disparity	I1
Idle/Disparity OK	I2
~	~
Idle/Disparity OK	I2



Link_Startup

- The link is not considered usable until Link_Startup is complete.
  - To perform Link_Startup, a station transmits the Link_Not_Available sequence and simultaneously attempts to acquire bit and transmission word synchronization from the received signal.
  - Link_Not_Available is followed by the Link_Configuration sequence.

Invalid Transmission Words

- The detection of Invalid Transmission Words is an indication that a receiver is out of synchronization.
- An invalid Transmission Word is recognized by the receiver when one of the following conditions is detected:
  - A Code Violation is detected within a Transmission Word;
  - A Special Character Alignment Error is detected. (e.g., a K28.5 character is received as an odd-numbered character, a non-K28.5 special character immediately follows a K28.5 character, a non supported special character is detected, etc.);
  - An ordered set with improper Beginning Running Disparity is received.

Running Disparity

- 8B10B code recognizes the idea of a Running Disparity (the difference between the number of 1's and 0's transmitted).
  - The sender keeps the running disparity around zero, the receiver checks the sender.
    - After powering on or exiting diagnostic mode, the transmitter assumes the negative value for its initial running disparity.
    - Upon transmission of any transmission character, the transmitter calculates a new value for its running disparity.
    - After powering on or exiting a special mode, the receiver assumes either a positive or negative initial running disparity.
    - Upon reception of any transmission character, the receiver determines whether the character is valid and calculates a new value for its RD.
    - All Idle ordered sets end with negative RD. The I2 ordered set also begins with negative running disparity.
    - The I2 ordered set may be removed or added from an encoded bit stream by an inter-station unit to compensate for differences in clock frequencies, one word at a time, without altering the beginning RD of the immediately preceding transmission word.

Synchronization States

- A receiver shall enter the Synchronization_Acquired state when it has achieved both bit and transmission word synchronization.
  - Signal_Status=OK when a receiver is in the Synchronization_Acquired state.
- The following four conditions shall cause an operational receiver to enter the Loss_of_Synchronization state.
  - Signal_Status=NOT_OK when a receiver is in the Loss_of_Synchronization state:

  1) Completion of the Loss_of_Synchronization procedure;
  2) Transition to power on;
  3) Exit from receiver reset condition;
  4) Detection of Loss_of_Signal.

Loss_of_Synchronization Procedure States

- The following five detection states are defined as part of the Loss_of_Synchronization procedure:
  - State 1: **No** invalid transmission word has been detected (the No_Invalid_Transmission_Word detection state).
  - State 2: The **first** invalid transmission word is detected (the First_Invalid_Transmission_Word detection state).
  - State 3: The **second** invalid transmission word is detected (the Second_Invalid_Transmission_Word detection state).
  - State 4: The **third** invalid transmission word is detected (the Third_Invalid_Transmission_Word detection state).
  - State 5: The **fourth** invalid transmission word is detected (the Fourth_Invalid_Transmission_Word detection state).
- A receiver in the Synchronization_Acquired state may be in any of the first four detection states listed above. A receiver in State 5 shall enter the Loss_of_Synchronization state.

Loss_of_Synchronization Procedure Transitions

- The following state transitions are defined as part of the Loss_of_Synchronization procedure:

  a) The first invalid transmission word is detected;

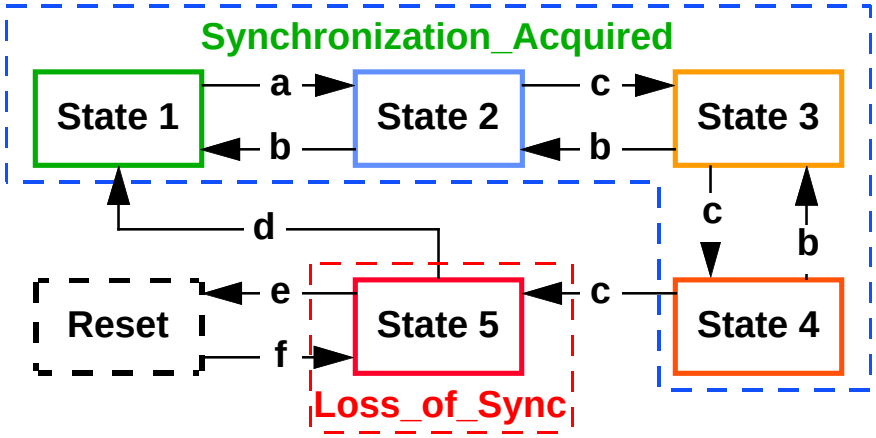
  b) An additional invalid transmission word is not detected in the next two or fewer consecutive transmission words;

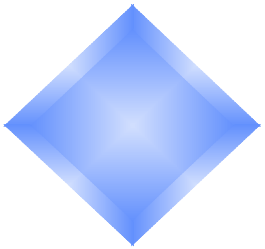
  c) An additional invalid transmission word is detected in the next two or fewer consecutive transmission words;

  d) The receiver regains synchronization;

  e) The receiver is reset;

  f) The receiver exits a previously established reset condition.





Summary

- 8B10B codes and the proposed coding structure is **efficient, robust, and flexible** enough to meet **Gigabit Ethernet PCS requirements**.
- Future activities:
  - Continue to follow direction from Task Force;
  - Revise serial link codes pending proposal acceptance.