Control character notation of ordered sets

Addressing D1.0 Comments 5, 6, 52, 44

Adee Ran, Cisco

Introduction

- The notion of ordered set goes back to early days of Ethernet
- PCS clauses of different encodings: BASE-X (8B10B), BASE-T (64B65B), and BASE-R (64B66B) share text and notations
  - Often carried over without much consideration
- Until this project, BASE-R PCSs had effectively only one type of ordered set – the sequence ordered set
  - It is explicitly denoted /Q/
- The notation /O/ is also used in PCS clauses 82, 119, 172, and 175
  - This double notation is unclear unless one studies its history
- In this project, a new ordered set is defined – the control ordered set
  - Functionally different
  - The same notation is still used
- This presentation explores the history of ordered set notation and suggests changes for clarity

Background

1.4.439 ordered set: A single special code-group, a combination of special and data code-groups, or a combination of a control character and data characters that are used to send control and status information such as remote fault and local fault status over the link. Also used by the 1000BASE-X and 10GBASE-X PCS for delineation of a packet and synchronization between the transmitter and receiver circuits at opposite ends of a link. (See IEEE Std 802.3, Clause 36, Clause 48, Clause 49, Clause 55, and Clause 82.)

Ordered sets initially appeared in Clause 36, 1000BASE-X PCS, but were quite different from the BASE-R ones. Clause 48, 10GBASE-X PCS, is similar. Both are not discussed in this presentation.

It is worth noting that these clauses defined many types of ordered sets, not just 1 or 2.

Table 46–5—Sequence ordered sets

Lane 0	Lane 1	Lane 2	Lane 3	Description
Sequence	0x00	0x00	0x00	Reserved
Sequence	0x00	0x00	0x01	Local Fault
Sequence	0x00	0x00	0x02	Remote Fault
Sequence	0x00	0x00	0x03	Link Interruption
Sequence	≥ 0x00	≥ 0x00	≥ 0x04	Reserved

NOTE—Values in Lane 1, Lane 2, and Lane 3 columns are in hexadecimal, most significant bit to least significant bit (i.e., <7:0>). The link fault signaling state diagram allows future standardization of reserved Sequence ordered sets for functions other than link fault indications

Clause 46 (RS/XGMII) uses only one type of ordered set, the sequence ordered sets, for link fault signaling and link interruption (not used in BASE-R).

Background

49.2.4.5 Ordered sets

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. Ordered sets consist of a control character followed by three data characters. Ordered sets always begin on the first octet of the XGMII. 10 Gigabit Ethernet uses one kind of ordered set: the sequence ordered set (see 46.3.4). The sequence ordered set control character is denoted /Q/. An additional ordered set, the signal ordered set, has been reserved and it begins with another control code. The 4-bit O field encodes the control code. See Table 49–1 for the mappings.

Table 49–1—Control codes

Control character	Notation	XGMII Control Code	10GBASE-R Control Code	10GBASE-R O Code	8B/10B Code[a]
Sequence ordered set	/Q/	0x9C	Encoded by block type field plus O code	0x0	K28.4
Signal ordered set[d]	/Fsig/	0x5C	Encoded by block type field plus O code	0xF	K28.2

[d]Used by INCITS T11 Fibre Channel.

Note that the signal ordered set /Fsig/ is not used in Ethernet (not defined in RS/XGMII) and arguably should not have been defined at all.

Its definition in Table 49-1 shows it is created from XGMII control code 0x5C rather than 0x9C – so it is effectively a control ordered set as defined in 802.3dt. **This contradicts the claim in comment #52 that ordered sets used by other organization are changed by this amendment.**

Background

49.2.4.10 ordered set (/O/)

The ordered set control characters (/O/) indicate the start of an ordered set. There are two kinds of ordered sets: the sequence ordered set and the signal ordered set (which is reserved). When it is necessary to designate the control character for the sequence ordered set specifically, /Q/ will be used. /O/ is only valid on the first octet of the XGMII. Receipt of an /O/ on any other octet of TXD indicates an error. Block type field values implicitly encode an /O/ as the first or fifth character of the block. The 4-bit O code encodes the specific /O/ character for the ordered set.

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and shall delete only one of the two. Only Idles may be inserted for clock compensation. Signal ordered sets are not deleted for clock compensation.

Clause 49 has 3 notations for ordered sets: /Q/, /Fsig/, and /O/ !

The designation of /O/ as standing for either /Q/ or /Fsig/ is a poor choice; this notation is only used once in the clause (in the definition of T_BLOCK_TYPE) and could have been replaced by “either /Q/ or /Fsig/”.

The ordered set notation and description text was carried over to multiple PCS clauses, including BASE-T, which never uses any ordered set other than /Q/; /Fsig/ is noted in later clauses as “reserved”.

Background

Table 82–1—Control codes

Control character	Notation	XLGMII/ CGMII control code	40/100GBASE-R O code[a]	40GBASE-R and 100GBASE-R control code
idle	/I/	0x07		0x00
LPI	/LI/	0x06		0x06
start	/S/	0xFB		Encoded by block type field
terminate	/T/	0xFD		Encoded by block type field
error	/E/	0xFE		0x1E
Sequence ordered set	/Q/	0x9C	0x0	Encoded by block type 0x4B plus O code, control codes are set to 0x00

[a] INCITS T11 Fibre Channel uses O code 0xF for the signal ordered set. OIF uses O code 0x5 for the FlexE [B59] ordered set. ITU-T uses O code 0xC for the G.8312 [B47] MTN path ordered set.

82.2.3.9 ordered set (/O/)

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. Ordered sets consist of a control character followed by three data characters followed by four zero data characters on the XLGMII/CGMII. Ordered sets always begin on the first octet of the XLGMII/CGMII. 40 Gigabit and 100 Gigabit Ethernet use one kind of ordered set: the sequence ordered set (see 81.3.4). The ordered set control character /Q/ indicates the start of an ordered set. The block type field plus the O code encode the specific control character for the ordered set. See Table 82–1 for the mapping.

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and only one of the two ordered sets may be deleted. Only idle control characters may be inserted for clock compensation.

Clause 82 PCS has only /Q/ and does not list /Fsig/ anymore. It mentions the signal ordered set (and other ordered sets not used in Ethernet) in the table footnote which seems to be informative.

The text still has the duplicate notation /O/ which is now meaningless and confusing!
All later BASE-R PCS clauses (119, 172, 175) point to 82.2.3.9.

More on non-Ethernet ordered sets

- Comment #52 states that encoding control ordered sets using the same block type as sequence ordered sets (0x4B) alters the specifications of other organizations.
- The non-Ethernet ordered sets mentioned in Table 82–1 footnote a) are not listed as having a different xMII control code. However, based on Clause 49 (and other clauses that kept /Fsig/) it is likely that they use 0x5C rather than 0x9C, which is identical to the control ordered sets defined in this amendment.
- Therefore, the encoding of both types of ordered sets by the same block type 0x4B is consistent with previous use
- Comment #52 can be AIP with a change of the anchor of the footnote to be the “0x1-0xF” cell in the “Control ordered set” row - to signify that the control ordered set definition encompasses these non-Ethernet ordered sets.
    - Alternatively, reject the comment.

Background - summary

- The text of Clause 82 inherits the double notation /O/ and /Q/ from older PCS clauses, although it has only one ordered set (/Q/), so /O/ and /Q/ are identical.
- In all PCS clauses, the notation /O/ is used only once after it is defined, in the definition of T_TYPE (or earlier T_BLOCK_TYPE). In Clause 82, this definition includes “A valid /O/ is any character with a value for O code in Table 82–1” – which means exactly /Q/.
- This is very confusing (especially in first reading).
  - Ideally it should have been fixed in maintenance.

Where are we now…

Change 82.2.3.9 as follows:

82.2.3.9 Oordered set (/O/)

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. There are two types of ordered sets: the sequence ordered set (see 81.3.4) and the control ordered set (see 195.4.1.2). Sequence oOrdered sets consist of ~~a~~ the /Q/ control character followed by three data characters followed by four zero data characters on the XLGMII/CGMII. Control ordered sets consist of the /C/ control character followed by seven data characters. The notation /O/ is used to represent either the /Q/ or /C/ control characters. All Oordered sets always begin on the first octet of the XLGMII/CGMII. ~~40 Gigabit and 100 Gigabit Ethernet use one kind of ordered set: the sequence ordered set (see 81.3.4). The ordered set control character /Q/ indicates the start of an ordered set.~~ The fifth octet of an ordered set (D_4) contains the O code value in bits<3:0>. The block type field plus the O code encode the specific control character for the ordered set. See Table 82–1 for the mapping.

Support for the sequence ordered set is required. Support for the control ordered set is optional. If the control ordered set is supported, it is indicated by the assertion of the control_ordered_set_ability variable.

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and only one of the two ordered sets may be deleted. Only idle control characters may be inserted for clock compensation. Control ordered sets are never deleted by the PCS.

802.3dt D1.0 rewrites 82.2.3.9 significantly and adds the /C/ notation.
Instead of perpetuating the bad content in the original Clause 82, there is opportunity to fix it.

Proposed changes (1)

82.2.3.9 ordered set (~~/O/~~/Q/ and /R/)

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. There are two types of ordered sets: the sequence ordered set (see 81.3.4) and the control ordered set (see 195.4.1.2). Sequence o~~O~~rdered sets consist of ~~a~~ the /Q/ control character followed by three data characters followed by four zero data characters on the XLGMII/CGMII. Control ordered sets consist of the ~~/C/~~/R/ control character followed by seven data characters. ~~The notation /O/ is used to represent either the /Q/ or /C/ control characters.~~ All ~~O~~ordered sets always begin on the first octet of the XLGMII/CGMII. ~~40 Gigabit and 100 Gigabit Ethernet use one kind of ordered set: the sequence ordered set (see 81.3.4). The ordered set control character /Q/ indicates the start of an ordered set.~~ The fifth octet of an ordered set (D4) contains the O code value in bits<3:0>. The block type field plus the O code encode the specific control character for the ordered set. See Table 82–1 for the mapping.

Support for the sequence ordered set is required. Support for the control ordered set is optional. If the control ordered set is supported, it is indicated by the assertion of the control_ordered_set_ability variable.

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and only one of the two ordered sets may be deleted. Only idle control characters may be inserted for clock compensation. Control ordered sets are never deleted by the PCS.

Note: Comment #44 suggests changing /C/ to /R/ to avoid confusion with the “C” output of the T_TYPE function (see next slide). The proposed change in this slide assumes that #44 is accepted.

Proposed changes (2)

Bring in the full definition of T_TYPE in 82.2.19.2.3 and change it as follows:

T_TYPE~~=~~(tx_raw<71:0>)

This function classifies each 72-bit tx_raw vector as belonging to one of the following types depending on its contents. The classification results are returned via the t_block_type variable.

Values:

C; The vector contains one of the following:

a) Eight valid control characters other than ~~/O/~~/Q/, /S/, /T/, /LI/, and /E/;

b) One valid ordered set.

LI; For EEE capability, this vector contains eight /LI/ characters.

S; The vector contains an /S/ in its first character, and all characters following the /S/ are data characters.

T; The vector contains a /T/ in one of its characters, all characters before the /T/ are data characters, and all characters following the /T/ are valid control characters other than ~~/O/~~/Q/, /S/ and /T/.

D; The vector contains eight data characters.

E; The vector does not meet the criteria for any other value.

A tx_raw character is a control character if its associated TXC bit is asserted. A valid control character is one containing an XLGMII/CGMII control code specified in Table 82–1. A valid ordered set consists of a valid ~~/O/~~/Q/ character in the first character and data characters in the seven characters following the ~~/O/~~/Q/. ~~A valid /O/ is any character with a value for O code in Table 82–1.~~

NOTE—A PCS that does not support EEE classifies vectors containing one or more /LI/ control characters as type E. A PCS that does not support the control ordered set classifies vectors containing the ~~/C/~~/R/ control character as type E.

Note that this function is only used by the state-diagram encoder; thus, it should not be changed to recognize the new /C/ character, which is only supported when the PCS uses the stateless encoder. Thus, the original /O/ meant /Q/ in this context. However, **with the addition of /C/, using /O/ here changes the definition of the state diagram** – which is out of scope for this project!

Proposed changes (3)

Bring in the full definition of T_TYPE in 119.2.6.2.3 and change it as follows:

T_TYPE~~=~~(tx_raw<71:0>)

This function classifies each 72-bit tx_raw vector as belonging to one of the following types depending on its contents. The classification results are returned via the t_block_type variable.

Values:

C; The vector contains one of the following:

a) Eight valid control characters other than ~~/O/~~/Q/, /S/, /T/, /LI/, and /E/;

b) One valid ordered set.

LI; For EEE capability, this vector contains eight /LI/ characters.

S; The vector contains an /S/ in its first character, and all characters following the /S/ are data characters.

T; The vector contains a /T/ in one of its characters, all characters before the /T/ are data characters, and all characters following the /T/ are valid control characters other than ~~/O/~~/Q/, /S/ and /T/.

D; The vector contains eight data characters.

E; The vector does not meet the criteria for any other value.

A tx_raw character is a control character if its associated TXC bit is asserted. A valid control character is one containing an 200GMII/400GMII control code specified in Table 82–1. A valid ordered set consists of a valid ~~/O/~~/Q/ character in the first character and data characters in the seven characters following the ~~/O/~~/Q/. ~~A valid /O/ is any character with a value for O code in Table 82–1.~~

NOTE—A PCS that does not support EEE classifies vectors containing one or more /LI/ control characters as type E. A PCS that does not support the control ordered set classifies vectors containing the ~~/C/~~/R/ control character as type E.

This definition has the exact same text as in 82.2.19.2.3, so the same changes are required (or alternatively, replace it with a reference to 82.2.19.2.3).

Newer PCS clauses refer back to 119.2.6.2.3 (see 172.2.6.2.3 in 802.3df-2024 and 175.2.6.2.3 in 802.3dj) so no further changes are needed.

Summary

- The history of the /O/ notation was presented.
  - This notation became unnecessary (and confusing) in Clause 82, and was carried over to Clause 119.
- Since 802.3dt is making changes in 82.2.3.9 and 82.2.19.2.3 anyway, it would be good to improve the text and get rid of the confusing /O/ notation.
- With the addition of the /C/ ordered set, /O/ becomes “/Q/ or /C/”, and this changes the definition of the T_TYPE function, which is out of the project’s scope.
- The proposed changes in slides 9-11 resolves both issues.

That’s all!

Questions?