

Task Group	Congestion Management
Revision	1.0
Author	Hugh Barrass (Cisco)

Definition for new PAUSE function

Project Headline

Definition of a new PAUSE function that can halt traffic according to priority tag while allowing traffic at other priority levels to continue.

Modification History

Rev	Date	Originator	Comment
1.0	5/30/2007	Hugh Barrass	Initial Submitted Version

Table of Contents

1. INTRODUCTION	3
2. FRAME FORMAT DEFINITION.....	3
2.1 BASIC FRAME FORMAT.....	3
2.2 NEW CODEBLOCK DEFINITION.....	4
2.3 INTERPRETATION OF 802.3X PAUSE FRAMES	4
3. CLASS BASED PAUSE OPERATION	5
3.1 PAUSE DESCRIPTION (31B.1)	5
3.2 PARAMETER SEMANTICS (31B.2).....	5
3.3 TRANSMIT OPERATION (31B.3.1).....	5
3.4 RECEIVE OPERATION (31B.3.3)	5
3.5 STATUS INDICATION (31B.3.5).....	5
3.6 TIMING CONSIDERATIONS (31B.3.7)	5
4. MANAGEMENT	6
4.1 AMACCONTROLFUNCTIONSUPPORTED (30.3.3.2)	6
4.2 UNCHANGED OBJECTS	6
4.3 ACBFPCPAUSEFRAMESTRANSMITTED_0..7 (NEW OBJECTS)	6
4.4 ACBFPCPAUSEFRAMESRECEIVED_0..7 (NEW OBJECTS)	6
4.5 ACBFPCPAUSETRANSITIONS_0..7 (NEW OBJECTS)	6
5. HIGHER LAYER FUNCTION.....	6

1. Introduction

This document contains a proposal for a new MAC control frame for the purpose of per-priority or class-based PAUSE.

The function is very closely related to the PAUSE function (802.3x) defined in IEEE 802.3 Clause 31, Annex 31A and Annex 31B.

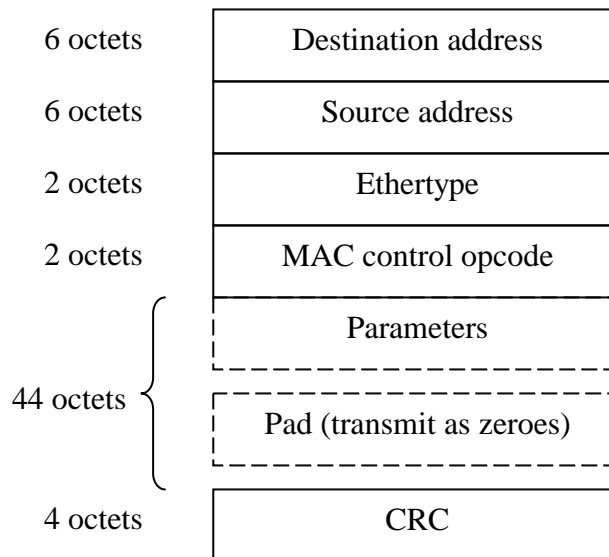
The frame semantics support a MAC control function similar to that defined in presentation “barrass_2_0505” (given to the 802.3ar Task Force during the May 2005 session).

2. Frame format definition

The basic format of a MAC control frame is defined in IEEE 802.3, Clause 31. The opcodes used are defined in Annex 31A and the format of an 802.3x PAUSE frame is defined in Annex 31B. The new PAUSE function is referred to as Class Based Flow Control (CBFC).

2.1 Basic frame format

The MAC control frame format is described in subclause 31.4.1 of IEEE 802.3 with the following diagram:



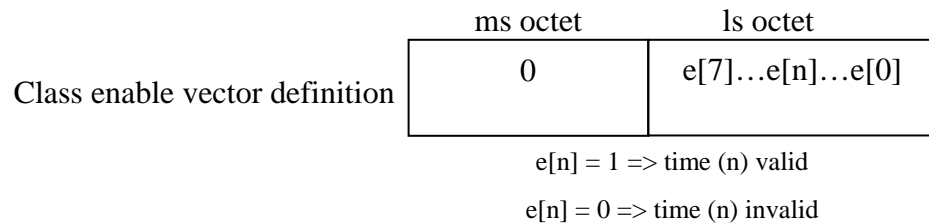
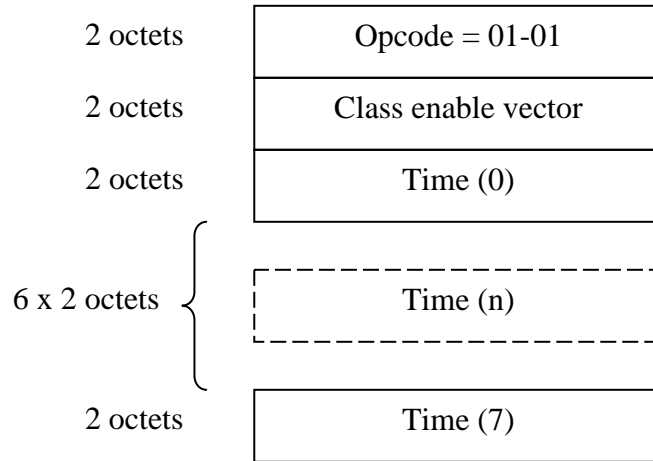
The fields contain the following values:

- Destination address: 01-80-C2-00-00-01
- Source address: sending station address
- Ethertype: 88-08

Note that MAC control frames are never tagged or envelope frames.

2.2 New codeblock definition

The CBFC PAUSE frame is defined with a new opcode, the following semantics are used:



Time (n) is defined as the pause timer for class n, defined in the same manner as in subclause 31B.2 of 802.3

2.3 Interpretation of 802.3x PAUSE frames

After the use of CBFC has been negotiated, there shall be no use of 802.3x format PAUSE frames. If an 802.3x format PAUSE frame is received, the receiving MAC may ignore the frame or pause classes in an undefined manner.

3. Class based PAUSE operation

The class based PAUSE function is an extension to the MAC control PAUSE function defined in IEEE 802.3, Annex 31B. This section describes the additions to Annex 31B to support the class based PAUSE function. For each subsection, the corresponding subclause of Annex 31B is shown.

3.1 PAUSE description (31B.1)

The class based PAUSE function includes a MA_CONTROL.request primitive specifying:

- a) The globally assigned 48-bit multicast address 01-80-C2-00-00-01;
- b) The CBFC PAUSE opcode;
- c) A request_operand indicating the set of classes addressed and lengths of time for which it wishes to inhibit data frame transmission of the corresponding classes. (See section 3)

3.2 Parameter semantics (31B.2)

The pause_time(n) operands are defined in an identical manner to the pause_time operand of Annex 31B.2.

3.3 Transmit operation (31B.3.1)

The response to a MA_CONTROL.request is similar to the basic PAUSE function, with the appropriate set of format of operands for CBFC PAUSE.

The MAC control sublayer does not interfere with the transmission of data frames (initiated by MA_DATA.request) as part off the CBFC PAUSE function.

3.4 Receive operation (31B.3.3)

Upon receipt of a valid CBFC PAUSE frame, the MAC control sublayer starts 1 to 8 separate counters (depending on the class enable vector). These timers operate in an identical manner to the single pause timer of 31B.3.3.

3.5 Status indication (31B.3.5)

The MA_CONTROL.indication primitive contains a vector of “paused / not paused” indications corresponding to the state for all 8 classes.

3.6 Timing considerations (31B.3.7)

All of the timing considerations for CBFC PAUSE are identical to those defined for basic PAUSE operation.

Note that for whatever classes are paused, the traffic for those classes must stop within the allowed time. In terms of absorption buffers: if you have entirely separate buffers per class then each one needs to be large enough to absorb the slack; however, if you have perfectly shared buffers then the total shared buffer needs only to be large enough to absorb the slack. In both cases the slack is defined in terms of time at the egress transmit rate.

4. Management

The following management attributes are required to support the CBFC PAUSE function (in addition to those already defined in IEEE 802.3 Clause 30):

4.1 aMACControlFunctionsSupported (30.3.3.2)

The definition of aMACControlFunctionsSupported requires an additional SEQUENCE definition:

CBFC_PAUSE CBFC PAUSE command implemented

Note that support of the CBFC PAUSE command implies support for the original PAUSE command.

4.2 Unchanged objects

Objects aMACControlFramesTransmitted; aMACControlFramesReceived; aUnsupportedOpCodesReceived; aPAUSELinkDelayAllowance; aPAUSEMACCtrlFramesTransmitted; aPAUSEMACCtrlFramesReceived are all supported as defined without change.

4.3 aCBFCPAUSEFramesTransmitted_0..7 (new objects)

This is a set of 8 objects recording the number of CBFC PAUSE frames that have been transmitted with the same definition as aPAUSEMACCtrlFramesTransmitted (30.3.4.2).

4.4 aCBFCPAUSEFramesReceived_0..7 (new objects)

This is a set of 8 objects recording the number of CBFC PAUSE frames that have been received with the same definition as aPAUSEMACCtrlFramesReceived (30.3.4.3).

4.5 aCBFCPAUSETransitions_0..7 (new objects)

This is a set of 8 objects recording the number of times that each of the status indications, MA_CONTROL.indication, has transitioned from “not_paused” to “paused.”

5. Higher Layer function

The CBFC PAUSE function is supported by a modification to the scheduling of frames for transmission defined in clause 7 of 802.1D. Data frames from each class of traffic may be scheduled for transmission if, and only if the MA_CONTROL.indication status for that class of traffic is “not_paused” for the destination port.

Note that the implementation of fewer classes of traffic may be supported by combining two or more classes into a single queue. In this case data frames may only be scheduled for transmission if, and only if the MA_CONTROL.indication status for all of the classes of traffic sharing the queue is “not_paused” for the destination port. In this case, an implementation may be optimized to contain fewer than 8 pause timers without changing the externally visible behavior of the system.