

Unconfirmed Minutes of the IEEE 802.3ad Link Aggregation Meeting

July 7-8th, 1998

La Jolla, CA

Mr. Geoff Thompson called the meeting to order at 1:30PM on July 7th. He reported that the Link Aggregation (LA) activity was now an official 802.3 Task Force. As his first order of business, he opened nominations for Chairman of the Task Force.

Rich Seifert nominated Steve Haddock (Extreme Networks, shaddock@extremenetworks.com), seconded by Bill Quackenbush.

No other nominations were made.

Steve was elected Chairman by unanimous voice vote.

Immediately following the election, Mr. Geoff Thompson turned the meeting over to Mr. Steve Haddock. Steve indicated that he would postpone selection of editor and recording secretary until end of the meeting on Wednesday. He called for someone to take minutes for this meeting. Jeff Lynch agreed to take the minutes.

Steve presented the following agenda for the meeting:

- Introductory remarks and introductions
- Report from Merrimack TSG meeting
- Work Plan/Timetable
- Presentations
- Discussion on Link Aggregation Control
- Call for patents
- Plans for next meeting
- Minutes from the Merrimack meeting

After introductions were made, Steve:

- Reviewed the E-mail Reflector & Web site
- Gave a brief summary of the Merrimack interim study group meeting
- Proposed an 802.3ad schedule based on historical experience from other 802.3 activities. Some key dates from the timeline were:
 - Proposal cutoff, November 1998
 - First Draft, January 1999
 - Standard, March 2000

A brief discussion took place concerning whether there needed to be time allocated in the proposed schedule for a re-circulation of the sponsor ballot. It was noted that there was adequate time in the schedule to do one, if one was needed. No other concerns were raised on the proposed schedule.

A copy of Steve's foils can be found on the IEEE 802.3 Link Aggregation Web page: <http://grouper.ieee.org/groups/802/3/ad/public/july98/index.html>

The following presentations were made:

Tony Jeffree & Alan Chambers
Walter Thirion

Link Aggregation Control Protocol
Link Aggregation Operations

Luc Pariseau	Link Aggregation: Support for More Limitations
Jeff Lynch, Loren Larsen, Arush Kumar & Mike Siegel	Flush Scenarios and Requirements
Jeff Lynch, Loren Larsen, Arush Kumar & Mike Siegel	Frame Proposal
Rich Seifert	IEEE 802.3 Service Interface Issues

A PDF version of each of these presentations can be found on the IEEE 802.3 Link Aggregation Web page: <http://grouper.ieee.org/groups/802/3/ad/public/july98/index.html>

Link Aggregation Control Presentation

Tony Jeffree presented an overview of a consolidated state-exchange protocol based upon his presentation and the Finn/Wakerley/Fine presentation from the interim Merrimack meeting. The protocol description was divided into the following states: RX, Desirable, Nervous, Crowd, Match, Aggregate, MUX, and TX. Tony summarized each of these states and their high-level state machines. After Tony finished describing the protocol, Alan Chambers reviewed two example protocol scenarios.

The presentation introduced new terminology and addressed an expanded set of objectives. As such, there were quite a few questions and much discussion was generated.

Some of the discussion points were:

- The ability to detect "Crowds" on shared media links
 - The NIC should know it is in half duplex mode.
- Whether to accommodate hardware that is non-standard
 - Such as hardware that can not detect the physical link configuration
- The perceived complexity of the overall protocol and whether several of the objectives and boundary conditions needed to be satisfied.
- Distinction of desirable and automatic. It was observed that there should be at least one shot to aggregate even if both ends were set to automatic. (i.e., Possibly an initial probing before going quiet)
 - There was a suggestion to change the term "desirable" to "advertising."
- The ability to use the "Need to Know" indication from Tony's previous presentation to help speed up the initial exchange.

Link Aggregation Operations Presentation

Walter Thirion briefly reviewed the concept of logical and physical MACs and then made the following general assertions:

- Each logical MAC is initially associated with at least one Physical MAC.
- Logical MAC's address is persistent for the duration of the group existence.
- Aggregation Control sublayer is associated with one and only one LAG.
- Decision of when to add/delete Physical MACs from a group is beyond the scope of the standard.

- Moving Physical MAC's between groups is beyond the scope of the standard.

Walter also suggested a command/response-based protocol. In the ensuing discussion, several comments and questions compared and contrasted this presentation with the Jeffree/Alan presentation. Walt emphasized his main objective was not to focus attention on the command/response versus state-based operation of the protocol but rather to highlight the need for a simple standard with an appropriate scope (i.e., the Link Aggregation sublayer is per group not across groups and the number of commands sent back and forth should remain small).

He indicated his main concern about the Jeffree/Alan presentation was the statement about searching other groups to determine if a link should be included in a LAG.

IEEE 802.3 Link Aggregation: Support for More Limitations Presentation

Luc Pariseau pointed out that dynamically changing the "Port Capability" to overcome "system specific" configuration limitations may not always result in an optimal link aggregation grouping between two systems. In some cases, it may not converge at all. Luc concluded the presentation by asking if this was a problem that we wanted the LA protocol to address?

Norm Finn suggested that to solve this the LA protocol must either advertise all possible link combinations or select a simplifying algorithm such as "aggregatable links can only be reduced in number during capability exchange." This always converges to a solution but not necessarily the most optimal one.

Proposals on how to address this were requested for future meetings.

There was a significant interest from participants in attending the 802.1 Plenary to participate in the discussion on Flow Control extensions. This discussion was scheduled for 8:30 - 10:00 on Wednesday. To accommodate this, Steve adjourned the 802.3ad meeting until Wednesday at 10:00.

Meeting adjourned at 5:50 PM.

LINK AGGREGATION (802.3ad)

Wednesday 7/8/98

Steve called the meeting to order at 10:10, Wednesday 7/8/98, and handed out the minutes from the interim meeting in Merrimack, NH.

First thing on the agenda was finishing up the presentations.

Flush Scenarios and Requirements Presentation

Jeff Lynch highlighted that any time an active flow is moved from one physical link to another, a flush operation (or some other mechanism such as a timeout) is required to ensure that frames are not delivered out of order. He summarized a flush operation and several potential flush scenarios. He concluded his presentation with the following operational requirements for the flush command:

- Flush command must be optional to send.
- Response to the flush command is mandatory.
- Flush command can be sent anytime.
- Flush command is processed on a per link basis.

The majority of the discussion was on the operational requirements.

There was a discussion of whether the flush response must be returned on the same physical link as it was transmitted. The pros and cons were debated.

There was vigorous discussion on the span of the flush function. Does the flush operation include entry to the transmit queue to exit from the receive queue? A debate about whether queues are within the scope of 802.3 and whether the LA standard needed to acknowledge queues occurred. Norm Finn strongly urged the Task Force to solve the queue-to-queue problem. Steve Haddock suggested that the flush operation be between the "distributor" and "collector".

There was also a discussion of whether the flush transmission capability should be mandatory or if it was adequate to leave it as optional to send.

Link Aggregation Frame Proposal Presentation

Jeff Lynch explored the question of what mechanism should be used to convey Link Aggregation messages. From the perspective that the mechanism chosen should be consistent with previous 802.3 extensions and positioned for eventual incorporation into hardware, the presentation suggests that Link Aggregation should follow the precedent of the flow control Pause message and use MAC Control frames.

The pros and cons of using the MAC Control frame were debated.

Some expressed concern that the 64 byte MAC Control frame may not be long enough, while others felt it should have ample payload.

A question was raised whether existing MAC chips would pass the MAC Control Link Aggregation frames. Some thought that the current MAC Control state machine specifies that frames received with unknown op codes are bit bucketed. If vendors have implemented this, then we may need to choose something other than MAC control. One possibility that was mentioned is to define a new MAC frame that is different from the MAC Control (new code point rather than 88-08).

IEEE 802.3 Service Interface Issues

Rich Seifert highlighted the problem that the 802.3 MAC interface definition does not currently allow the passing of the source destination address to the MAC from the MAC client. He indicated that this was not a new problem. Since it is expected that the LA standard will need to pass the source address to the physical MACs, 802.3 needs to address this problem.

Rich suggested the following possible changes:

- Change MA_DATA.request (and possibly MA_CONTROL.request) to allow the client to provide a source address
- Change MA_DATA.request and MA_DATA.indication to allow FCS to be passed to, and provided by, the client
- Provide a new service interface to allow MAC Control frames to be passed to a Control Client

Review of Patent Policy

Mr. Geoff Thompson reviewed the IEEE patent policy and requested that all holders of patents related to work of the study group should become familiar with the policy and submit letters if appropriate. He also informed the group that it was the policy of 802.3 to also request letters for patent applications. A description of the policy and example letters can be found at the IEEE web site (<http://grouper.ieee.org/groups/802/3/patent.html>).

Interim Meeting

An interim meeting hosted by Walter Thirion from JATO Technologies is scheduled for August 31 - September 4th at the Marriott at the Capital hotel in Austin. 802.3ad will meet September 1st and 2nd. A meeting notice containing the logistical information for the meeting will be posted to the Link Aggregation Web page.

A tentative count of expected attendees was taken. 51 individuals indicated that they would attend the interim meeting.

Selection of Chief Editor for the 802.3ad Specification

Steve asked for nominations for editor.

Tom Dineen nominated Rich Seifert, seconded by Walter Thirion.

Mick Seaman nominated Tony Jeffree, seconded by David Lanie.

Tony Jeffree (Independent consultant sponsored by 3COM) and Rich Seifert (Independent consultant sponsored by ICUBE) discussed their qualifications.

Vote:	Tony Jeffree	25
	Rich Seifert	45

Steve declared Rich Seifert the editor of 802.3ad.

Approval of Merrimack Minutes

The minutes from the interim meeting held in Merrimack, NH were approved (without amendment) by voice vote.

Call for Secretary.

Steve asked if there was anyone interested in being recording secretary. There were no takers. Steve indicated that he will continue to draft a volunteer at each meeting.

General Discussion

Steve opened the meeting to general discussion concerning the various proposals that had been presented.

A question was asked regarding the relationship of the System ID and capability groups.

Norm Finn explained that a capability group was a system specific ID (label) identifying potential links that may be grouped.

The term "KEY" was suggested as an alternative to "capability" group.

When asked to recap his presentation, Walt Thirion indicated that he is Interested in restricting the scope of the standard. He is looking for the standard to address a set of primitives that control adding a link, deleting a link, etc., but not necessary how a link moves among different groups.

This led to a long discussion on supported models and the relationship of logical MACs to physical MACs/physical ports.

Two schools of thought still exist concerning the scope of the proposed Link Aggregation standard:

1. The standard should define the operation of a single client model (single aggregator) and leave how it is extended to multiple aggregators/groups to implementation specific details
2. The standard must explicitly define the operation of the multiple aggregator model; otherwise interoperability problems will arise.

The meeting broke for lunch at 1:30. Since a significant number of participants wanted to participate in the Call For Interest meeting on Flow Control from 2:00 - 4:00 PM, Steve indicated that the LA meeting would start again at 4:00.

After resuming at 4:00 PM, the group concluded that rather than continue with the general discussion, we should break and allow individuals a chance to review the presentations in detail and formulate proposals for the next meeting.

With no further business to conduct and without objection, the 802.3ad meeting was adjourned at 4:20PM.