Link Aggregation Control

Frame Format and Syntax

IEEE 802.3ad Interim Mtg.
Austin, TX
Rich Seifert
1 Sept 98

I-Cube

Networks and Communications Cons

Goals and Objectives

- Correctness
  - Restrict Link Aggregation Control frames to the link in question
- Backwards Compatibility
  - Devices implementing 802.1D/Q (BPDUs)
  - Devices implementing 802.3x (MAC Control/Flow Control)
- Ease of Implementation
  - Hardware parsing
- Flexibility, Expandability
- Ease of Standards Specification
  - Use existing mechanisms where possible
  - Don’t invent new protocols unless there is a real need

Protocol Requirements

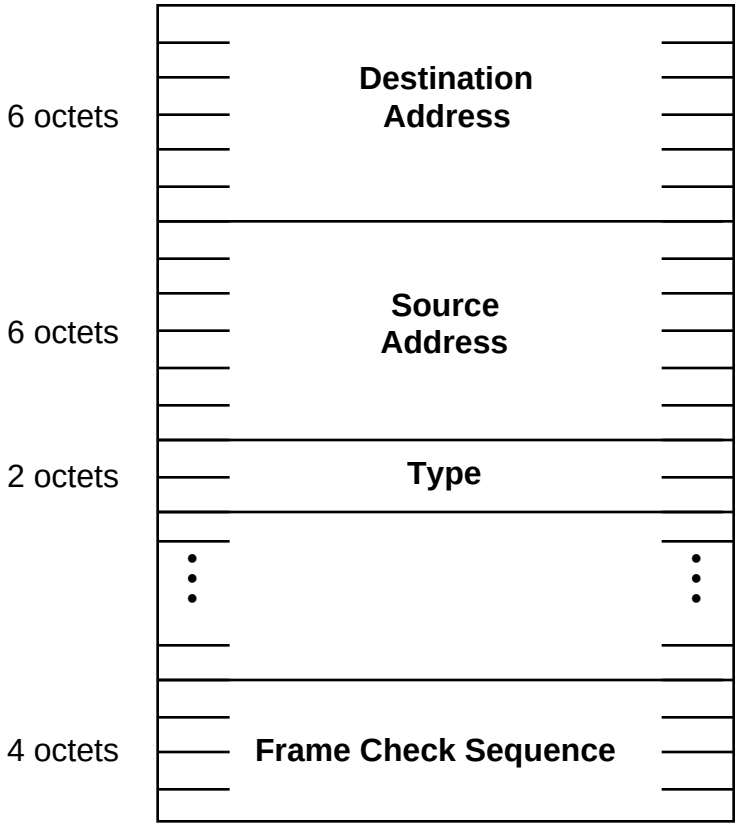
Restrict Link Aggregation Control frames to the link

- Don’t propagate through repeaters
  - Link must be Full Duplex
  - Auto-Negotiate if possible (A repeater will never negotiate to full-duplex operation)
  - Manual configurations force the user to accept responsibility (There are bigger problems than LAgC if a full duplex device is connected to a repeater!)

Protocol Requirements (cont)

- Don't propagate through 802.1D/Q bridges
  - Use DA in reserved multicast space
    01-80-C2-00-00-00 through 01-80-C2-00-00-10
  - Alternative: Use the unicast address of the port
    - Implies some means of learning the address
    - Unnecessary complexity
- Don't propagate through non-standard bridges
  - Non-standard device must have full-duplex capability (rare)
  - Buffered Distributors
    "Dumb" Bridges (no Spanning Tree support)
  - Can use Auto-Negotiation to determine capability (if supported)
  - This is an existing problem for 802.3x if manually configured
  - Can solve at higher layer (within the LAgC protocol), if necessary

Frame Format



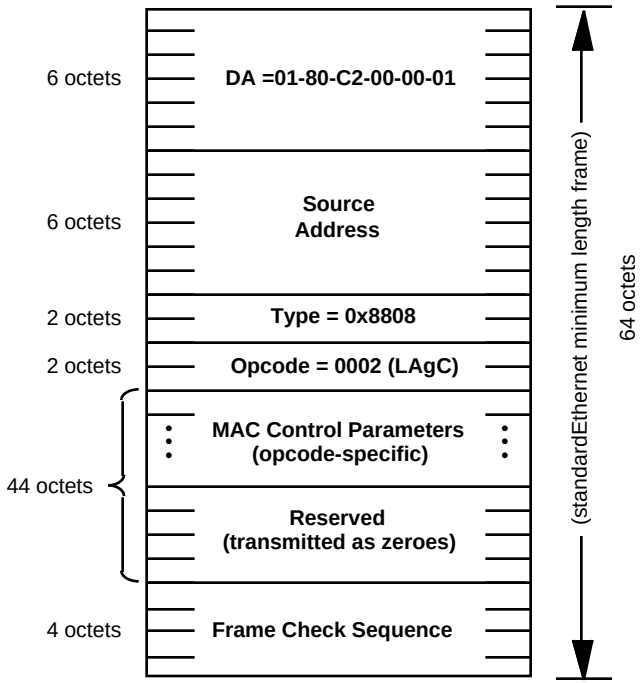
Can Multiplex (identify LAgC) based on DA and/or Type Field

Hardware Parsing

- 802.3x/PAUSE-compatible devices must already parse the DA and Type field in hardware
- Re-use this capability as much as possible
- Easiest to parse in received order
  - DA first, then Type field (and later fields, if necessary)
  - While 802.3x “architecturally” inspects the Type field first, this is a specification artifact, and not typically done in practice
- Hooks are already in many implementations for extensions to MAC Control/PAUSE

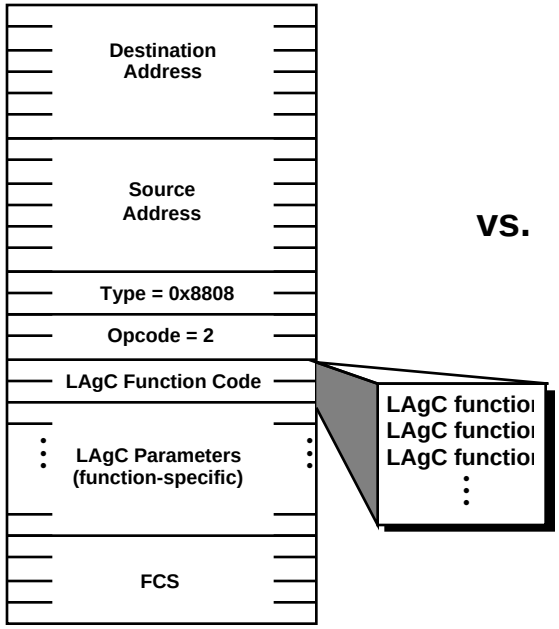
LAgC Format Proposal

- Redefine the PAUSE reserved multicast ID to mean “link restricted” MAC Control client)
  - Same address for PAUSE and LAgC saves additional hardware compare logic vs. assigning a new address
- Use MAC Control Type for LAgC
  - Saves additional hardware compare logic vs. assigning a new Type field
- Use fixed length frames
  - Minimizes hardware complexity
  - Predictable LAgC storage for hardware
- Make LAgC protocol a MAC Control client
  - Specification structure already exists
  - Multiplex from the opcode space
  - Room for 44 bytes of parametric info
  - Can use >1 frame if more is needed (unlikely)

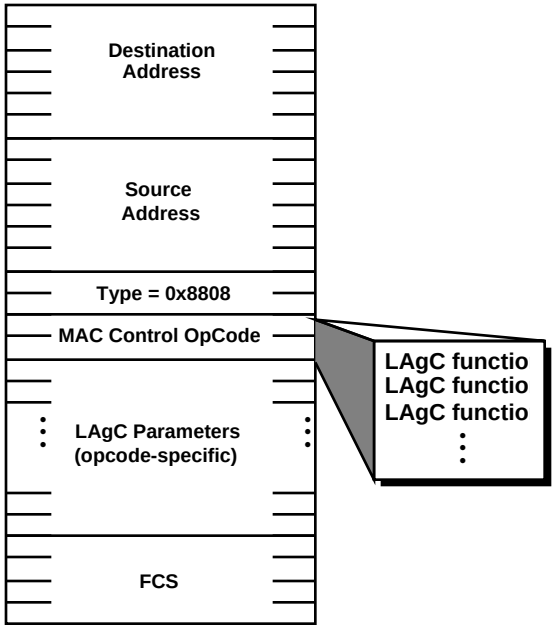


Use a Single MAC Control OpCode for all LAgC functions

- Single Hardware decode
- Can decode LAgC functions in software
- Much more flexible, expandable



vs.



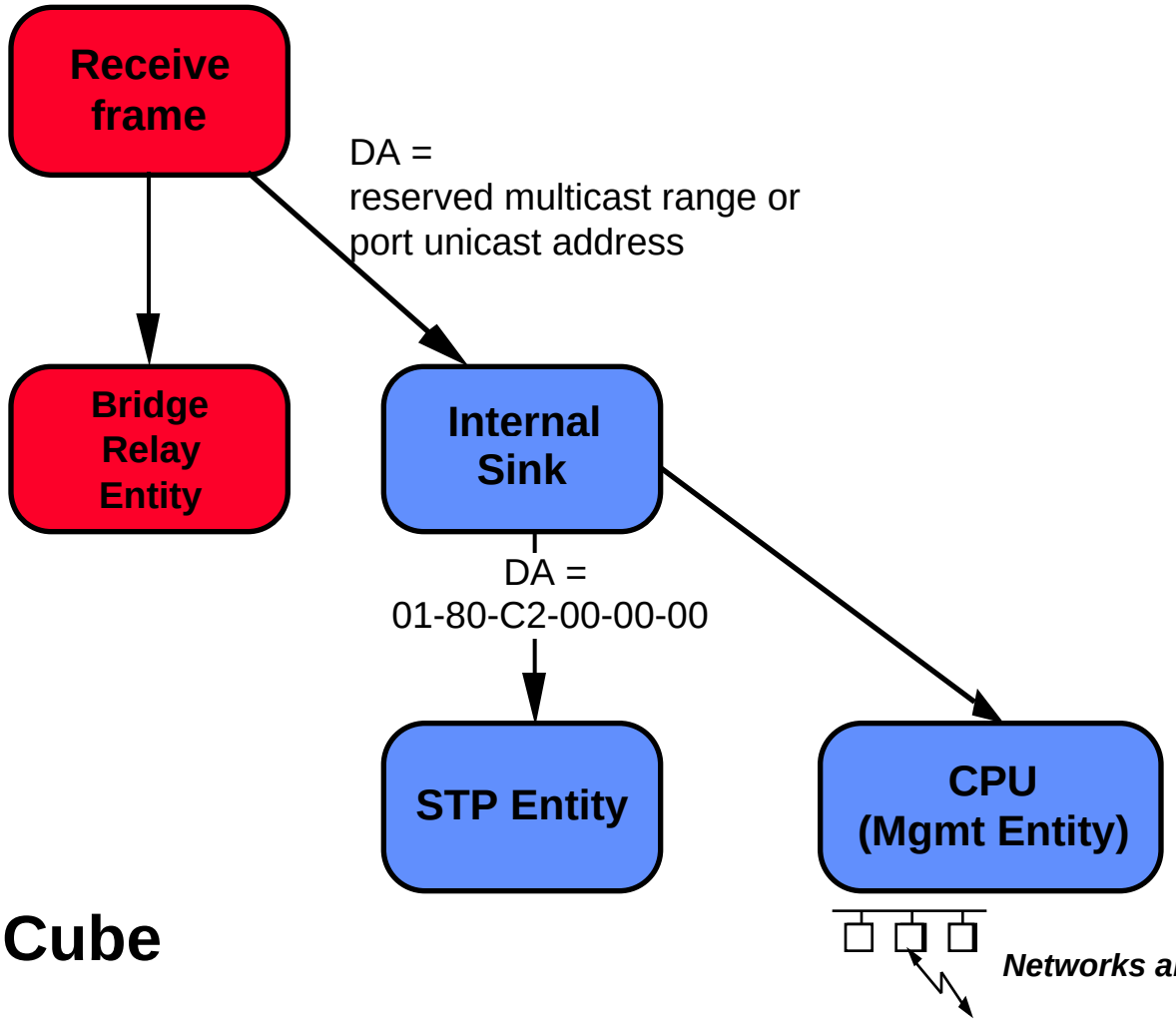
Backwards Compatibility

- Any device implementing either 802.1D, 802.3x, or both will properly handle frames
  - LAgC frames are never forwarded through a bridge
  - 802.3x will report “Unrecognized Opcode”
- Non-standard devices will likely flood LAgC frames
  - Avoid through proper configuration (manual or Auto-Negotiated)
  - Handle boundary cases within the LAgC protocol (software)

802.1D Compatibility \ 802.3x Compatibility	Y	N
Y	LAgC frames are discarded (802.1D reserved address range and 802.3x Unrecognized Opcode)	LAgC frames are discarded (802.1D reserved address range)
N	LAgC frames are discarded (802.3x Unrecognized Opcode)	Non-Std device Will likely flood (unspecified)

Parsing

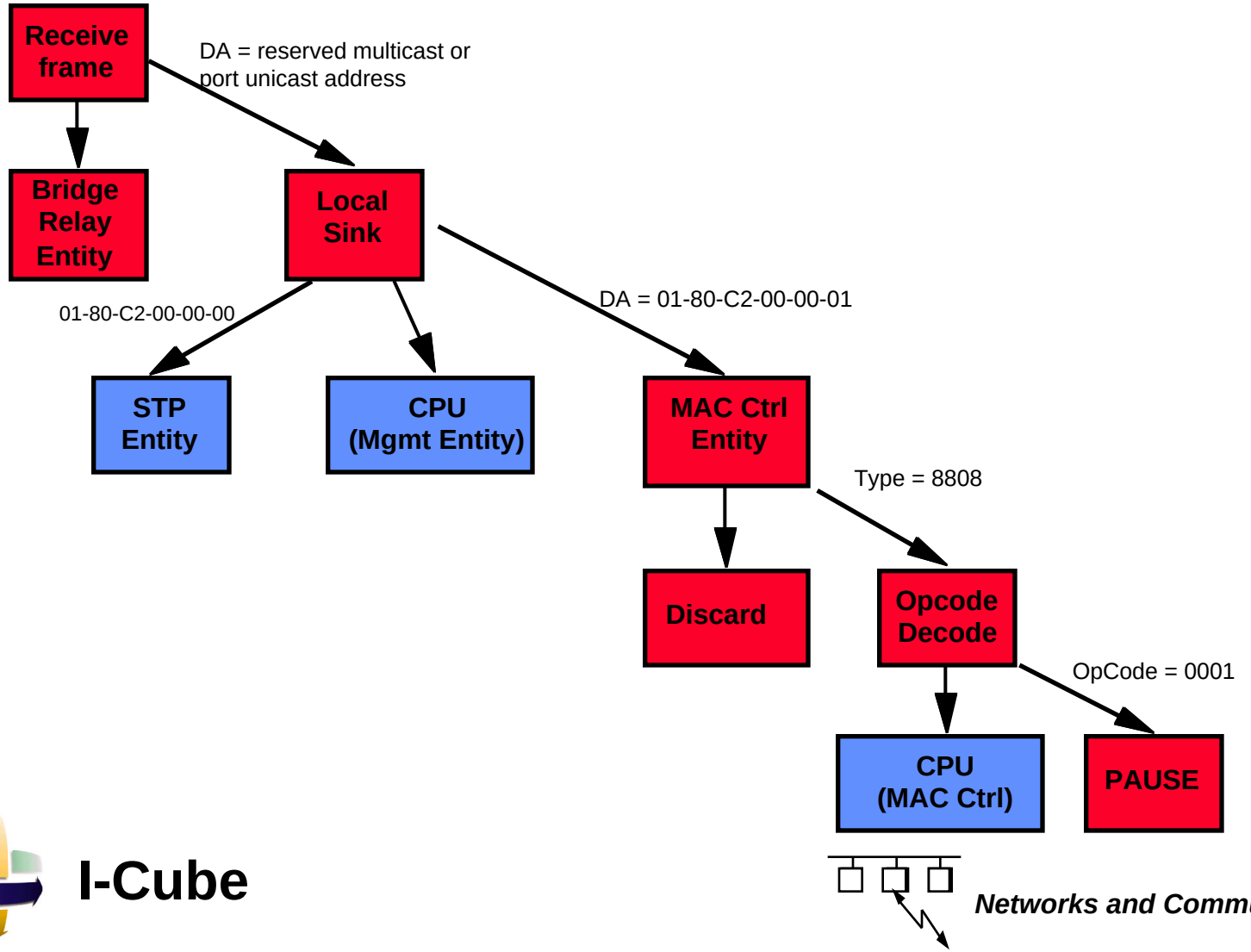
Case 1: Pre-802.3ad, Pre-802.3x, 802.1D-compliant device



I-Cube

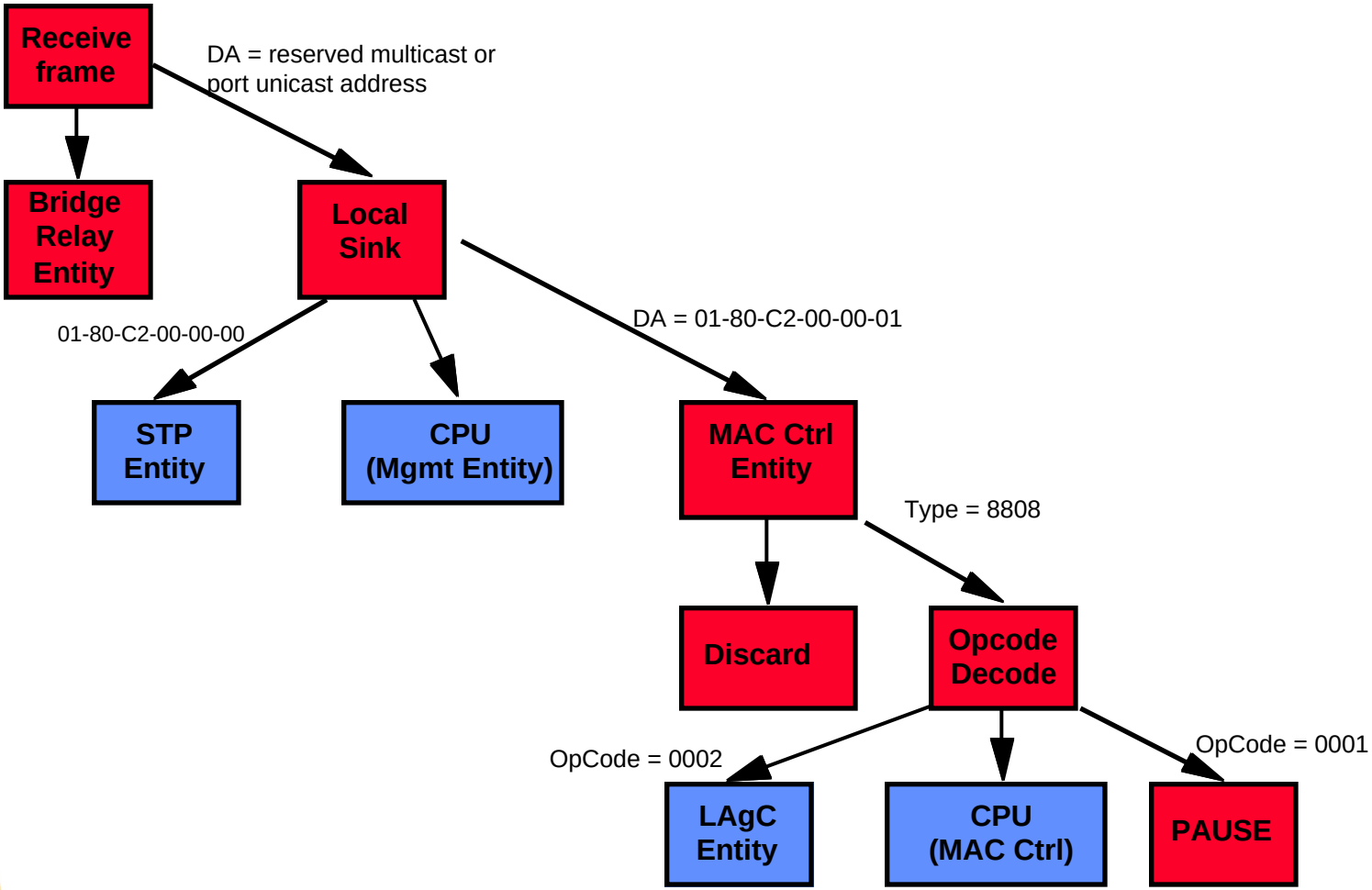
Parsing

Case 2: Pre-802.3ad, 802.3x/802.1D-compliant device



Parsing

Case 3: 802.3ad/802.3x/802.1D-compliant device



Advantages

- Parsing is straightforward in hardware (or software)
  - Each element can be used ONCE ONLY in the decision tree (no need to fork in multiple places on the same branch criteria)
- Pre-defined, fixed-length frame makes hardware easier
- May be possible to add to existing 802.3x-compatible devices with only a software change

Implications for the Standard

- Would need to redefine the semantics of the PAUSE reserved multicast address
  - Currently defined for PAUSE operation only (not generic, link-restricted protocol usage)
  - Technical change to 802.3, Editorial change to 802.1D/Q
- Link Aggregation Control Protocol becomes an Annex to Clause 31