802.1AX -- Link Aggregation:

Editor's Report: September 2017

Version 1

Stephen Haddock

September 7, 2017

Draft 0.1 / First Task Group ballot

- Draft generated for first Task Force ballot
  - Change bars relative to draft 0.0 (merge of 802.1AX-2014 and 802.1AX-2014-Cor1-2017)
  - Bulk of changes in Clause 6. Other clauses, in particular Clause 9 (DRNI), have very few updates as of yet.
- Significant changes to Clause 6:
  - Replace Mux state machine with proposal from http://www.ieee802.org/1/files/public/docs2017/ax-rev-seaman-wait-to-restore-0117-v01.pdf
  - Integrate the per-Aggregation Port portions of Conversation-sensitive Distribution and Collection (CSDC) into existing LACP state machines.
  - Modify CSDC Update Mask state machine to correct behavior when defaulted, specify behavior when management changes administrative variables, and (hopefully) clarify information flow.
  - Added informative text on how CSDC works (and a little on why you might want it).

Task Group Ballot results

- Ballot completed June 23, 2017

- Ballot Results
  - Approve: 4
  - Disapprove: 4
  - Abstain: 12

- Comments
  - Commenters: 5
  - Comments: 199
    - TR: 67 T: 34
    - ER: 24 E: 74

Comment Resolution

- Proposed dispositions resulting from July meeting:
  - http://www.ieee802.org/1/files/private/ax-rev-drafts/d0/802-1AX-Rev-d0-1-pdis-v02.pdf
- Resolution
  - 40+ comments discussed in July meeting
  - Large number of comments were “Propose Accept” or “Propose Accept in Principle” with minor variations on the suggested remedy.
    - July Editor’s report stated that these would be accepted as is unless someone specifically requests discussion of them.
    - This has now been done.
- Current status
  - One unresolved comment (189 – “element of determinism”)
  - Two outstanding issues on Mux machine

Outstanding Issues

1. Comments 139, 177
   - Mick proposed a change to make non-revertive links go to the DETACHED state in the Mux machine (currently go to WTR_ATTACHED).
   - Editor disagrees
2. Comment 146
   - Editor's note proposed a change to make coupled-control implementations go through the COLLECTING state of the Mux machine (currently go directly to COLLECTING_DISTRIBUTING).
   - Mick disagrees

For each of these the proposed change will not implemented in draft 0.2, and the issue will be described in editor's notes and/or Annex Z.

Plans for draft 0.2

- Incorporate resolutions from ballot on draft 0.1.
- Update clause 9 (DRNI)
- (Hopefully) conduct second Task Group ballot prior to the November meeting.

DRCP version number

- Plan is to roll version number from 1 to 2.
- Version 2 DRCP implementations will not work with version 1 (actually version 1 will not work at all).
  1. Version 1 DRCPDUs do not include the Aggregation Partner's System Identifier.
     - Means there is no way for Portal Systems to confirm that they are creating a LAG with the same Partner.
  2. Version 1 DRCPDUs contain a list of active Aggregation Port Numbers, instead of the Link Numbers
     - The Link Numbers are necessary for Conversation-Sensitive Distribution and Collection (CSCD) to work.
  3. Version 1 DRCPDUs do not include digests of the Port and Gateway Service ID-to-Conversation ID maps.
     - Constrains the distribution algorithms that can be used for CSCD.

TLV_Type values

- Can the contents of a TLV be redefined without changing the TLV type value?
  - In this case I would argue that the answer should be YES.
    - Since the version number is changing and there is no expectation that a version 2 implementation will work with a version 1 implementation, I don't think this is a problem.
  - Can mitigate the risk of a version 1 implementation attempting to parse a version 2 DRCPDU by making sure the "Portal Check" fails.
    - The Portal Check will fail if we change the definition of the "Portal System" field from the System Identifier of the Portal to the System Identifier of the Portal System.

DRNI System Identifiers

System Identifier = concatenation of Priority + MAC Address

Portal System System Identifier
(address unique to each Portal System)

Portal System Identifier
(address common to all Portal Systems in the Portal)

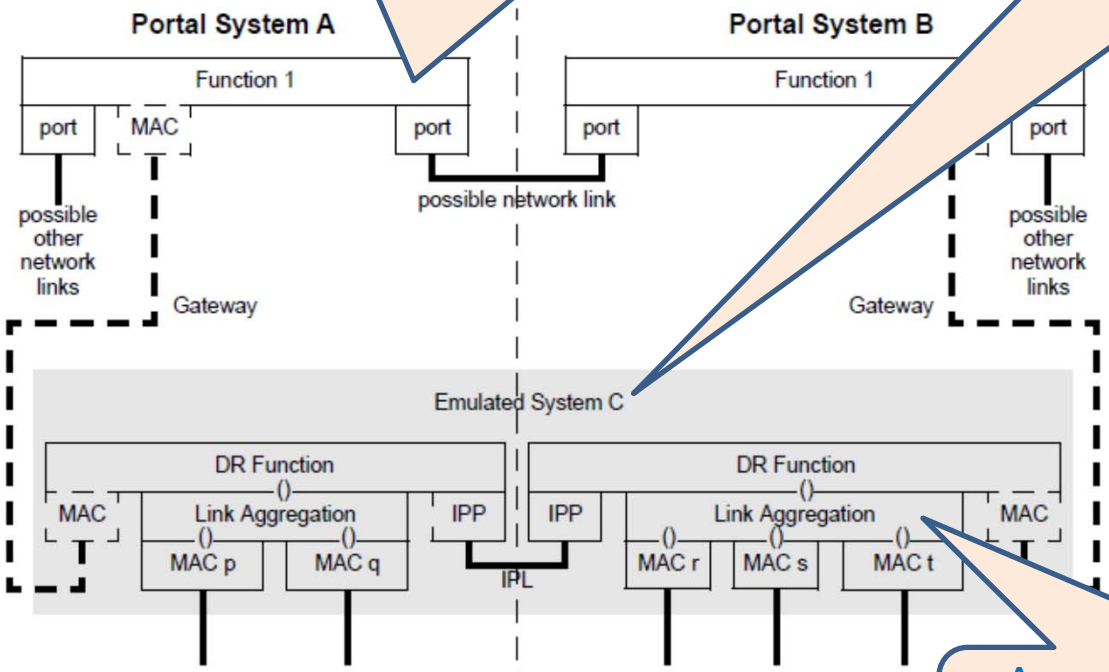


Figure 9-3—Distributed Relay: as seen by Systems A and B

Can be the same

- Reason for having both dates back to old ideas about how to resolve split brain issue(?)

Aggregator System Identifier
(address common to all Portal Systems in the Portal)

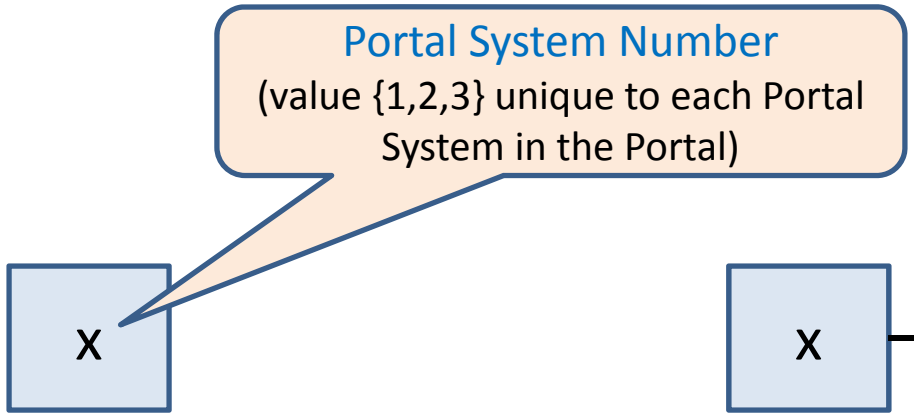
System Identifiers in DRCPDUv1

- First TLV in DRCPDUv1 contains two System Identifiers (concatenation of a priority field and a MAC address):
  1. Aggregator: System Identifier of the LACP "Actor"
     - Must be the same in each Portal System of the Portal
  2. Portal: The System Identifier of the "Portal"
     - Must be the same in each Portal System of the Portal
- The two System Identifiers can be the same
  - In most cases they will be the same.
  - Both were included in case we discover that complex Systems (such as Backbone Edge Bridges) need distinct addresses for the Emulated System when viewed from the "network" versus the "interface".
    - For version 2 I propose that we deal with this possibility, when it becomes necessary, by defining an optional TLV containing the second address.

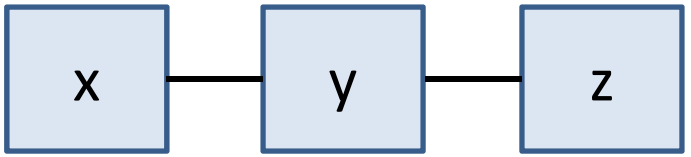
System Identifiers in DRCPDUv2 (Proposal)

- First TLV in DRCPDUv2 contains two System Identifiers (concatenation of a priority field and a MAC address):
  1. Aggregator: System Identifier of the LACP "Actor"
     - Must be the same in each Portal System of the Portal
  2. Portal System: The System Identifier of the "Portal System"
     - Must be different in each Portal System of the Portal
- Guarantees that if a version 1 Portal System is connected to a version 2 Portal System, both will fail the "Portal Check".
- Including the identifier of the Portal System in the DRCPDU is not strictly necessary, but it is helpful in detecting some (obscure and highly unlikely) Portal topology errors.

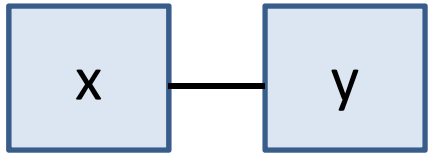
Valid Portal Topologies



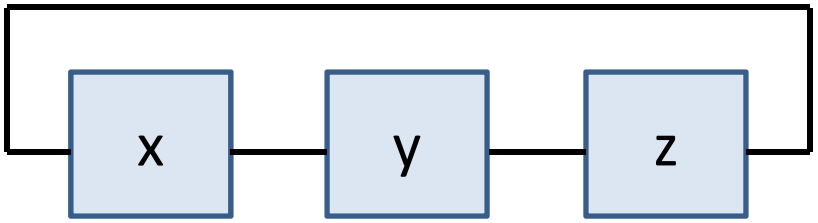
1. Single Portal System



3. Chain of three Portal Systems



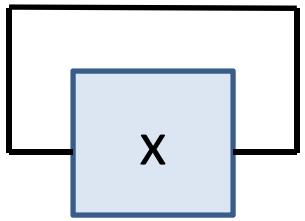
2. Two Portal Systems



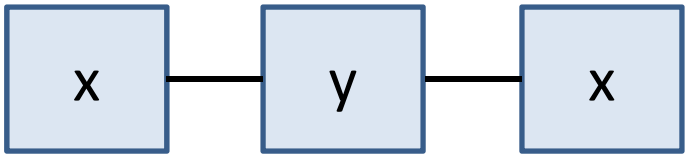
4. Ring of three Portal Systems

Portal Topology Errors

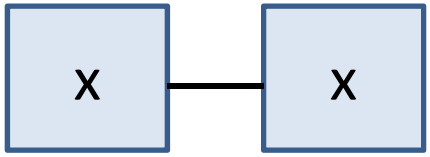
(Detectable based on Portal System Numbers)



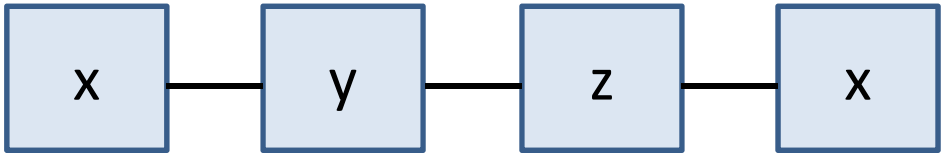
1. Loopback Portal System



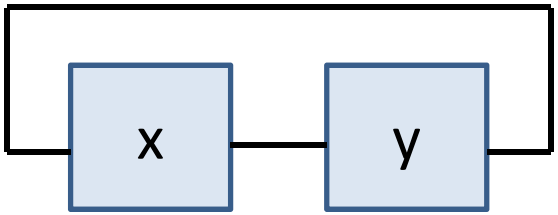
4. Duplicate Portal System Number



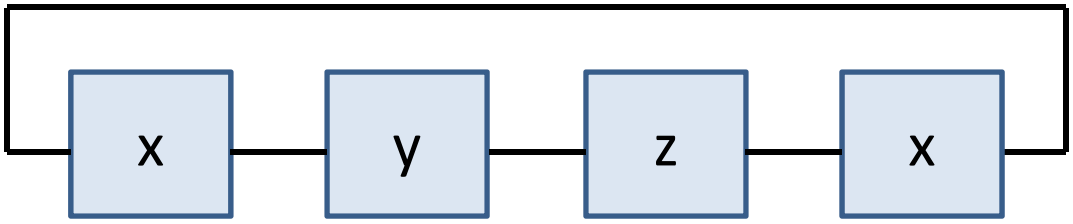
2. Duplicate Portal System Number



5. Chain of four or more Portal Systems



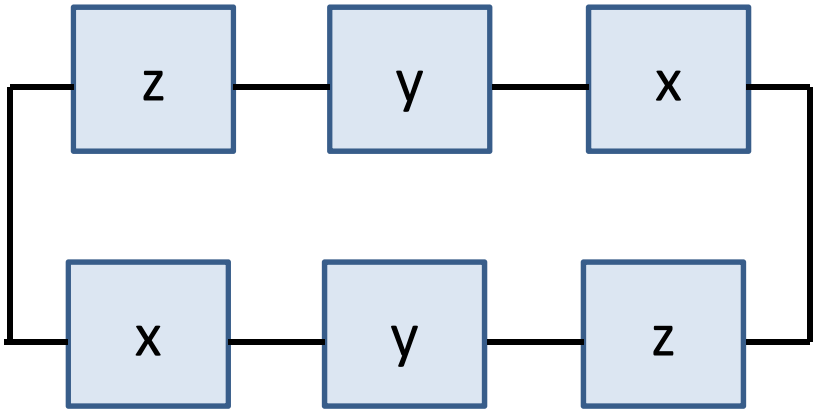
3. Ring of two Portal Systems



6. Ring of four or more Portal Systems

Portal Topology Errors

(Requires System Identifier of Portal System to detect)



7. Ring of multiples of three Portal Systems

Backup Slides

Comments to Discuss

- Port Algorithms: 16
- TLV length: 133
- Bit order: 129
- Point-to-point support: 91, 92, 178
- Addressing: 97
- ISS MAC_Enable: 171, 191
- Digests: 77, 78
- Parameter names: 26, 27, 11, 161, 39
- Determinism: 189
- Mis-ordering: 145, 151, 157
- Loopback: 184
- Individual: 185
- Conformance: 17, 18, 19
- Misc.: 20, 82, 47, 135, 162, 188, 132

Items in blue discussed
July 11-12 2017 in Berlin

Mux/WTR comments

- Comments to be discussed
  - Transition with “|| (Wtr_timeout == 0)”?: 29, 192
  - NTT in DETACHED?: 9
  - Optional WTR?: 35, 38
  - Timer semantics and names?: 29, 33, 30, 31
  - Detached when non-revertive?: 139, 177, 72
  - Coupled_control and COLLECTING?: 146

Items in blue discussed
July 11-12 2017 in Berlin

Individual Ports

- A port designated as “Individual” (i.e. not “Aggregateable”) will not form a Link Aggregation Group with any other link.
- What does this mean in terms of selecting an Aggregator?
  a) An Individual Aggregation Port always has an associated Aggregator that is dedicated to that port. This Aggregation Port cannot select any other Aggregator and no other Aggregation Port can select this Aggregator.
     - This is explicitly the case in the context of “Recommended default operation” in revisions prior to 802.1AX-2014-Cor-1.
     - This functionality can be achieved by assigning the Aggregation Port and Aggregator a key not used by any other port or Aggregator, independent of designating the port Individual
  b) An Individual Aggregation Port selects an Aggregator using the same selection rules as any other port, with the added constraint that no other Aggregation Ports can select the same Aggregator. The fact that a port is designated “Individual” does not, in itself, give that port a higher claim to any particular Aggregator.
     - This seems to be what is envisioned in the “Selection Logic requirements”

July 12 2017: Prefer b)

Possible diagram for 6.6.3

re: Comment #161