Resilient Network Interconnect:  
D-LAG Models  
Version 1

Stephen Haddock  
Extreme Networks  
October 5, 2010
Introduction

• At the September Interim two models for Distributed Link Aggregation were presented:
  – Distributed Bridge Model
  – Distributed Port Model

• Concerns were raised with respect to the Distributed Port model.

• This presentation modifies the model to address those concerns.
‘Baggy Pants’ Representation

Higher Layer Entities
Relay
Link Aggregation
MAC
MAC

Higher Layer Entities
Relay
Link Aggregation
MAC
MAC
Distributed Bridge Model

• Emulate a single bridge
  • Create illusion that there is a single relay, single instance of all higher layer entities, and a single Bridge Port representing entire Distributed Link Aggregation Group.

• In normal operation neither the NNI nor My Network can distinguish this from a single bridge.

• Failure of the DLACC (“split brain” scenario) potentially causes a significant change in operation as viewed from My Network.
Distributed Port Model

- All unique behavior confined to the Ports that are part of the D-LAG.
  - Each Node operates as a separate bridge on all ports that are not part of the D-LAG.
- Distributed LAG creates a single Bridge Port on the Relay of each bridge.
  - LAG Distributor and Collector functions control frame forwarding between the D-LAG links and the Bridge Relays.
  - In some cases may require “tunneling” frames on the DLACC to the other Node.
- May need special behavior in port specific portions of some L2 protocols to maintain single Bridge Port illusion across D-LAG:
  - Probably xSTP and MxRP (if run these over D-LAG); maybe CFM LinkTrace
Concerns on Distributed Port Model

• Panos:
  – Generally uncomfortable with a single Link Aggregation Group looking like a Bridge Port on each of two distinct Bridges.

• Mick:
  – Specifically concerned with the idea that from the NNI the D-LAG looks like a single Bridge Port, while from My Network it looks like two distinct Bridge Ports, each on a distinct Bridge.
  – Means it is impossible for any control plane protocol operating over both My Network and the Other Network to have a consistent world view.
  – Presents an insoluble problem to any routing protocol (and perhaps to any control protocol?).

• Need a model where the D-LAG looks like a single Bridge Port from both the NNI and My Network.
Distributed Component Model

- Distributed LA Sublayer comprises a logical VLAN-aware component that:
  - Spans all physical bridges.
  - Has a single Bridge Port for all external links in the Distributed-LAG.
  - Has internal links/ports to the bridge component in each physical bridge.
  - Distributed Relay acts as a VLAN multiplexer (no MAC address learning).
Network Representation

Device View:

Logical View:

My Network

DLACC

D-LAG

LAG
Distributed Component Model: Data Plane

– FDB of Distributed Relay configured as a VLAN multiplexer.
  • Member set of an VID includes only the D-LAG Bridge Port and one of the internal Bridge Ports (same constraints as a PEB C-VLAN component).
  • No MAC address learning.

– Results in same behavior as the Gateway function described in the Distributed Port Model of

– Network data flows are the same as those described in the Distributed Port Model.

– Still have situations where a frame needs to be transferred between physical bridges in the Distributed Link Aggregation Sublayer:
  • Frames received (or to be transmitted) on a D-LAG link terminating at one physical bridge, while the frame’s VID is in the member set of a Bridge Port on another physical bridge.
  • Such frames may be transferred on a dedicated physical link, or tunneled on a physical link shared with the normal active topology.
Distributed Component Model: Control Plane

- Distributed Component runs an instance of all supported control applications (e.g. RSTP/MSTP).
  - Since Bridge Port and VLAN configuration have same constraints as a PEB C-VLAN component, can use the RSTP enhancements described in 13.38. This allows the Distributed Component to have multiple Root Ports when the D-LAG Bridge Port is Designated.
- As with Distributed Port Model, still need a Distributed Link Aggregation Communications Channel (DLACC):
  - to convey Distributed Link Aggregation Sublayer state and control information between physical devices.
  - to transfer data plane frames in the Distributed Link Aggregation Sublayer between physical devices.
Distributed Component Model: Observations

• Model presents D-LAG as supporting a single Bridge Port when viewed from NNI or My Network.
  – Provides a “consistent world view” from any point in network.
• Model provides clear behavioral reference for any higher layer application, control protocol, or protocol shim.
• Model easily accommodates more than two physical bridges in the D-LAG.
• Model easily accommodates bridges supporting multiple D-LAGs and overlapping D-LAGs.
• Model easily accommodates D-LAGs on bridges that are already multi-component.
  – E.g. Provider Edge Bridges and Backbone Edge Bridges
Some Thoughts on Standardizing Distributed Link Aggregation
Distributed Link Aggregation: Standardization

- Amendment to 802.1AX Link Aggregation
  - Add a new Distributed Link Aggregation Sublayer clause (or two)
  - Allow either Distributed Bridge or Distributed Component as conformant behavioral reference models.

- No changes to 802.1Q
  - Can just refer to 802.1Q for component definitions and specifications.

- Minimal specification if assume single vendor for all bridges in D-LAG:
  - Require that external behavior must match the Distributed Bridge or Distributed Component Model.
  - Specify constraints on VLAN configuration of Distributed Component Model.
  - All details of how to create Distributed Bridge or Distributed Component, including the DLACC, left to the implementer.
  - No standardized management model.
  - Will probably need to specify or constrain the uniqueness versus re-use of identifiers for logical ports and components.
  - Will need to specify what the model looks like when the DLACC fails.
Distributed Link Aggregation: Standardization

- Specification if do not assume single vendor for all bridges in D-LAG:
  - Specify how functionality is distributed between physical devices for Distributed Component Model only (not Distributed Bridge).
    - Distributed Relay probably best specified as a Gateway function in each device.
    - Specify whether control protocols are to be distributed, or run in a selected device with PDUs tunneled to/from Bridge Ports in other physical devices using the DLACC.
    - Specify Distributed Component management model, and which managed objects are implemented by which physical device.
      - Could follow 802.1ah precedent where all objects/parameters of a full-up component are specified, or 802.1ad precedent where only pertinent objects/parameters are specified.
  - Specify frame formats for the DLACC.
    - Simplest approach for interoperable DLACC may be to assume a directly connected dedicated link. Single vendor solutions may optionally implement a way for the DLACC to share links with the active topology.