

# Resilient Network Interconnect using Distributed Link Aggregation

## Case for Distributed Port Model

Prakash Kashyap  
Extreme Networks

IEEE Interim Meeting  
Sep, 2010

# Distributed Bridge & Distributed Port Model

- Example implementation of both models exist in the industry.
- Examples of Distributed Bridge model:
  - Switch Clustering of physically different switches
  - Switch Stacking of physically different switches.
- Both these emulate Single Bridge.
- Involves full control plane integration. Forwarding plane under the control of one controlling entity.

## Distributed Bridge & Distributed Port Model (cont'd)

- Examples of Distributed Port model:
  - D-LAG/M-LAG implementations.
- Involves integration of only specific controlling entities in the switches. Specifically, LAG entities.
- More amenable for implementation as a loosely coupled control entities.
- More amenable as a software upgrade feature.

## Complexity of Distributed Bridge vs. Distributed Port Model

- Clustering can be considered as a “superset” among the 3 features.
- However:
  - Each feature has it’s primary problem space it is trying to solve.
  - Generally, all these 3 features are supported equally. i.e. one has not displaced others completely.
  - Reason for this due to different complexity & applicability of these features.
- Hope, it is not too controversial to say:
  - On complexity Scale: High -→ Low
  - Clustering → Stacking → D-LAG/M-LAG

## Complexity of Distributed Bridge vs. Distributed Port Model

- Higher complexity translates to:
  - Availability on limited/newer platforms.
  - Availability only among homogeneous switch architectures.
  - Just as a reference: Qbh “Port Extension” work found the necessity to define new tags. Even though Qbh itself, is trying to provide a more limited functionality than clustering for e.g.
- Note: This is true even if it is all a single vendor solution.

# ENNI & D-LAG Model

- I think general feeling is:
  - D-LAG & Single Bridge emulation model would work as ENNI.
- Yes, ENNI specification can perhaps just chose “Distributed Bridge” as the only model.
  - Based on the fact that industry already knows how to build “Distributed Bridge” model.
  - Nothing special may need to be mentioned about interface to region nodes.
  - But, this may also mean ignoring implementation realities.
- Mandating just one model for D-LAG nodes would impose unnecessary burden on quick adaptation & availability of solution for ENNI.

## LACP changes for the two models?

- Proposed LACP changes applicable even to Distributed Bridge model.
  - Can benefit from optimal use of D-LAG links.
  - However, will work without any LACP changes.
- If both sides of ENNI are doing Distributed Bridge D-LAG model:
  - Just assume “Split-Brain” never happens?
  - Not very realistic, however.
- LACP changes covering optimal use of D-LAG links & handling of “Split-Brain” scenario may have to be specified anyway.
- Such LACP changes would be similar for both models.
- Distributed port model would also work without LACP changes involving D-LAG link use preferences, maybe, sub-optimally.
- Split-brain handling would have to be deemed mandatory for Distributed Port model.

## Impact on Region nodes?

- From discussions so far, no changes in region nodes have been found necessary to interface to the region's D-LAG nodes.
- Haddock's presentation identifies some changes in xSTP of D-LAG nodes, if the desire is to run one xSTP covering all region and D-LAG nodes.
- I think (Norm Finn) proposed the possibility of putting D-LAG nodes in a MSTP region of their own. This would avoid changes in xSTP of D-LAG nodes as well.
- So, it seems that there would be number of ways for region nodes to inter-work with D-LAG nodes implementing "Distributed Port" model.

## Recommendations

- Include in ENNI work:
  - LACP changes to achieve optimal use of D-LAG links, & handling of split-brain scenarios in LACP.
  - Allow both Distributed Bridge & Distributed Port model for D-LAG.
  - Specify ways for region nodes to inter-work with D-LAG nodes adhering to “Distributed Port” model. At least, as informational text.
  - Helps with inter operability of region nodes and D-LAG nodes.