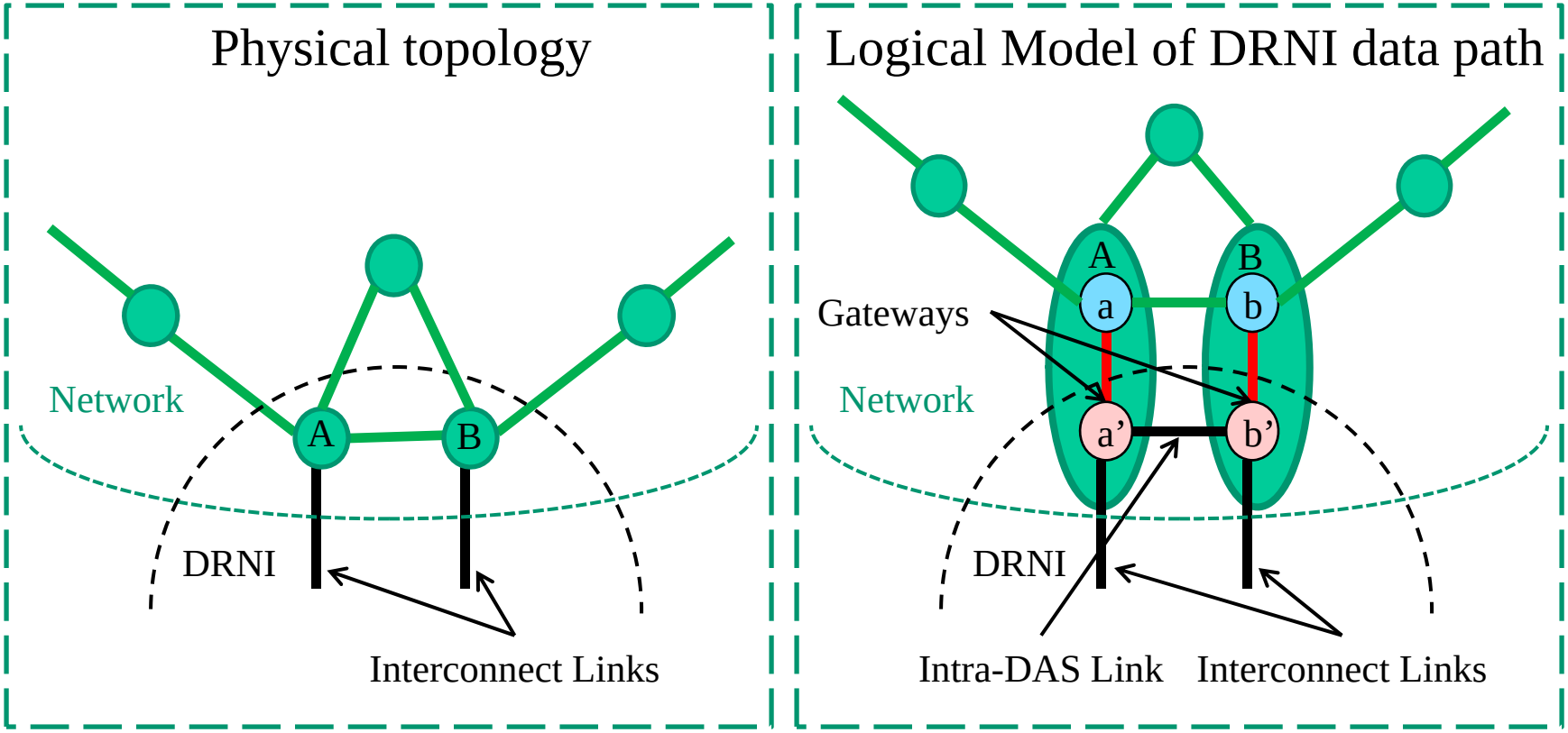
DRNI – Intra-DAS Link

Version 02

Stephen Haddock

August 30, 2011

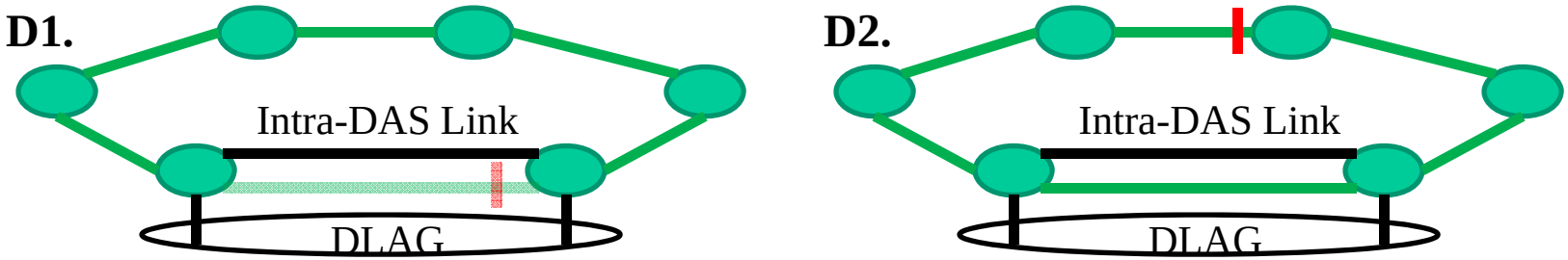
Review of DRNI Gateways and Links



- The functionality of the Gateways, a', and b' are all in the Distributed Link Aggregation Sublayer.

- Notice that in the logical model the Intra-DAS Link between a' and b' is distinct from the network link between a and b. This presentation explores three options (D, V, and S) for implementing the logical model of the Intra-DAS Link and network link on the physical topology.

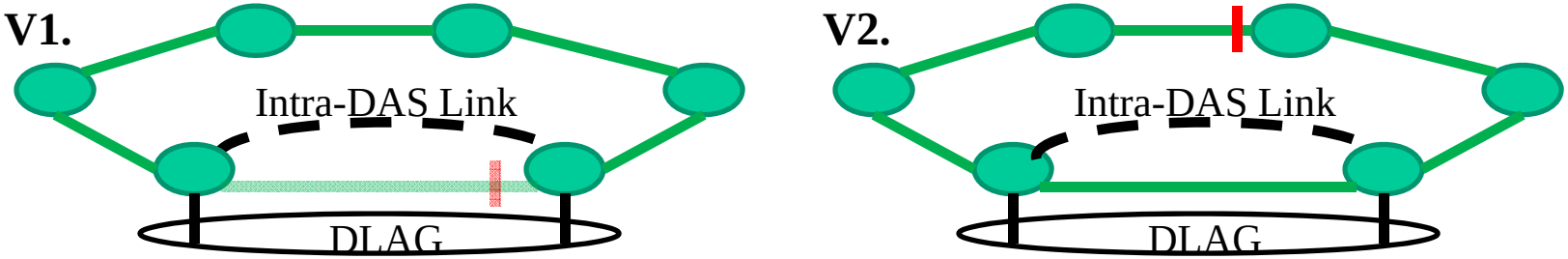
Option D: Dedicated Intra-DAS Link



In this case the Intra-DAS Link is a direct physical link between the DRNI nodes dedicated to the Intra-DAS Link function.

- In Case D1 there is no direct physical network link between the DRNI nodes that is part of the active topology (either there is no link, the link is broken, or it is blocked by a loop prevention algorithm). In Case D2 there is a direct physical network link between the DRNI nodes that is part of the active topology.
- The Dedicated Intra-DAS Link is separate from the network link (if the network link is present) and is only used for DRNI traffic.
- DRNI data frames traversing the Dedicated Intra-DAS Link do not require encapsulation.

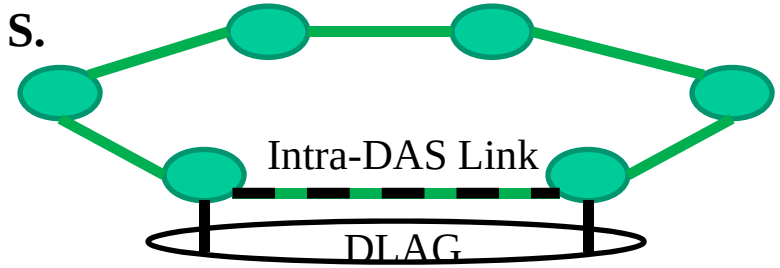
Option V: Virtual Intra-DAS Link



In this case the Intra-DAS Link is a virtual connection (a.k.a. tunnel) between the DRNI nodes.

- In Case V1 there is no direct physical network link between the DRNI nodes that is part of the active topology (either there is no link, the link is broken, or it is blocked by a loop prevention algorithm). In Case V2 there is a direct physical network link between the DRNI nodes that is part of the active topology.
- The Virtual Intra-DAS Link overlays the active topology of the network links.
- DRNI data frames traversing the Virtual Intra-DAS Link must be encapsulated to distinguish them from network data frames (e.g. using PBB for encapsulation).
- See http://www.ieee802.org/1/files/public/docs2011/new-farkas-RNI-data-plane-0111-v02.pdf

Option S: Shared Intra-DAS Link



In this case there is a direct physical link between the DRNI nodes that is used both as a network link and as the Intra-DAS Link.

- When the physical link between the DRNI nodes is not in the active topology of the network, the link is used only as an Intra-DAS Link (Case S1).
- When the physical link between the DRNI nodes is in the active topology of the network, the link is used only as a network link (Case S2).
- DRNI data frames traversing the Shared Intra-DAS Link do not require encapsulated to distinguish them from network data frames.

Objectives for Gateway and Link Selection

- Gateway selection objectives/requirements
  1. Assure that at most one copy of any frame is delivered between the network and the DRNI.
  2. For bridges, gateway selection must be compatible with the learning processes within the network.
- Link selection objectives/requirements
  1. The same link must be selected for all frames belonging to any given “conversation” (traditional Link Aggregation constraint).
  2. A DRNI must support a means to select the same link in both directions for all frames belonging to a single service (reverse path congruent per-service link selection).
- Constraint added by time-sharing proposal
  - When the link between devices is part of the active topology of the network, the selected link and selected gateway must be on the same device.

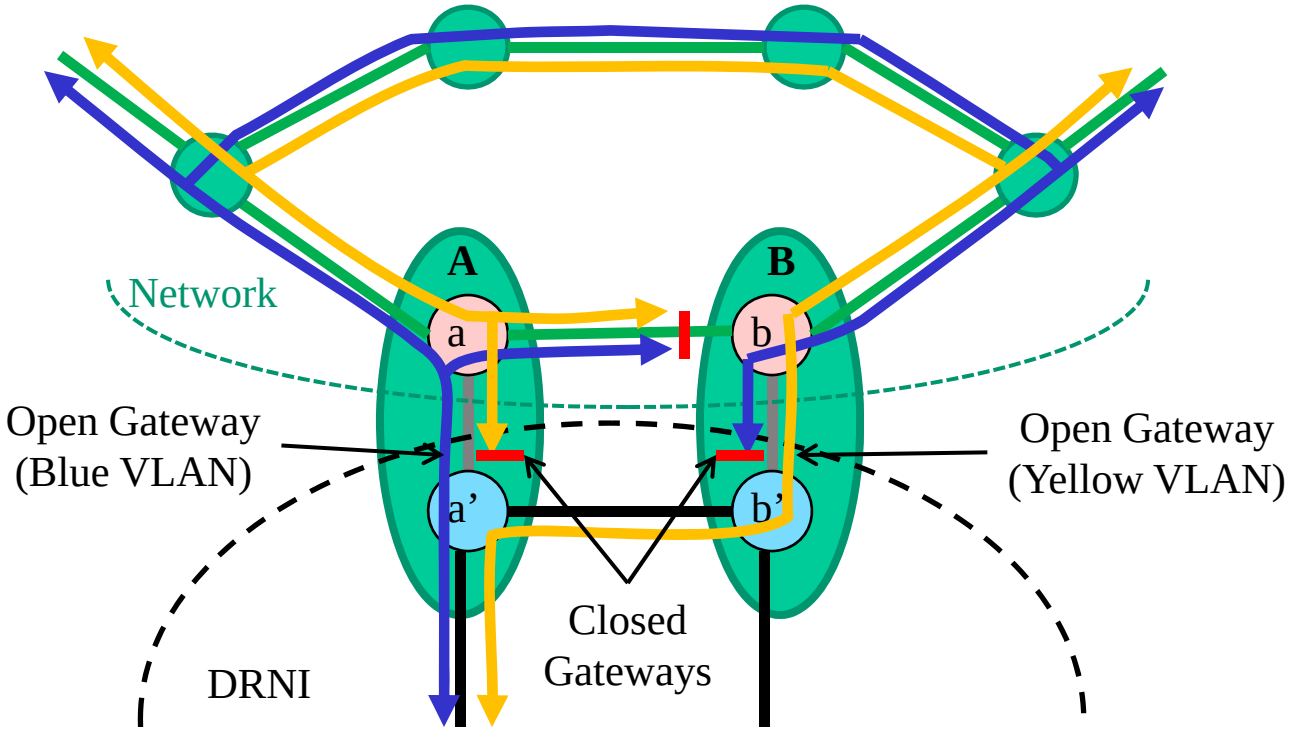
DRNI Data vs Control Frames

- DRNI data frames and frames that are received at the gateway in one network, and forwarded across a DRNI link to the gateway in the other network, possibly traversing Intra-DAS Links on the way.
  - This presentation is primarily concerned with how data frames utilize the Intra-DAS Link.
- DRNI control frames are sourced and sunk within the DRNI for the purposes of maintaining the “Distributed Link Aggregation Sublayer” and the “Distributed Forwarding and/or Upper Layers”.
  - This presentation assumes that DRNI control frames are formatted so that they can take any available path between DRNI nodes, including:
    - A Dedicated Intra-DAS Link if present, or
    - A directly connected network link that is part of the active topology of the network, or
    - A directly connected network link that is blocked, i.e. not part of the active topology of the network, or
    - An indirect path following the active topology of the network, or
    - (possibly) An indirect path utilizing DRNI links.

Comparing Dedicated and Shared Intra-DAS Link Options

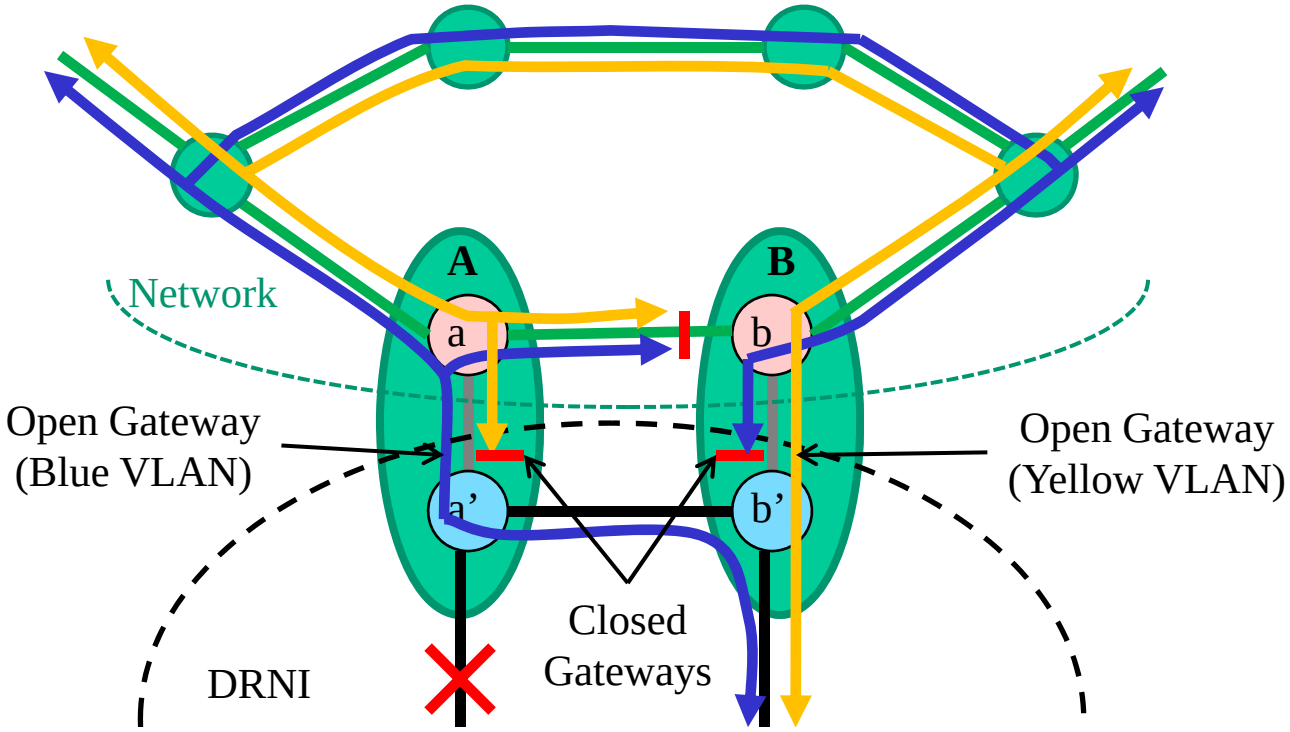
- When the Shared Intra-DAS Link concept was presented in July, the biggest concern that I heard was whether it resulted in undesirable propagation of recovery actions or topology changes between the network and the DRNI.
- To address this concern, the primary focus of this presentation is to examine the response of the Dedicated Intra-DAS Link option to various failure scenarios, and then compare that to the Shared Intra-DAS Link option.

Case D1 examples: Starting Condition



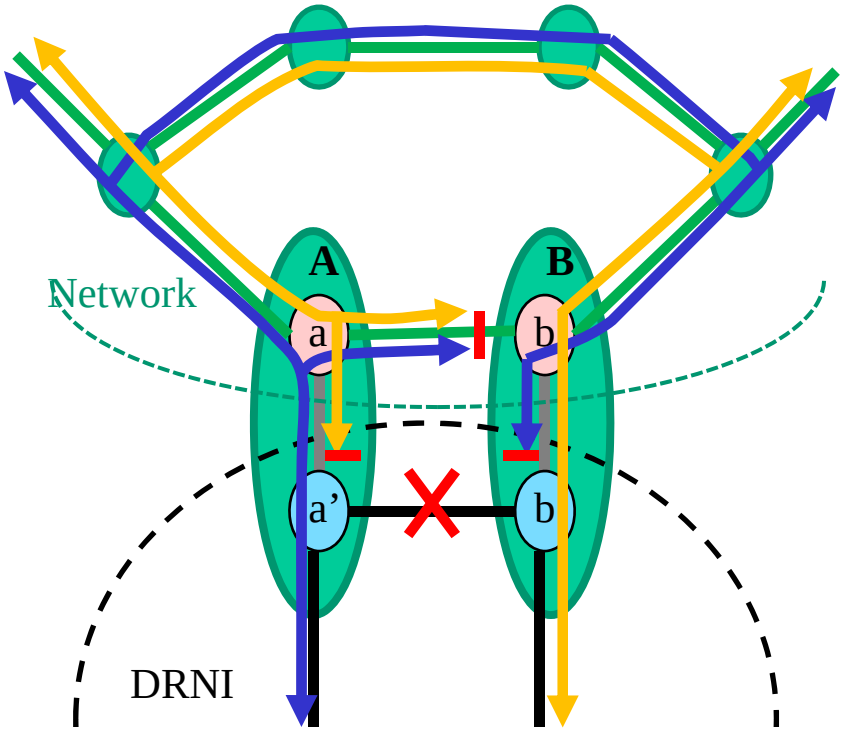
In Case D1 the network link between the DRNI nodes is either not present, is broken, or is blocked. This example shows the broadcast path for two VLANs (Blue and Yellow). The link selection for both VLANs is the same, but the gateway selection is different.

Case D1 examples: DRNI Link Failure

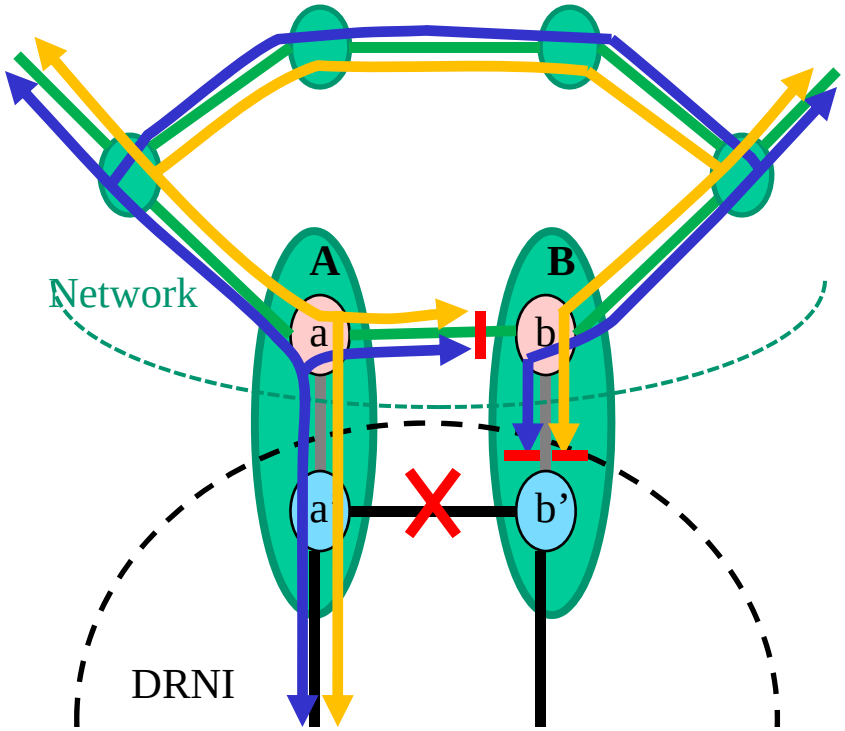


A failure of the DRNI link forces a change in the link selection for both VLANs. The gateway selection does not change, and therefore there is no change within the network.

Case D1 examples: Intra-DAS Link Failure

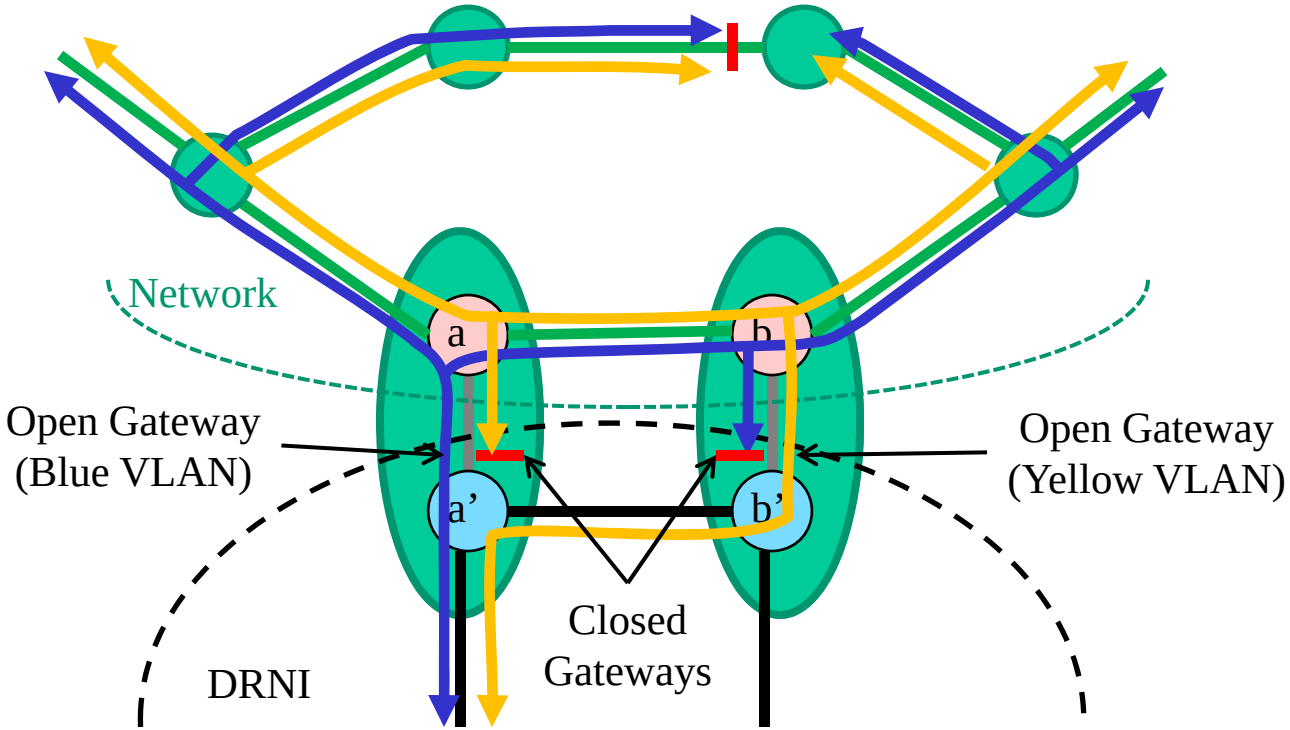


If the Intra-DAS Link failure is considered a failure in the DRNI, then the recovery method is to change the link selection for any VLANs using the Intra-DAS Link.



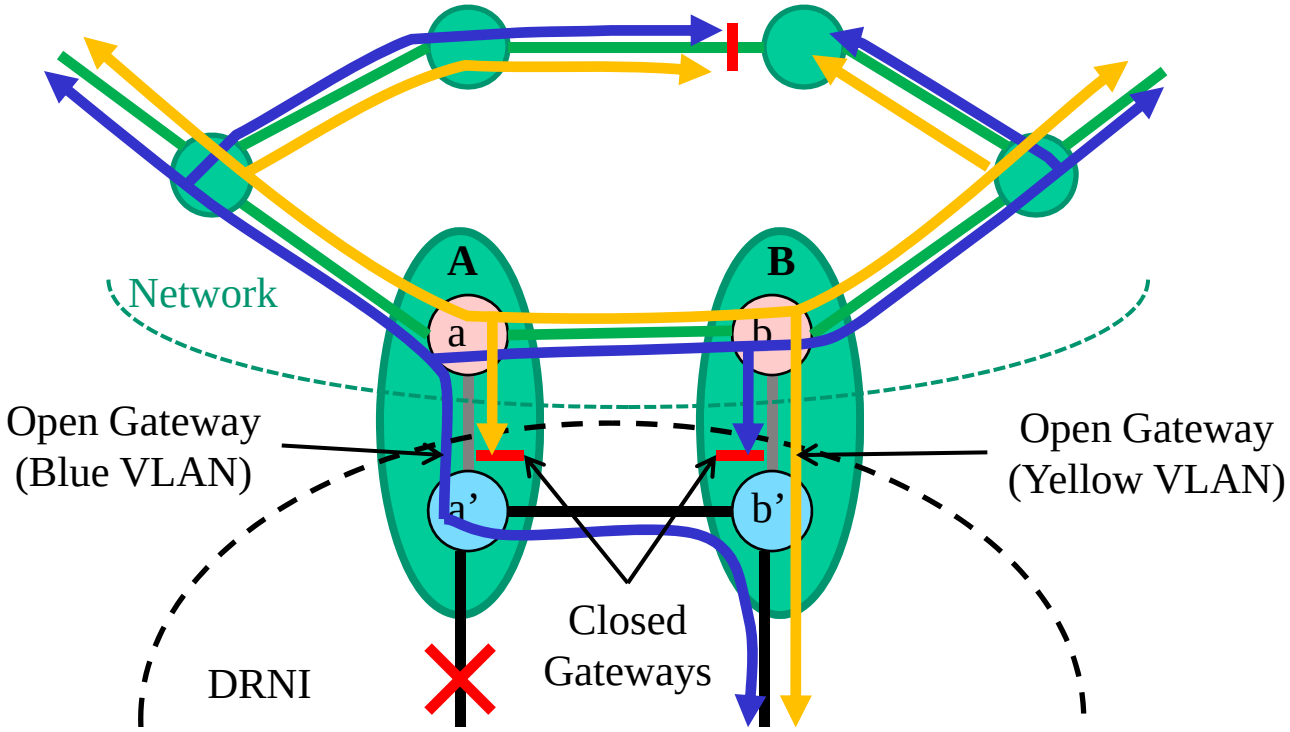
If the Intra-DAS Link failure is considered a failure in the network, then the recovery method is to change the gateway selection for any VLANs using the Intra-DAS Link. Changing the gateway reverses the direction of received DRNI frames on the path between A and B, which is a topology change in the network.

Case D1 to Case D2 transition



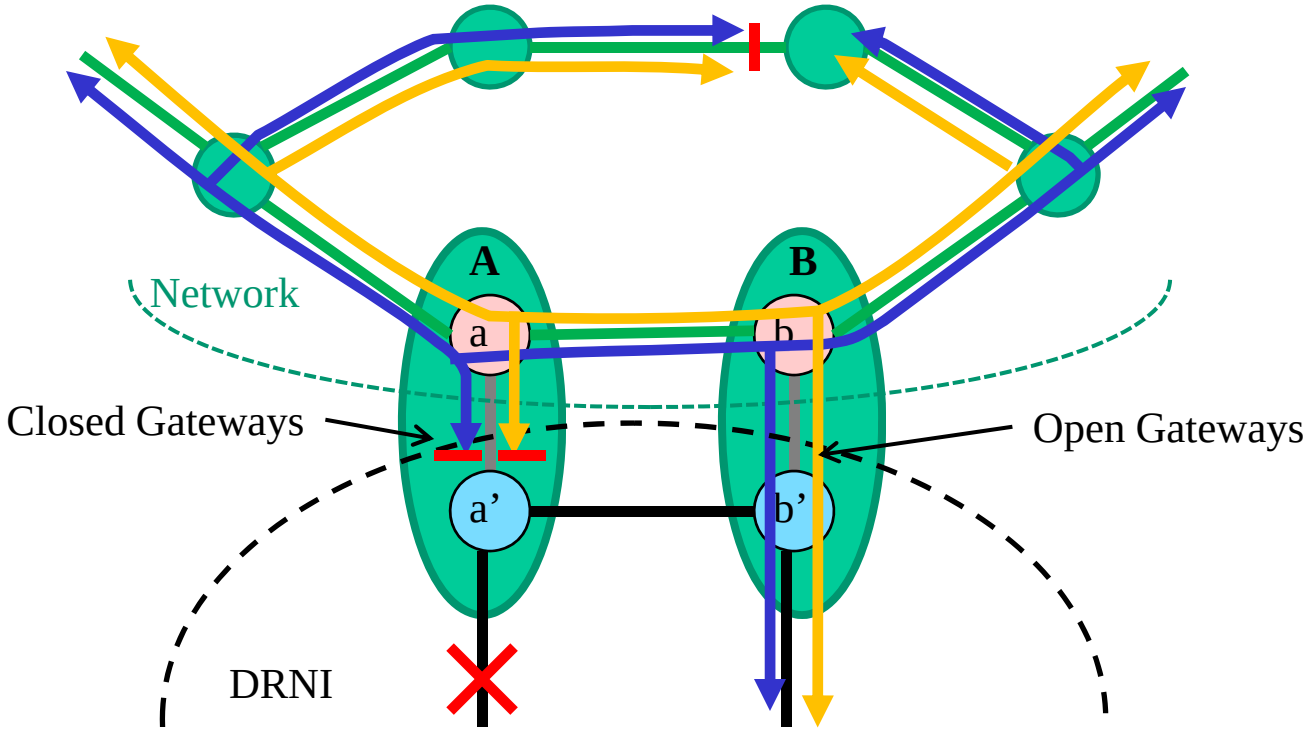
A topology change in the network, possibly resulting from a failure somewhere in the network, causes the network link between the DRNI nodes to become part of the active topology. The gateway selection does not change, however note the inefficiency in the new data path of the Yellow VLAN since frames traverse both the network link and the Intra-DAS Link between the DRNI nodes. This could be optimized by changing the gateway selection for the Yellow VLAN as part of the topology change. The link selection does not change, so the topology change does not propagate to the other network.

Case D2 examples: DRNI Link Failure



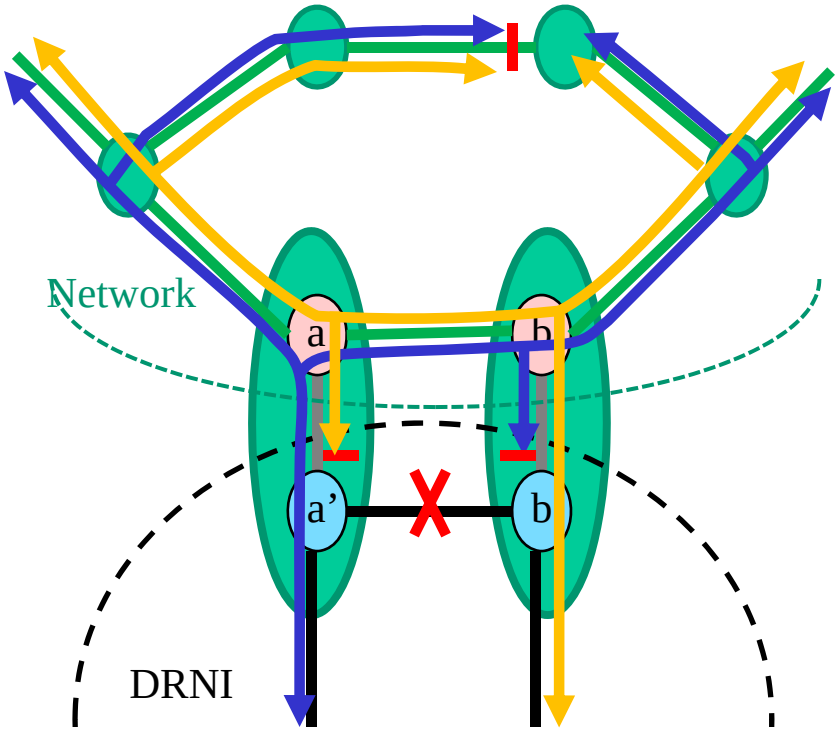
DRNI link failure forces a link selection change in the selected link. Neither the failure itself nor the resulting link selection change force a gateway selection change, however the resultant data path of the Blue VLAN is inefficient.

Case D2 examples: Optimization

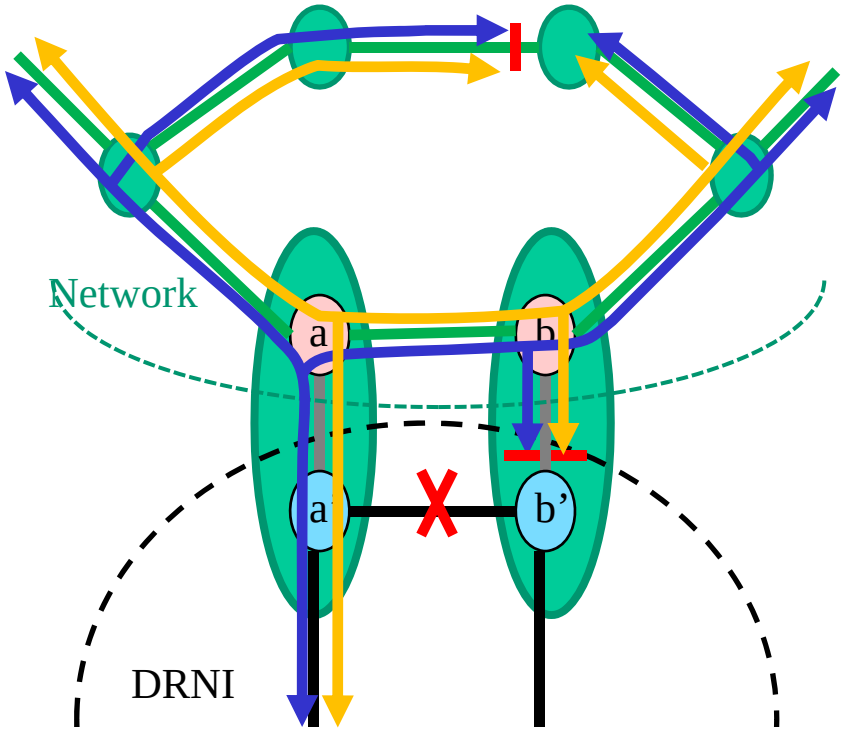


The network operator may choose to allow a change in the gateway selection as a result of the change in the link selection to make the data path more efficient. Changing the gateway reverses the direction of received DRNI frames on the path between A and B, however since the network link between A and B is in the active topology, the scope of the topology change is limited to the DRNI nodes and does not propagate to any other part of the network.

Case D2 examples: Intra-DAS Link Failure



If the Intra-DAS Link failure is considered a failure in the DRNI, then the recovery method is to change the link selection for any VLANs using the Intra-DAS Link.



If the Intra-DAS Link failure is considered a failure in the network, then the recovery method is to change the gateway selection for any VLANs using the Intra-DAS Link. Changing the gateway reverses the direction of received DRNI frames on the path between A and B, which in this case contains the topology change to the DRNI nodes.

Comparing Case D to Case S

- The data flow for Case S1 is identical to the data flow for Case D1, including recovery from all failure cases.
  - The underlying reason for this is that the network link between DRNI nodes is not part of the active topology, so it is never used to carry data frames. In Case D1 there are two physical links, but only the Intra-DAS Link is used. In Case S1 there is only one link, and it is used as the Intra-DAS Link.
- The data flow for Case S2 is identical to the optimized data flow for case D2, including recovery from all failure cases.
  - In case D2, any time the Intra-DAS Link is used the frames traverse both the network link and the Intra-DAS Link. The network operator can unilaterally choose to move the gateway, so the frames traverse only the network link. The result is that only the network link, never the Intra-DAS Link, is used. In case S2 there is only one link, and it is used as the network link.
  - The distinction is that in case D2 the gateway change is a choice to make the data flow more efficient, whereas in case S2 the gateway change is necessary. Because the network link between DRNI nodes is in the active topology, the effect of the gateway change is limited to just the DRNI nodes and has no effect on the rest of the network.

Shared Intra-DAS Link Summary

	Negotiated, symmetric Link Selection	Unilateral Link Selection
Link IS in active topology	CASE S2a Gateway selection based on Link Selection	CASE S2b Gateway selection based on ingress port
Link IS NOT in active topology	CASE S1 Gateway selection based on VID	

NOTE 1: The frame forwarding of Case S1 is the same as for a Dedicated Intra-DAS Link (case D1), and is basically the same as what is currently described in the draft.

NOTE 2: The frame forwarding of Case S2b is the same as for most proprietary Multi-chassis LAG implementations.

NOTE 3: The applicable column is determined by configuration. The applicable row changes with topology changes in the Network, but does not cause changes in the Interconnect.

Comparing Case D2a/S2a to Case D2b/S2b

- The difference between case 2a and 2b is that there is no predetermined link selection.
  - When transmitting to the DRNI this means the link selected is on the first DRNI node encountered by the frame. In terms of gateway selection this means a DRNI node is the gateway for any frame received on any link that is not directly connected to the other DRNI node.
  - When receiving from the DRNI node that receives the frame is also the gateway for that frame.
- The recovery from DRNI or Intra-DAS Link failures does not change between cases 2a and 2b.
  - For a DRNI link failure, the other DRNI becomes the gateway for all frames (which is the case for S2 and the optimized data flow of D2).
  - For an Intra-DAS Link failure, there is no change to the data flow because the Intra-DAS Link is not used (selected gateway is always on the same DRNI node as the selected link).

Shared Intra-DAS Link

- A directly-connected physical link serves as both
  - a network link (available to be part of the active topology of the network), and
  - an Intra-DAS Link

  without requiring encapsulation of data frames.
- This should be specified in the standard.
- Support should be required by the standard.
  - A common required behavior enhances interoperability.
  - Not requiring encapsulation facilitates backwards compatibility goal and increases probability that it could be supported on existing equipment without a hardware change.
  - Proprietary multi-chassis LAG implementations have set market expectation that a dedicated link is not necessary.
  - Can still have option to implement a virtual Intra-DAS Link.
  - Can still have option to implement a dedicated Intra-DAS Link.