Generalized LAN Emulation

Integrating "LAN Emulation - like" media into the heart of a bridge

Norman Finn, Cisco Systems

Emulated LANs

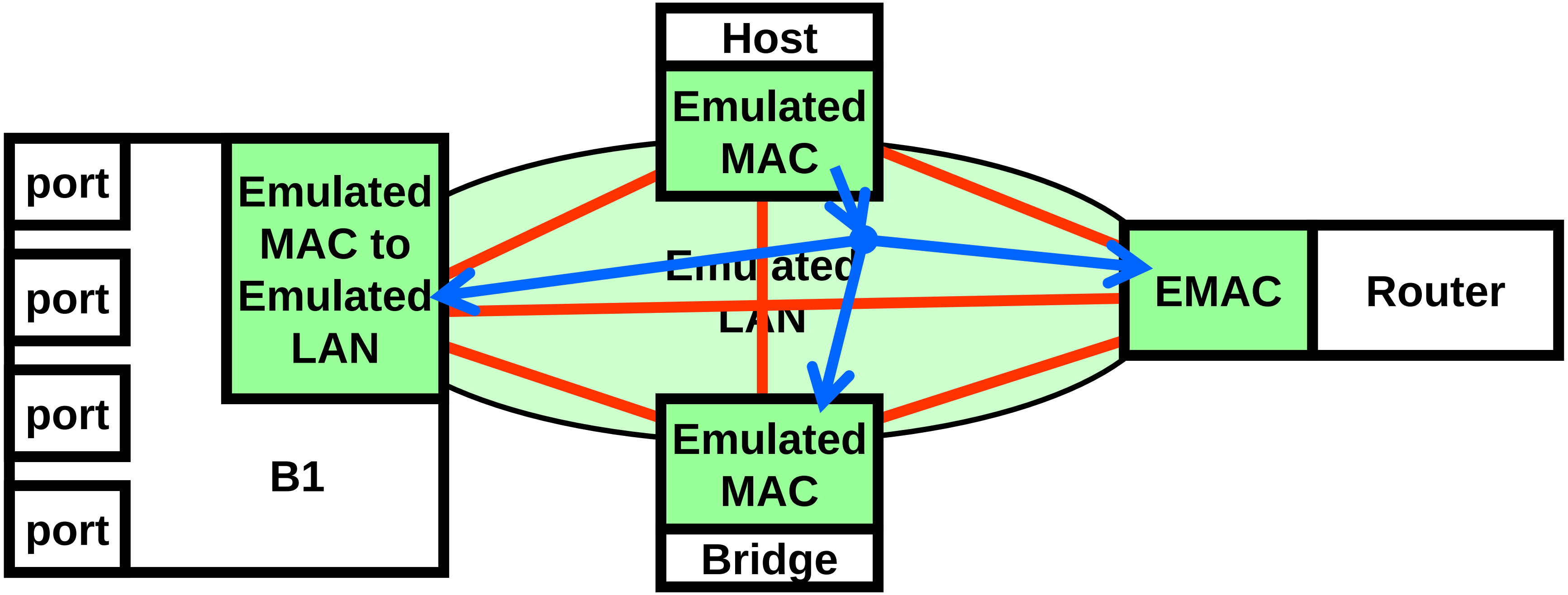
What is an “Emulated LAN”?

- **An “Emulated LAN” is an attempt to make a bunch of point-to-point and/or point-to-multipoint links look, to the bridge, like a single shared medium.**
- **It has ports to L2 endstations and/or bridges that look like ports on a shared medium.**
- **It wants to look like a shared medium to the bridge, router, or host, but have better performance, based on its knowledge of MAC addresses and VLANs.**

What is an “Emulated LAN”?

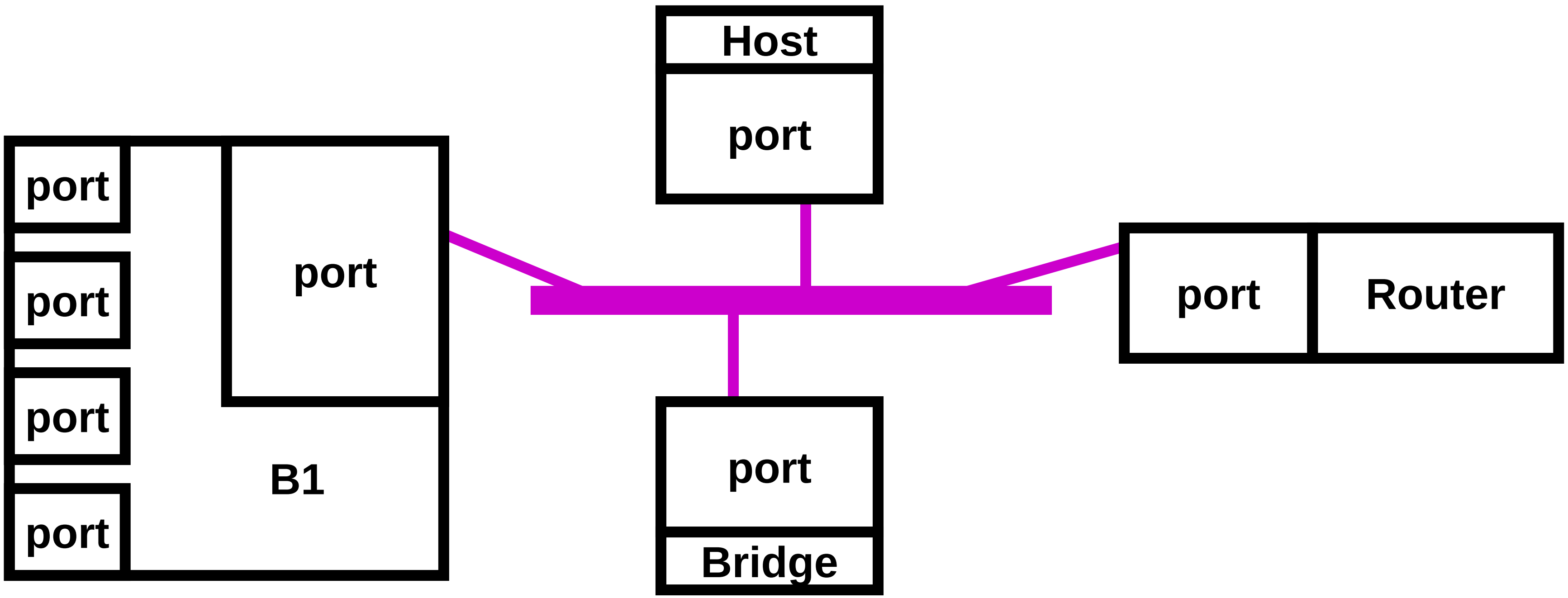
- **Examples:**
  - **ATM Forum LAN Emulation**
  - **IEEE 802.11 Wireless Access Point**
  - **IEEE 802.3ah Ethernet Passive Optical Network**
  - **IEEE 802.17b Spatial Reuse**
  - **IETF L2VPN Virtual Private LAN Service (VPLS)**
  - **IEEE P802.1ah Backbone Bridge (MAC-in-MAC)**
  - **IEEE P802.1?? Secure Hub (Phone + PC)**

What is an "Emulated LAN"?



- "Emulated LAN" is really a combination of point-to-point and point-to-multipoint links.

What is an “Emulated LAN”?



- **But it looks to the upper layers of each device like a shared medium: a good old fat yellow coax.**

WHY?

- **In some cases, the reasoning is:**
  - **I have a special point-to-point capability, a special multipoint capability, or both.**
- **Bridges do not understand my special capabilities, so I must either:**
  - **Emulate a shared medium LAN, so that bridges can understand me (hard); or**
  - **Accept the performance penalty of not using all of my capabilities (unacceptable).**

WHY?

- In other cases, the reasoning is:

  “I don’t understand bridges and don’t want to, or perhaps I just don’t like/understand/want Spanning Tree, so I’m going to do something similar to a bridge, perhaps with a way-cool new feature or two, only ***much*** simpler than a bridge!”

- The result: a chain of discoveries as the designers painfully reinvent bridges.

  Eventually, this “simple” solution is either
  a) even more complex than a bridge, or
  b) bypassed by the market.

WHY?

- Sometimes the reason is a better one: to offload the central bridging engine.
- If I can confine knowledge of a large class of MAC addresses to a particular Emulated LAN, I can relieve the central bridging engine of some load, and thus scale up the total number of MAC addresses my bridge can handle.

Solving the Problem

Two ways to solve the problem

- Create an Emulated LAN.

    And do it again and again and again.

    This is a very attractive solution, as witnessed by the number of times it has been employed.

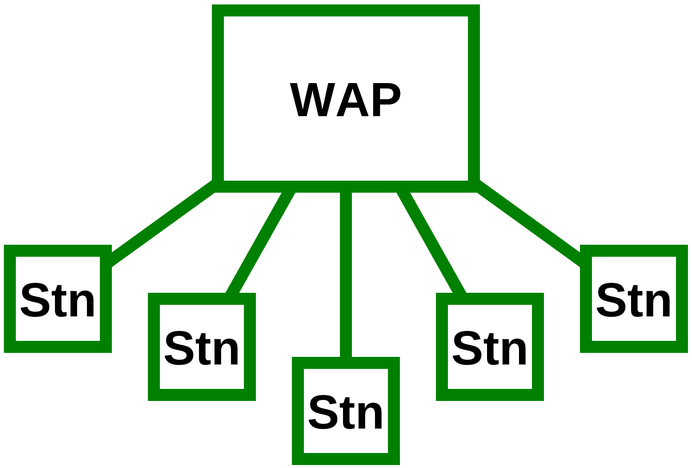
- Teach bridges to handle this common case.

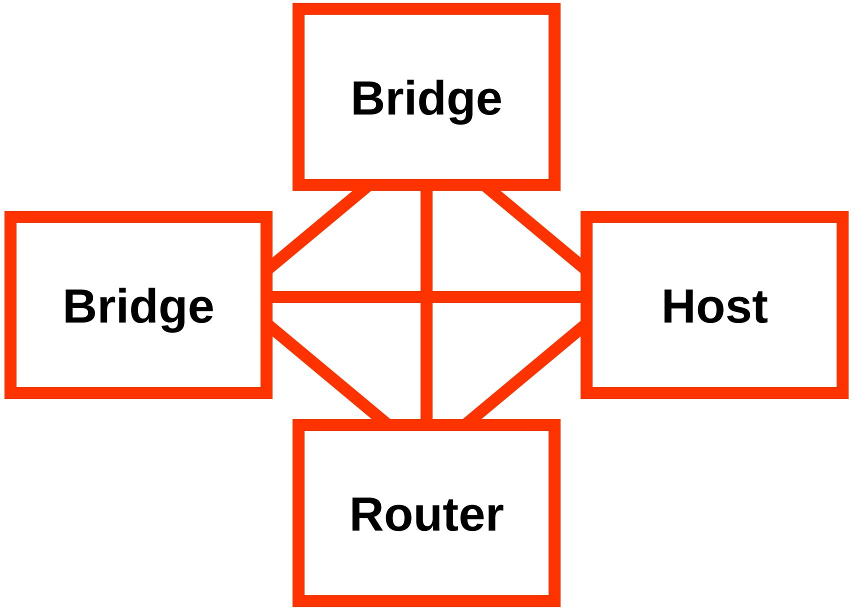
    Extend the standard Bridge Relay Function so that it can directly use the point-to-point and point-to-multipoint elements from which all "Emulated LANs" are built.

    And do it once.

Create a Generalized LAN Emulation Module in Bridges

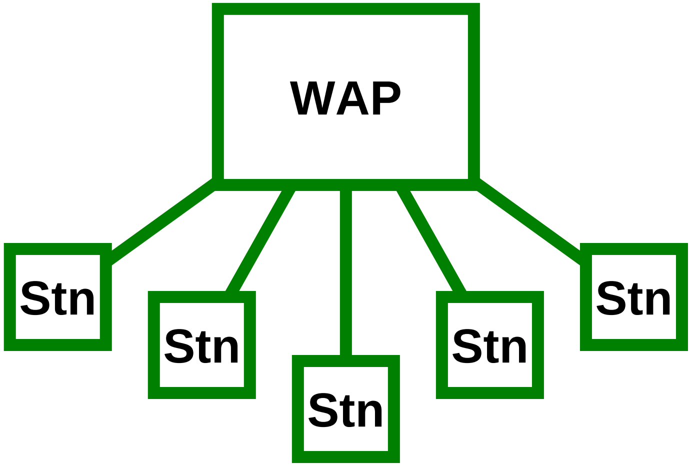
Observation

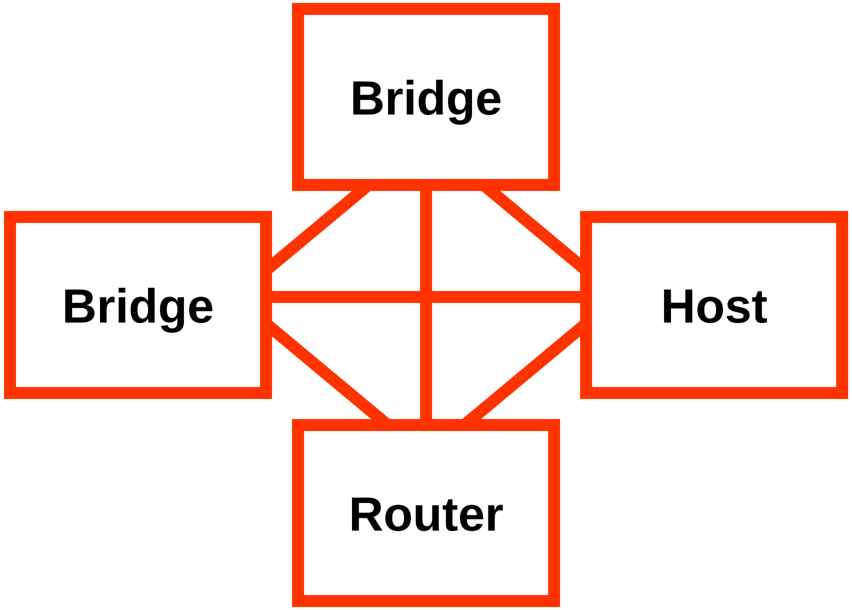




- **There are two very different models for "Emulated LANs": The Master/Slave model (EPON, 802.11) and the Full Mesh model (ATM LANE, VPLS).**

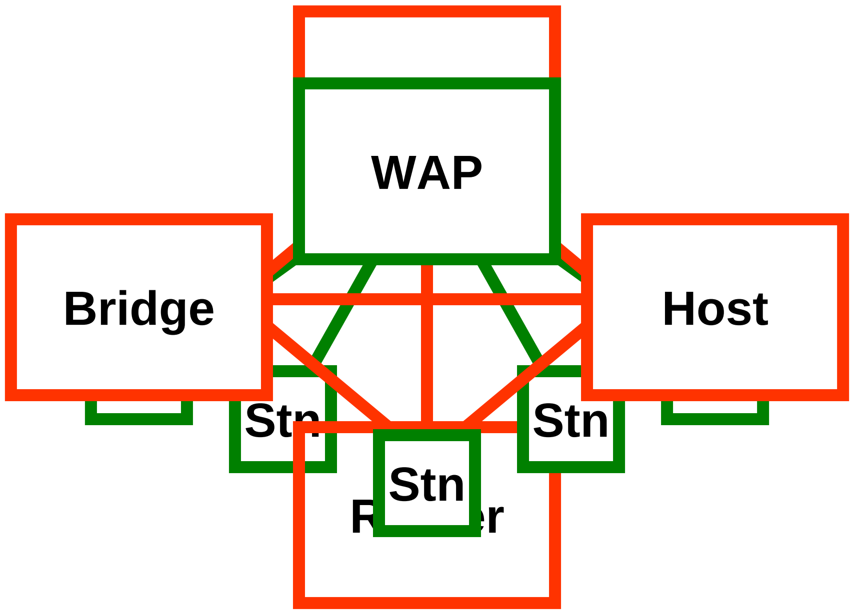
Observation





- The Master/Slave depends on the Master to relay all traffic among the Stations.
- The Full Mesh model never relays traffic inside the mesh (split horizon).

Observation



- **If you can integrate those two models, you can cover all such technologies!**
- **The remainder of this section describes the details of this integration.**

All “ELANs” have things in common.

- In every case, an “Emulation MAC” wants:

  To send frames for known unicast addresses to special point-to-point interfaces;

  To learn associations between source MAC addresses in received frames, and the special point-to-point interfaces on which they were received.

- And in many cases, it also wants to:

  Send frames for unknown unicast or multicast addresses to another special multicast interface.

A better way?

- So, suppose we teach the bridge’s Relay Function to deal with:
  - A special multicast interface serving a set of point-to-point interfaces; and
  - Different kinds of point-to-point interfaces?
- Then, we can leave MAC address learning to the Relay Function, which knows it best.

Generalized LAN Emulation: Port Associations and Portlets

- Two or more Bridge Ports in the same bridge may be associated together into a Port Association.
- Those associated Bridge Ports are then called, “Portlets”.
- A given Portlet belongs to exactly one Port Association.
- To avoid confusion, we reserve the term “Bridge Port” for non-associated ports.

Generalized LAN Emulation: Port Associations and Portlets

- **Some of the current behaviors of bridges with respect to their Bridge Ports apply directly to Portlets; the Port Associations become invisible to those behaviors.**
- **Other bridge behaviors apply to Port Associations; the Portlets become invisible to those behaviors.**

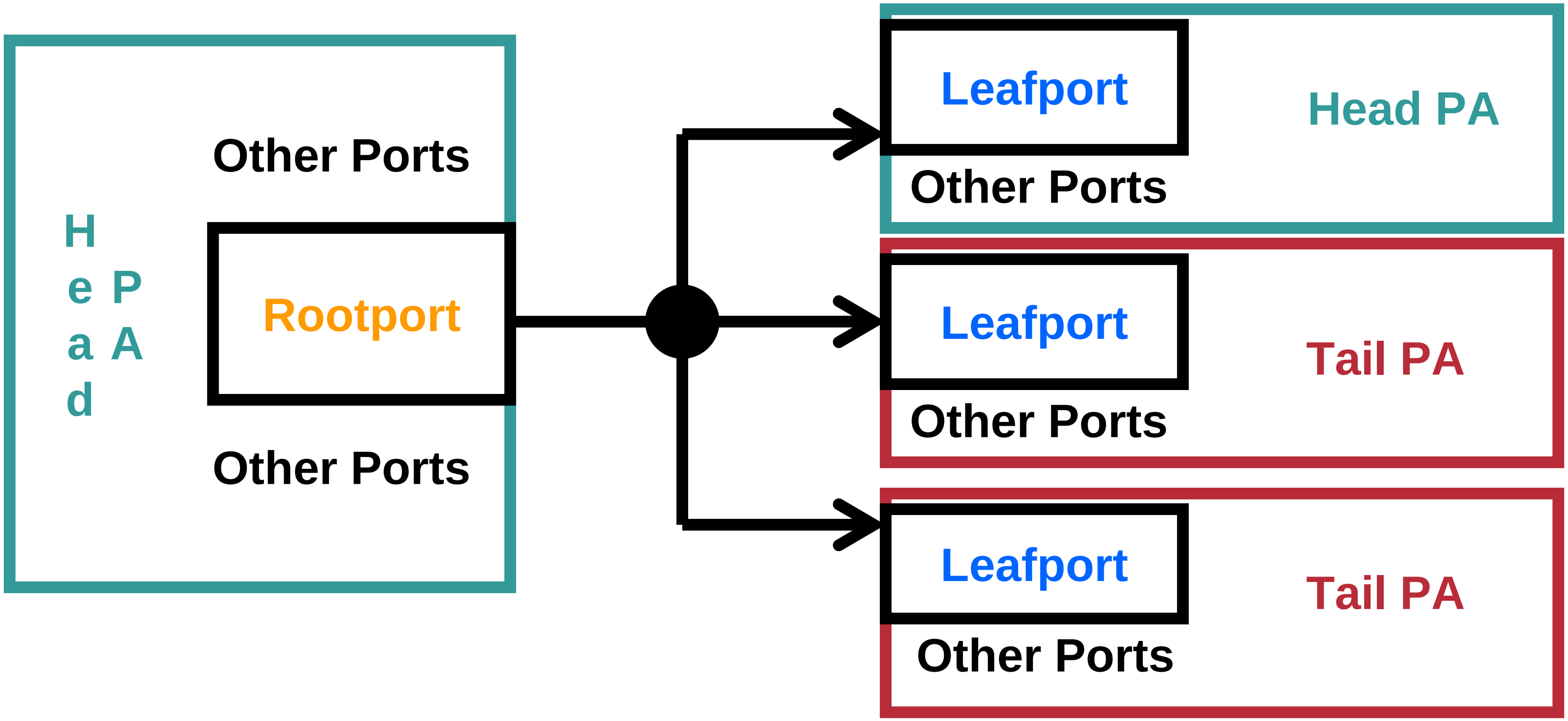
Generalized LAN Emulation: Port Associations and PA Identifiers

- **Every Port Association has a Port Association Identifier (PAID). The PAIDs must be unique over the extent of an “Emulated LAN”.**
- **So, they must either be globally unique, or the underlying medium must guarantee that interconnections that would cause confusion are impossible.**
- **There are two types of Port Associations: Head PAs and Tail PAs.**

Generalized LAN Emulation:
Headport Associations

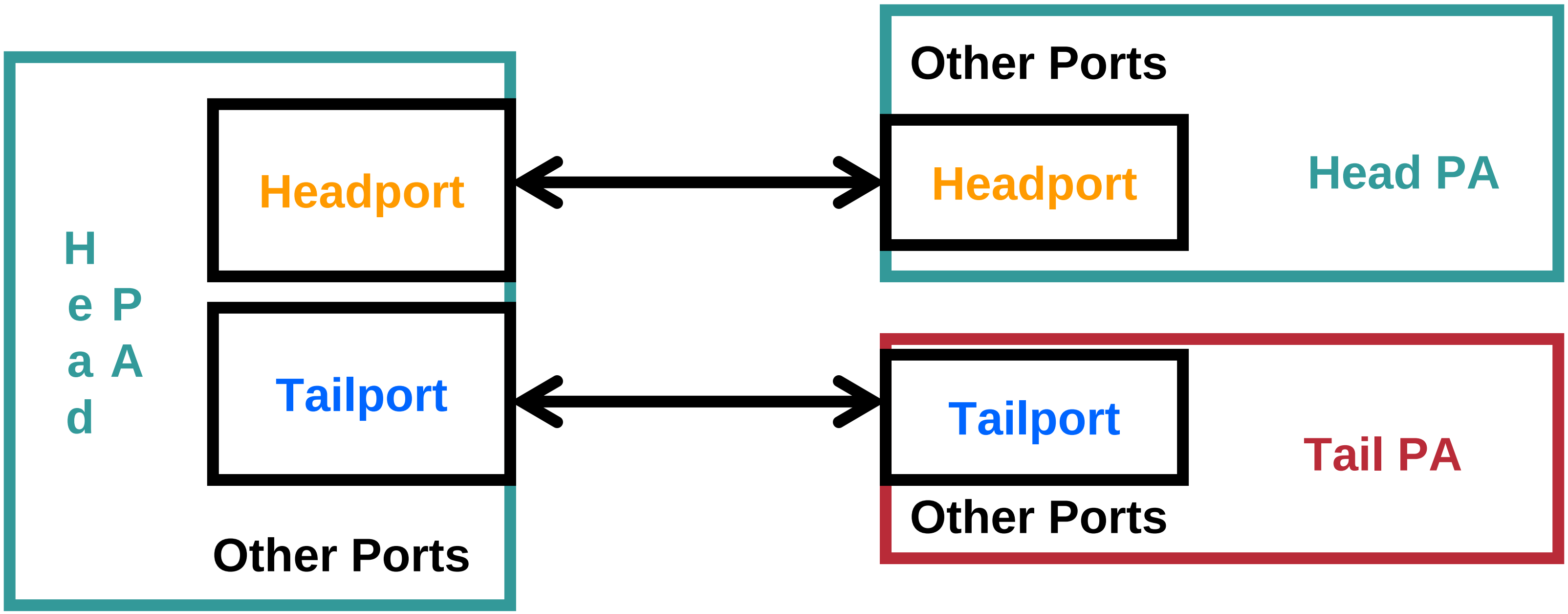
- **A Portlet in a Head PA is classified into one of four types:**
  - **A “Tailport” terminates a bidirectional point-to-point link (a “Taillink”) to a Tailport in a Tail PA.**
  - **A “Headport” terminates a bidirectional point-to-point link (a “Headlink”) to a Headport in another Head PA.**
  - **A “Rootport” takes frames outward to any number of Leafports in other Head PAs and/or Tail PAs via a unidirectional point-to-multipoint link.**
  - **A “Leafport” brings frames inward from a single Rootport in another Head PA via a unidirectional point-to-multipoint “Rootlink”.**

Generalized LAN Emulation



- **Rootports and Leafports.**
- **Unidirectional Rootlink.**

Generalized LAN Emulation



- **Headports and Tailports.**
- **Bidirectional Headlink and Taillink.**
- **There are no Tail PA–to–Tail PA links.**

Generalized LAN Emulation: Headport Associations

- A **Head PA** has:

  No more than one Tailport or Headport link to the same PAID.

  No more than one Rootport, to which is linked a single Leafport for each PAID to which the Head PA has a Tailport or Headport link.

  For each Headport link to another Head PA, at most one Leafport linked to the Rootport in that same Head PA.

Generalized LAN Emulation: Tailport Associations

- A Portlet in a Tail PA is classified into one of two types:
  - A “Tailport” terminates a bidirectional point-to-point link to a Tailport in a Head PA.
  - A “Leafport” brings frames inward from a single Rootport in a Head PA via a unidirectional point-to-multipoint link.
- A Tail PA has:
  - Exactly one Tailport linked to a Rootport in a Head PA.
  - At most one Leafport linked to that same PAID.

Generalized LAN Emulation:
Head PAs vs. Tail PAs

- **Note that “Head” and “Tail” are <u>not</u> synonymous with “Bridge” or “Station”**

Generalized LAN Emulation: Head PAs vs. Tail PAs

- **The Head PAs must have a full mesh of Headports.**
- **The Head PAs <u>may</u> have a full mesh of point-to-multpoint Rootport-to-Leafport links.**
- **Note that these full meshes are difficult to ensure when the underlying medium is point-to-point, and when there are more than one or two interconnected Head PAs.**
- **A Tail PA is linked to exactly one Head PA.**

Generalized LAN Emulation: PAID Marking and Learning

- **The frame format used on the point-to-multpoint links must include a provision for carrying the PAID <u>in addition</u> to the standard IEEE 802.1D ISS MAC addresses.**
- **Source MAC addresses received by a Bridge on a Portlet are learned according to the standard rules, except that frames received on a Leafport are learned as if received from the Headport connected to that same PA.**

Control Protocols

Generalized LAN Emulation:
Spanning Tree BPDUs

- **As far as the Spanning Tree algorithms are concerned, the Port Association is one big Bridge Port.**
- **Identical BPDUs are transmitted on either:**

  The Rootport; or

  All Tailports and Headports. (Different Hello Timers may be used in this case.)

Generalized LAN Emulation: Spanning Tree BPDUs

- **BPDUs received on Tailports are *both*:**

  Accepted by the receiving Bridge; and

  Relayed just like a normal multicast among the Portlets of the Port Association (but of course, not to other Bridge Ports or other Port Associations).

  The Headport may be used for relaying BPDUs.

- **BPDUs received on Headports or Leafports are accepted by the receiving Bridge and applied to the whole Port Association, but are never relayed.**

Generalized LAN Emulation: Forwarding Rules

- **A frame received on a Headport or Leafport must never be forwarded on any Portlet belonging to the same Port Association.**
- **A frame received from a Tailport and forwarded to the Rootport of the same Head PA is marked with the identity of the Tail PA from which it was received.**
- **All other frames transmitted to the Rootport are marked with the identity of the Head PA.**
- **A frame receive from a Leafport marked with the identity of the receiving PA must be discarded.**

Generalized LAN Emulation: GMRP, GVRP, IGMP, etc.: “Smart distribution”

- Smart distribution (via IGMP, GVRP, etc.) of broadcasts, multicasts, and unknown unicasts is handled in one of two modes, PA Mode or Individual Mode, as determined by configuration.
- In PA Mode:

  A single database of distribution state information is maintained for all of the Portlets in the PA.

  All received control information is handled as if it arrived on the Rootport.

  All multicasts and unknown unicasts are transmitted either to the Rootport, if present, or to all Tailports and Headports, if not.

Generalized LAN Emulation: GMRP, GVRP, IGMP, etc.: “Smart distribution”

- **In Individual Mode:**

  A separate database of multicast and VLAN distribution state information is maintained for each of the Headports and Tailports in the Port Association.

  All control protocol information received from a Leafport is handled as if it arrived from the corresponding Headport or Tailport.

  Control protocol transmissions tied to these databases are always transmitted on the corresponding Tailport or Headport.

  Propagation of information follows rules for data forwarding, e.g. GVRP registration received on a Headport is not relayed to other Headports, but is relayed to Tailports.

Generalized LAN Emulation: GMRP, GVRP, IGMP, etc.: “Smart distribution”

- **In both modes:**

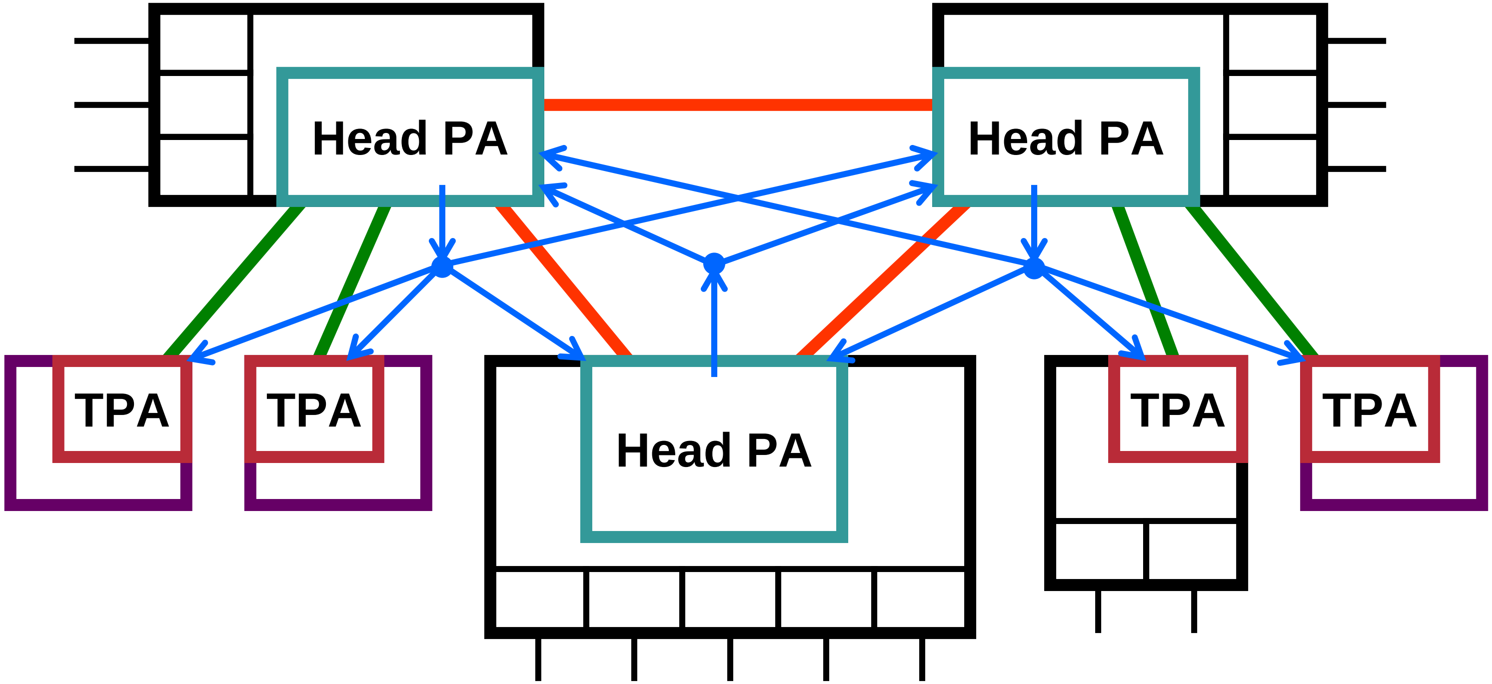
  “Access” Tailports may be exempted from the transmission of unknown unicast frames.

  Heuristics may be employed to transmit frames *(other than the multicast control protocols)* through the Rootport, if the nature of the underlying medium makes this more efficient than transmitting frames through multiple point-to-point links.

  The Rootport and its point-to-multipoint link is an optional optimization for each individual Head PA.

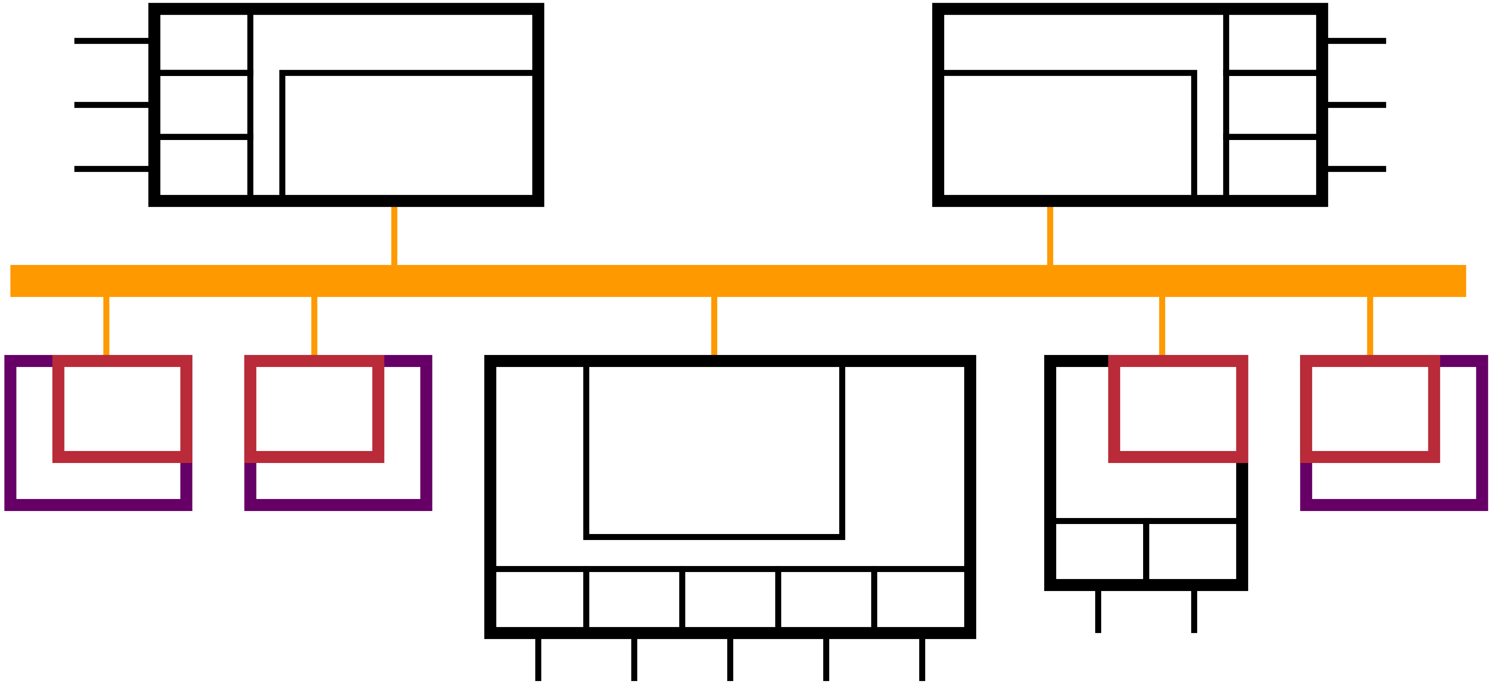
1000 Words

A small "Emulated LAN"



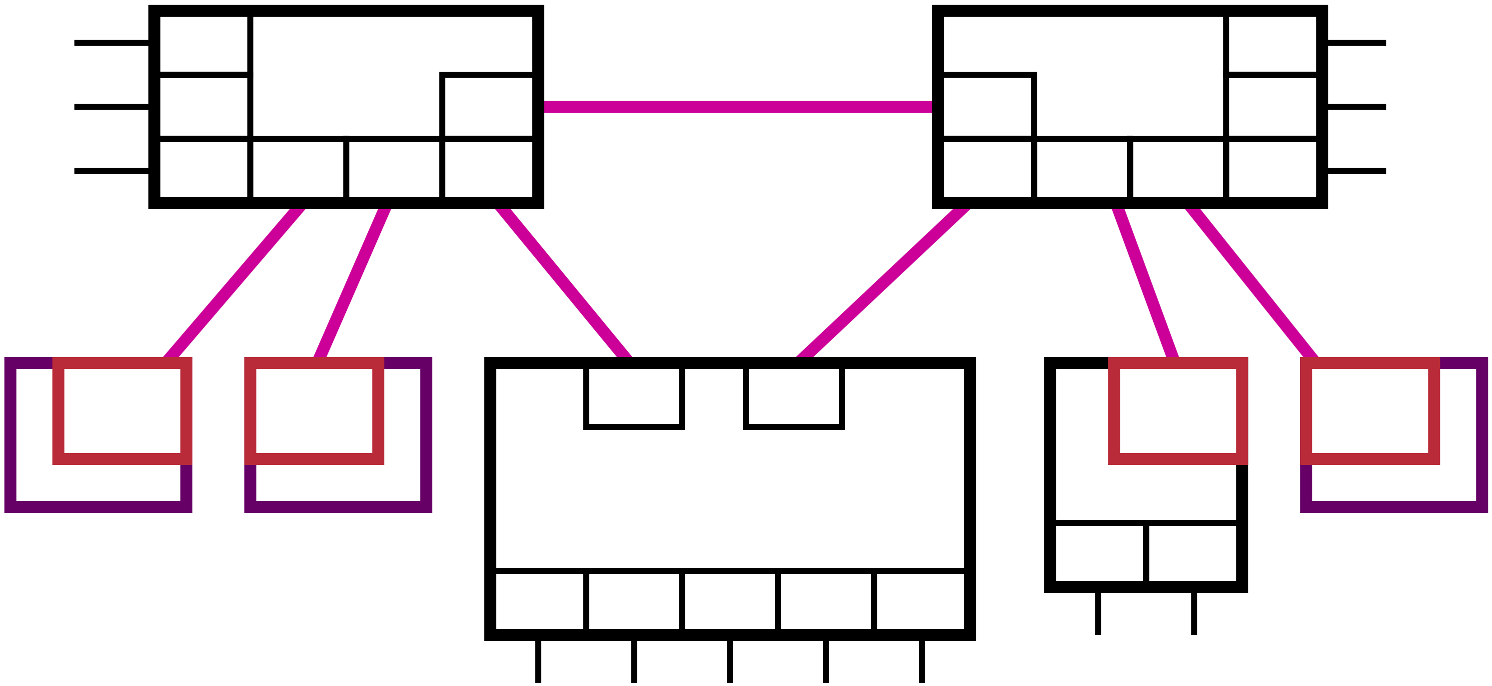
- **Bridges, Endstations, Head Port Associations, Tail Port Associations, Headlinks, Taillinks, and Rootlinks.**

A small “Emulated LAN”



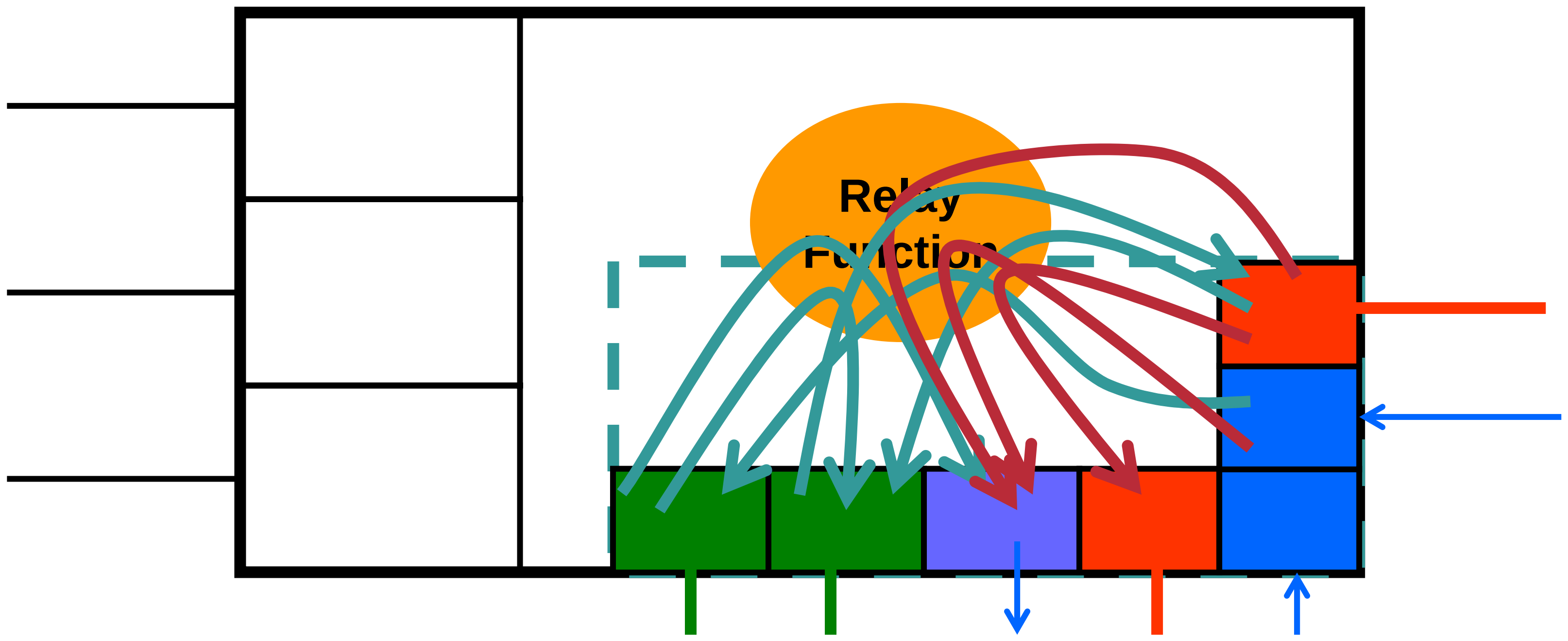
- To BPDUs or GARP in PA Mode, this looks like a Fat Yellow Coax.

A small “Emulated LAN”



- To GARP in Individual Mode, this looks like a lot of point-to-point links.

Head PA Forwarding Rules



- **Some forwarding is allowed, some is not allowed.**

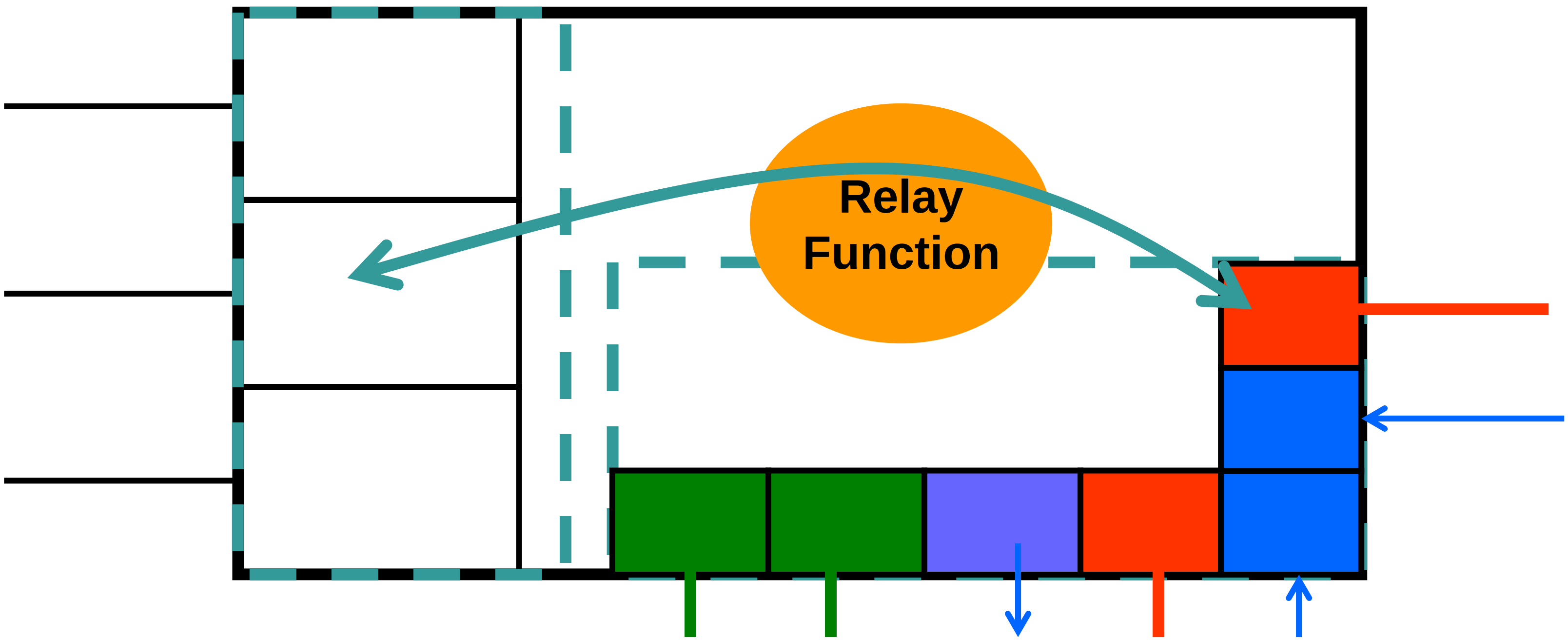
Head PA Forwarding Rules

TO

FROM

	Root	Head	Tail	Other
Leaf	NO	NO	YES	YES
Head	NO	NO	YES	YES
Tail	YES	YES	YES	YES
Other	YES	YES	YES	normal bridging

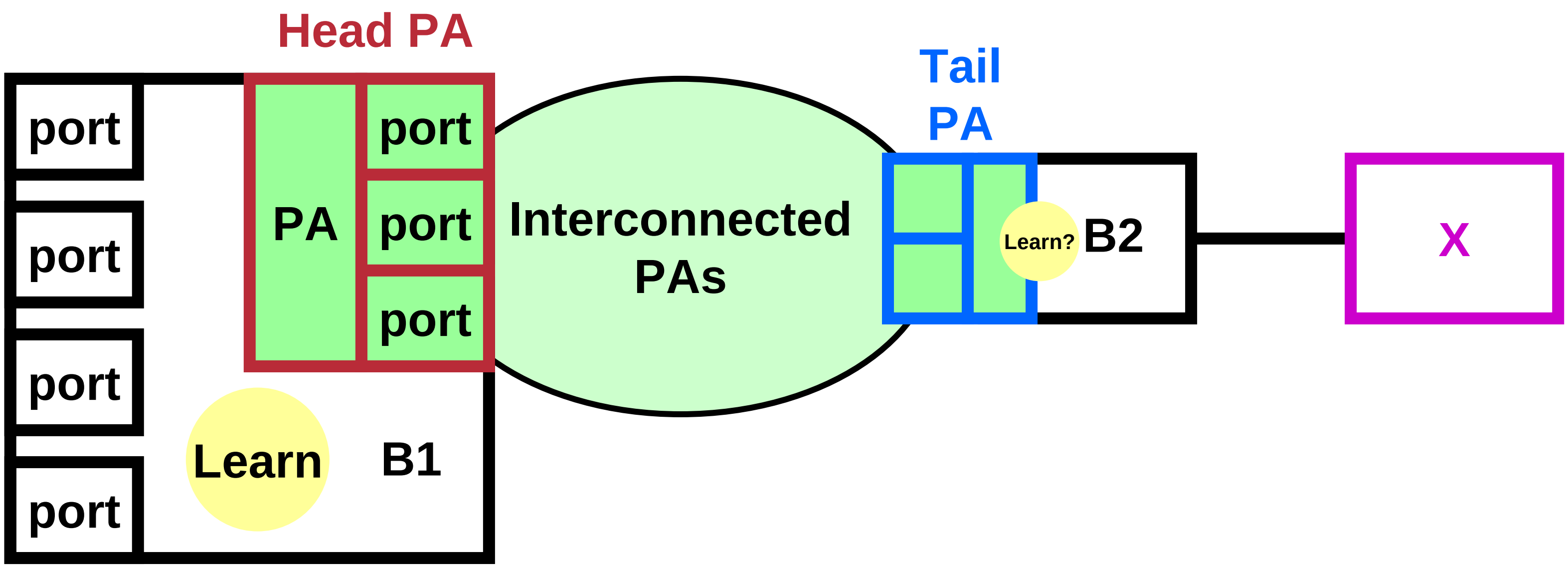
Multiple Port Associations?



- **To any given PA, all other PAs are ordinary Ethernet ports, and the “other” columns in the table apply.**

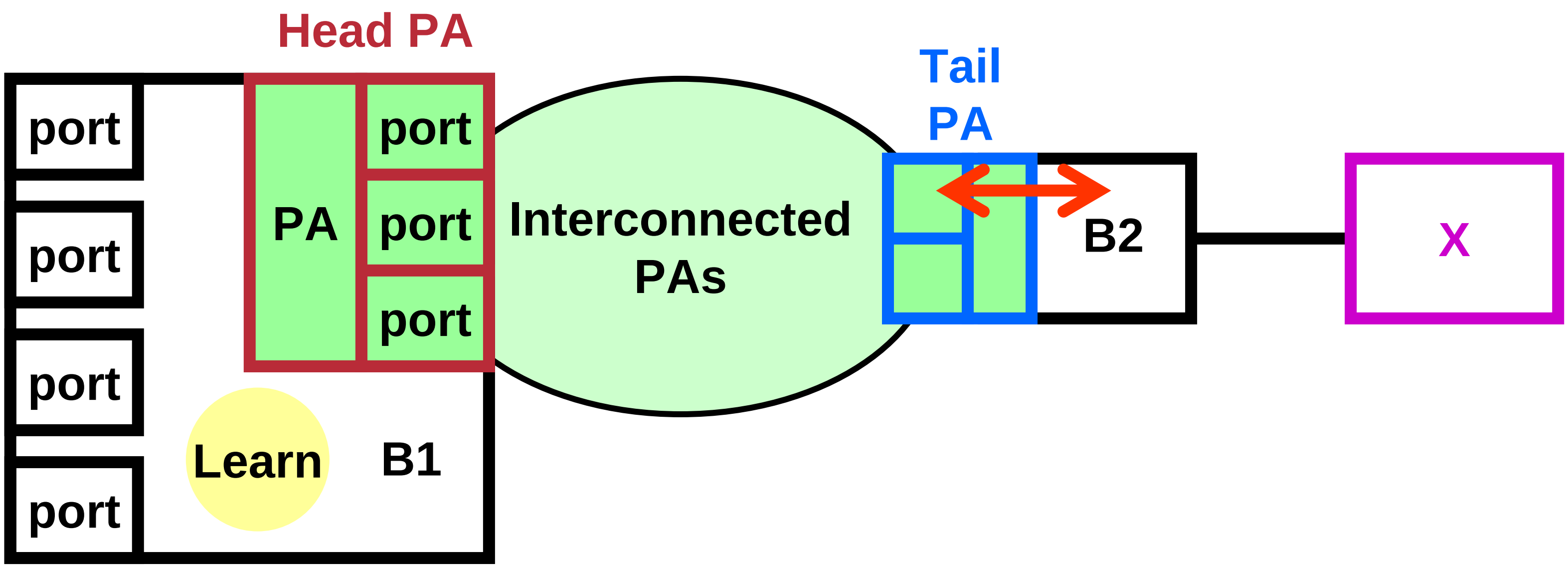
LAN Emulation vs. Port Associations

LAN Emulation or Tailport Association?



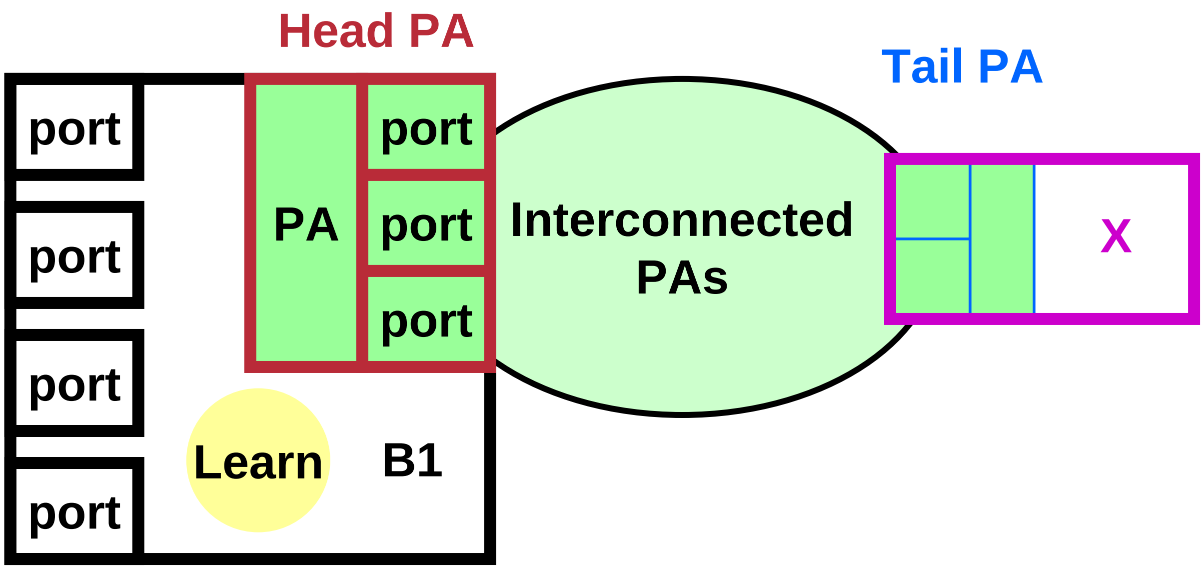
- **Two Bridges with Port Associations, one Head PA and one Tail PA.**
- **The Tail PA is connected to Endstation X.**

LAN Emulation or Tailport Association?



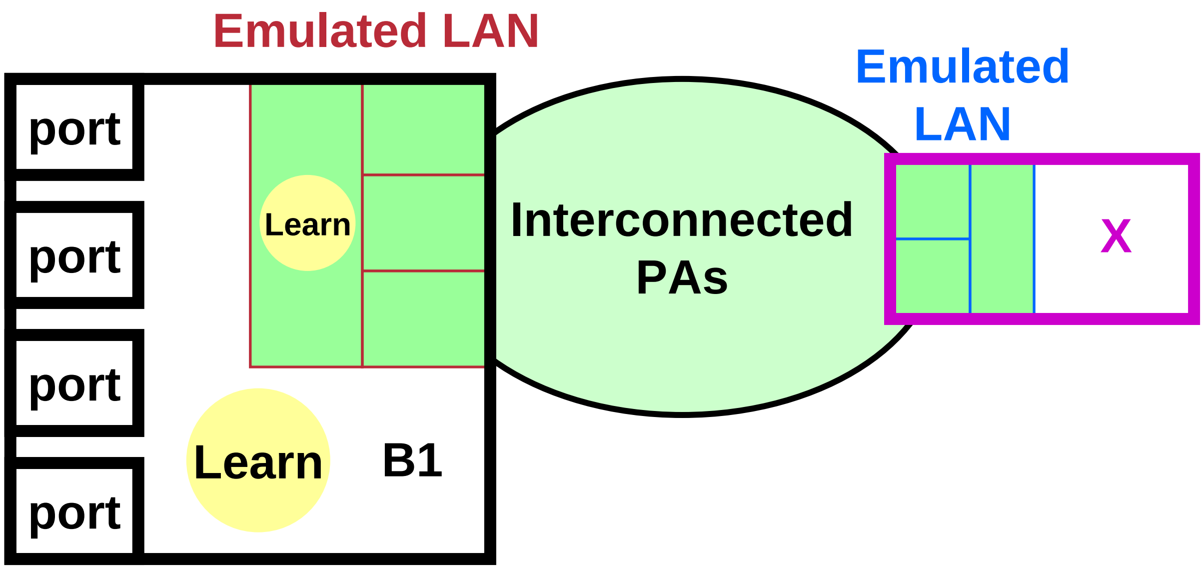
- **Bridge B2 doesn't care on which of the Tailport or Leafport it receives frames.**
- **It never has to learn any MAC addresses.**

LAN Emulation or Tailport Association?



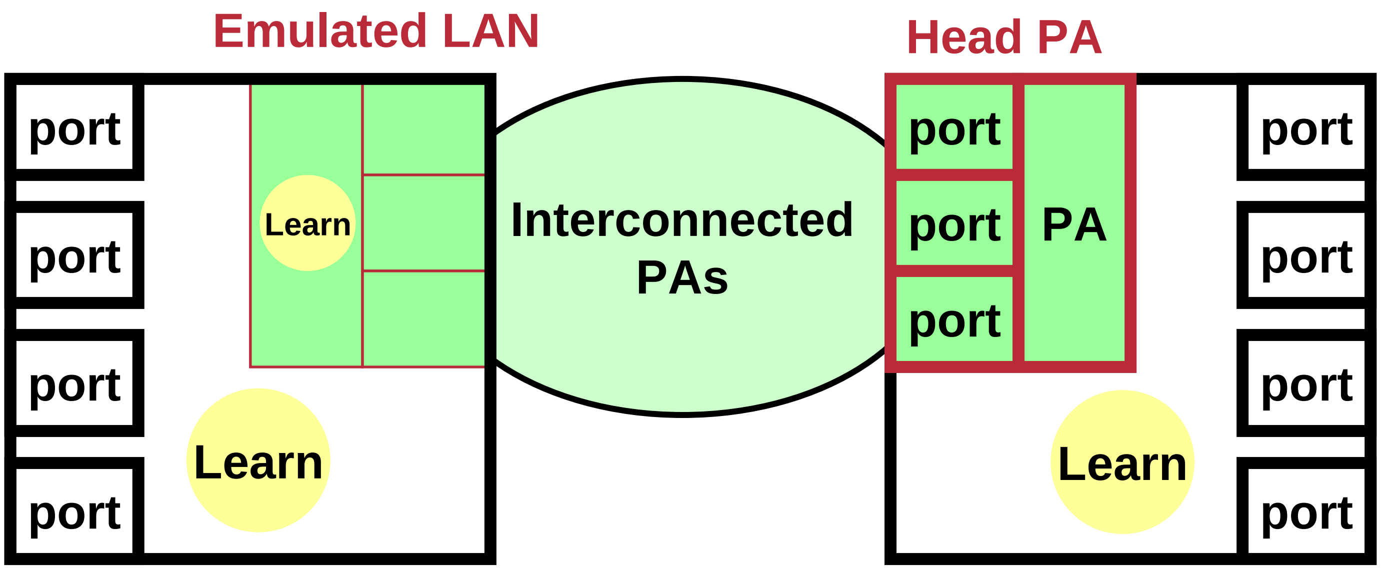
- In fact, if you collapsed X into B2, it would work just fine.
- Looks rather like LAN Emulation!

LAN Emulation or Port Association?



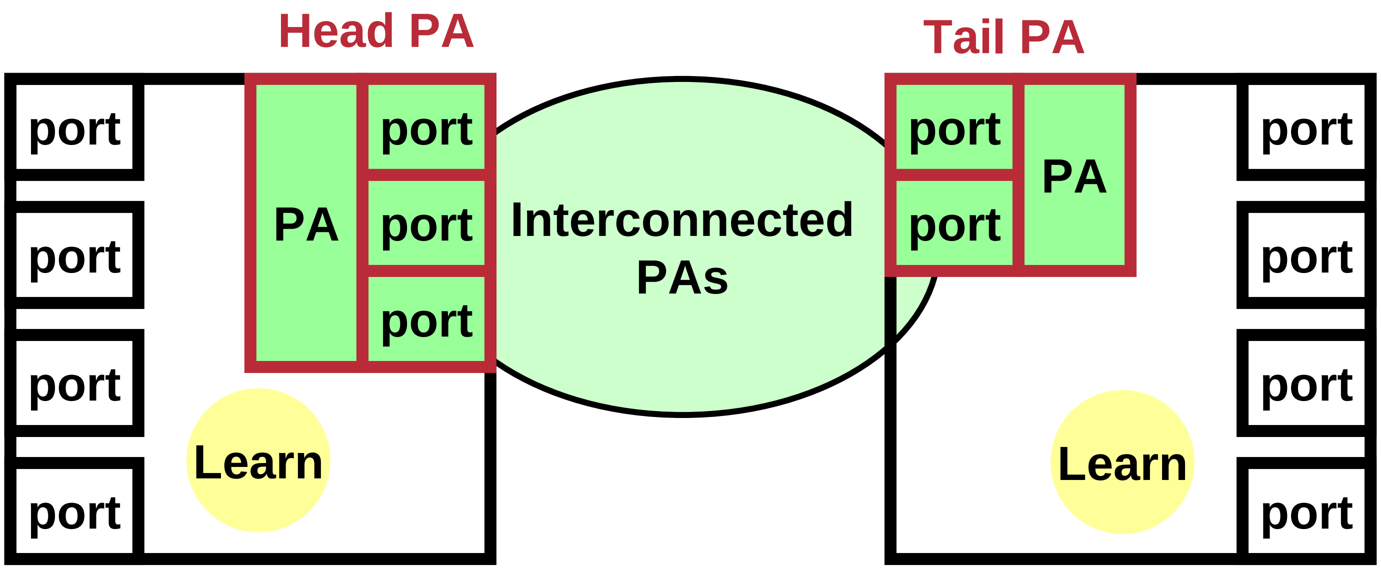
- And, as the handling of BPDUs might make one suspect,
- The only difference between Port Associations and “classic” LAN Emulation is the placement of the MAC Learning function!

LAN Emulation or Headport Association?



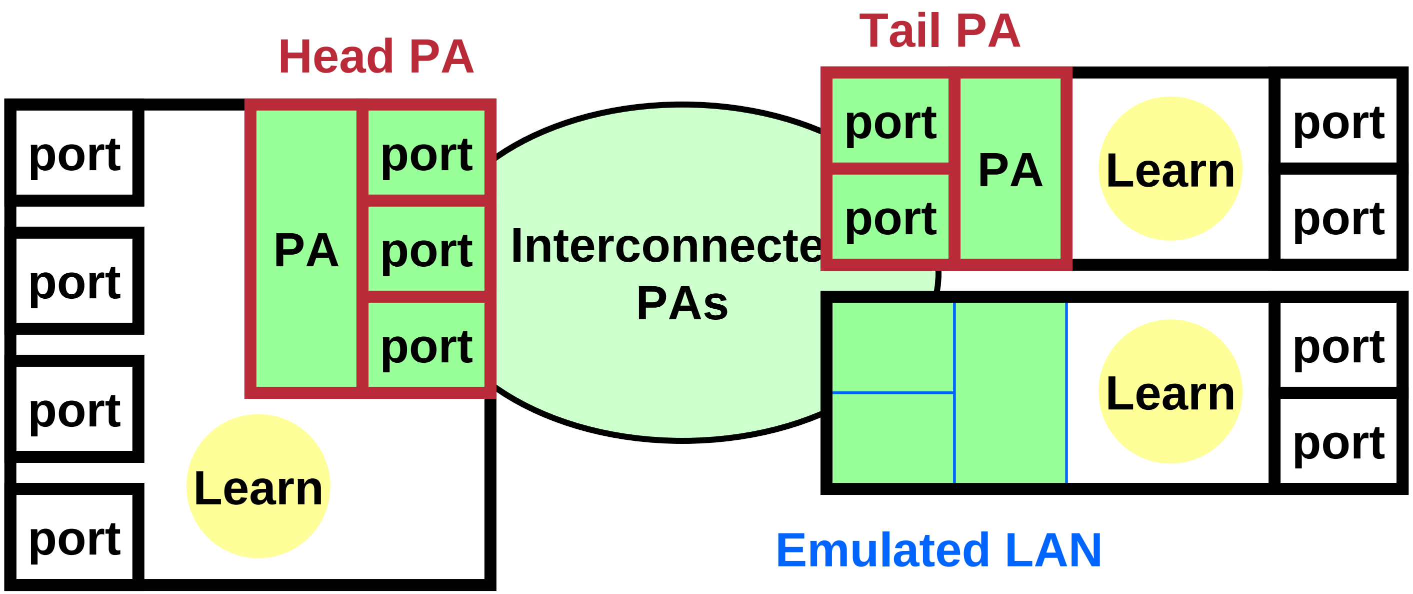
- Classic LAN Emulation has two MAC Learning Functions.
- A Headport Association needs only one.

Tailport Associations and Bridges



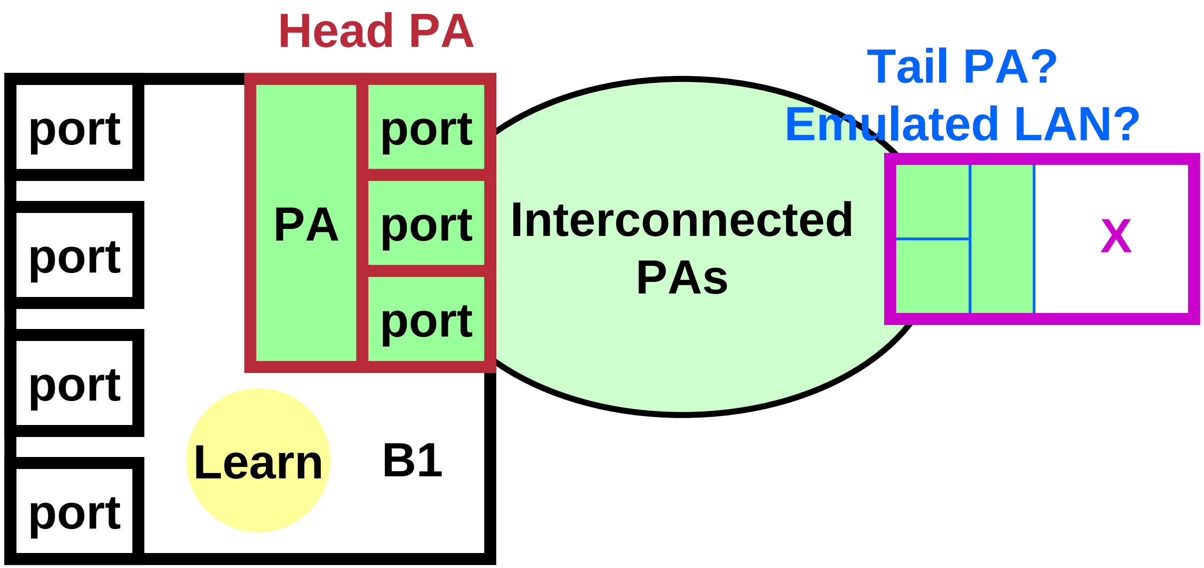
- A Bridge may have a Tailport!
- But, some bridge on the Emulated LAN must have a Headport.

Tailport Associations and Bridges



- In a Bridge, a **Tail PA is indistinguishable from an Emulated LAN port**, because the Tail PA needs no additional Learning function.

LAN Emulation or Port Association?



- **A Tailport is particularly suited to an Endstation, because no Learning function is needed, at all.**

Mapping the existing (and planned) emulations to Generalized LAN Emulation

ATM LAN Emulation

- Every node in the ELAN is a Head PA.
- A full mesh is created and maintained dynamically using the ATM LAN control protocols.
- There are no Tail PAs nor Tailports.
- The Rootport is the point-to-point link to the BUS.
- The mesh of point-to-multpoint links is emulated via the BUS, and a single point-to-multpoint link.

802.11

- **WAPs have Head PAs, Stations Tail PAs.**
- **MAC addresses serve as PAIDs.**

  The four-address format is required to allow a Bridge to be a Station.

- **Inter-WAP channels are ordinary point-to-point Bridge Ports, not PAs, and are subject to normal spanning tree rules.**

EPON

- The IDs in the preamble serve as PAIDs.
- There is one Head PA (the OLT) and some number of Tail PAs (the ONUs).
- Heuristics to optimize multicast transmission are clearly in order, as are the suppression of unknown unicast transmissions.
- The Rootport is not used by a bridge at present, but should be.

802.17b Spatial Reuse

- Every node is a Head TA.
- The full mesh is implied by ring connectivity.
- Non-SR nodes:

  All frames sent Rootport.

- Non-bridge SR nodes:

  Filtering Database determined by ring exploration.

  PA Mode used for multicast and unknown unicast.

- Bridge SR nodes:

  Filtering Database determined by learning.

  Either PA Mode or Individual Mode is acceptable.

VPLS

- Every node is a Head PA, so a full mesh is both required, and difficult to guarantee.
- There is no Rootport, so every multicast, broadcast, or unknown unicast must be sent via replication on all Headports.
- Since all links are Headports, intra-PA forwarding is disallowed. (This is also called, “split horizon”.)

Backbone Bridges (MAC-in-MAC)

- (Encapsulate-on-egress form.)
- Every MAC-in-MAC node is a Head PA.
- The M-M nodes' MAC addresses are the PAIDs.
- Outer destination multicast addresses are used to form the point-to-multpoint networks.
- The backbone network provides a shared medium, so the full mesh comes for free.

Secure Hub (Phone + PC)

- This will make a Bridge Port attached to a hub operate much like an 802.11 WAP.
- One (or two bridges) could be Head PAs, the rest (if any) Tail PAs. One is easier to set up, two quicker to failover.
- MACsec in null mode provides the PAIDs.
- Rootports may be allowed, but if present, would reduce the level of security.

What is left for "Emulated LAN" designers?

Still plenty of details to solve

- **Obvious things:**
  - **Encapsulation formats.**
  - **Rootlink topology, e.g. broadcast server vs. mesh of point-to-multipoints.**
  - **Creating backup Head PAs for important Tail PAs, and switching over when appropriate.**

Still plenty of details to solve

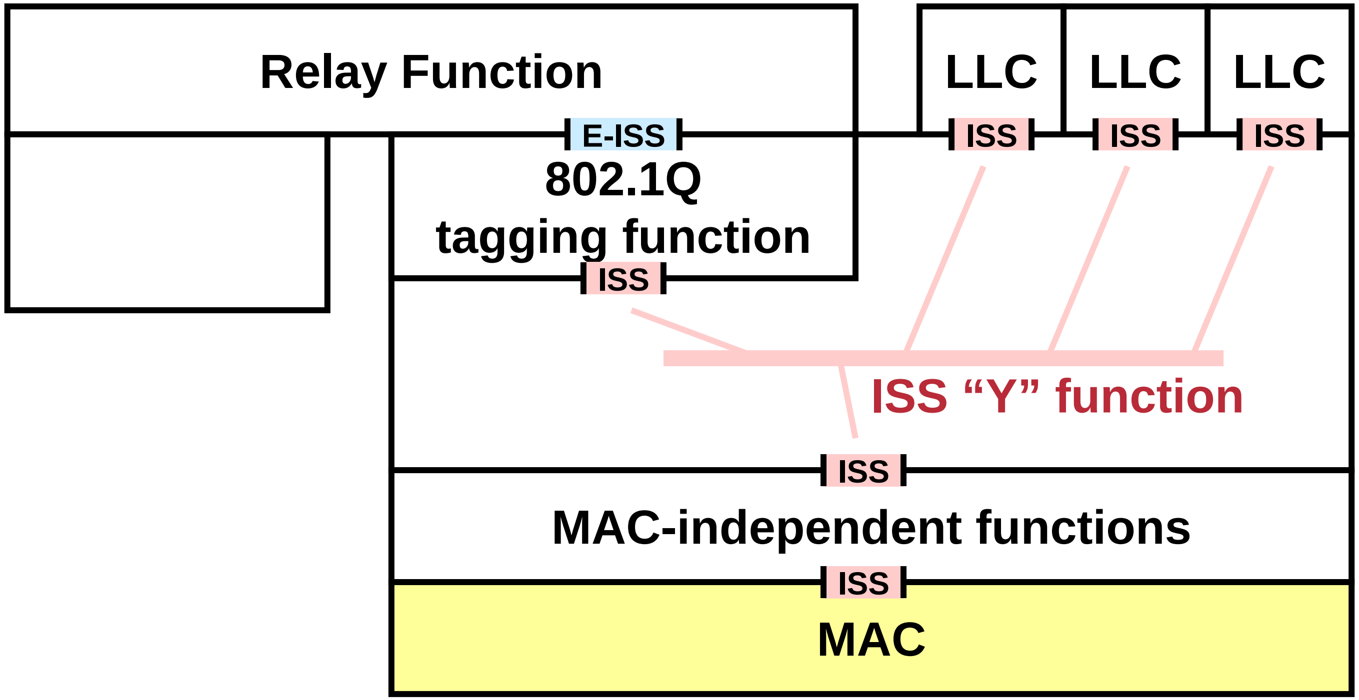
- **Not-so-obvious things:**
  - Frame ordering, e.g. replicate unknown unicasts or use "flush" mechanism on Rootlink.
  - Either guarantee full mesh among Head PAs, or create Headlinks dynamically and use static Rootlinks.
  - Guarantee that unidirectional Headlinks and Taillinks cannot persist.
  - Either ensure that a Tail PA can identify and discard its own multicasts received on the Leafport, even if the Tail PA is part of a bridge, or use Tailports to distribute multicasts from other Tailports.

What Emulated LANs do not need to solve

- The relationships between MAC addresses and VLANs.
- The relationships between Emulated LANs and the Spanning Tree Protocols.
- When to forget learned MAC addresses.
- How to signal VLAN distribution.

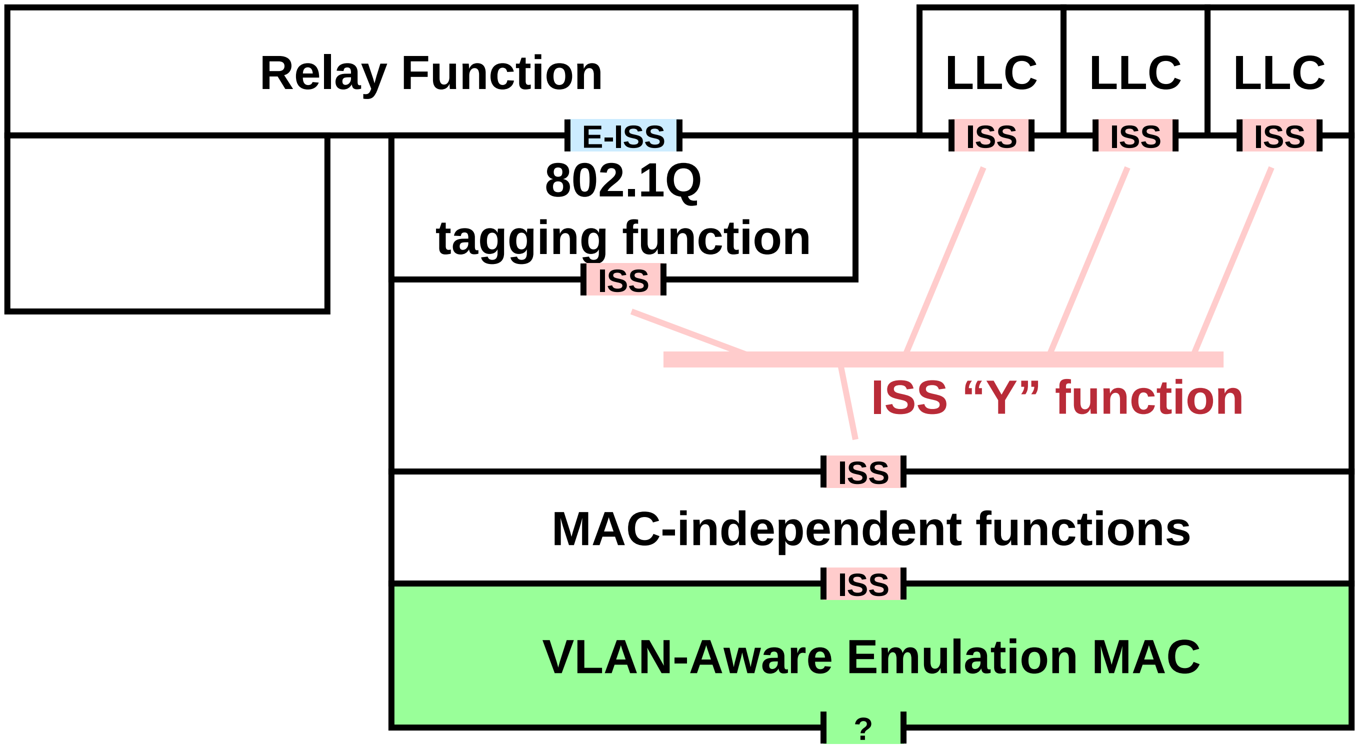
Bridge Architecture for VLAN-Aware Emulation MACs

Baggy Pants Diagram



- This is the current P802.1Q-REV model.

Baggy Pants Diagram



- Can't put the VLAN-Aware Emulation MAC here, because it can't see the VLAN info.

So, what do we do?

- Given the number of “Emulated LANs” invented, they are not difficult to build using current technology.
- There are very good reasons (as well as bad) for using an Emulated LAN design, rather than an integrated design, in one’s bridge.
- The bridge architecture in IEEE 802.1 standards should support this and show how it can be done.

Summary

Summary

- **Adding “Generalized LAN Emulation” rules to IEEE 802.1Q, integrating ELANs into the Relay Function, would most of the fun out of reinventing bridges via LAN Emulation, but still do the world a great service.**
- **But, the “Emulated MAC to an Emulated LAN” model must also be supported.**

Cisco Systems