Trunking Study Group

Ariel Hendel

hendel@eng.sun.com

Sun Microsystems

4-February-1998

Goals

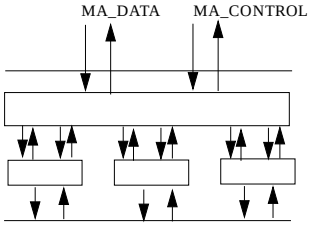
- Some broad goals:
  1. Specify trunk configuration and failover
  2. Support all four combinations of host/bridge endpoints
  3. Host trunk must cover at least single MAC address per trunk model (in order to interoperate properly with ARP and resemble a single link to the outside world)
  4. No changes below MAC service interface (i.e. allow S/W implementations with existing NICs)
  5. Leverage 802.3x MAC Control infrastructure
- Goals for load balancing:
  1. Satisfy packet ordering constraints
  2. Allow flexibility to optimize

Architecture

- Add the trunking sub-layer above the MAC Control Sublayer.

- Use the MA_CONTROL.request / MA_CONTROL.indication for link initialization and failover.

- A single instance of the trunking sub-layer serves as client to all MAC Control sublayers in a trunk.



Ramifications:

1) The MAC Control sublayer becomes mandatory for trunking.

2) PAUSE frames affect a single trunk link (good).

3) Link initialization and failover uses MAC Control frames.

Load Balancing

The appeal of link aggregation trunks depends on the ability to load balance traffic across different links.

For example, trunking bridges would load balance while preserving the following external properties:

1. Unless load balancer “knows better” no frame misordering for a given priority level between a given MAC source and destination.

2. No frame duplication.

Additional endpoint and state information may be available at the server to support a more sophisticated downstream load balancing.

Each end of the trunk may **independently** attempt to load balance across the segments using information from a given layer(s), while remaining transparent to protocols operating above these layers.

A load balancer that “knows better” would be one that has upper layer knowledge / state and can safely relax the ordering rule.

Flexibility to optimize

- Unreasonable to specify a load balancing algorithm when the optimization metric, the application, and the equipment capabilities are unknown.
- Interoperability does not require the algorithm to be the same or even known by:
  - **the receiving side**
  - **the reverse direction load balancer**

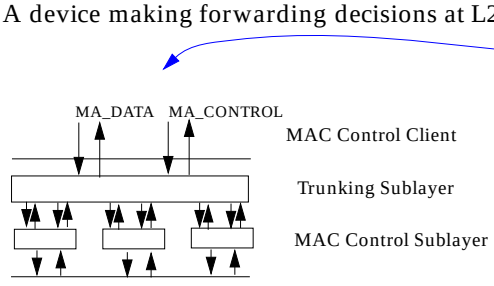
On the other hand there is a comfort factor in specifying boundaries for what is allowed.

A possible approach:

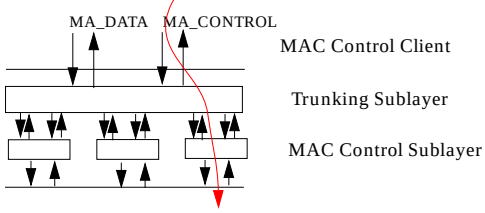
Spell out the L2 properties but retain access via the trunking layer to the individual MAC service interfaces.

Example

A device making forwarding decisions at L2 can safely use this interface



While a device with higher layer functionality may direct packets to a specific MAC instance of the trunk



Summary

- Incorporate behavior rules for load balancing into the trunking sublayer (see slide 4, and examples in Paul Congdon's presentation)

802.1d bridging for example may forward packets as usual via the MA_DATA portion of the MAC Control Client Interface and benefit from the trunking layer load balancing function.

- Allow direction of packets through the trunking sublayer to specific MACs.

Additional information ("Flows", RSVP session, connection, etc.) may be used for load balancing by devices possessing such information.