DetNet Data Plane using PseudoWires

Jouni Korhonen
Shahram Davari
Norm Finn

IETF#94,
Yokohama

Background

- DetNet QoS is expressed in terms of:
  - Minimum and maximum end-to-end latency from talker to listener;
  - Probability of loss of a packet, assuming the normal operation of the relay systems and links;
  - Probability of loss of a packet in the event of the failure of a relay system or link

This presentation reuses draft-finn-detnet-architecture-01 in many places!

Background cont’d

- Three techniques are employed by DetNet to achieve these QoS parameters:
  - Zero congestion loss. Network resources such as link bandwidth, buffers, queues, shapers, and scheduled input/output slots are assigned in each relay system to the use of a specific DetNet stream or group of streams.
  - Pinned-down paths. Point-to-point paths or point- to-multipoint trees through the network from a talker to one or more listeners can be established, and DetNet streams assigned to follow a particular path or tree.
  - Packet replication and deletion. End systems and/or relay systems can sequence number, replicate, and eliminate replicated packets at multiple points in the network in order to ensure that one or more equipment failure events still leave at least one path intact for a DetNet stream.

Data Plane for DetNet

- Citing the charter:
  - “This work will document how to use IP and/or MPLS to support a data plane method of flow identification and packet forwarding over Layer 3.”

- Proposal:
  - Leverage PseudoWires for Detnet Data Plane.

What there is in the toolbox?

- Zero congestion loss: checked
  - Deterministic resource allocation using multiple existing mechanisms.
- Pinned-down paths: checked
  - E.g. statically provisioned label paths.
- Packet replication and deletion: almost-there
  - Assuming no congestion loss the most important causes of packet loss are random media and/or memory faults and equipment failures
  - How to do IEEE 802.1CB or ISO/IEC 62439-3 like Seamless Redundancy over IP/MPLS?

Seamless Redundancy

- Seamless redundancy involves three capabilities:
  - Adding sequence numbers to the packets of a DetNet stream.
  - Replicating these packets and, typically, sending them along at least two different paths to the listener(s).
  - Discarding duplicated packets.

- Note that data streams can be multicast or unicast.

- Apart from ’Discarding part’ needed bits & pieces seem to be there in current MPLS & PseudoWire toolbox.

Seamless Redundancy approaches

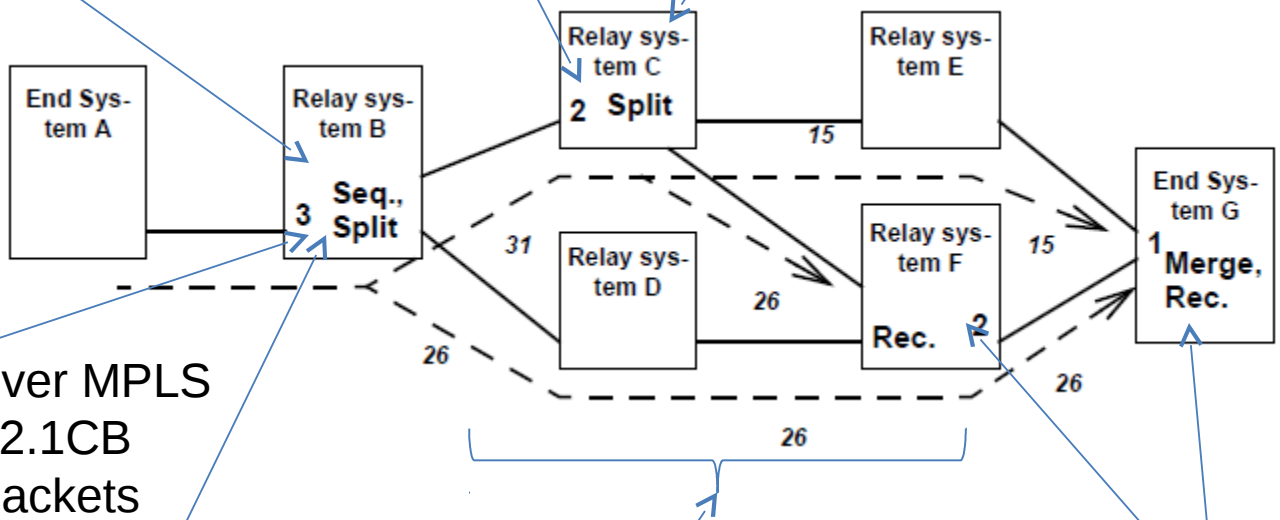
- Three cases to cover:
  - (T-)PE implements 802.1CB.
  - (T-)PE provides ”stream split”, ”merging and recovery” and ”stream identification” for non-802.1CB Ethernet traffic.
  - (T-)PE provides ”stream split”, ”merging and recovery” and ”stream identification” for any IP traffic.

Mapping Seamless Redundancy to PseudoWires

Ethernet over MPLS for 802.1CB packets

1+1 PW/LSP protection like approach for 8021.CB stream split

1+1 LSP protection like approach for other stream split



Ethernet over MPLS for non-802.1CB Ethernet packets

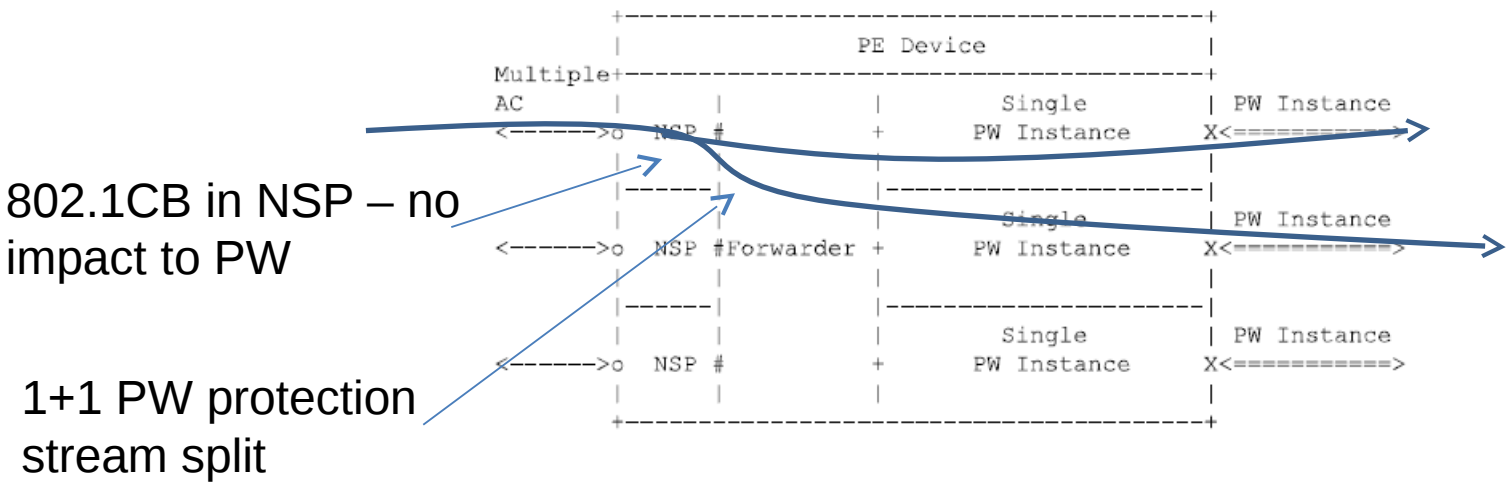
Packet PW encapsulation over an MPLS PSN for IP packets

Segmented PW & PW switching for relay systems

New functionality for merging and recovery

#1 Talker/Sender side

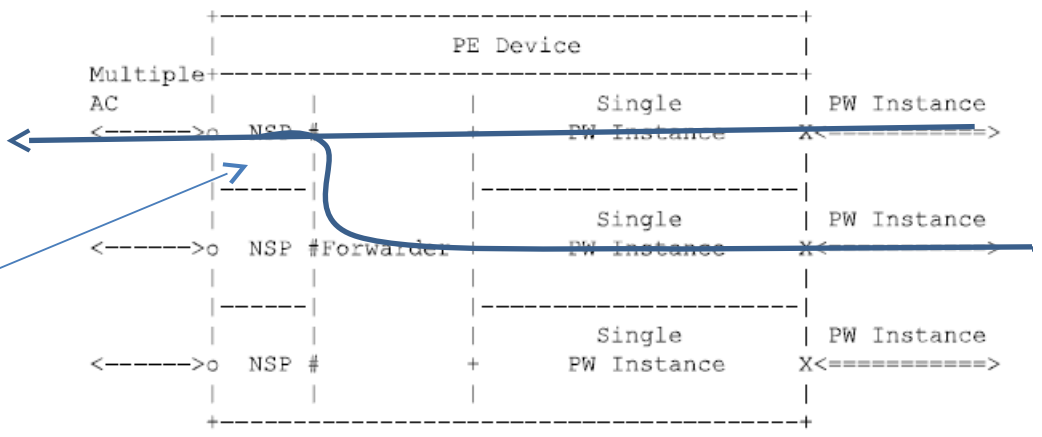
- PE implements 802.1CB natively:
  - Handle in NSP.
  - No impact to current PW architecture.
  - Existing PW and LSP protection mechanisms available for ”stream split”



#1 Listener/Receiver side

- PE implements 802.1CB natively:
  - Handle in NSP.
  - No impact to current PW architecture.

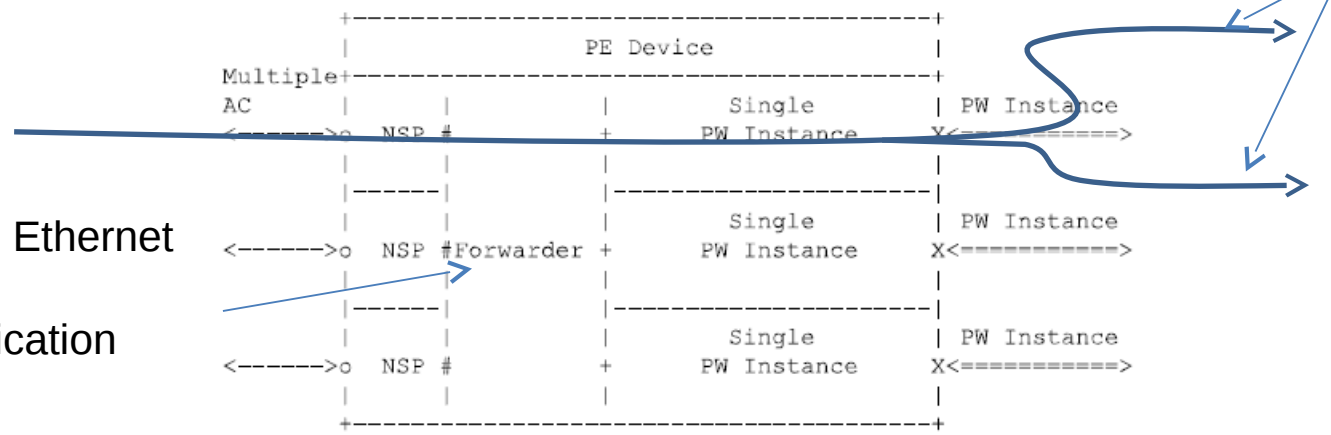
802.1CB in NSP – no impact to PW



#2 Talker/Sender side

- PE supports seamless redundancy for non-802.1CB Ethernet traffic:
  - ”stream identification” in forwarder.
  - ”stream splitting” at LSP level -> streams have the same PW label but different outer labels on the PSN tunnel.

Stream split
- Two LSPs
- One PW

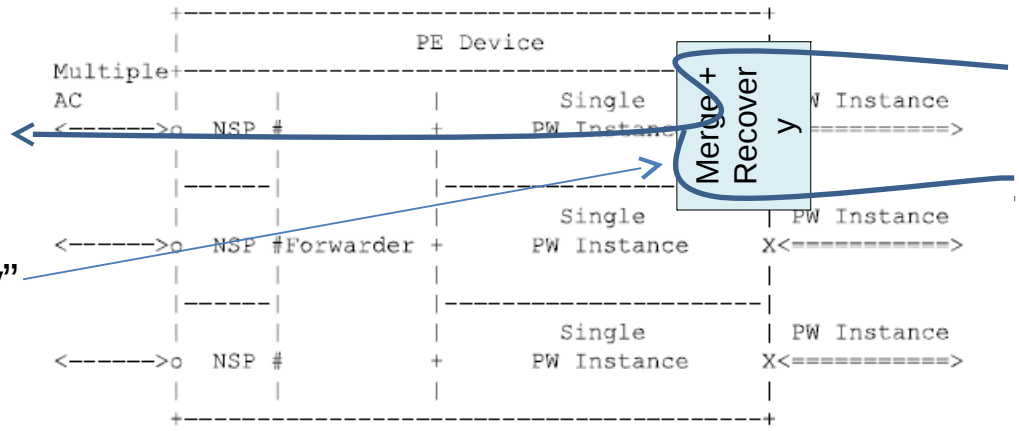


For non-802.1CB Ethernet Forwarder does:
- Stream Identification

#2 Listener/Receiver side

- PE supports seamless redundancy for non-802.1CB Ethernet traffic:
  - PW Instance receives packets from multiple LSPs that have the same PW label.
  - ”stream merge” and ”stream recovery” as a new functionality.

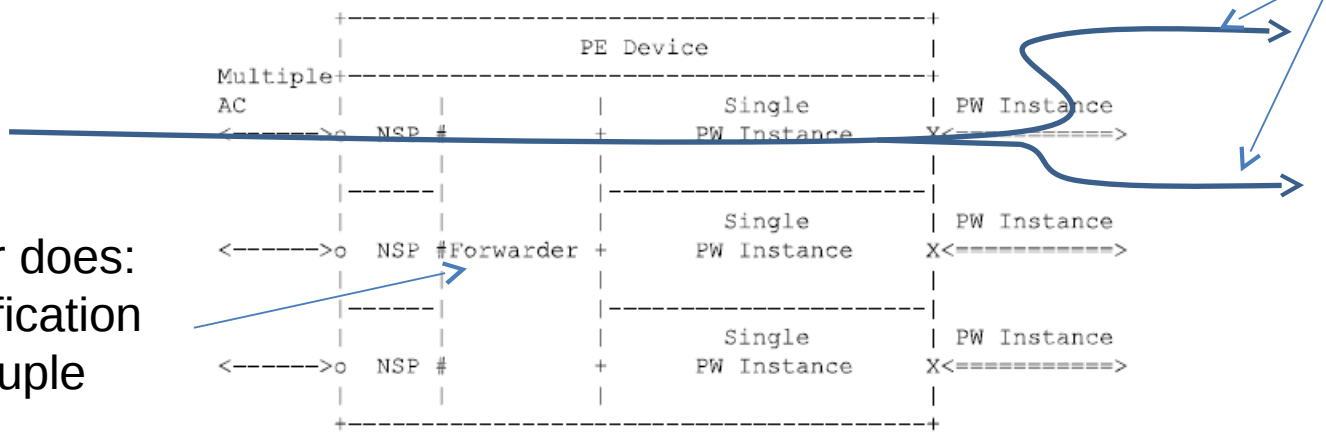
non-802.1CB Ethernet – new ”merge and recovery” function needed.



#3 Talker/Sender side

- PE supports seamless redundancy for IP traffic:
  - ”stream splitting” at LSP level -> streams have the same PW label but different outer labels on the PSN tunnel.
  - ”stream identification” in forwarder.

Stream split
- Two LSPs
- One PW

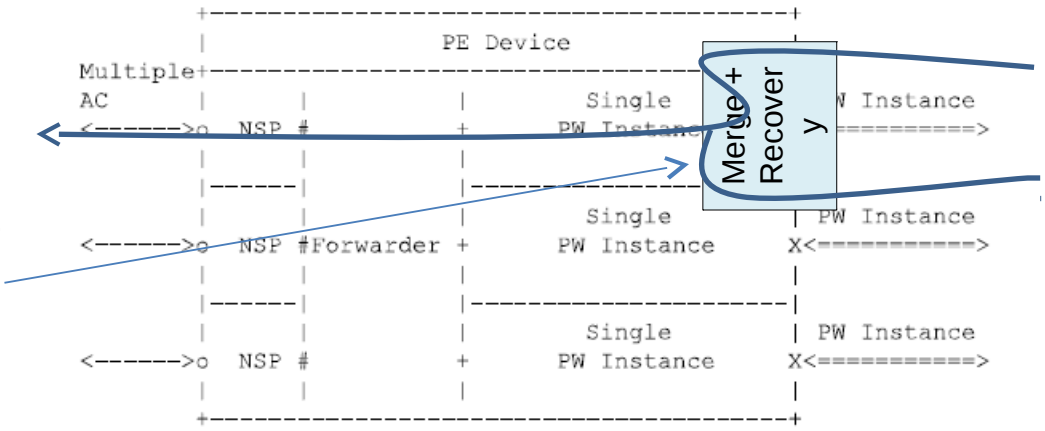


For IP Forwarder does:
- Stream Identification e.g., using 5-tuple

#3 Listener/Receiver side

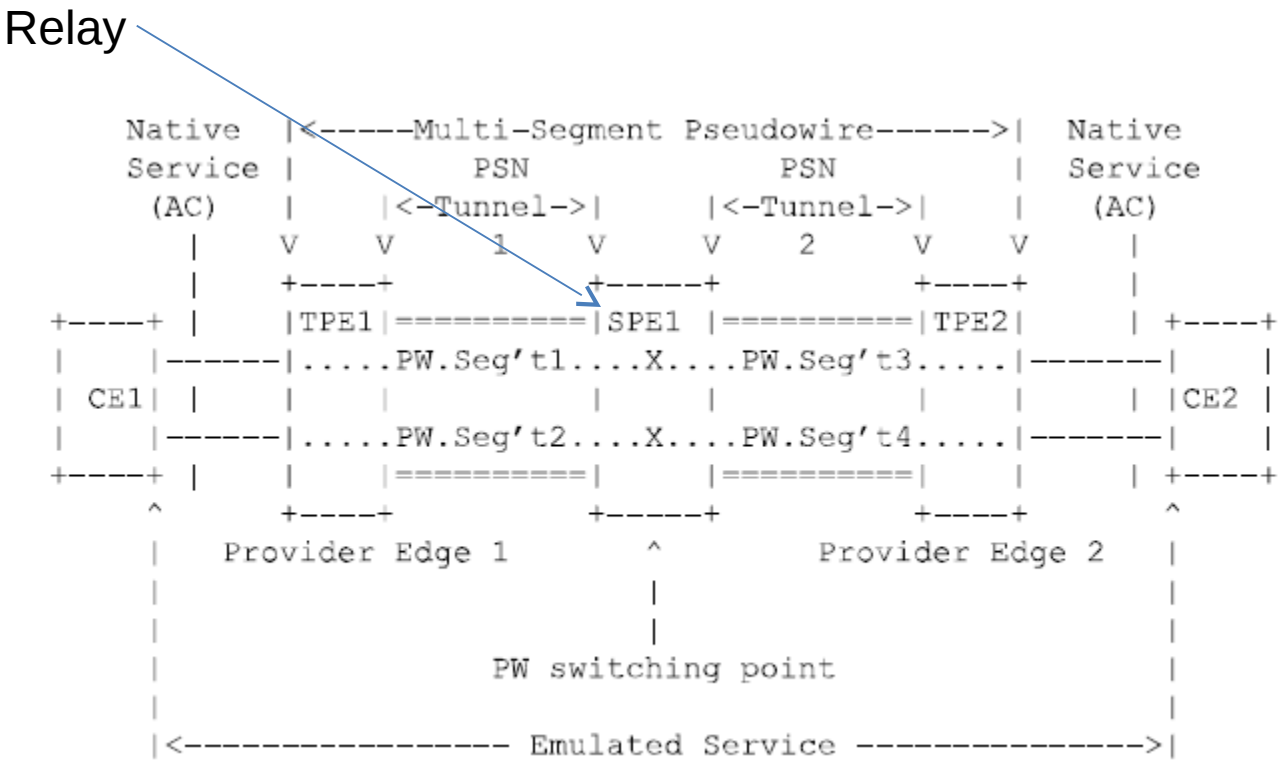
- PE supports seamless redundancy for any IP traffic:
  - PW Instance receives packets from multiple LSPs that have the same PW label.
  - ”stream merge” and ”stream recovery” as a new functionality.

IP trafficx – new ”merge and recovery” function needed.



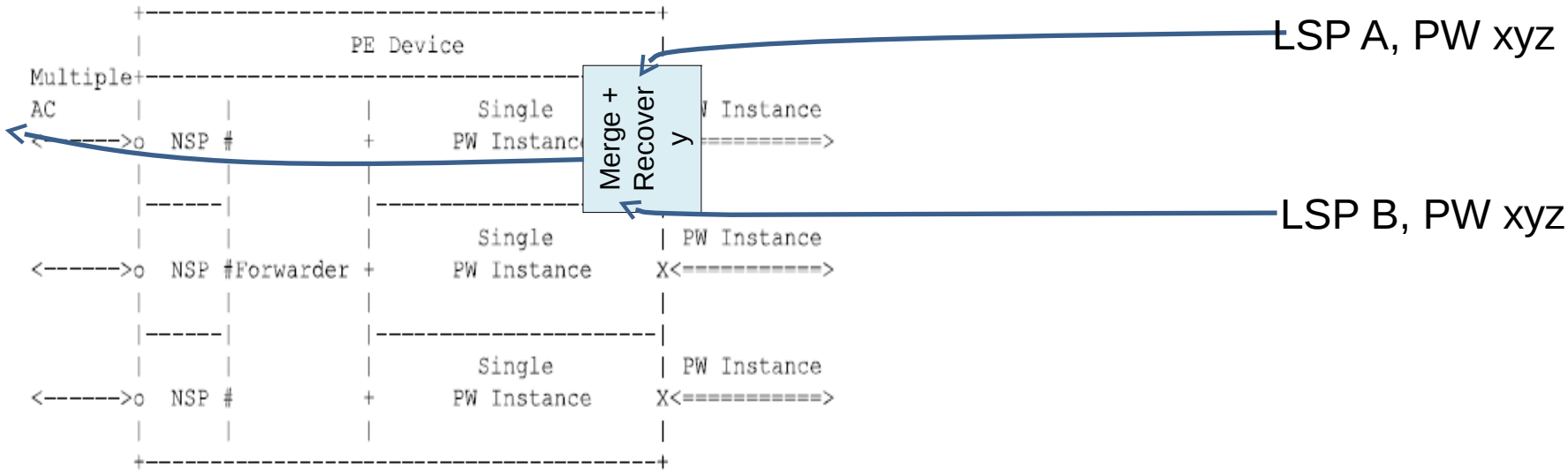
Relay system

- Leverage MS-PW model..



”Merge and recovery” function

- New functionality on the receiver side PE :
  - Receive PW Instance packets from multiple LSPs.
  - All packets on a specific PW Instance share the same PW Control Word Sequence Number space.
  - All packets to a specific PW Instance have the same PW label but may have different outer PSN tunnel labels.



Destination MAC address

- A pinned-down may pass through non-MPLS-aware bridges, as well as LSRs.
- Bridges require that every stream have its own unique {VLAN ID, destination MAC address} pair, and have a suitable L2 priority.
- That VLAN ID may be dedicated for the use of TSN/DetNet packets on pinned-down paths, and not used by normal bridged/routed/MPLS traffic.
- Therefore, LSRs need to generate and accept packets with L2 addresses assigned e.g., by the PCE.

Other considerations

- Possible optimizations:
  - No need for double sequence numbering. One to one mapping between the 802.1CB Tag Sequence Number and the PW Control Word Seqeunce number.
  - Eliminate the 802.1CB Ethernet header and regenerate that on the (T-)PEs. Would be a new ”DetNet” PW -> saves 6 (SA) + 6 (DA) + 4 (CB Tag) octets.

Other considerations cont’d

- Control channel.. TDB.
  - LDP possible, PCE possible, netconf possible, static configuration possible, ..

Other discussion

- PseudoWires and MPLS considered heavy..
  - True in a sense that MPLS typically in devices/chips that are provider class..
- Other approaches still to be considered.

Questions & Comments?

Source: http://www.politicsforum.org/images/flame_warriors/flame_67.php