PACKET PWE3 – EFFICIENT FOR IP/MPLS

IETF80 (PRAGUE) – MAR 27- APR 1, 2011

SRIGANESH KINI

DAVID SINICROPE

Problem statement

- Need for a packet service that can carry **any** protocol (similar to an Ethernet pseudowire)
- Service should be efficient for the most common protocol carried by the service
- IP and MPLS are pre-dominant protocols in many deployments
- Packet service efficient for IP/MPLS is useful

Service model

- Service is modeled as a VPWS that can carry packets of any protocol
  - Henceforth referred to as Packet Pseudowire (**PPW**)

- Efficient encapsulation defined for PPW over an MPLS PSN
  - Henceforth referred to as “Packet Pseudowire – Efficient for IP/MPLS” (**PPW-EIM**)

Network Reference model

- A single layer-2 (virtual) circuit is an access circuit (AC) to a PPW-EIM
- An AC of a PPW-EIM must not encapsulate another layer-2 circuit. E.g. in a Q-in-Q scenario, S-tag cannot be an AC to a PPW-EIM since it has multiple C-tags

Solution – with Control-Word

- CW is used to signal whether the packet is of type - IP, MPLS or 'other'
- IP and MPLS packets encapsulated in PW without layer-2 header
- For non IP/MPLS packets, the layer-2 header is included. The protocol type in the layer-2 header indicates the layer-3 protocol type.

Solution – without CW

- Packet following bottom of label-stack is always IP
- Since there is a single bottom-of-stack bit in MPLS label stack, MPLS packets don’t need special identification
- For non IP/MPLS packets, an IP header encap (GRE) is used for the entire packet (including layer-2 header). A non-routable IP address is used as destination IP address to indicate that packet is non IP/MPLS.
- Even if intermediate nodes hash based on IP header there is no re-ordering.

Example: Router interconnect

- R1, R2 – routers running LLDP, ISIS on the inter-connecting p2p IP/MPLS interface
- PE1, PE2 – PEs providing VLL service using PPW-EIM
- IP/MPLS traffic encapsulation in MPLS PSN has no layer-2 header
- PE1 and PE2 encapsulate LLDP, ISIS packets (including layer-2 header) into GRE (if no CW is used) or following the CW (when CW is used)

Conclusion

- Lesser bandwidth used.
- Fragmentation is reduced for jumbo IP/MPLS packets
- Multi-layer network in-efficiency reduced
- Enables flow based applications to parse packets efficiently even if there are multiple layers.
  - This includes ability to do ECMP based on IP (a widely deployed capability today)

v/s draft-ietf-pwe3-packet-pw (virtual ethernet)

- Advantages
  - Lesser bytes on the wire (Bandwidth efficient)
  - Less chance of fragmentation (throughput efficient)
  - IP ECMP is possible (even for multi-layer networks)
  - FAT-PW is not necessary for ECMP

- Disadvantages
  - Not possible to carry layer-2 circuit encapsulated inside a layer-2 circuit
  - For the no CW case – GRE encapsulation is more involved

Draft positioning

- This draft is positioned as an enhancement to using the ethernet PW as a packet-PW.

Efficiency analysis

		Total packet size in PSN					Efficiency as %age gain in end-				Efficiency as %age of bandwidth			
		PWE3-ETH			PPW-EIM		PPW-EIM (no CW) v/s PWE3-ETH -			PPW-EIM (with CW)+flo	PPW-EIM (no CW) v/s PWE3-ETH -			PPW-EIM (with CW)+flo
EVC		without CW	with CW	with CW + flow_lbl	without CW	with CW + flow_lbl	without CW	with CW	with CW + flow_lbl	with CW + flow_lbl	without CW	with CW	with CW + flow_lbl	with CW + flow_lbl
IP Payload size (end-user data)	40													
AC Payload Size - IP	60													
Ethernet Header size (untag)		96	100	104	82	90	35%	45%	55%	35%	15%	18%	21%	13%
802.1q single tag		100	104	108	82	90	45%	55%	65%	45%	18%	21%	24%	17%
802.1ad double tag		104	108	112	82	90	55%	65%	75%	55%	21%	24%	27%	20%
IP Payload size (end-user data)	128													
AC Payload Size - IP	148													
Ethernet Header size (untag)		184	188	192	170	178	10.9%	14.1%	17.2%	10.9%	7.6%	9.6%	11.5%	7.3%
802.1q single tag		188	192	196	170	178	14.1%	17.2%	20.3%	14.1%	9.6%	11.5%	13.3%	9.2%
802.1ad double tag		192	196	200	170	178	17.2%	20.3%	23.4%	17.2%	11.5%	13.3%	15.0%	11.0%
IP Payload size (end-user data)	512													
AC Payload Size - IP	532													
Ethernet Header size (untag)		568	572	576	554	562	2.7%	3.5%	4.3%	2.7%	2.5%	3.1%	3.8%	2.4%
802.1q single tag		572	576	580	554	562	3.5%	4.3%	5.1%	3.5%	3.1%	3.8%	4.5%	3.1%
802.1ad double tag		576	580	584	554	562	4.3%	5.1%	5.9%	4.3%	3.8%	4.5%	5.1%	3.8%

Comments
Welcome