NVGRE and VXLAN Encapsulation for L3 Overlay

draft-yong-l3vpn-nvgre-vxlan-encap-03

Lucy Yong     Xiaohu Xu

Nov. 2013   Vancouver BC
Problem Statement

• Both NVGRE \((\text{draft-sridharan-virtualization-nvgre-03})\) and VXLAN \((\text{draft-mahalingam-dutt-dcops-vxlan-05})\) are originally specified for L2 vitalization overlay data encapsulation

• Network Virtualization Overlay (NVO3) states the need of the L2 and L3 virtualization overlays

• Simple NVGRE and VXLAN enhancement can achieve the L3 virtualization overlay
About this draft

• Propose NVGRE enhancement for L3 virtualization overlay data encapsulation
• Propose VXLAN enhancement for L3 virtualization overlay data encapsulation
NVGRE Enhancement

• Propose 0x0800/0x86dd as the protocol type for IPv4/v6 payload in NVGRE header
  — 0x6558 is the protocol type for Ethernet payload [NVGRE]
• No change to other fields in NVGRE header
  — The usage of other fields remains the same too
• No change to outer header
• MUST be IP payload in the inner header if 0x800/086dd in the protocol type

GRE Header:

```
+-----------------------------+
| 0 | 1 | 0 | Reserved | Ver | Prot Type=0x6558/0x0800/0x86dd |
+-------------------------------+
| Virtual Subnet ID (VSID) | Reserved |
+-------------------------------+
Inner Header
+-------------------------------+
| IP Header |
+-------------------------------+
```
VXLAN Enhancement

• Use 16 reserved bits in VXLAN header as protocol type field
  – 0x0800/0x86dd for IPv4/v6 payload and 0x6558 as Ethernet payload
  – For the backward compatibility, value 0x0000 is treated as Ethernet payload

• No change to other fields in VXLAN header
  – The usage of other fields remains the same too

• No change to outer header

• Inner header may be Ethernet or IP depending on the value in protocol type

VXLAN Header:

```
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
| R | R | R | R | R | R | R | Reserved | Protocol Type |
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
|               |               |               |               |               |               |               |               |
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
|               | VXLAN Network Identifier (VNI) | Reserved |               |
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
|               |               |               |               |               |               |               |               |
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
```

Inner Header:

```
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
|               | Ethernet or IP header |
|               |               |               |               |               |               |               |               |
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
```
Others in Draft

• Backward compatibility in both methods
  – i.e. if tunnel egress only supports original method
• Benefit of these enhancements
  – Enable both encapsulation methods to support L3 virtualization overlay
  – To be a generalized network virtualization overlay data plane encapsulation format
  • The application for other payload type is for future study
Open Discussion

• Do we need two network virtualization overlay data encapsulation methods?
  – This draft and draft-yong-tsvwg-gre-in-udp make the enhanced NVGRE and VXLAN encapsulations very similar in the formats, the difference between two:
    • Use different standard UDP port number
    • Use different bit (3 or 5) to indicate overlay header existence

• Should IETF standardizes one or both?
  ✔ One: no need interworking or supporting both
  ✔ Both: used in industry already, if two are very similar, hardware supports both at no cost
Next Step

• Welcome comment and feedback on this
• Ready for the WG adoption