

# Geneve: What Is It and Why Is OVN Using It?

---

Jesse Gross, VMware



OVS Conference

November 16, 2015

# Existing Formats (VXLAN, NVGRE, STT)

---

- Tied to control plane, linking unrelated components.
  - Flood and learn, or other control plane semantics
- Lack extensibility to allow data plane innovation.

Possible uses of extensibility:

- Checksums/Encryption
- Security - ACLs
- OAM
- Unknown Future



# Co-Authors

---

- Jesse Gross, VMware
- T. Sridhar, VMware
- Pankaj Garg, Microsoft
- Chris Wright, Red Hat
- Ilango Ganga, Intel
- Puneet Agarwal, Innovium
- Ken Duda, Arista
- Dinesh Dutt, Cumulus
- Jon Hudson, Brocade

*<https://tools.ietf.org/html/draft-ietf-nvo3-geneve-00>*

# Geneve Overview

---

- Extensible encapsulation format to allow for future innovation
- Decouple control plane and data plane components to allow different rates of evolution
- Continue to use standard IP fabrics as an underlay
- Support for multiple encapsulated protocols and OAM

Geneve combines a UDP shim, small base header, and TLV options to achieve these goals.

# Header Format

---

## UDP:

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Source Port = xxxx           |           Dest Port = Fixed Port           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           UDP Length                    |           UDP Checksum                    |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

## Geneve Base Header:

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Ver|  Opt Len |O|C|  Rsvd.  |           Next Protocol           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Virtual Network Identifier (VNI)           |           Reserved           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Variable Length Options                   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

## Options:

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Option Class           |           Type           |R|R|R| Length |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Variable Option Data                   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

# Implementations

---

## **Controller:**

- Open Virtual Networking (OVN)

## **Software Endpoint:**

- Open vSwitch
- Linux

## **Debugging Tool:**

- Wireshark
- tcpdump
- libpcap

## **NIC:**

- Intel XL710
- Mellanox ConnectX-4
- Broadcom NetXtreme
- QLogic 578xx
- Netronome NFP-6xxx

## **Switching ASIC:**

- Broadcom Trident 2+/DNX
- Cavium XPliant
- Mellanox Spectrum
- Intel Red Rock Canyon
- Centec GoldenGate
- Marvell Prestera

# Geneve In OVN

---

OVN uses Geneve metadata to direct traffic through the virtual network.

This enables common ACLs and efficient implementation of concepts like multicast groups.

## Elements:

- Datapath ID (24-bits)
- Ingress Port (15 bits)
- Egress Port (16 bits)



# Future - Telemetry

---

In-band Network Telemetry (INT) is a framework to allow collection of fabric state.

Stores metadata in Geneve headers upon request.

*<http://p4.org/wp-content/uploads/fixed/INT/INT-current-spec.pdf>*

Examples:

- Packet path
- Hop-by-hop latency
- Buffer status





