

GENEVE Tunnels For Linux Endpoints

John W. Linville

LinuxCon North America

17 August 2015

Introduction
Motivation
Current Status
Future
Conclusion

Who am I?
What is this?

Who am I?



What is this?

Let's talk about GENEVE!

- GENEVE is a relatively new network tunneling protocol
- OVS (~ 1 year), netdev (~ 3 months) in kernel
- Some hardware support is appearing too
- Discussion needed of control plane and other options

If we keep building it, will they come?

Why tunnels?

Tunnels are a fundamental technology for connecting endpoints across a foreign network...

- VLANs are limited to physical L2 networks
- Some applications demand an L2 environment
- Deployments can span diverse physical networks
- Virtualized (i.e. logical) topology – SDN

Tunnels make the network fit how you want to use it!

Why GENEVE?

Why not use VXLAN, NVGRE, STT or other tunneling protocols?

- Each alternative has its own drawbacks
- GENEVE provides a superset of the capabilities from the other protocols (including non-Ethernet frame tunneling)
- Readily extensible through variable length option headers
- Not tied to any control plane (e.g. no need for multicast)

GENEVE is best-of-breed and the most flexible tunneling option...

Why a netdev?

Doesn't OVS already support GENEVE tunnels? Yes, but...

- Not everyone wants to run OVS
- Ultimately, a netdev is needed somewhere in the path anyway
- A netdev allows for fewer internal bridge hops at endpoint
- Configuring a tunnel netdev is simpler too

A netdev provides the familiar networking configuration point that users expect!

Standardization

GENEVE is still working its way through the IETF standardization process...

- Initial draft dated 14 February 2014
- 4th draft – Network Virtualization Overlay WG (NVO3)
- Current draft expires 9 November 2015
- Draft approval – ???

GENEVE is coming...

Infrastructure

Low-level GENEVE infrastructure shared between OVS and netdev...

- Developed to enable OVS vport
- Kept separate from vport implementation (thanks!)
- Currently limited to IPv4 support (some code for IPv6)
- Simplifies netdev implementation – cleaner code!

Going, going, gone...

Network Driver

GENEVE netdev merged for 4.2

- Unicast, IPv4 endpoints
- Pre-configured point-to-point tunnels
- Support a handful of tunable parameters (e.g. TOS, TTL)
- IPROUTE2 patches merged as well

Suitable for use in place of other tunneling protocols within those limitations...maybe...

Surprise!

Last week (8/11), Pravin posted 6-part series “Geneve: Add support for tunnel metadata” ...surprise!

- Uses new technology for OVS tunnels (tunnel metadata API)
- Some overlap with locally queued patches (configurable UDP destination port)
- Refactors geneve_core and geneve into single module

Slow pace and failure to communicate has it consequences... :-)

New Features

So, what is to come?

- IPv6 support (in testing)
- More tunables (similar to VXLAN)
- Options processing (?)
- Features to enable control planes

Need user stories to guide development!

Control Planes

Static configuration is error prone and boring...

- VXLAN control planes should work on GENEVE
- Some VXLAN control planes are little used
- Standardized control planes?
- How about GENEVE support in Flannel? Others?

LPC networking mini-conference discussion scheduled for Friday...

Hardware Offload

Tunnel encapsulation is great, but it comes at a cost...

- CPU usage increases, limits throughput
- Hardware offload can eliminate excess CPU usage
- Vendors have shown NICs with GENEVE hardware offload
- Linux could take advantage of such NICs (similar to `ndo_add_vxlan_port`)

Push for GENEVE offload when purchasing NICs...

Wrap-Up

So what's the point?

- Containers and virtual hosts need tunnels to plumb virtual networks
- GENEVE is the best-of-breed technology for tunneling L2/L3 traffic over IP
- Linux kernel support for GENEVE is here and improving

Let's decide how to make the best use of GENEVE!

Questions?



Contact

Feel free to contact me!

- Email linville@tuxdriver.com
 - ...@redhat.com
 - ...@gmail.com
 - ...@kernel.org
- IRC [linville](#) on [FreeNode](#)
- Facebook as “[John W. Linville](#)”

Links

Current GENEVE Draft

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

Open vSwitch

<http://www.openvswitch.org/>

Flannel

<https://github.com/coreos/flannel>