Generic Protocol Extension for VXLAN (VXLAN GPE)

draft-quinn-vxlan-gpe-04

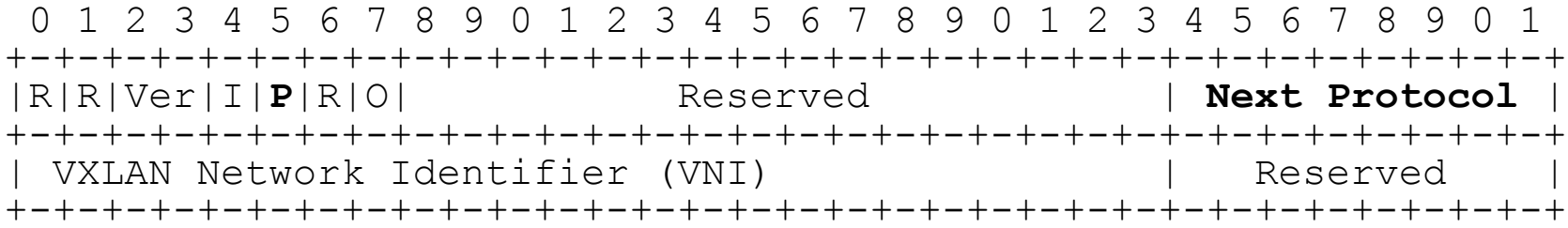
Paul Quinn, Lawrence Kreeger, Darrel Lewis, Fabio Maino, Michael Smith (Cisco)
Uri Elzur (Intel)
Pankaj Garg (Microsoft)
David Melman (Marvell)
Rajeev Manur (Broadcom)
Lucy Yong, Xiaohu Xu (Huawei)
Puneet Agarwal

VXLAN GPE Goals

- Add the ability to carry additional protocols beyond Ethernet (e.g. IPv4 and IPv6)
- Keep the Network Virtualization header lean while allowing for a transport independent shim header to be added for any optional extensions/metadata
- Improve virtual network insight and troubleshooting by enabling VN level OAM messaging
- Minimize impact on existing deployed hardware supporting VXLAN and leverage existing VXLAN hardware designs
- Allow any future incompatible changes to VXLAN GPE to be supported without requiring a new UDP Port

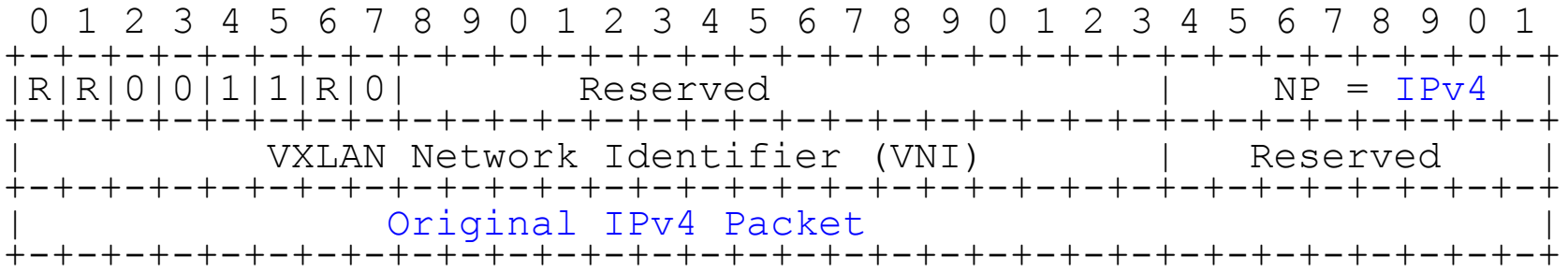
VXLAN GPE Next Protocol Field

- Presence of the Next Protocol field is signaled by the P-bit
  - This allows the same parsing logic to be used for both VXLAN GPE and VXLAN
- Next Protocol can signal the payload type or the next header. Values are defined for IPv4, IPv6, Ethernet and Network Service Header (NSH)

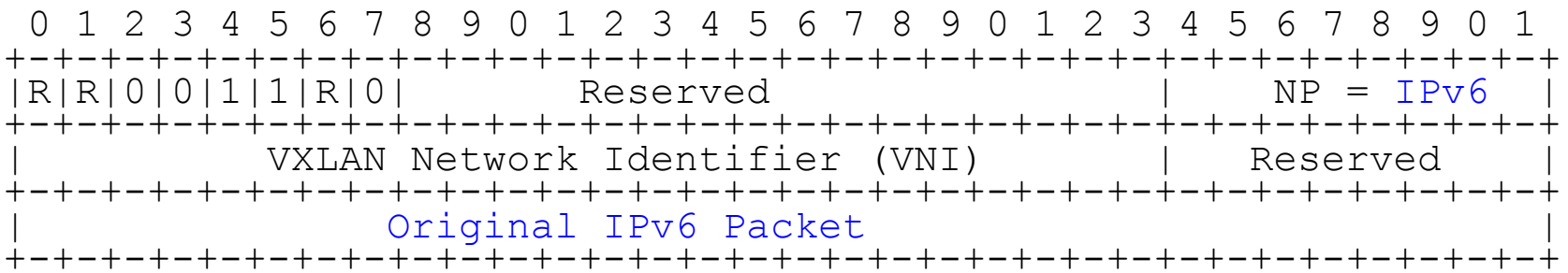


Next Protocol Encapsulation Examples

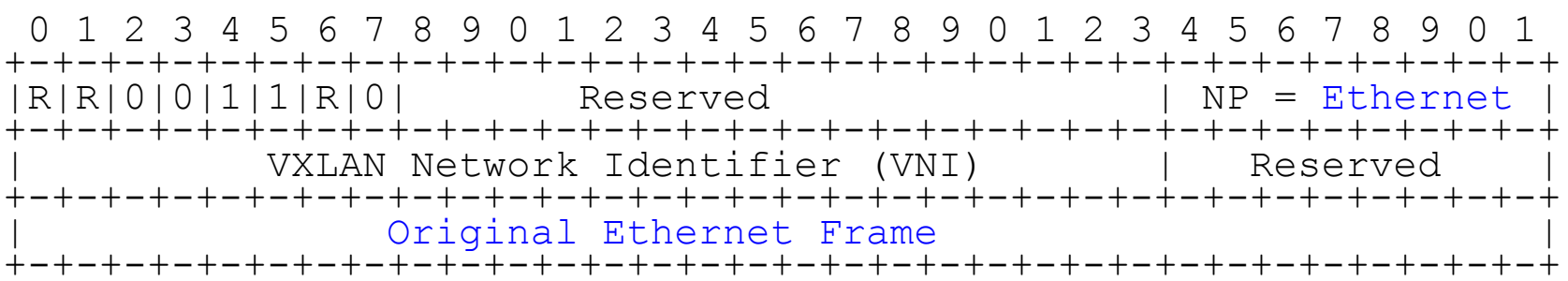
IPv4



IPv6

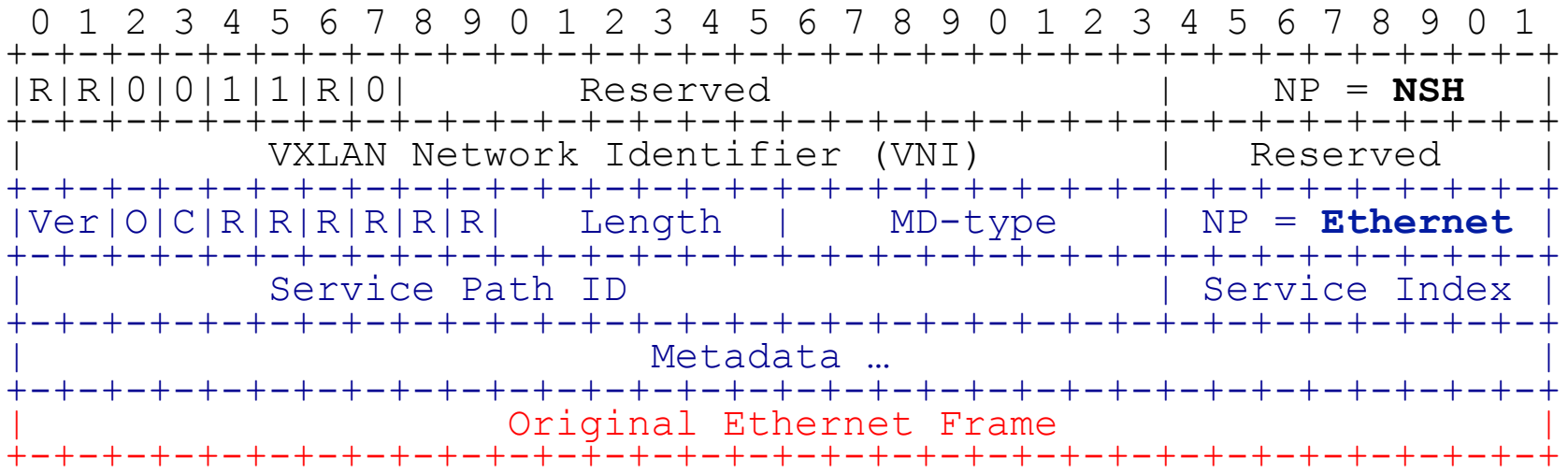


Ethernet



Extensibility using the Next Protocol

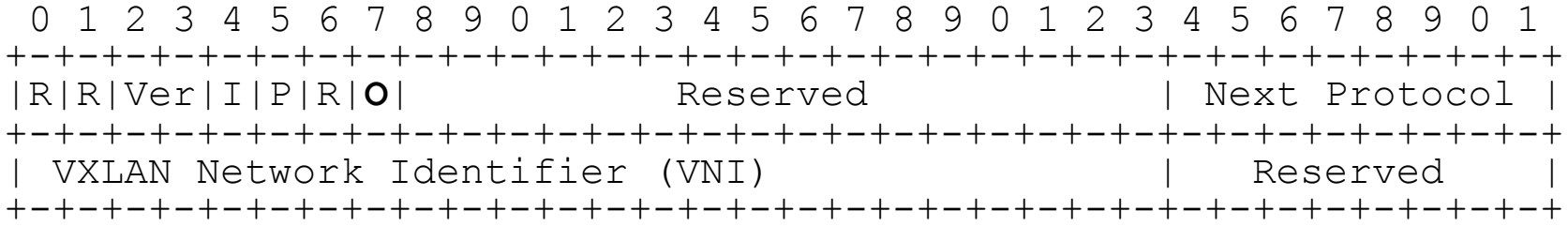
- Use the Next Protocol field to insert a Network Service Header after the VXLAN GPE header



VXLAN GPE + Network Service Header with Ethernet Payload

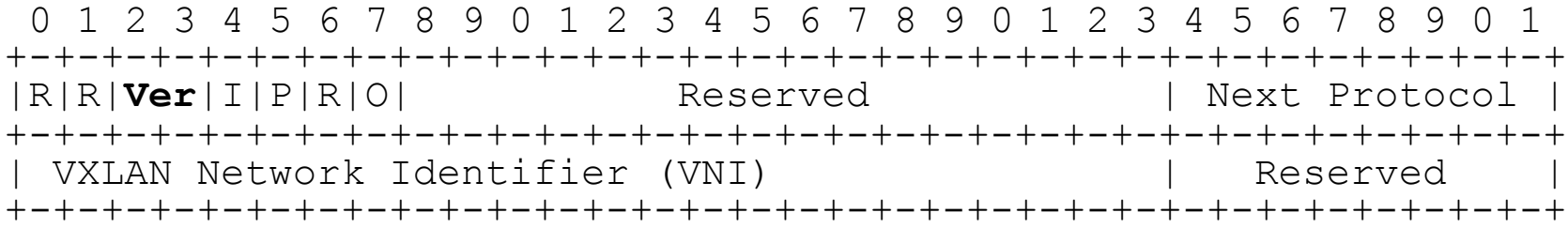
OAM Support

- OAM bit signals the payload carries an OAM message
  - Packets with the OAM bit set MUST not be delivered to a tenant end system
  - Packets with the OAM bit should be consumed for OAM message processing



Future VXLAN GPE Enhancements

- Reserved flags/fields MUST be set to zero on transmit and *ignored* on receipt
  - This allows the fields to be used in a forward compatible way as long as the new usage can be safely ignored by existing VXLAN GPE VTEPs
- The version field MUST be set to zero on transmit and dropped if non-zero on receipt
  - The version must be incremented if reserved fields are using in a non-forward compatible way, or other fields are changed



New UDP Port

- A new UDP port was assigned for VXLAN GPE
- Using the new port value (4790) prevents accidental delivery of packets to existing VXLAN VTEPs from being misinterpreted

Changes in version 04

- Added IANA assigned UDP port (4790)
- Moved Version field into the 8 bit flags byte and clarified usage
- Added text/figures to allow VXLAN GPE to stand on its own without having to read RFC 7348 (VXLAN)
- Enhanced introduction to explain motivations
- Reorganized section 3 for clarity
- Changed VXLAN reference from VXLAN draft to VXLAN RFC
- Added a new co-authors
- Miscellaneous minor editorial fixes

Work Group Adoption

- Multiple independent implementations are currently underway
  - Software (including vendor and open source)
  - Hardware (multiple vendors, both NIC and switch)
- VXLAN is a de facto industry standard and VXLAN GPE builds on that momentum
- The authors feel the draft is ready and suitable for NVO3 WG adoption