

Encapsulation Considerations

Design team report
draft-rtg-dt-encap-01.txt
Selected topics for NVO3

Design team members

Albert Tian

Erik Nordmark

Jesse Gross

Jon Hudson

Larry Kreeger

Pankaj Garg

Pat Thaler

Tom Herbert



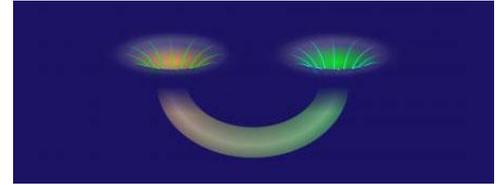
Charter <http://www.ietf.org/mail-archive/web/rtgwg/current/msg04715.html>

Motivation for design team



- IETF doing new encaps - NVO3, SFC, BIER
 - And multiple might be used in the same packet
- Each encap has its own information, but also needs to handle common issues
 - Explore more common ways to handle those issues
 - Each proponent/WG doesn't need to reinvent
- Focus is on encaps packet format - *not* on control plane

What this IS



- A look across the three new encapsulations
 - While taking lots of previous work into account
- Focus on encaps that run over IP/UDP
 - Many encaps desire to run at least over IP
 - Avoided diving into control-plane interaction
- Turns out some “transport” independence fell out as a result
 - E.g., MPLS entropy label fits in

What this is NOT

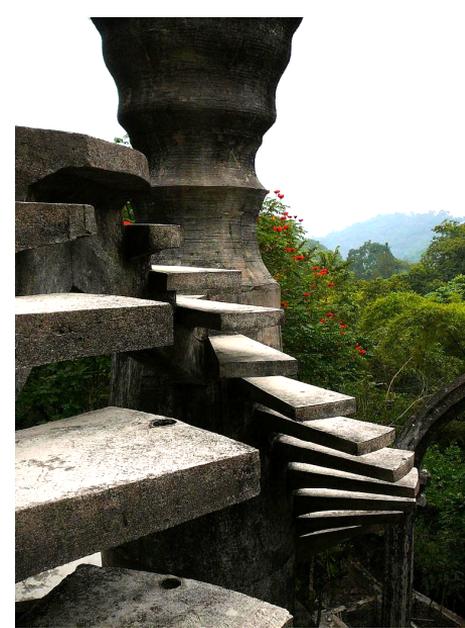


- A design of a new encaps to rule them all
- A design of a new NVO3 encaps
- A selection from existing encapsulations
- An evaluation of existing and proposed encapsulations
- A floor wax and/or dessert topping

Set of common issues

A twelve-step program

1. How to provide entropy for ECMP
2. **Next header indication**
3. Packet size and fragmentation/reassembly
4. OAM - what support needed in an encapsulation format?
5. **Security and privacy**
6. QoS
7. Congestion Considerations
8. **Header and data protection - UDP or header checksums**
9. **Extensibility - for OAM, security, and/or congestion control**
10. Layering of multiple encapsulations
11. Service model
12. **Hardware Friendly**

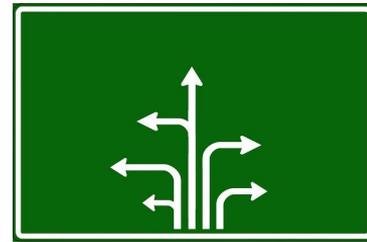


Different encaps - different information



- NVO3 needs to carry at least a VNI-ID
 - Carried edge-to-edge unmodified
 - Optional OAM info like timestamps modified?
- SFC carries service path and meta-data
 - Something modified at each hop for loop prevention
 - Service meta-data may be modified by SF
- BIER carries a bitmap of egress routers
 - Bitmap modified as packet is forwarded

Next header indication



- Each encapsulation wants to carry different payloads
 - Use Ethernet types? IP protocol number? Create new numbering space?
- When layering multiple encapsulations headers?
 - Define a common approach?
 - Define a common numbering space?
- But also needs to fit with existing schemes
 - UDP uses port numbers; GRE Ethernet types; etc.
 - Used to indicate the (first) encapsulation header

Security and privacy



- At least three considerations for security
 - Anti-spoofing - prevent packet injection
 - Interaction with and use of IPsec
 - Privacy
- Different possible anti-spoofing mechanism
 - Cookie in encaps header - against off-path attacks
 - Secure hash of header fields (excluding fields modified in transit)

Header protection



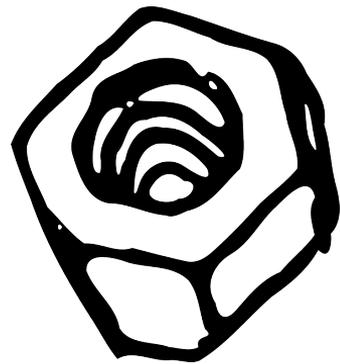
- RFC 6936 Applicability Statement for the Use of IPv6 UDP Datagrams with Zero Checksums
- Need checksum for the encaps header?
 - Misdelivery if e.g. VNI ID, BIER bitmap is corrupted
 - Using pseudo-header for important IP fields?
- Ties in with higher assurance for security
 - No need for checksum if secure hash is used?

Extensibility

- Needed semantics
 - New incompatible version
 - Stuff which can be ignored by the egress
 - Error/drop if egress doesn't support
 - Handle on-path parsing (BIER routers, middleboxes)
- Different encodings
 - Use reserved bits/fields
 - TLVs; extension header chains
 - Flag-fields as in GRE
- Use it or lose it?



Hardware Friendly



- Not required, but impacts deployment
 - Using existing chips; facilitate design of new chips
- Different hardware concerns for
 - Switch/router chips, vs. NIC offload
- Encap header checksum OK - not whole
 - However, NIC offload can do whole pkt checksum
- Put important info at fixed offsets
 - Unconstrained TLVs seem hard
 - Limit number of header combinations

Middlebox Considerations



- As encapsulations get widely deployed middleboxes might do more
 - Not just drop based on UDP port number
 - Gateways stitching could have similar effect
- Example would be to filter VNI IDs for NVO3
 - Better defense in depth
- Should the IETF document what not to do?
 - Avoid accidentally blocking OAM but not payload
 - Avoid interfering with ECMP?

Open Issues

- Common OAM error reporting protocol?
 - Useful or a distraction?
- Next protocol indication - common across different encapsulation headers?
- In-order-delivery service layer on top vs. sequence numbers and timestamps for OAM and CC?



Next Steps

- Present in RTGWG in Dallas
- Gather feedback from different groups in the IETF

