GRE-in-UDP Encapsulation

draft-ietf-tsvwg-gre-in-udp-08

Lucy Yong lucy.yong@huawei.com

November 2015 , Yokohama JP

Update in this version

- Explicitly specify two distinct usages for GRE-in-UDP encapsulation protocol,
  - GRE-in-UDP usage on an IP network including the Internet
  - GRE-in-UDP usage within a well-managed operator network
    - Give the definition of the well-managed operator network
- List GRE-in-UDP tunnel implementation requirements under the Internet and a well-managed operator network.
- Specify GRE-in-UDP tunnel implementations under each network
  - Some are the same and some are different
  - The main differences between two networks are in use of UDP checksum and congestion control

A well-managed operator network

- An IP network that is under single administrative control and meet at least one of following:
  a) known that packet corruption is exceptionally unlikely and where the operator is willing to take the risk of undetected packet corruption.
  b) Judged through observational measurements that the level of packet corruption is tolerably low and where the operator is willing to take the risk of undetected packet corruption.
  c) Known that carrying applications that are tolerant of mis-delivered or corrupted packets where the operator is willing to rely on the applications using the tunnel to survive any corrupt packets.
- The network has congestion control via traffic-engineering
  – network operators may avoid congestion by careful provisioning of their networks, by rate limiting of user data traffic, and traffic engineer according to path capacity.
- As a result, when GRE-in-UDP is used within the network
  – UDP zero-checksum is allowed in IPv6 with additional provided tunnel implementation,
  – no need congestion control function at the tunnel endpoints

Tunnel Requirements for the Internet

1. SHOULD perform UDP checksum when over an IPv4 network.
2. MUST perform UDP checksum when over an IPv6 network.
3. IP-traffic can be assumed to be congestion-controlled; other tunneled protocol/payload SHOULD implement an appropriate congestion control method because the GRE/UDP tunnel does not itself provide any congestion control. If GRE-in-UDP tunnel MUST NOT to traffic that has no congestion control over the general Internet.
4. UDP src port that is used for flow entropy SHOULD be set to a UDP ephemeral port (49152-65535).
5. For IPv6 delivery network, if IPv6 flow label load balancing is supported [RFC4638], the flow entropy SHOULD also be placed in the flow label field.
6. If a tunnel ingress fragments the incoming packet (before encapsulation), the UDP checksum MUST be used so that the receiving endpoint can validate reassembly of the fragments, and the same src UDP port SHOULD be used for all packet fragments to ensure that the transit routers will forward the packet fragments on the same path.
7. If the incoming packet needs to be fragmented, it SHOULD be done before the encapsulation [RFC7588] and calculate the size of fragments based on the MTU and including the size of the UDP header.

The draft refers the tunnel implementation for the Internet as of the default GRE-in-UDP tunnel implementation

Tunnel Requirements for a well-managed operator network

1. When over an IPv4 network, SHOULD set UDP zero-checksum to improve the tunnel performance.
2. When over an IPv6 network, MUST perform UDP checksum as default but MAY be configured with UDP zero-checksum with additional implementation requirements that are specified in Section 5.2.
3. A tunnel may encapsulate a protocol/payload that does not provide congestion control if the delivery network is traffic-engineered and/or operated by the network operator to avoid congestion, e.g. use of pre-provision capacity or utilize a circuit breaker [CK].
4. UDP src port that is used for flow entropy SHOULD be set to a UDP ephemeral port (49152-65535).
5. For IPv6 delivery network, if IPv6 flow label load balancing is supported [RFC4638], the flow entropy SHOULD also be placed in the flow label field.
6. If a tunnel ingress fragments the incoming packet (before encapsulation), the UDP checksum MUST be used so that the receiving endpoint can validate reassembly of the fragments, and the same src UDP port SHOULD be used for all packet fragments to ensure that the transit routers will forward the packet fragments on the same path.
7. If the incoming packet needs to be fragmented, it SHOULD be done before the encapsulation [RFC7588] and calculate the size of fragments based on the MTU and including the size of the UDP header.

The draft refers the tunnel implementation for this network as of the conditional GRE-in-UDP tunnel implementation

Next Steps

- Address comments
- Ready for WG LC?