Tunnel congestion Feedback
(draft-wei-tunnel-congestion-feedback-04)

Xinpeng Wei

Lei Zhu

Bob Briscoe

Lingli Deng

IETF 93 Prague, Czech

Tunnel-based Congestion Management Overview

- The aim of tunnel-based congestion management is to provide network operator with a better control on network congestion status, and prevent the network fall into persistent congestion state.

- Congestion management could be divided into three stages:
    - Congestion level measurement
    - Congestion level information transmission
    - Congestion management action

①, ② are where we focus on

Feedback Models

```
                        Feedback
        |-------------------------------------------|
        |                                           |
        |                                           |
        |                                           V
+----------+          tunnel             +-----------+
|Egress    |============================ |Inress     |
|(Exporter)|                             |(Collector)|
+----------+                             +-----------+

(a) Direct Feedback Model.
```

Egress feeds information directly to ingress, In this model, egress collects network congestion level information and feedback the information to ingress for congestion management

```
 Feedback      +-----------+
  --------->|Controller |####################
   |        |(Collector)|                   #
   |        +-----------+                   #
   |                                        #
+----------+          tunnel         +------V-+
|Egress    | ========================|Ingress|
|(Exporter)|                         +-------+
+----------+

(b) Centralized Feedback Model
```

After egress collects network congestion level information, it feeds back the information to controller instead of ingress, and then the controller makes congestion management decision and sends the decision to ingress.

Congestion Level Measurement (1)

- It assumes that the interior routers in the tunnel supports ECN;
- Network congestion level could be indicated through the ratio of CE-marked packet and the ratio of packet drop.
- If packet loss is detected, it could be assumed that severe congestion has occurred in the tunnel
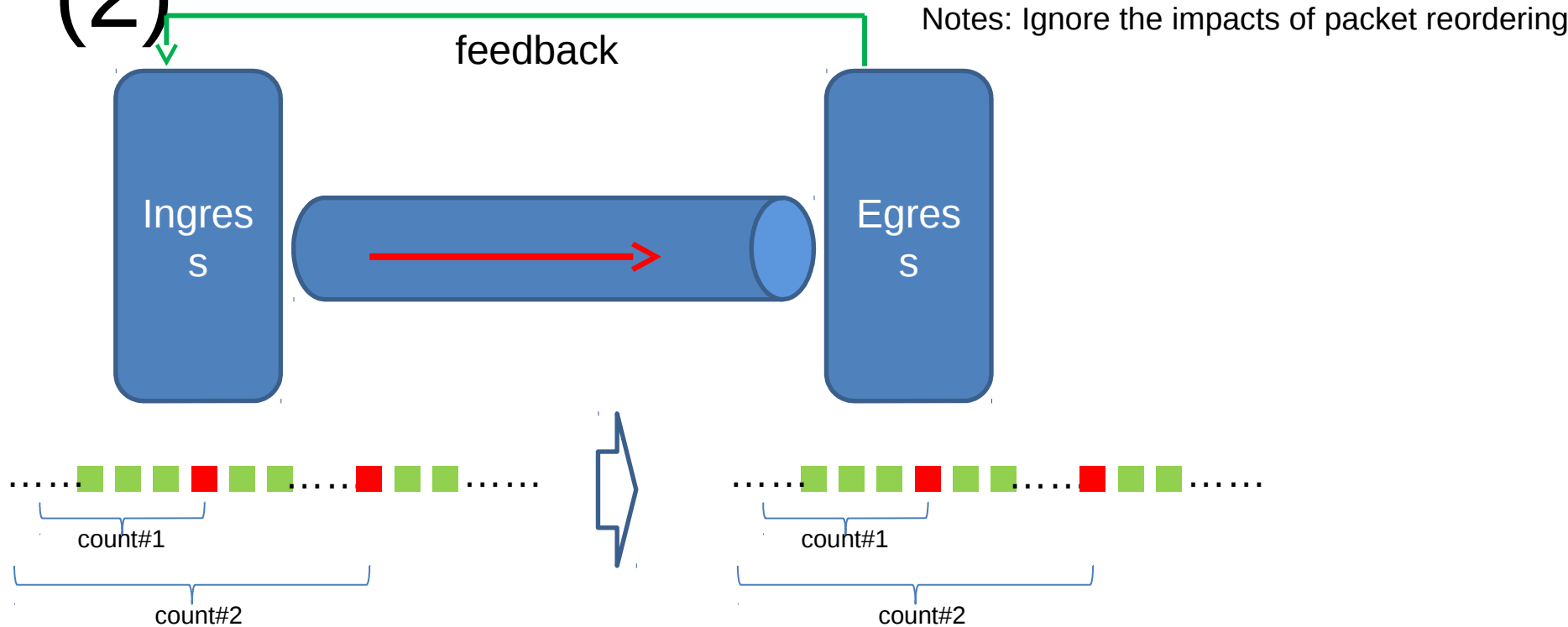- Faked ECT is used to defer packet loss to egress.

Ingress

Egress

Ingress Marking (outer ECN| inner ECN):
CE|CE
ECT|N-ECT
ECT|ECT

Egress Marking (outer ECN| inner ECN):
CE|CE
ECT|N-ECT
CE|N-ECT
CE|ECT
ECT|ECT

Congestion Level Measurement (2)

feedback

Notes: Ignore the impacts of packet reordering

Ingress

Egress

count#1

count#2

count#1

count#2

Legends:

network traffic packet

signal message packet

☐Contents of signal message packet (Ingress→Egress): <Ingress node id, cumulative packets count of each ECN combination>

☐Ingress cumulatively collects packet counts and inserts signal message packet into network traffic.

☐After egress received a signal message packet, it will add cumulative packet counts of each ECN combination to the signal message packet and feed it back to ingress or controller.

IPFIX Extensions

- IPFIX is selected as feedback protocol, and a list of new Information Elements are defined to convey packet number of each ECN combination:
  - ce-cePacketTotalCount
  - ect-nectPacketTotalCount
  - ce-nectPacketTotalCount
  - ce-ectPacketTotalCount
  - ect-ectPacketTotalCount

Congestion Management

- Congestion management could be taken based on the result of congestion measurement to prevent network from falling into persistent congestion.

- Different congestion management method could be deployed, here are two possible candidates:
  - Circuit breaker as defined in draft-ietf-tsvwg-circuit-breaker.
  - Congestion Policing as defined in draft-briscoe-conex-policing.

Next Steps

- Ready for call for adoption!

Questions?