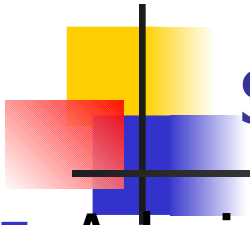
LC-PCN – The Load Control PCN solution

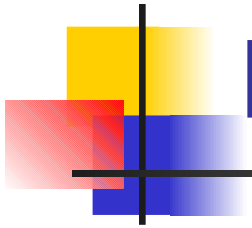
draft-westberg-pcn-load-control-03.txt

Lars Westberg, Anurag Bhargava,
Attila Bader, Georgios Karagiannis



Supported solutions for issues

- Admission control support
  - With data marking
  - With probing (solves ECMP during admission control)
  - Combination of the two
- Flow Termination:
  - Base mode
  - Optimization mode
- ECMP solutions:
  - Admission control => using probing
  - Flow termination => using Affected marking



ECMP problem

- Occurs in admission control and flow termination:
  - Flows can belong to congested ingress-egress-aggregate, but due to ECMP routing, packets belonging to these flows might not pas through congested node
  - Any measures taken on such flows will not solve congestion problem, since such flows do not contribute to congestion

ECMP solution in admission control

- Use probing (probe is rejected)
  - When certain ingress-egress-aggregate in egress operates in admission control state AND receives a probe packet belonging to flow associated with same ingress-egress-aggregate AND probe packet is PCN_marking encoded, then egress knows for sure that probe packet passed through one or more congested PCN-interior-nodes
  - => probe packet rejected

ECMP solution in admission control

- Use probing (probe is admitted)
  - When certain ingress-egress-aggregate in egress operates in admission control state AND receives a probe packet belonging to flow associated with same ingress-egress-aggregate AND probe packet is **NOT** PCN_marking encoded, then egress knows for sure that probe packet **HAS NOT** passed through one or more congested PCN-interior-nodes
  - => probe packet accepted

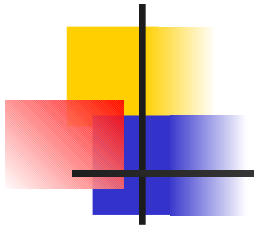
ECMP solution in admission control

- Use probing
  - **Requirement**: if interior node congested THEN ALL probe packets MUST be marked.
  - **Problem**: when excess rate marking is used, if interior node in admission control, then NOT CERTAIN that all probe packets are PCN-marking encoded.
  - **Solution**: using a router alert option for probe packets to make sure that PCN nodes always PCN_marking encode them when corresponding links are pre-congested with regard to configured-admissible-rate (C-A-R)

ECMP solution in flow termination

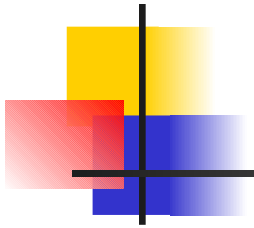
- Only flows that are passing through congested node are selected for termination
- Requires:
  - flow termination state at PCN-interior-nodes
  - additional encoding state: Affected Marking
  - when PCN-interior node operating in flow termination state, then all packets passing through PCN-interior-node and are NOT PCN_marking encoded:
    - => PCN_Affected_Marking encoded
  - when PCN-egress-node operates in flow termination state it selects for termination only flows that contain:
    - PCN_marking encoded packets
    - PCN_Affected_Marking encoded packets

PCN-interior-node (admission control with data marking)



- U >= 1, defined as in SM draft; equal in PCN domain
- N >= 1, proportionality excess rate and remarked rate; equal in PCN domain
- When no ECMP solution is supported then admission control and flow termination uses only one encoding state
- Admission control with data marking:
  - Identical to SM draft, but for optimization/accuracy purposes:
    - Marked excess rate = measured excess rate / N
    - Excess rate measurements should be done before dropping and marking after dropping
    - PCN_marking encoded packets should not be preferentially dropped

PCN-interior-node (admission control with probing)

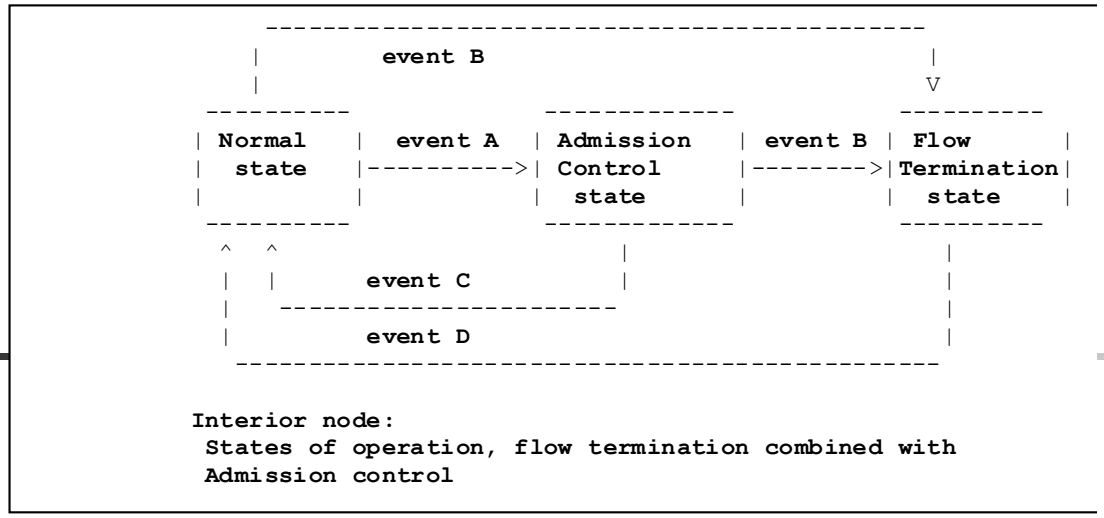


- Admission control with probing:
  - Used to solve ECMP problem
  - Uses the same admission control state as admission control with data marking
  - When in admission control state required to PCN_marking encode packets that carry RAO
  - The two admission control mechanisms can be used independently or combined.

PCN-interior-node (flow termination)

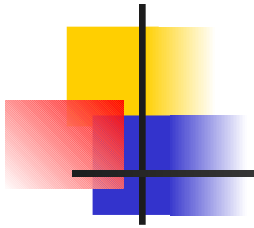
- Base mode:
  - operates same as admission control state with data marking
- Optimization mode:
  - Used to solve innaccuracies in measurements due to existing delays between metering and marking events, decisions made at egress, but flows and their packets are stopped by the ingress
  - Flow termination state is required
  - Sliding window used to store rates of packets that were PCN_marking encode done in previous intervals

Interior Node



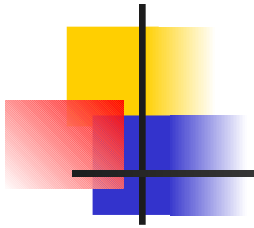
- Event A: Measured Rate per PHB (MR) > (C-A-R))
  ("encoded DSCP" rate = 1/N * excess rate (rate above C-A-R))
- Event B: used only in Flow Termination optimization mode and when Flow Termination ECMP solution used
  - MR > U * C-A-R
  - ("encoded DSCP" rate = 1/N * excess rate (rate above C-A-R))
- Event C: MR ≤ C-A-R
- Event D: used only in Flow Termination optimization mode and when Flow Termination ECMP solution used
  - MR ≤ (U * C-A-R)

PCN-egress-node (admission control with data marking)



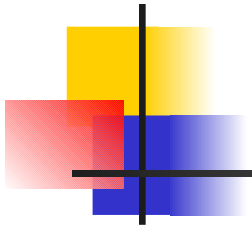
- Detection of admission control state is identical to SM draft:
  - ratio between incoming_PCN_marked_rate and total received PHB aggregated PCN traffic higher predefined value, e.g., 1%
  - incoming_PCN_marked_rate = N * measured excess rate
- Admission control by combining PCN operational state and admission control request provided by external to PCN, signaling protocol
  - If ingress-egress-aggregate at egress operates in admission control state then received admission control request: => rejected
  - If ingress-egress-aggregate at egress operates in normal state then received admission control request: => accepted

PCN-egress-node (admission control with probing)



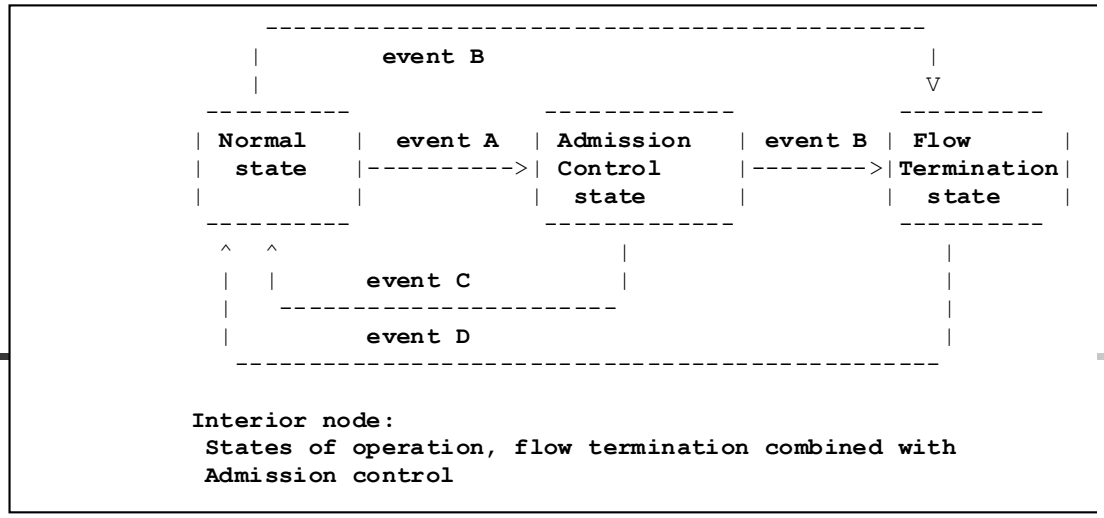
- Used to solve ECMP problem
- Uses the same admission control state as admission control with data marking
  - It could however, operate even if no ingress-egress-aggregate state is available at egress
- Arrived probe packet is PCN_marking encoded: => reject
- Arrived probe packet is NOT PCN_marking encoded:=> accept
- Send feedback to PCN-ingress-node

PCN-egress-node (flow termination)



- Detection of flow termination state is identical to SM draft:
  - When ratio between incoming_PCN_marked_rate and (total received PHB aggregated PCN traffic or PCN_unmarked rate) higher than predefined value, e.g., (U-1)
    - Go into flow termination state
    - Store value of incoming_PCN_marked_rate => => configured-termination-egress-rate
- Excess rate above configured-termination-egress-rate is used to calculate number of flows to be terminated
- Send feedback for flows to be terminated to ingress

Egress Node



- Event A: IMR/(IUR +IMR) > 1%

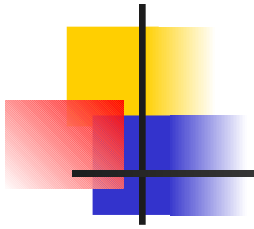
  where, IMR = Measured rate of PCN_marking packets * N,

  Where IUR: measured rate of NOT PCN_marking packets
- Event B: IMR/(IUR +IMR) > (U-1)
- Event C: IMR/(IUR +IMR) ≤ 1%
- Event D: IMR/(IUR +IMR) ≤ (U-1)

PCN-ingress-node (admission control with data marking)

- if feedback received from egress => notifies accept, then request accepted
- if feedback received from egress => notifies reject, then request rejected

PCN-ingress-node (admission control with probing)



- probe packets can be either user packets or packets used by signalling messages, e.g., RSVP PATH.
- probe packets must use the same flow ID as packets belonging to the same flow
- if not available, include Router Alert option into the probe packets
- if feedback received from egress => notifies accept, then request accepted
- if feedback received from egress => notifies reject, then request rejected

Conclusions and next steps

- LC-PCN at ingress:
  - Different than SM
- LC-PCN at interior:
  - Admission control with data marking:
    - Same as SM draft, but to increase accuracy small modifications needed
  - Admission control with probing (additional option to solve ECMP)
  - Flow termination:
    - Base mode, same as features used in admission control with data marking
    - Optimization mode (optional feature that is required in order to increase accuracy of algorithm)
    - ECMP solution (additional and optional feature that requires a flow termination state and additional encoding state)

Conclusions and next steps

- LC-PCN at Egress:
  - Admission control with data marking: identical to SM
  - Admission control with probing (additional option used to solve ECMP problem)
  - Flow termination:
    - Detection feature: identical to SM
    - Selection of the flows to be terminated: different than SM
    - Feedback to ingress: different than SM
    - ECMP solution (additional and optional feature)

Conclusions and next steps

- Integrate LC-PCN with SM (and possibly other PCN WG schemes)