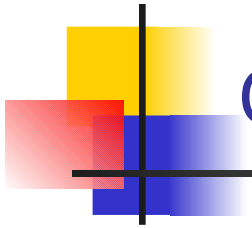
LC-PCN – The Load Control PCN solution

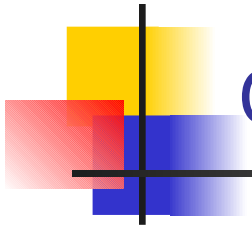
draft-westberg-pcn-load-control-00.txt

Lars Westberg, Attila Bader,
David Partain, Georgios Karagiannis



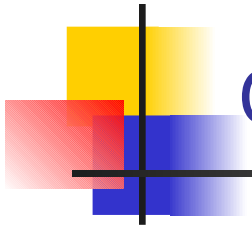
Outline

- Overview of LC-PCN solution
- Comparison with other PCN schemes
- Conclusions and next steps



Overview of LC-PCN solution

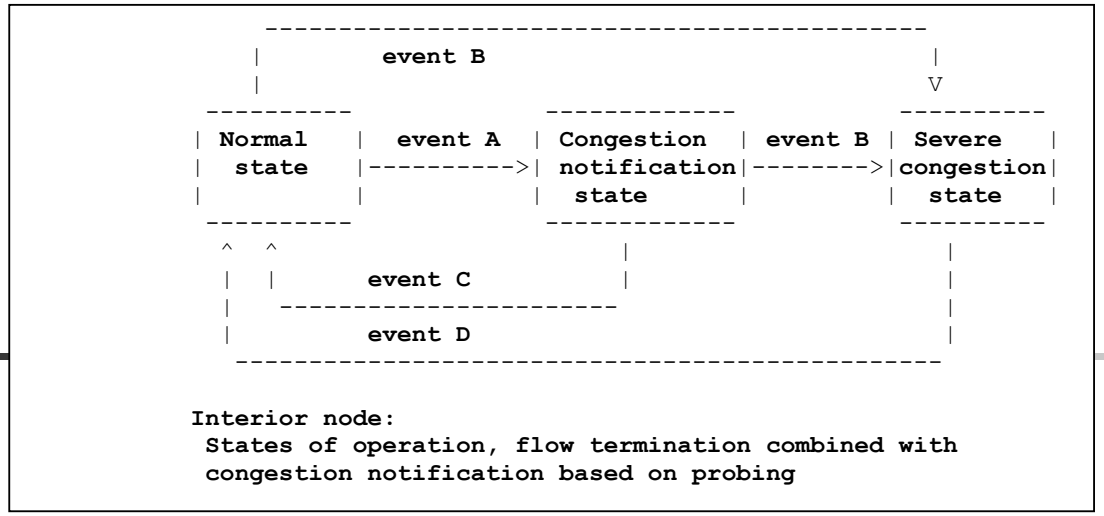
- Applied in a PCN domain and used for unidirectional and bidirectional flows
- Supports admission control (based on probing), flow termination and ECMP handling during admission control and flow termination



Overview of LC-PCN solution

- Diffserv configuration:
  - Interior: Meter, Marking Action, Packet Classification:
    - Marked excess rate = (Metered excess rate / N), where N >1 and same in whole PCN domain Configuration
  - Egress: Identifies probe packets and measures excess rate and defines which new flows should be rejected and which ongoing should be selected for termination
  - Ingress generates probe packets and uses information from egress to reject/admit the new flow and to stop selected on-going flows

Interior Node

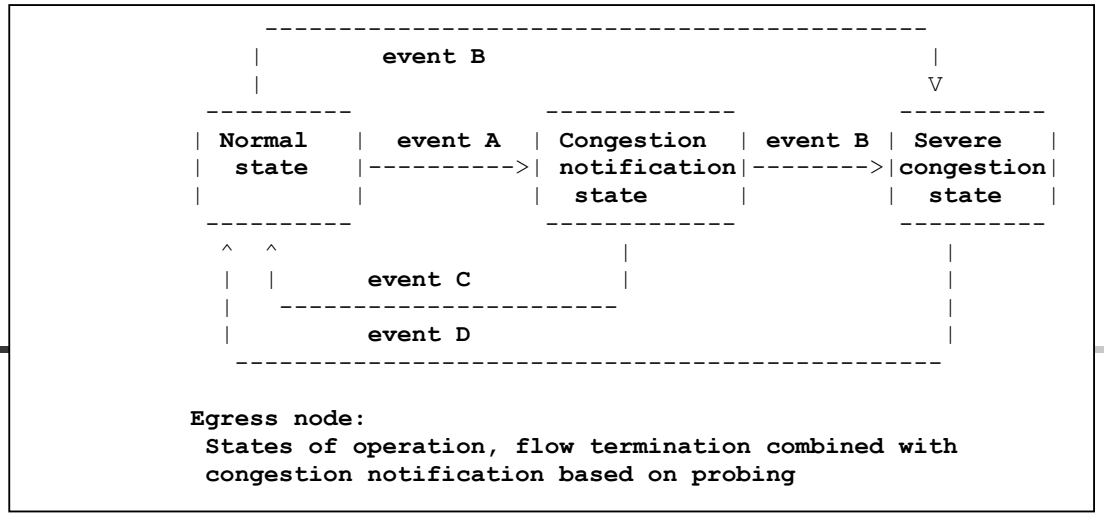


- Normal state: no congestion
- Severe congestion state = Flow Termination (FT) state
- Congestion notification state = Admission control (AC) state

Events when one encoding state used for admission control and flow termination

- Event A: Measured Rate per PHB (MR) > congestion notification rate (i.e., configured admissible rate (C-A-R))
  ("encoded DSCP" rate = 1/N * excess rate (rate above C-A-R))
- Event B: MR > severe congestion detection (i.e., congestion termination rate: C-T-R)
  ("encoded DSCP" rate = 1/N * excess rate (rate above C-T-R))
- Event C: MR ≤ C-A-R
- Event D: MR ≤ severe congestion restoration rate (C-T-R)
- Event E: (same as event D) but not in the figure and only used when two encoding states are used for AC and FT states

Egress Node



- Normal state: no congestion
- Severe congestion state = Flow Termination (FT) state
- Congestion notification state = Admission control (AC) state

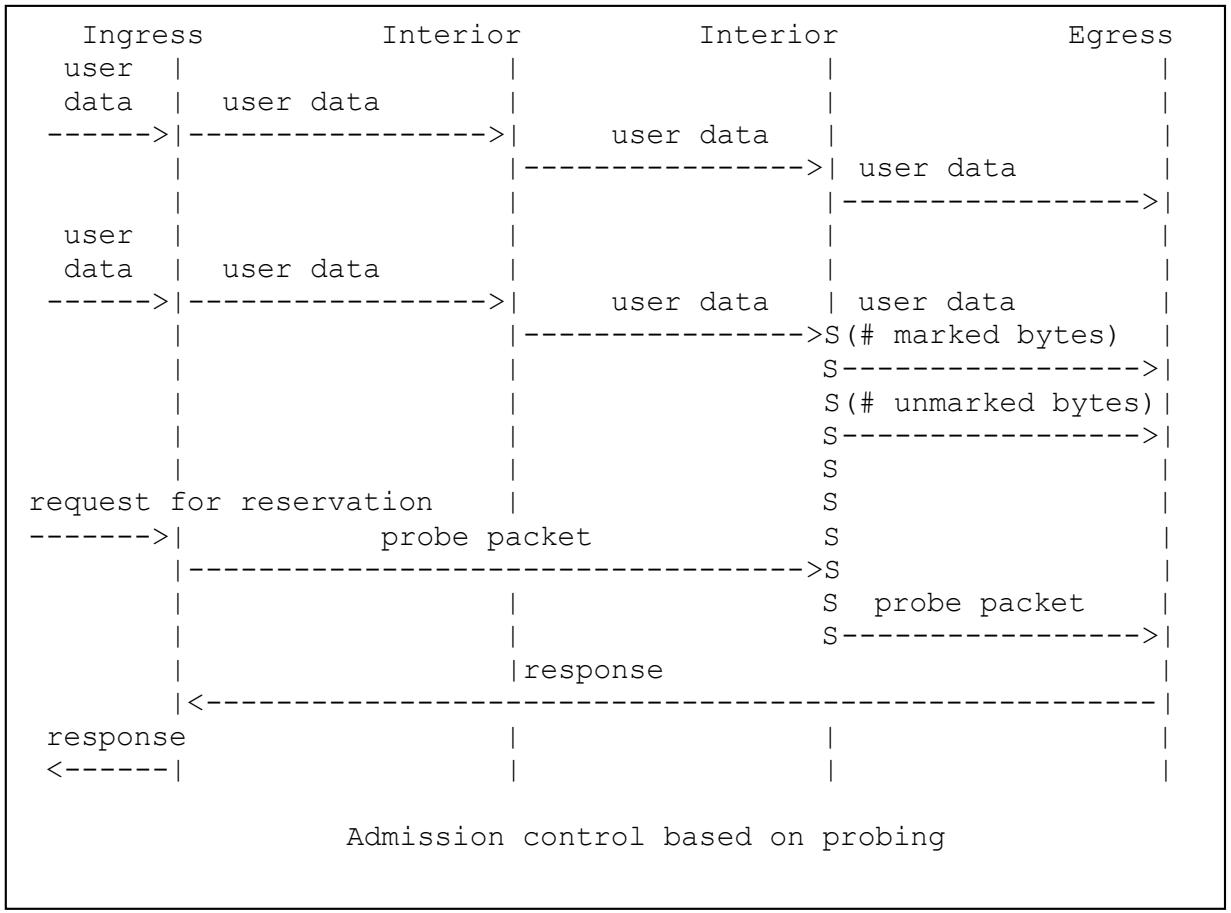
Events when one encoding state used for admission control and flow termination

- Event A: (MRE > C-A-R) AND (MRE ≤ C-T-R)

where, MRE = Measured rate of "encoded DSCP" * N, C-A-R = congestion notification rate, C-T-R = severe congestion detection

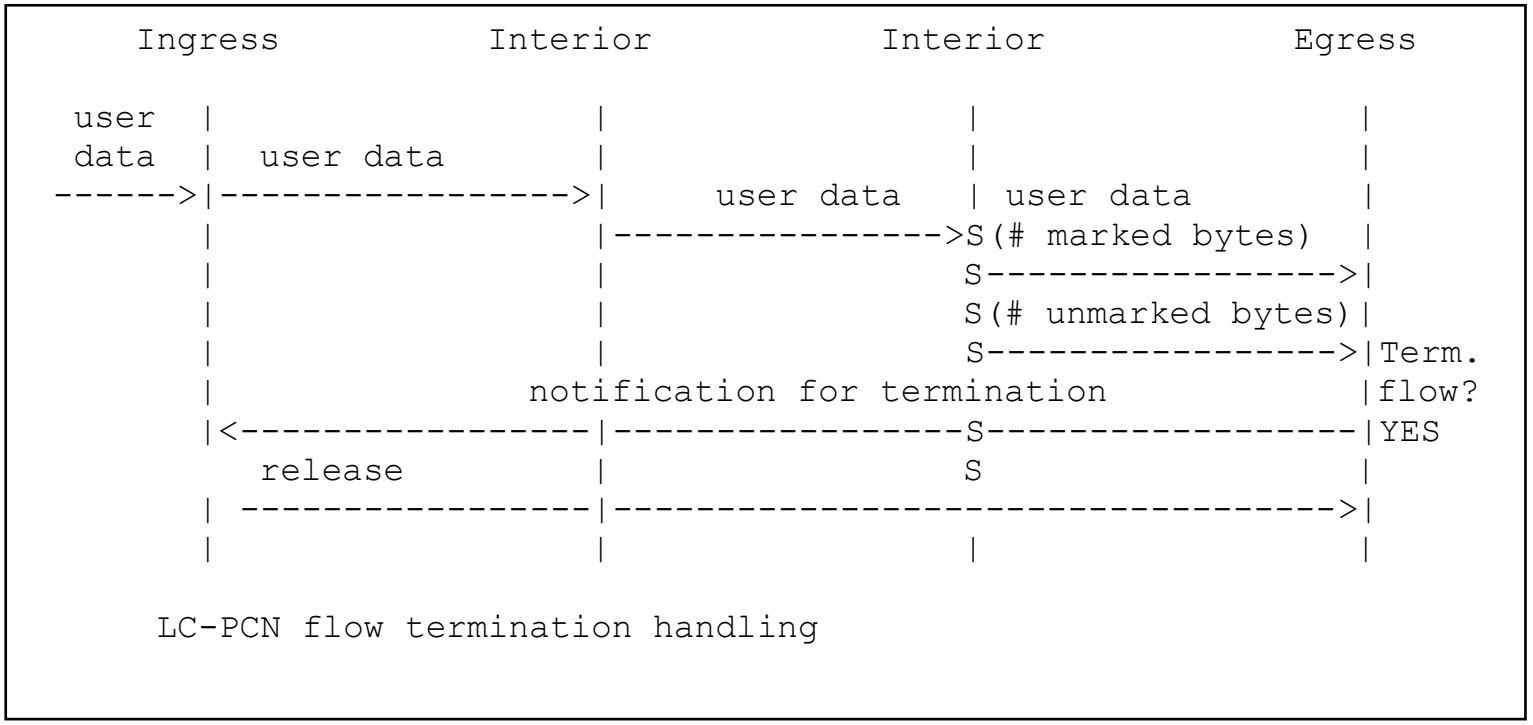
- Event B: MRE > C-T-R
- Event C: MRE ≤ C-A-R
- Event D: MRE ≤ C-T-R
- Event E: (same as event D) but not in the figure and only used when two encoding states are used for AC and FT states

Overview of LC-PCN solution



Admission control based on probing

Overview of LC-PCN solution



LC-PCN flow termination handling

Comparison with other PCN schemes

	LC-PCN	CL-PHB	Single Marking	3SM
PCN features	NC, AC, FT, ECMP-AC, ECMP-FT (measurements per byte)	NC, AC, FT	NC, AC, FT	NC, AC, FT, ECMP-AC, ECMP-FT
Operation at Ingress	AC: Generate probes and reject if negative notification FT: Terminate selected flows	AC: CLE > C-A-R => reject FT: Termination BW (TBW)= Input load –SAR Terminate flows according to TBW	AC: same as CL-PHB FT: SPR = u* SAR Termination BW (TBW)= Input load –SPR Terminate flows according to TBW	AC: Either due to negative notification or generate probes and reject if negative probe notification FT: If S = 0, same as CL-PHB If S>0 => terminate selected flows

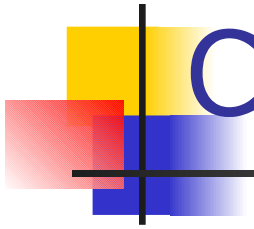
- NC = Not congested, AC = Admission Control, FT = Flow Termination
- ECMP-AC = ECMP solution used during AC, ECMP-TC =ECMP used during FT
- CLE = Congestion Level Estimation, SAR = Sustainable Admission Rate, SPR = Sustainable Preemption Rate

Comparison with other PCN schemes

	LC-PCN	CL-PHB	Single Marking	3SM
INTERIOR	Option 1 (two encodings): MR>C-A-R=>AM MR>C-T-R=>TM Option 2 (one encoding): MR>C-A-R=>TM MR>C-T-R=>TM TM= 1/N Excess MR (applied when MR even above Maximum Capacity)	MR>C-A-R=>AM MR>C-T-R=>TM (applied when MR not above Maximum capacity)	MR>C-A-R=>AM (applied when MR not above Maximum capacity)	MR>C-A-R=>AM MR>C-T-R=>TM (applied when MR not above Maximum capacity)
EGRESS	Option 1 (two encodings): AC: MRE-AC = AM FT: TBW = TM*N: (MRE-TM> C-T-R) Option 2 (one encoding): AC: MRE-AC=TM*N reject: probe marked+MRE-AC>C-A-R FT:TBW=TM*N: (MRE-TM> C-T-R) Select flows according to TBW, send TBW to ingress	AC: CLE=(AM+TM)/total Send CLE to ingress FT: SAR= rate unmarked packets Send SAR to ingress	AC: CLE= AM/total Send CLE to ingress FT: SAR = rate unmarked packets Send SAR to ingress	AC: Reject either MRE-AC>C-A-R or probe marked. Send notification to ingress FT: If S= 0, see CL-PHB; If S>0 => Select all TM marked flows to terminate

Conclusions and next steps

- LC-PCN at ingress:
  - Generate probe packets and and reject if probe is marked, accept otherwise
  - Terminates selected flows
- LC-PCN at interior:
  - packets TM marked according to excess rate
  - All packets that are not TM marked are Affected Marked (used for ECMP)
  - Probing used to solve ECMP during AC
  - Similar to 3SM and single marking
- LC-PCN at Egress:
  - Excess rate measurements and probing is used to admit a reservation request or not
  - Selects only (TM and Affected Marked) marked flows to be terminated according to the calculated termination bandwidth (TBW)
    - Solves ECMP problem
  - similar to 3SM when S>0 and when S the same in whole PCN domain



Conclusions and next steps

- Evaluate if and how the LC-PCN scheme can be combined/integrated with the other PCN WG schemes