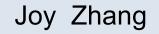


### PCN with Single Marking draft-charny-pcn-single-marking-01

### Anna Charny

acharny@cisco.com



joyzhang@cisco.com

Francois Le Faucheur

flefauch@cisco.com

Vassilis Liatsos

vliatsos@cisco.com

# Outline

- Basic algorithmic overview http://www.ietf.org/internet-drafts/draft-charny-pcn-si
- Summary of performance evaluation
- Tradeoffs
- Next steps

### **Basic Idea: Admission**

- Only <u>one</u> (admission) rate threshold set at each link
- Traffic exceeding the admission threshold <u>rate</u> is marked
  - <u>Excess rate</u> marking rather than "virtual queue occupancy" marking (e.g. implemented by a token bucket)
- Egress measures marked traffic rate (aka CLE) on a per-ingress basis and sends it to ingress
  - Just as in draft-briscoe-cl-architecture
- Ingress stops admitting when CLE exceeds a given threshold
  - Operation just as in draft-briscoe-cl-architecture
    - But the semantics of CLE is different
      - Hence the setting of the CLE parameter at which admission stops may also be different

# **Basic Idea:** Flow Termination (2)

 Nothing is done at the core for termination (admissionmarking only)

-Flow termination threshold is *implicitly* derived from admission threshold

 Egress measures <u>unmarked</u> traffic ("sustainable admission rate") on a per-ingress basis and sends it to ingress along with the CLE

Just as in draft-briscoe-cl-architecture, but based on <u>admission</u> marking

- Ingress computes sustainable flow termination rate by multiplying sustainable-admission rate by a network-wide constant K
- Effectively, works as if flow termination threshold were configured at all links at K\*admission-threshold and the separate marking for flow termination were used

### **Performance Comparison**

- Both VQ and excess-rate admission work well at reasonable bottleneck and per-ingress-egress pair aggregation levels
- Excess-rate admission is more sensitive to low ingress-egress aggregation levels, especially for bursty on-off traffic
- Even for small levels of ingress-egress aggregation a range of parameters with reasonable performance across all topologies and traffic models tested exists

Multi-bottleneck topologies and real video traces have now been simulated

• See draft-charny-pcn-single-marking for more info

### Tradeoffs - the good

Saves one codepoint

**Especially important with MPLS** 

 Requires only one metering/marking mechanism in the core instead of two

Important for data path performance

• Ease of use with existing hardware Everyone implements a token bucket!

### Tradeoffs: The Bad

Limits the flexibility of choosing termination thresholds

-Bad if two networks with different system-wide K merge

Excess-rate admission control

-More sensitive to parameters and traffic patterns than virtual-queue based Admission Control

•But still quite tolerable in simulations

 Appears to conflict with the anti-cheating mechanism of draft-briscoe-tsvwg-re-ecn-bordercheat

### The Ugly (?) : Does this conflict with draft-briscoecl-architecture?

Different marking semantics from base CL architecture

-Different metering admission-marking mechanisms at the core

-Different decision-making process at the ingress for admission

 Open question: Does this mean that the WG group needs to choose between the two (or more) mechanisms or is there a way to define the standard to allow them all?

-Interoperability concerns if all are allowed?

# Possible approach for allowing both

- Could single-marking behavior be defined as a subset of twothreshold marking behavior?
- "Excess-rate" marking (Type 1) and "VQ" marking (Type 2)
- A core device may support Type 1 or both
- An ingress edge always computes sustainable termination rate as sustainable rate times a constant K

K=1 for "two-threshold" CL

K is the system-wide constant for single-marking

An egress edge

always measures sustainable rate for Type 1 marking

Configuration decides whether CLE is measured for Type 1 or Type 2 marking

Always sends both CLE and sustainable rate to the ingress

### Interoperability

- If Core devices of Type 1 only and Type 1&2 are in the same network, all must revert to Type 1
- All ingresses must be configured with the same value of K
- All egresses must be configured to measure CLE for the same type marking
- Allows for incremental implementation and deployment

Type 1 first, Type 2 next

at the cost of 1 configuration knob at the ingress (value of K) , and 1 at the egress (which marking is used for CLE measurement)

## **Conclusions and Next Steps**

 Single-marking approach appears a technically viable alternative based on evaluation so far

Simulations/analysis will continue to gain further confidence

 Allows for incremental implementation and deployment as a first step to baseline approach of draft-briscoe-cl-architecture

-Fewer implementation changes to existing core equipment

- -Smaller performance impact in the data path of core routers
- Need feedback from the WG:
  - -on the scheme details/properties?
  - -Allow multiple options or choose one?
  - -Allow single marking as an option or not?

# CISCO SYSTEMS