

# Controlled Load (CL) Service using distributed measurement-based admission control (D-MBAC)

**Bob Briscoe**, Gabriele Corliano, Phil Eardley,  
Peter Hovell, Arnaud Jacquet, Dave Songhurst

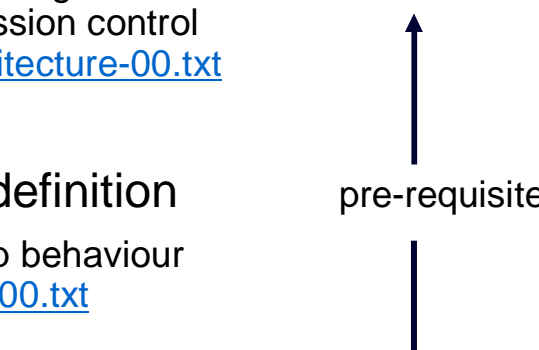
**BT Research**

IETF-63 tsvwg Aug 2005

Original idea:  
Martin Karsten  
then of TU Darmstadt



# drafts

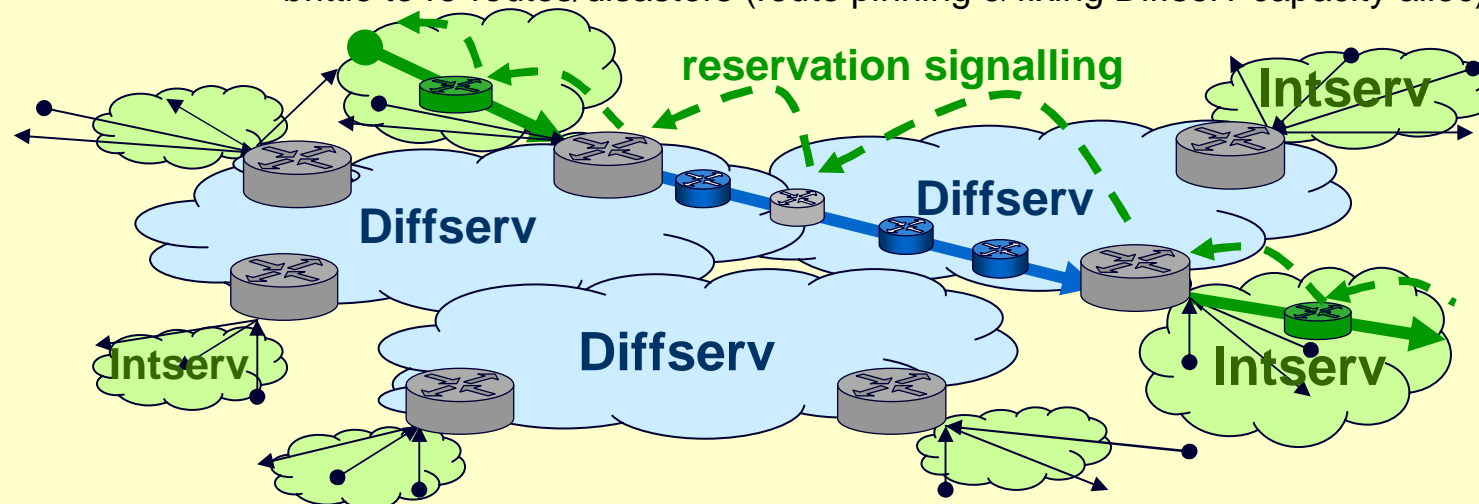
- use-case
    - An architecture for edge-to-edge controlled load service using distributed measurement-based admission control  
[draft-briscoe-tsvwg-cl-architecture-00.txt](#)
    - intention: **informational**
  - per-hop behaviour (PHB) definition
    - The controlled load per hop behaviour  
[draft-briscoe-tsvwg-cl-phb-00.txt](#)
    - intention: **standards track**
    - advice sought on best working group (assume tsvwg)
  - related to:
    - RTECN drafts from Joe Barbiarz/Kwok Chan & co, Nortel (tsvwg)
    - Load control of real-time traffic, RMD framework, Lars Westberg & co, Ericsson (nsis)
  - distinguishing features of our work
    - principled design, based on sound theoretical foundations
    - **uses standard IETF wire protocols**, but not their (informational) architectures
- 
- A diagram consisting of a horizontal line extending from the 'standards track' intention of the PHB definition, a vertical line going up from the end of that horizontal line, and another vertical line going up from the top of that vertical line to the 'informational' intention of the use-case architecture. The word 'pre-requisite' is written to the right of the vertical line.




# the problem:

controlled load service  
end to end

- voice bits initially ~50% in BT's converged network
  - presumably similar for converged internetwork
- problems in cores/backbones rare
  - unexpected traffic matrix
  - disasters/re-routes
- end-to-end admission ctrl without costly core or border mechanisms

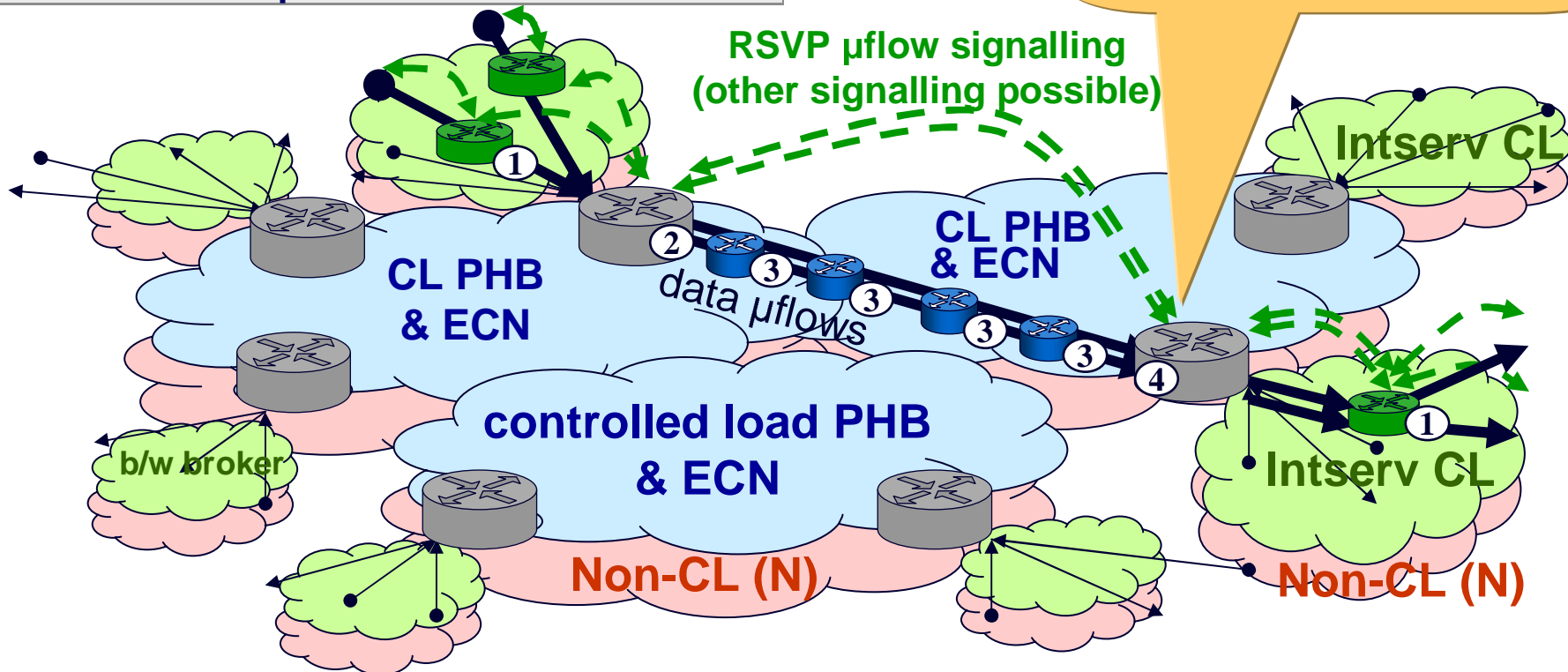
- build on Intserv over Diffserv [[RFC2998](#)], but solve hidden fudge
  - for long topologies describes how some interior nodes do CAC
  - scaling problem returns, esp at borders
  - brittle to re-routes/disasters (route pinning & fixing Diffserv capacity alloc)



IP routers	Data path processing
Reservation enabled 	① Reserved flow processing
RSVP/ECN gateway 	② Policing flow entry to CL ④ Meter ECN per aggregate
CL PHB & ECN only 	③ Bulk ECN marking CL prioritised over N

end to end controlled load (CL) service  
**system arrangement**  
 RSVP example

data aggregate identification only at egress gateway – per previous RSVP hop



- absolutely no flow state or processing within Diffserv region
- more robust than Intserv CL to re-routes/disasters

# don't jump to conclusions

- uses standard IETF wire protocols & most semantics
  - but not their (informational) architectures

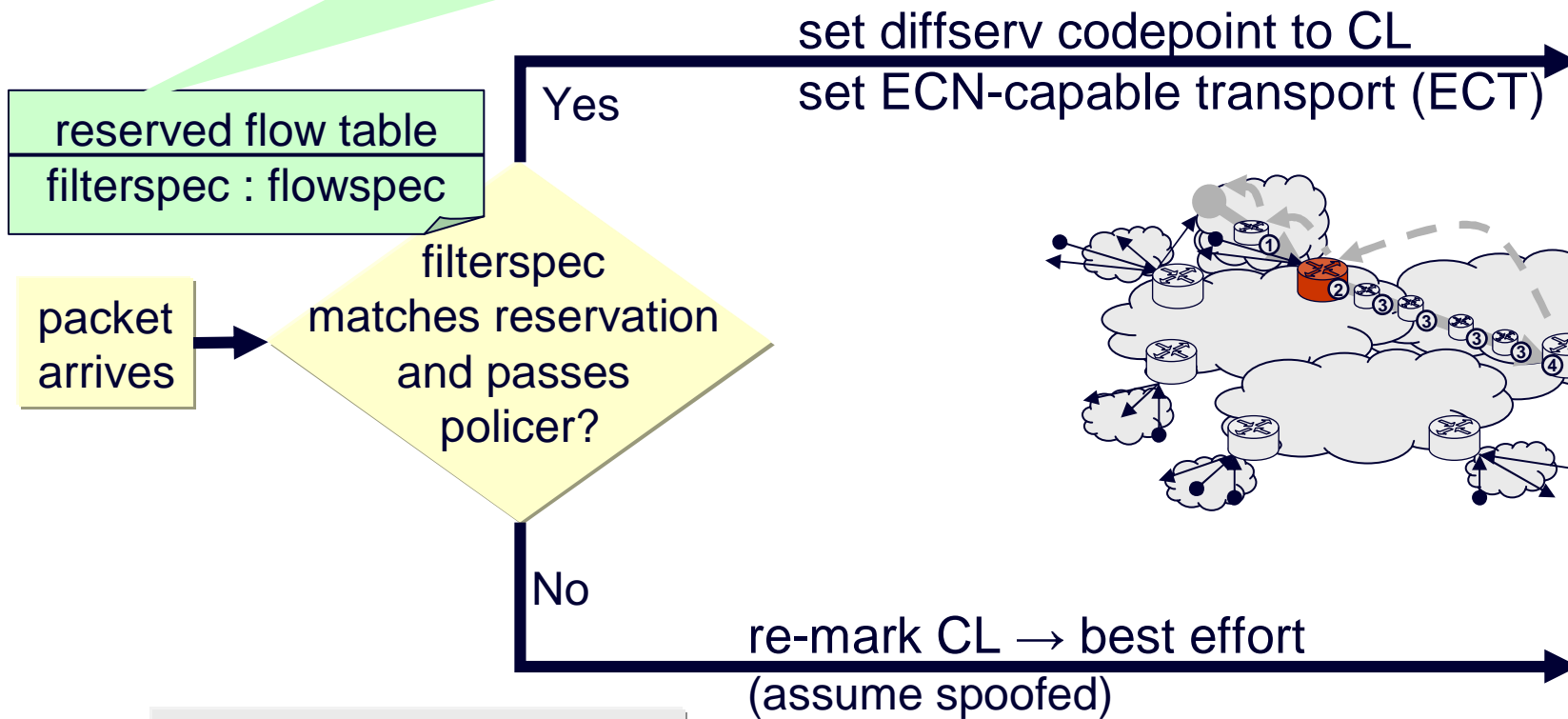
RSVP [ <a href="#">RFC2205</a> ]	DSCP [ <a href="#">RFC2474</a> ]	ECN [ <a href="#">RFC3168</a> ]
not Intserv core & borders (other signalling poss.)	not Diffserv policing & not fixed capacity alloc	edge-to-edge not end-to-end

- when you hear the words RSVP, DSCP or ECN they mean just that – the wire protocols & semantics
- BTW, this edge-to-edge scenario chosen as first step
  - to encourage ECN deployment



# data plane functions: ingress gateway

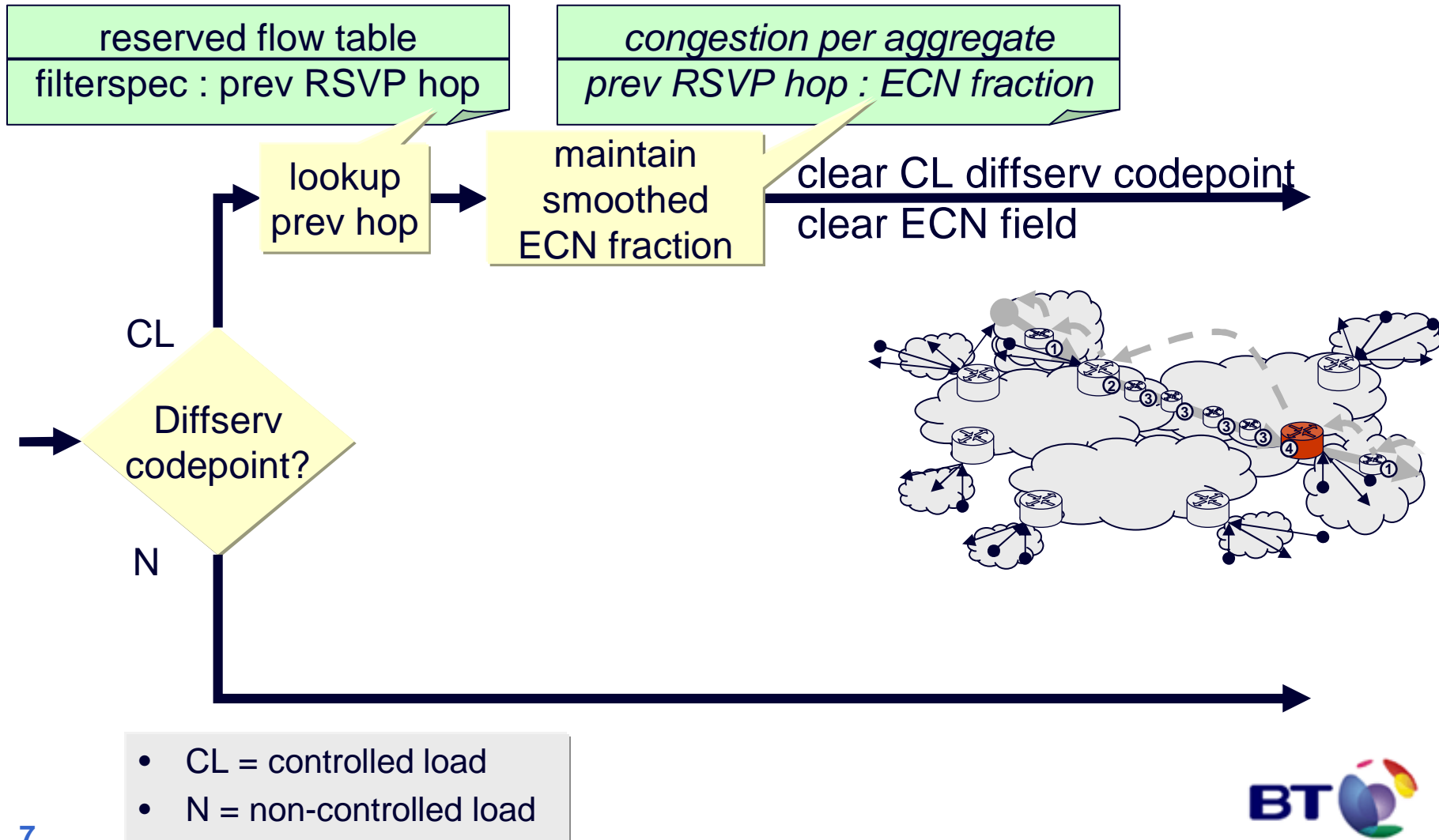
explanation easier if we start by assuming we have already admitted a flow



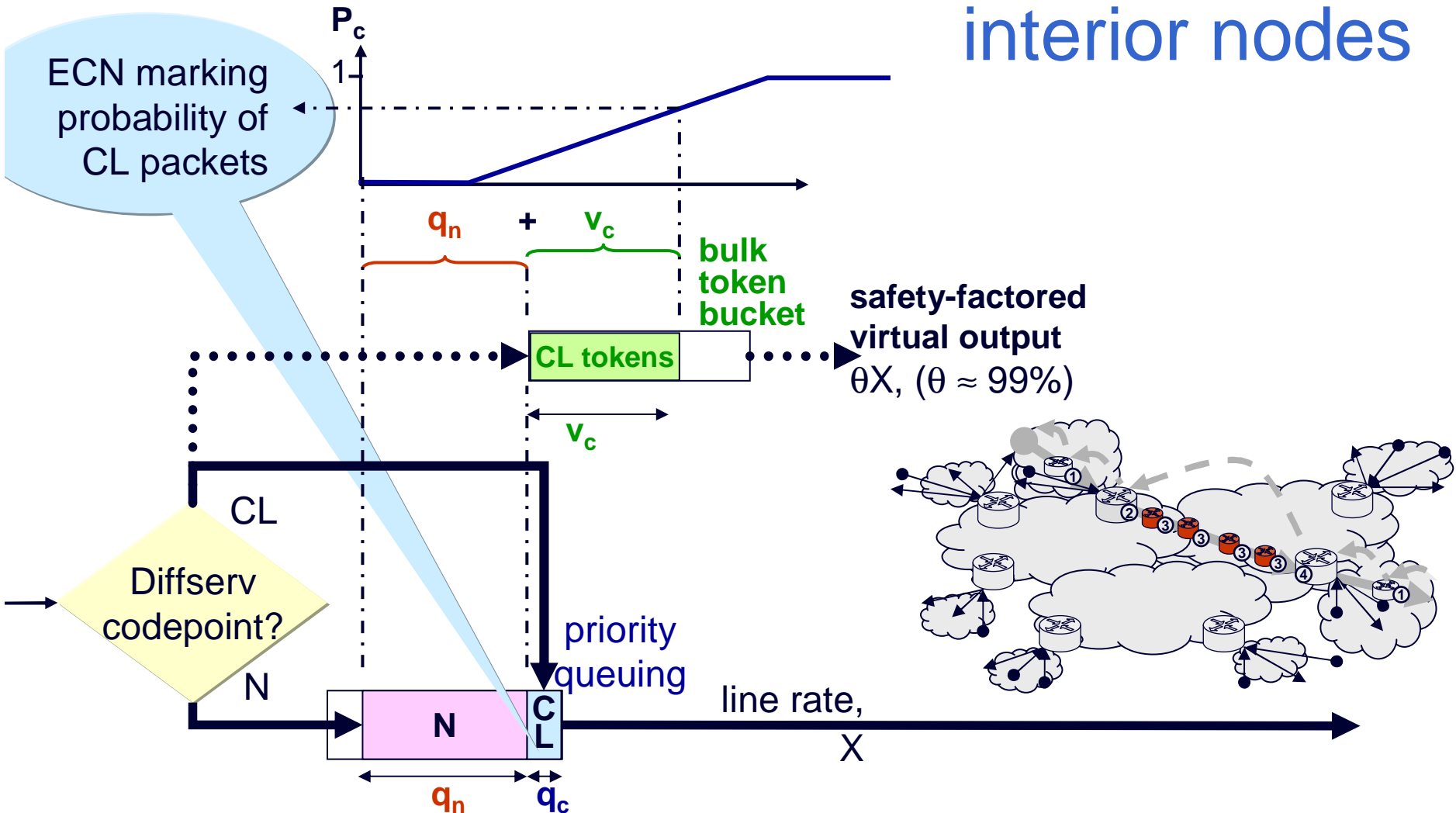
- CL = controlled load
- N = non-controlled load



# data plane functions: egress gateway






# data plane functions: interior nodes



- CL = controlled load
- N = non-controlled load

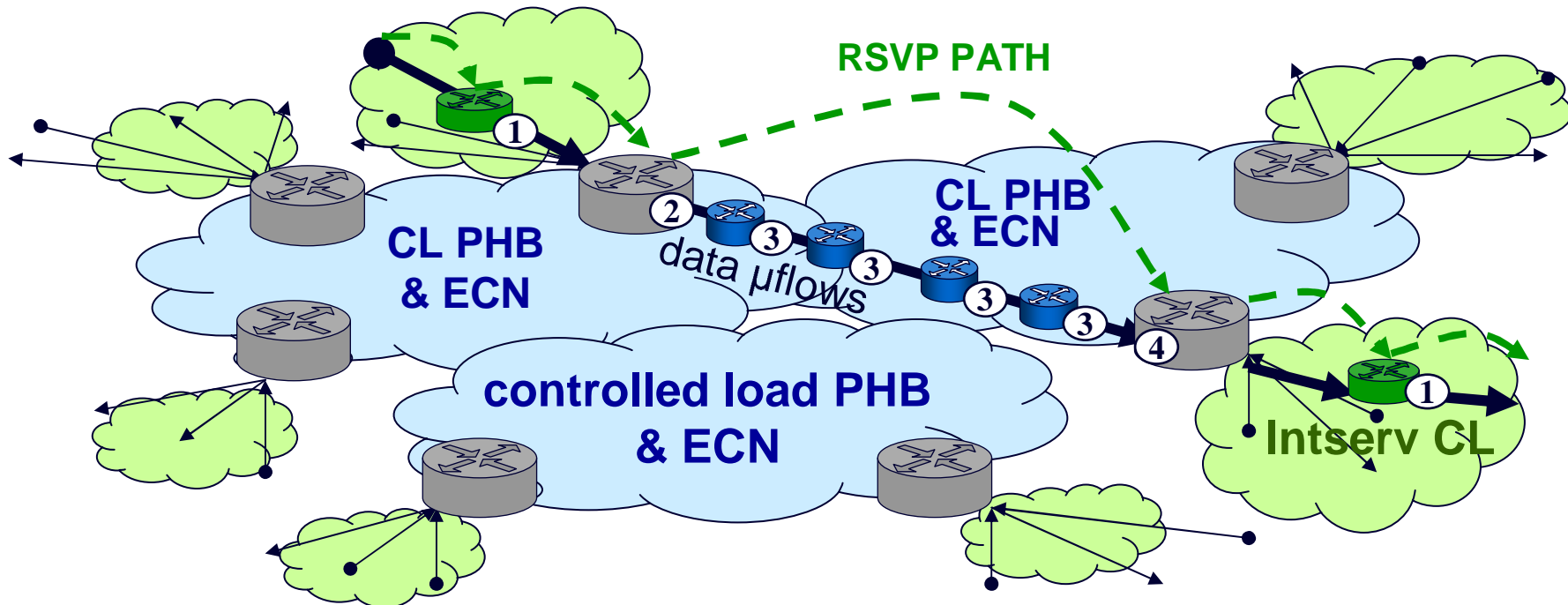







IP routers	Control signalling
Reservation enabled 	① standard RSVP PATH
RSVP/ECN gateway 	② standard RSVP PATH ④ standard RSVP PATH
CL PHB & ECN only 	③ RSVP unaware

# admission control

RSVP example (others possible)



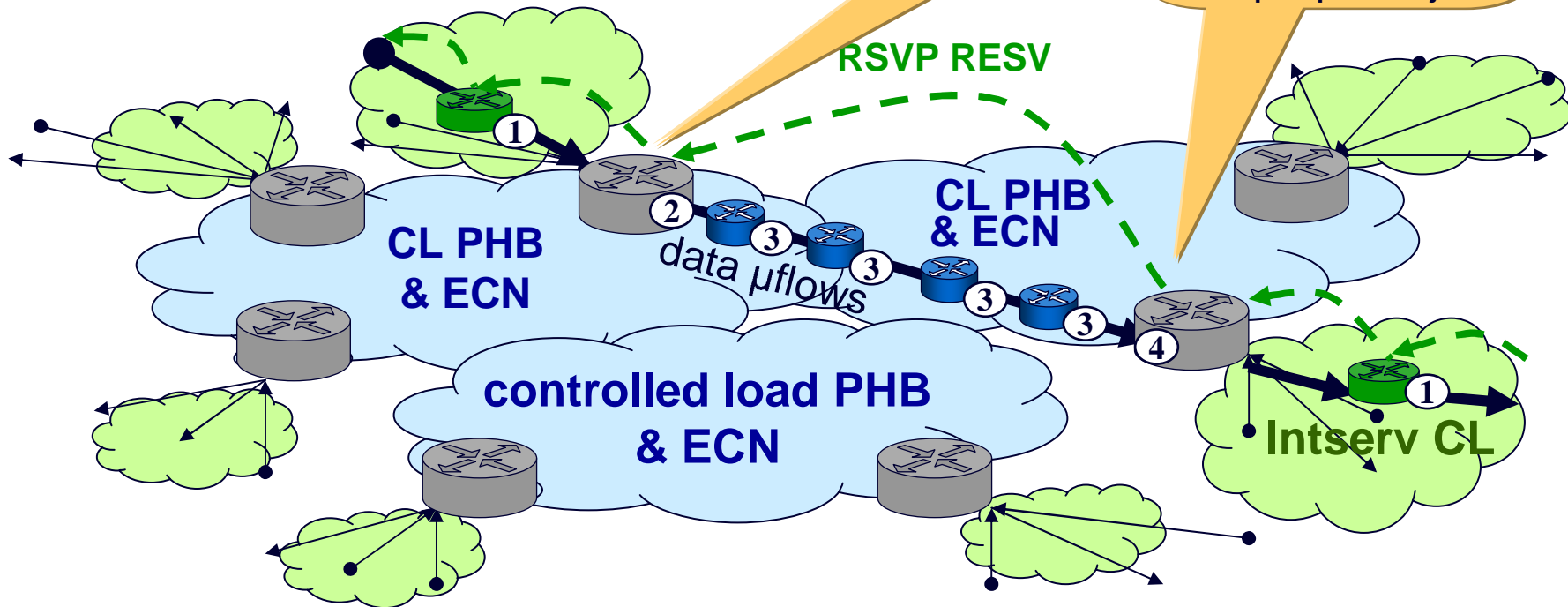
IP routers	Control signalling
Reservation enabled 	① standard RSVP RESV
RSVP/ECN gateway 	② extended RSVP RESV
CL PHB & ECN only 	③ RSVP unaware
	④ extended RSVP RESV

# admission control

RSVP example (others possible)

admit each μflow to aggregate across region

piggy-back ECN fraction as opaque object



# summary

- controlled load (CL) service
- more robust than Intserv CL
  - preserve CL service to admitted flows during re-routes
  - then allocations gracefully adapt
- no flow signalling nor state...
  - ...on core AND border routers
  - but correct admission control wherever congestion arises

## no time for...

- more cool features
  - ECN-based anti-cheating mechanism
    - passive inter-domain policing
  - incremental deployment
    - scales better as networks join
  - re-route/disaster scenarios
- design details
  - bootstrap of aggregates (probing)
  - silence suppression & VBR
  - interaction with other PHBs
    - esp. preventing starvation
  - various commercial contexts
    - charging, policy etc
- design motivations
- extensive simulation
  - most challenging simulations ever
  - scheduler, RTT & session timescales
  - many scenarios, up to 1G core
  - sudden traffic shifts
- all the above documented

# plans at IETF

## 1. controlled load (CL) PHB

- first PHB to define non-default ECN semantics

as allowed by ECN [[RFC3168](#)]:

```
...The above discussion of when CE may be set instead of dropping a packet applies by default to all Differentiated Services Per-Hop Behaviors (PHBs) [RFC 2475]. Specifications for PHBs MAY provide more specifics on how a compliant implementation is to choose between setting CE and dropping a packet, but this is NOT REQUIRED. ...
```

- administrative scoping of ECN semantics satisfies "Specifying Alternate Semantics for the ECN Field", [draft-floyd-ecn-alternates-00.txt](#)
- aiming for consensus with RTECN, RMD & others
- intended for standards track
- add ECN semantics to EF PHB [[RFC3246](#)] without changing scheduling?

## 2. extension to RSVP for opaque ECN fraction object

- is tsvwg working group appropriate (for both)?
- working group items?

# Controlled Load (CL) Service

spare slides



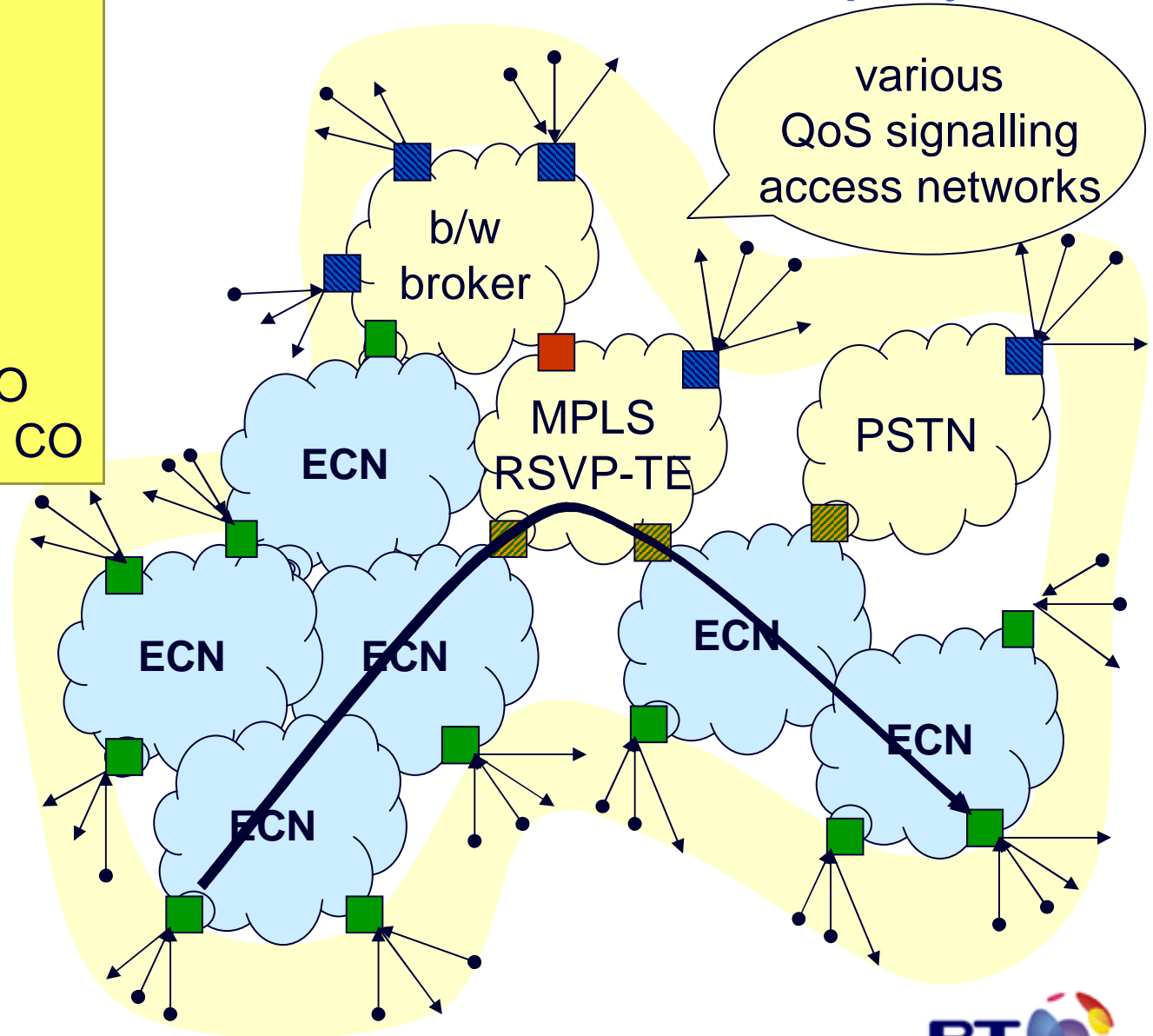
legend

- connection-oriented (CO)
- connectionless

gateways

- CL/access CO
- CL/core CO
- core CO/core CO
- access CO/core CO

# incremental deployment

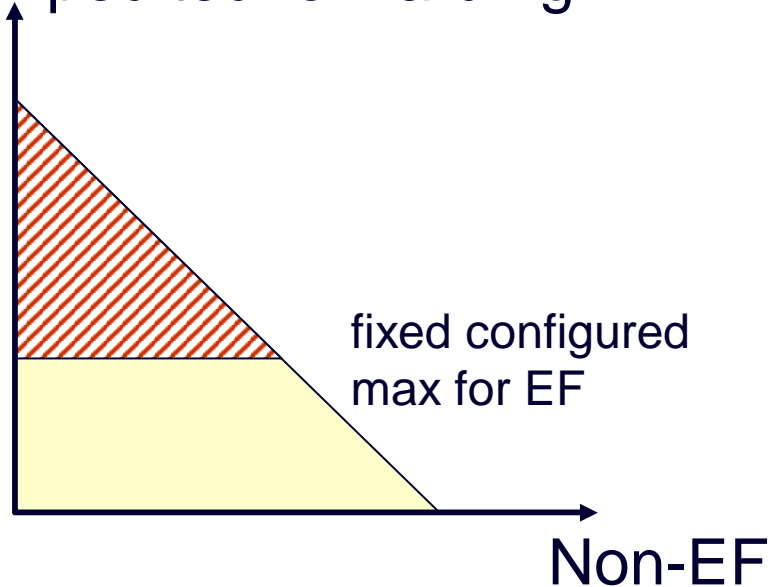


assume app layer signalling (SIP) initiates out of band



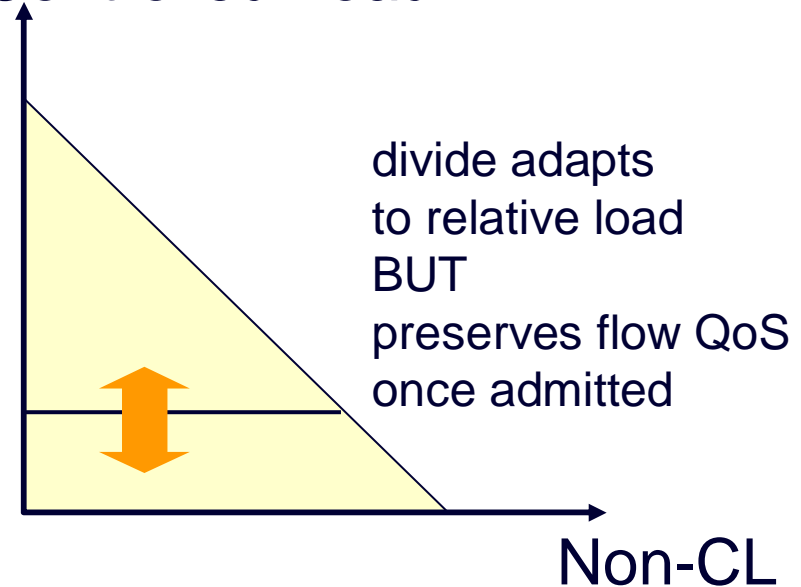
# robustness during re-routes: comparison

Expedited forwarding PHB



- fixed max
  - maps to many industry business models

Controlled Load PHB



- adaptive max
  - exactly the behaviour required for robustness during re-routes/disasters

## proposed definition of explicit congestion notification

- The congestion caused by a packet at single resource is the probability that the event  $X_i$  will occur if the packet in question is added to the load, given any pre-existing differential treatment of packets.
- Where  $X_i$  is the event that another selected packet will not be served to its requirements by the resource during its current busy period.
- This definition maps directly to economic cost
  - also usefully approximated by algorithms like RED



## congestion of capacity configured for a class or the whole resource?

- operator should be able to configure either
- fixed max (e.g. EF)
  - higher class is confined to its own resources
  - 'congestion' should mean of the class
- adaptive max (e.g. CL)
  - higher class can adapt to use lower resources
  - 'congestion' should mean of the resource the traffic could use