Communications Innovation Institute

downstream knowledge upstream re-feedback

Bob Briscoe

Arnaud Jacquet, Carla Di Cairano-Gilfedder, Andrea Soppera & Martin Koyabe

BT Research

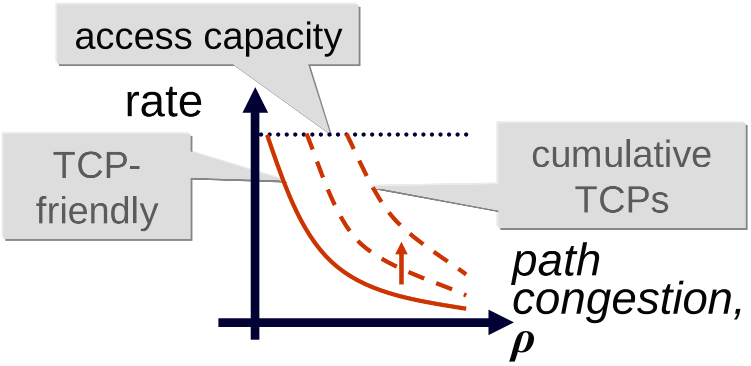
goals & non-goals, approach

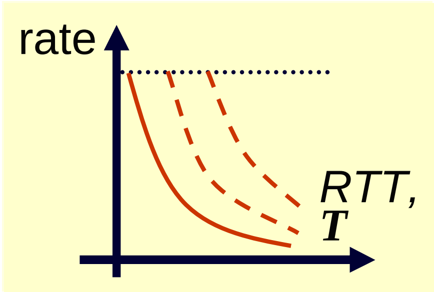
- goals & non-goals
  - goal: fix Internet's resource allocation and accountability architecture
  - non-goal: solve the whole DoS problem
  - non-goal: solve app-layer/user-space flooding
  - goal: foundation for wider DoS solution(s)
- approach

  - part of effort to determine new Internet architecture
  - mechanism for non-co-operative end-game in case things get nasty
  - network economics & incentives, but no fiddling with retail pricing
  - network operators (not users) assumed to be rational
- work in progress
  - simulations in progress
  - not even submitted yet

the problem: rate policing

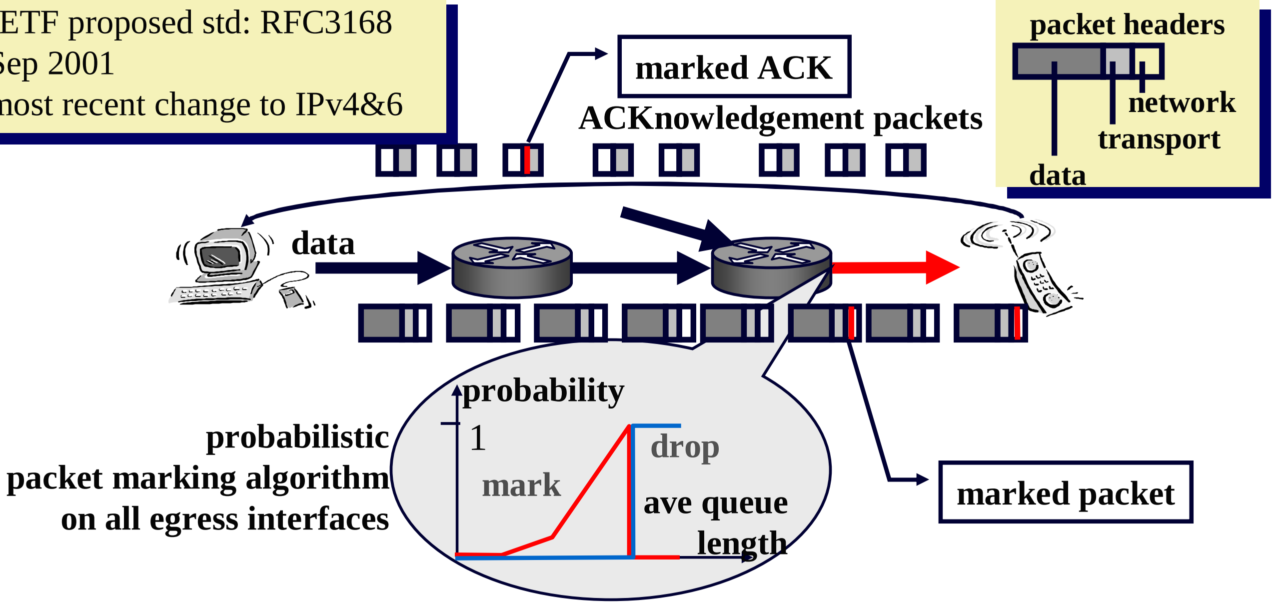
- short & long term congestion
  - short: e.g. policing TCP-friendliness (or any agreed response)
  - long: e.g. policing zombie hosts, p2p file-sharing (selfish not malicious)
- user congestion response voluntary
  - why is TCP compliance stable? what shouldn't we do to keep it?
  - TCP-friendly malware?? imagine a TCP virus
- network congestion response voluntary
  - why care if my users cause congestion in downstream networks?





pre-requisite knowledge:
explicit congestion notification (ECN)

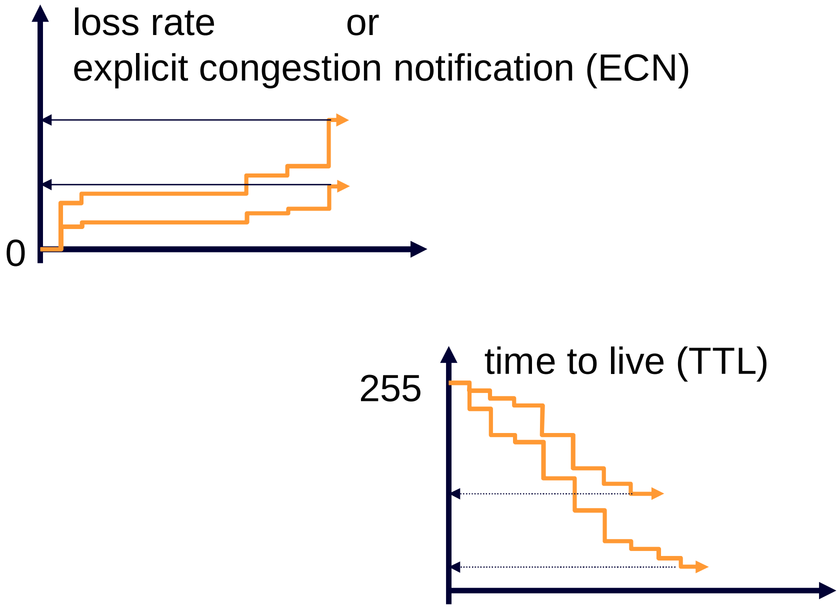
IETF proposed std: RFC3168
Sep 2001
most recent change to IPv4&6



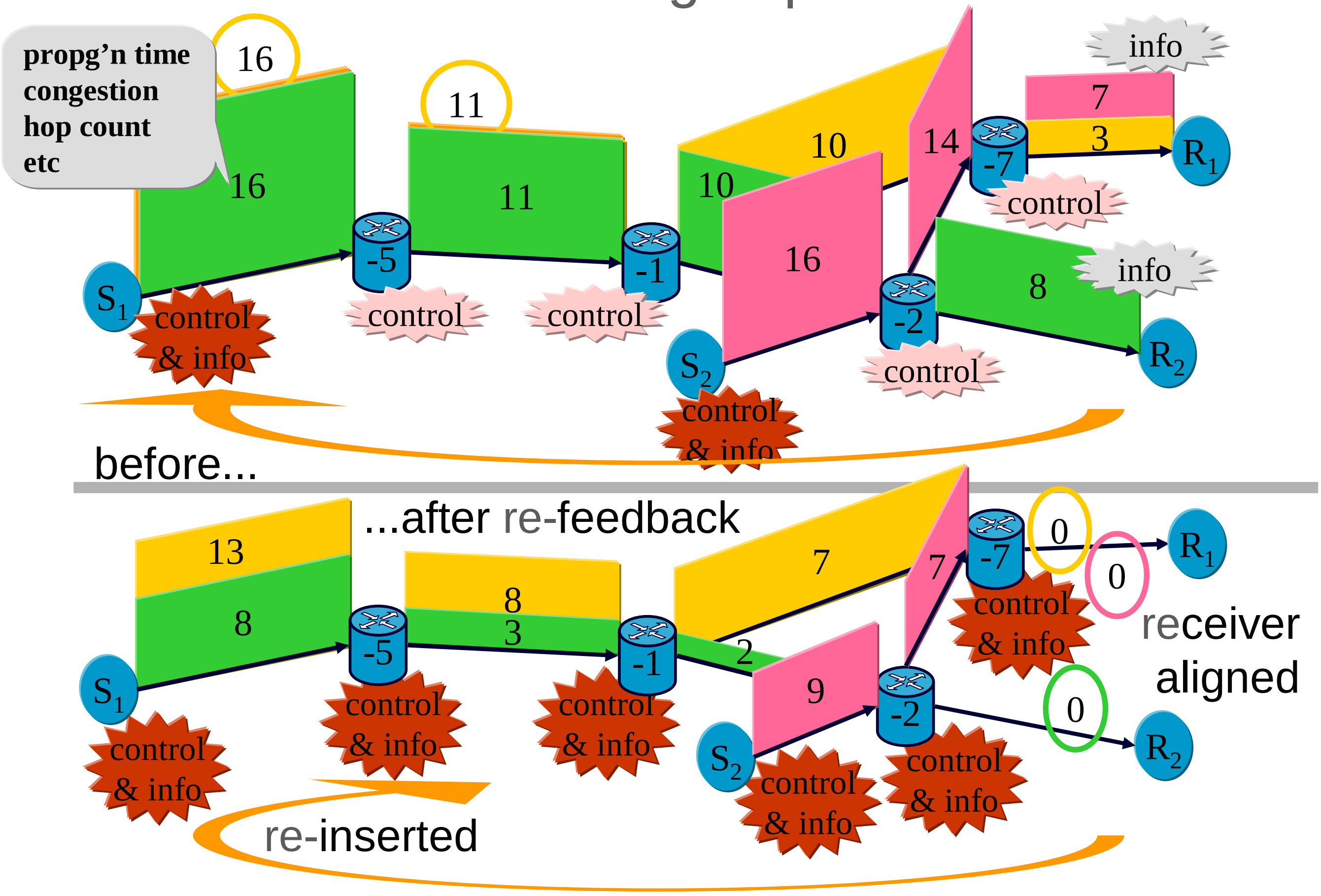
00:	Not ECN Capable Transport (ECT)
01 or 10:	ECN Capable Transport - no Congestion Experienced (sender initialises)
11:	ECN Capable Transport - and Congestion Experienced (CE)

0 5 6 7
DSCP | ECN
bits 6 & 7 of **IP** DS byte

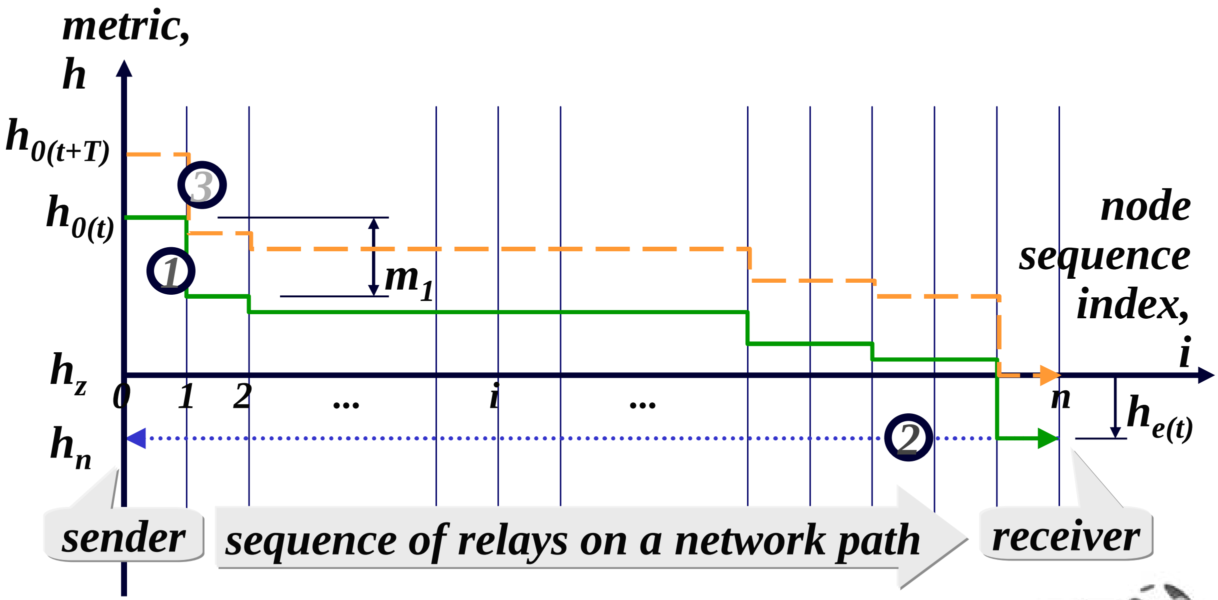
path characterisation via data headers



downstream knowledge upstream: the idea



downstream knowledge upstream — re-feedback

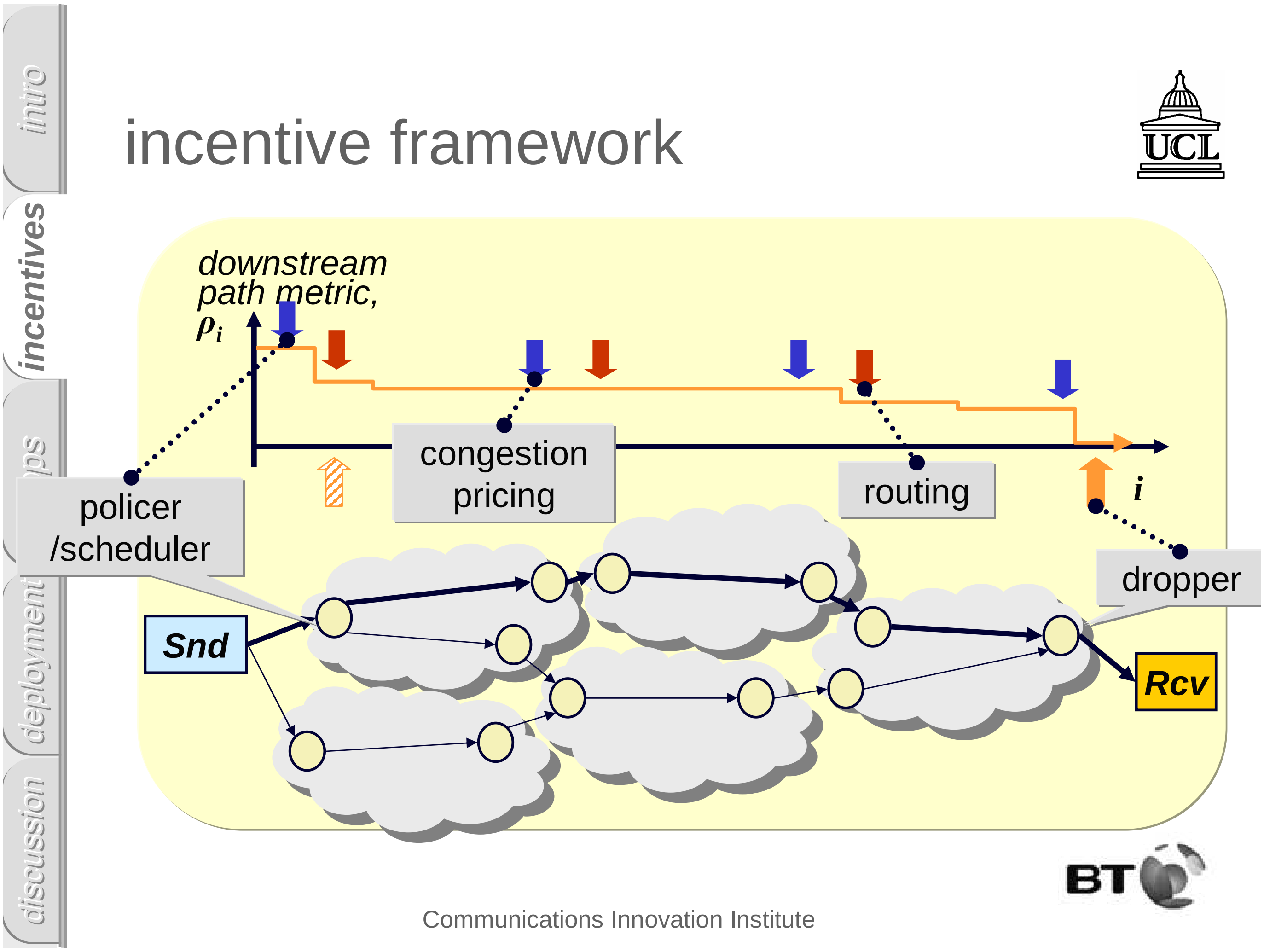


congestion protocol terms

- ECN = Explicit Congestion Notification
- ECL = Explicit Congestion Level (my term)
- ‘re-’ = receiver aligned
  (or re-inserted)

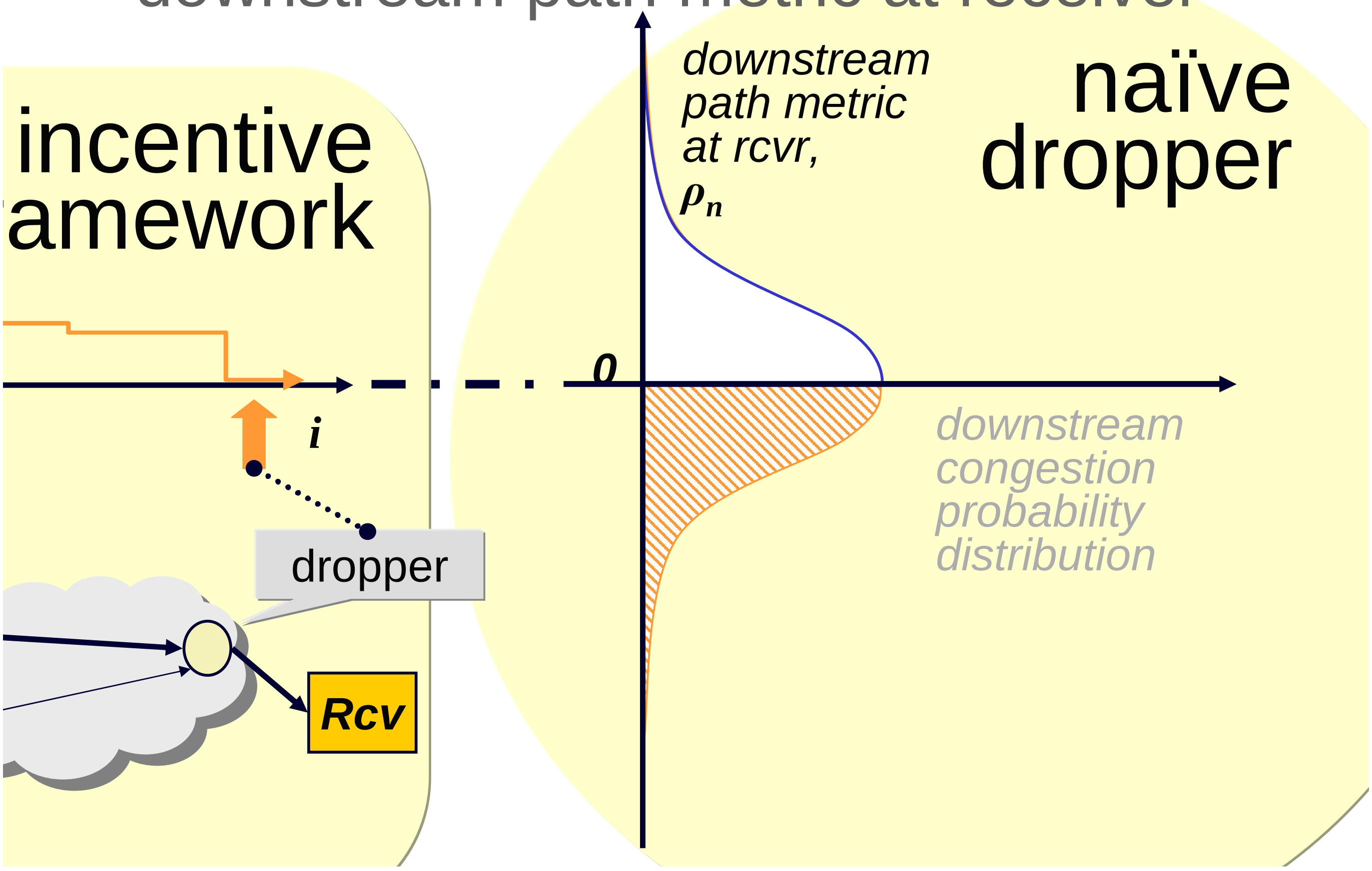
aligned at	binary	multi-bit
sender	ECN	ECL
receiver	re-ECN	re-ECL

incentive framework

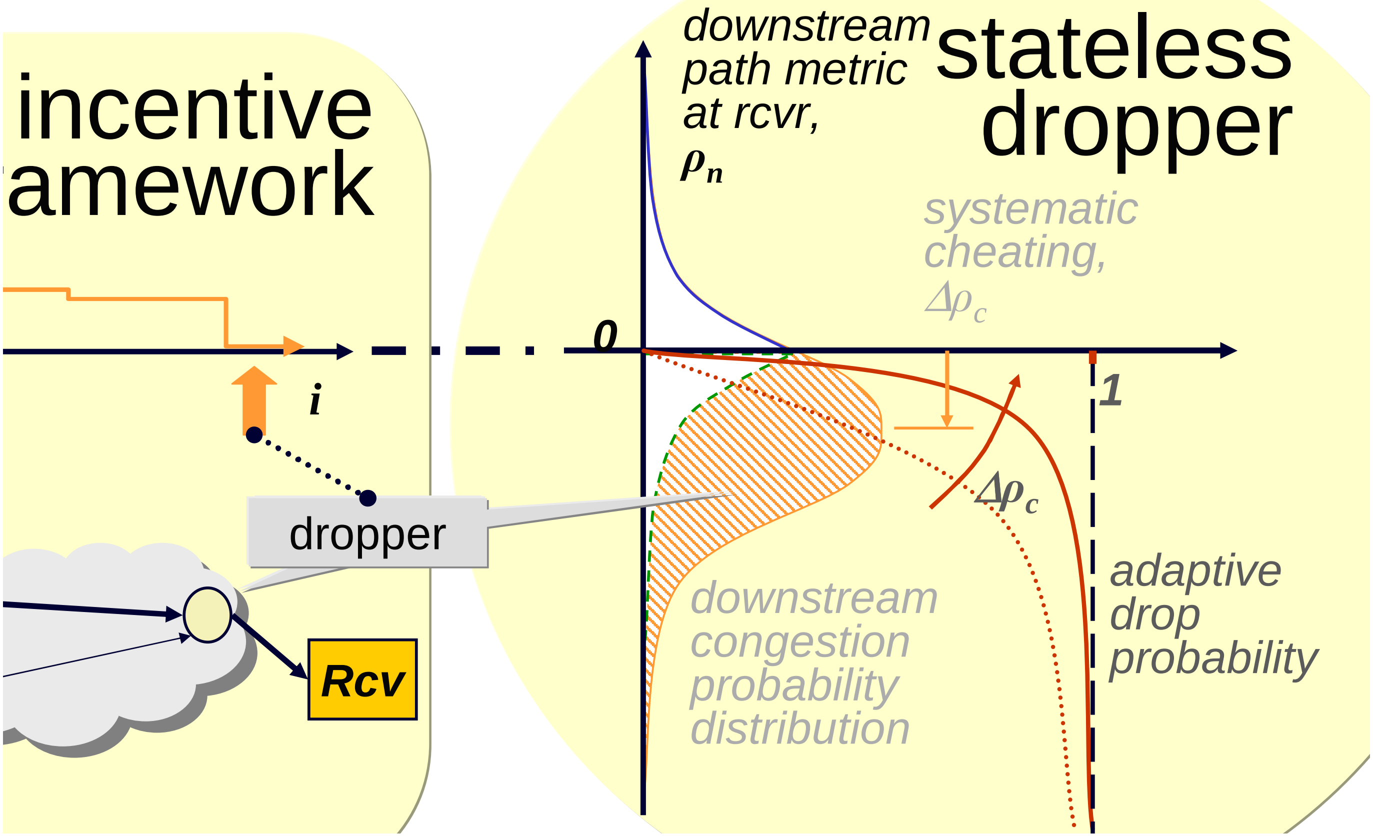


Communications Innovation Institute

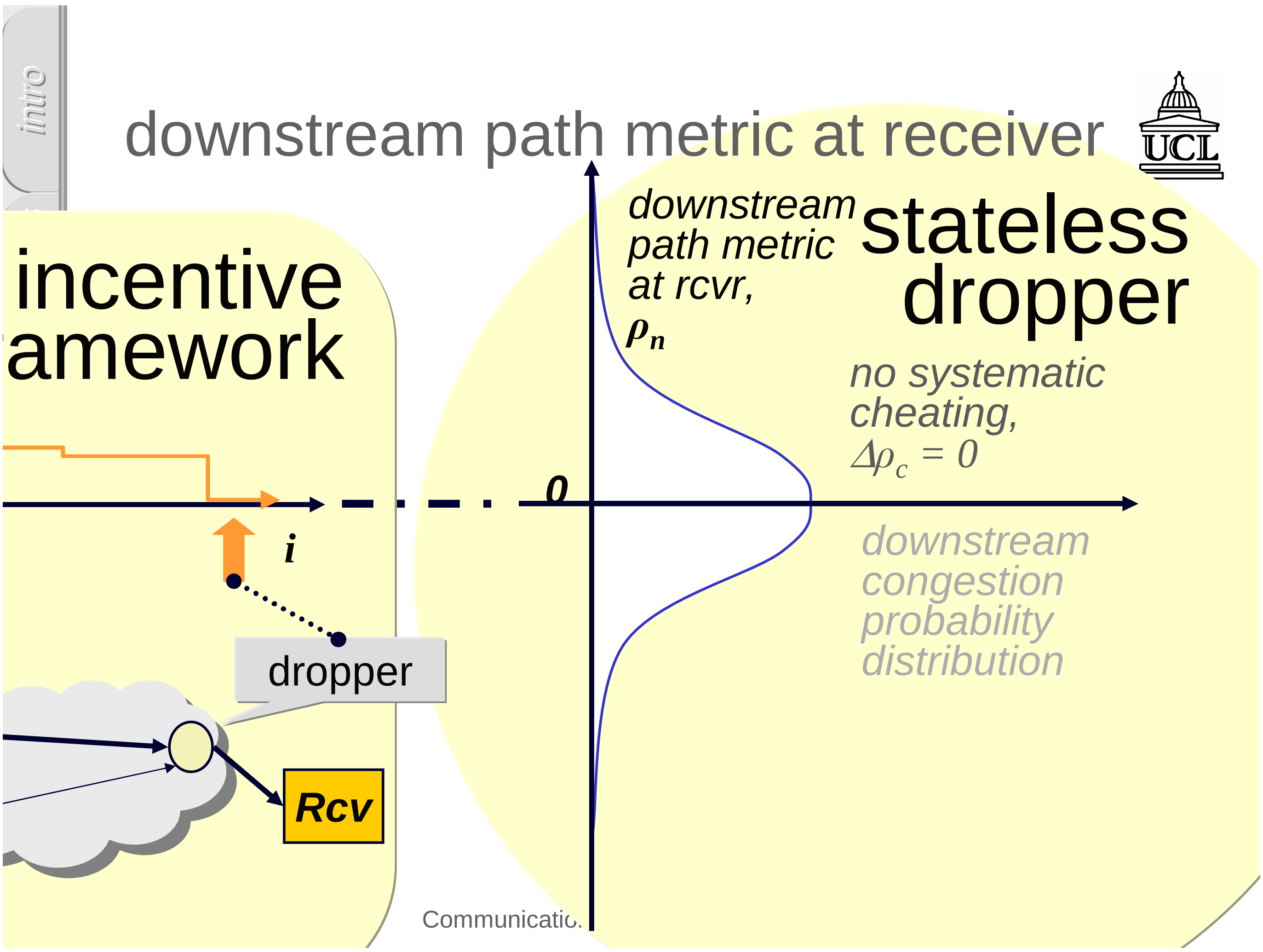
downstream path metric at receiver



penalising uncertain misbehaviour



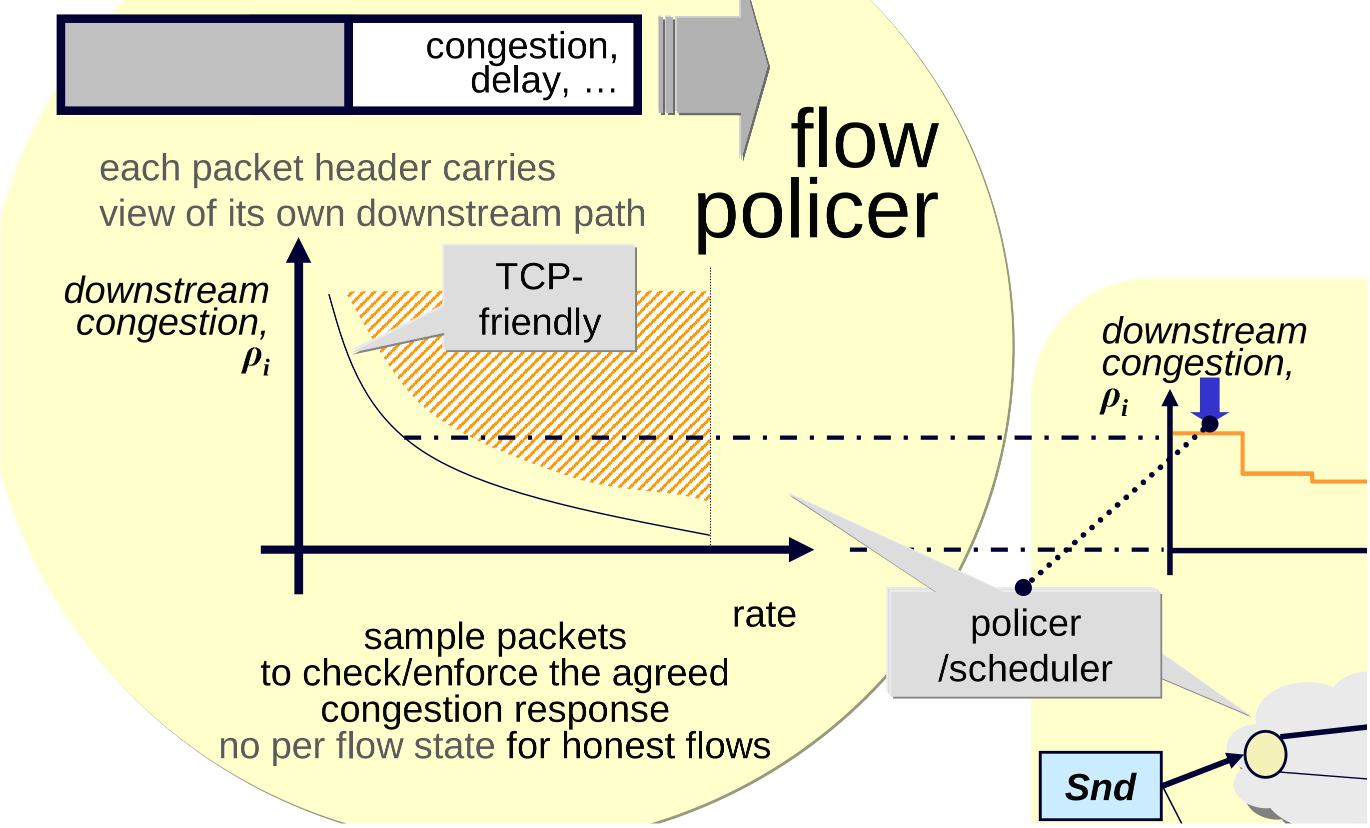
downstream path metric at receiver



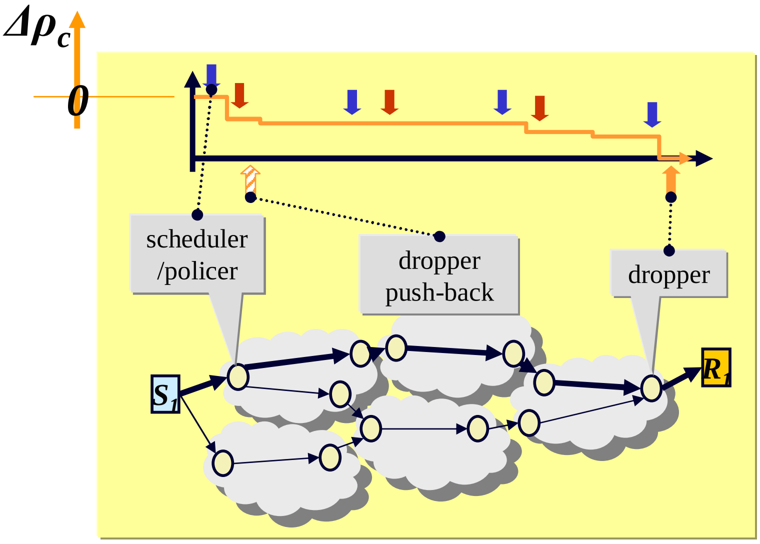
spawning focused droppers

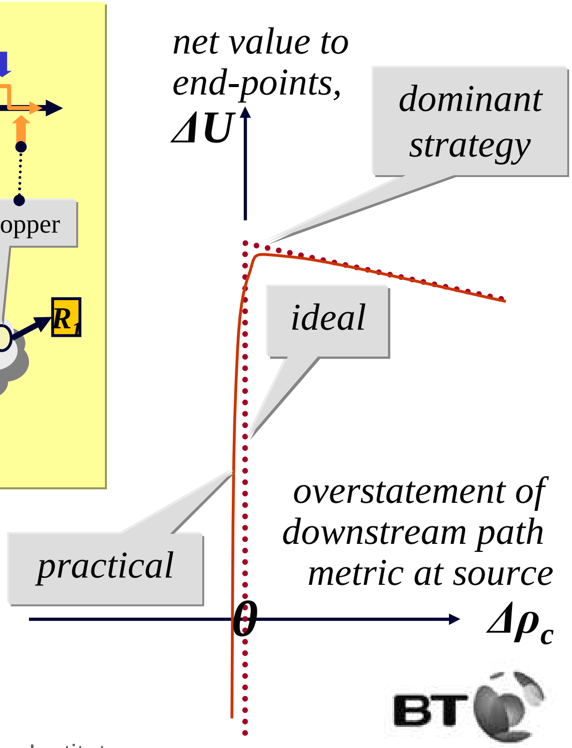
- use penalty box technique [Floyd99]
  - examine (candidate) discards for any signature
  - spawn child dropper to focus on subset that matches signature
  - kill child dropper if no longer dropping (after random wait)
- push back
  - send hint upstream defining signature(s)
  - if (any) upstream node has idle processing resource
    test hint by spawning dropper focused on signature as above
- cannot DoS with hints, as optional & testable
  - no need for crypto authentication – no additional DoS vulnerability

TCP policer – no state for well-behaved flows



incentive compatibility – hosts



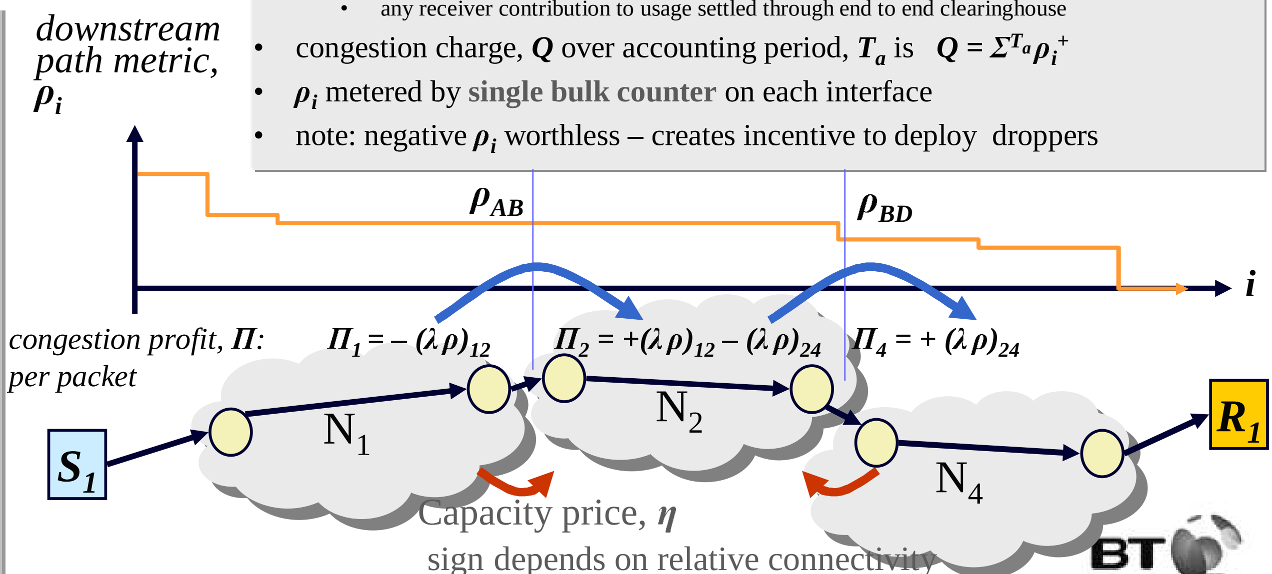


- incentivise:
  - responsible actions
  - honest words

intro
incentives
apps
deployment
discussion

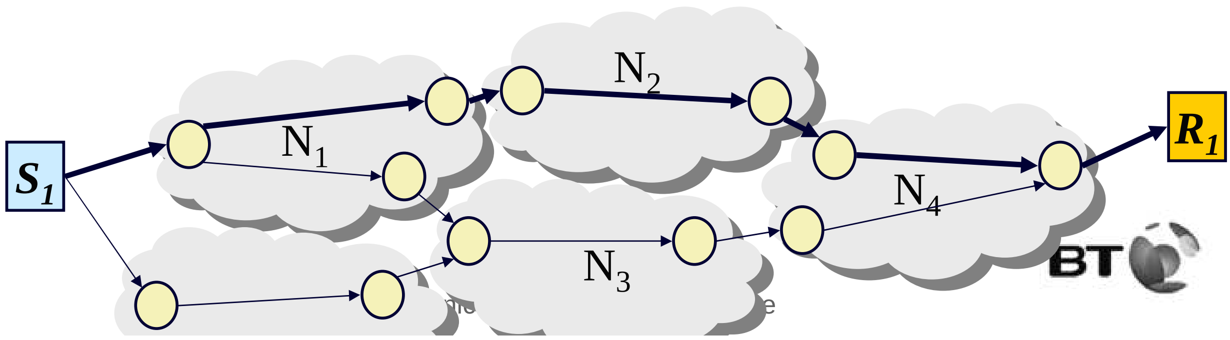
inter-domain policing

- **bulk congestion charging emulates policing: passive & simple**
- capacity charge modulated by congestion charge
- sending domain pays $C = \eta X + \lambda Q$ to receiving domain (e.g. monthly)
- η, λ are (relatively) fixed prices of capacity, X and congestion, Q resp.
  - 'usage' related price $\lambda \geq 0$ (safe against 'denial of funds')
  - any receiver contribution to usage settled through end to end clearinghouse
- congestion charge, Q over accounting period, T_a is $Q = \Sigma^{T_a} \rho_i^+$
- ρ_i metered by **single bulk counter** on each interface
- note: negative ρ_i worthless – creates incentive to deploy droppers

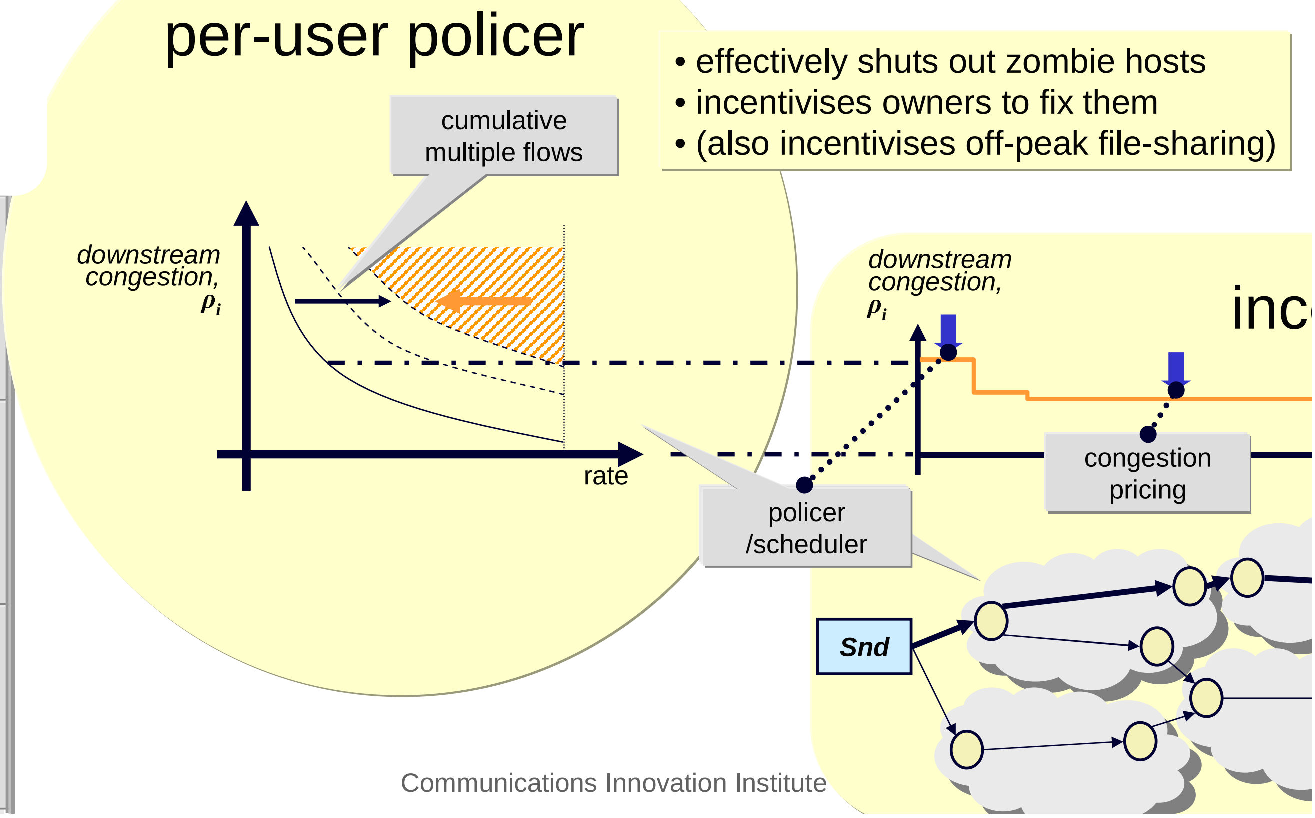


incentive compatibility – inter-domain routing

- why doesn't a network overstate congestion?
  - **msecs**: congestion response gives diminishing returns (for TCP: $\Delta\Pi \propto \sqrt{\Delta\rho}$)
  - **minutes**: upstream networks will route round more highly congested paths
    - by sampling data N_1 can see relative costs of paths to R_1 thru N_2 & N_3
  - **months**: persistent overstatement of congestion:
    - artificially reduces traffic demand (thru congestion response)
    - ultimately reduces capacity element of revenue
- also incentivises provision, to compete with monopoly paths



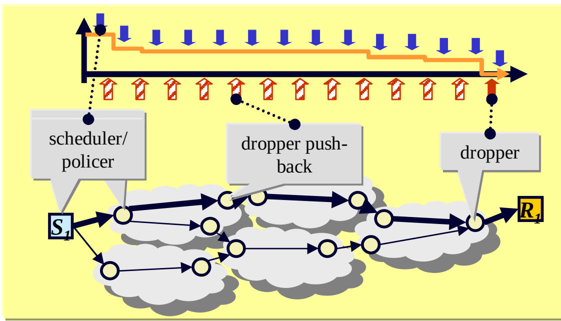
long term congestion incentives



incentives for other metrics

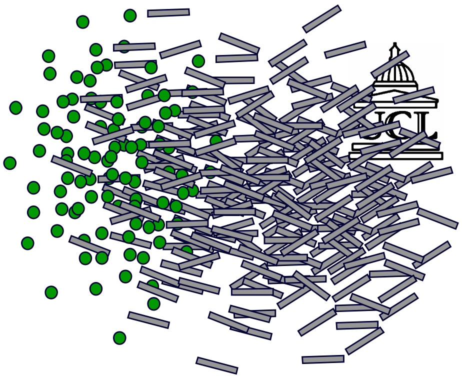
- downstream unloaded delay (emulated by TTL)
  - approximates to ½ feedback response time (near source) ⇒ RTT
  - each node can easily establish its local contribution
  - identical incentive properties to congestion
    - increasing response time increases social cost
    - physically impossible to be truthfully negative
  - incentive mechanism identical to that of congestion
- assess other metrics case-by-case

slow-enough-start



- initial value of metric(s) for new flows?
  - undefined – deliberately creates dilemma
  - if too low, may be dropped at egress
  - if too high, may be deprioritised at ingress
- without re-feedback (today)
  - if congested: all other flows share cost equally with new flow
  - if not congested: new flow rewarded with full rate
- with re-feedback
  - risk from lack of path knowledge carried solely by new flow
  - creates slow-start incentive
  - once path characterised, can rise directly to appropriate rate
  - also creates incentive to share path knowledge
  - can insure against the risk (see differentiated service)

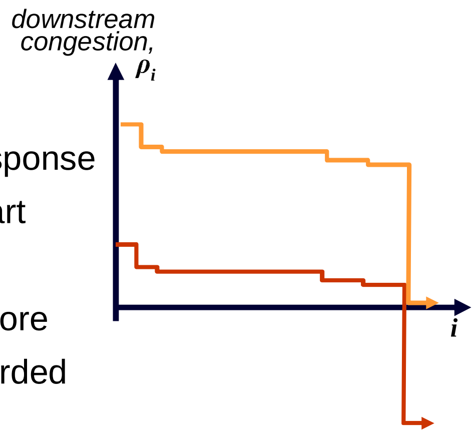
single datagram-dominated traffic mix



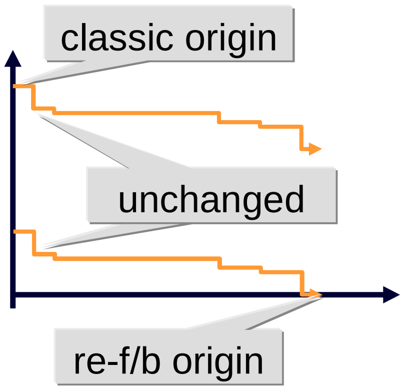
- current Internet would collapse
  - not designed for all eventualities
  - 10^{12} devices, 10^{9} users, RPCs, sensor nets, event avalanches
- with re-feedback
  - service protected against completely uncorrelated traffic mix
  - demanding users can still insure against risk
- for brief flows, TCP slow start sets rate limit
  - …not technology performance advances
  - with re-feedback, once characterised path, can hit full rate

distributed denial of service

- merely enforcing congestion response
- honest sources
  - increase initial metric & reduce rate
- malicious sources
  - if do increase initial metric
    - policer at attacker's ingress forces rate response
    - have to space out packets even at flow start
  - if don't increase initial metric
    - negative either at the point of attack or before
    - distinguished from honest traffic and discarded
    - push back kicks in if persistent



migration



- approach
  - realign metrics by modifying sender and/or receiver stack only
  - unchanged router path characterisation (protocol & routers)
  - re-ECN possible without contravening existing ECN code-points
  - reason: changing hosts: incremental; changing routers: flag day
- deployment path
  - network operators add incentive mechanisms to edge routers
  - add policers & droppers, but permissively configured
  - increasing strictness incentivises incremental host upgrades

summary
re-feedback incentive framework