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

marked ACK
ACKnowledgement packets

marked packet

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)

bits 6 & 7 of IP DS byte
path characterisation via data headers

- loss rate
- explicit congestion notification (ECN)

255 time to live (TTL)

Communications Innovation Institute

downstream knowledge upstream: the idea

prop’n time congestion hop count etc

before...

...after re-feedback

receiver aligned

re-inserted
downstream knowledge upstream — re-feedback

Communications Innovation Institute

congestion protocol terms

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

<table>
<thead>
<tr>
<th>aligned at</th>
<th>binary</th>
<th>multi-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>sender</td>
<td>ECN</td>
<td>ECL</td>
</tr>
<tr>
<td>receiver</td>
<td>re-ECN</td>
<td>re-ECL</td>
</tr>
</tbody>
</table>
incentive framework

downstream path metric, $\rho_i$

congestion pricing

routing

dropper

policer/scheduler

incentives

apps

deployment

discussion

Communications Innovation Institute

incentive framework

downstream path metric at receiver

naïve dropper

incentive framework

$\text{Snd}$

$\text{Rcv}$

dropper

downstream congestion probability distribution

downstream path metric at $\text{rcvr}$, $\rho_n$
penalising uncertain misbehaviour

incentive framework

adaptation drop probability

systematic cheating, $\Delta \rho_c$

stateless dropper

downstream congestion probability distribution

downstream path metric at receiver, $\rho_n$

incentive framework

no systematic cheating, $\Delta \rho_c = 0$
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
incentive compatibility – hosts

- incentivise:
  - responsible actions
  - honest words

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)
- \( \eta, \lambda \) are (relatively) fixed prices of capacity, \( X \) and congestion, \( Q \) resp.
  - ‘usage’ related price \( \lambda \geq \theta \) (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 \( Q = \sum_{T_a} \rho_i \)
- \( \rho_i \) metered by single bulk counter on each interface
- note: negative \( \rho_i \) 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 around 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 (through congestion response)
    - ultimately reduces capacity element of revenue
- also incentivises provision, to compete with monopoly paths

![Diagram of inter-domain routing](image)

long term congestion incentives

**per-user policer**

- effectively shuts out zombie hosts
- incentivises owners to fix them
- (also incentivises off-peak file-sharing)
incentives for other metrics

- downstream unloaded delay (emulated by TTL)
  - approximates to $\frac{1}{2}$ feedback response time (near source) \( \Rightarrow \text{RTT} \)
  - 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