Aggregate Based Congestion: Detection and Control

http://www.aciri.org/pushback

Ratul Mahajan, Steve Bellovin,
Sally Floyd, John Ioannidis,
Vern Paxson, and Scott Shenker

What is an Aggregate?

Collection of packets from one or more flows with a common property.

Examples:

- all TCP SYN packets going to 128.95/16
- all DNS packets

What is Aggregate Based Congestion?

Congestion attributable to a particular aggregate

- Flash crowds: non-malicious
  - Victoria’s Secrets webcast, Starr report, Pathfinder launch
- DoS attacks: malicious
  - February 2000, today (?)

Congestion signature: description of aggregate responsible for congestion

What is *not* Aggregate Based Congestion?

Congestion caused by undifferentiated overall increase in traffic

- under-provisioned link
- hardware failures (e.g. fiber cuts, router crashes)

Why Traditional Congestion Control Schemes are not an answer?

- An aggregate can be composed of any number of flows
- No well-defined definition of an aggregate
- No well-defined fairness goal for aggregates

It has to be a reactive scheme. Comes into play when congestion occurs.

Goal/Motivation

Protect the rest of the traffic from the adverse impact of aggregate based congestion

Two types of “rest of traffic”

- unrelated traffic (local ACC)
- innocent traffic within the aggregate (pushback)

Related Work

- Source filtering
  - ingress, egress
- Traceback
  - IP Traceback, ICMP Traceback ….
- Input Debugging
- Web Caching, CDNs

Aggregate Based Congestion Control (ACC)

1. Kicks in on severe sustained congestion
2. Try to identify the responsible aggregate(s)
3. Compute bandwidth limit
4. Apply rate limiting
5. Decide whether to invoke pushback
6. Periodically, review situation

Severe Congestion

1. Monitor the queue’s packet drop rate over an interval
2. Severe congestion: when drop rate goes above a threshold

Identification

- Congestion signature has a destination component

- Fact: most routing table entries are 24 bits
- Observation: most sites can be described by one or more 24+ bit envelopes

Identification(2): Algorithm

1. Consider all high bandwidth 32 bit addresses
2. Now consider their 24 bit prefixes
3. Go down the prefix tree to find a longer prefix
4. Go up the prefix tree to find a smaller prefix for multiple prefixes
5. Sort the clusters obtained

Deciding the Bandwidth Limit

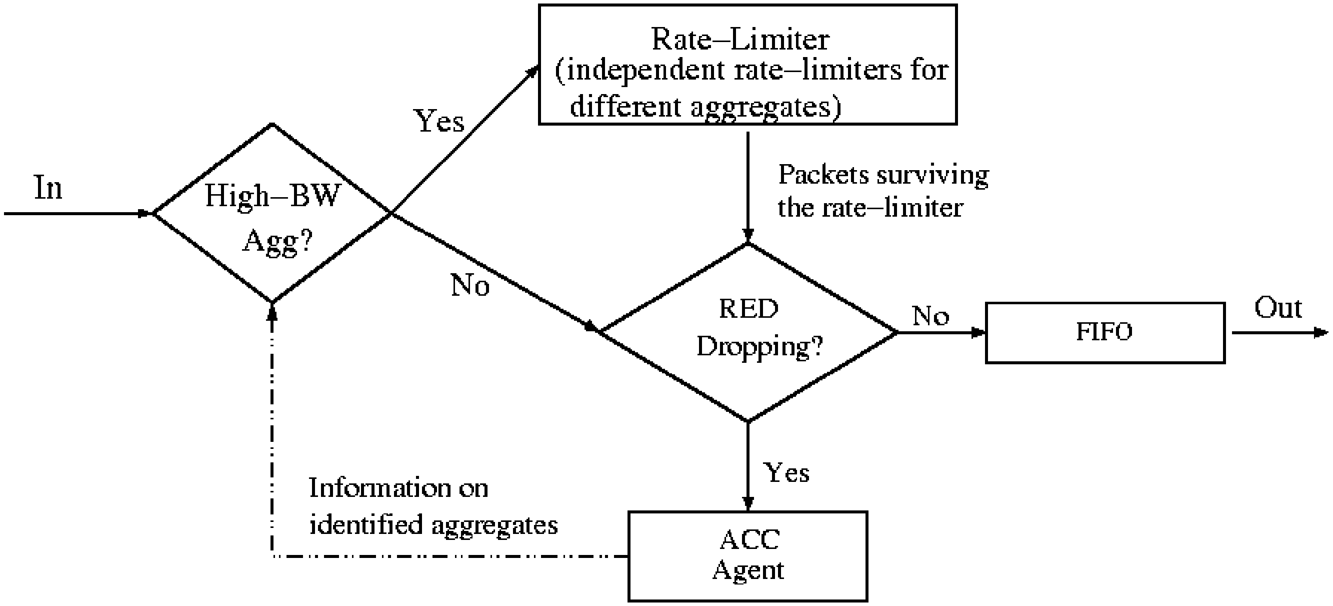
- From the sorted list of clusters pick the top *i* clusters such that rate-limiting them to the sending rate of *(i+1)st* cluster brings down the drop rate at the queue to acceptable levels
- Highly policy dependent
- How to distinguish undifferentiated congestion from aggregate-based congestion?

Rate Limiting

- As a *virtual queue* before output queue
  - a FIFO virtual queue is a token bucket
  - present in CISCO routers as CAR

- Can focus more on some subset of packets within an aggregate

Architecture



In
High–BW Agg?
Yes
No
Rate–Limiter (independent rate–limiters for different aggregates)
Packets surviving the rate–limiter
RED Dropping?
No
FIFO
Out
Yes
ACC Agent
Information on identified aggregates

- Never starves an aggregate
- Never favors an aggregate

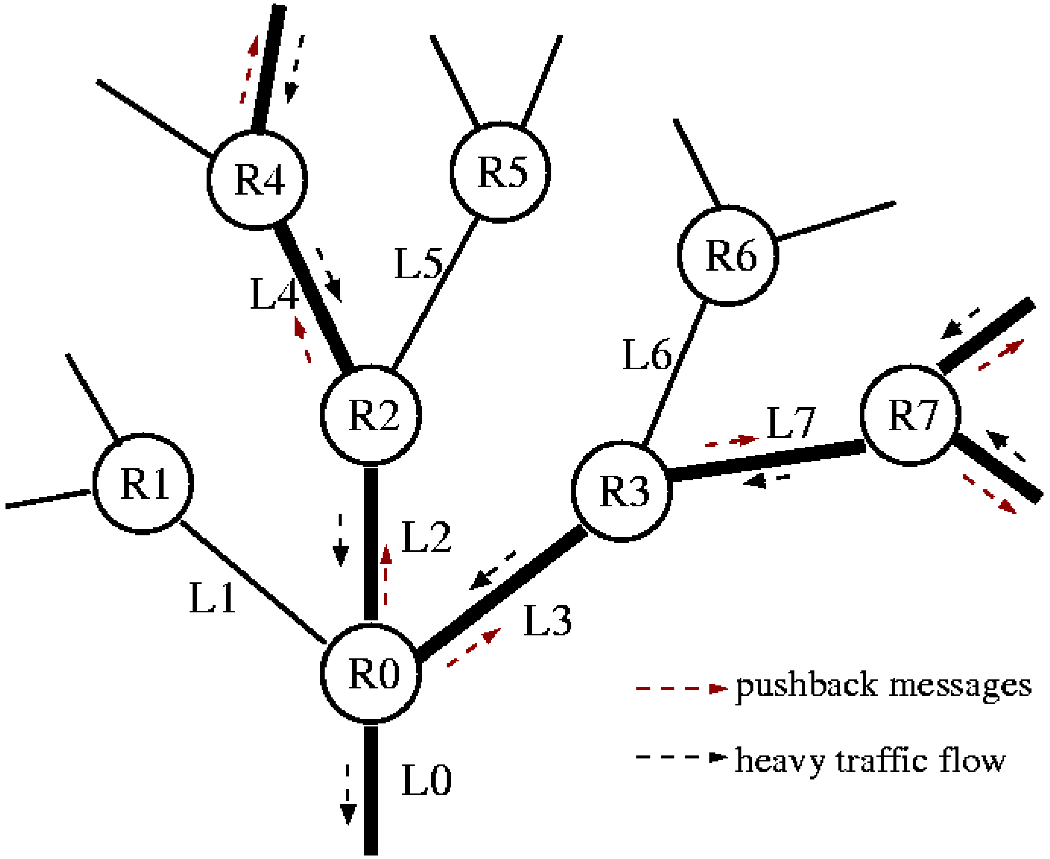
Pushback: Motivation

- Local ACC protects only the unrelated traffic
- Need to protect the innocent traffic within the aggregate

Pushback: Mechanism

- Main idea: Spatial narrowing of congestion signature
- Congested routers request high-sending upstream routers to rate-limit the aggregate
- Recursively propagates upstream towards the source(s) of the aggregate

Pushback in Action



Pushback: Advantages

- Concentrates rate-limiting on malicious traffic within an aggregate (works even with spoofed source addresses)
- Improves network utilization

Pushback: Limitations

- Overcompensation
- Less effective for diffuse attacks
- A new DoS attack?
  - prevent a source from sending to a destination

When to Invoke Pushback?

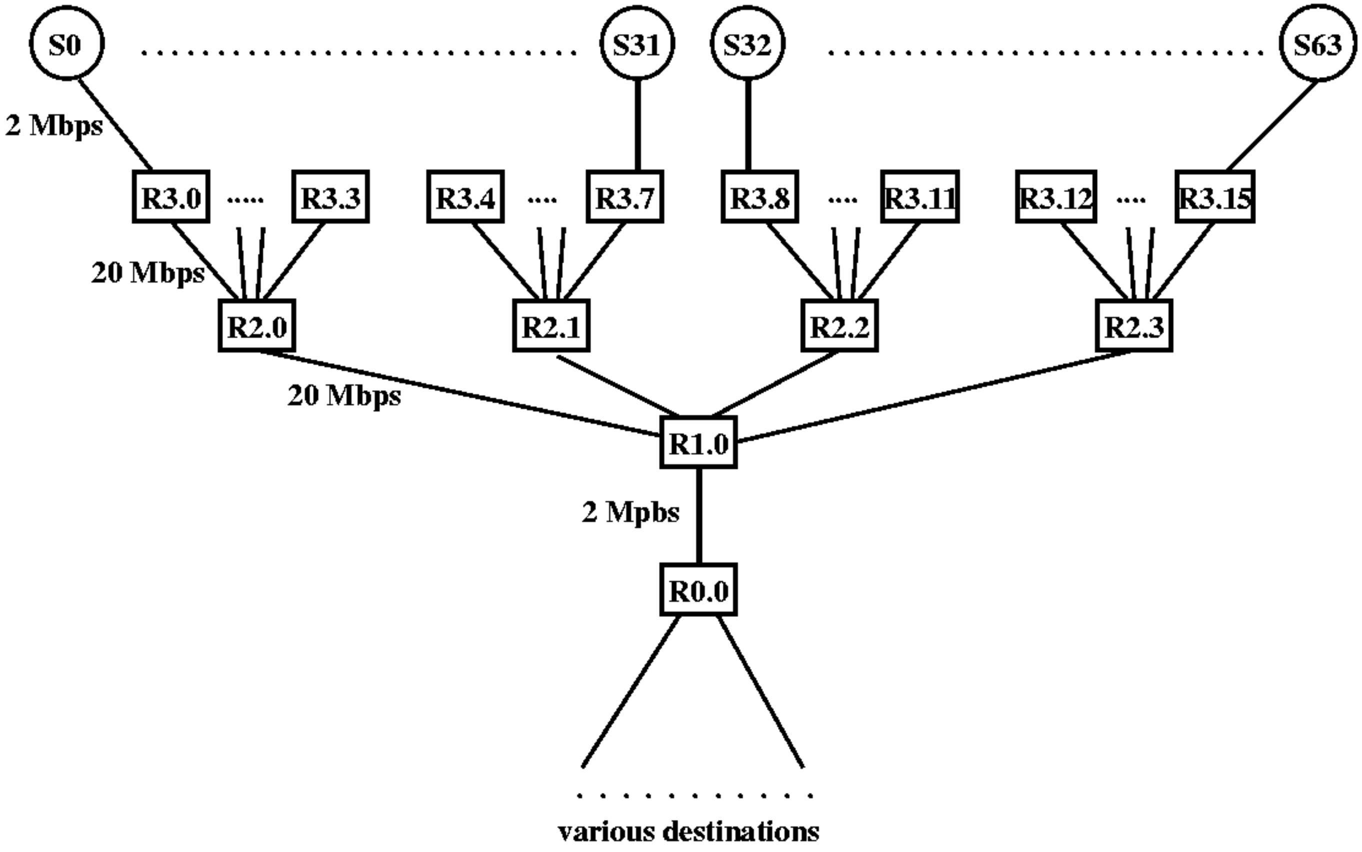
- High drop rates in the rate-limiter for the aggregate
- Out-of-band information

Pushback: Messages

- Request: goes one hop upstream
- Status: goes one hop downstream
  - arrival rate estimate
  - helps the congested router take decisions
  - aggregate received status messages from all upstream neighbors and pass downstream

1. All messages between directly connected routers
2. Very little control message overhead

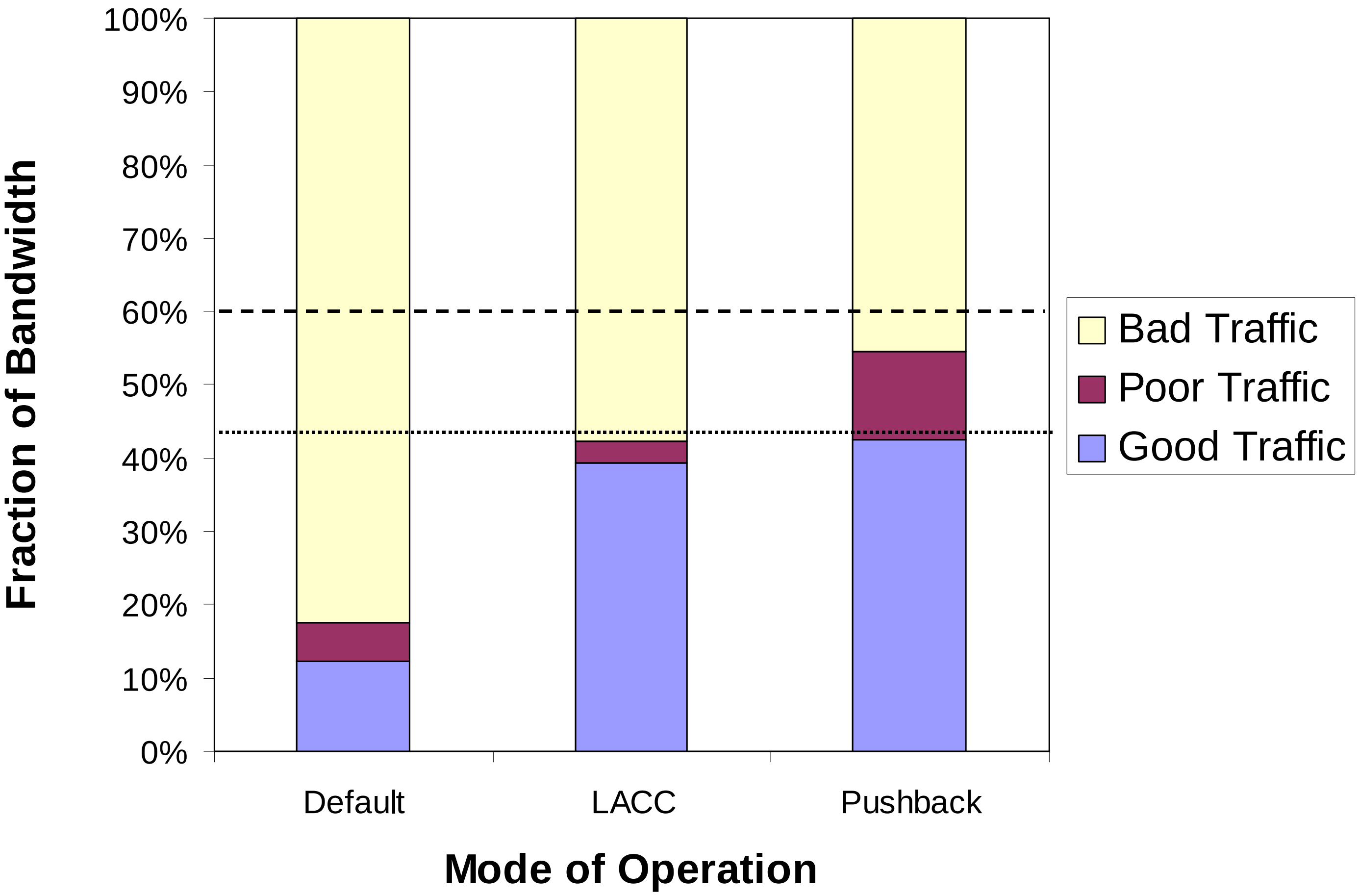
Simulation Topology



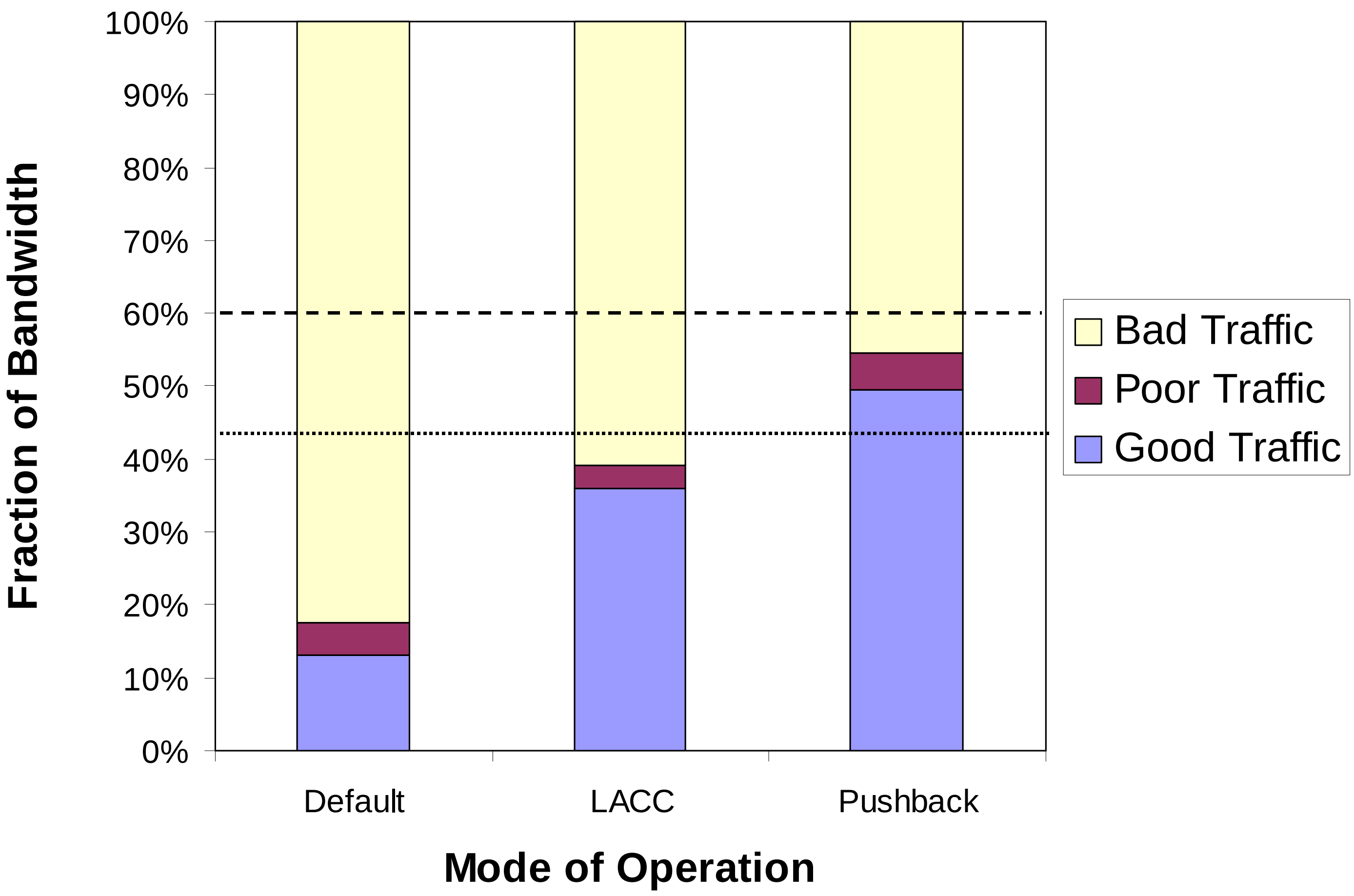
Terminology

- *Bad*: attack traffic
- *Good*: unrelated traffic
- *Poor*: innocent traffic within the aggregate

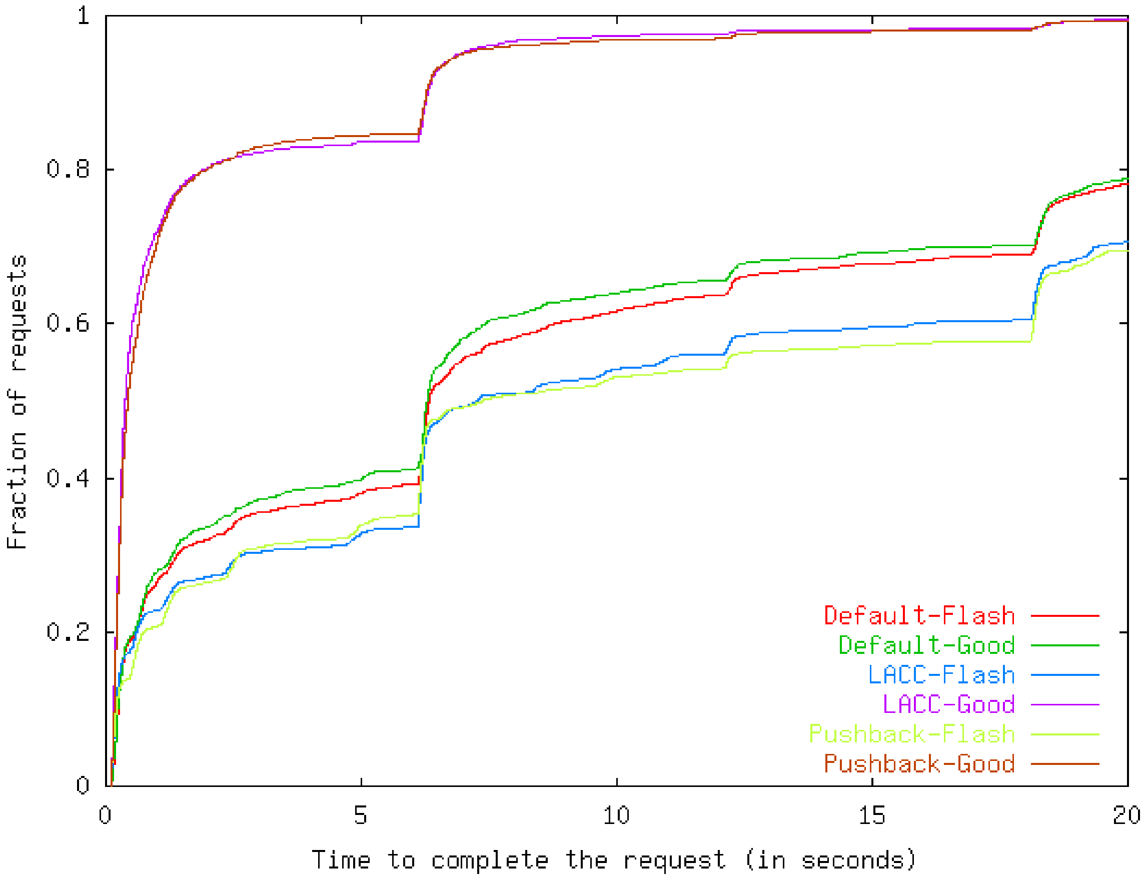
Sparse DoS Attack (4 hosts)



Diffuse DoS Attack (32 hosts)



Flash Crowd



Fraction of requests
Time to complete the request (in seconds)
0
0.2
0.4
0.6
0.8
1
0
5
10
15
20
Default-Flash
Default-Good
LACC-Flash
LACC-Good
Pushback-Flash
Pushback-Good

Issues

- Implementation
  - only rate-limiter is in forwarding path
- What is upstream for routers connected by a LAN?
  - *dummy* pushback messages
- Incremental deployment
- Policy
  - intra and inter AS
- Time constants involved in pushback

Open Questions

- What do attack topologies look like?
- Packet header traces from attacks
- Data on sustained periods of congestion
  - how long and what causes them?
- What else can pushback be used for?
  - dynamic peering pipewidth, multipath routing, traffic matrix