Measurement

Lecture 8, Computer Networks (198:552)

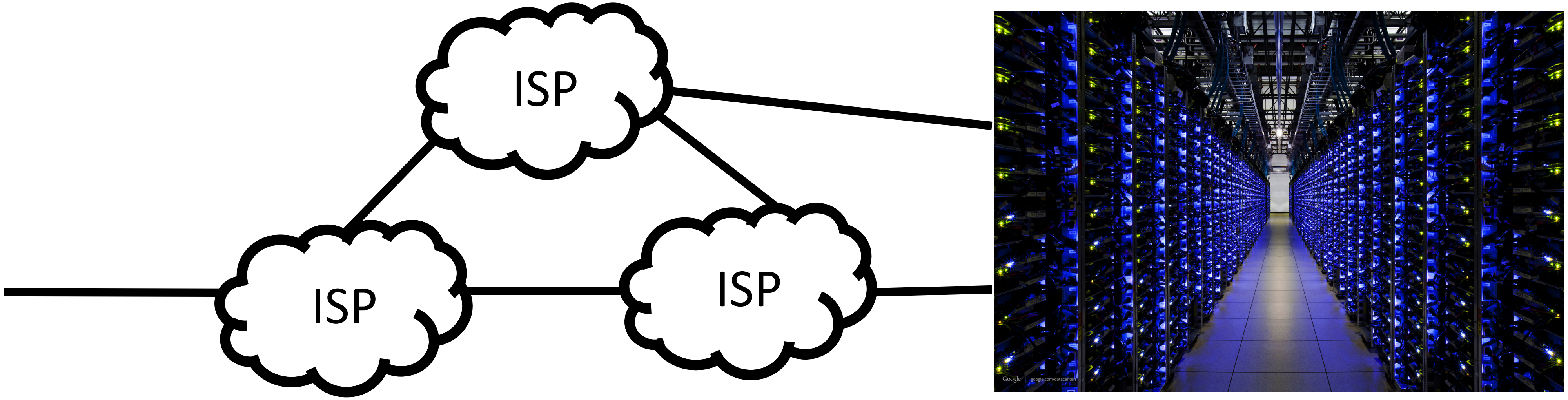
RUTGERS
UNIVERSITY | NEW BRUNSWICK

Why measure networks?

Application QoS
Throughput
Delay
Loss
Jitter

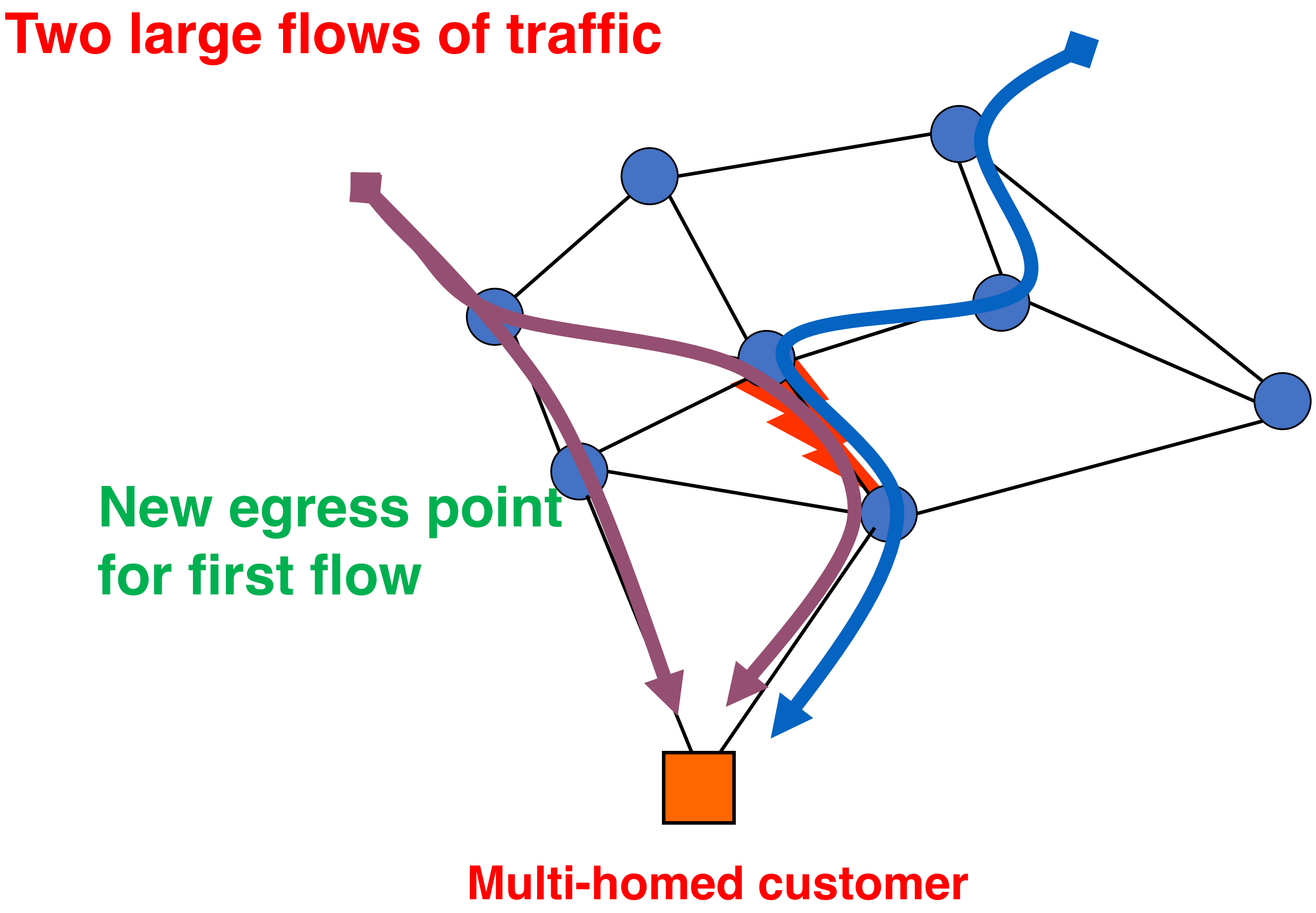
Availability
Congestion/overload
Long-term demands
SLO violations

Application QoS
Problematic ISPs
Problematic CDNs

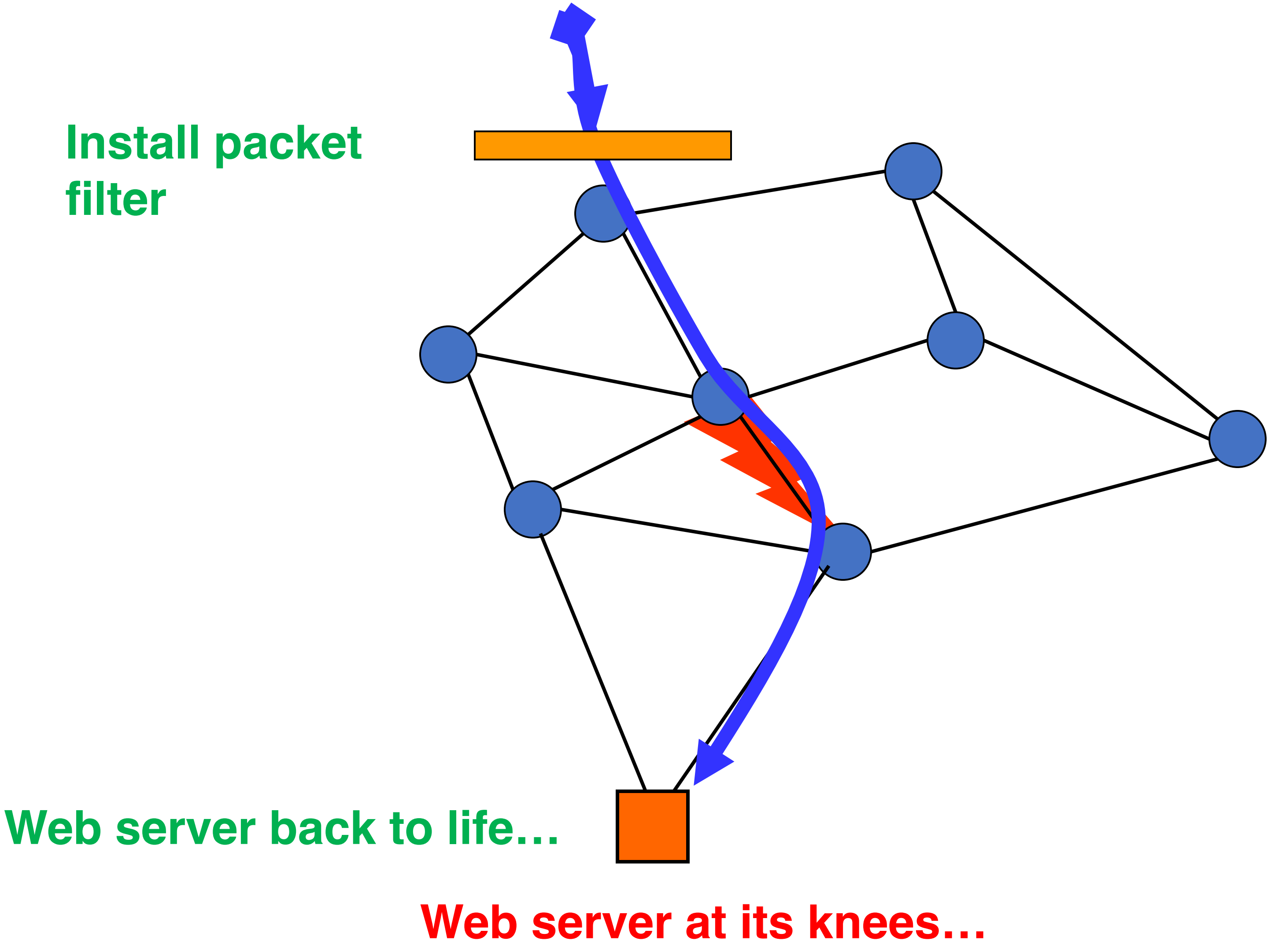


Measurements for
ISP Network Operators

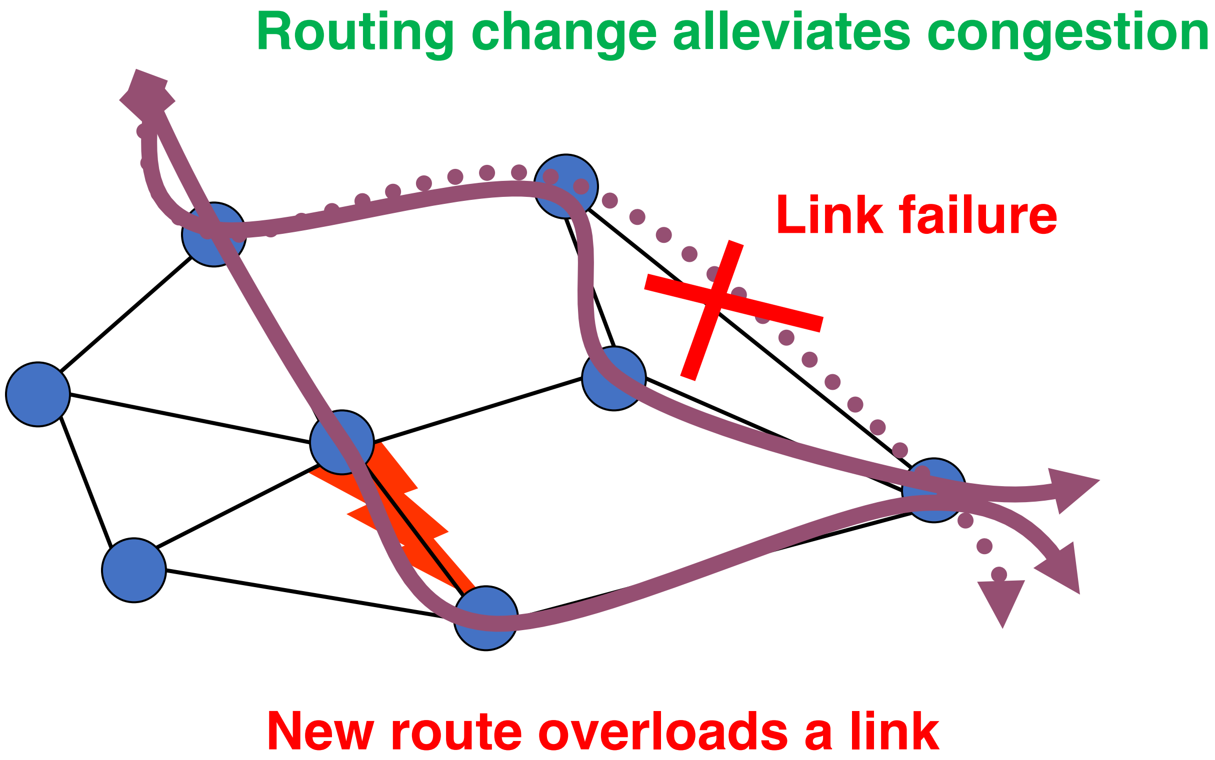
Example (1): Excess Traffic



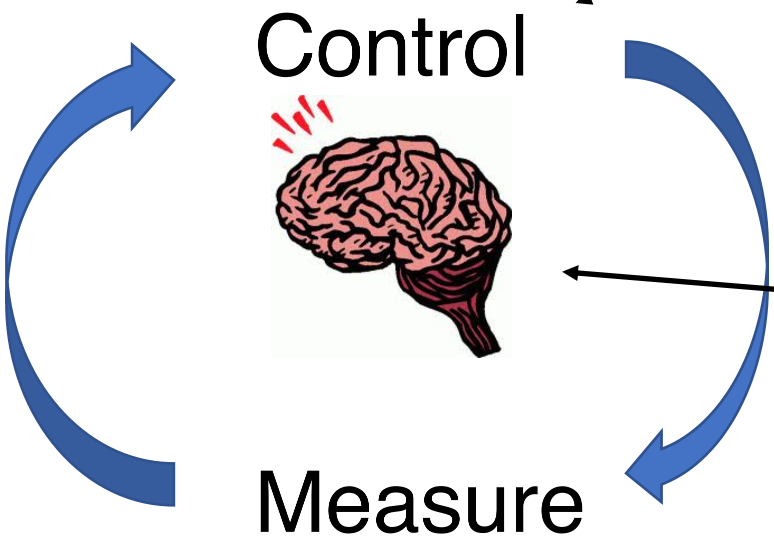
Example (2): DoS Attack



Example (3): Link Failure



Measurements for ISP network operators



- Route and schedule traffic
- Filter traffic
- Provision additional capacity

- Diagnose root cause
- Determine how to route traffic

- Detect link or path-level problems
- Measure incoming traffic demands
- “Measure” forwarding updates!

How do ISPs measure today?

- Periodic link statistics
  - SNMP counters
  - Example: port1: 500 packets transmitted, 13 dropped
- Periodic flow statistics
  - NetFlow, sFlow, IPFIX
  - Example: src: 10.0.0.1, dst:8.8.8.8, inport: 4, count: 45
- Active end-to-end probes
  - Ping: 64 bytes from 128.6.68.140: icmp_seq=0 ttl=55 time=6.575 ms
  - Traceroute: more to come
- User complaints!
  - Customer phone calls, NANOG posts

Diagnosis & Traffic engineering

- Control plane issues
  - New routes
  - Link failures
  - Network upgrades!
- Data plane issues
  - DoS attack
  - Flash crowds
  - Poor demand prediction, in general
- “Decision plane” issues
  - Poor provisioning
  - Lack of peering

Lot of neat algorithms
& measurement systems

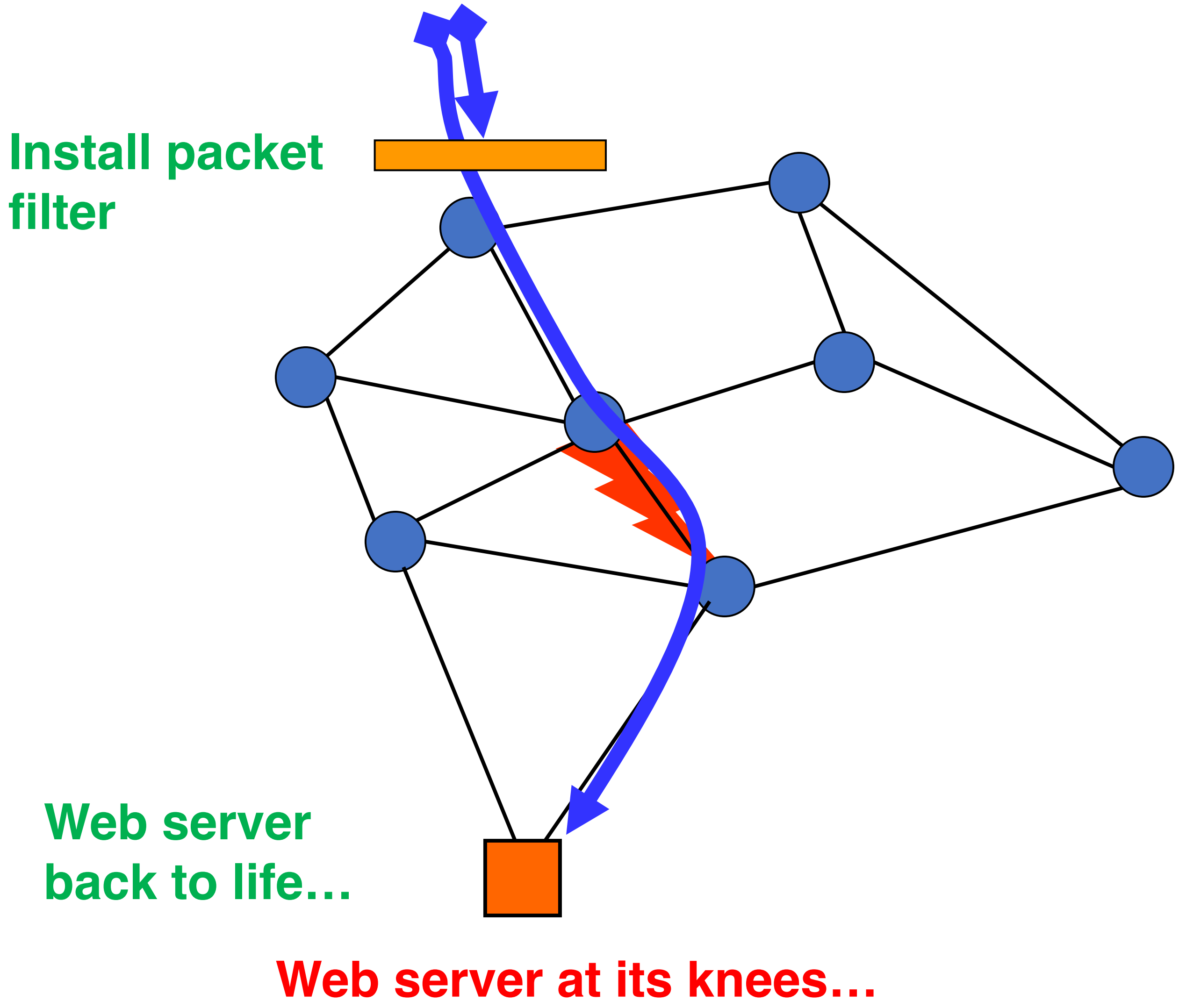
Quality of input data matters!
Scope to do a lot more…

Challenge: Measurement data reduction

- A network can't capture every packet with timestamps
    - Too much data!
- *Filter* to restrict to data of interest
    - Ex: by source, by app, by (physical) port, …
- *Sample* to thin the data stream for exact computations
    - Systematic, random, stratified
    - "Consistently" sample same/distinct packet at each hop
- *Aggregate* (ex: by flow) to summarize data over many packets
    - One problem: too many flows
    - *Sketches:* aggregation that approximates with limited memory

Challenge: Joining traffic with forwarding

- Where is DoS traffic entering the network?
- How do I know which traffic is DoS traffic?
- Are there other links that are affected?
- Should you reroute other traffic that is affected?



End-to-End Measurements

Why end-to-end measurements?

- Endpoints could directly measure what matters to users

- ISPs may not be willing to share data
  - Proprietary design, net neutrality, …
  - Data shared improperly may violate user privacy!

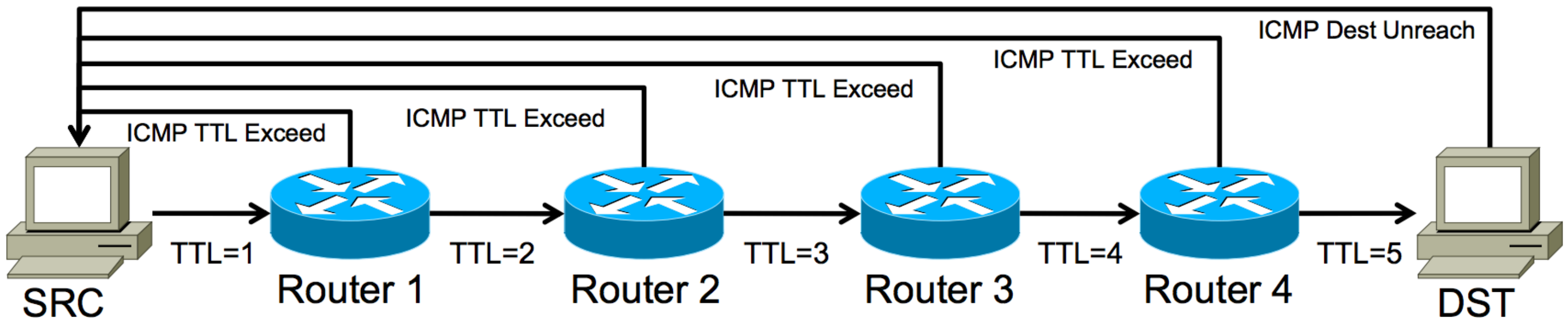
- Indirect view: can’t say for sure why something happens
  - Hard to corroborate with ground truth
  - Possible to use multiple endpoints and span ISP boundaries!

Metrics and tools

- Reachability: ping & its variants
- Path: traceroute & its variants
- Available bandwidth: speedtest, iperf, pathrate, …
- Delays and loss rate: a selection of the above tools

Traceroute

1. Launch a probe packet towards DST, with a TTL of 1
2. Each router hop decrements the TTL of the packet by 1
3. When TTL hits 0, router returns ICMP TTL Exceeded
4. SRC host receives this ICMP, displays a traceroute “hop”
5. Repeat from step 1, with TTL incremented by 1, until…
6. DST host receives probe returns ICMP Dest Unreach



Traceroute: Example output (1/2)

```
[552]$  traceroute google.com
traceroute to google.com (172.217.10.78), 64 hops max, 52 byte packets
 1  fios_quantum_gateway (192.168.1.1)  1.628 ms  1.537 ms  1.506 ms
 2  lo0-100.nwrknj-vfttp-354.verizon-gni.net (74.102.79.1)  2.093 ms  2.486 ms  1.835 ms
 3  b3354.nwrknj-lcr-21.verizon-gni.net (100.41.137.110)  4.962 ms  2.935 ms  3.985 ms
 4  * * *
 5  0.et-10-1-5.gw7.ewr6.alter.net (140.222.2.233)  3.864 ms
    0.et-11-1-0.gw7.ewr6.alter.net (140.222.239.27)  3.503 ms
    0.et-10-1-5.gw7.ewr6.alter.net (140.222.2.233)  3.581 ms
 6  209.85.149.208 (209.85.149.208)  3.949 ms  4.222 ms  4.669 ms
 7  * * *
 8  108.170.226.198 (108.170.226.198)  9.154 ms
    108.170.237.214 (108.170.237.214)  7.080 ms
    72.14.234.64 (72.14.234.64)  10.782 ms
 9  lga34s14-in-f14.1e100.net (172.217.10.78)  4.097 ms
    108.170.248.66 (108.170.248.66)  5.462 ms
    108.170.248.20 (108.170.248.20)  9.410 ms
```

Traceroute: Example output (2/2)

```
[552]$  traceroute rutgers.edu
traceroute to rutgers.edu (128.6.68.140), 64 hops max, 52 byte packets
 1  fios_quantum_gateway (192.168.1.1)  1.536 ms  1.083 ms  1.098 ms
 2  lo0-100.nwrknj-vfttp-354.verizon-gni.net (74.102.79.1)  2.343 ms  1.932 ms  1.948 ms
 3  b3354.nwrknj-lcr-21.verizon-gni.net (100.41.137.110)  3.124 ms
    b3354.nwrknj-lcr-22.verizon-gni.net (100.41.137.112)  4.026 ms  2.766 ms
 4  * * *
 5  * * *
 6  0.ae1.gw1.phil.alter.net (140.222.0.221)  6.599 ms
    0.ae6.gw1.phil.alter.net (140.222.0.223)  5.401 ms  5.670 ms
 7  rutgers-gw.customer.alter.net (63.65.75.238)  5.061 ms  6.937 ms  6.205 ms
 8  172.29.8.17 (172.29.8.17)  5.321 ms  5.475 ms  10.577 ms
 9  172.29.6.63 (172.29.6.63)  6.500 ms  7.154 ms  7.254 ms
10  172.29.6.45 (172.29.6.45)  6.808 ms  6.799 ms  6.612 ms
11  172.28.193.138 (172.28.193.138)  8.201 ms  7.956 ms  8.180 ms
...
64  * * *
```

Some problems with traceroute

- Control traffic (ICMP) and data traffic may see different behavior
  - Router CPU versus forwarding table
  - Probes load-balanced differently
- A different packet observes each hop
  - Route changes while packet “in transit”
- Not all routers may respond to ICMP messages
  - Hidden routers
  - Anonymous routers
  - Improper processing
- One-way measurement

End-to-End Routing Behavior in the Internet

LBNL Technical Report (1996)

Vern Paxson

Methodology

- Traceroute between NPDs distributed worldwide (add pic)
- Exponential sampling/PASTA property
  - Why?
  - What might happen otherwise?
- D1: unidirectional traceroutes
- D2: “paired” traceroutes
- Confidence intervals for probability that an event occurred
- Measurements sample half of the Internet by AS weight

Pathologies in Internet routing

- Forwarding loops!
  - Persistent and temporary
- Circuitous routing
- Routing transients
  - Recovery times are bimodal
- Route fluttering
- Partitioned network
- Temporary outages, some $>$ 30 seconds
- Too many hops
- Pathologies correlated with operator change and congestion

Summary of pathologies

Pathology	Probability	Trend	Notes
Persistent loops	0.13–0.16%		Some lasted hours.
Temporary loops	0.055–0.078%		
Erroneous routing	0.004–0.004%		No instances in $\mathcal{D}_2$.
Mid-stream change	0.16% // 0.44%	worse	Suggests rapidly varying routes.
Infrastructure failure	0.21% // 0.48%	worse	No dominant link.
Outage $\geq$ 30 secs	0.96% // 2.2%	worse	Duration exponent. distributed.
Total pathologies	1.5% // 3.4%	worse	

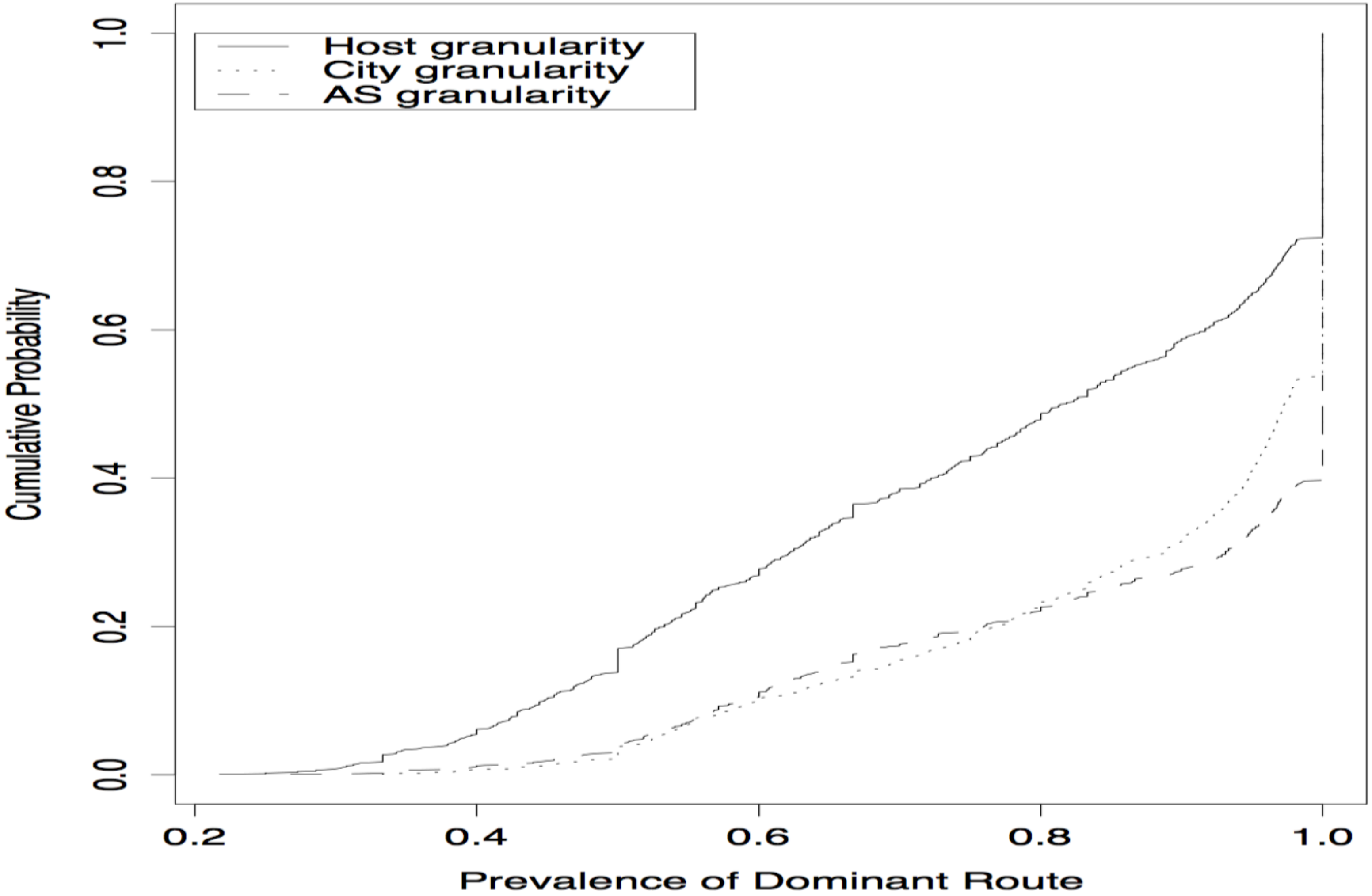
Routing stability

- Why does routing stability matter?

- Prevalence: how frequently do you see a route?
  - PASTA ensures that samples see “true” stable behavior

- Persistence: how long does a given route persist over time?
  - Challenging to measure!
  - Example: R1, R2, R1, but samples miss the intermediate R2

Routing prevalence



Routing persistence

Time scale	%	Notes
seconds	N/A	“Flutter” for purposes of load balancing. Treated separately, as a pathology, and not included in the analysis of persistence.
minutes	N/A	“Tightly-coupled routers.” We identified five instances, which we merged into single routers for the remainder of the analysis.
10's of minutes	9%	Frequent route changes inside the network. In some cases involved routing through different cities or AS's.
hours	4%	Usually intra-network changes.
6+ hours	19%	Also intra-network changes.
days	68%	Bimodal. 50% of routes persist for under 7 days. The remaining 50% account for 90% of the total route lifetimes.

Routing asymmetry

- 49% of D2 measurements saw asymmetric paths!
    - visiting a different city each way around
    - 30% with a different AS!

- Trend worsening over time

A summary

- No guarantees on where your traffic might end up
  - A black-hole!
  - Somewhere unintended (US east→London goes through Israel)
- Routes are dominated by single winner but can be quite flappy
  - Implications on what performance apps might expect
  - What measurement tools provide
- Asymmetry makes a lot of things complex
  - Diagnosis: Assumptions about where problems lie
  - Flow state in the core: can’t assume you’ll see return traffic

Limitations of the study

- Representativeness:
  - Routes within an AS may not have similar characteristics!
  - Sample a really small subset of actual Internet paths

- Methodology:
  - PASTA doesn’t hold when the network is down
  - Hard to extrapolate trends in Internet evolution with just 2 points

- E2E measurements:
  - Fundamentally hard to corroborate with ground truth

Reverse Traceroute

Usenix NSDI 2010

Ethan Katz-Bassett, Harsha V. Madhyastha, Vijay Kumar Adhikari, Colin Scott, Justine Sherry, Peter van Wesep, Thomas Anderson, and Arvind Krishnamurthy

Can we find the reverse path?

- Routes aren’t always symmetric!
- What are reverse routes useful for?

Main techniques

- Distributed set of vantage points issuing forward traceroutes
  - Create an “atlas” of nodes and paths to the source
- Incrementally stitch reverse path until you hit an atlas node
- IP record route: grab first (few) router IP address(es) on return path
  - Recursively reverse traceroute from there!
- Timestamp option: verify whether a router is on reverse path
- Source spoofing: sample reverse path without forward path
  - Use prior mapping of vantage points “closest” to the destination
- When all else fails, assume symmetric routing

How accurate is reverse traceroute?

- Ground truth: actual traceroutes from D to S
- Overlap in hops of reverse and (ground truth) traceroute
  - Close to 87% in the median
- Why are there differences between the two?
- Reverse paths used undiscovered peering links

(E2E) Measurement research challenges

- Ground truth
  - Explaining empirical observations
  - Aliasing, router identification, AS identification, …
- Representativeness
- Measuring without bias
  - PASTA
- Coordinating distributed vantage points
- Probing overheads
- Detailed knowhow of the Internet and its quirks!
  - Ex: IP timestamp marked only when router sees itself on top
- How will the conclusions evolve over time?