On the Characteristics and Origins of Internet Flow Rates

Yin Zhang
Lee Breslau
AT&T Labs – Research
{yzhang,breslau}@research.att.com

Vern Paxson
Scott Shenker
ICIR
{vern,shenker}@icir.org

ACM SIGCOMM 2002

Motivation

- Limited knowledge about flow rates
  - Flow rates are impacted by many factors
    - Congestion, bandwidth, applications, host limits, …
  - Little is known about the resulting rates or their causes
- Why is it important to understand flow rates?
  - Understanding the network
    - User experience
  - Improving the network
    - Identify and eliminate bottlenecks
  - Designing scalable network control algorithms
    - Scalability depends on the distribution of flow rates
  - Deriving better models of Internet traffic
    - Useful for workload generation and various network problems

Two Questions

- What are the characteristics of flow rates?
  - Rate distribution
  - Correlations
- What are the causes of flow rates?
  - T-RAT: TCP Rate Analysis Tool
    - Design
    - Validation
    - Results

Characteristics of Internet Flow Rates

Datasets and Methodology

- Datasets
  - Packet traces at ISP backbones and campus access links
    - 8 datasets; each lasts 0.5 – 24 hours; over 110 million packets
  - Summary flow statistics collected at 19 backbone routers
    - 76 datasets; each lasts 24 hours; over 20 billion packets
- Flow definition
  - Flow ID: <SrcIP, DstIP, SrcPort, DstPort, Protocol>
  - Timeout: 60 seconds
- Rate = Size / Duration
  - Exclude flows with duration < 100 msec
- Look at:
  - Rate distribution
  - Correlations among rate, size, and duration

Flow Rate Characteristics

- Rate distribution
  - Most flows are *slow*, but most bytes are in *fast* flows
  - Distribution is skewed
    - Not as skewed as size distribution
    - Consistent with log-normal distribution [BSSK97]
- Correlations
  - Rate and size are strongly correlated
  - Not due to TCP slow-start
    - Removed initial 1 second of each connection; correlations increase
  - What users download is a function of their bandwidth

Causes of
Internet Flow Rates

T-RAT: TCP Rate Analysis Tool

- Goal
  - Analyze TCP packet traces and determine rate-limiting factors for different connections
- Requirements
  - Work for traces recorded anywhere along a network path
    - Traces don't have to be recorded near an endpoint
  - Work just seeing one direction of a connection
    - Data only or ACK only → there is no easy cause & effect
  - Work with partial connections
    - Prevent bias against long-lived flows
  - Work in a streaming fashion
    - Avoid having to read the entire trace into memory

TCP Rate Limiting Factors

Factor	***Description***
Application	Application doesn't generate data fast enough
Opportunity	Flow never leaves slow-start
Receiver	Flow is limited by receiver advertised window
Sender	Flow is limited by sending buffer
Bandwidth	Flow fully utilizes and is limited by bottleneck link bandwidth
Congestion	Flow responds to packet loss
Transport	Flow has left slow-start & no packet loss & not limited by receiver/sender/bandwidth

T-RAT Components

- MSS Estimator
  - Identify *Maximum Segment Size (MSS)*
- RTT Estimator
  - Estimate RTT
  - Group packets into flights
    - Flight: packets sent during the same RTT
- Rate Limit Analyzer
  - Determine rate-limiting factors based on MSS, RTT, and the evolution of flight size

What Makes It Difficult?

- The network may introduce a lot of noise
  - E.g. significant delay variation, ACK compression, ...
- Time-varying RTT is difficult to track
  - E.g., handshake delay and median RTT may differ substantially
- Delayed ACK significantly complicates TCP dynamics
  - E.g. congestion avoidance: 12, 12, 13, 12, 12, 14, 14, 15, …
- There are a large number of TCP flavors & implementations
  - Different loss recovery algorithms, initial cwnd, bugs, weirdness
- Timers may introduce behavior difficult to analyze
  - E.g. delack timer may expire in the middle of an RTT
- Packets missing due to packet filter drop, route change
  - They are not lost!
- There may be multiple causes for a connection
- Some behaviors are intrinsically ambiguous
- *And a lot more …*

MSS Estimator

- Data stream
  - MSS ← largest data packet payload
- ACK stream
  - MSS ← *"most frequent common divisor"*
    - Like GCD, apply heuristics to
      - avoid looking for divisors of numbers that are not multiples of MSS
      - favor popular MSS (e.g. 536, 1460, 512)

RTT Estimator

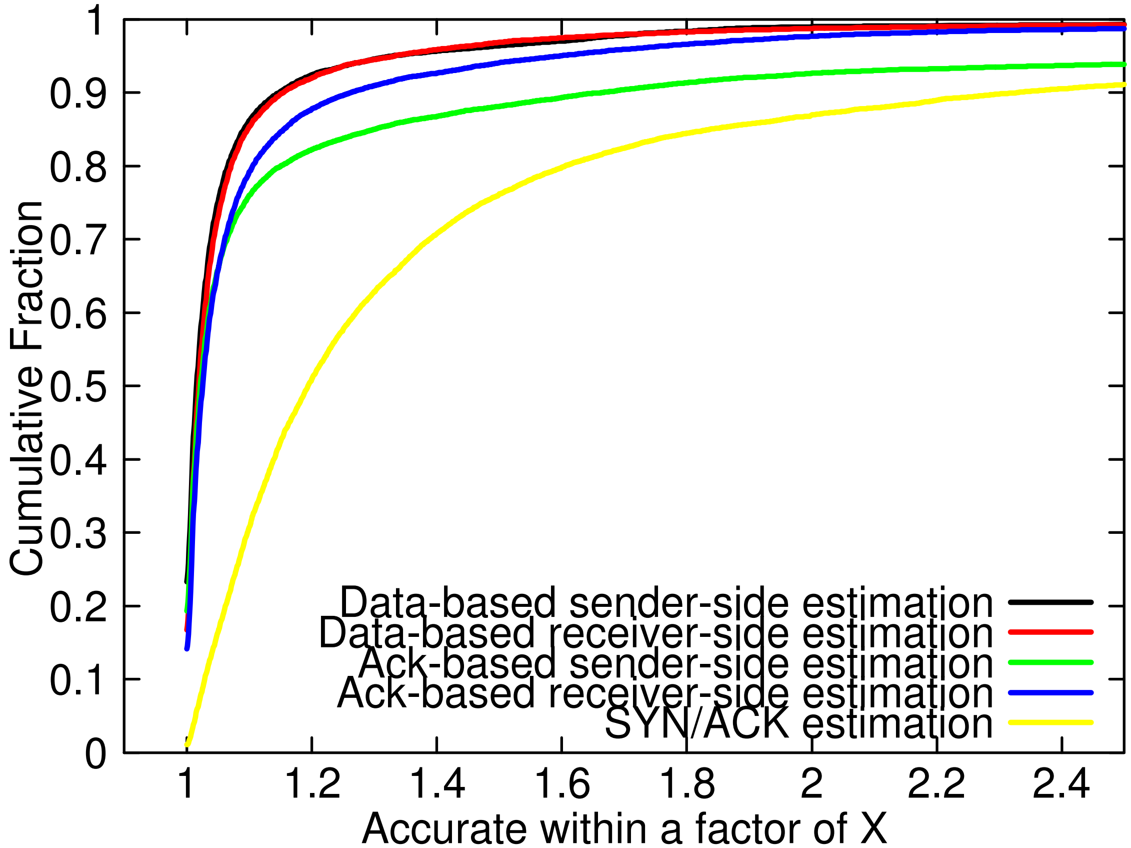
- Generate a set of candidate RTTs
  - Between 3 msec and 3 sec: 0.003×1.3^{K} sec
- Assign a score to each candidate RTT
  - Group packets into flights
    - Flight boundary: packet with large inter-arrival time
  - Track evolution of flight size over time and match it to identifiable TCP behavior
    - Slow start
    - Congestion avoidance
    - Loss recovery
  - Score ← # packets in flights consistent with identifiable TCP behavior
- Pick the top scoring candidate RTT

Rate Limit Analyzer

Cause	***Test***
Application	A sub-MSS packet followed by a lull > RTT, followed by new data
Opportunity	Total bytes < 13*MSS, or it never leaves slow-start
Receiver	3 consecutive flights with size > $awnd_{max}$ – 3*MSS
Sender	Flight size stays constant, and not receiver limited
Bandwidth	A flow keeps achieving same amount of data in flights prior to loss; or it has nearly equally spaced packets
Congestion	Packet loss and not bandwidth-limited
Transport	Sender has left slow-start and there is no packet loss
Host	Flight size stays constant, but we only see data packets and can't tell between sender and receiver
Unknown	None of the above

RTT Validation

Validation against *tcpanaly* [Pax97] over NPD N_2 (17,248 conn)



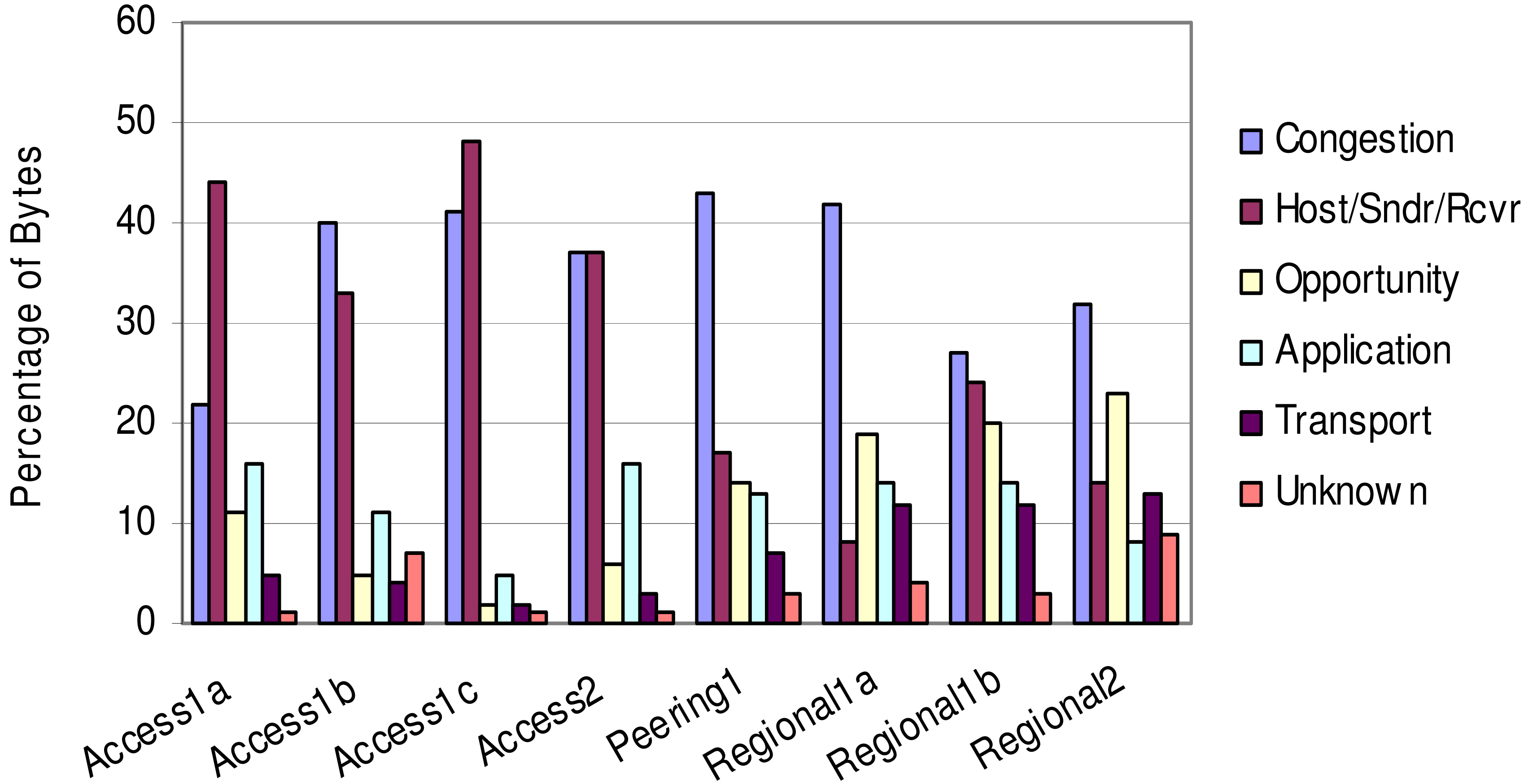
RTT estimator works reasonably well in most cases

Rate Limit Validation

- Methodology
  - ns2 simulations + dummynet experiments
- T-RAT correctly identifies the cause in vast majority of cases
- Failure scenarios

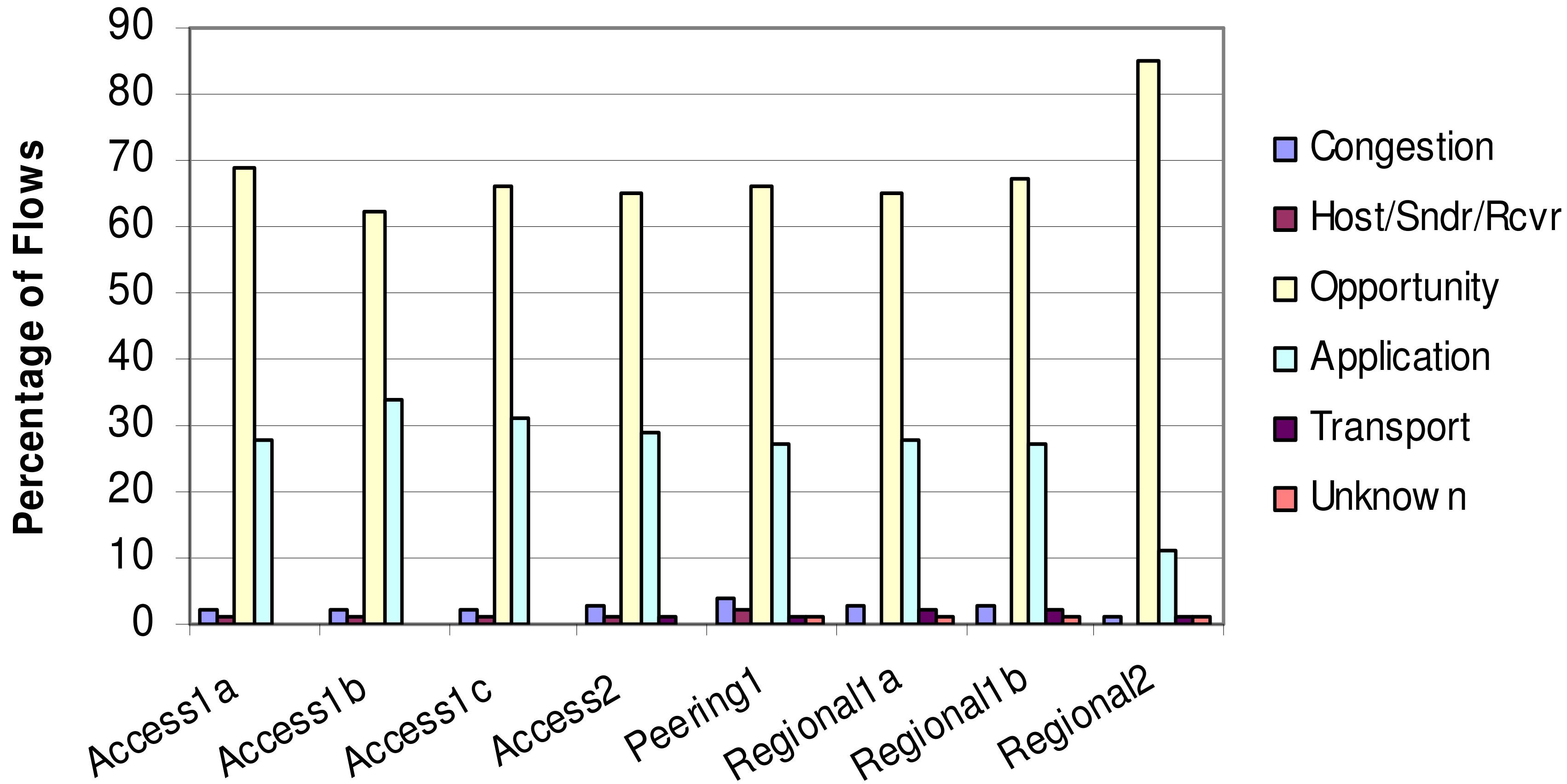
Rcvr/Sndr	Window = 2 packets; or link utilization > 90%
Transport	High background load (esp. on ACK stream)
Congestion	"transport limited" w/o loss → NOT failure!
Bandwidth	RTT wrong, but rate limit correct
Opportunity	Connection size is very large
Application	Less accurate with Nagle's algorithm

Rate Limiting Factors (Bytes)



Dominant causes by bytes: Congestion, Receiver

Rate Limiting Factors (Flows)



Dominant causes by flows: Opportunity, Application

Flow Characteristics by Cause

- Different causes are associated with different performance for users
  - Rate distribution
    - Highest rates: Receiver, Transport
  - Size distribution
    - Largest sizes: Receiver
  - Duration distribution
    - Longest duration: Congestion

Conclusion

- Characteristics of Internet flow rates
  - Fast flows carry most of the bytes
    - It is important to understand their behavior.
  - Strong correlation between flow rate and size
    - What users download is a function of their bandwidth.
- Causes of Internet flow rates
  - Dominant causes:
    - In terms of bytes: congestion, receiver
    - In terms of flows: opportunity, application
  - Different causes are associated with different performance
- T-RAT has applicability beyond the results we have so far
  - E.g. correlating rate limiting factors with other user characteristics like application type, access method, etc.

Thank you!

http://www.research.att.com/projects/T-RAT/