KNOM Tutorial 2003

Internet Traffic Matrix Measurement and Analysis

Sue Bok Moon

Dept. of Computer Science

KAIST

Overview

- Definition of Traffic Matrix
  - Traffic demand, delay, loss
- Applications of Traffic Matrix
  - Engineering, research, SLAs
- Challenges in Obtaining Traffic Matrix
  - Limitation of NetFlow and active probes
  - Challenges in measurement and modeling
- Summay & Future Work

Definition of Traffic Matrix

- What is a traffic matrix?
  - A matrix built on metric of interest
  - Traffic demand matrix
    - How much traffic flows from point A to point B
    - Granularity: PoP, router, link, prefix
  - Delay matrix
    - How much delay from point A to point B
    - Granularity: PoP, router, link, end hosts
  - Loss matrix
    - How many packets are dropped from point A to point B
    - Granularity: PoP, router, end hosts

Example : AT&T Latency Matrix

Current Average : 35 msec

	ATL	CAM	CHI	DAL	DEN	LA
CAM	26					
CHI	37	22				
DAL	14	42	20			
DEN	35	64	40	21		
LA	50	69	47	36	32	
NY	18	06	19	34	54	61

NY

Latency in milliseconds

Traffic Demand Matrix

- Not part of SLAs
  - Hard to obtain
  - Few available publicly

Delay Matrix

- Usually a matrix of average delay of *ping*s between routers of random selection per PoP
  - Average of all PoP-to-PoP delays => SLA
- At end hosts
  - Easy to get using *ping*s between hosts of interest

Loss Matrix

- Usually a matrix of average loss rate of *ping*s between routers of random selection per PoP
  - Average of all PoP-to-PoP loss rates => SLA
- At end hosts
  - Easy to get using *ping*s between hosts of interest

Applications of Traffic Matrix

- Marketing/Sales
  - How much traffic does customer A send from point #1 to point #2?
    - Where should customer A buy more capacity from us?
  - Is most traffic originating in Korea stay within Korea?
    - What is the trend in international traffic growth?
  - What is the performance that customer A sees?
    - Do we have an edge over our competitors?

Applications of Traffic Matrix

- Network Operators
  - Capacity Planning
    - How much traffic do we have from point A to point B?
    - How much capacity should we add?
    - When should we add more capacity?
  - Network Engineering
    - Where is the hot spot? – From SNMP
    - What if a link fails from point A to point B?
    - What if we move traffic from point A to point B?

Applications of Traffic Matrix

- Customers: SLAs
  - What quality of service am I getting?
    - How much delay do I get from ISP A?
    - How much loss do I experience from ISP A?
    - Can I get delay under X ms from ISP A?
    - What is the most popular destination of my traffic?

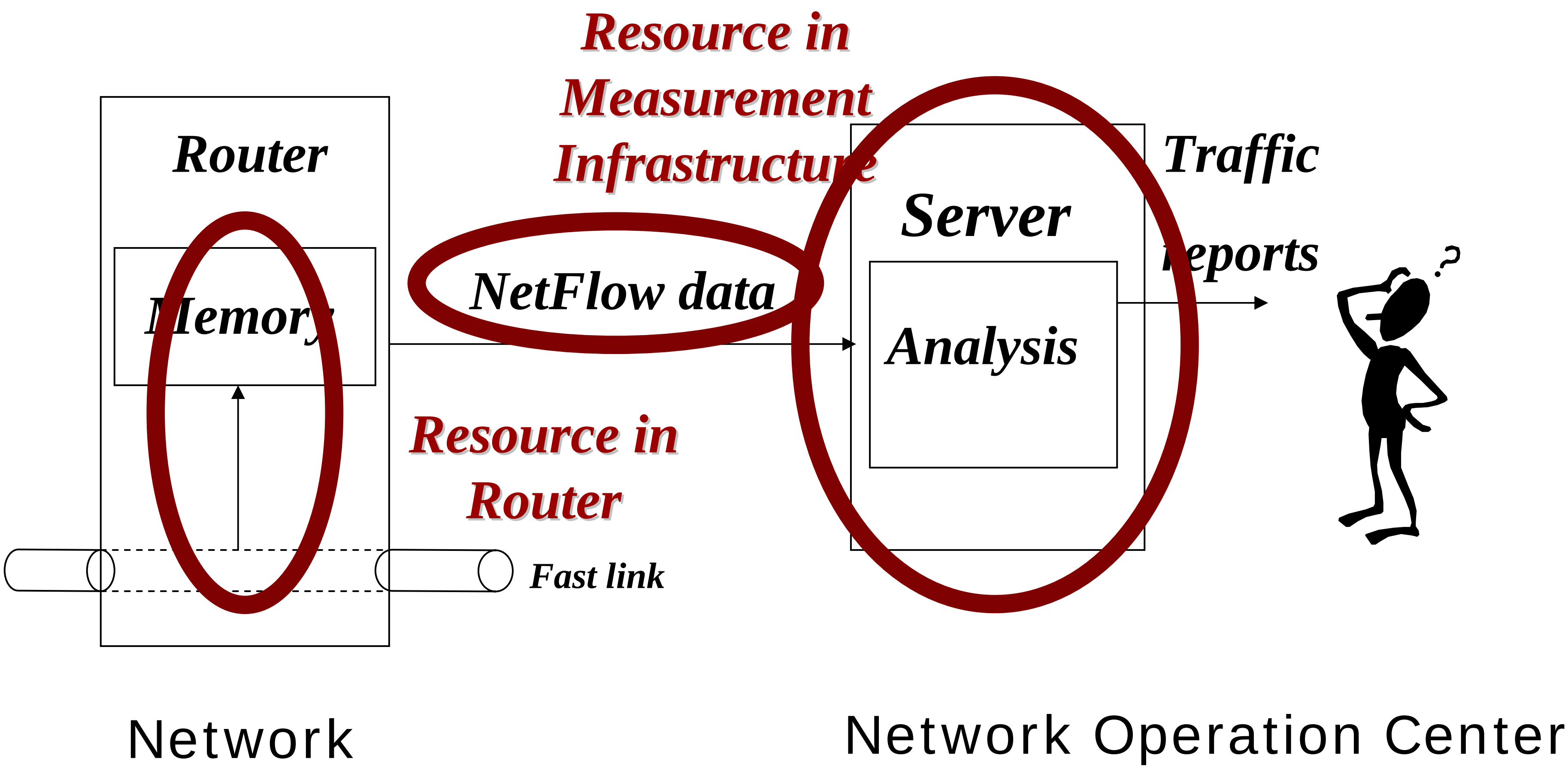
Applications of Traffic Matrix

- Researchers
  - Traffic modeling
    - How does TM evolve over time?
    - What is the fanout factor of traffic?
    - How much more capacity do we expect between point A and point B?
  - Example: IP over WDM
    - Given physical topology of routers and optical nodes, what is the “best” virtual topology?
    - Based on traffic demand matrix

Challenges in Obtaining Traffic Matrix

- Traffic Demand Matrix
  - resource requirements in routers
    - # of concurrently active flows
  - resource requirements in measurement infrastructure
    - production rate of flow statistics
  - traffic characterization
    - packet/byte rate of original traffic
    - rate o f occurrence of original flows
    - average packet/bytes per original flow

Resource Requirements



Most Popular Tools of Choice?

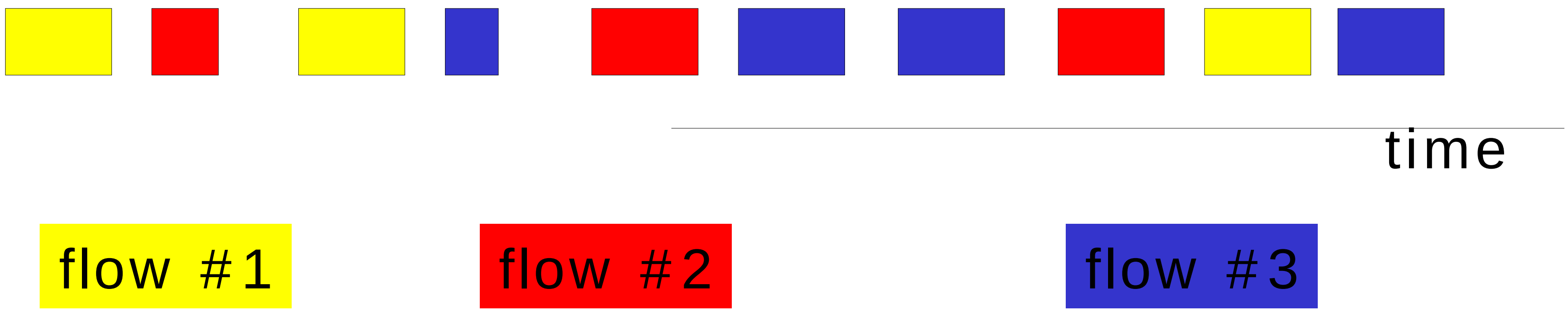
- NetFlow for traffic demand matrix
- ping for delay and loss matrix

NetFlow

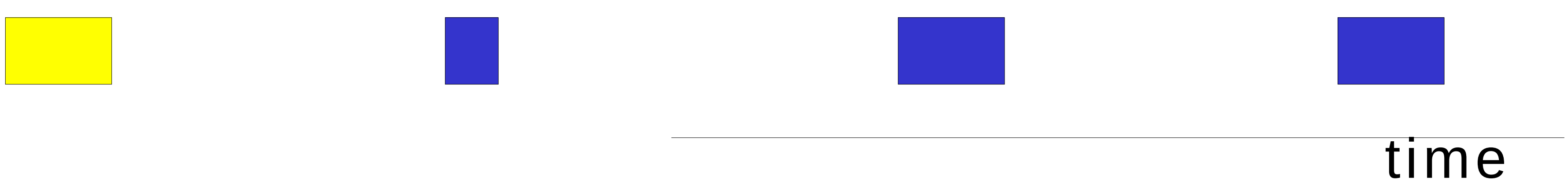
- Cisco’s “proprietary” tool
  - Not an IETF standard
- Basic idea
  - Based on (src ip, src port, dst ip, dst port, proto)
  - Records byte/packet/duration per flow
  - Cannot keep up with high speed links
  - Can sample every N^{th} packet

NetFlow Sampling

Original Packets



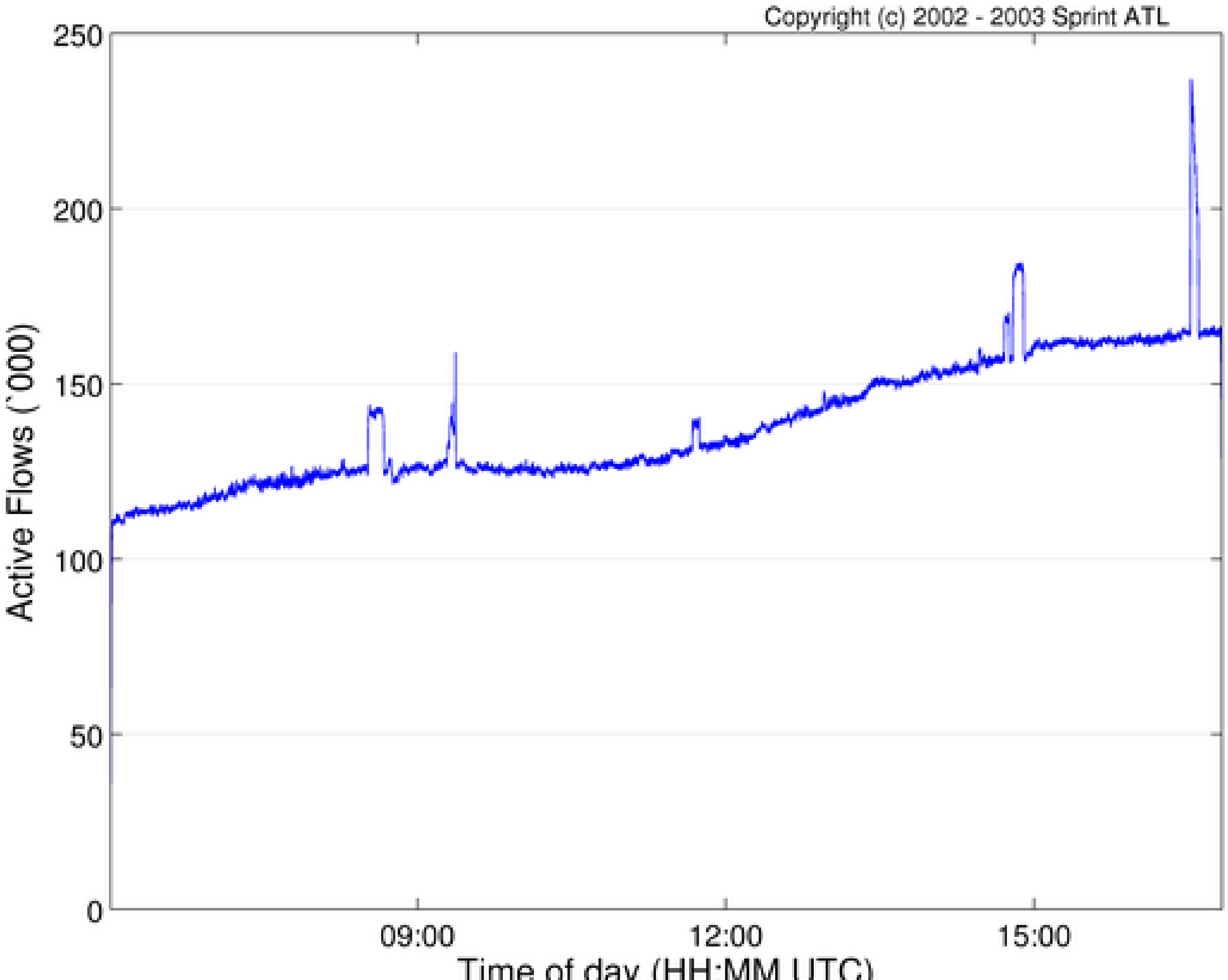
Sampled Packets (every 1/N, N=3)



Limitation of NetFlow

- Scalability
  - Historically NetFlow had a “performance issue”
  - Never deployed at the core
  - Number of flows in case of DDoS attacks beyond capacity
    - Network melt down

Number of Active Flows on a OC-48 Link

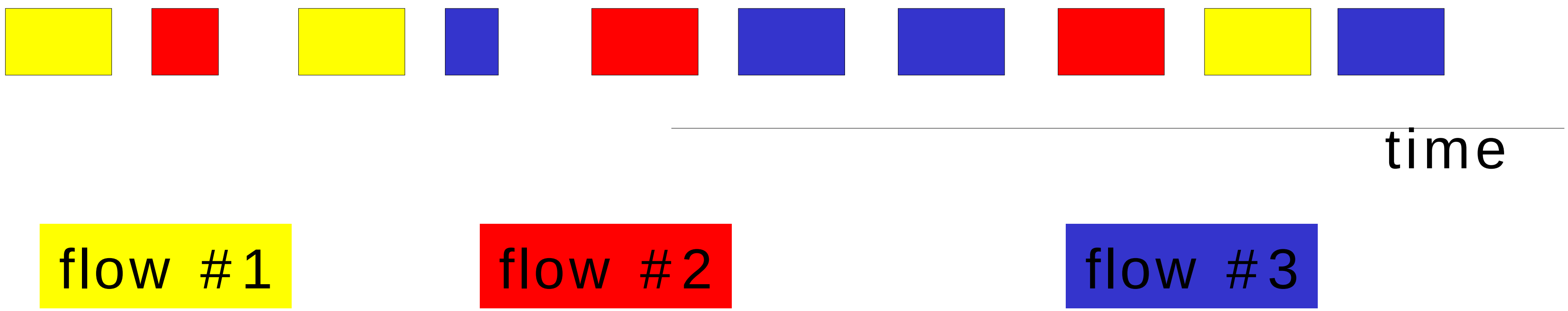


Limitation of NetFlow

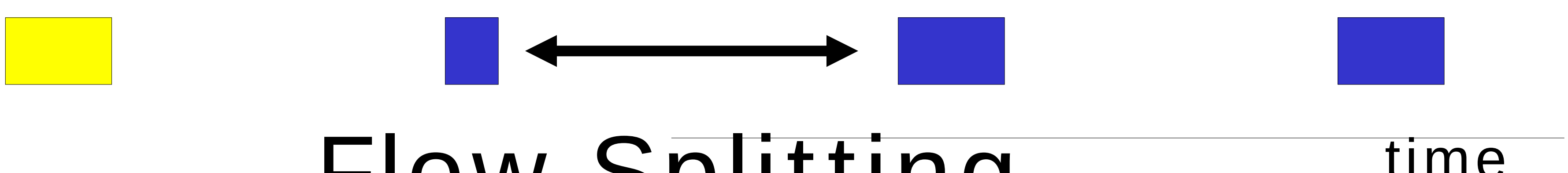
- Representativeness
  - Can we estimate # of total flows from # of sampled flows accurately?
  - Can we estimate # of total WWW flows from # of sampled WWW flows accurately?
  - Metrics of interest:
    - # of flows, flow rate,
  - Packet sampling
    - reduce effective packet rate
    - save cost: slower memory sufficient (DRAM vs SRAM)

NetFlow Sampling

Original Packets



Sampled Packets (every 1/N, N=3)



Flow Splitting

Comparison of sparse and non-sparse applications

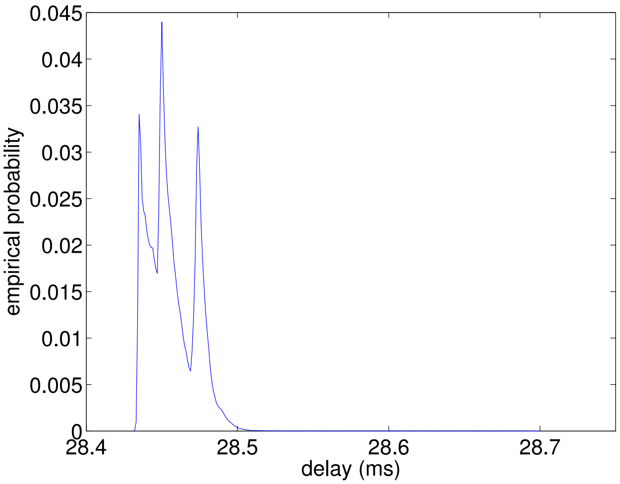
- Flow definition
  - 5-tuple = (src ip, src port, dst ip, dst port, proto)
  - interflow timeout = T
- Increase timeout T
  - potentially less splitting
  - fewer measured flows, more active flows
- Sparse vs. non-sparse flows
  - napster vs. www
  - # of mean active flows change differently over T
  - No simple model of rate and # active flows based on aggregate traffic rates
  - Model sparse and non-sparse flows separately [Duffield03]

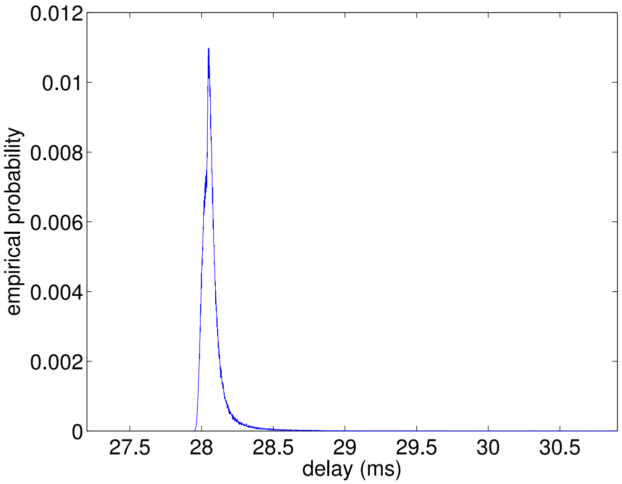
Challenges in Delay Monitoring

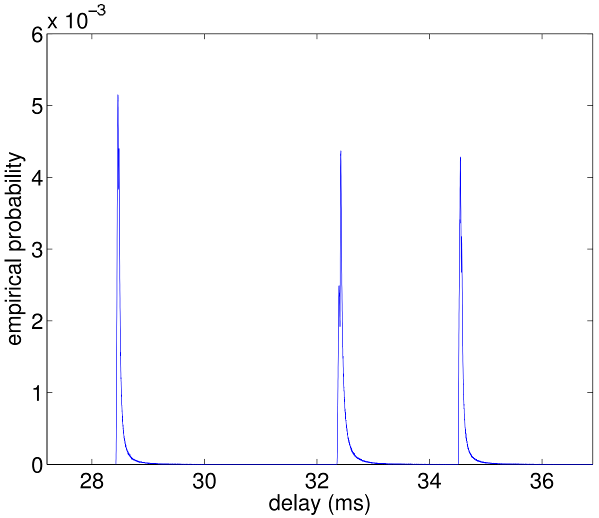
- Not much is known about delay within ISP
  - People think they know delay, but ...
  - Cisco SAA implementation on GSR did not consider clock synchronization, and outputs meaningless numbers
- Too many paths to cover
  - hop-by-hop addition not yet possible

Limitation of Active Probes

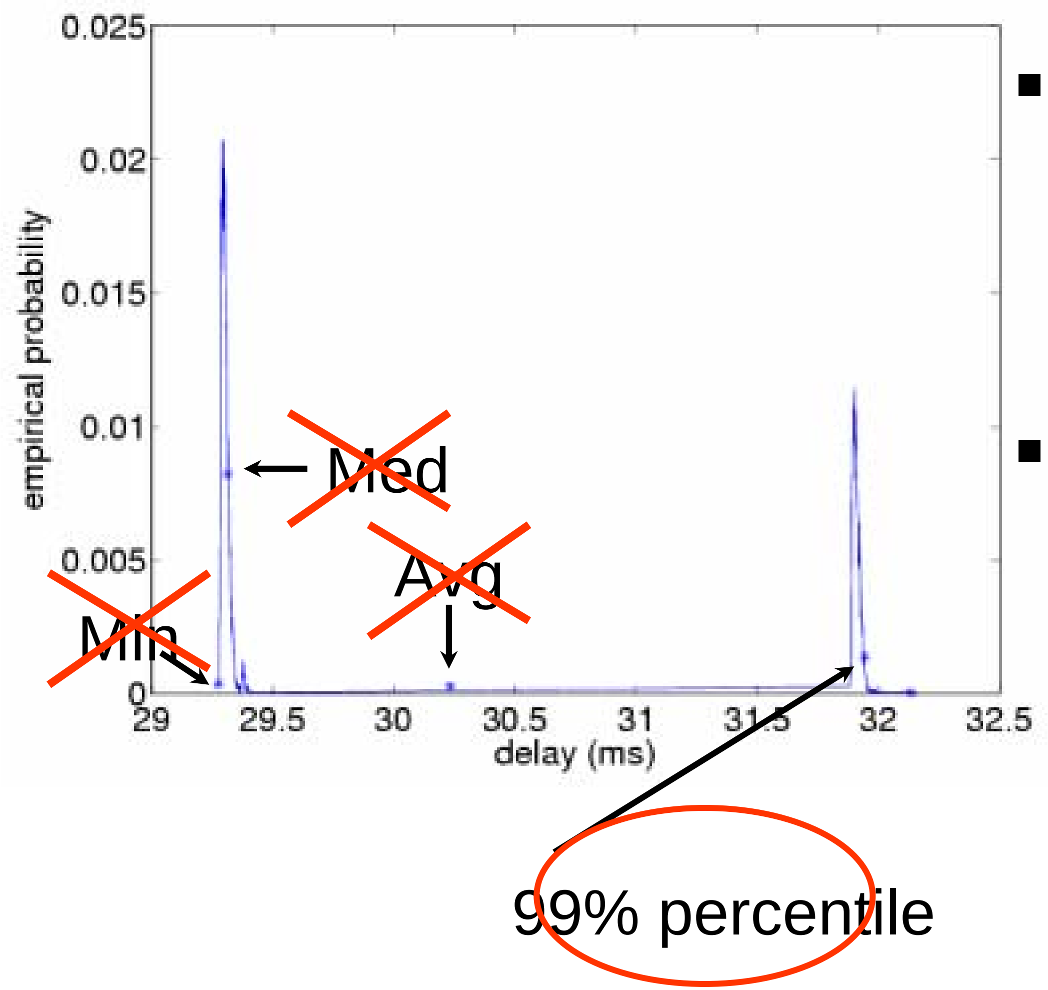
- Representativeness [Choi04]
  - Average? Median?







Suitable Statistic: Percentile!



- ~~Mode~~ detection is hard
  - Difficult to distinguish *small from big*
  - Don’t know *how many ahead* of them
- High-percentile
  - represents upper bound for “most” delay
  - requires a very small number of probes to estimate

Sampling for Demand Matrix

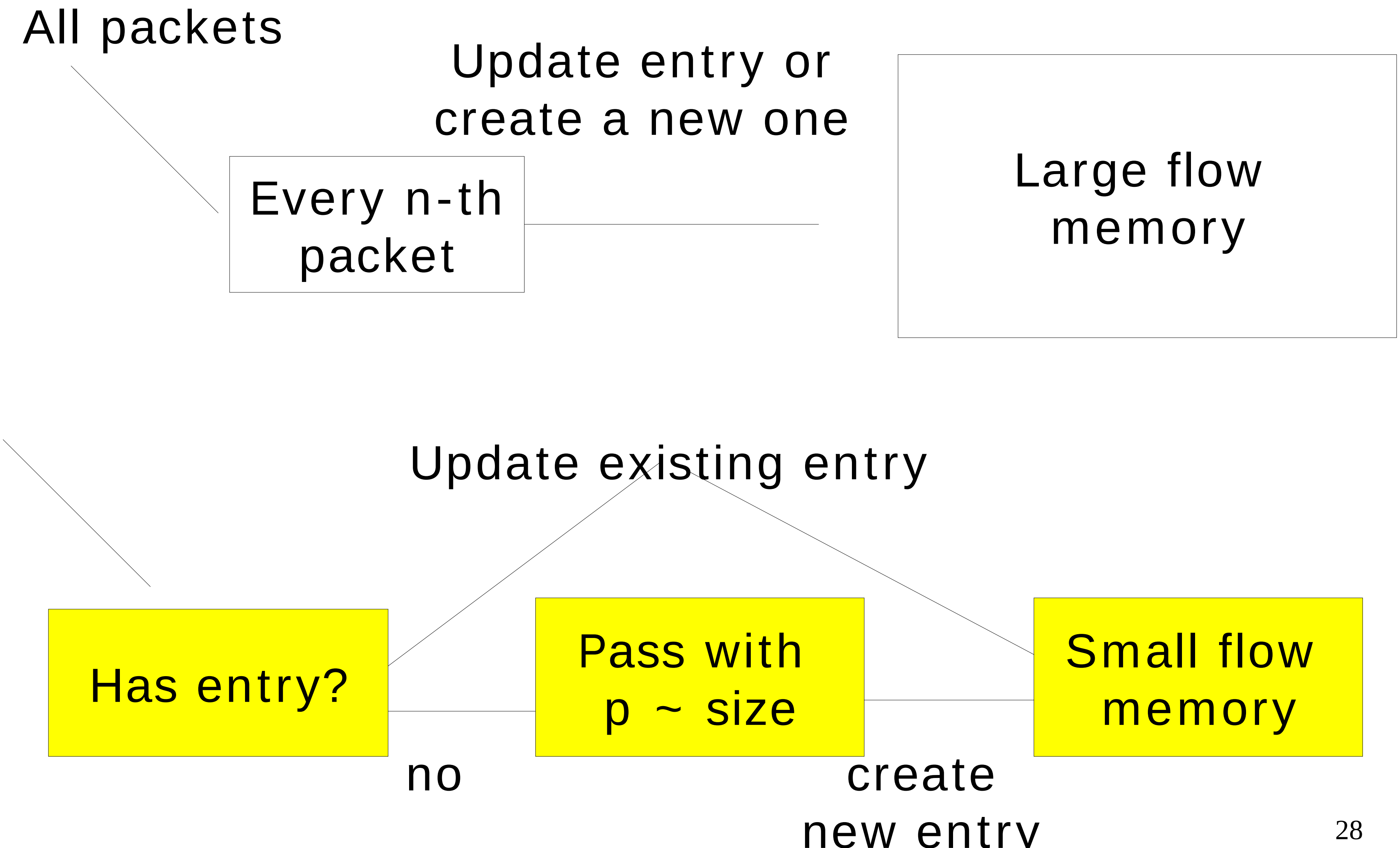
- Periodic sampling does not answer:
  - What are the top 10 flows?
  - What is the most dominant application and who is the heaviest user?
  - What is the total # of packet for every flow?

Hash Function

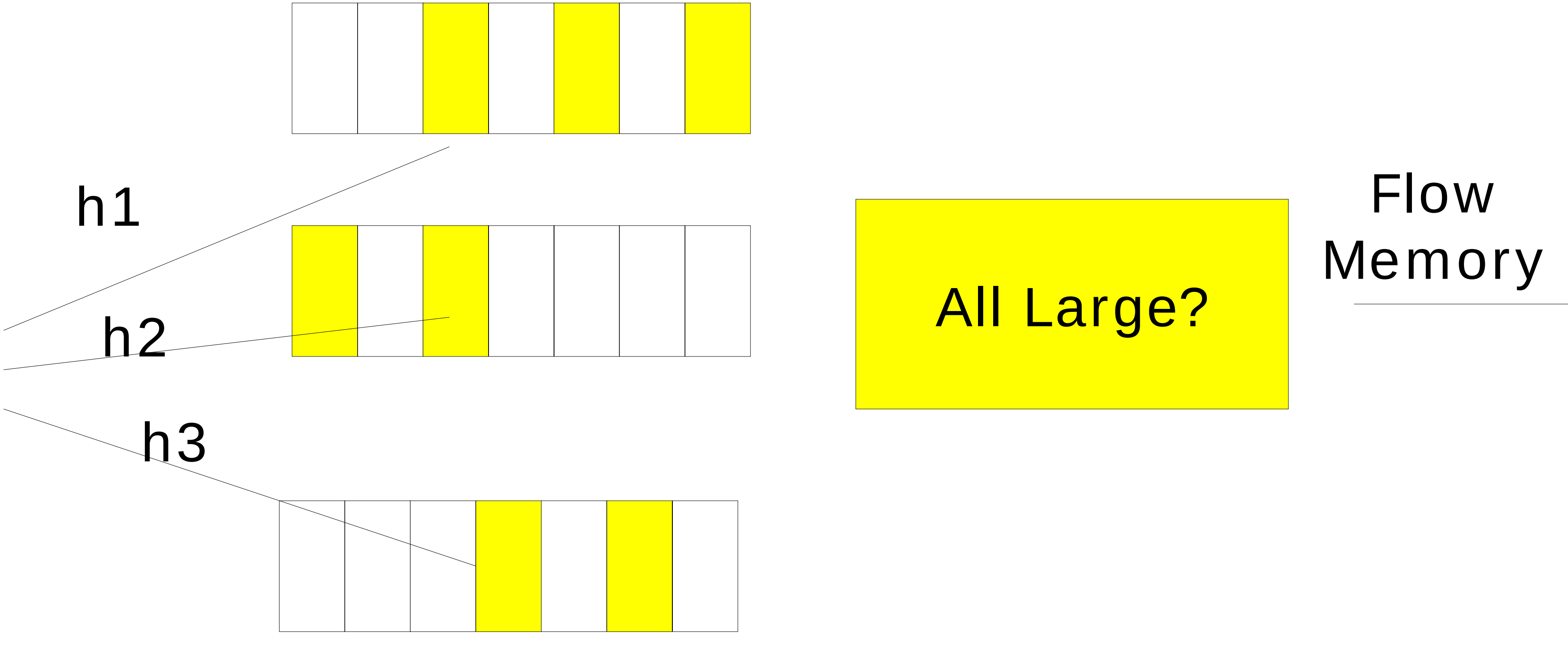
- Mapping from a very large space to a smaller space
  - h: X →Y where |X| >> |Y|
  - IP address to 10-bit hashed key
  - 5-tuple address to 30-bit hashed key
- Load factor = collision probability

What are the top 10 flows?

Sampling for Elephants [Estan02]

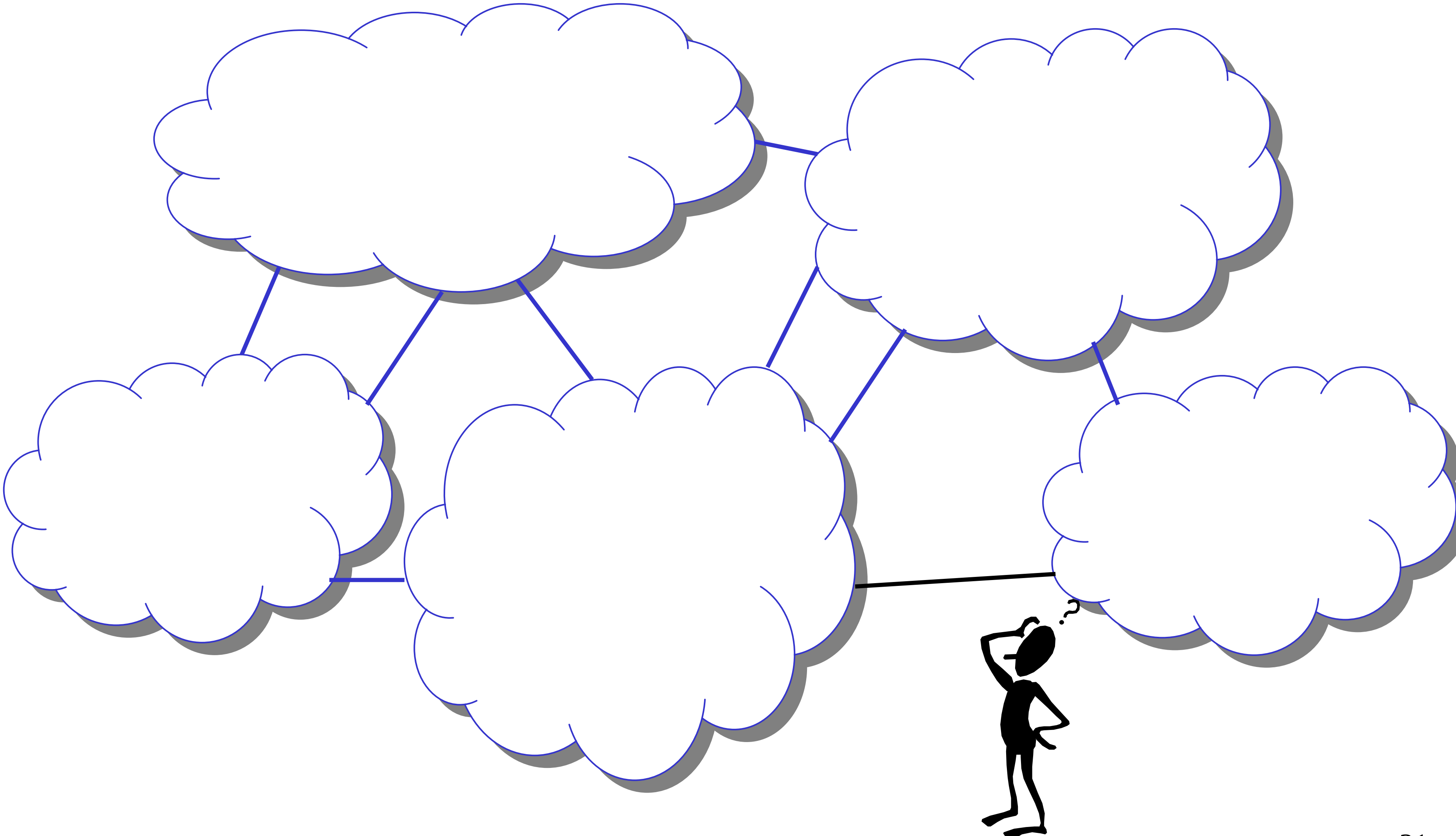


Sampling for Elephants [Estan02]

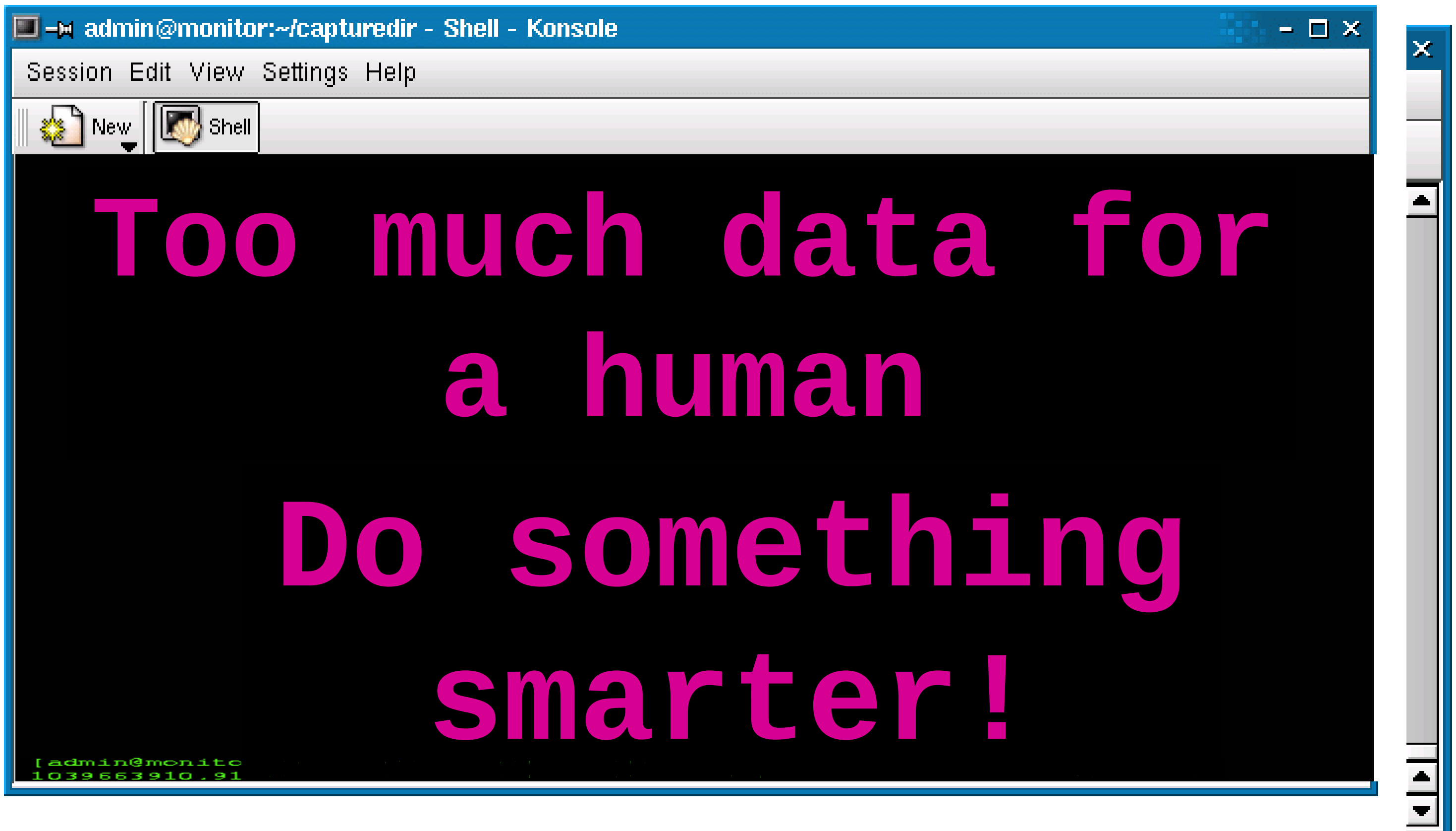


What is the most dominant application and who is the heaviest user?

Who is using my link? [Estan03]

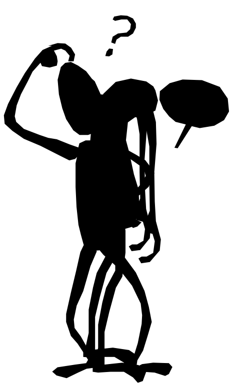


Looking at the traffic



Too much data for a human

Do something smarter!



Looking at traffic aggregates

Src. IP	Dest.	Protoc	Src.	Dest.

Src.	Dest.

Ran	Destination IP	Traffic

Ran k	Source	
3	dorm.bigU.edu	17.8%

- Aggregating o… gives useful re…
  - Traffic report… (e.g. individual IP addr…)
  - Cannot show aggregate… (e.g. which network uses…)
- The traffic analysis to… find aggregates over… granularity

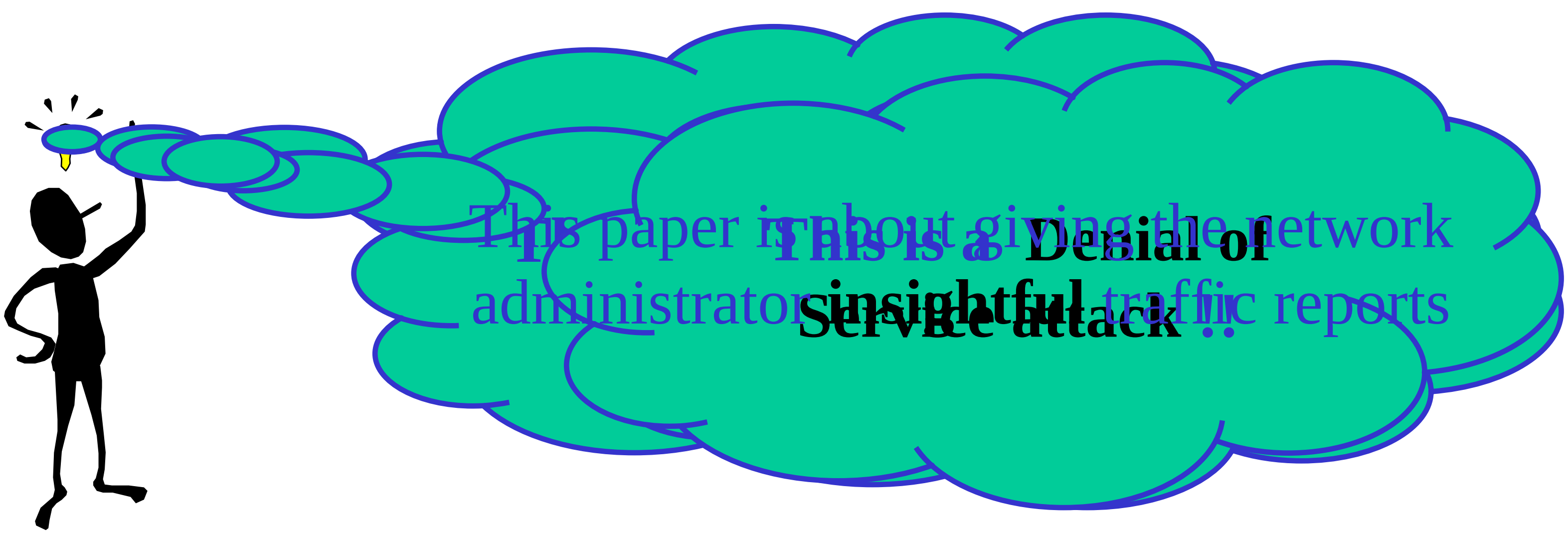
Which network uses web and which one kazaa?

…at apps …e used?

~~Most traffic goes to the dorms …~~

Ideal traffic report

Traffic aggregate	Traffic
Web traffic	42.1%
Web traffic to library.bigU.edu	26.7%
Web traffic from www.schwarzenegger.com	13.4%
ICMP traffic from sloppynet.badU.edu to jeff.dorm.bigU.edu	11.9%

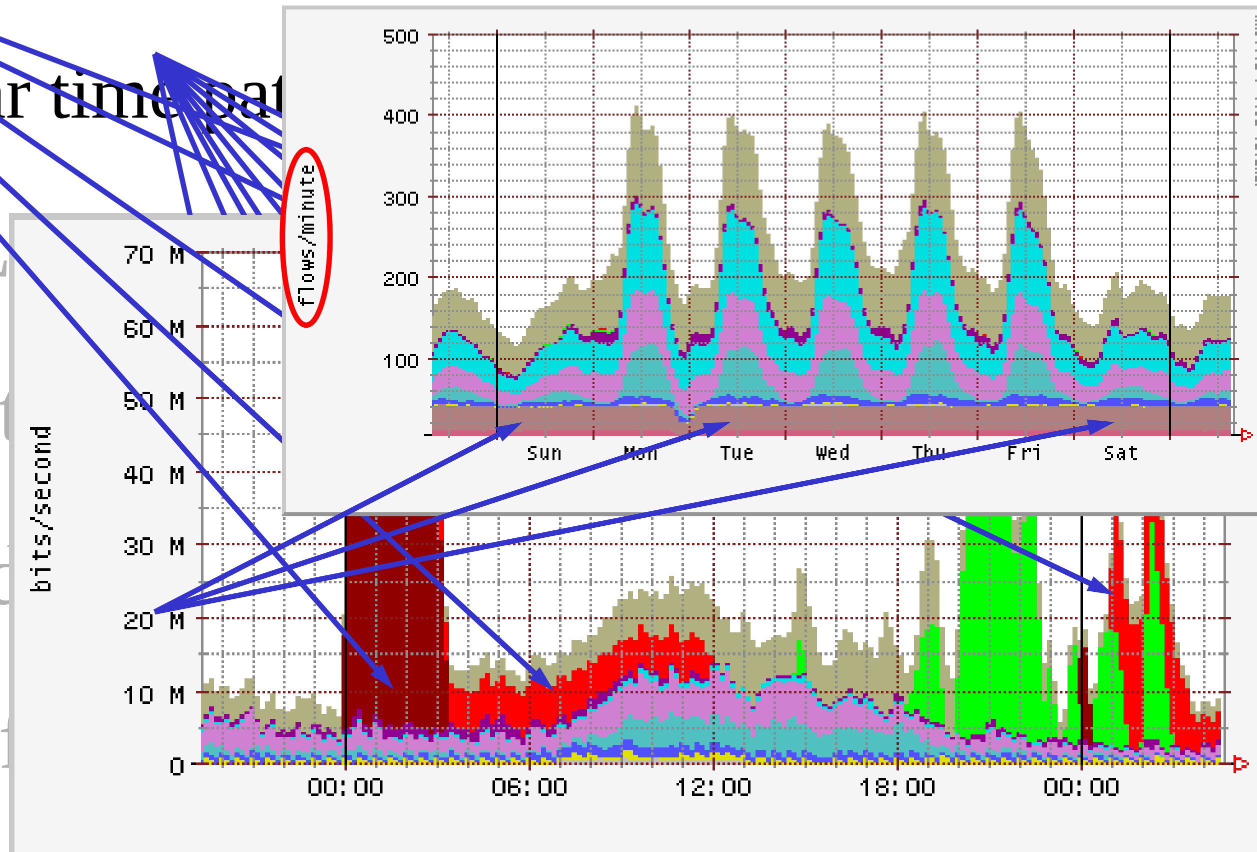


Traffic Clusters and Reports

- Traffic clusters are multidimentional aggregates.
- Traffic reports give volume of chosen clusters
- Only those over threshold are reported
- To avoid redundant data, compress inferrable data (up to error H)
- Highlight non-obvious aggregates with unexpectedness label

Structure of regular traffic mix

- Backups from CAIDA to tape server
  - Semi-regular time pat
- FTP from SL
- Scripps web t
- Web & Squid
- Large ssh traf



- Steady ICMP probing from CAIDA

What is the total # of packet of every flow?

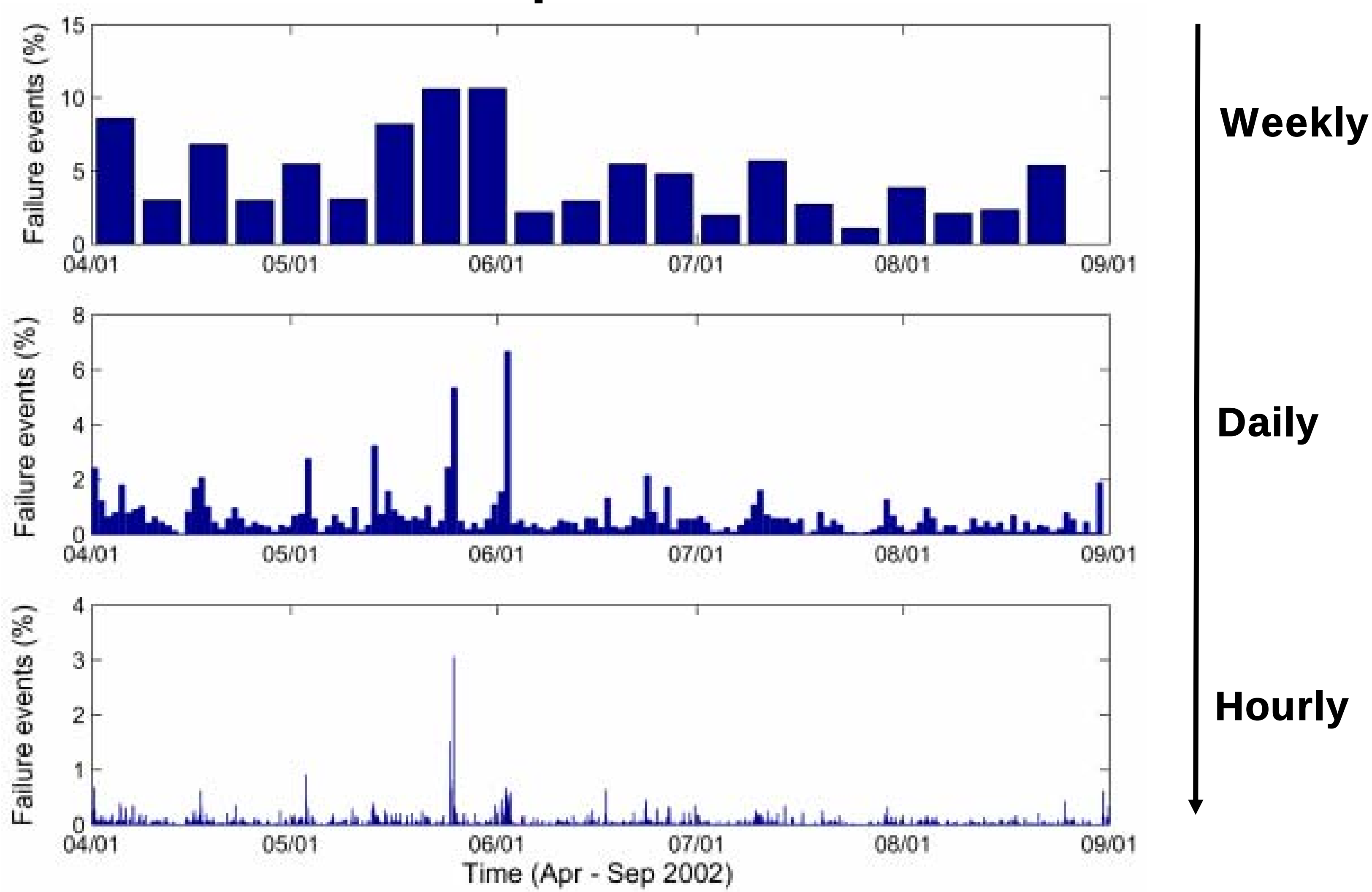
Space-Code Bloom Filter

- Bloom filter answers set-membership.
- Space-code bloom filter answers multiset-membership
- Use a number of “virtual Bloom-filters, spread multiplicity information over space.
- Write-only
- At OC768, it can work at 5ns SRAM
- What about storage space at the router?

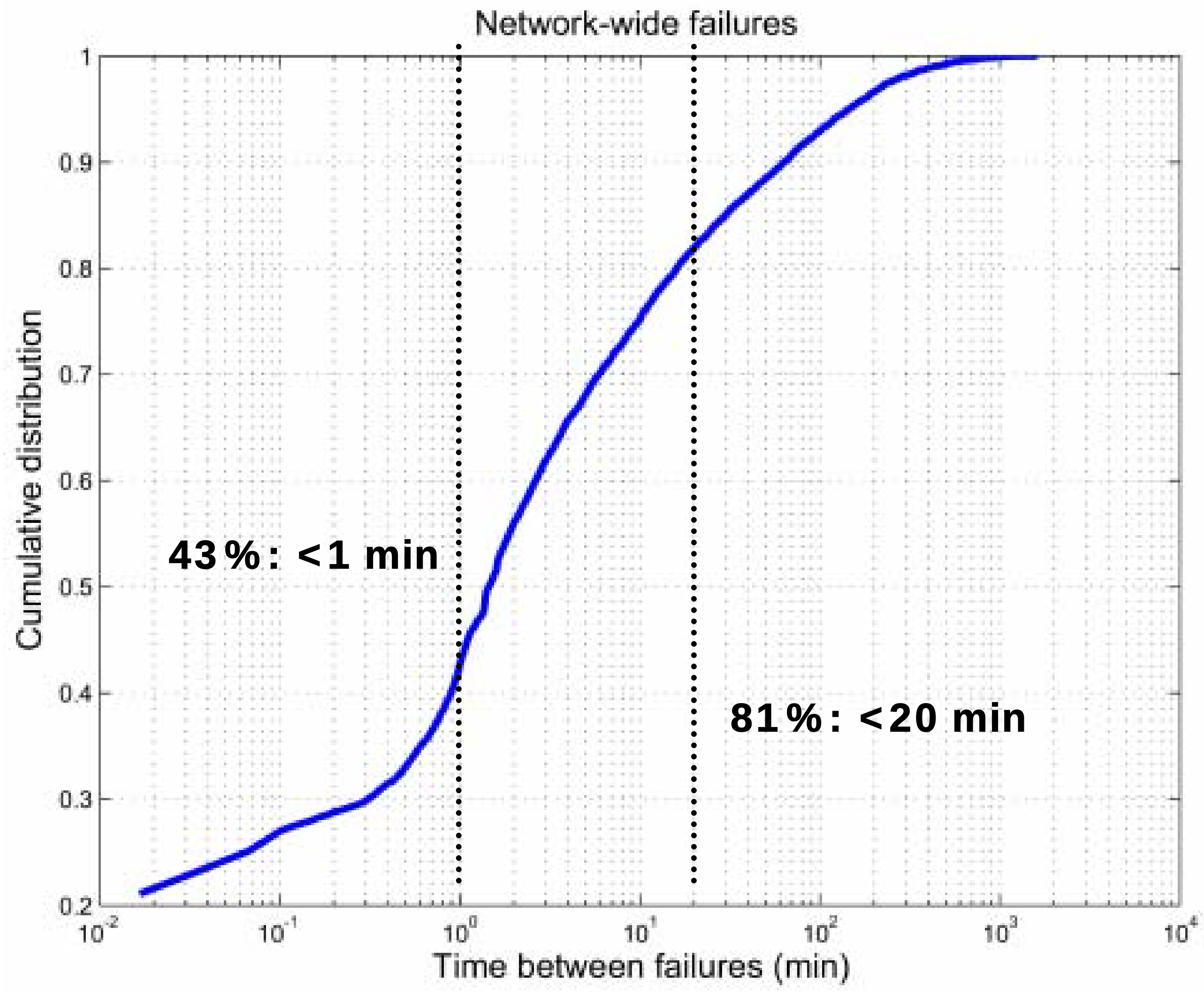
Future Work

- One traffic matrix to rule?
  - Can we answer all questions with one matrix?
- Continuous monitoring
  - data export in real-time
  - query over streaming data
- Availability/survivability
  - Impliations in SLAs?

Failures are part of everyday operations



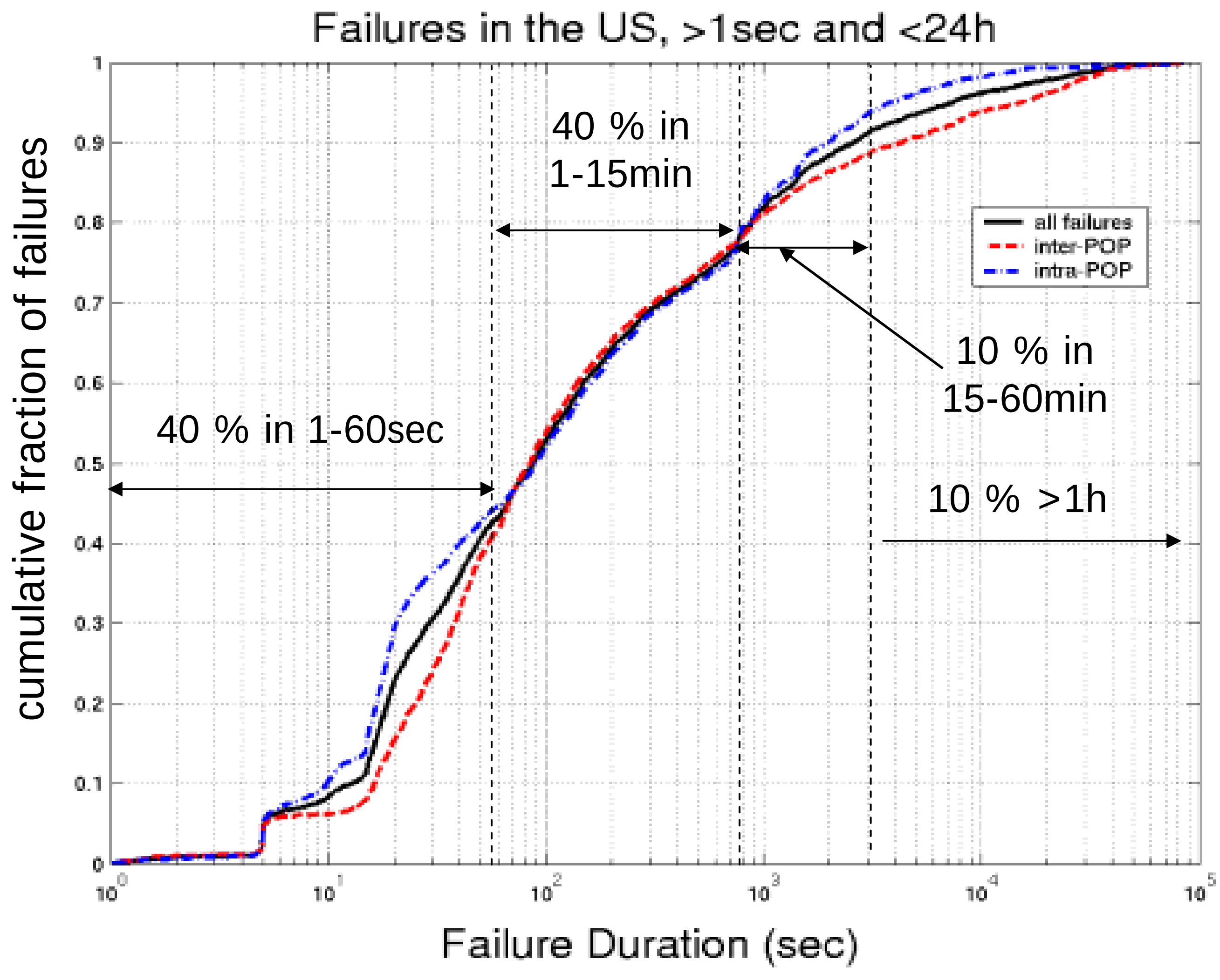
Time between Failures (network-wide)



Sources of failures

- Duration can provide hints, e.g.,
  - long (>1hour): fiber cuts, severe failures
  - medium (>10min): router/line card failures
  - short (>1min): line card resets
  - very short (<1min): software problems, optical equipment glitches
- Other hints
  - shared equipment (routers, optical)
  - router logs (e.g., SONET alarms), etc.

Network-wide Failure Duration



References

[Duffield03] N. Duffield, C. Lund, M. Thorup, "Properties and Prediction of Flow Properties from Sampled Packet Streams," ACM SIGCOMM IMC, Miami, Oct., 2003

[Choi04] B.Y. Choi, S. Moon, Z.L. Zhang, C. Diot, "Analysis of Point-to-Point Packet Delay in an Operational Network," IEEE INFOCOM, Hong Kong, Mar., 2004

[Estan03] C. Estan, S. Savage, G. Varghese, "Automatically Inferring Patterns of Resource Consumption in Network Traffic," SIGCOMM 2003

[Estan02] C. Estan, G. Varghese, "New Directions in Traffic Measurement and Accounting," SIGCOMM 2002

Acknowledgements

- C. Estan’s SIGCOMM 2002 talk.
- S. Bhattacharyya and G. Iannaconne’s ICNP 2003 Tutorial.