IP Network Monitoring and Measurements: Techniques and Experiences

Philippe Owezarski

LAAS-CNRS
Toulouse, France
Owe@laas.fr

IDMS-PROMS, Coimbra, Portugal, November 26th-29th 2002

Outline

- Introduction
- Monitoring problematic
  - Only based on network administration tools
  - Problematic example
- Description of monitoring / measurement systems and projects
- …

Outline (cnd)

- Some results, analysis and trends
  - Traffic characterization
  - Traffic modeling
  - losses
  - Delays in routers
  - Traffic matrices
  - Routing table explosion
  - Summary on QoS
- Threats for the Internet
- METROPOLIS

Introduction

- Deals with both monitoring results and effects on network design, research and management
- Framework of METROPOLIS
- Topic under the spotlight

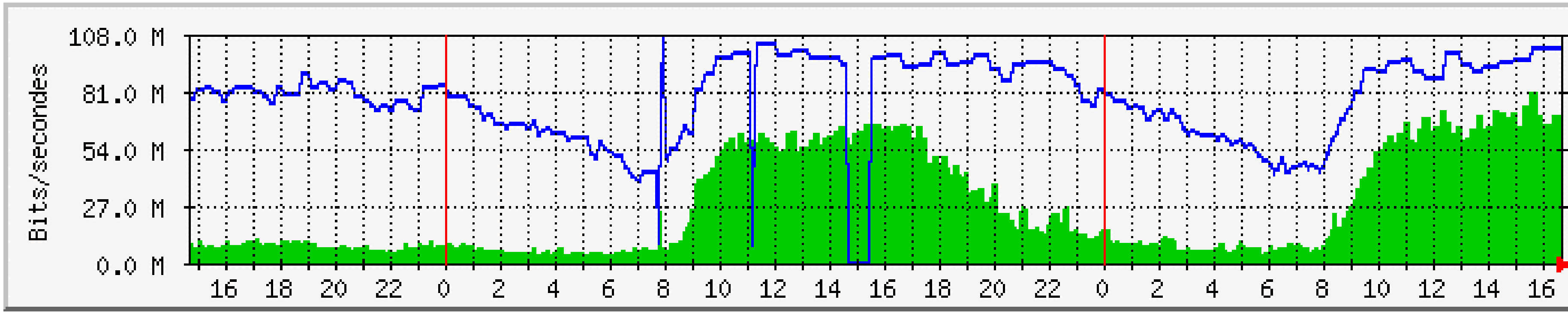
Common solutions for network monitoring

What to use for network monitoring?

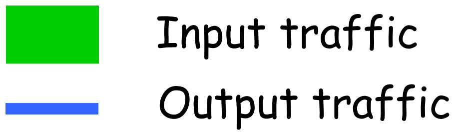
- Administration / operation tools based on SNMP
  - Topology of networks / configuration
  - Some statistics measurements
    - Granularity is too coarse: min = 5 s (but can be 1 hour, 1 day, 1 week or whatever)
    - Measured parameters are more or less the amount of traffic sent and received

Some examples of SNMP results

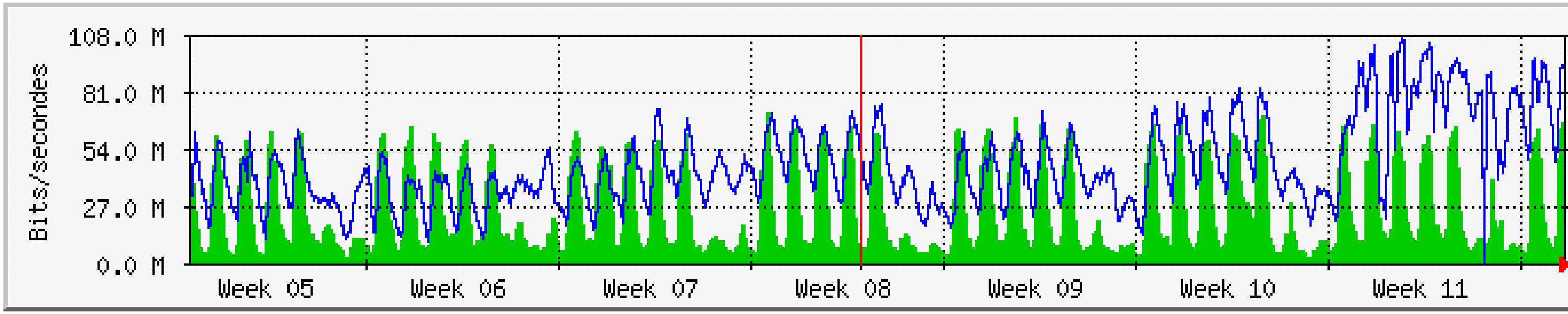
RAP ↔ RENATER interconnection



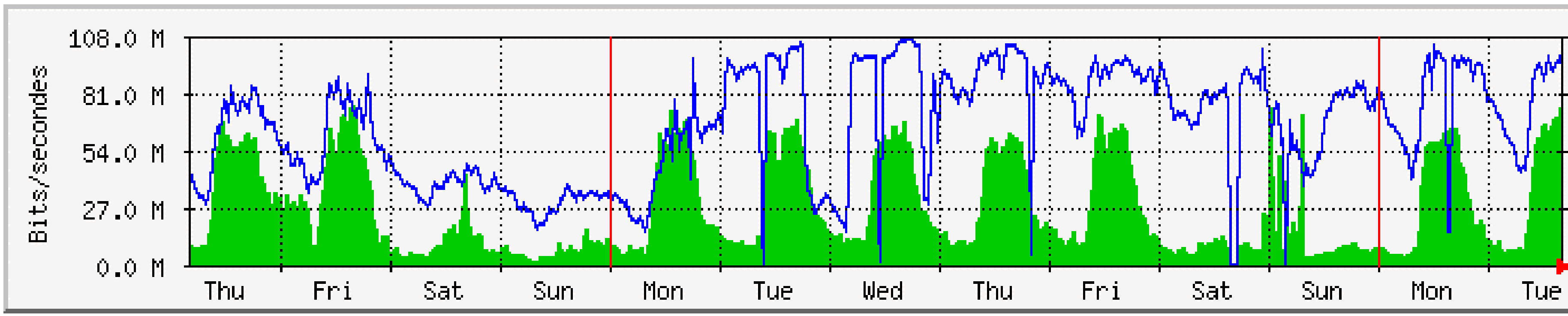
Per hour trace



Some examples of SNMP results (2)



Per Month trace



Per Week trace

Input traffic

Output traffic

Problems for monitoring networks

- Impossible to monitor traffic dynamics (second order values as variability auto-covariance for instance)
- Impossible to monitor traffic QoS (user point of view – goodput)
- Impossible to get a (formal) traffic model

Example on network provisioning

- Common beliefs tell us traffic is Poisson:
  - $E[X]=\lambda$
  - $V[X]=\lambda$
  - Provisioning should be 2λ
- Actually, provisioning has to be at least 1:3 (i.e. 3λ)
  - RENATER 1:3
  - Sprint 1:3
  - WorldCom 1:5
  - AT&T 1:10

Questions on the example

- How explaining this over-provisioning requirement ?
- How to predict the traffic that will be supported by a new network to design ?

?

IP monitoring: goals and importance

- Network and traffic exist and is full of information
- Help to predict what will be the traffic in the future based on some current trends
- Help to design and provision a network and Internet protocols

IP monitoring: goals & importance (2)

⇒ Monitoring changes the network engineering and research process

⇒ Monitoring is a new service that must be provided by vendors, carriers and ISP (technical and commercial adds) and strongly requested by users

Monitoring concerns

- Network design
- Traffic engineering / routing tables
- Network management
- Provisioning
- Pricing / charging
- QoS monitoring
- Assessment and tuning of mechanisms as
  - QoS (IntServ, DiffServ, IPv6, MPLS, ...)
  - Traffic engineering (OSPF, MPLS...)

IP Monitoring and Research

- New protocols and architectures for:
  - Traffic characterization and modeling
  - Multi-domains QoS guaranty
  - Service and network utilization optimization
  - Network or VPN or CoS provisioning
  - QoS routing
  - Network security (?)

- Techniques and mechanisms for:
  - pricing

State of the art
(as far as I know)

Active vs. Passive Measurements
Some Monitoring Projects

Active measurements

- Active measurements
  - Consists in sending packets on a network and observing results (Delay, RTT, Throughput, etc.)
  - User point of view
  - Best solution to evaluate the service you can get from the network you're connected to
- Drawbacks
  - Probe packets change the state of the network
    - →IETF IPPM WG is working on the definition of probing scenarios minimizing the effects on the network state

Some active measurement tools

- Ping
- Traceroute
- MGEN
- RIPE equipments
- Etc.

⇒ Importance of clock synchronization: most of the time GPS is required

Projects based on active measurements

- Surveyor (NSF) based on ping and GPS clocks
  - Delays
  - Loss
- NIMI (Paxson/ACIRI)
  - Worldwide (USA+CH) measurement infrastructure
  - Distance matrix in the Internet
  - QoS parameters (delays, loss, throughput, …)

Projects based on active measurements

- RIPE (Europe)
  - Similar to NIMI
  - GPS clock on every measurement point
  - Statistics on QoS on some links
    - General analysis or on demand
- MINC (Multicast INC)
  - Use NIMI infrastructure
  - Generate multicast probe packets
  - Infer internal structure of the network
  - Tomography

Projects based on active measurements

- UINC (Unicast INC)
  - As MINC but using unicast probe packets
  - Multicast is not always available
  - Links and traffic are not symetrical
- Netsizer (Telcordia ex Bellcore)
  - Measure the increase of the Internet
  - Detect points of congestion
  - Delays
- AMP (NLANR)
  - Active probing

Passive measurements

- Capture packets (or headers)
- Not intrusive at all
- Carrier / ISP point of view
- Best solution for a carrier to measure traffic

- Drawbacks
  - Sampling issues
    - Creation of a new IRTF WG (IMRG)
  - Difficult to get a user point of view
  - Technical limits (speed of components, capacity)

On line vs. Off line measurements

- On line
  - Packets are analyzed in real-time
  - Analysis on very long periods
  - But complexity of analysis is quite limited
- Off line
  - Packets are stored on hard drives / SAN for later analysis
  - Possibilities of analysis are endless
  - Possibility of correlating several traces
  - But amount of stored data is really huge (small periods only)

Passive measurement tools

- TSTAT
- NTOP
- LIBCAP
- Tcpdump
- Tcptrace
- QOSMOS
- IPANEMA
- CISCO's Netflow
- OCxMON (mainly ATM)

Projects based on passive measurements

- Netscope (AT&T)
  - Based on Netflow
  - Relations between traffic crossing network nodes and routing table
  - Tomography
  - To improve routing policies
  - To improve load balancing
  - To increase QoS

Projects based on passive measurements

- Paxson/ACIRI
  - Proposal for a model of flows and packets arrivals
  - A reference since 1995 !
- CAIDA
  - Based on OCxMON
  - Monitoring of vBNS
  - Evolution of traffic on long periods (new applications, behavior of users, etc.)
  - Analysis tool: CoralReef
  - Representing the Internet

Monitoring efforts … in France …

- NetMet
  - Analysis tool suite for CISCO's NetFlow traces
  - Designed by and for network administrators (Nancy)
  - 2 approaches:
    - Flows trace
    - Macroscopic view of the traffic
  - Used by Renater, Remip2000, etc.

Monitoring efforts … in France … (2)

- NetSEC
  - Used for attacks detection
  - Off-line
  - Based on NetMet
- METROPOLIS

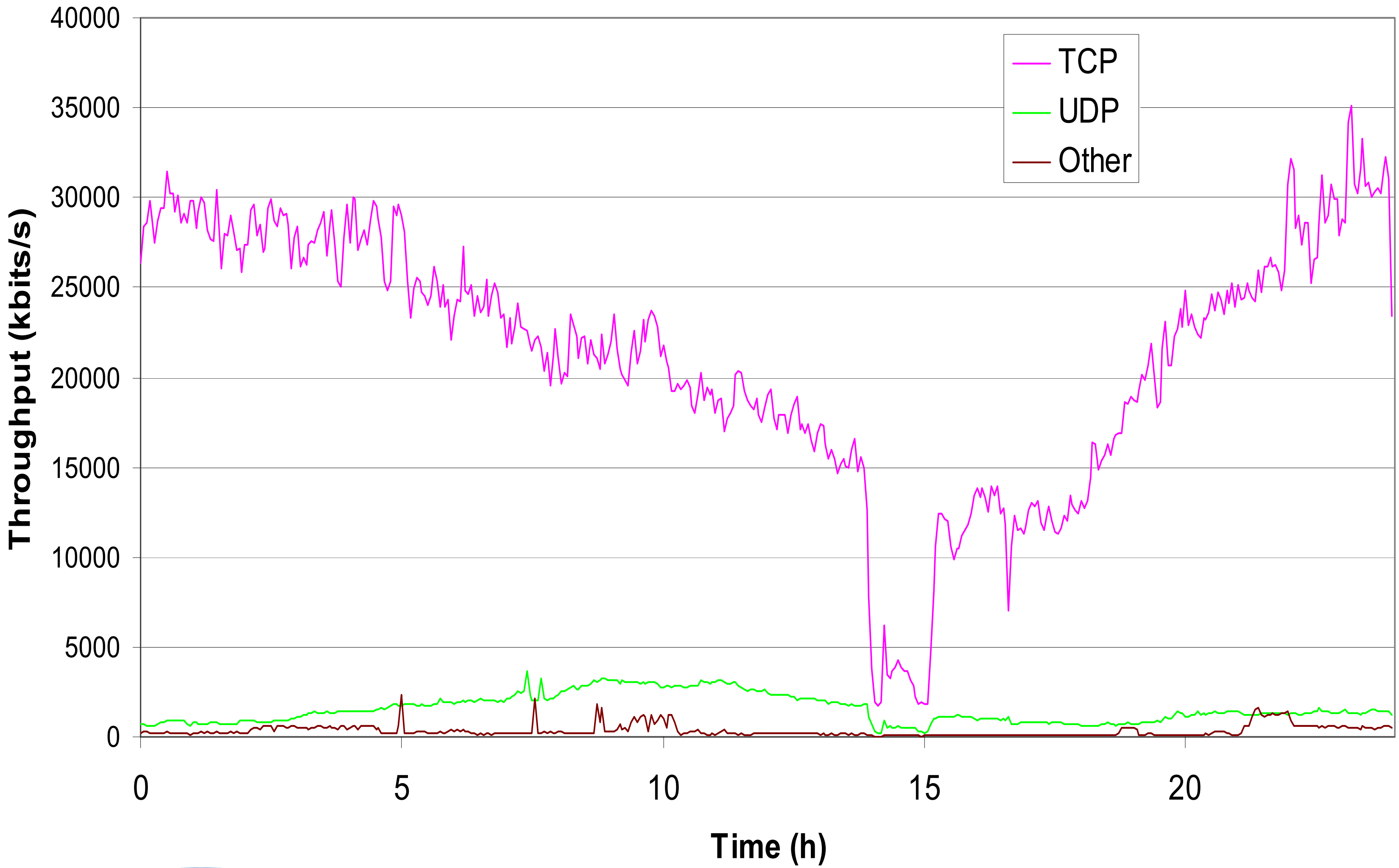
SPRINT and METROPOLIS (passive measurements)

- Insert optical splitter on network links
  → transparent system, not intrusive
- Data from an operational IP backbone
- Integrated system to collect packet-level, flow-level, and routing measurements
  - Collect and timestamp all IP headers (44 bytes) with GPS timestamps (accuracy > 2 µsec)
  - POS/ATM/Ethernet PCI network interface (DAG: University of Waikato /Endace, NZ)
  - Collect routing information (IS-IS, OSPF, BGP)

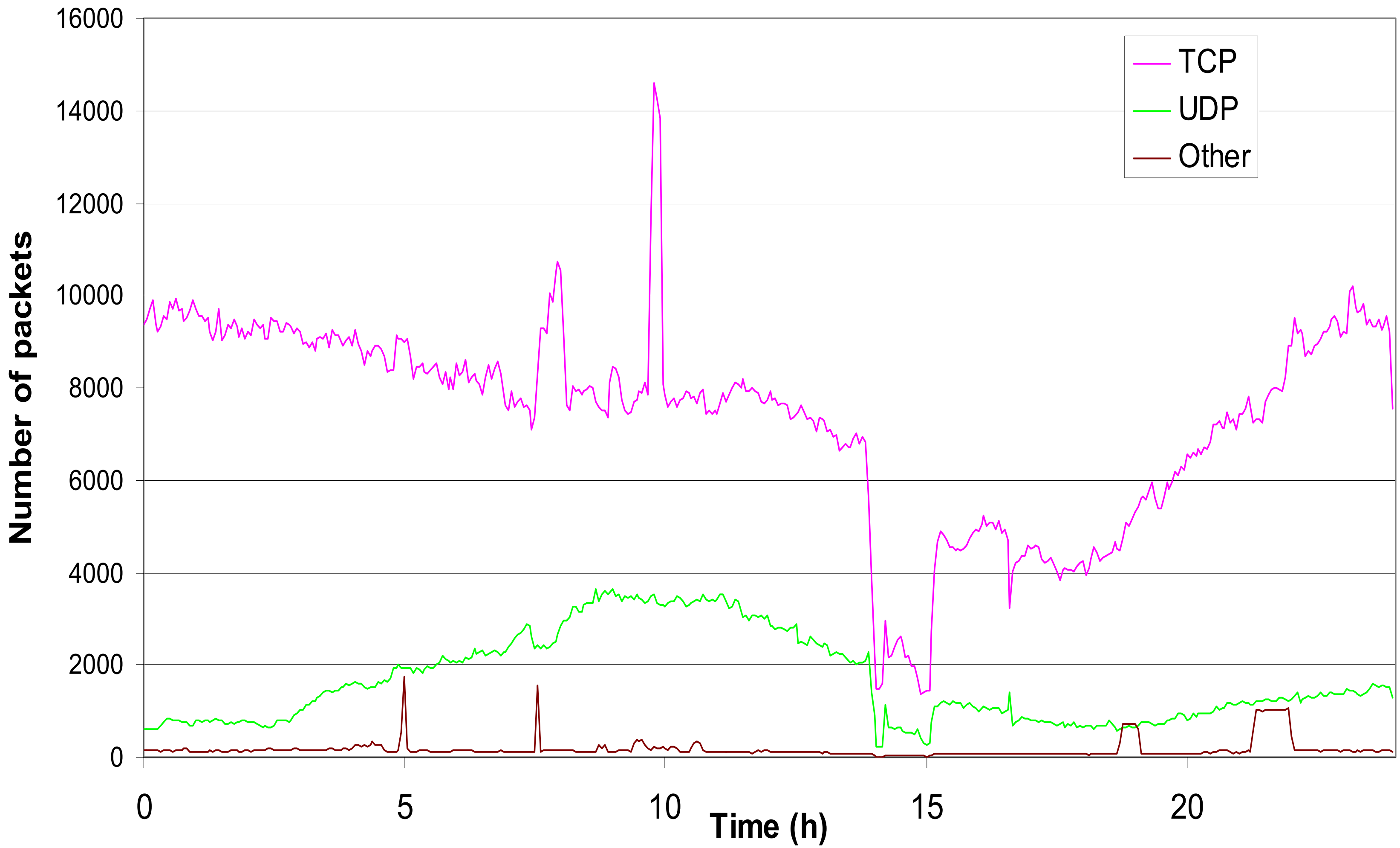
Some Results, analysis and trends

Traffic characterization

Link Utilization: bandwidth

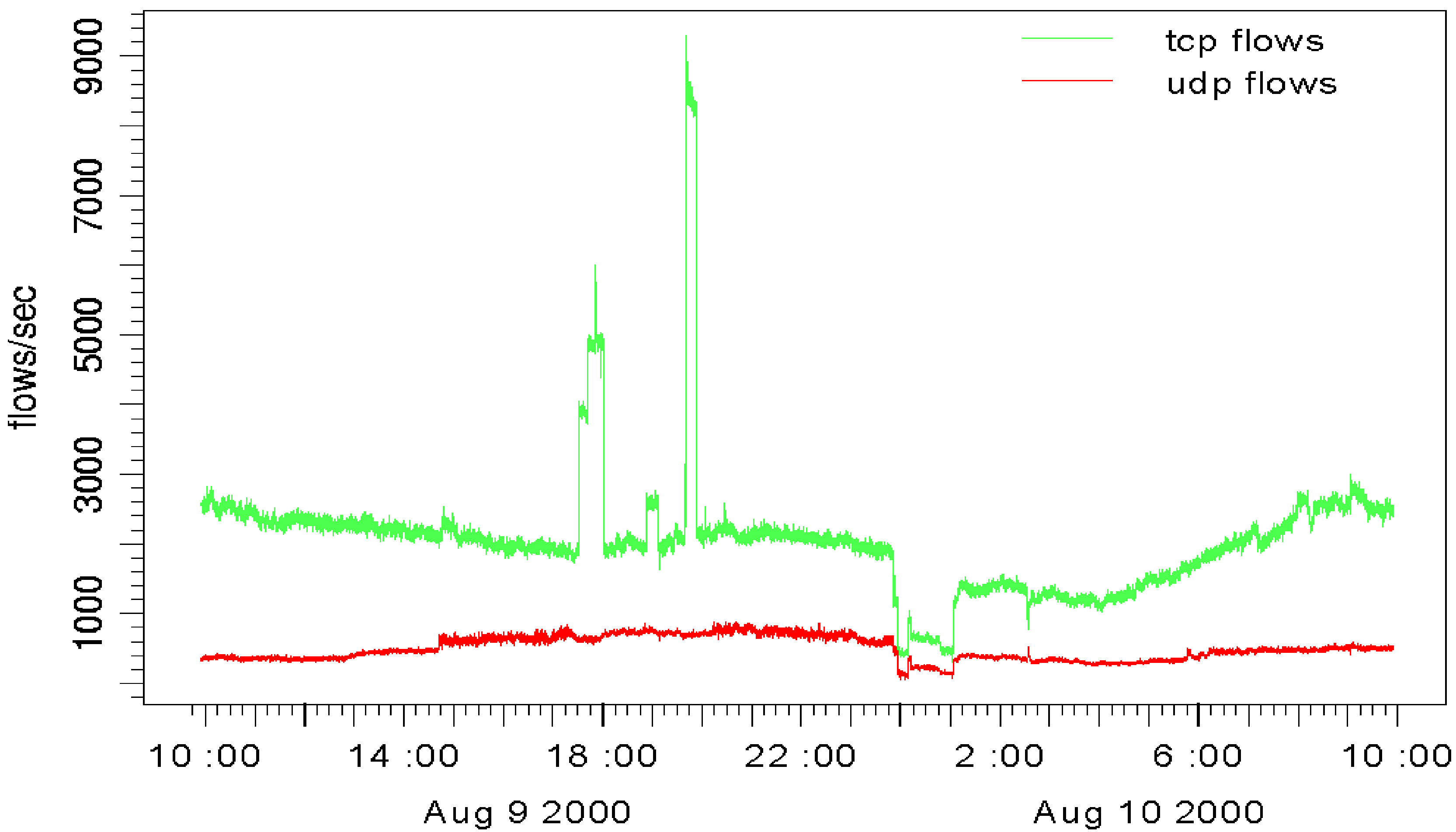


Link utilization: packets

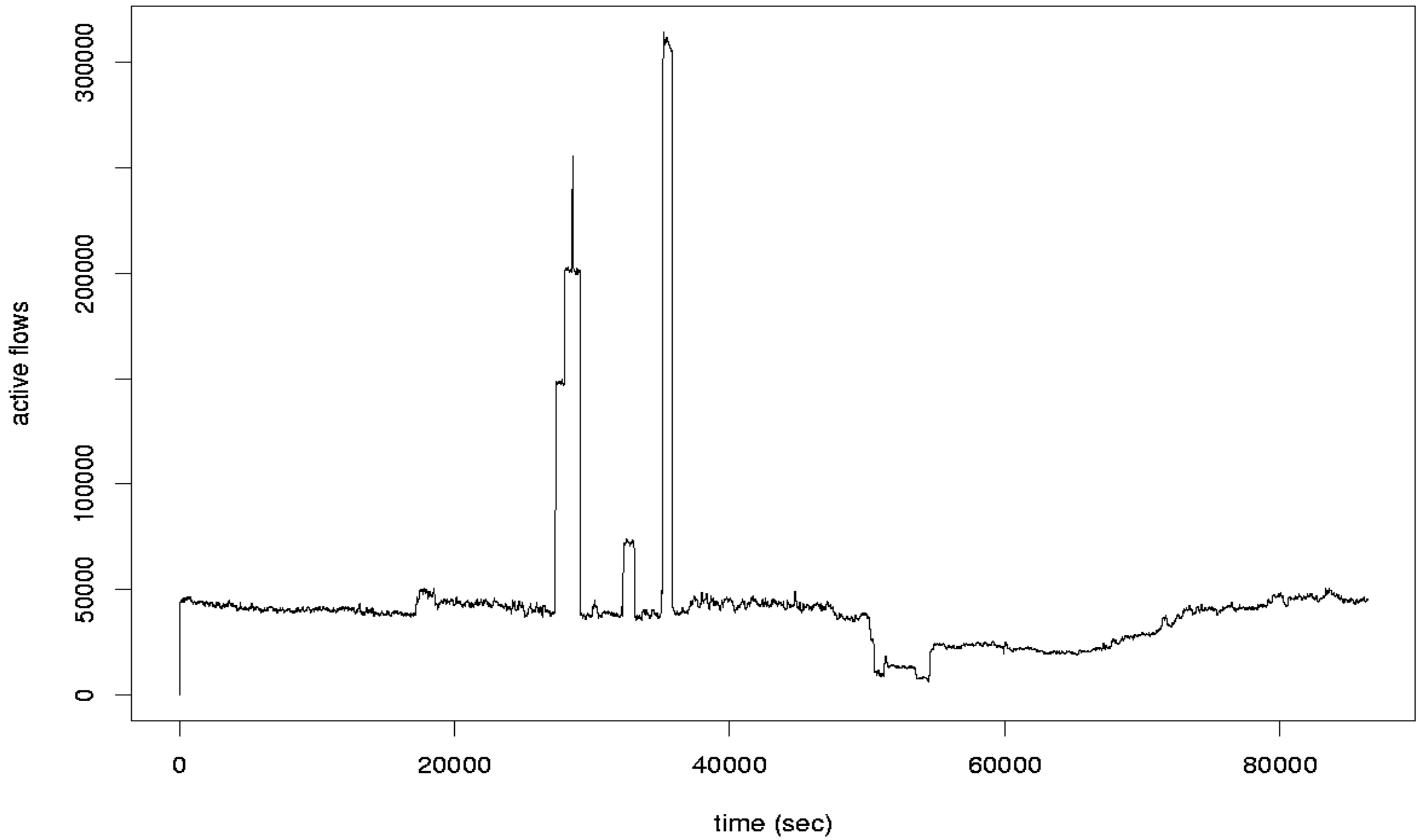


IDMS-PROMS, Coimbra, Portugal, November 26th-29th 2002

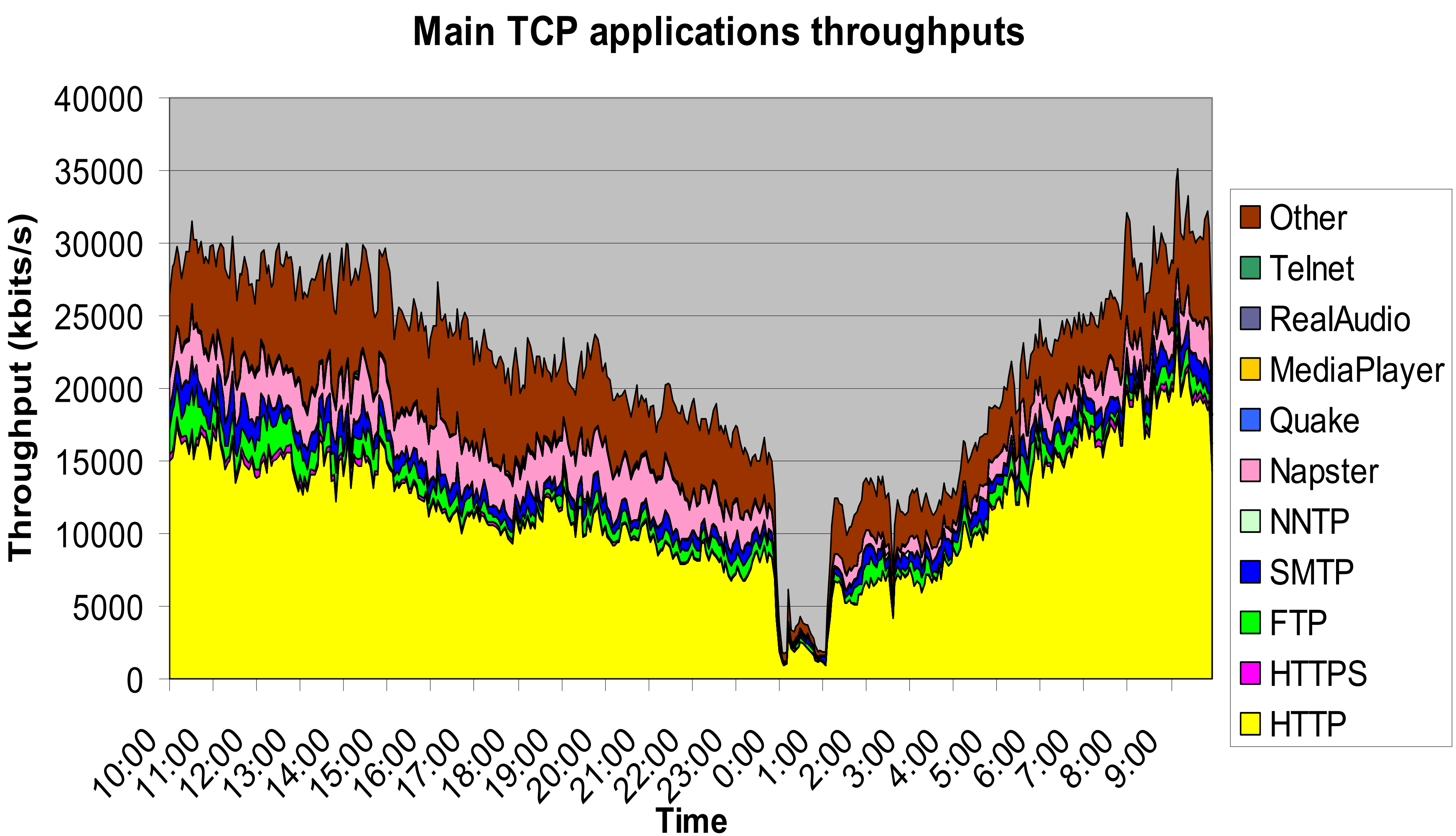
Link utilization: instantaneous flows



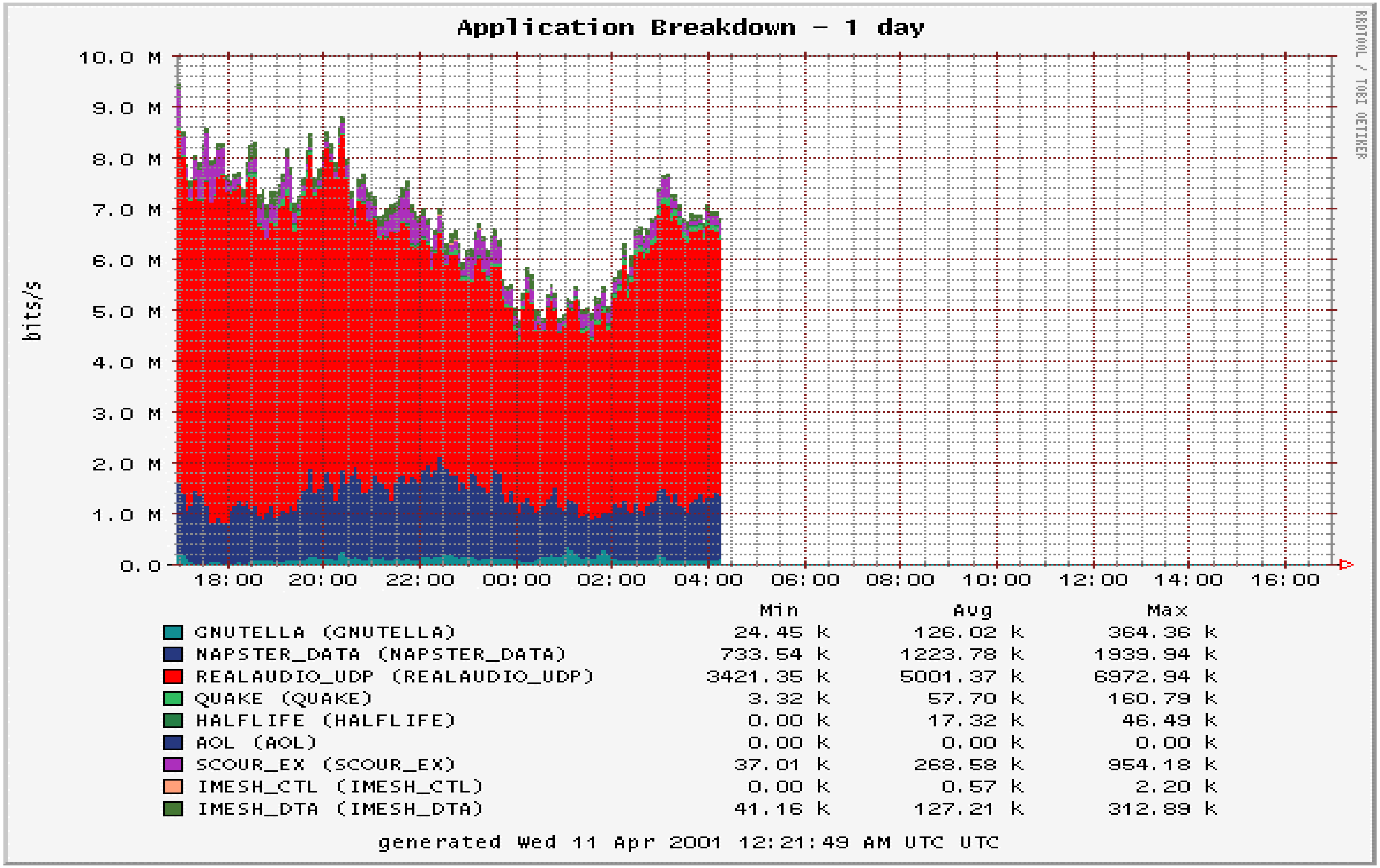
Link utilization: active flows



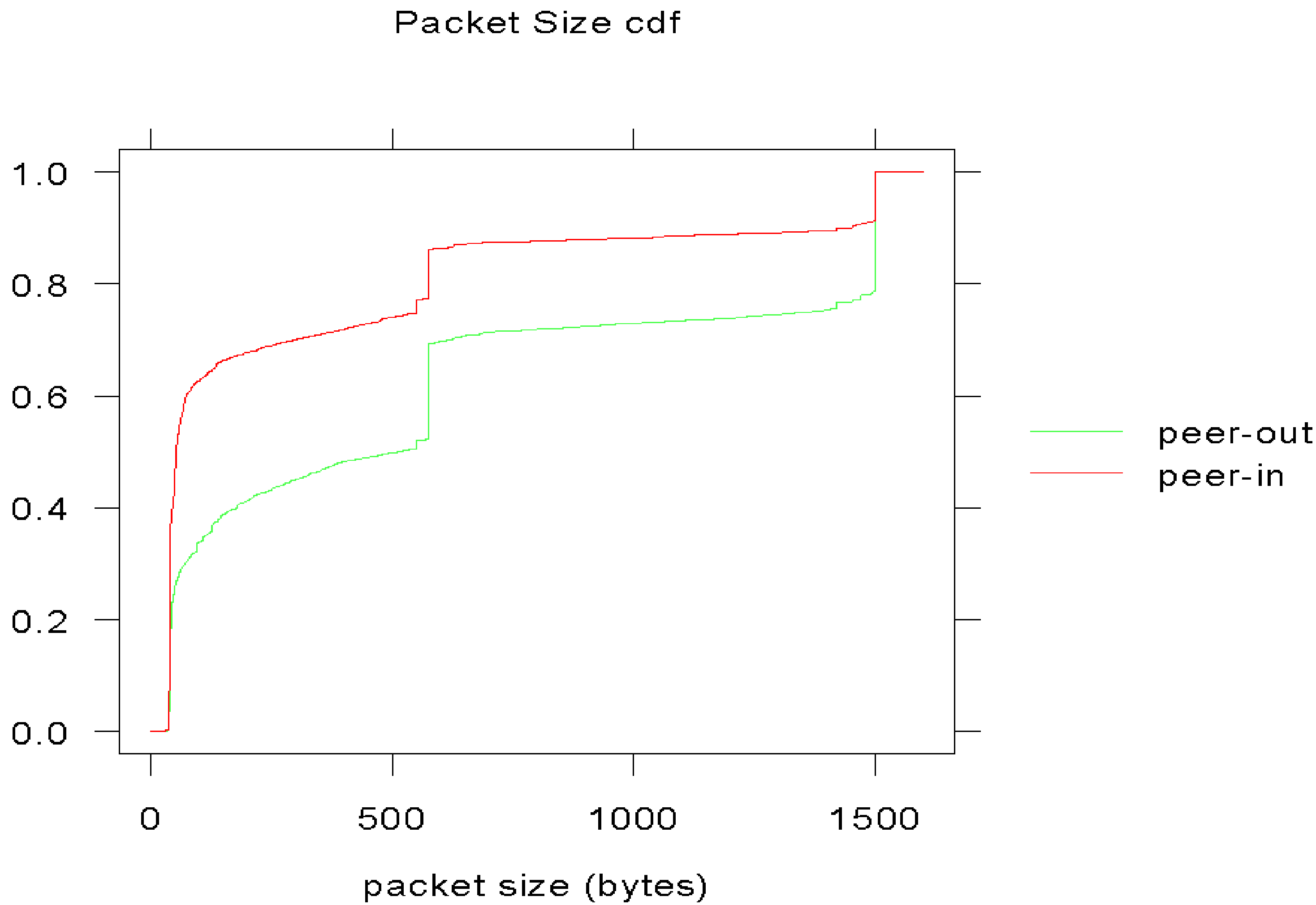
Link Utilization: applications (→sociology?)



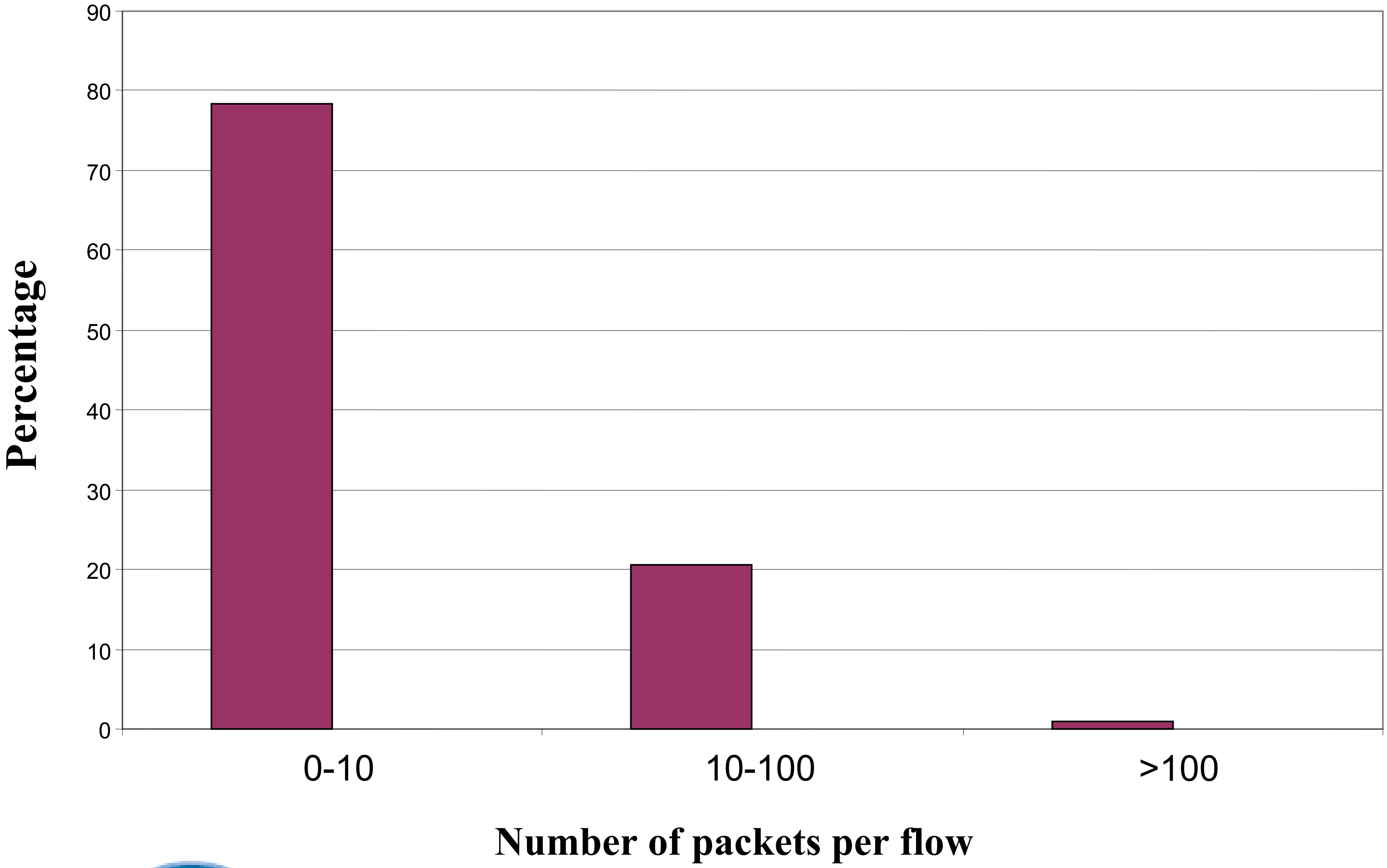
Link Utilization: emerging applications



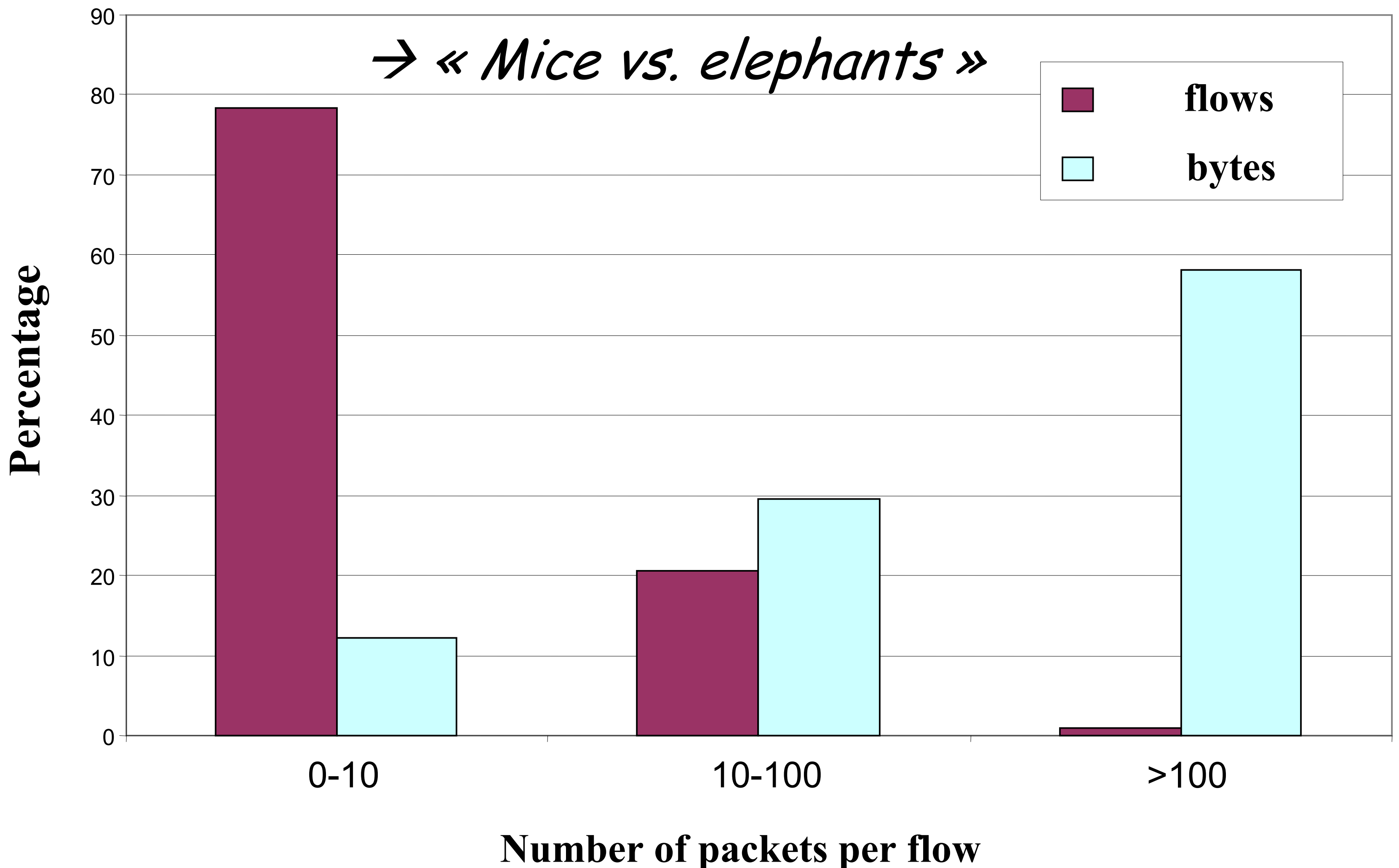
Packet size cumulative distribution



TCP flow size

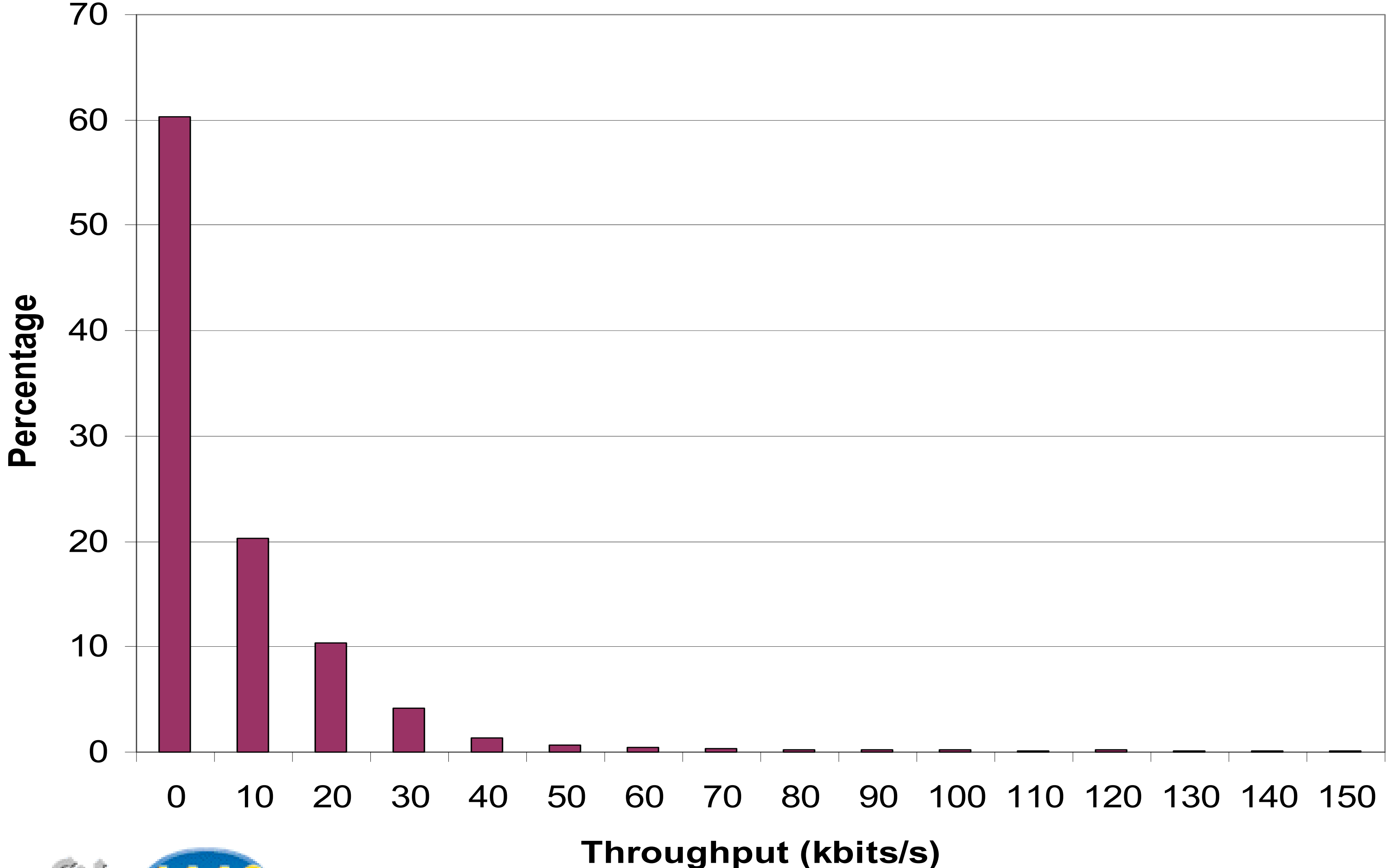


TCP flow size vs. total bandwidth

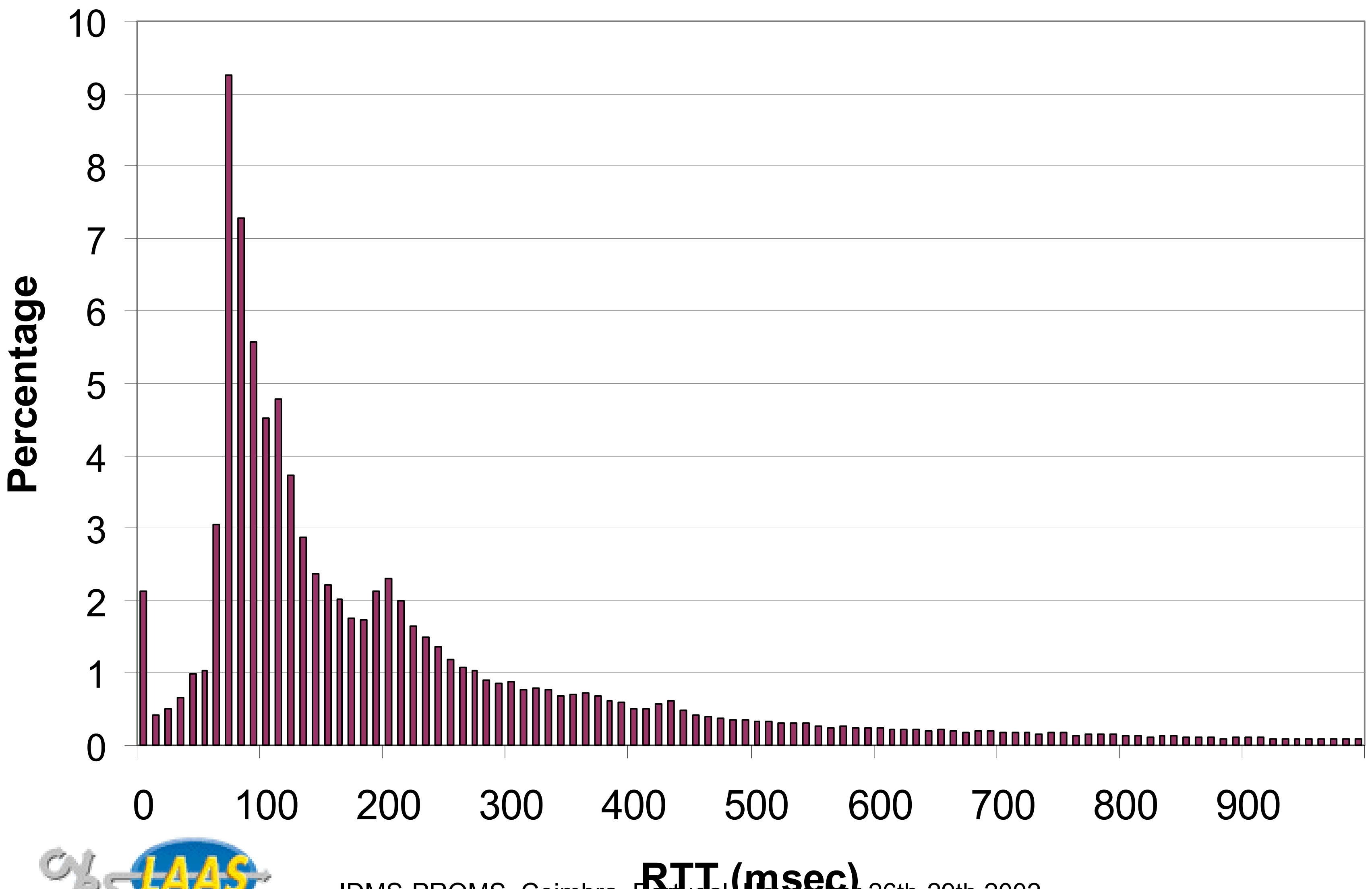


IDMS-PROMS, Coimbra, Portugal, November 26th-29th 2002

TCP throughput



TCP flows RTT



Traffic modeling

Why modeling Internet (TCP) traffic ?

- Different from common thinking i.e. telephone model (Poisson, Gilbert)
- Give information on how designing, managing, ***provisioning*** and operating an IP network
- Give information on future research directions
- Allows researchers to simulate new technical proposals

Previous work on traffic modeling

- **Self-similar**
- *Multi-fractal*
- *LRD*

- Due to:
  - Heavy tailed distribution of flow size
  - TCP-like congestion control
  - Routers
  - Human and application behavior

Self-similarity

- Internet traffic is said to be self-similar
- Self-similar ? What does it mean ?
- Is it bad ?

Self-similar process

- A process X(t) is self-similar, with self-similarity parameter H (the Hurst parameter), iff for any $c > 0$, $c^H X(t)$ and $X(ct)$ have the same joint distributions of all orders. That is, for any integer n, $t_1, \ldots, t_n, x_1, \ldots, x_n$

  $$P(X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n) = P(X(ct_1) \leq c^H x_1, \ldots, X(ct_n) \leq c^H x_n)$$

- Sample paths look qualitatively the same at different time scales.

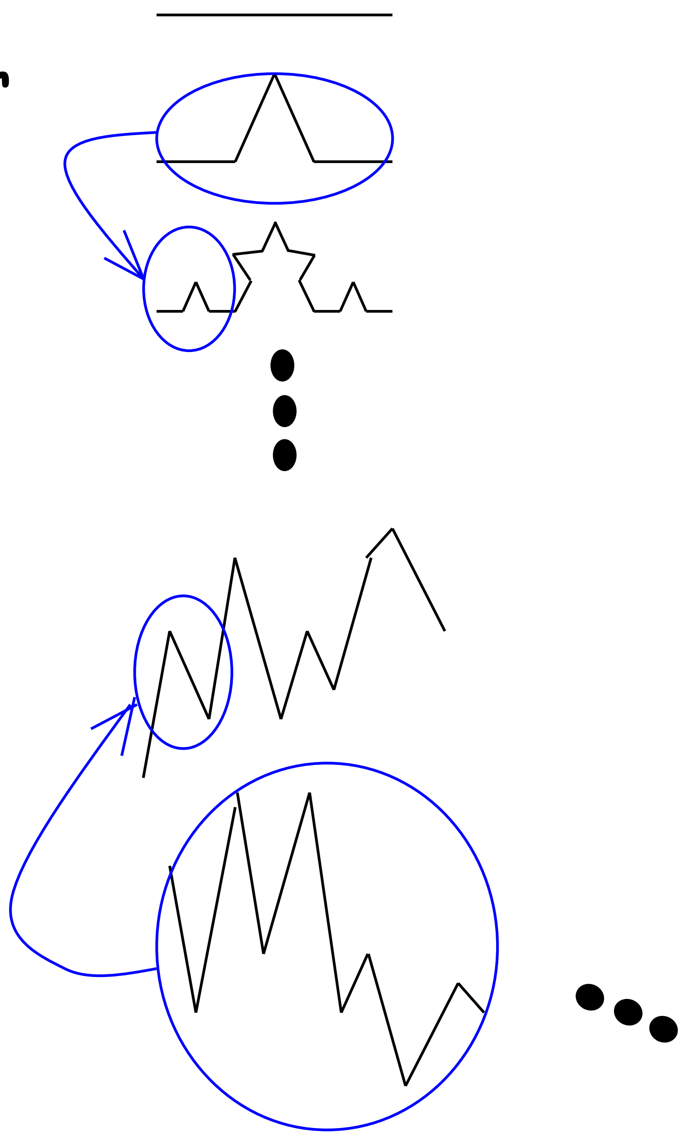
LRD

A covariance stationary model $\{X_t\}_{t \in Z}$ is said to be a long memory process if

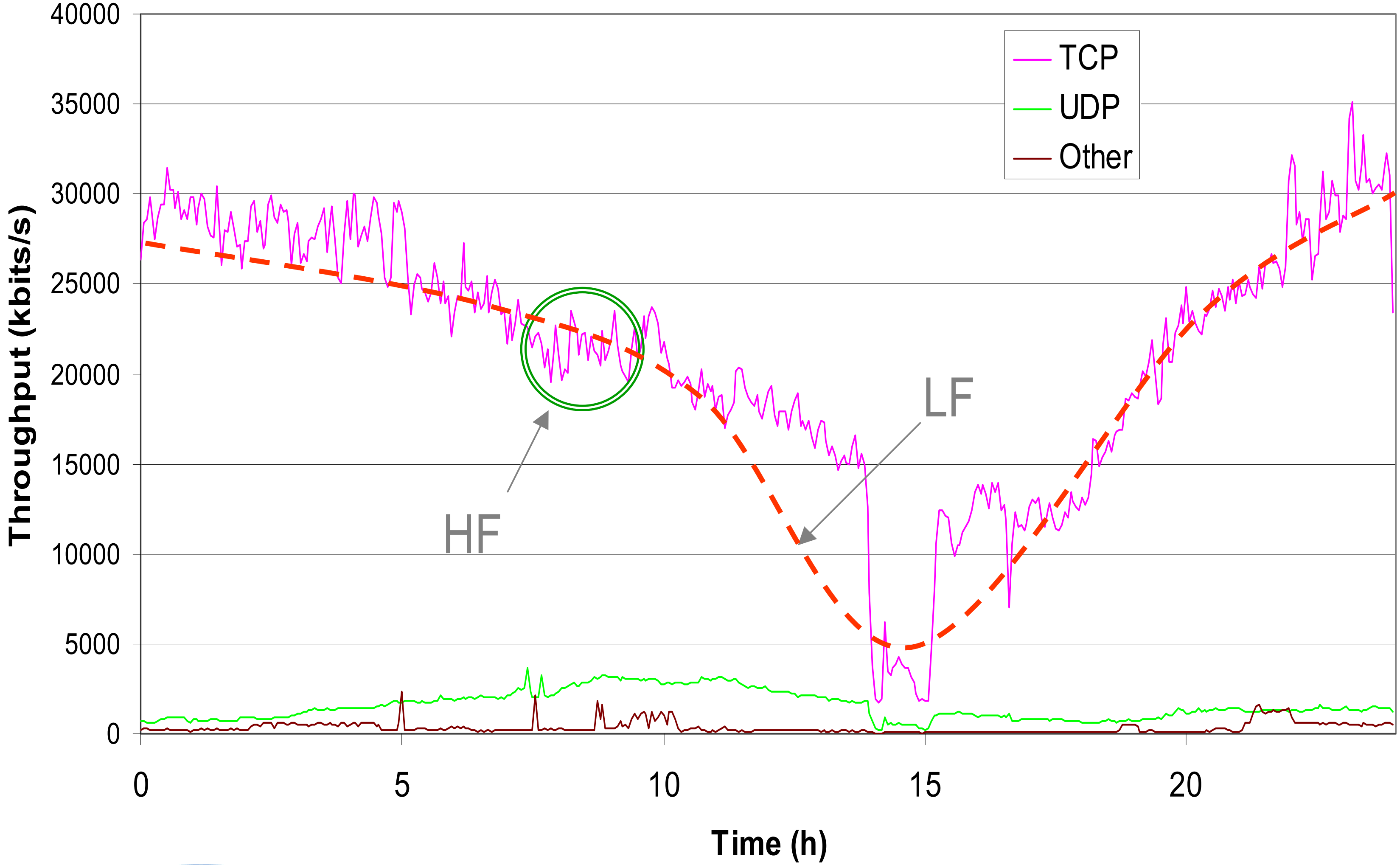
$$\sum_{\tau} \left| \mathrm{cov}(X_\tau, X_0) \right| = \infty$$

Fractals

- Deterministic self-similar shapes repeat themselves exactly as we go closer
- Sample paths of a self-similar process look qualitatively the same, irrespective of the distance at which we look at them. Does not mean that the same picture repeat itself.



Actual traffic



Actual traffic visual analysis

- Suspicion of Self-similarity

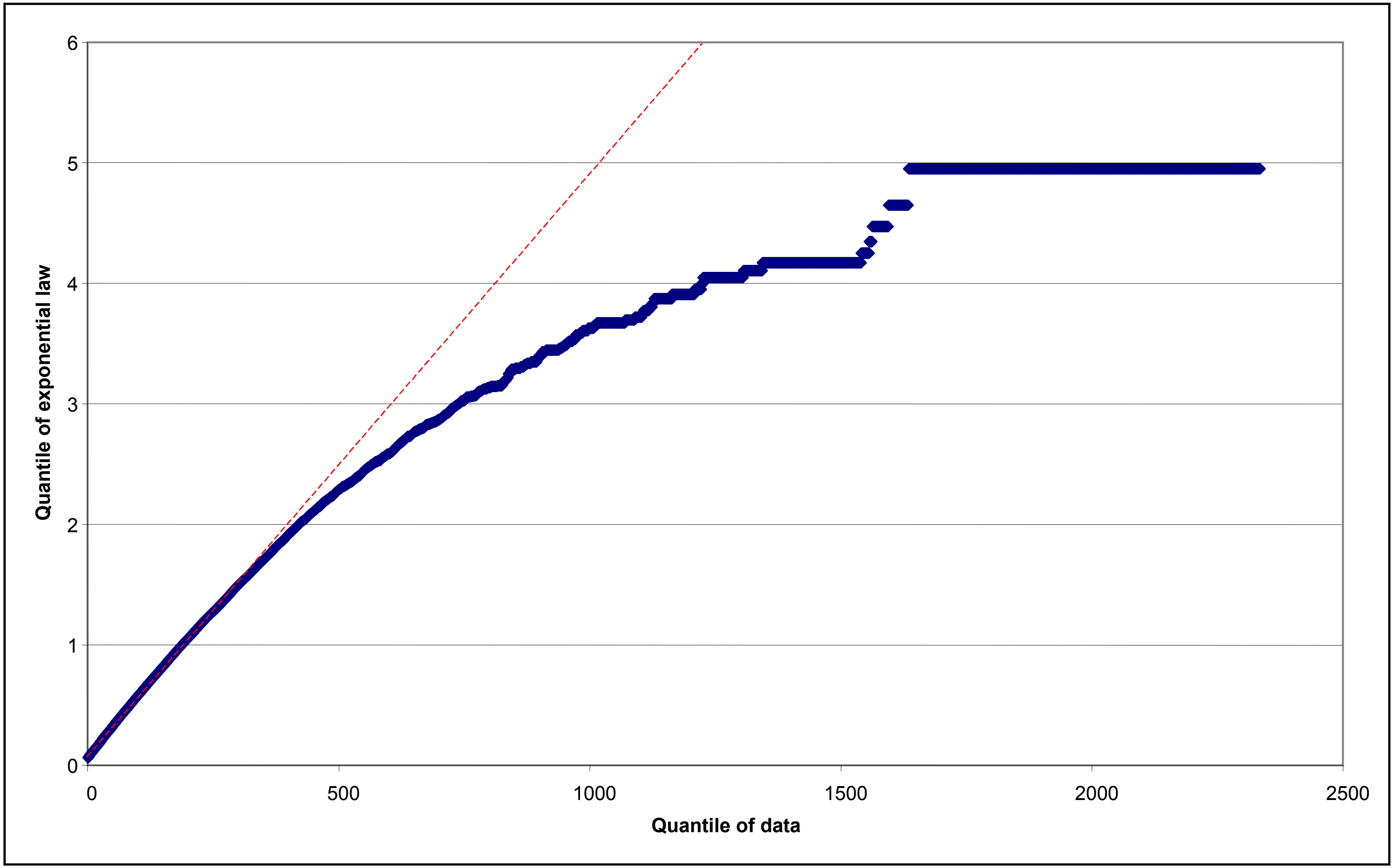
- Variability of traffic profile at all scale is a major matter for:
  - QoS
  - Stability
  - Performance
  - …

How to evaluate H

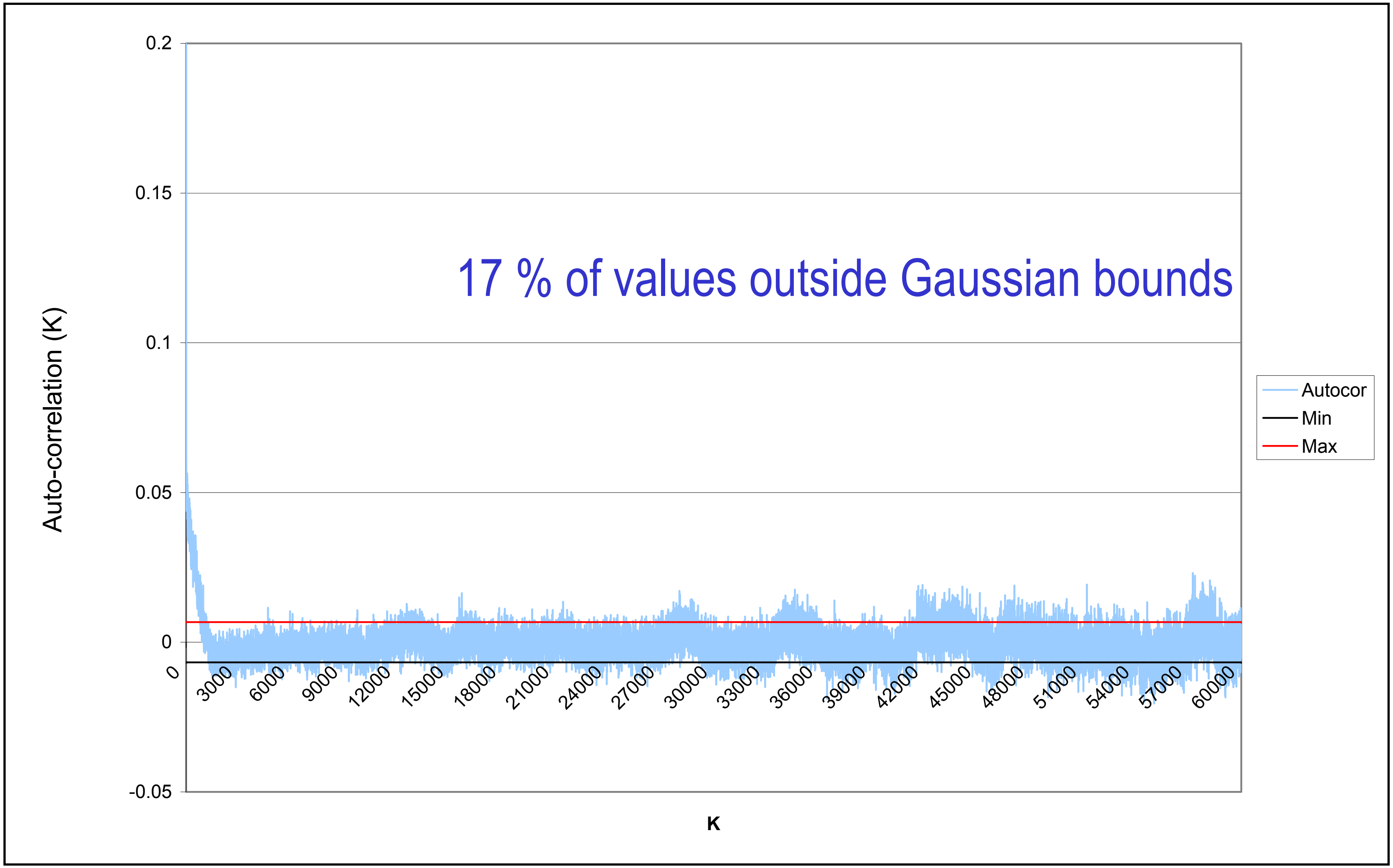
- Related to auto-correlation
- Some tools based on semi-parametric technique exist
  - Periodogram
  - Spectral based analysis
  - Wavelet based analysis (→ LDestimate)

Access link traffic analysis
10 Mbps

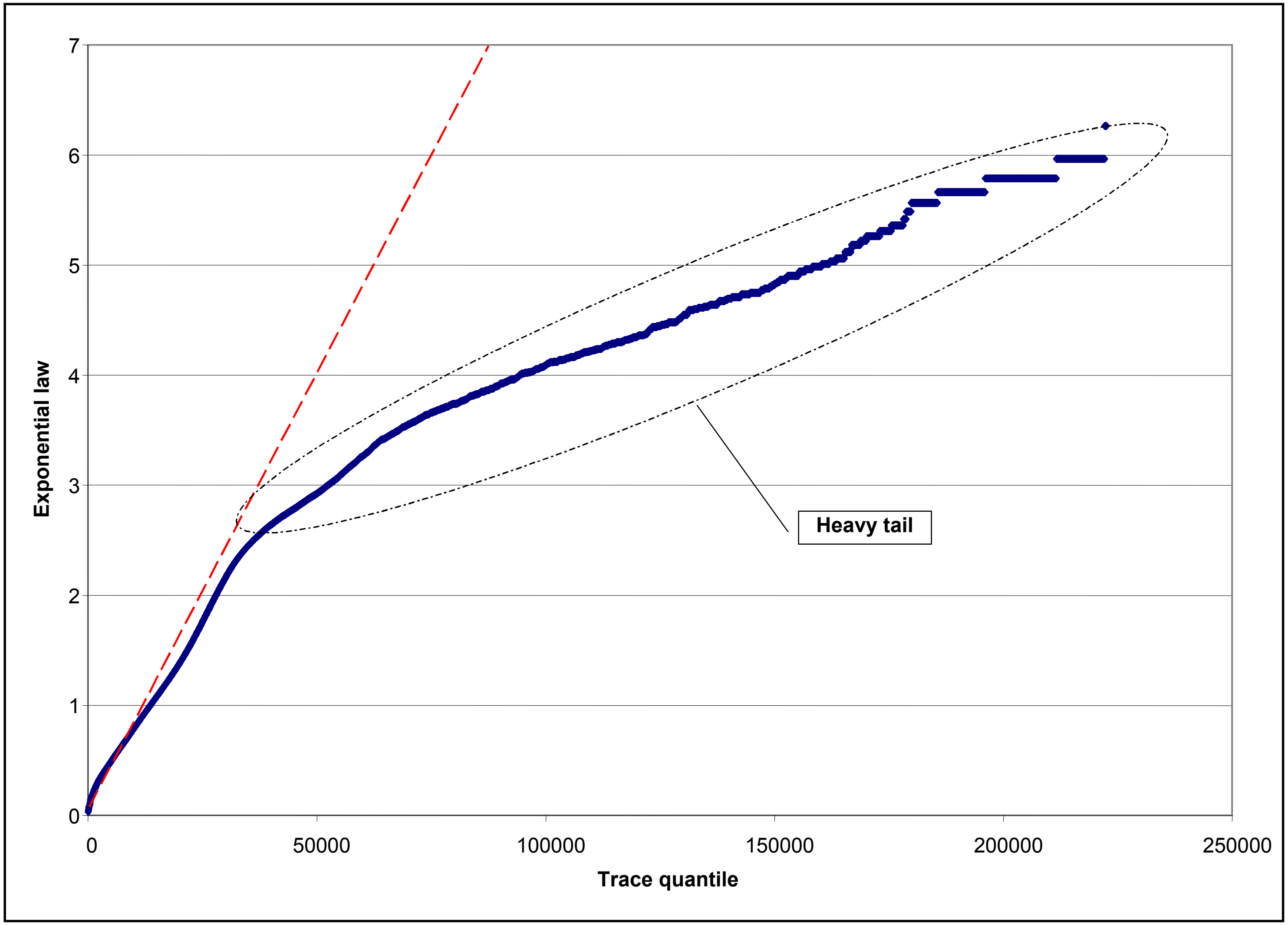
QQ-Plot of flow arrivals



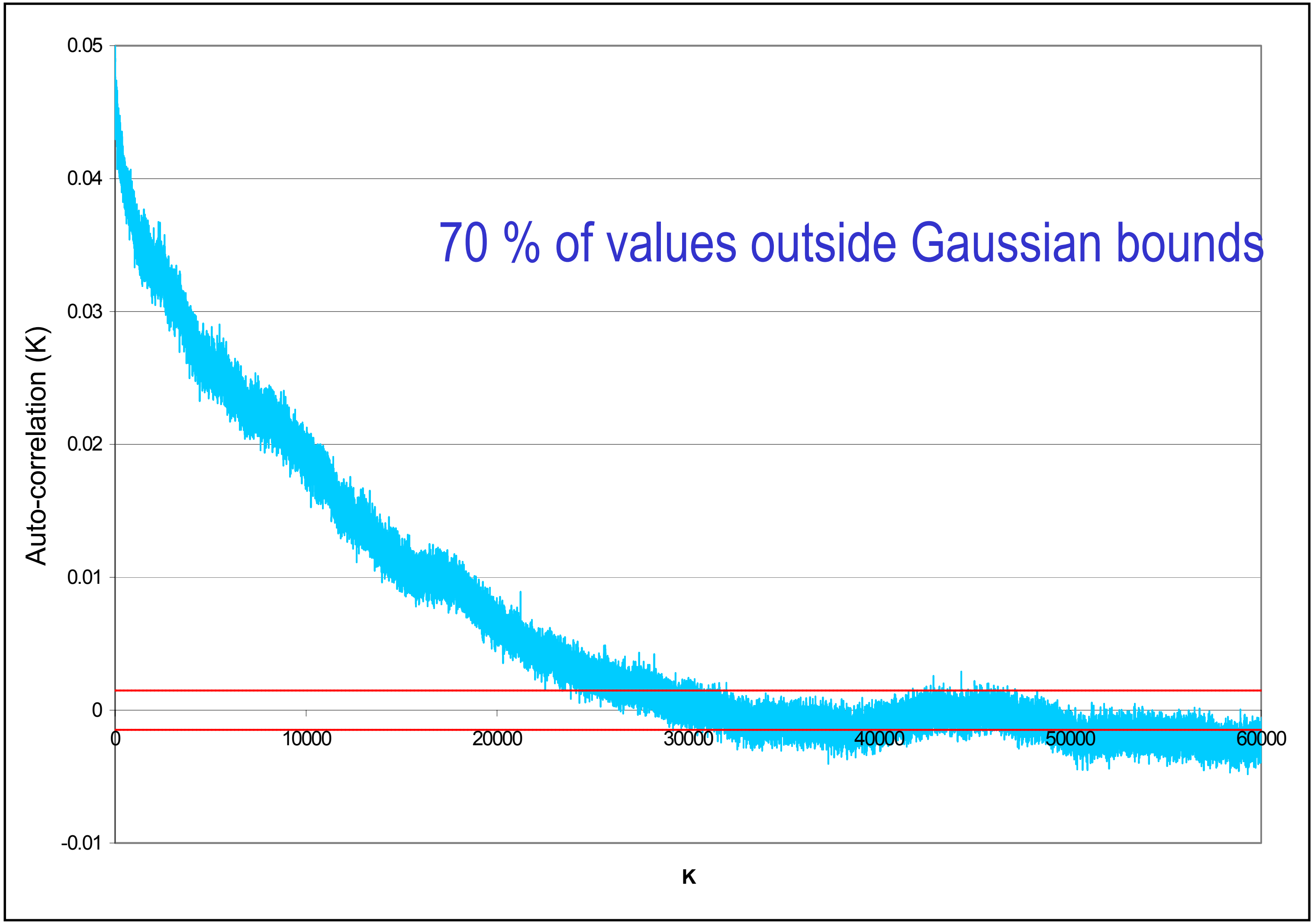
Auto-correlation of flow arrivals



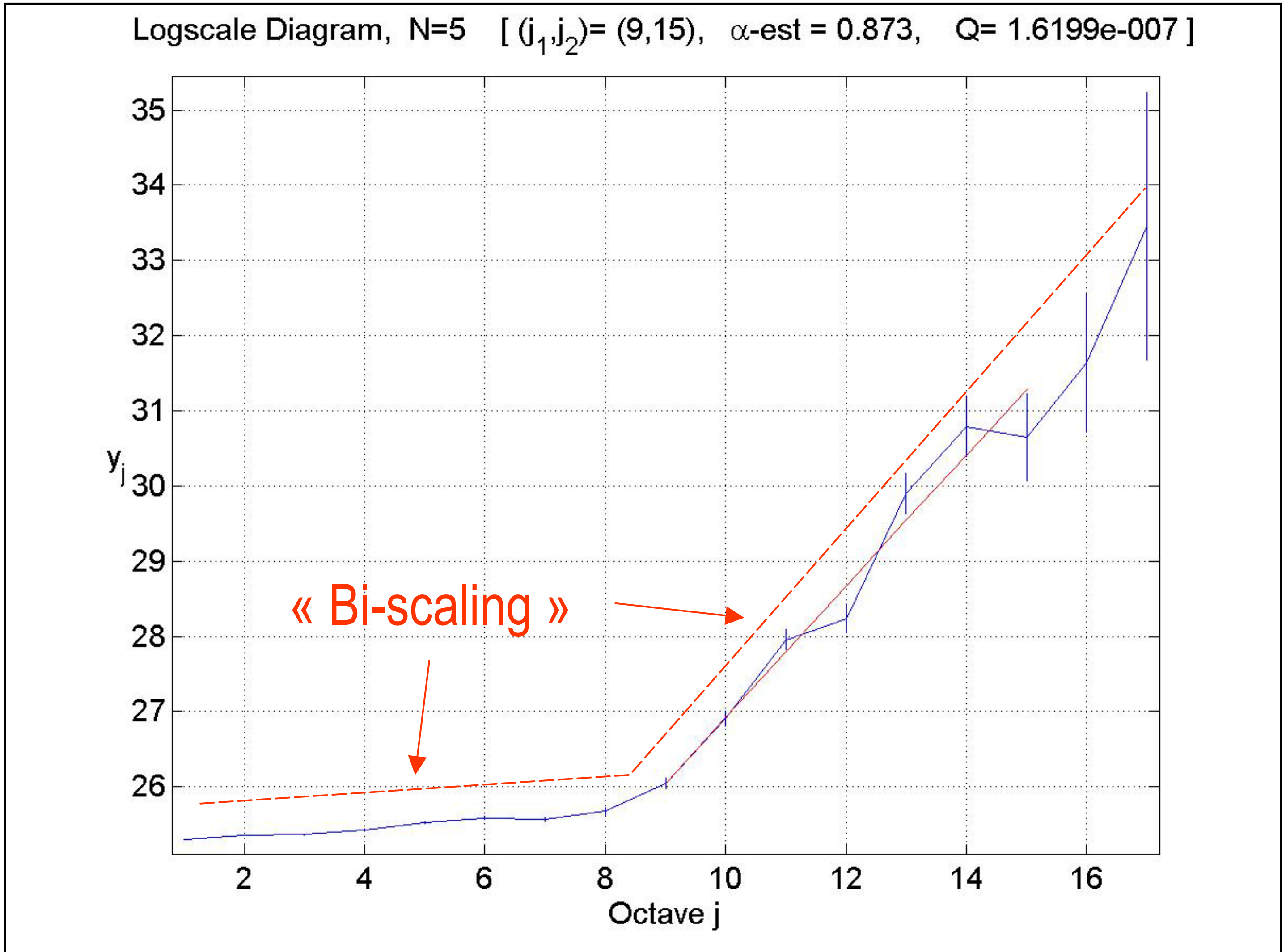
QQ-Plot of packet arrivals



Auto-correlation of packets arrivals

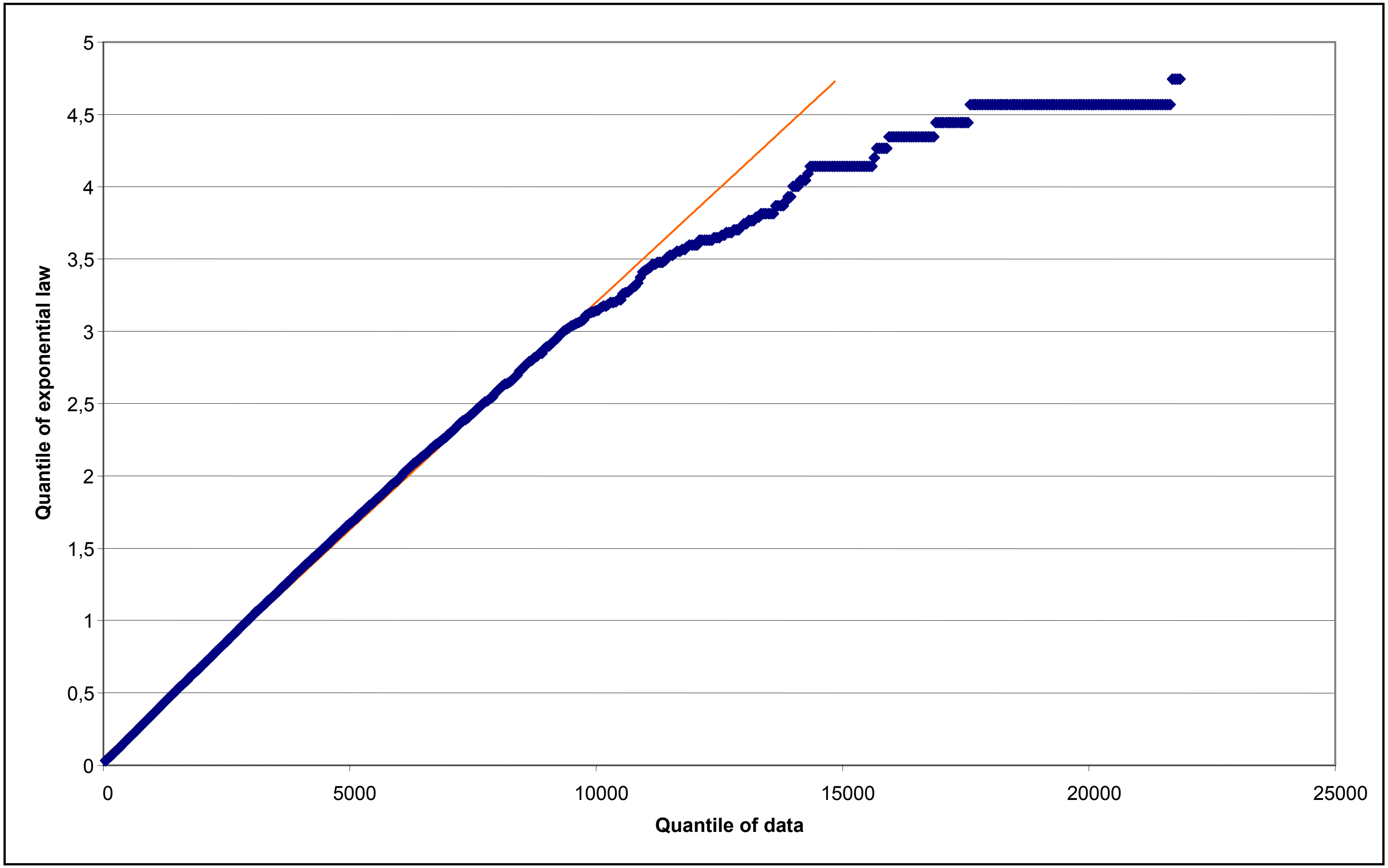


H (LRD) measurements

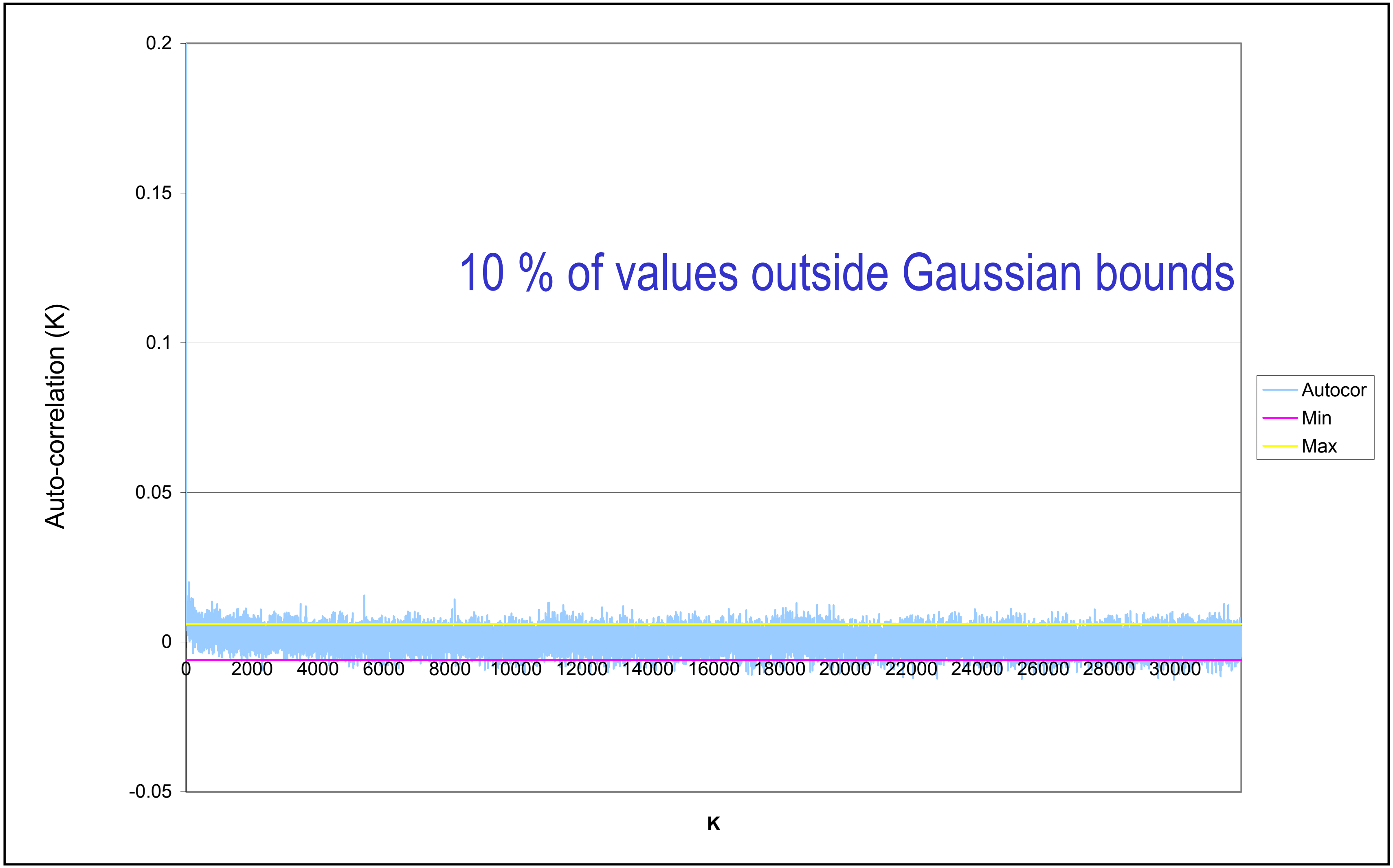


Backbone link traffic analysis
155 Mbps

QQ-Plot of flow arrivals

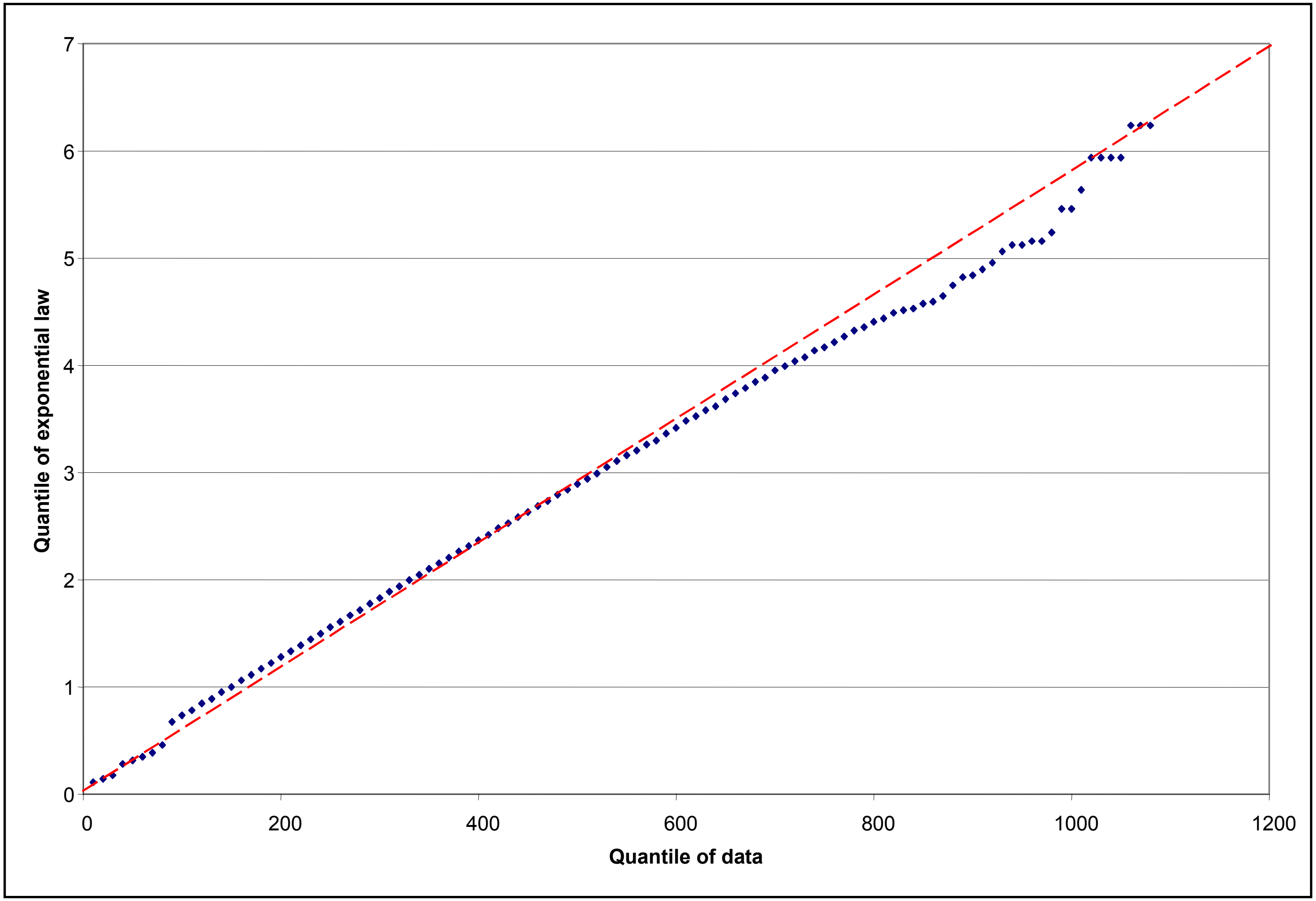


Auto-correlation of flow arrivals

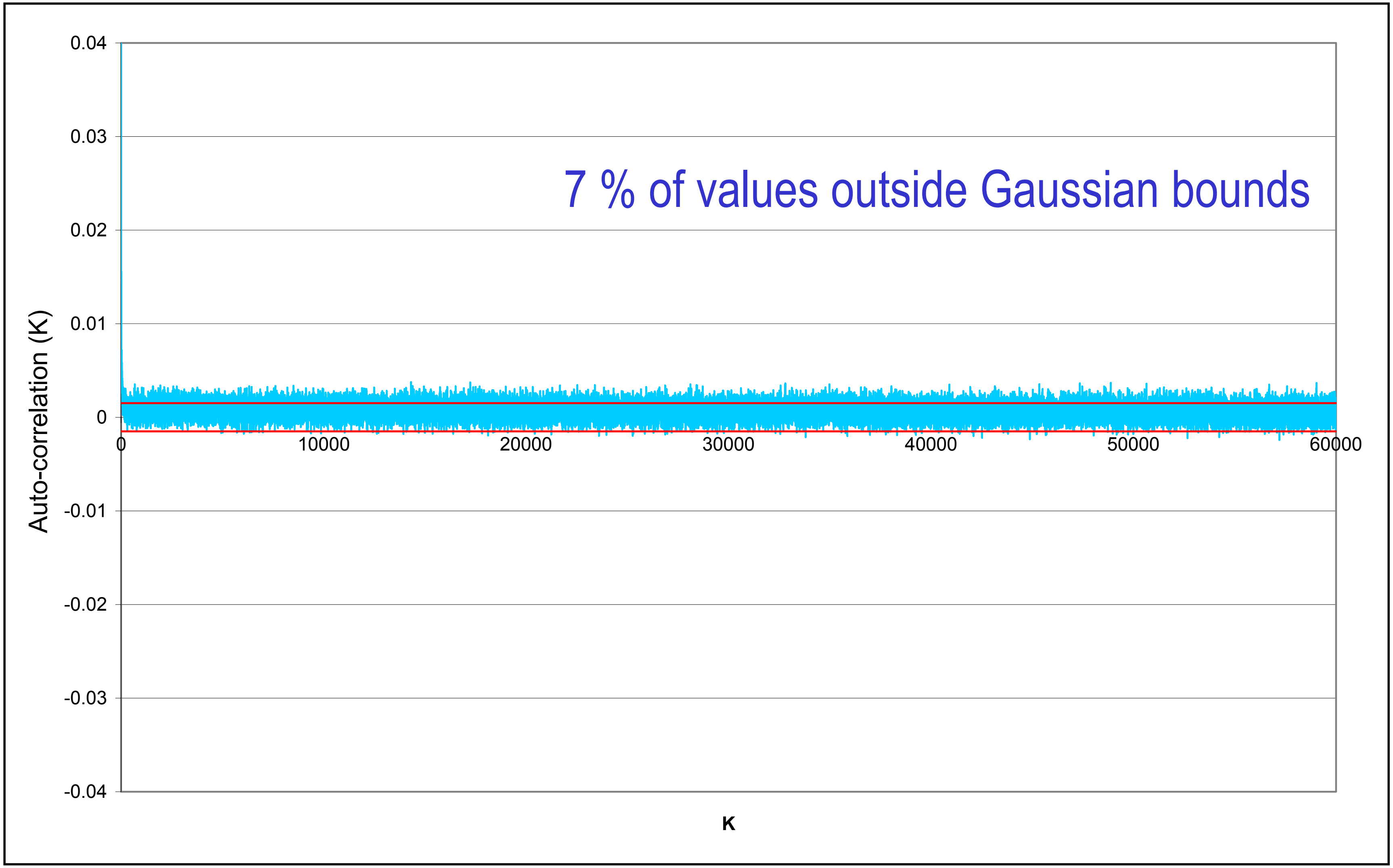


IDMS-PROMS, Coimbra, Portugal, November 26th-29th 2002

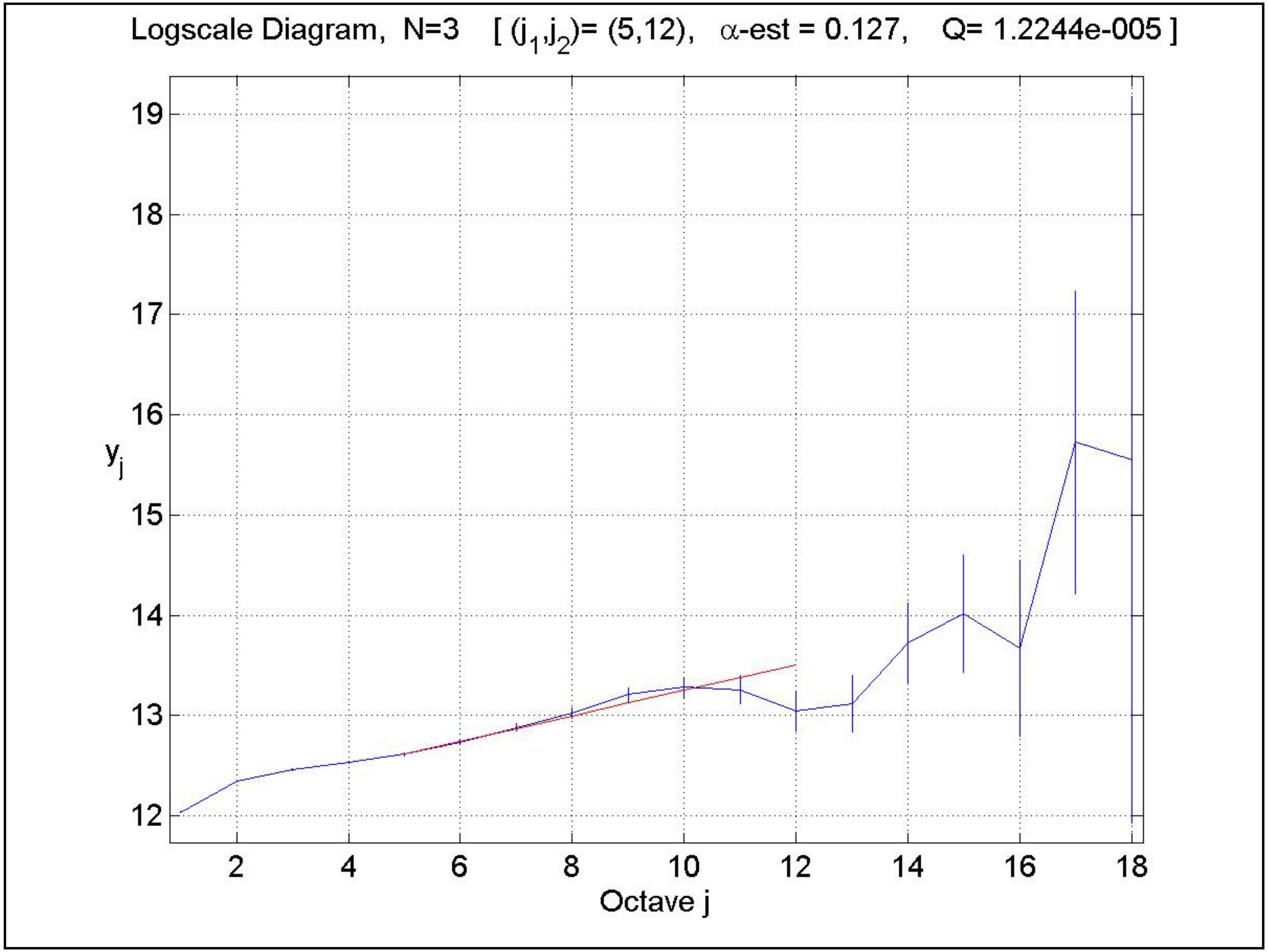
QQ-plot of packet arrivals



Auto-correlation of packet arrivals



H (LRD) measurements



Conclusion on traffic

	Access traffic	Backbone traffic
Hurst (H) parameter	H = 0,915 [0.868, 0.962]	H = 0,561 [0.556, 0.565]

- Access traffic is very complex
- Backbone traffic is smoother
- Networking main issues (QoS, performance decrease,…) mainly appear on edge and / or access links

Backbone traffic modeling

Recalls

- Backbone traffic is quite smooth (maybe close from Poisson ?)
- More than 80% of flows do not enter congestion avoidance
  - Most flows are mice / Few are elephants

- What are the effects of slow-start / congestion avoidance on traffic characteristics ?

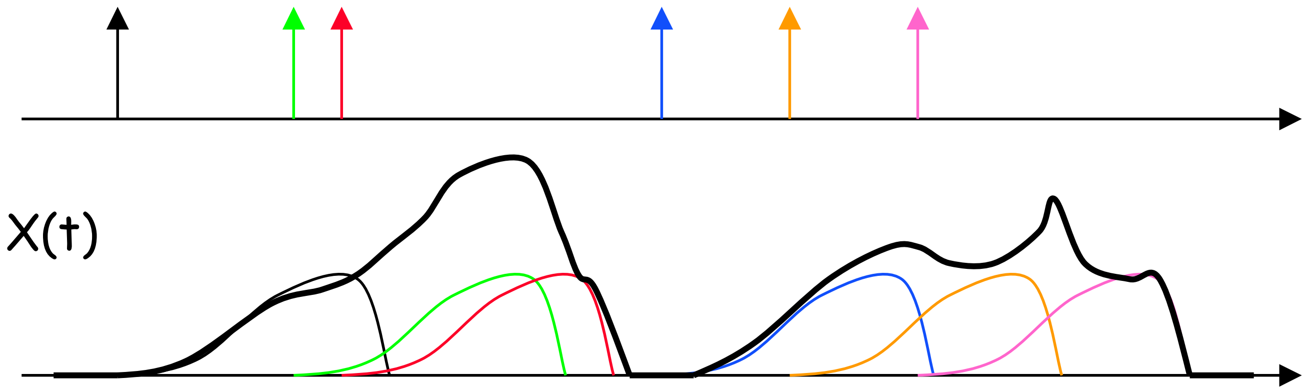
Poisson shot noise model

« A flow based model for Internet backbone traffic »
Barakat, Thiran, Iannaccone, Diot, Owezarski (IMW'2002)

- Evaluate the matching between Sprint's backbone traces and Poisson shot noise model
  - Throughput analysis
  - Variance of Throughput analysis
  - Average and covariance of model traffic and real traffic

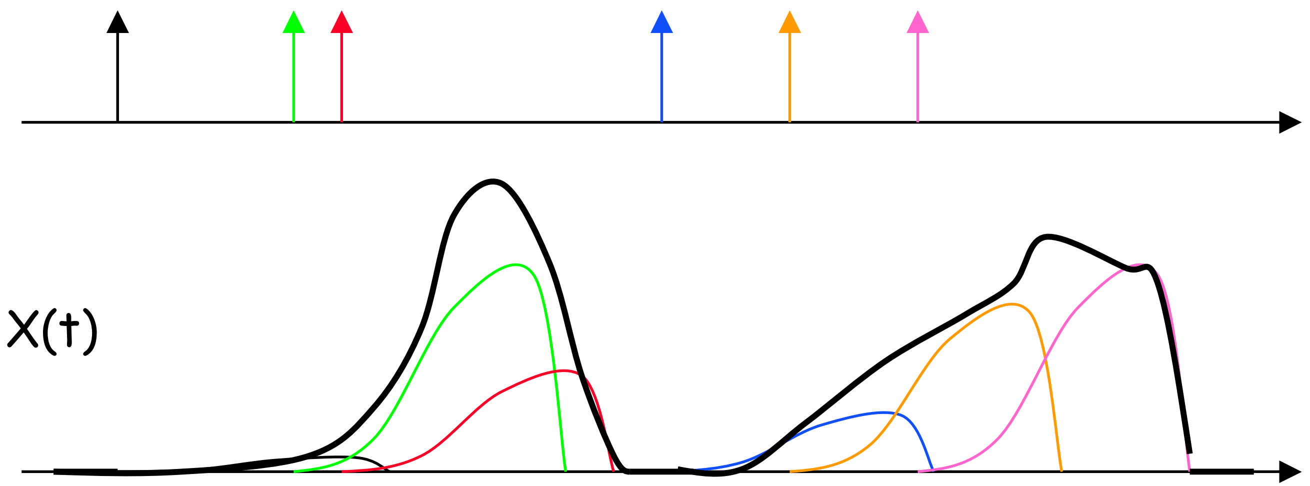
Poisson shot noise

- X(t) is a Poisson shot noise (rate λ, pulse shape g(t)):
  - Starting times of pulses is a Poisson process
  - Pulse has shape g(t), with $\int |g(t)|\, dt < \infty$



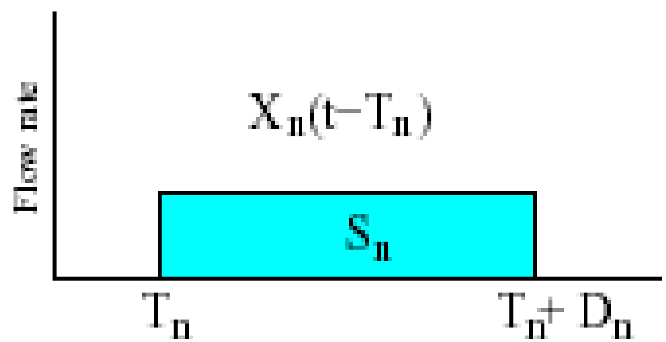
Generalized Poisson shot noise

- X(t) is a Poisson shot noise (rate λ, pulse shape g(t)):
  - Starting times of pulses is a Poisson process
  - Pulse has shape g(t)
  - Amplitude multiplied by random variable *A*



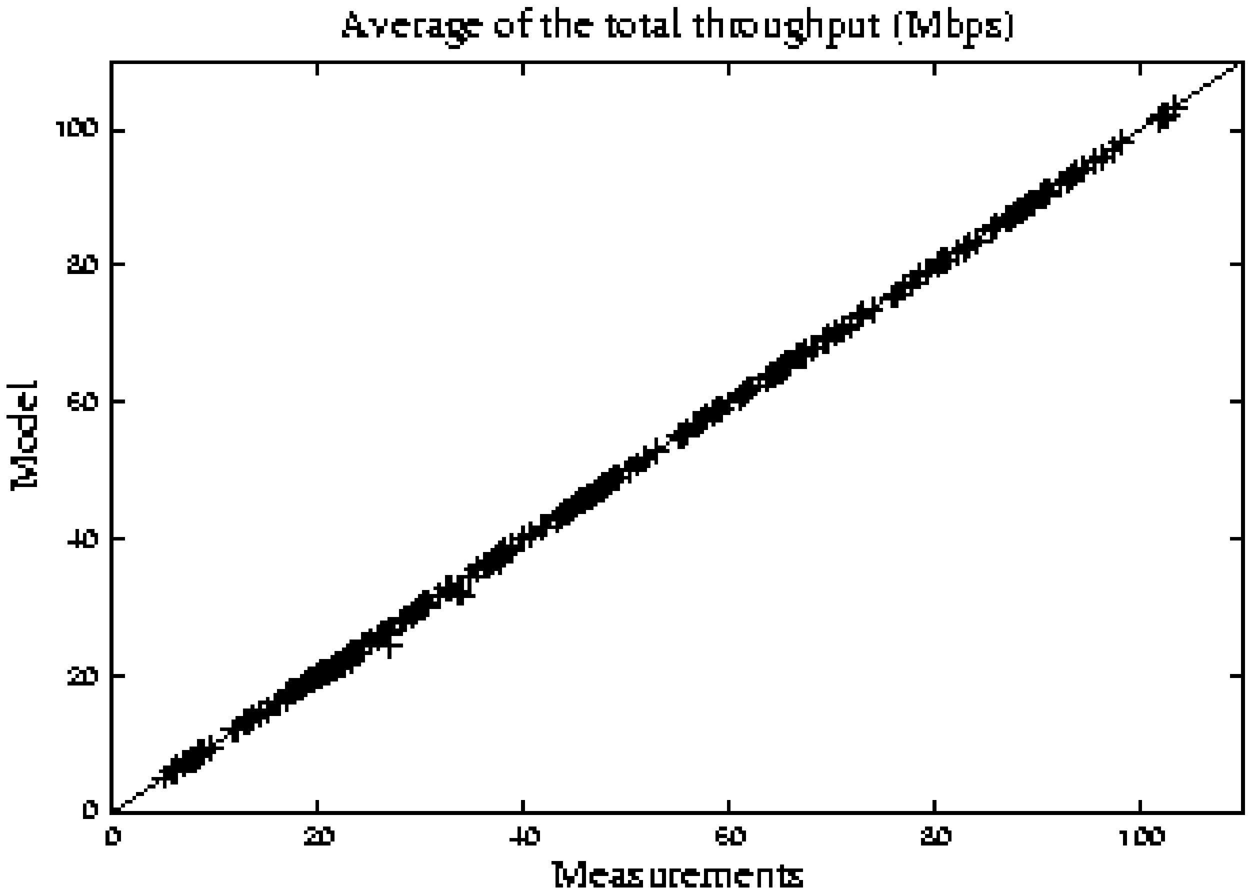
First Model

- → Backbone traffic is the result of aggregating many access links traffics that are bounded (rectangle shape of flow traffics ?)
- → Flows arrival model on access link is not very far from Poisson

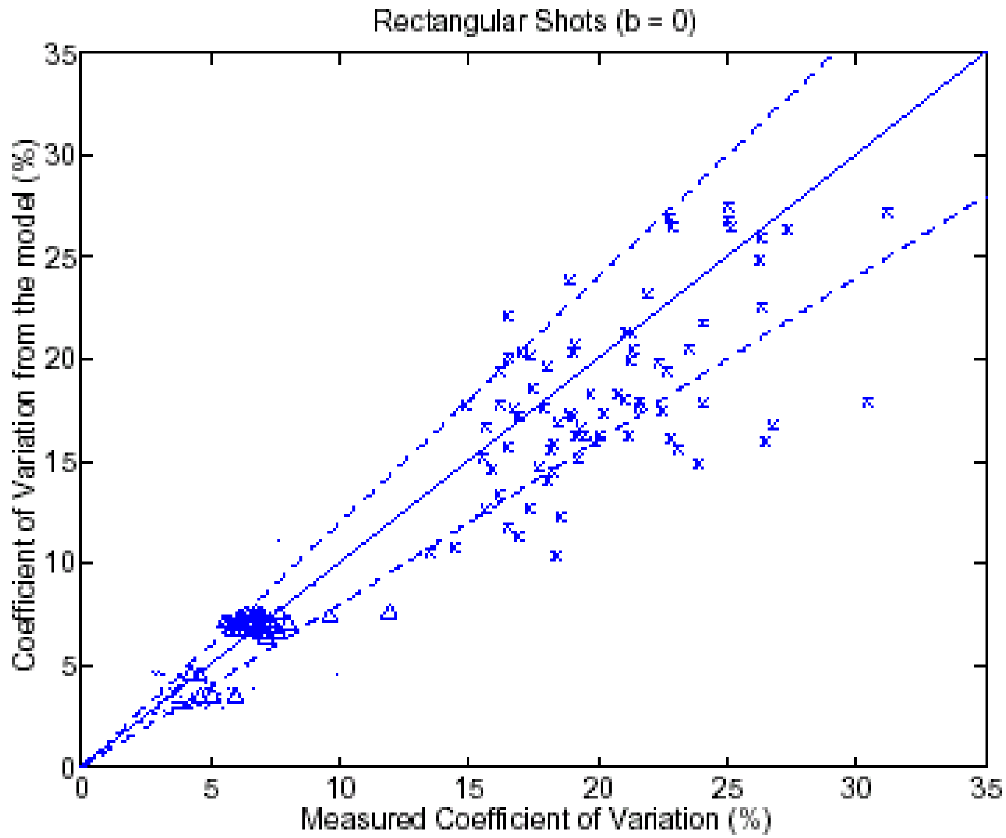


(a) Rectangular shot (b = 0)

Constant approximation (Rectangles)



Constant approximation (2)

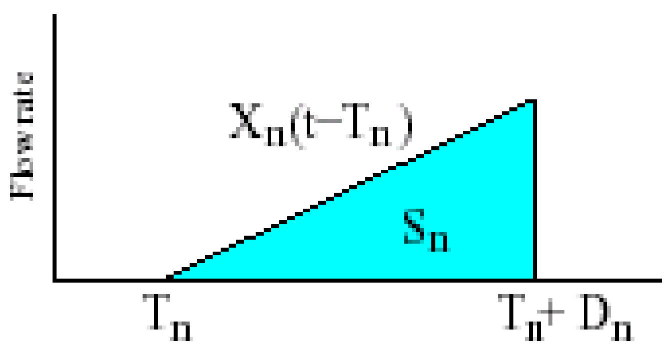


→ Bounds due to access links is not the key parameter for traffic modeling / analysis

Second model

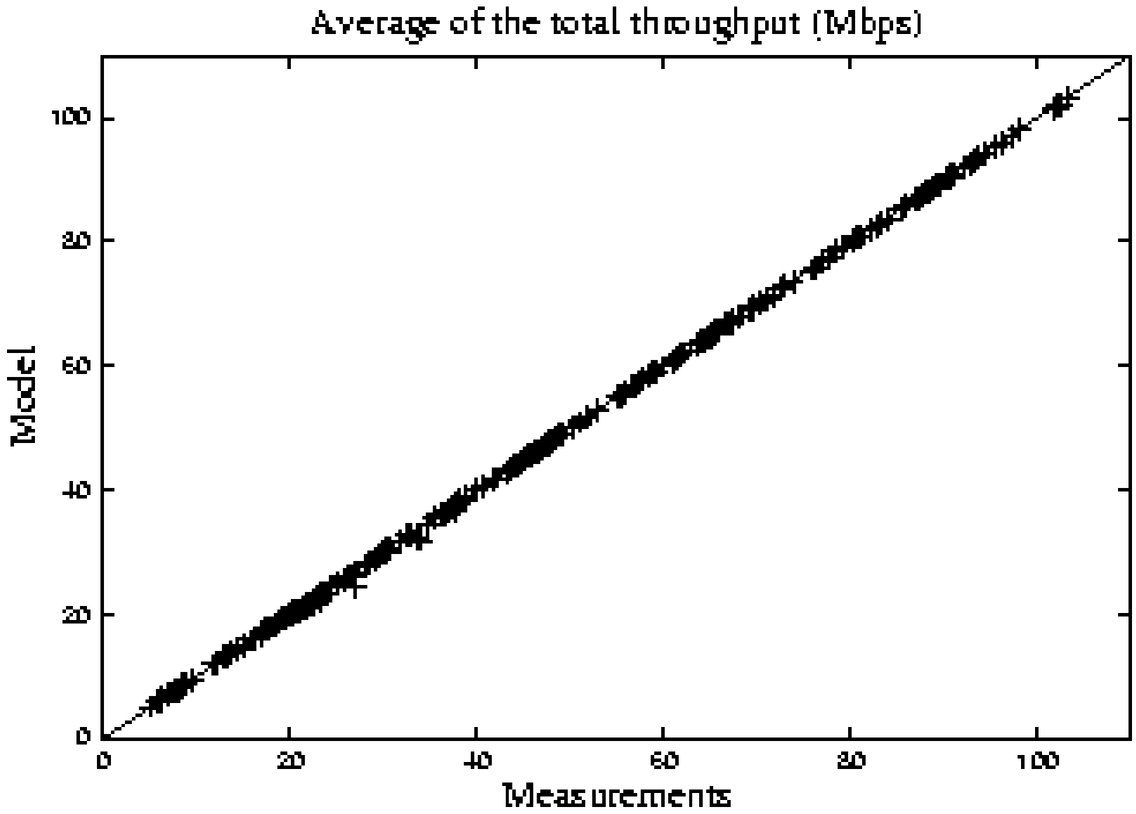
→ Flows arrival model on access link is not that far from Poisson

→ Elephants dominates on the traffic → i.e. congestion avoidance (Triangular flow traffic)

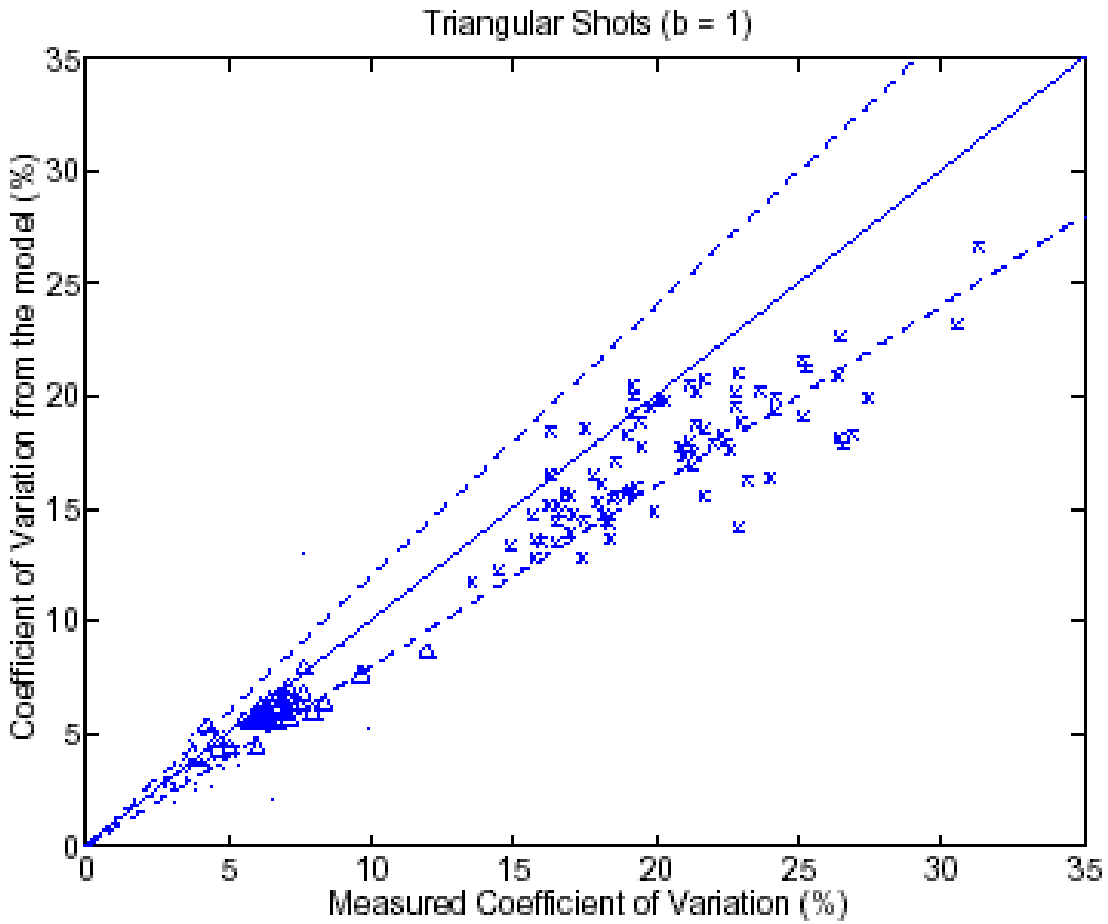


(b) Triangular shot (b = 1)

Linear approximation (triangles)

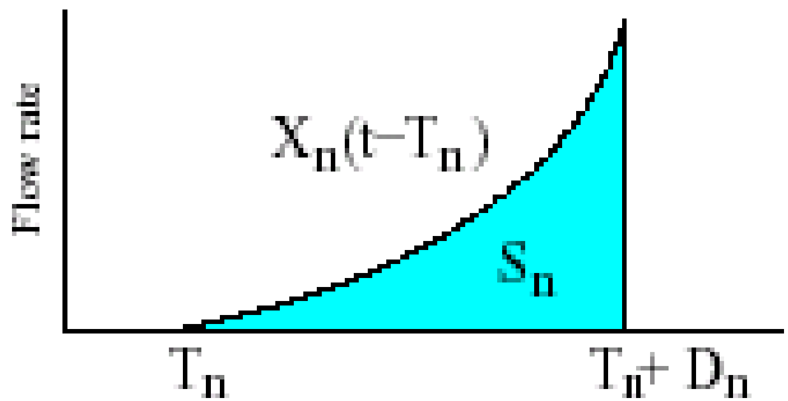


Linear approximation (triangles) (2)



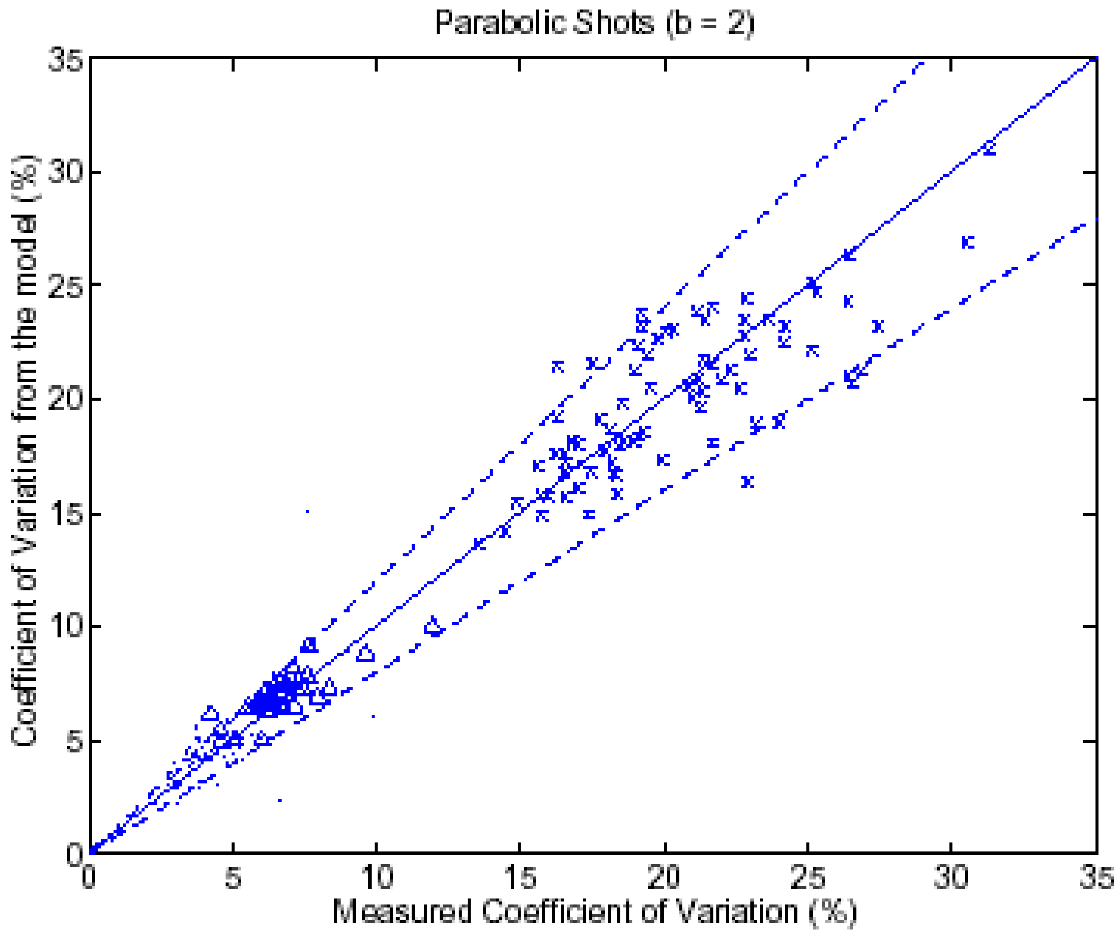
Third model

→Mice dominates on the number of flows → i.e. slow start (parabolic flow traffic)



(d) Superlinear shot (b > 1)

Parabolic approximation



Conclusion on congestion

- Results of linear and quadratic approximation are quite impressive
- TCP congestion control mechanisms are dominant
- Some mismatches
  - Poisson impulsions ?
  - Presence os some elephants

- Individual flow characteristics can be observed in the backbone (after several aggregation steps)

Losses

Loss

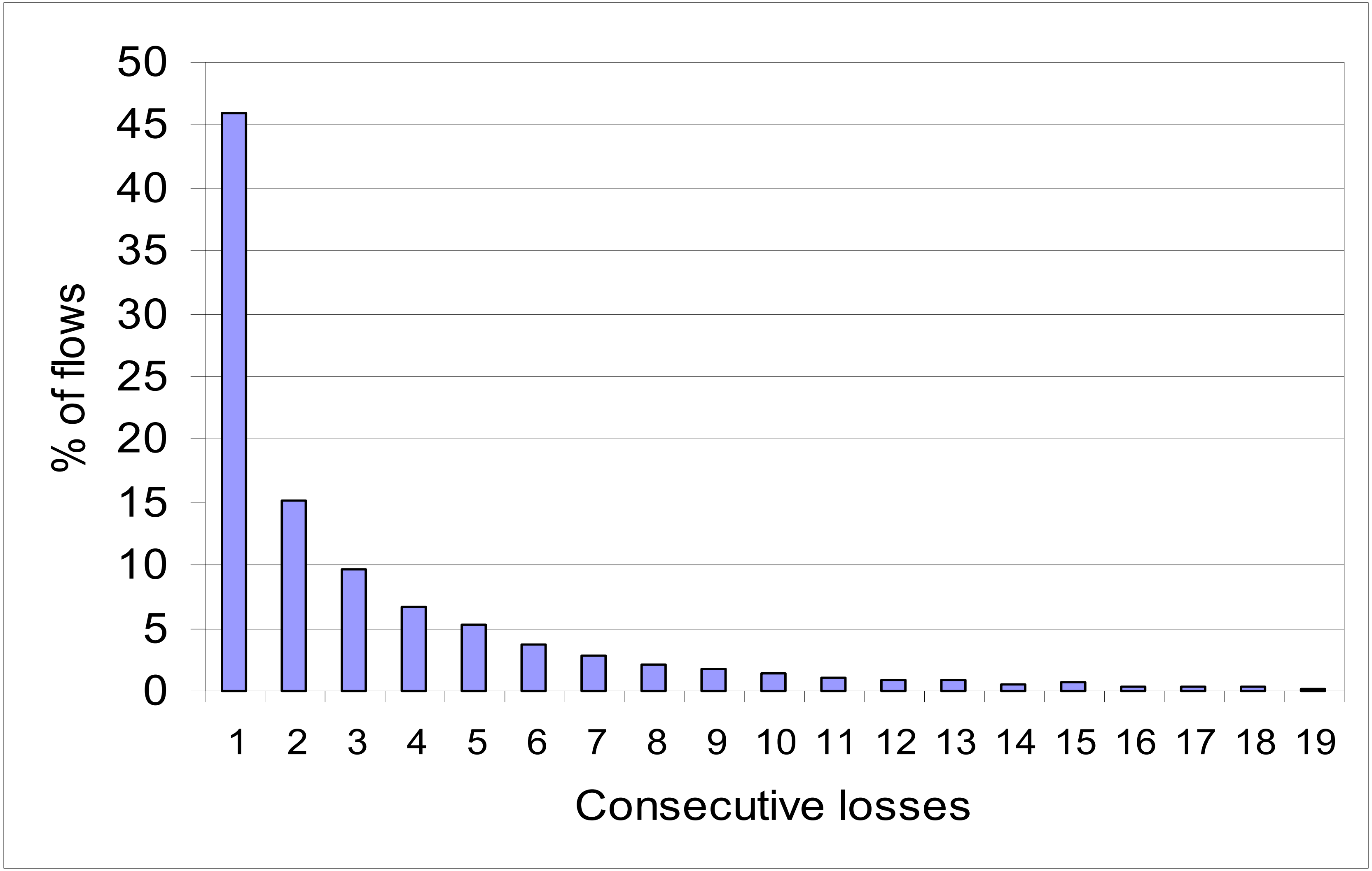
- Losses are defined according to the TCP meaning (end to end)

- Loss packets: 4.48% (between 3.5 and 5.5%)

- Flows experiencing at least 1 loss: 27.1 % (between 10 and 35 %)

➔ To compare to the physical loss ratio ≈ 0

Consecutive losses



Conclusions on loss

- Loss process is not Gilbert
  → not independent

- Loss process is stable

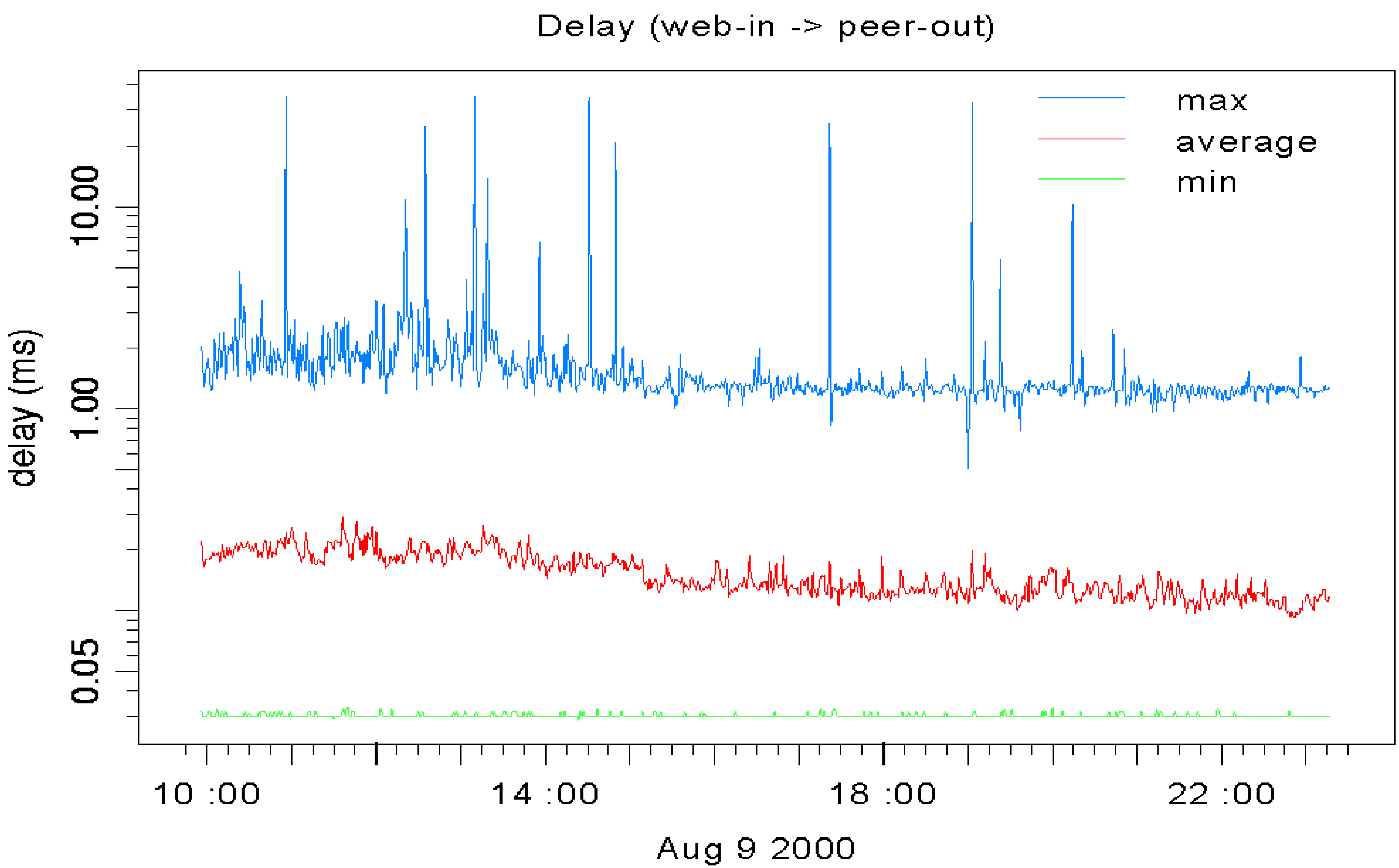
- Losses are dependent (and series of losses exist)
  → TCP loss recovery mechanisms are not suited to the actual model of loss

Delays in routers

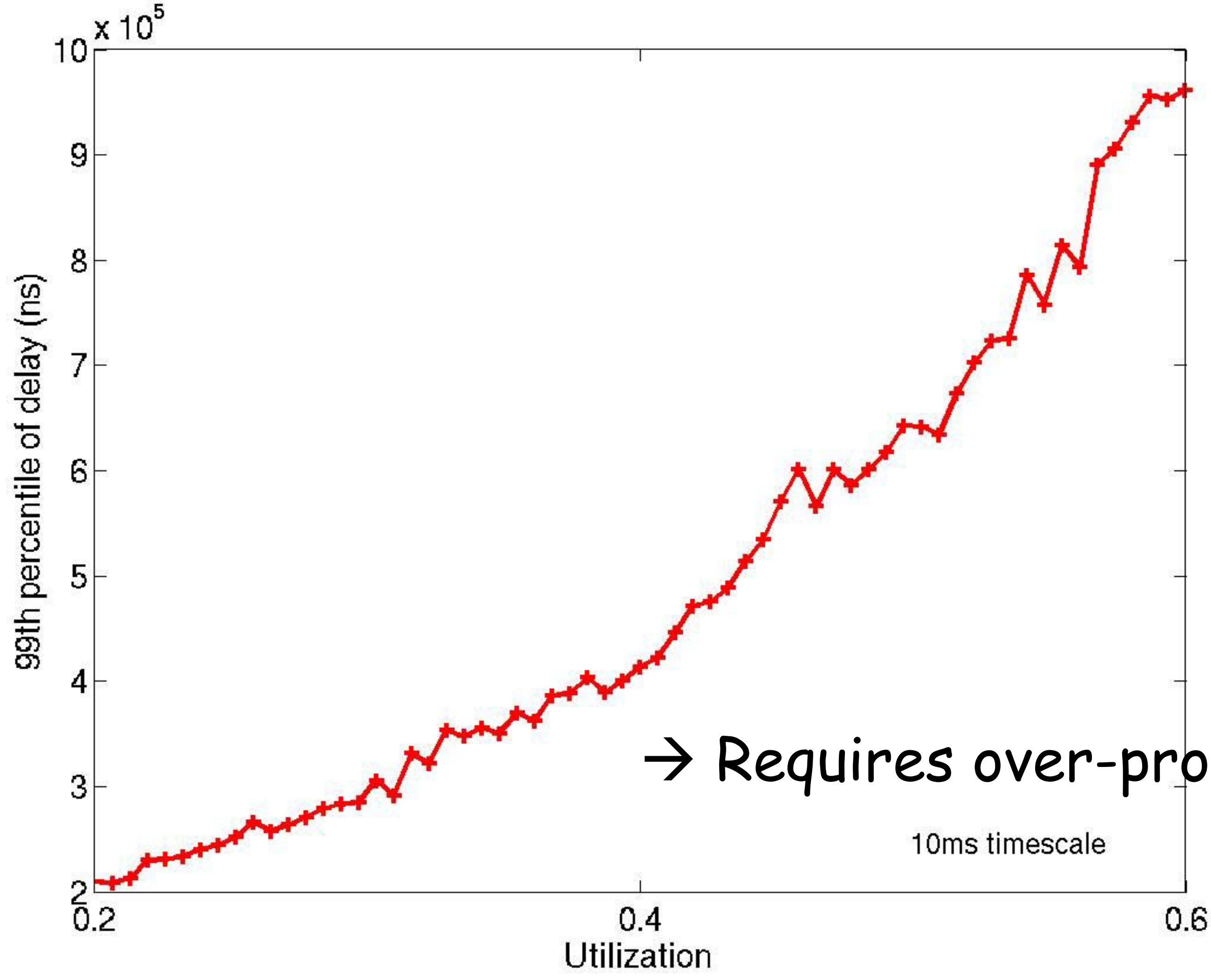
Transit time through a router

- Key metric in network performance
  - Critical to delay-sensitive applications
  - Adds up to end-to-end delay
  - Important in the QoS control
- How to calculate transmission time
  - need to match packets on incoming and outgoing links

Delay vs. time



Delay vs. Link utilization



→ Requires over-provisioning

Traffic matrices

Traffic dynamic

- Where does the traffic come from?
- Between any two POPs:
  - What is the volume of traffic?
  - What are the traffic patterns?
- How to design traffic matrices ?

→ IS-IS is used... Is it a good choice ?

Traffic Dynamic

- Map each packet entering our backbone to its egress POP
- Method :
  - Map each BGP next hop to a POP
  - Extract destination address from each packet
  - Use longest prefix match with (BGP destination, POP) table

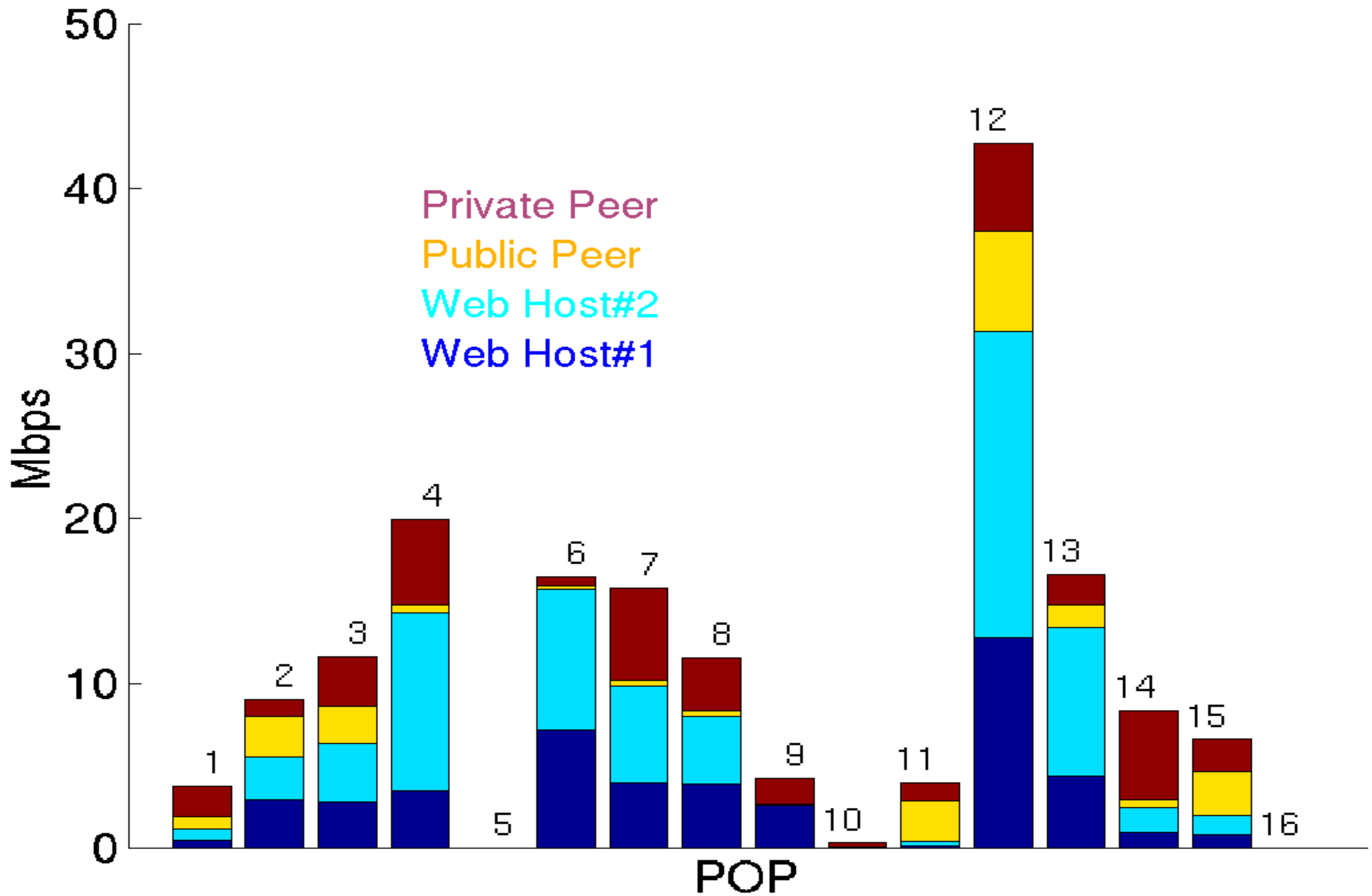
Traffic Matrix

- For each ingress POP :
  - identify traffic to each egress POP
  - further analyze this traffic

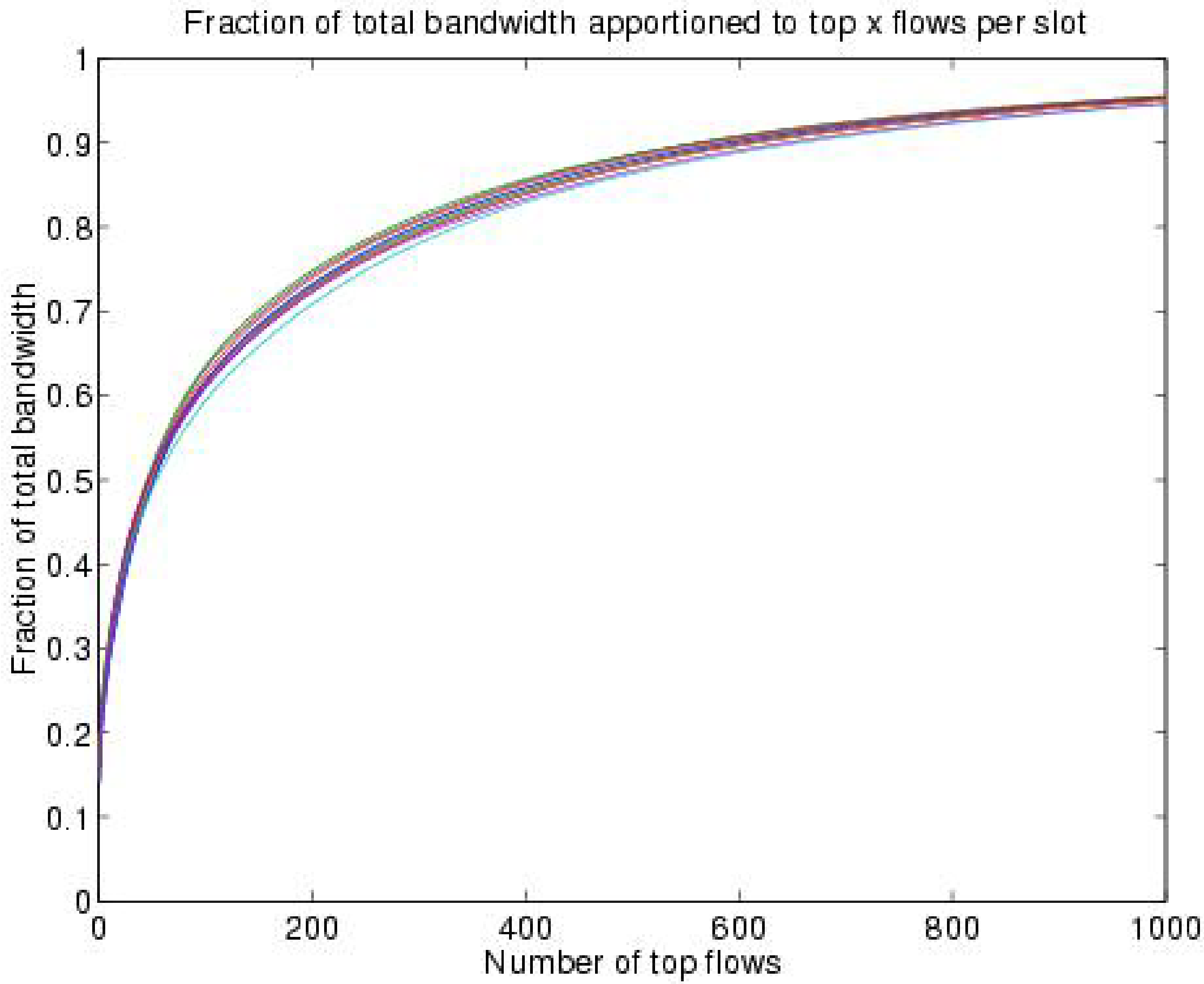
	City A	City B	City C
City A			
City B			
City C			

•Measure traffic over different timescales
•Divide traffic per destination prefix, protocol, etc.

POP-to-POP Traffic Matrix



Routing per destination prefixes



Conclusion on traffic matrices

- Routing matrices are stable
- Elephants and mice
  - Routing tables are closely related to few elephants
- Traffic engineering
  - IS-IS load balancing (based on flows) is sufficient (no OSPF and MPLS)
  - IS-IS load balancing avoids misordered packets what is good for TCP performances
  - Traffic engineering → elephants hunting

    → Maybe Lambda switching is possible ?

Routing table explosion

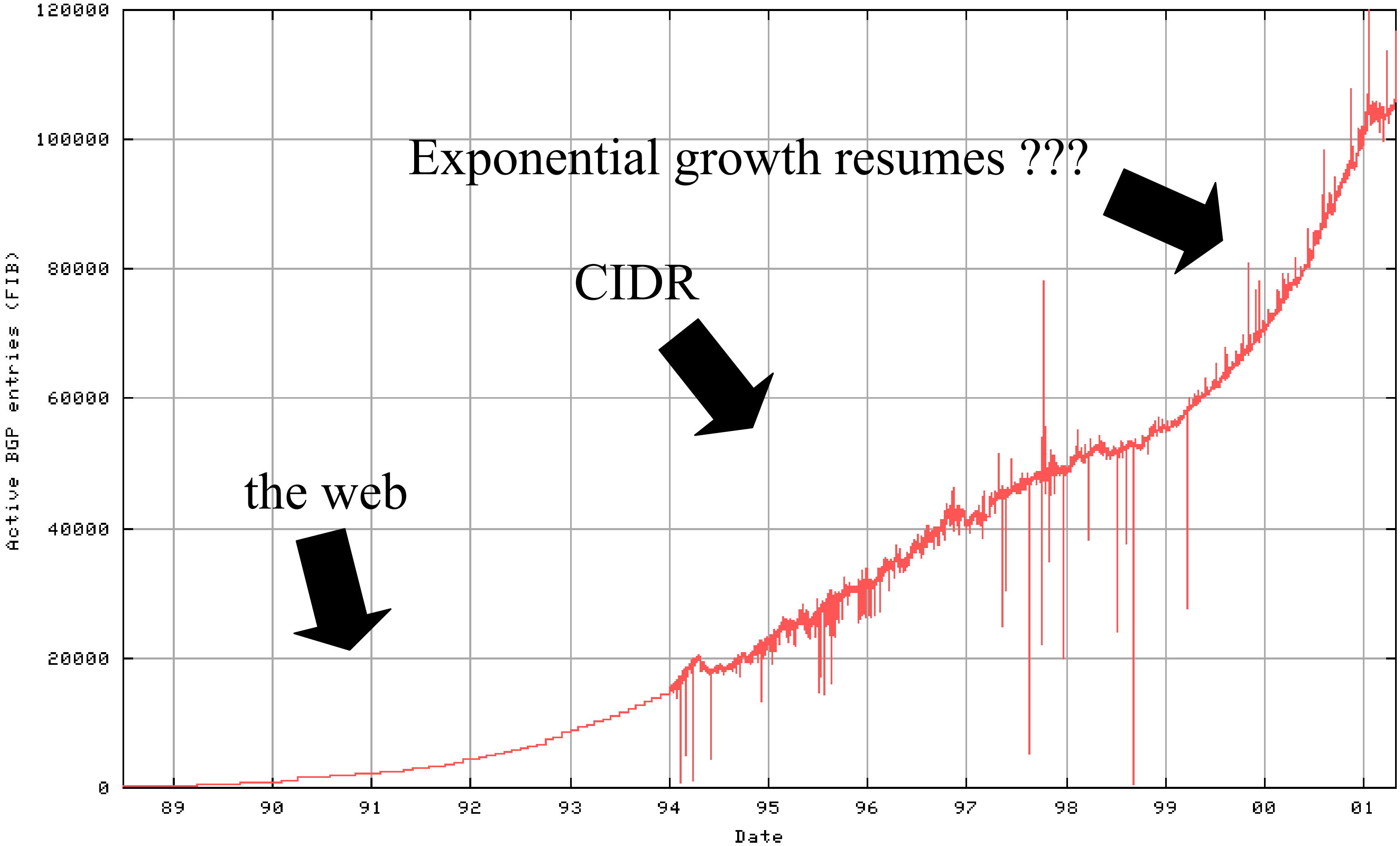
Scalability issues of the Internet

Addressed issues

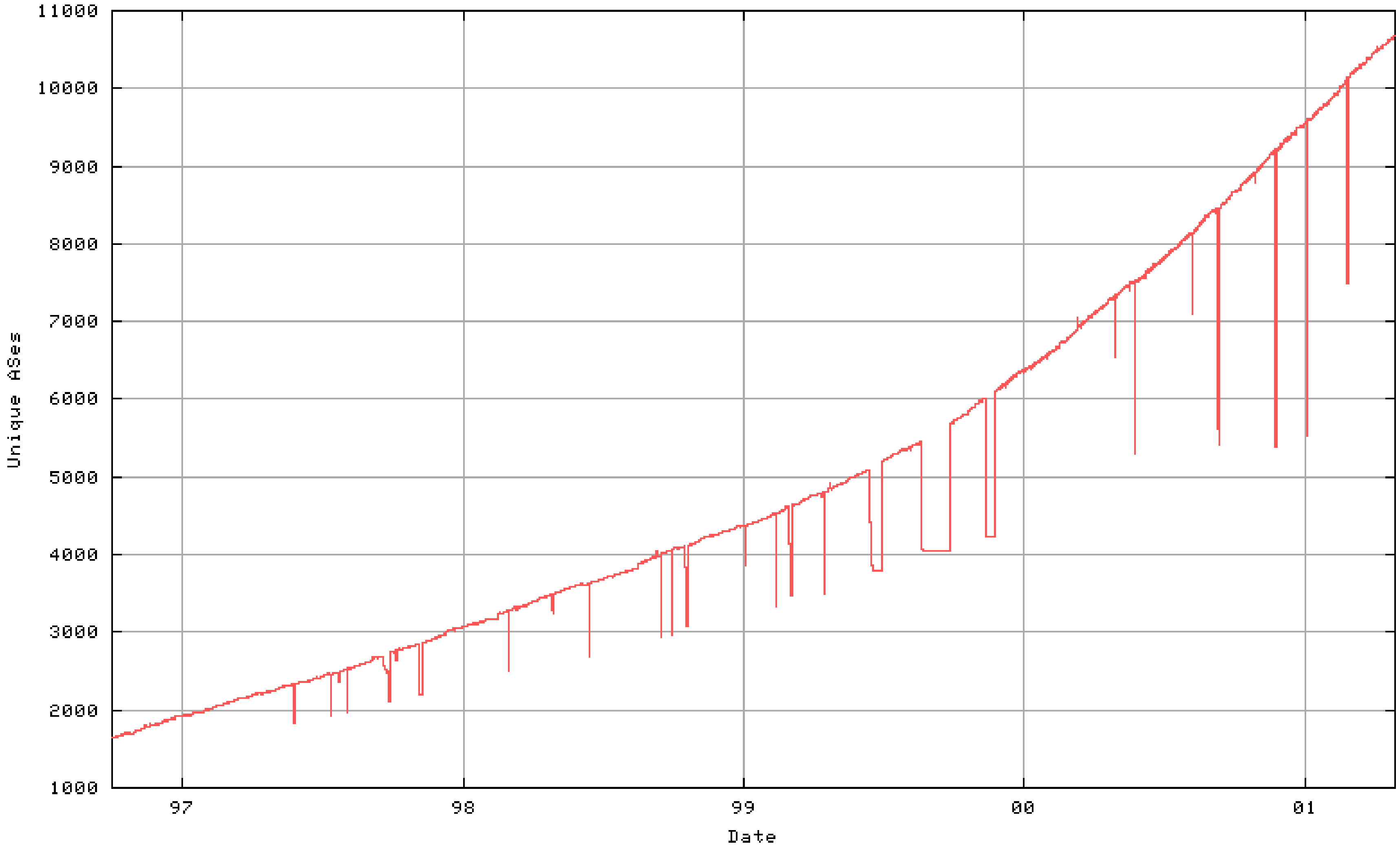
- Scalability issues of the current Internet (scalability becomes more and more an issue)
  - Explosion of routing tables
  - Effects on the Internet QoS

⇒ Routing table size growing from 15,000 to 150,000 entries in average in 6 years

BGP Table Growth (1989-2001)



AS number growth



What helps to reduce routing table size…

- CIDR (also to cope with IP addresses exhaustion)
  - Helps to « fight » routing table size increase related to the web explosion → address space aggregation

- Aggregate as much as possible

  ⇒ Black holes (in address space)

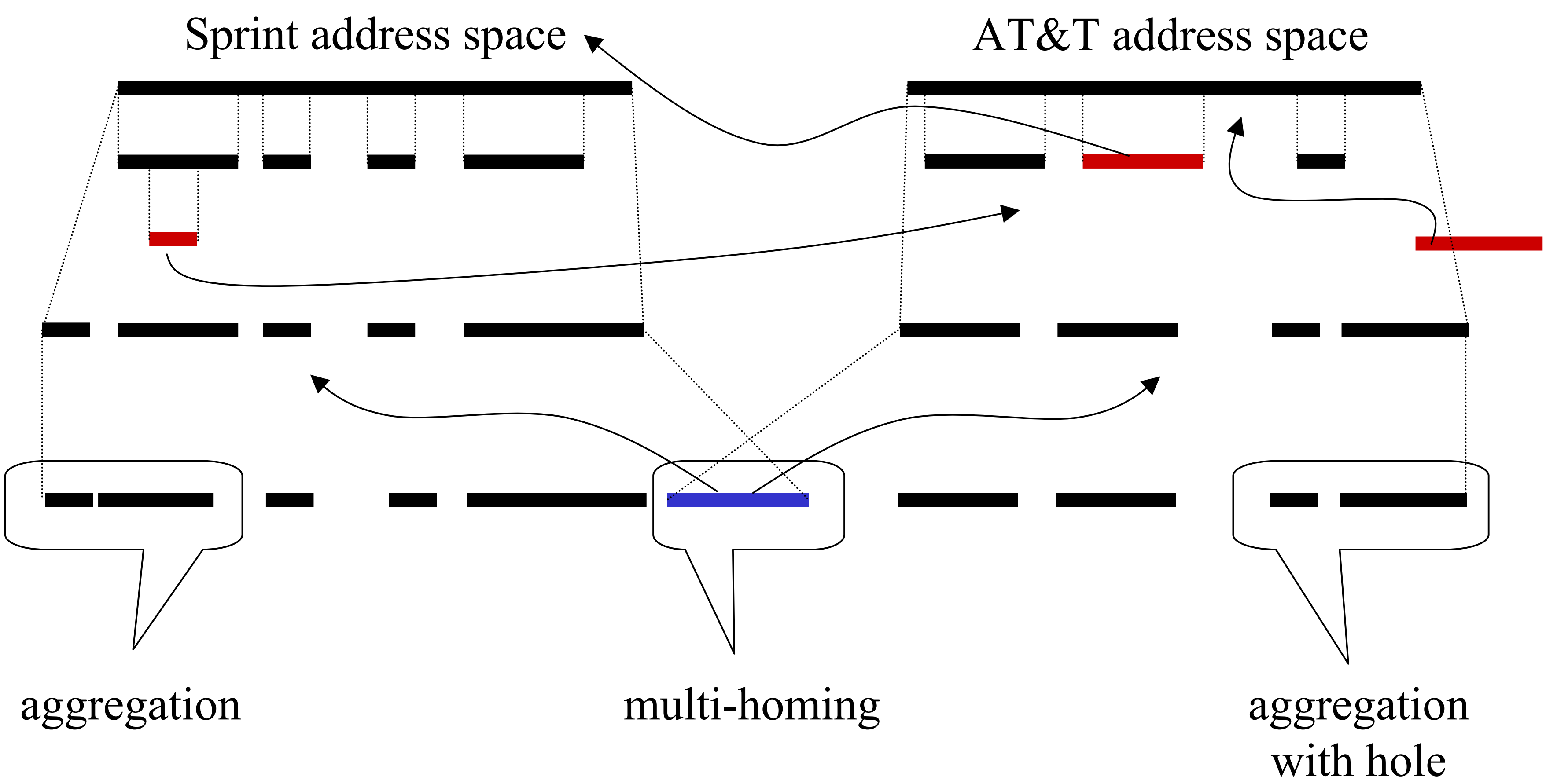
What improves scalability…

- Routers performances (new optical components, faster memories, new switch fabrics…)
- New flow based mechanisms in routers
  - CISCO's Netflow, CEF (CISCO Express Forwarding)
  - JUNIPER (Internet 2 processor)

What makes routing table size increase…

- NAT → makes /31 prefixes growing very fast
- Multi-homing
  → Makes hierarchical distribution of address spaces difficult to set-up
- Constancy of users that often change from Tier 1/ Tier 2/ ISP with their own address space
  → hierarchical address spaces difficult to change

IP addresses and BGP / black holes



Analyses of prefixes in routing table size

- Multi-homing : 20 – 30 %
- Failure to aggregate: 15 – 20 %
- Load balancing: 20 – 25 %

- Address fragmentation: > 75 %

Summary on QoS

Tier 1 concerns with QoS

- View of a tier 1 is intra-domain related
- Over-provisioning of core network and public peerings
  - To fight CDN
  - Problems with private peerings
- Increase scalability
  - Reduce routing table size (aggregation, black holes)
  - Trade off between delay / reliability

Tier 1 concerns with QoS (cnd)

- Traffic engineering
  - Tomography based (traffic matrices)
  - Elephants Hunting (→ Lambda switching ? / All optical solution ?)
  - IS-IS load balancing
    - Based on flows
    - No misordering to optimize TCP performance
- Speed of light dominates on delay
- Maybe, VPN service on the edge router for providing CoS

LAAS' concerns with QoS

- Over-provisioning (based only on the knowledge of average traffic throughput computed by tools using SNMP (measurement scale = 5 s)) is not a solution to the traffic LRD issue
  - It is not an optimized way for managing resources (→ resources waste)

→ Reduce LRD

→ No over-provisioning should be required

LAAS' concerns with QoS (2)

- TCP control loop / congestion control mechanisms have a strong impact on LRD and self-similarity
  - Over-provisioning is not realistic / not enough !

→ TCP has to be modified to limit LRD on the traffic (and loss) due to its mechanisms

→ Routers scheduling and discarding strategies have to limit LRD due to queuing (enhanced RED if it can reduce TCP loop synchronization)

METROPOLIS
(supported by RNRT)

Partners

- LIP6
- LAAS
- FT R&D
- GET
- INRIA Rocquencourt
- EURECOM
- RENATER

Objectives

- Defining a monitoring methodology
- Combining active and passive measurements
  - Active: IPANEMA, RIPE, QoSMOS
  - Passive: DAG

- A full set of networks
  - VTHD (high speed experimental network)
  - Renater (public operational network)
  - ADSL (private operational network)

Addressed issues

- Empiric and stochastic modeling (and more?)
- Provisioning and SLAs
- Classification
- Traffic, network and protocol analysis
- Sampling
- Pricing and charging

IP monitoring related topics

- QoS and performance optimization
- Realistic simulation system (replay of traces)
- Global IDS
- Emulation platform of the Internet
  - Multi-domains QoS
  - Network security
- …

More information about METROPOLIS

http://www-rp.lip6.fr/metrologie

http://www.laas.fr/~owe/METROPOLIS/metropolis.html