Monitoring for next generation Internet

Philippe Owezarski

LAAS-CNRS

owe@laas.fr

Disclaimer

This keynote contains the description of:

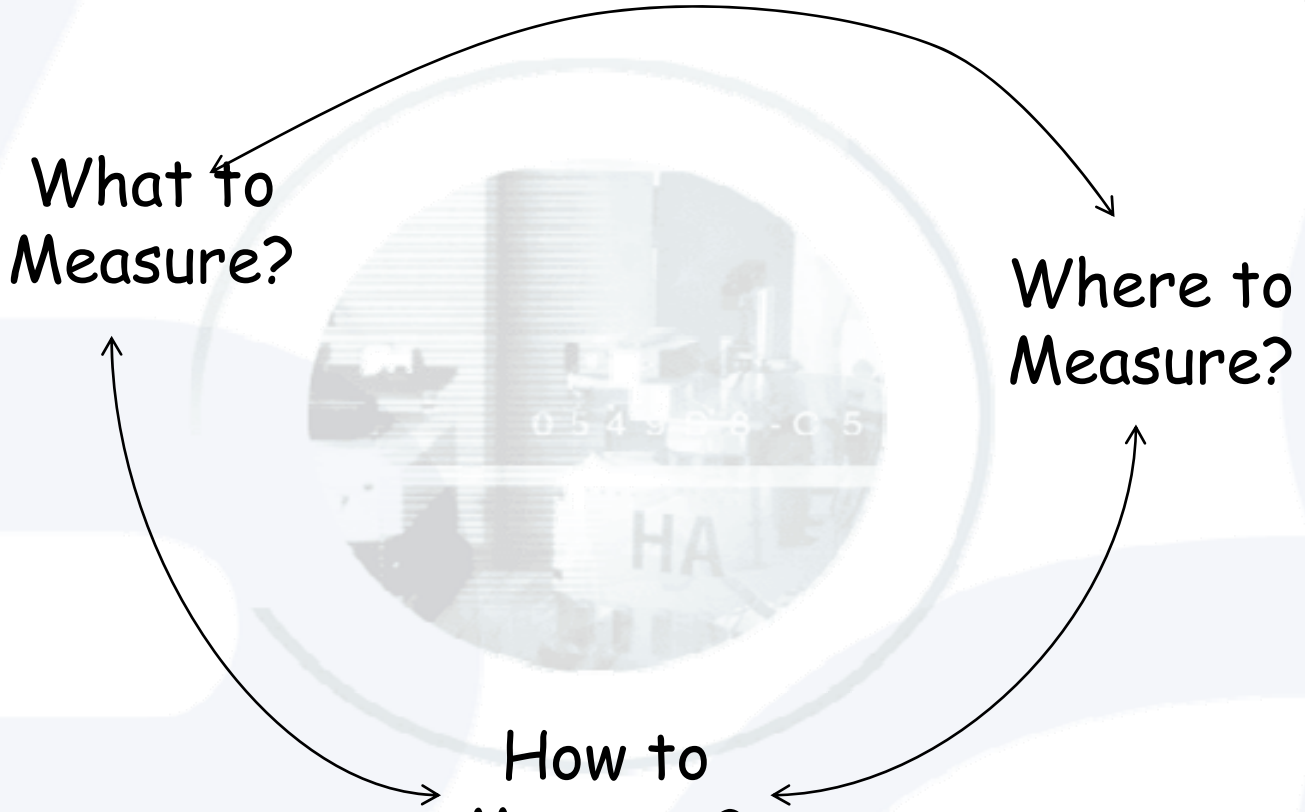
- Problems - Problems - Problems - Problems
- No solution

→ Just emphasizes research topics

Content

- Introduction on NGI and its monitoring
- Objectives and requirements of a global monitoring system
- Active measurements
- Passive measurements
- Analysis of traffic traces and network performance time series and its impact on monitoring
- Discussion

Questions to address



What
How
Where

impacted by

- Technical constraints
- Analysis results
- Domains organization

Next Generation Internet

- Multiple services with various QoS
- Wireless
- Multiple domains / technologies / management policies / …
- High speed
- Secure
- Autonomic (self-*)
- Seamless services

The old Internet engineering process

- Increasing complexity of the Internet
- No more control of the global Internet

- « Defeat » statement (For instance on QoS aspects)
  - Bad knowledge of the traffic

➔ « real » traffic exists and is full of information

Monitoring applications

- Traffic engineering
- Anomaly / intrusion detection
- Congestion control
- Self-management
- Charging
- And many others…

→ different requirements :
delays, latency, granularity (time scale), …

A global monitoring system

→ Because connections are from e2e !!!

Requirements:

- Scalable
- Works in multi domains / technologies environment
- Fast
  - Low delays / latency (according to ISP or user needs)
  - High speed
- transparent

Measurements: technical issues (1)

The network is very very large

- Many equipments are required
- Only a partial vew for transcontinental connections
- Night and day beahvior of the traffic (loaded/ congested points move in the network)

Measurements: technical issues (2)

Delay measurements

- All clocks of all measurement machines need to be synchronized with an accuracy of less than few μs

Measurements: technical issues (3)

How to inform all measuring equipments/agents about measurements made in other places ?

- Pb 1 : large volume of data
- Pb 2 : it is when the network is experiencing performance decreases (e.g. because of congestion) that we most need measurement information

  →but this information is lost or delayed

Measurements: juridic issues

Privacy (CNIL in France)

- 1 IP address identifies one person
  - Need to anonymize IP addresses
- We must not look at payloads

What to measure?

- Throughput
- Loss rate
- Temporal constraints
  - Delay
  - Jitter

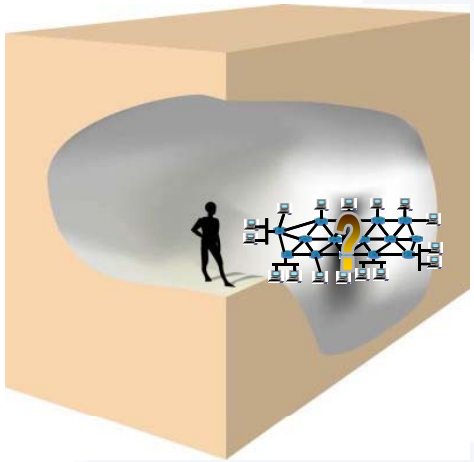
+ many **estimated** parameters from these basic ones

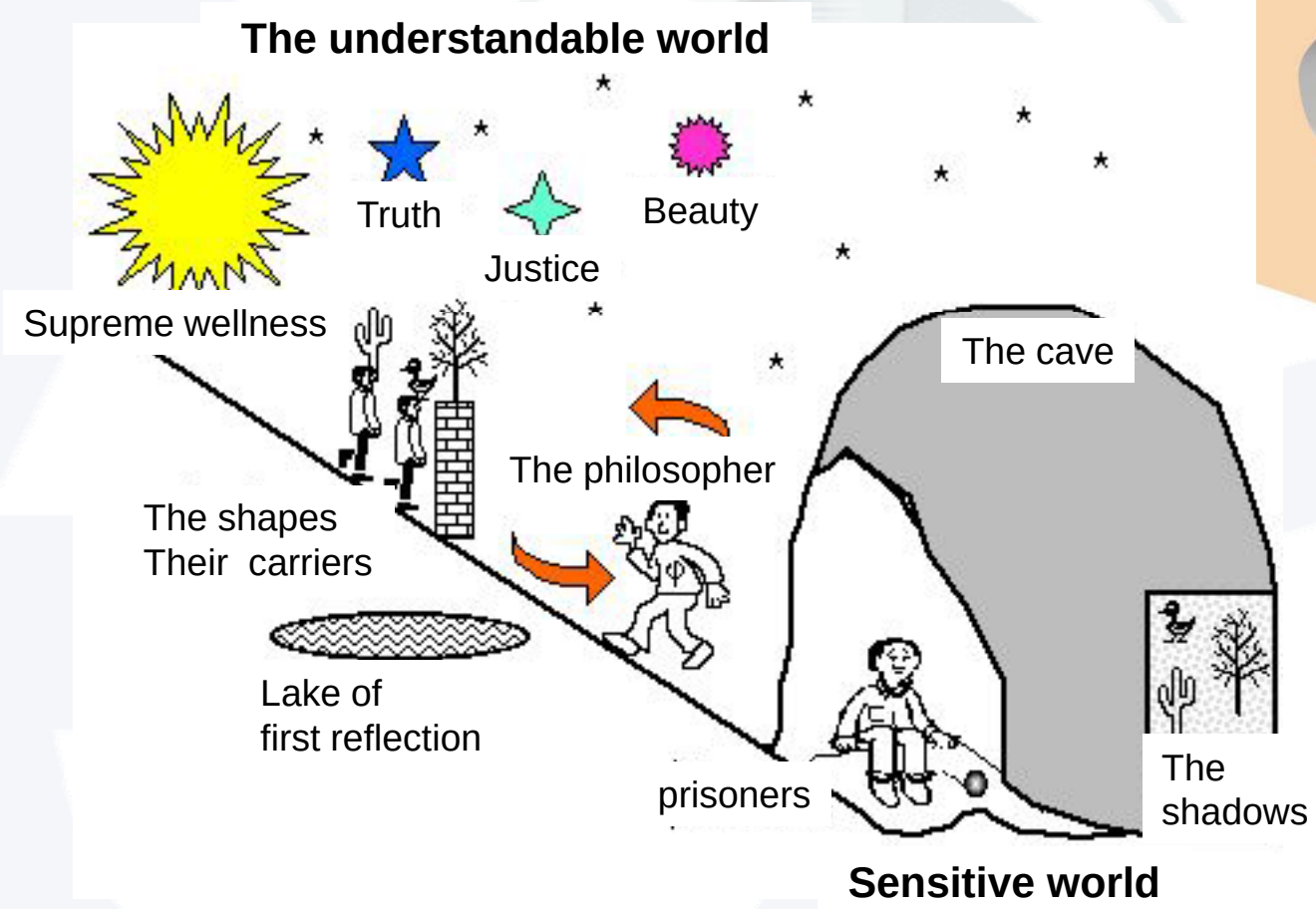
With different points of view:

- User vs. Carriers/ISP
- Point-to-point vs. End-to-end

Parameters estimations: A Platonian issue

« Allegory of the cave »





How to measure?

- Active vs. passive measurements
  - See further…
- Hardware vs. software
  - Performances
  - Costs
  - Accuracy
- Depends on traffic and network performances analysis results
  (for instance sampling)

Where to measure?

Routers - Gateways proxies

links

OK, if monitoring/measurements do not decrease forwarding performances

Transparent
BUT
Utility of performing monitoring / measurements in a passive component?

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

Active measurements (2)

- 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
  - It is stupid to use packets to measure the state of a packet network
    - Does the French « Bison Futé » (smart bull) use cars to estimate the traffic on French roads and highways ?

Some active measurement tools

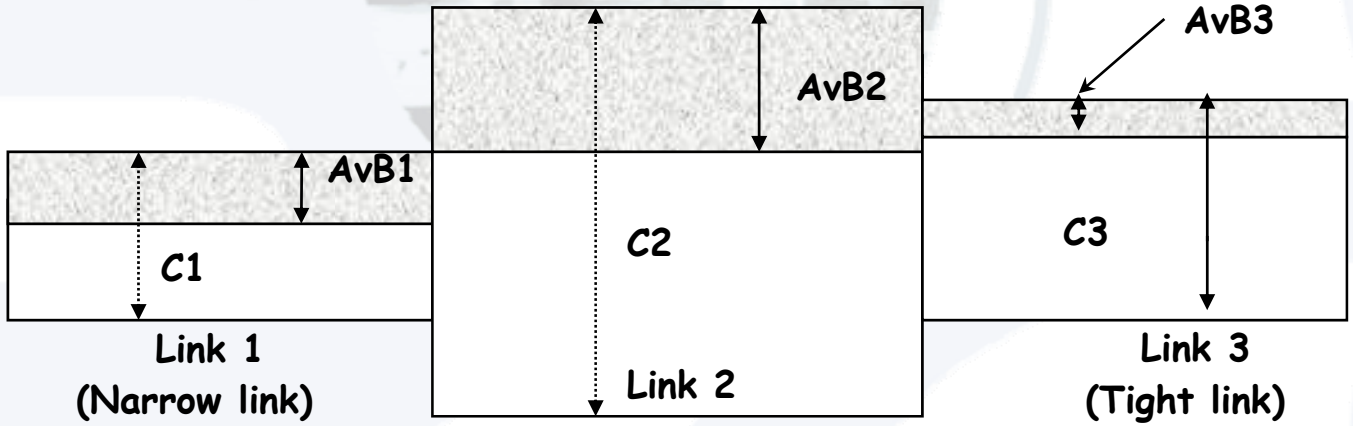
- Ping
- Traceroute
- MGEN
- RIPE equipments
- Etc.

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

(except if users/ISP don't have strong temporal requirements)

Evaluation of active estimation tools

- Bandwidth/Capacity (C_i) → Link/Path
- Utilization → Link
- Available bandwidth (AvB_i) → Link/Path
- Intrusiveness → Link



↳ How to <u>accurately estimate</u> these metrics in the <u>Internet</u>, <u>minimizing intrusiveness</u>?

Evaluated tools

Name	Authors	Methodology	Protocol	Path / Link	Root	Host
Clink	Downey	VPS	UDP	Link	yes	S
Pchar	Mah	VPS	UDP, ICMP	Link	yes	S
Pathchar	V. Jacobson	VPS	UDP, ICMP	Link	yes	S
Abing	Navratil	Packet Pair TD	UDP	Path	no	S & R
Spruce	Strauss	Packet Pair TD	UDP	Path	no	S & R
Pipechar	Guojon	Packet train TD	UDP	Link	yes	S
Pathchirp	Ribeiro	SLoPS	UDP	Path	no	S & R
IGI	Hu	SLoPS	UDP	Path	no	S & R

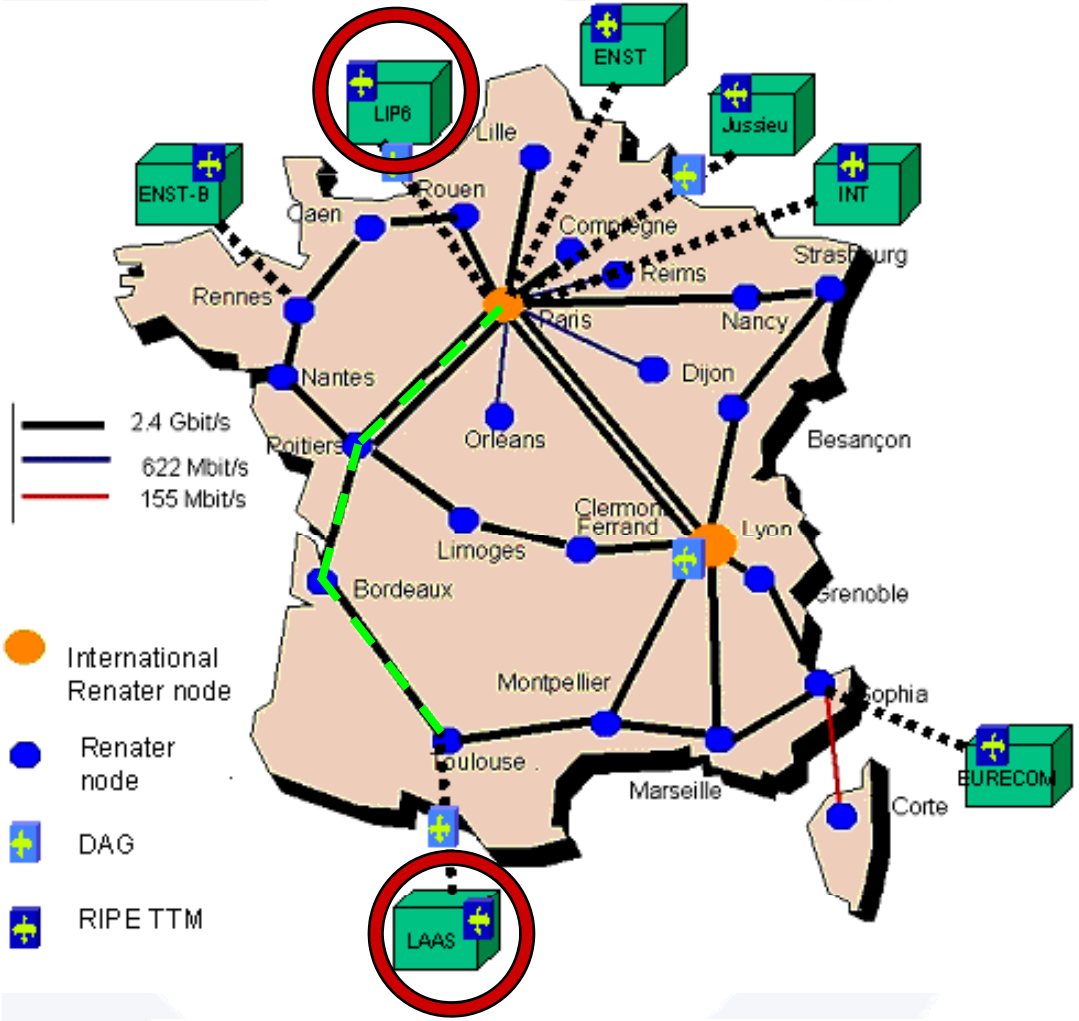
Two steps:

- Bandwidth/capacity
- Available bandwidth

} Metropolis monitoring & measurement platform

Evaluation methodology

- Metropolis: a real monitoring & measurement platform



- DAG cards
  - ➢ passive monitoring
  - ↳ Reference

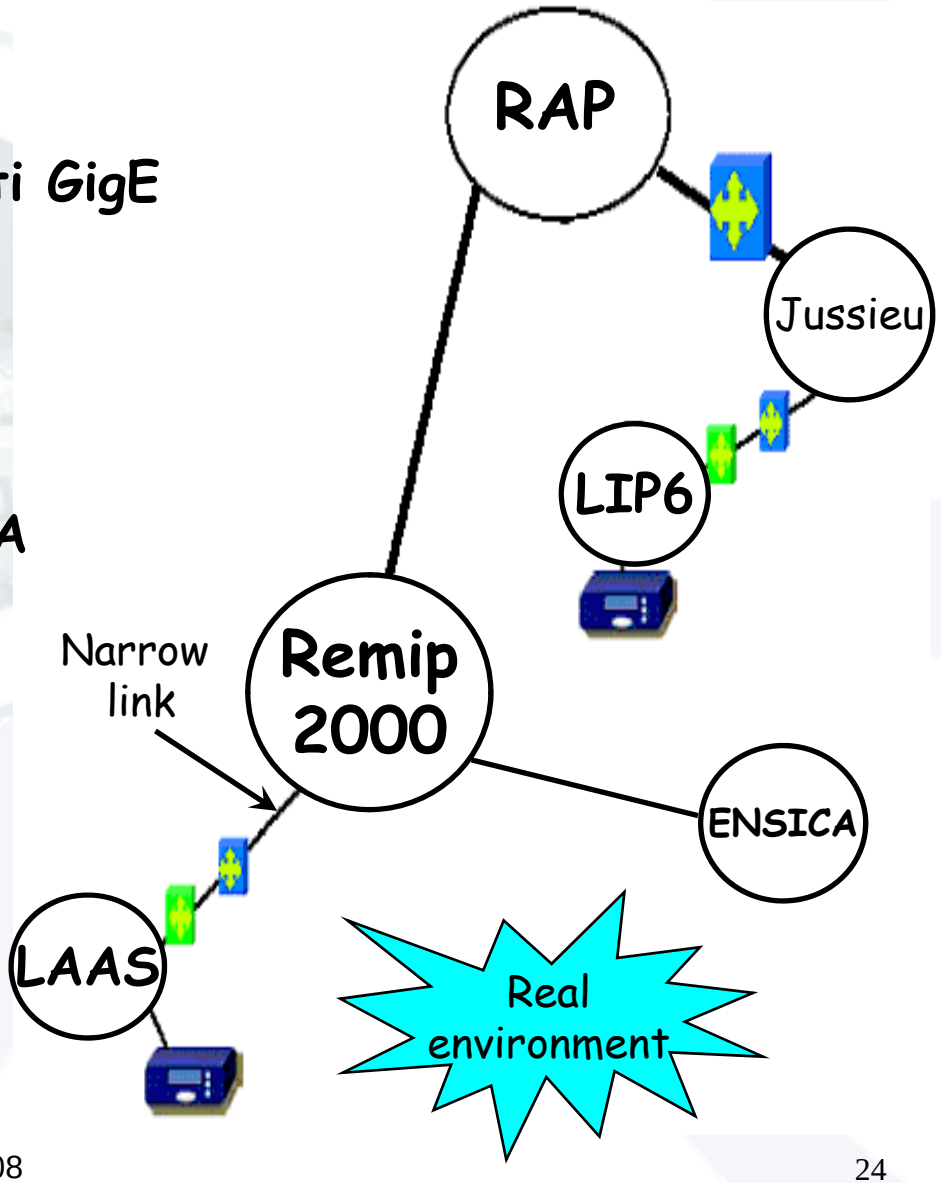
- RIPE boxes – GPS
  - ➢ active measurements
  - ↳ Bandwidth estimation

Evaluation methodology

- ✓ 15 hops between: LAAS → LIP6
- ✓ Link capacities: 100 Mbps → multi GigE
- ✓ Generated traffic: UDP constant (Iperf)
- ✓ Iperf destination traffic: ENSICA

➢ Tool hypothesis: 👍

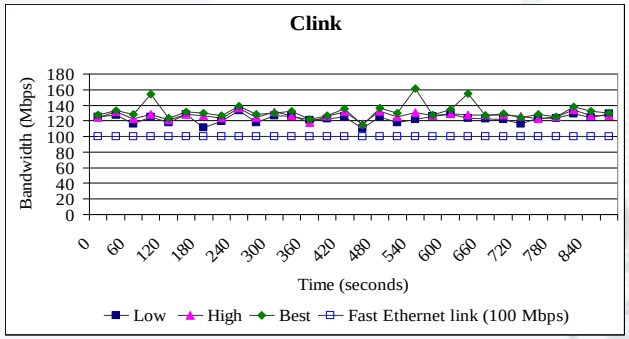
↳ Capacity/ Av. bandwidth:
On the LAAS' access link



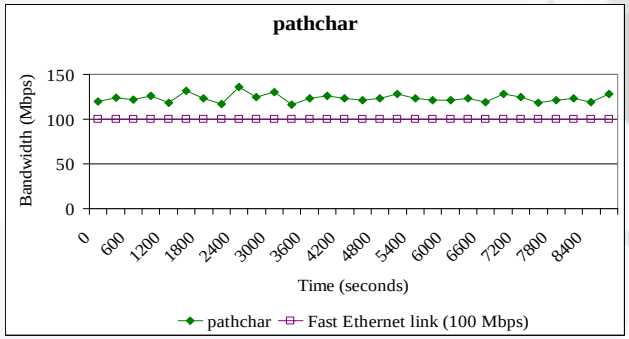
Capacity estimation

- Clink, Pchar & Pathchar

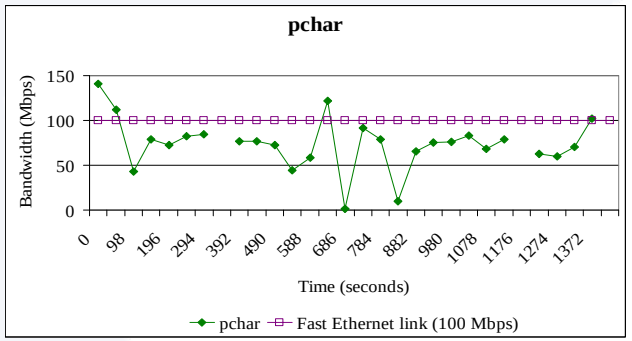
✓ no Iperf traffic



overestimation
negative values

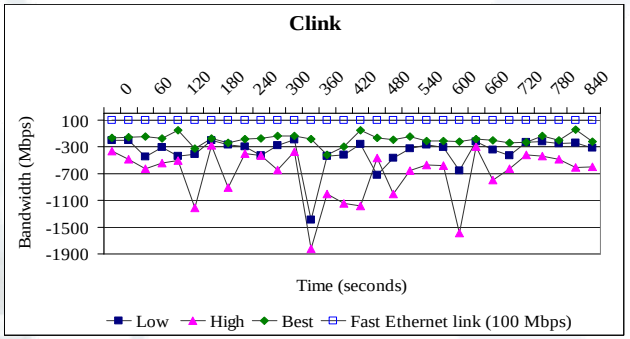


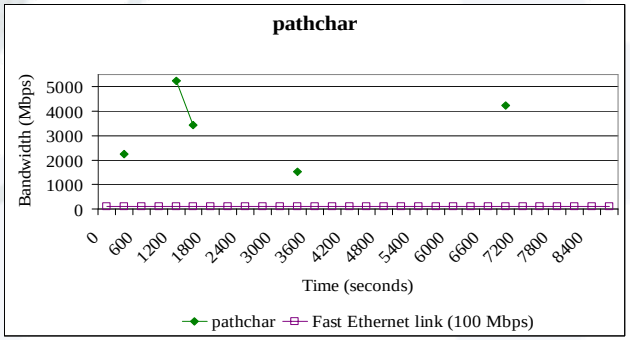
overestimation
crashes

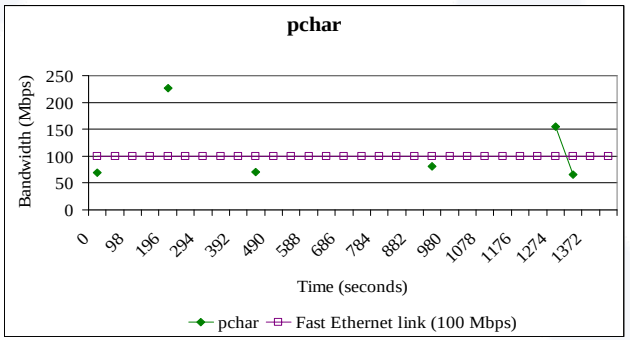


variable
crashes

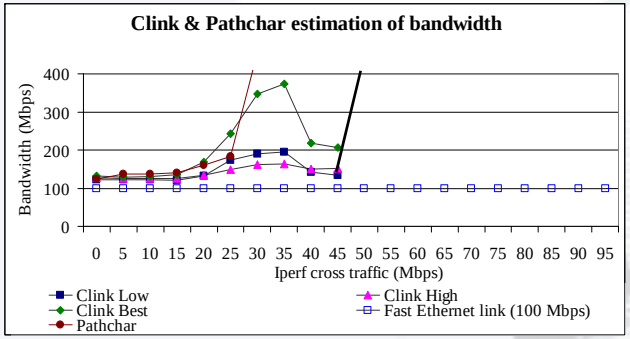
✓ 50 Mbps UDP Iperf traffic



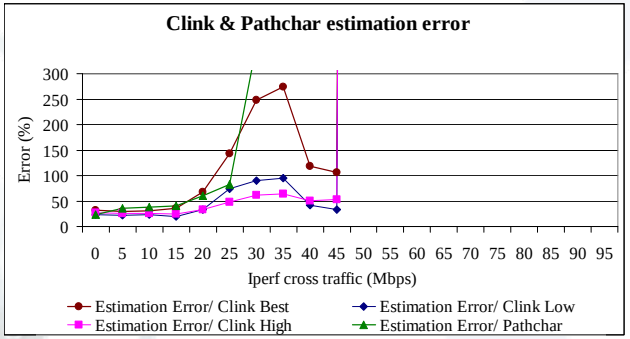




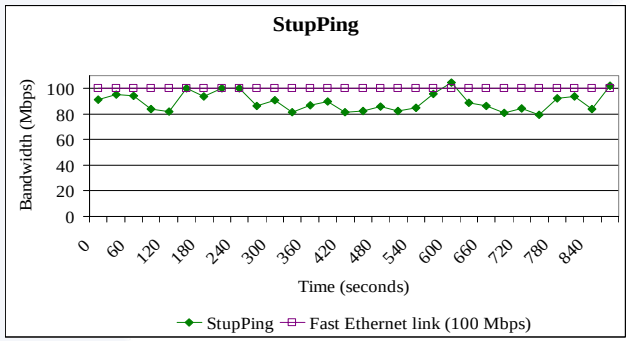
Capacity estimation



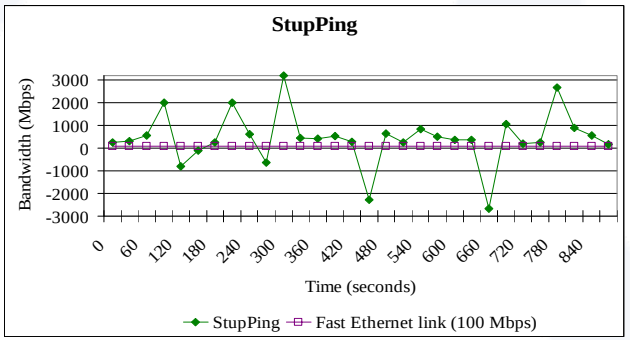
error ↗
crashes



- Tools are not accurate, reliable and robust
- *StupPing* – a rough estimation of the bandwidth with 4 *ping*: $C = 16(P_l - P_s)/((T2_l - T2_s) - (T1_l - T1_s))$



estimation
negative values



Capacity estimation

Conclusion:

- ✓ 600 runs / tool
- ✓ UDP constant traffic
- ✓ not accurate
- ✓ Intrusiveness & response time

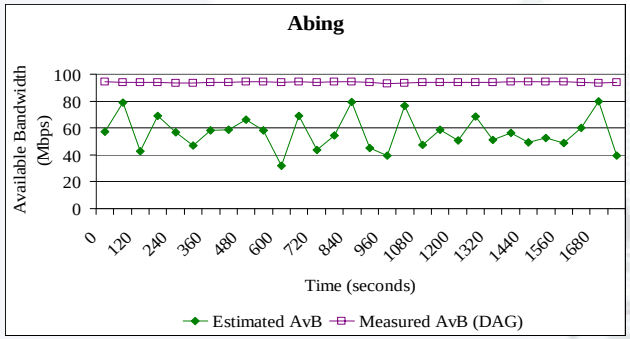
	Clink	*Pchar*	*Pathchar*	*StupPing*
Response time (s)	*1112*	*354*	*2400*	*16*
Probing Time (s)	*980*	*318*	*2232*	*14,3*
Laas response time (s)	*21*	*49*	*300*	*16*
Probing traffic Amount (Mbits)	*54,3*	*6,92*	*110,31*	*0,132*
Intrusiveness (%)	*0,0583*	*0,0217*	*0,05202*	*0,00974*

- ✓ not sufficient for most needs

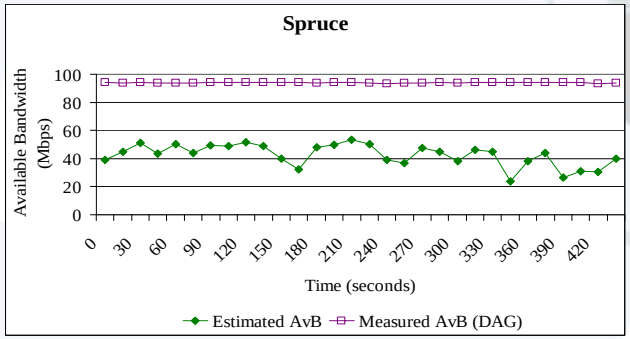
Available bandwidth estimation

- Abing, Spruce, Pipechar, Pathchirp, IGI

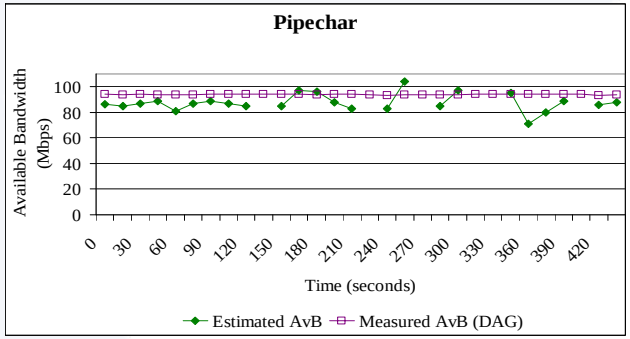
✓ no Iperf traffic



variable
wrong values

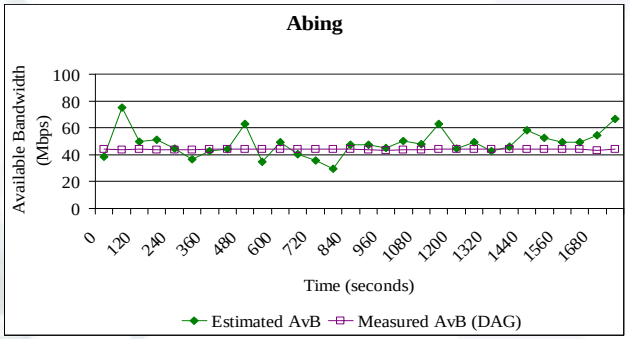


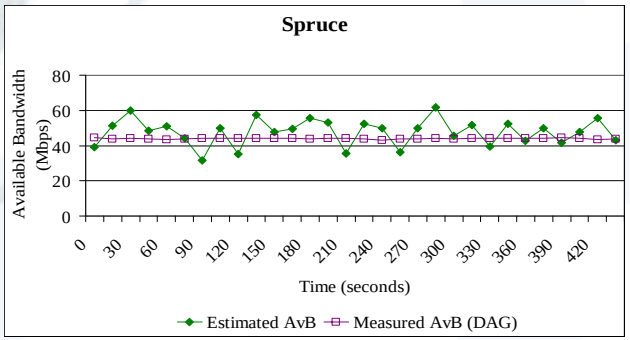
variable
wrong values

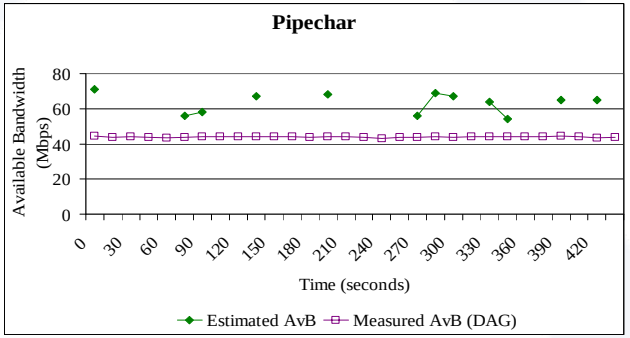


estimation
crashes

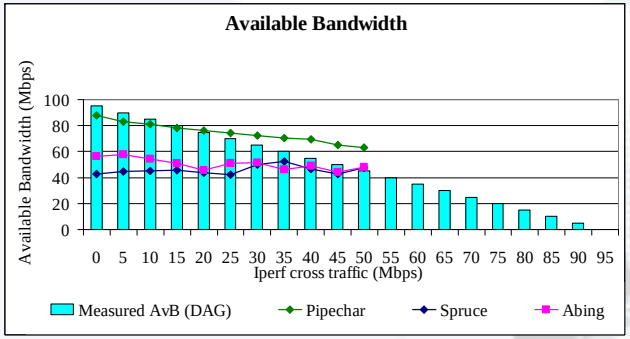
✓ 50 Mbps UDP Iperf traffic





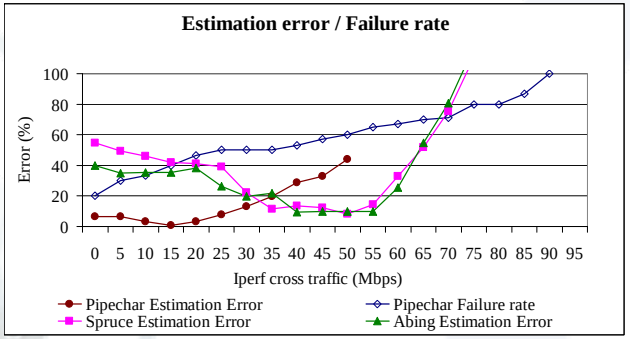


Available bandwidth estimation



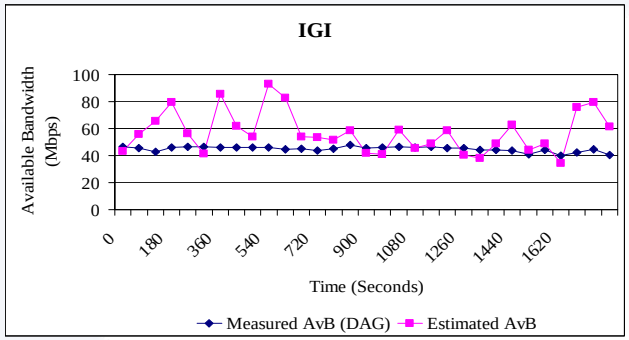
error ↗

crashes



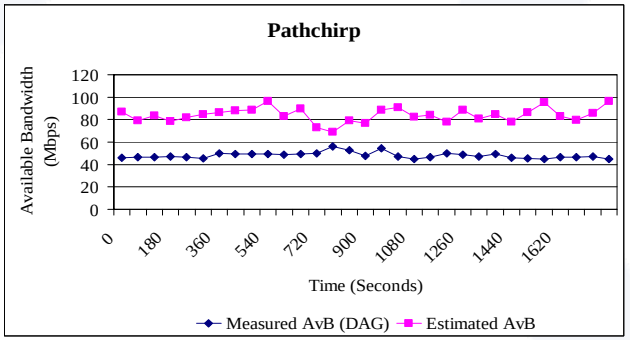
Tools are not accurate, reliable and robust

Another evaluation: IGI and Pathchirp



variable

overestimation



Available bandwidth estimation

Conclusion:

- ✓ 600 runs / tool
- ✓ UDP constant traffic
- ✓ not accurate
- ✓ Intrusiveness & response time

	Abing	*Spruce*	*Pipechar*
Response time (s)	***1,1***	***11***	***161***
Probing Time (s)	***0,96***	***9,8***	***141,2***
Laas response time (s)	***N/A***	***N/A***	***161***
Probe traffic Amount (Mbits)	***0,464***	***2,34***	***38,5***
Intrusiveness (%)	***0,509***	***0,251***	***0,286***

- ✓ not sufficient for most needs

Conclusion on active measurement tools

- overestimation of available bandwidth,
- variable results,
- many crashes,
- long response time…

↳ not convinced by any of these tools for our needs

Passive measurements

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
  - 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 (→ not really the scope of this keynote)
  - 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
- Wireshark (ethereal)
- CISCO's Netflow
- QOSMOS - IPANEMA
- OCxMON (mainly ATM)
- DAG

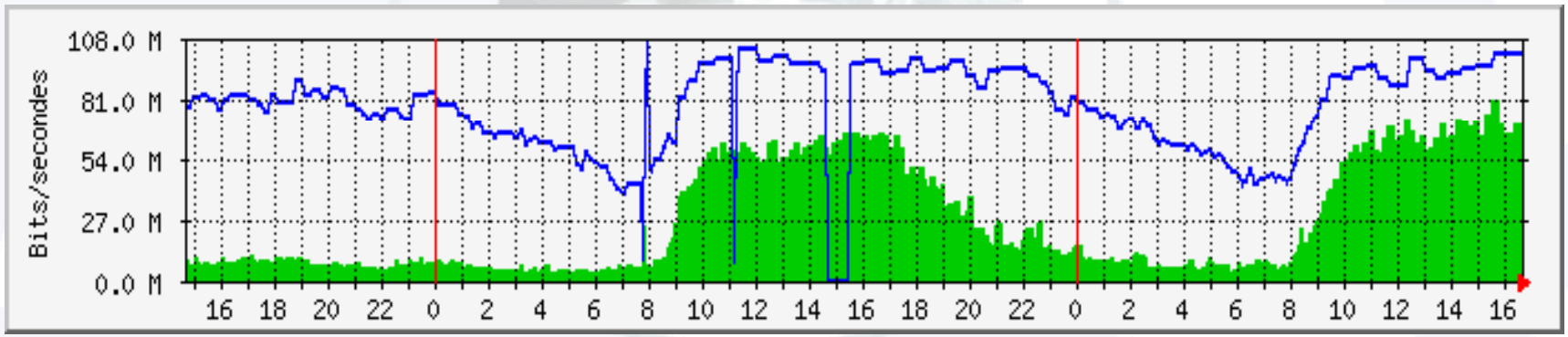
Software

Hardware

Traffic analysis

Examples of SNMP measurements

RAP ↔ RENATER interconnection

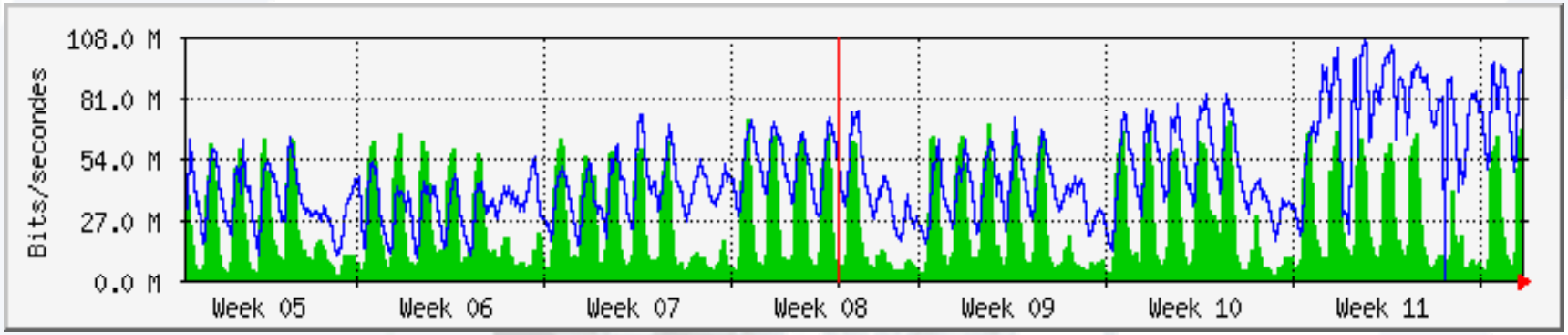


Per hour trace

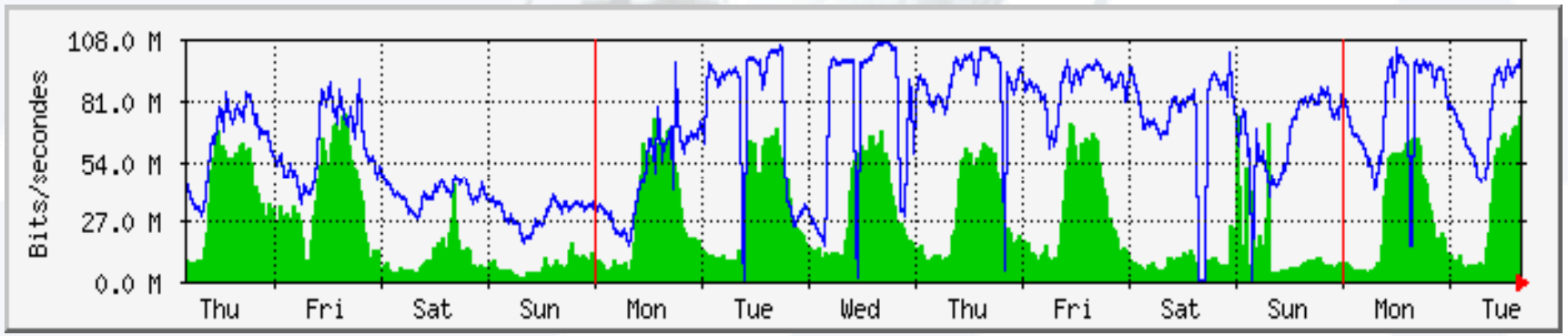
Input traffic

Output traffic

Examples of SNMP measurements (2)



Per Month trace



Per Week trace

Input traffic

Output traffic

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

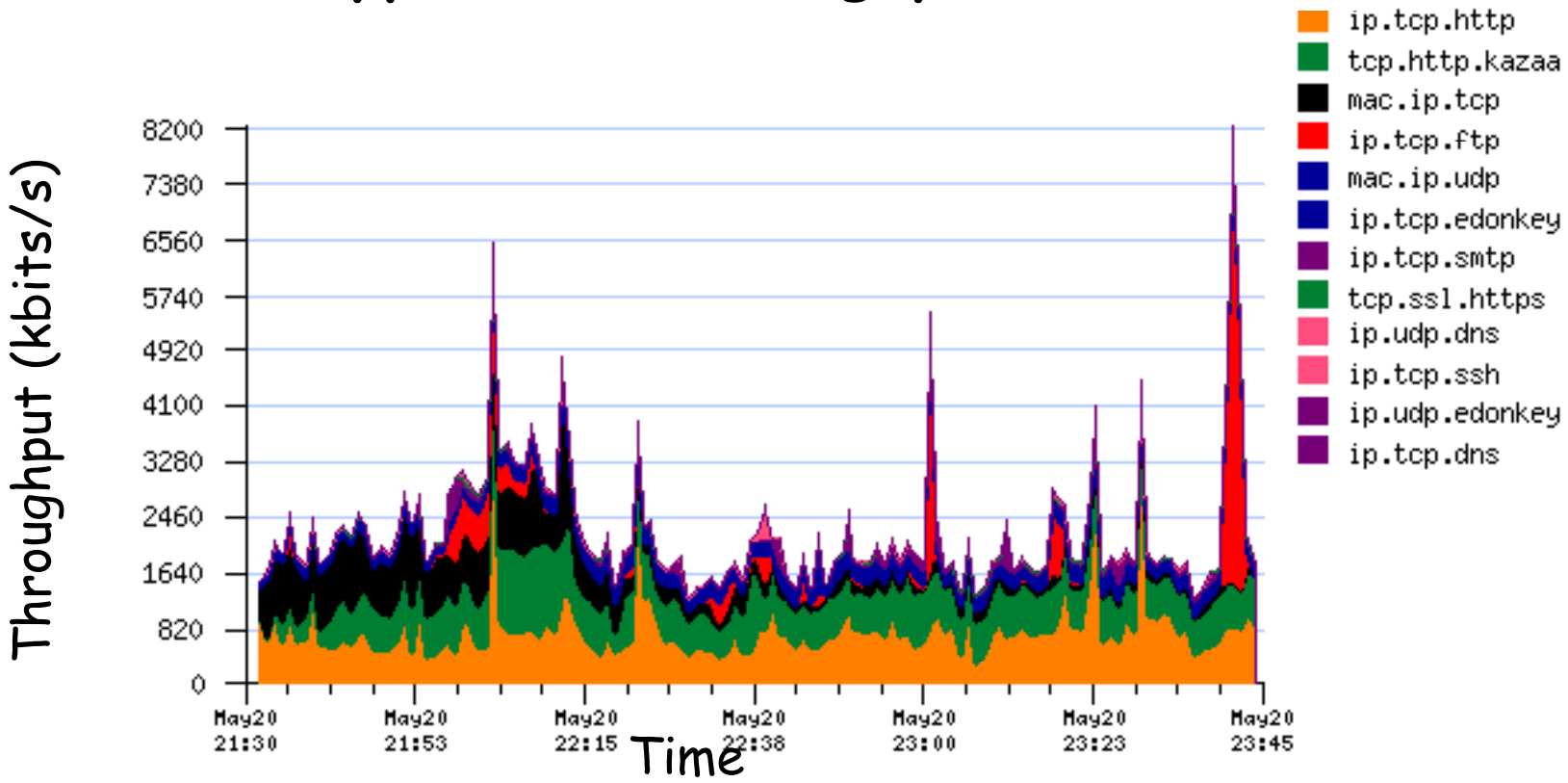
Problems for monitoring networks

- How to explain this over-provisioning requirement ?

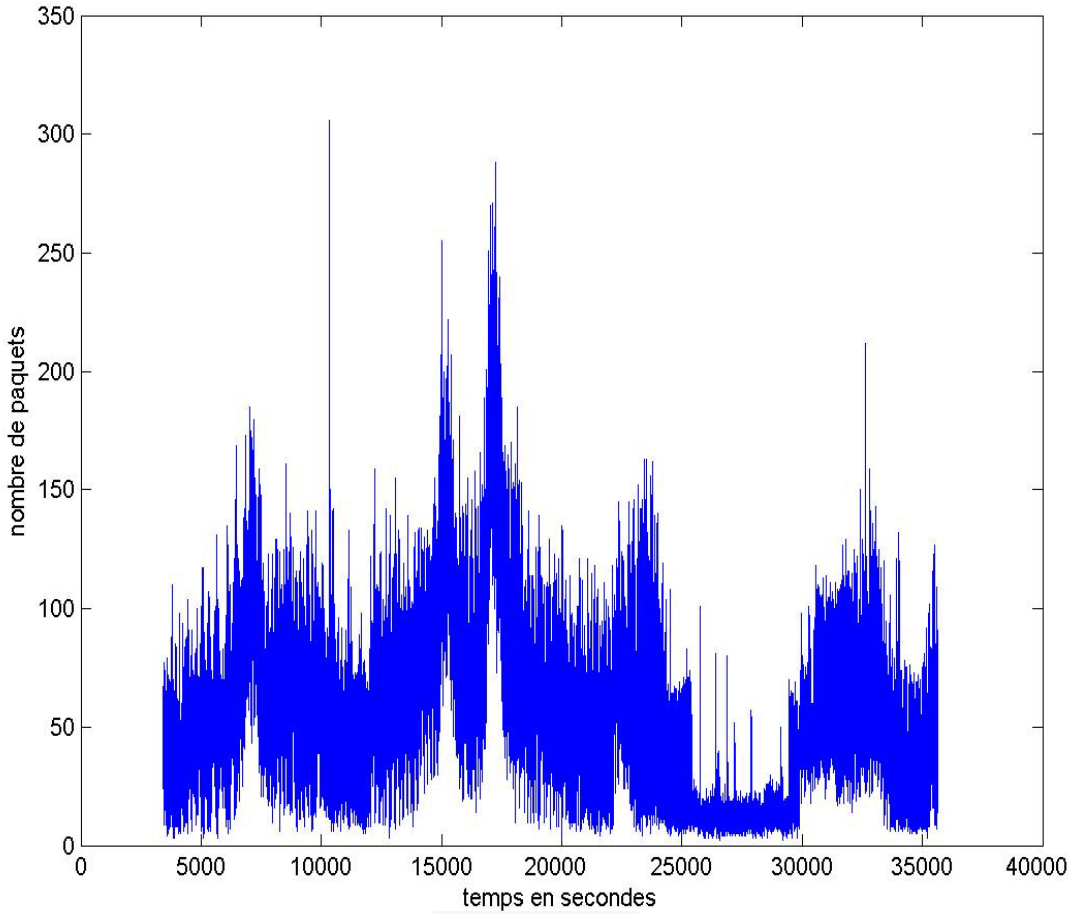
- Impossible to monitor traffic dynamics (second order values as variability auto-covariance for instance)
- Impossible to monitor traffic QoS (user point of view – goodput)
- Impossible to get a (formal) traffic model

Internet traffic variability (May 2003)

Main TCP applications throughputs (Renater)

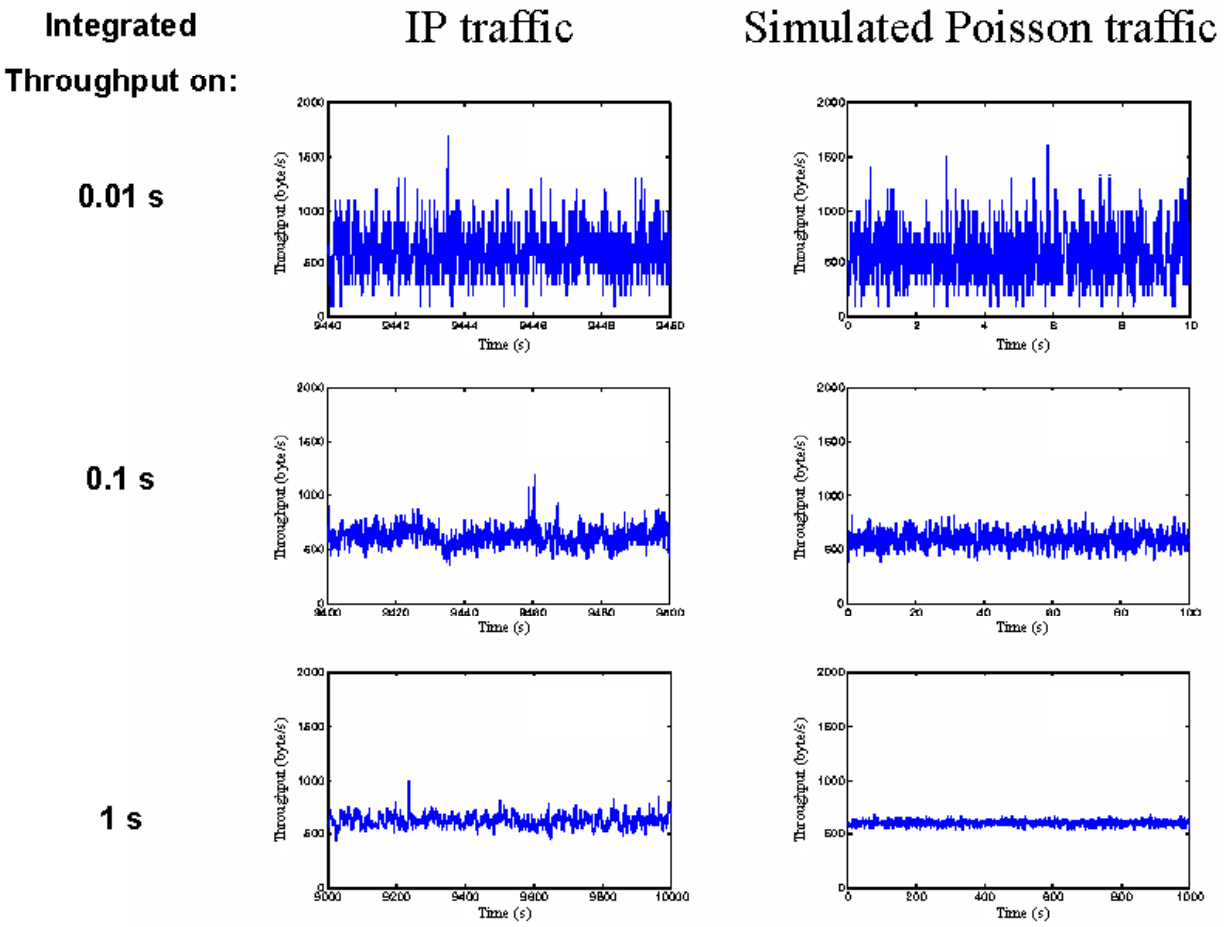


Internet traffic variability (May 2004)



Number of TCP/SYN packets on LAAS Internet access link

Traffic variability and aggregation scale

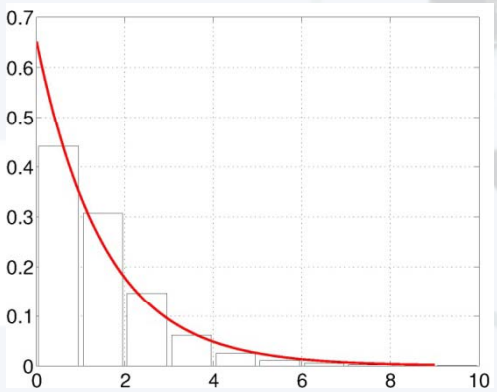


Marginal laws

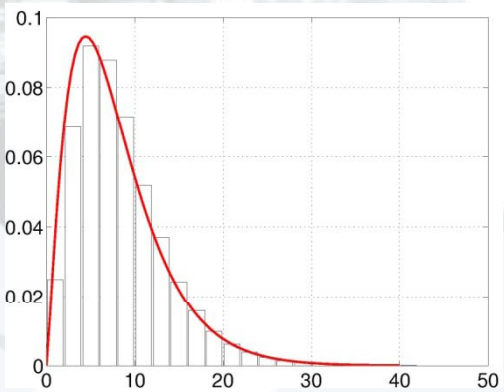
- Distributions of empirical probabilities LBL-TCP-3

for $X_\Delta(k)$ = #pkt during $[k\Delta, (k+1)\Delta]$

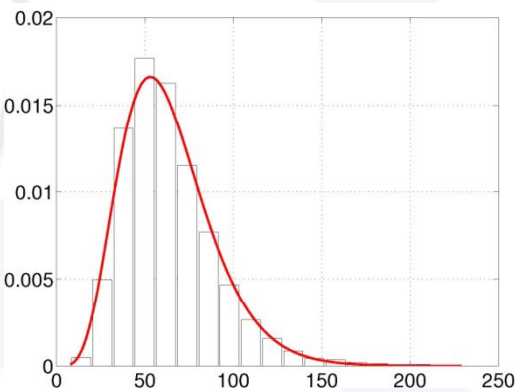
or $W_\Delta(k)$ = #bytes during $[k\Delta, (k+1)\Delta]$



Δ=4ms



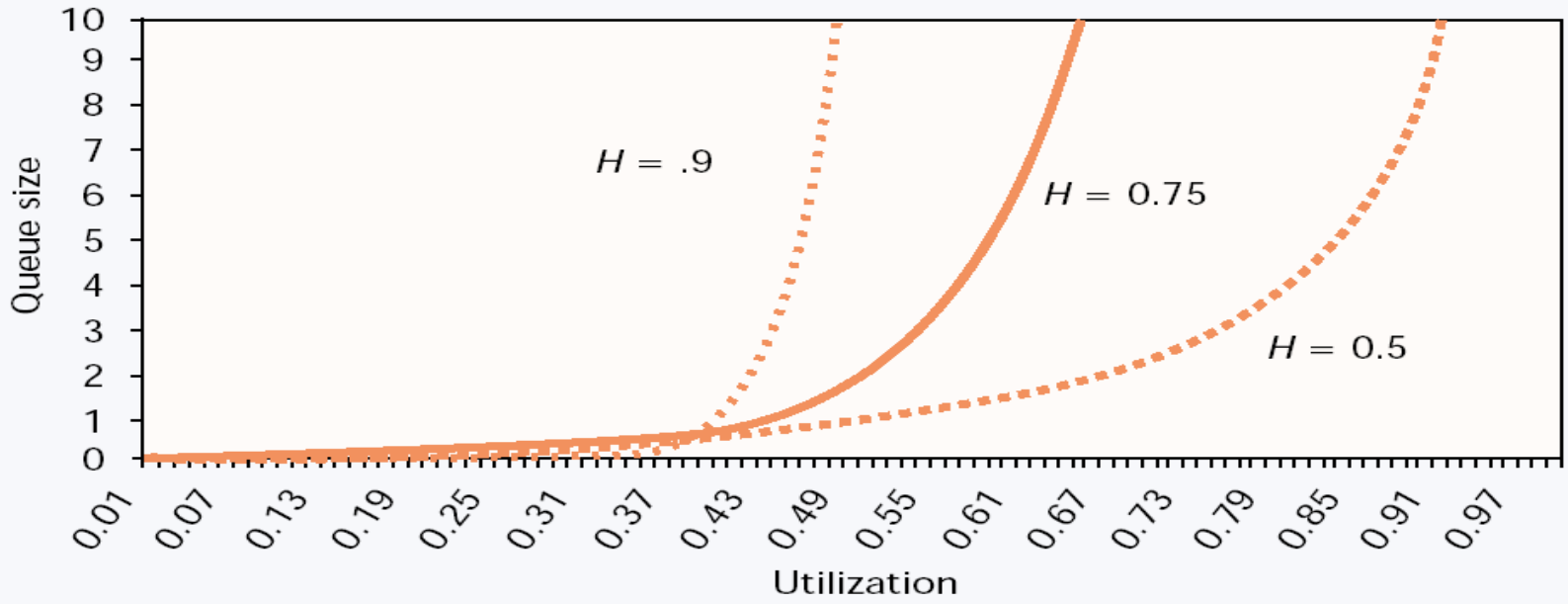
Δ=32ms



Δ=256ms

- Poisson model? Exponential law? Gaussian?
- What aggregation level to select? What sampling rule?

LRD and network performance



relation between LRD , network usage and queue sizes in routers

Illustration of LRD impact

- Worst case with a highly LRD traffic
  - $E[X] = \lambda$
  - $V[X] \approx O(\lambda^2)$
- Example:
  - If E[X] = 100 Mbps → Provisioning for a reliable network with low delays is of the order of 1 Gbps (size of bursts)
  - Reach the technological limits of networks (as well as monitoring equipments) with an average traffic of few hundreds of Mbps

Wireless networks monitoring

As far as I know

- Nothing exist for monitoring wireless traffic on the air medium
  - All current and past studies are being done on the collection wired network

- The interest would be to monitor wireless traffic at the MAC layer

Discussion

Guidelines for a global monitoring system (1)

- Combine active and passive measurements
- Passive measurements
  - For any intra-domain measurements
  - For inter-domains if ISP/carriers collaborate and exchange measurement information (e.g. IPFIX ?)
  - Need to develop multi-scales analysis procedures
    - → The help of signal processing and statisticians experts is welcome

Guidelines for a global monitoring system (2)

- Active measurements
  - Solution for inter-domain measurements (at least delays) and issuing other parameters estimates
  - Performances must be significantly improved
  - Probing sampling is necessary to avoid probes to be discarded by security equipments (looks like network scanning)

Guidelines for a global monitoring system (3)

- Clocks synchronization must use GPS (NTP only is not accurate enough)
- Computing of measured time series must be performed locally as much as possible
  - For scalability purposes
  - Then routers/gateways/proxies are certainly not the right place for monitoring – except if their architecture fully integrates hardware monitoring capabilities

Guidelines for a global monitoring system (4)

- Dedicated monitoring and measurement equipments must be designed
  - Combining hardware and software
  - Dedicated small OS for performance purpose and flexibility
- Wireless monitoring equipments working at MAC level must be designed

That's all folks