

General Outline

- Introduction to IP flows
- IP flow monitoring systems
- IP flow monitoring exporting standards
- An IP flow monitoring example: GÉANT2
- List of tools for IP flows processing
- Advanced stuff

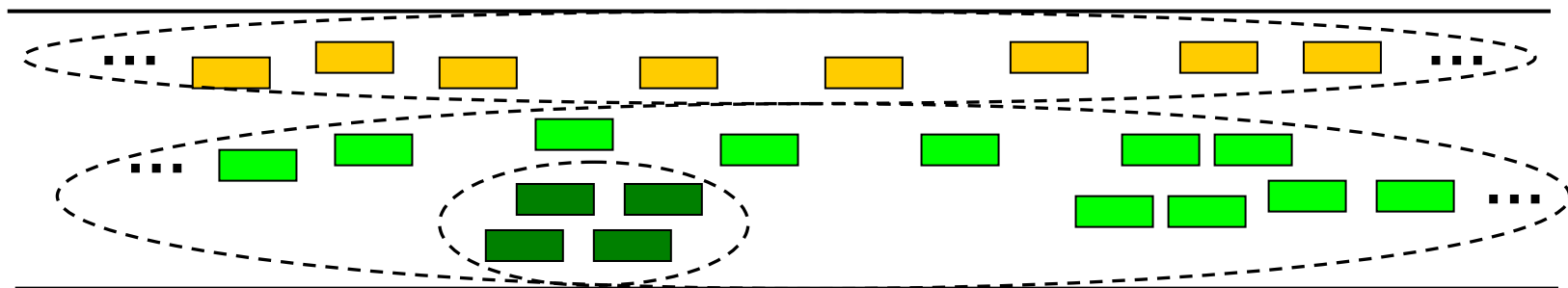
Part 1/6

- Introduction to IP flows
 - What are they?
 - How are they measured?
 - What applications use these measurements?
- IP flow monitoring systems
- IP flow monitoring exporting standards
- An IP flow monitoring example: GÉANT2
- List of tools for IP flows processing
- Advanced stuff

IP flows

- IP Flows are groups of IP packets sharing a common characteristic, e.g.
 - IP src/dst address
 - src/dst ports
 - Transport layer protocol
 - Type Of Service (TOS) field

Flows can be long lasting...

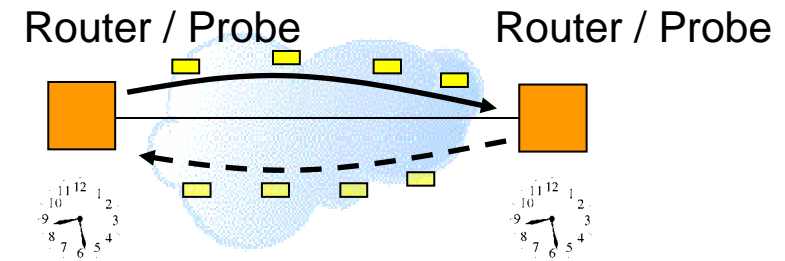
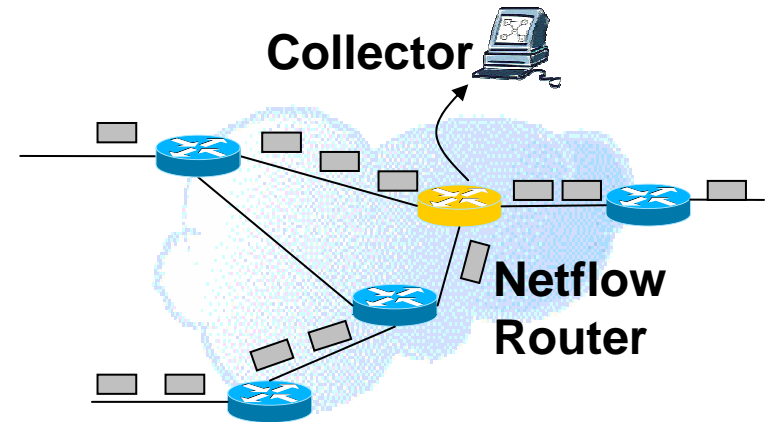


... or have a limited lifetime...

**... and packets may
belong to more than one flow**

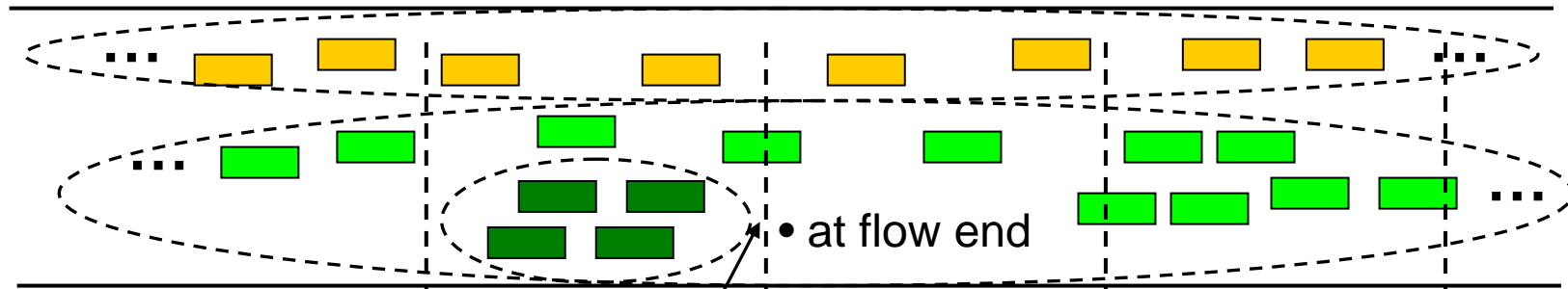
Measurement category

- IP flow monitoring is a *single point, passive* network measurement
 - Routers just “observe and report”
- In active measurements, test traffic is injected in the network
- In two point measurements, events at two points need to be correlated
 - E.g. packet transit time



IP flows measurement

Flows can be long lasting...



... or have a limited lifetime...

... and packets may belong to more than one flow

Reported flow information

-what:

- src IP, dst IP, ports
- Start time
- End time
- # packets
- # bytes
- Other

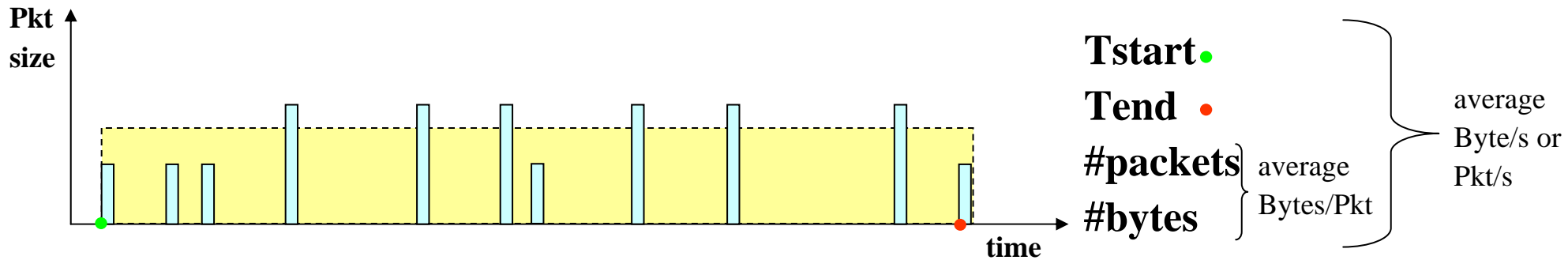
-when:

- Periodically for long lasting flows

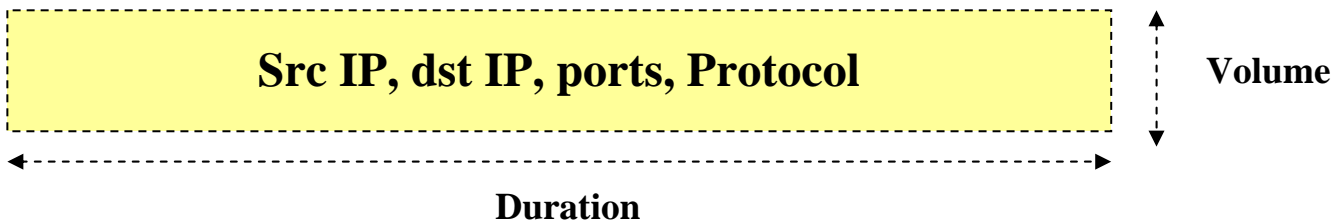
t

What IP flow monitoring gives, what not

- It's time and volume *summary* information

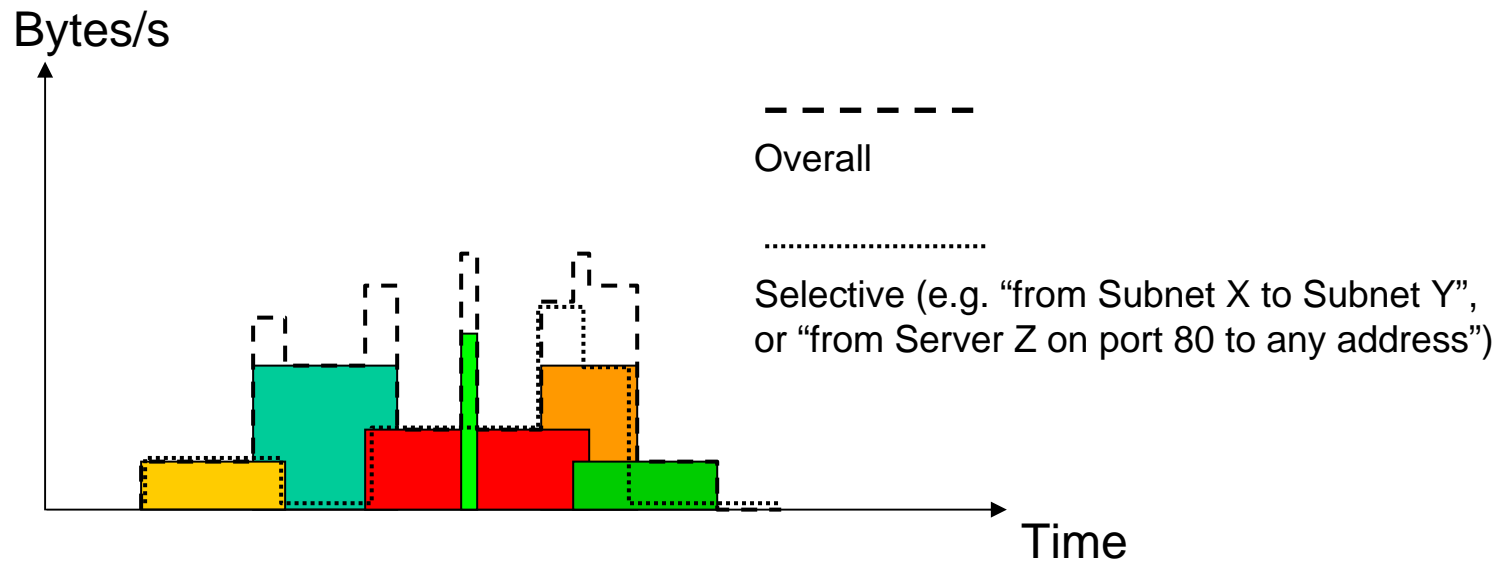


- No inter-Pkt arrival times
- No single Pkt sizes
- All you have is a “labelled brick”



So what can you do?






- Compose bricks...



Applications using IP flow info

- Traffic Engineering
- Billing / Accounting
- Network Planning
- Security
- Discovery of usage and application patterns
 - who talks to whom
 - E.g. AS/AS matrixes
 - what applications are used (if they can be recognised...)

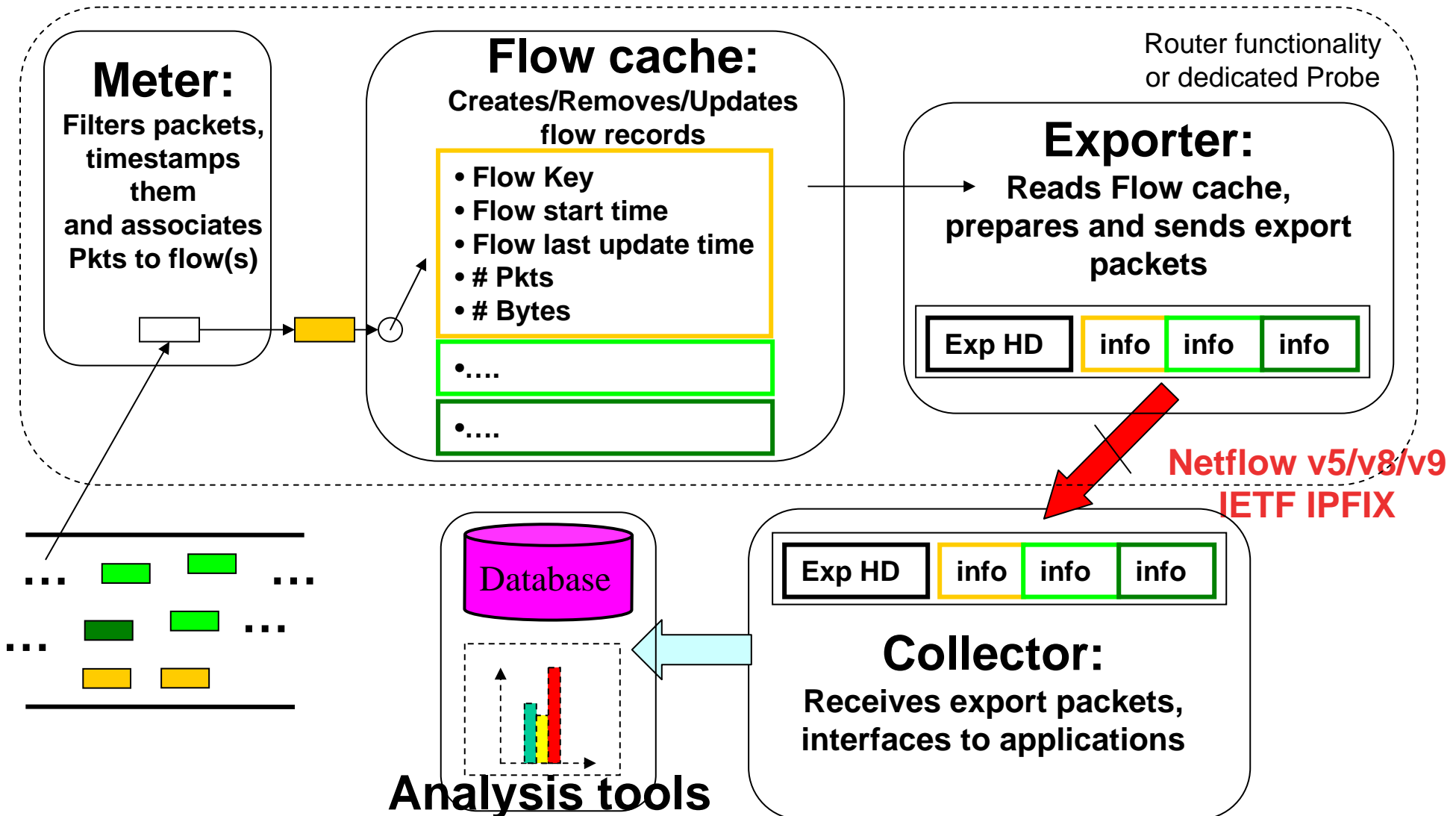
Applications using IP flow info (cont.)

Application	Time Granularity	Space Granularity
Traffic Engineering	(minutes)	
Billing / Accounting	(minutes-months)	
Network Planning	(months)	
Security	(minutes-days)	
Discovery of usage and application patterns	(months)	

Part 2/6

- Introduction to IP flows
- IP flow monitoring systems
 - General architecture
 - Challenges
- IP flow monitoring exporting standards
- An IP flow monitoring example: GÉANT2
- List of tools for IP flows processing
- Advanced stuff

General architecture of a IP flow monitoring system



IP flow monitoring system: challenges

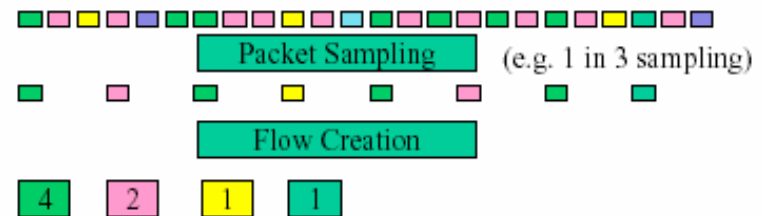
- Export side (router)

- A lot of flows, to be updated at each packet arrival: routers may not cope with that

- Dedicated hardware

- packet sampling

- Reduce flow cache size: aggressive export



- **Will require re-normalization!**
- **Small flows may be missed!**

- Transport

- UDP: easy but unreliable
- TCP/SCTP: reliable but heavy for NICs
- Security aspects (TLS/DTLS)

- Analysis

- Too much data: no “universal” tool. Different tools needed to do separate tasks well.

Part 3/6

- Introduction to IP flows
- IP flow monitoring in routers
- IP flow monitoring exporting standards
 - Netflow (and other given names ...): details
 - Netflow evolution: v5, v7, v8, v9
 - IPFIX
- An IP flow monitoring example: GÉANT2
- List of tools for IP flows processing
- Advanced stuff

Cisco Netflow: origin and evolution

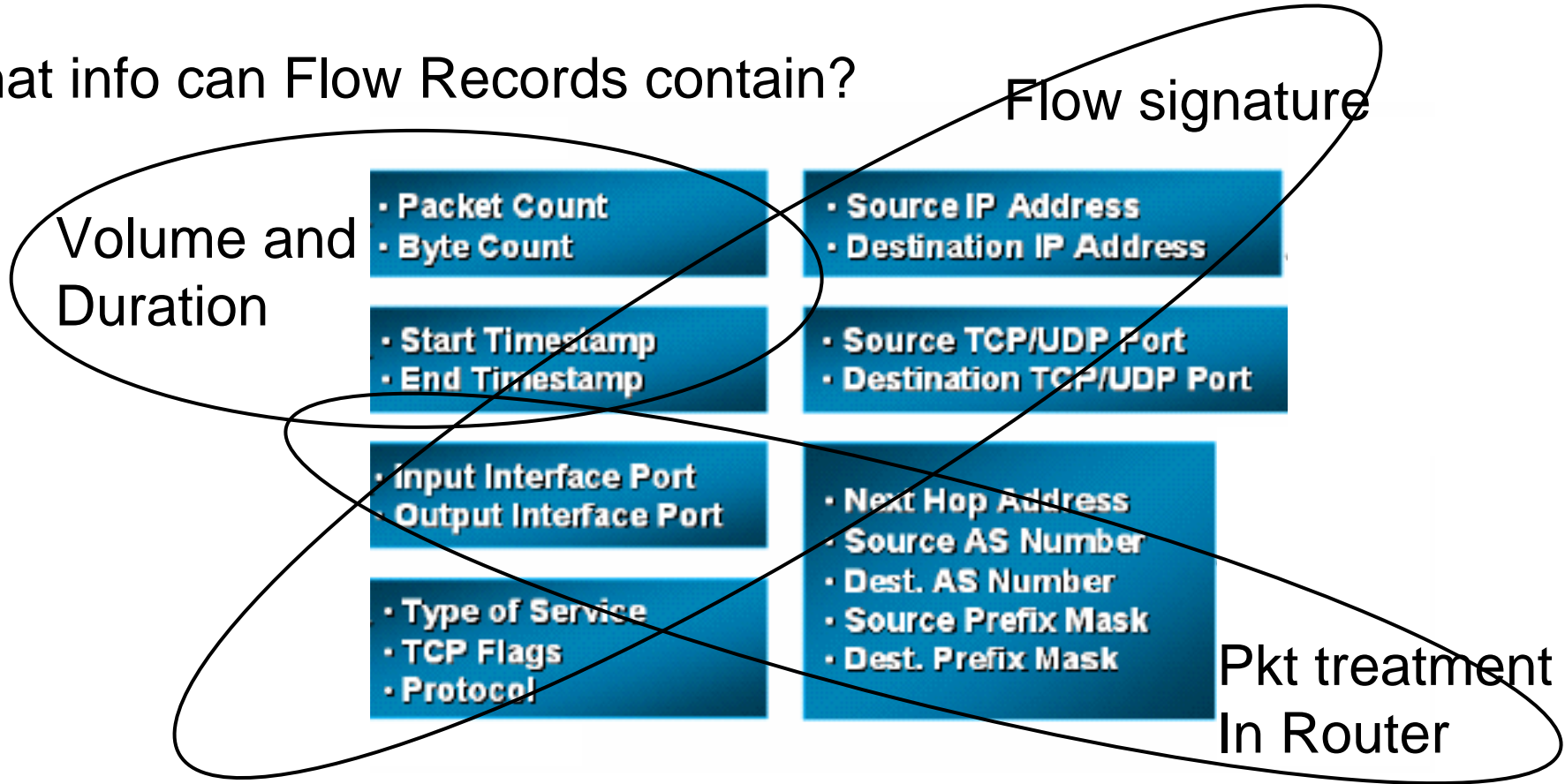
- 1996 Initially designed at Cisco (Daren Kerr and Barry Bruins) as a switching path speedup
 - Then realized that per-flow information had also other value
- **v5**: first widely implemented version
 - Fixed export format, no aggregation: each flow is reported separately
- **v7**: Specific to 6500 and 7600 Switches
- **v8**: 11 possible aggregation schema
- **v9**: flexible aggregation (template based).
Chosen as “baseline” for IPFIX

Netflow: other Given names

- Juniper
 - cflowd (v5, v8, v9 – recently!)
- Huawei
 - Netstream (v5, v8, v9)
- Avici
 - Supports v5 and v9
- Alcatel
 - Supports v5 and v8
- ...

Netflow record content

- What info can Flow Records contain?



- What identifies a Flow Record?

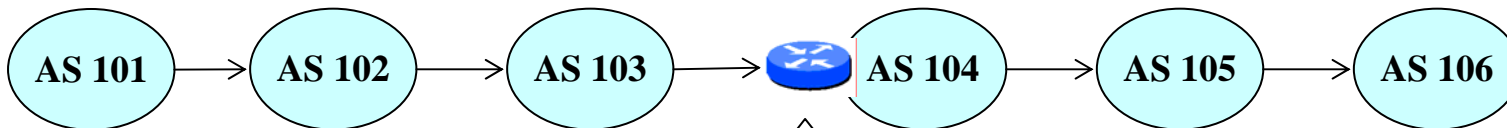
Src IP, Dst IP, Src Port, Dst Port, Protocol, Input If, TOS are *key fields*

5-tuple (most common definition)

7-tuple

Comments on Netflow record fields

- Start and end times are relative to first and last flow's packet (*not* to record's export time...)
- TCP flags (S,F,A,P,U,R) are *cumulative* for the flow
- AS can be either src/dst or prev/next, *not* both!
 - It's a configuration option
 - It's obtained in the router via a routing lookup (it's *not* in the IP packets) !



Netflow Enabled Router

- If “origin-as” is configured, it will report: Src AS->101, Dst AS->106
- If “peer-as” is configured, it will report: Src AS->103, Dst AS->105

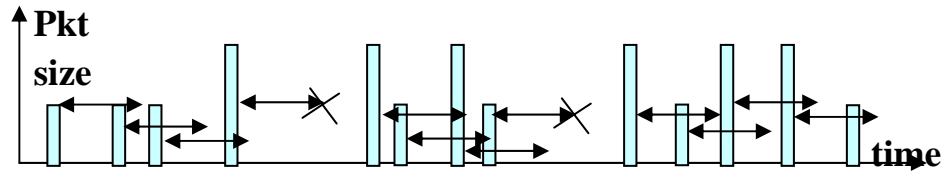
Controlling the exporting

- Four conditions govern the expiration of flows from flow cache (and their exporting)
 - Inactive timeout: if a flow has not been updated for more than IA_tout sec., export it
 - Active timeout: if a flow was created more than A_tout sec. ago, export it
 - End of flow detected: works for TCP only (FIN or RST Pkt)
 - Internal flow cache management: if flow cache has more than X flows, or is more than Y% full, start exporting flows (with some criteria)

Controlling the exporting (cont.)

- Inactive T_{out} :

- if too small, will “split” the same flow



- flows with low pkt rate R are more at risk: $1/(RS) \cong IA_{tout}$ (S : sampling rate)

- if too high, too many flows in cache
 - $N = \lambda \mu$ where μ (flow duration + IA_{tout}) is dominated by IA_{tout}
- Typical values of IA_{tout} : 10s-60s

- Active T_{out} :

- If too small, will “split” (too much...) the same flow
- If too high, collectors working on discrete time slots will show non-existing traffic peaks
- Typical values of A_{tout} : 5min-30min

- FIN or RST: will not be effective in case of sampling

Common configuration commands

- **Cisco (CLI)**

- `ip flow-export version <version> [origin-as|peer-as|bgp-next-hop]`
- `ip flow-export destination <address> <port>`
- `ip flow-cache timeout inactive <seconds>`
- `ip flow-cache timeout active <minutes>`
- `ip flow-cache entries <number>`

- **Juniper (conf-file)**

```
cflowd collector-host-address {  
  Autonomous-system-type (origin|peer);  
  port port-number;  
  version version-number;  
  (local-dump | no-local-dump);  
}
```

Visualizing the configuration and flow cache on routers

- Cisco

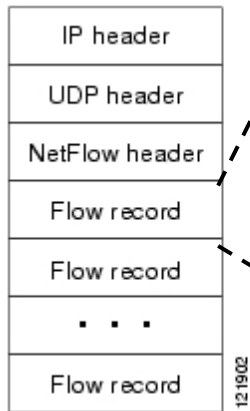
- `show ip cache [verbose] flow`
 - Will show flow cache configuration and statistics, and flow details
- `show ip flow export`
 - Will show exporting process statistics

- Juniper

- `show configuration forwarding-options sampling`
 - Will show flow collection configuration
- `monitor start sampled`
 - Equivalent of unix “tail -f” command on a file where the flow records are dumped (not advised to create this file in production, because of additional load on Routing Engine)

Netflow v5

- Most commonly deployed version, even today
- Flow records exported in UDP packets
- 30 flow records in a 1500 bytes pkt



121902

Content	Bytes	Description
srcaddr	0-3	Source IP address
dstaddr	4-7	Destination IP address
nexthop	8-11	Next hop router's IP address
input	12-13	Ingress interface SNMP ifIndex
output	14-15	Egress interface SNMP ifIndex
dPkts	16-19	Packets in the flow
dOctets	20-23	Octets (bytes) in the flow
first	24-27	SysUptime at start of the flow
last	28-31	SysUptime at the time the last packet of the flow was received
srcport	32-33	Layer 4 source port number or equivalent
dstport	34-35	Layer 4 destination port number or equivalent
pad1	36	Unused (zero) byte
tcp_flags	37	Cumulative OR of TCP flags
prot	38	Layer 4 protocol (e.g. 6=TCP, 17=UDP)
tos	39	IP type-of-service byte
src_as	40-41	Autonomous system number of the source, either origin or peer
dst_as	42-43	Autonomous system number of the destination, either origin or peer
src_mask	44	Source address prefix mask bits
dst_mask	45	Destination address prefix mask bits
pad2	46-47	Pad 2 is unused (zero) bytes

Netflow v7 and v8

- v7
 - Specific to 6500 and 7600 Switches
 - Similar to v5, but *without* AS, Interface, TCP flag and ToS info
- v8
 - Goal: reduce exported information, and primary flow cache size, with “aggregation”
 - 11 “aggregation schemes”: AS, Destination-Prefix, Prefix, Protocol-Port, Source Prefix, AS-ToS, Destination-Prefix-ToS, Prefix-ToS, Protocol-Port-ToS, Source Prefix-ToS, Prefix-Port

Netflow v9

- Previous versions have *all* a fixed export format
- To overcome the fixed format, one could always export “type, length, value” ⇒ *A lot of overhead! ...*
- ...or separate “type, length” from “value”
- Templates specify the type and length of carried info
- just the data is exported in “Data Flow Sets”
- Each Data Flow Set is preceded by an identifier *pointing* to the template needed to its decoding
 - If templates are lost, data flow sets cannot be decoded!
- v9 Can run over multiple transports (not just UDP)

Table 1 NetFlow Version 9 Export Packet

Packet Header	Template FlowSet	Data FlowSet	Data FlowSet	Template FlowSet	Data FlowSet
---------------	------------------	--------------	--------------	-------	------------------	--------------

IPFIX

- IETF standard, chartered in 2002 to
 - “Find or develop a basic common IP Traffic Flow measurement technology to be available on (almost) all future routers”
- Netflow v9 selected as a baseline for the IPFIX standard, but *without* backward compatibility constraints
- Cisco is the driving force behind IPFIX, but other vendors (NEC, Hitachi) are active (or observing)
- Status: main documents in RFC editor’s queue, i.e. the core protocol is “stable”
 - Still to be seen if/when Cisco will offer it!

IPFIX what's new

- Formal definition of a large number of “information elements” to carry the elementary information
 - “big extension” of the v5 table shown before
 - E.g. absolute and delta counters, timestamps with [s], [ms], [μs], [ns] resolution
 - Possibility to extend it and to define enterprise specific information elements
- Options templates and template flow records can be used to export configuration information about the metering process

IPFIX what's new (cont.)

- IPFIX can use Stream Control Transport Protocol (SCTP – RFCs 2960, 3309, 3758), TCP or UDP as transport protocols
 - Debate in the IETF, because
 - UDP is not congestion aware
 - TCP is heavy for line cards and exposes to Head of Line blocking
 - SCTP is new and not widely implemented
- PR-SCTP is the preferred transport because
 - “it is congestion aware...but with a simpler state machine than TCP”
- An SCTP association can contain multiple streams. At minimum, an IPFIX implementation **MUST** have two associations, one for data and one for templates
 - Reliable transport for templates, partly reliable (e.g. limited no of retransmissions) for data

IPFIX what's new (cont.)

- Simple devices can still use UDP as a transport
 - But templates must then be periodically refreshed
- Security:
 - If TCP is transport, use TLS
 - If UDP or SCTP, use DTLS
 - But mature implementation of DTLS over SCTP are missing, therefore
 - Either use TLS over TCP
 - Or use DTLS but without reliability
 - Always use mutual X.509 certificates based authentication

Part 4/6

- Introduction to IP flows
- IP flow monitoring in routers
- IP flow monitoring exporting standards
- An IP flow monitoring example: GÉANT2
 - Collection
 - Analysis
- List of tools for IP flows processing
- Advanced stuff

Netflow collection in GÉANT2

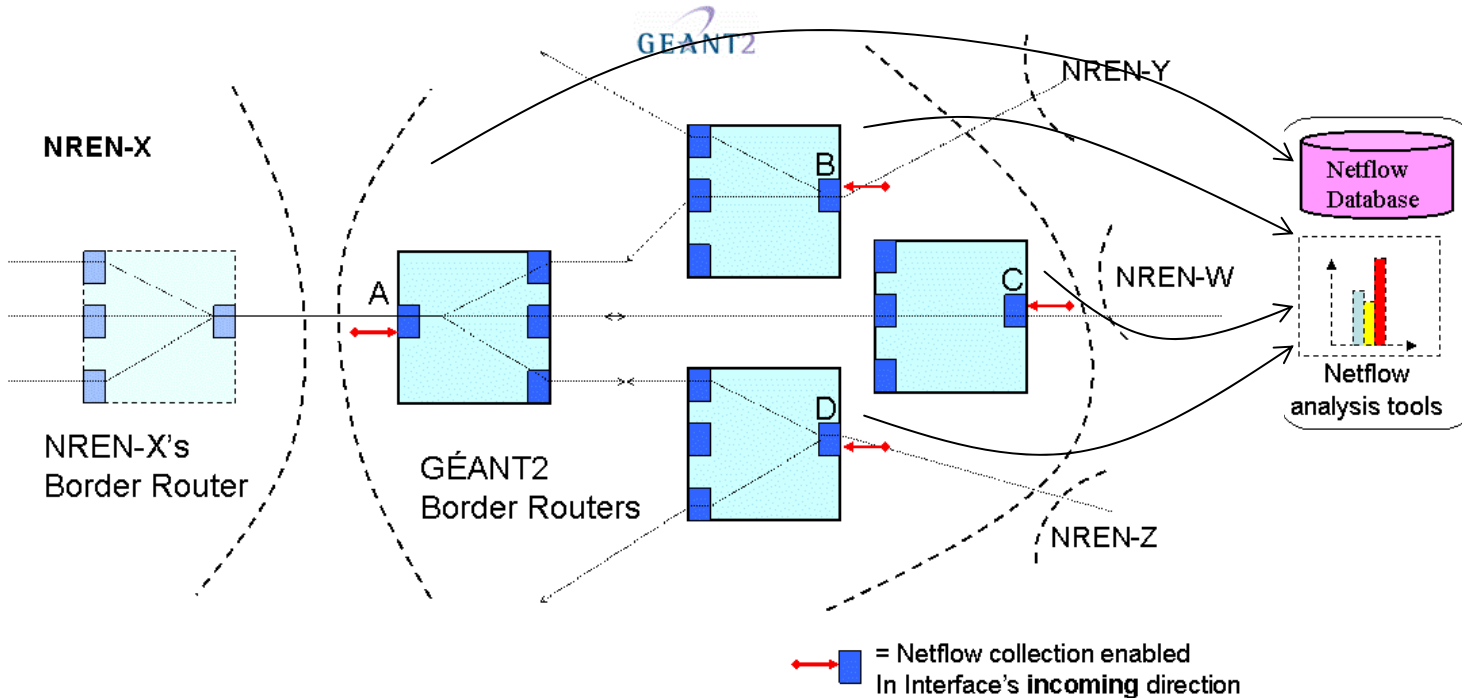
- In GÉANT2, we collect Netflow v5 at every peering point with an external Autonomous System
- We use 1/1000 sampling

- overall handled traffic is 25-30 Gbit/s

- This produces, with 1/1000 pkt sampling, 2-3 *sampled* Kflow/s

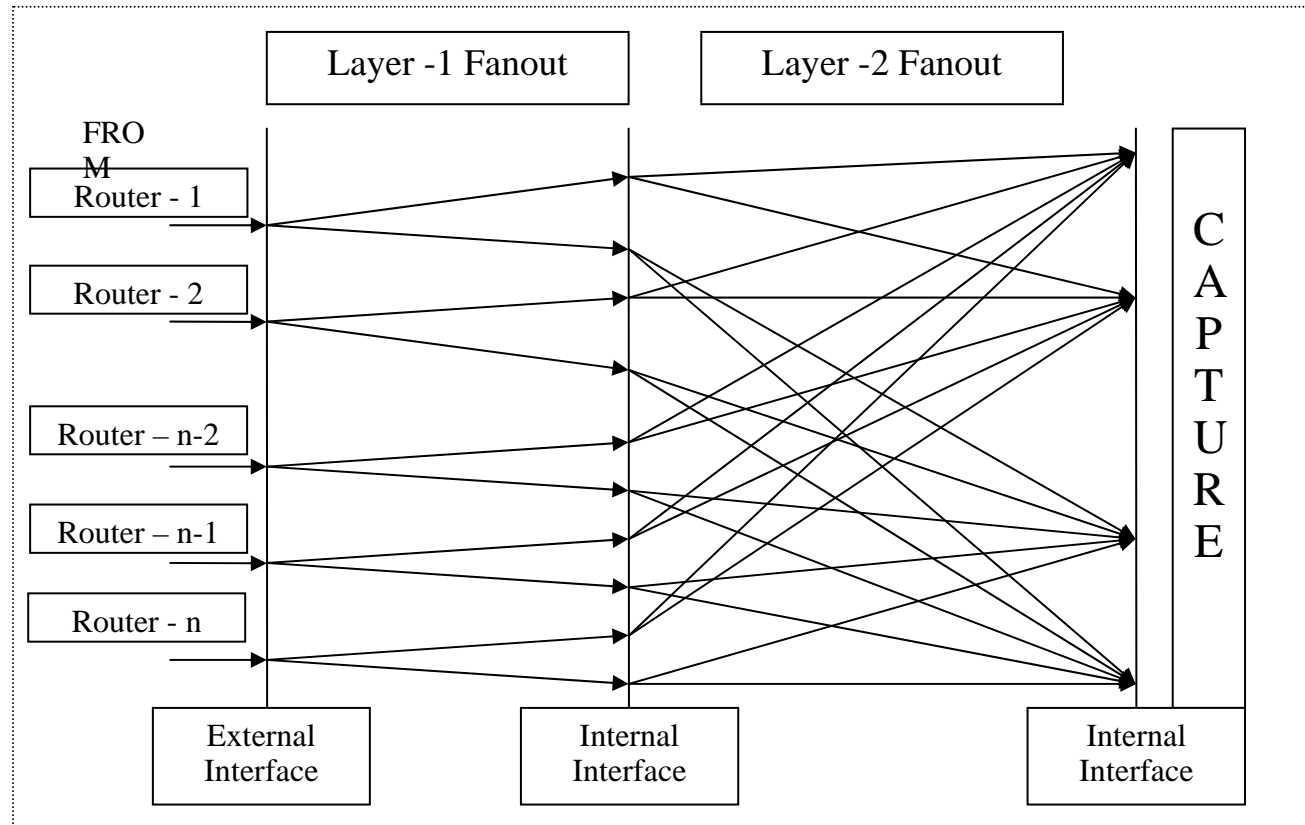
- and an overall Netflow traffic to the collector of 1-2 Mbit/s

- \cong 3Gbytes/day of disk space are needed to store that



Netflow analysis in GÉANT2

- Single collector, two environments
 - Test
 - Production
- Flowtools (flow-fanout) is used to separate environments...
 - (Layer-1 fanout)
- ...and applications
 - (Layer-2 fanout)



Netflow analysis in GÉANT2 (cont.)

- We create a 14-days flowtools archive that researchers can access...
 - ...*after* signing an NDA!
- We create two NfDump archives
 - Test
 - Production
- We look of at the overall traffic
 - NfSen’s “Live” profile
- And at all traffic to/from our NRENs

Netflow analysis in GÉANT2 (cont.)

NFSEN - Profile live Jul 02 2007 - 04:55 - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://62.40.115.20:8090/nfsen-snapshot-20070312/nfsen.php> Search Print

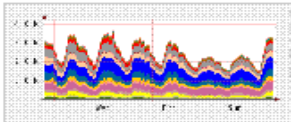
Home Bookmarks mozilla.org mozillaZine mozdev.org

NetFlow Services Solutions Guide NFSEN - Profile live Jul 02 2007 - 04:55

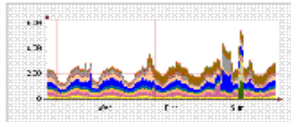
Home Graphs Details Alerts Stats Plugins live [Bookmark URL](#) Profile: live ▾

Profile: live

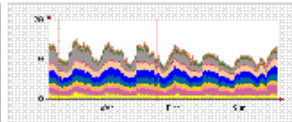
TCP



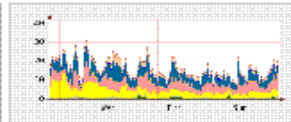
UDP



ICMP



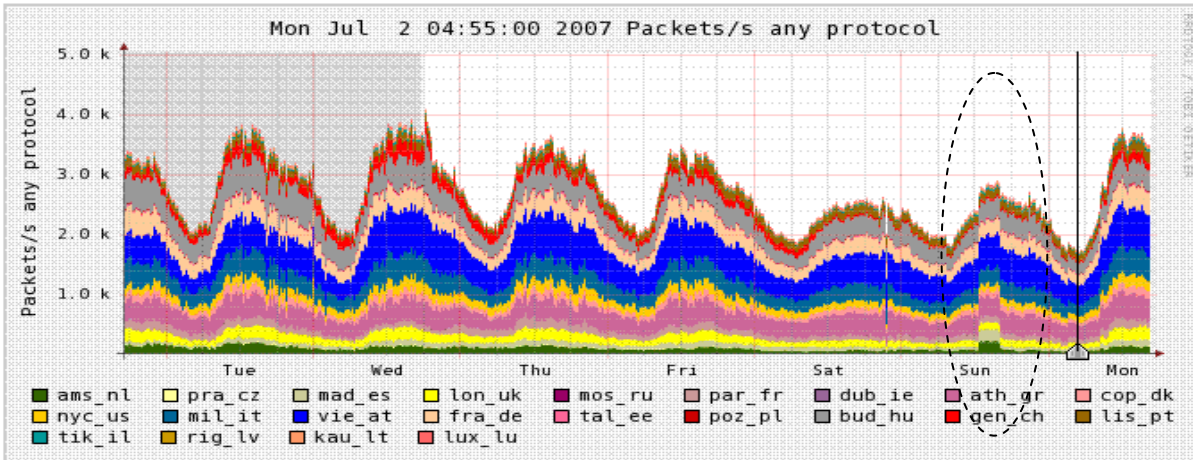
other



Profileinfo:

Type: live
Max: 80.0 GB
Exp: never
Start: Jun 27 2007 - 17:50 BST
End: Jul 02 2007 - 17:55 BST

Mon Jul 2 04:55:00 2007 Packets/s any protocol

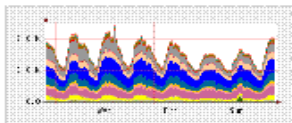


5.0 k
4.0 k
3.0 k
2.0 k
1.0 k

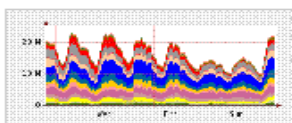
Tue Wed Thu Fri Sat Sun Mon

ams_nl	pra_cz	mad_es	lon_uk	mos_ru	par_fr	dub_ie	ath_gr	cop_dk
nyc_us	mil_it	vie_at	fra_de	tal_ee	poz_pl	bud_hu	gen_ch	lis_pt
tik_il	rig_lv	kau_lt	lux_lu					

Flows



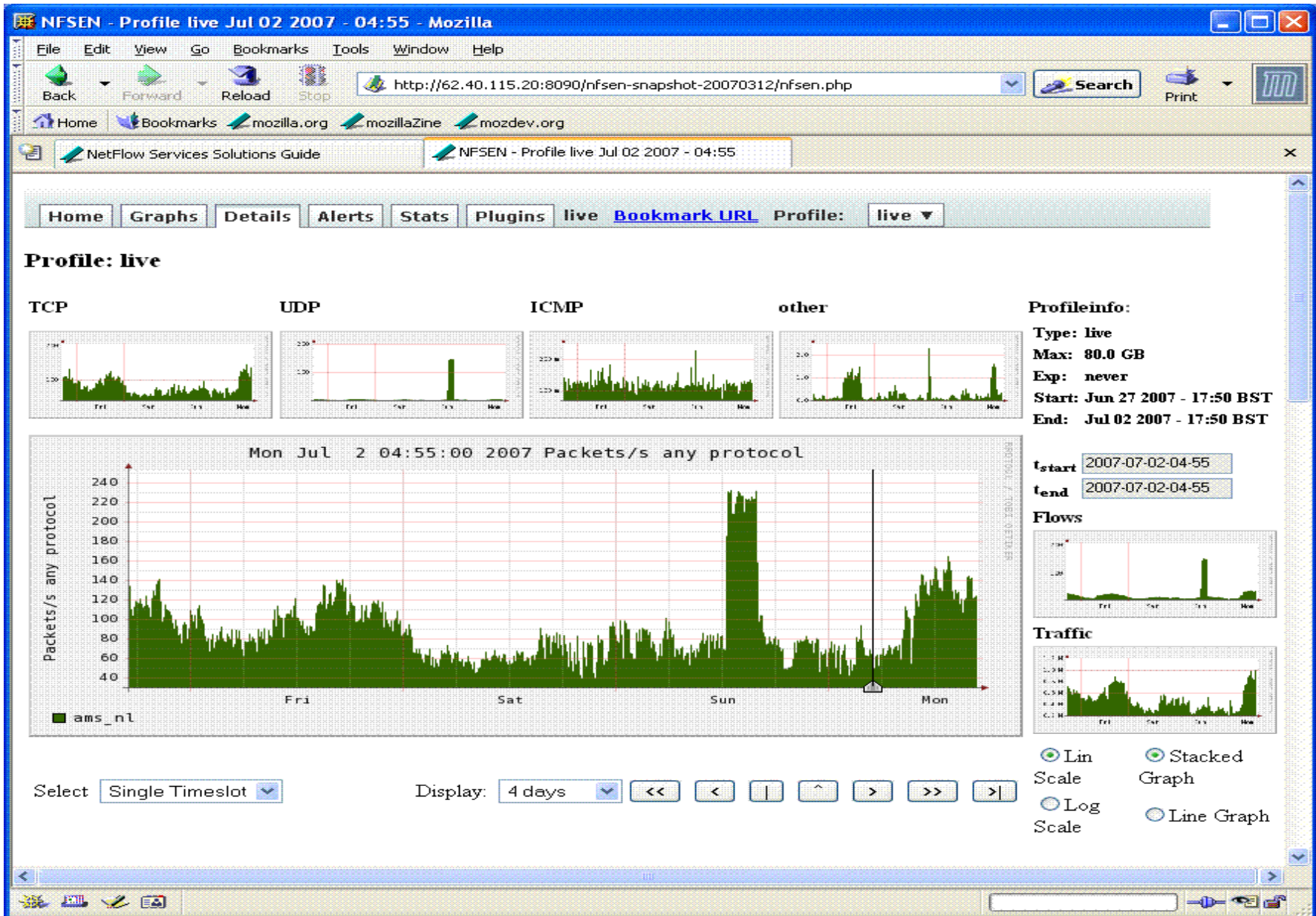
Traffic



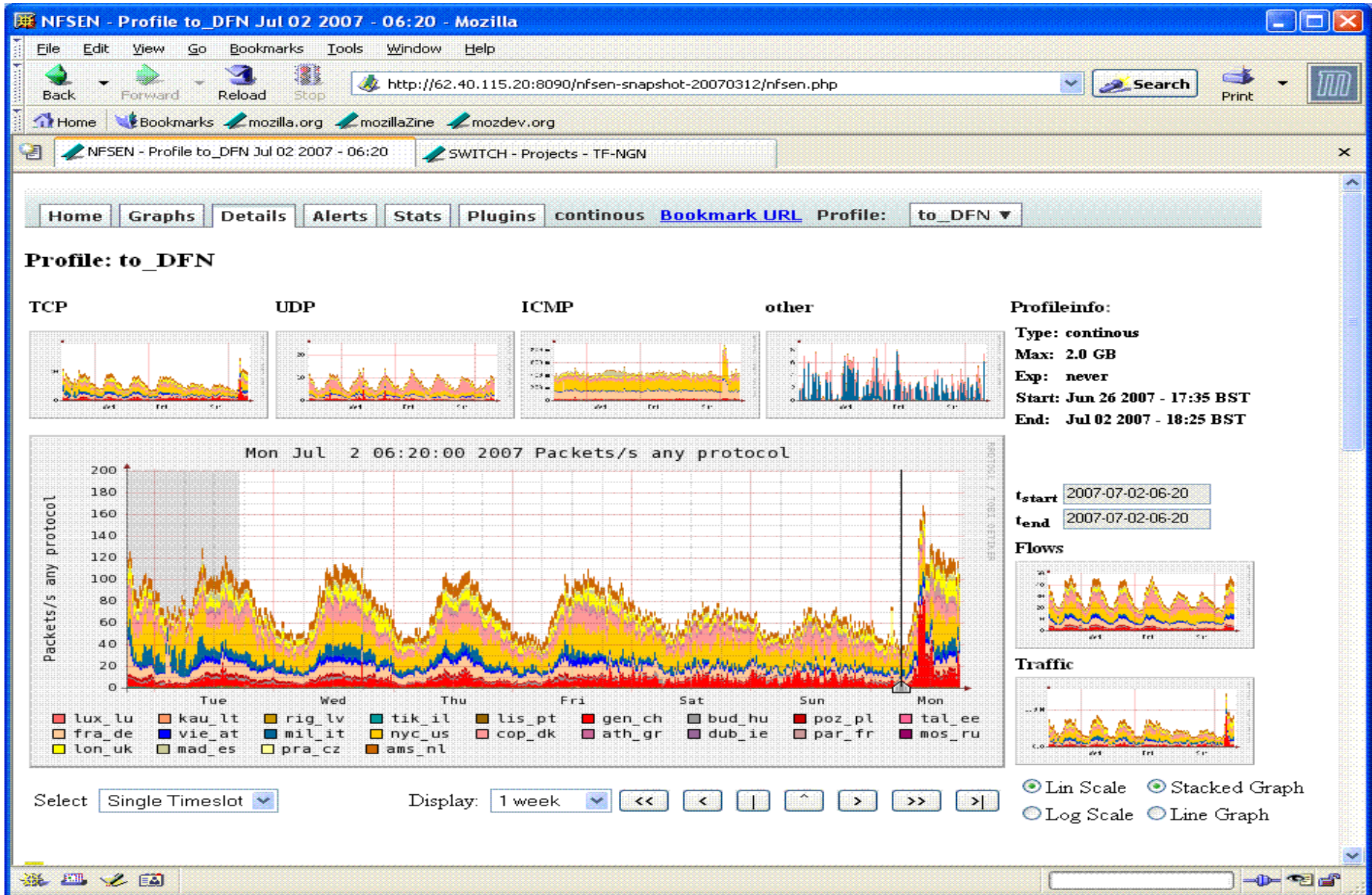
Select Display: 1 week <<< << || >> >>>

Lin Scale Stacked Graph
 Log Scale Line Graph

Netflow analysis in GÉANT2 (cont.)



Netflow analysis in GÉANT2 (cont.)



Netflow analysis in GÉANT2 (cont.)

new query

src AS => dst AS Bit/s matrix - day: 20070702 - time: 13:00-15:00

Click on cells to display time evolution of cell's traffic

TO=>	SURFnet	Abilene	TELIA	GARR	ACOnet	FCCN	HUNGARnet	ARNES	CARnet	RENATER	NORDUNET	RoEduNet	Level3	GRnet	SWITCH	DFN	RedIRIS	JANET	PSNC	Total
SURFnet	0	17.66M	0	16.88M	7.666M	3.888M	10.66M	3.888M	7M	85.22M	31M	1.222M	1.284K	41.88M	76.33M	98.22M	23.22M	83.11M	10.33M	518.2M
Abilene	28.11M	0	361.0K	31.55M	9.222M	6.666M	7.111M	6.555M	3M	46M	34.55M	2.222M	28.04K	8.555M	68.77M	87.33M	22.66M	84.66M	9.444M	456.8M
TELIA	83.69K	6.652K	218.5K	158.6K	15.16K	285.7M	26.85K	374.1M	360.6M	600.8K	69.93K	222.7M	82.35K	748.6M	6.615K	1.570M	269.7K	612.3K	243.8K	1.995G
GARR	13.77M	329.3M	0	520	94.55M	2.222M	1.222M	719.9K	1.444M	93.22M	5.666M	561.4K	93.33	5.444M	93.88M	32.77M	23.77M	16.66M	3.555M	718.8M
ACOnet	1.444M	3.555M	0	3.666M	0	191.2K	22.11M	19M	10.88M	1.444M	2.444M	11.22M	77.77	546.8K	20.44M	17.22M	651.5K	1.444M	831.1	116.2M
FCCN	1.666M	1.333M	388.1M	1.031M	352.3K	0	996.8K	159.5K	192.6K	525.0K	873.2K	31.18K	97.66M	1.121M	821.6K	2.222M	44.44M	1.444M	1.222M	544.2M
HUNGARnet	3M	6M	0	1.444M	6.555M	136.3K	0	2.111M	2.222M	4M	2M	1.131M	1.938K	7.444M	1.111M	5M	1.888M	3.444M	966.8K	48.45M
ARNES	2.111M	2.444M	242.4M	1.333M	2.333M	748.5K	1.104M	606.6	1.222M	413.6K	105.4M	364.6K	772.8M	1.888M	12M	2.444M	1.111M	1.222M	1.666M	1.153G
CARnet	4.333M	2.222M	272.1M	1.444M	833.1K	546.3K	1.111M	2.444M	2.4K	174.7K	3.666M	222.7K	861.4M	5.555M	364.8K	3.333M	319.5K	451.7K	2.333M	1.162G
RENATER	130.7M	43M	0	25.66M	21M	4.111M	1.095M	311.4K	209.7K	0	18.33M	203.4K	0	3M	28.88M	27.88M	40.88M	113M	861.7K	459.2M
NORDUNET	11.33M	24.55M	0	35M	6.333M	1.157M	4.444M	47.77M	4.666M	12.44M	0	2M	251.1	13.77M	19.77M	38.11M	8.888M	38.55M	5.444M	274.2M
RoEduNet	1.666M	1.222M	535.7M	396.7K	987.5K	487.6K	819.3K	2M	977.3K	513.2K	1.333M	0	189M	1.333M	617.9K	1.777M	193.7K	1.125M	3.222M	743.4M
Level3	152.2K	232.1K	14.65K	1.557M	231.0K	290.4M	60.54K	371.5M	316.5M	252.2K	141.5K	208.4M	5.506K	932.7M	28.09K	1.326M	82.02K	103.4K	274.6K	2.124G
GRnet	7.555M	6.111M	485.2M	4.222M	1.023M	1.111M	1.063M	2.555M	4.777M	4M	8.222M	718.0K	1.444G	0	8.111M	4.444M	1.666M	12.66M	3M	2.000G
SWITCH	95.33M	6.777M	368.2K	529M	32.77M	24.55M	1.555M	2.222M	5.222M	68.77M	49.66M	3.333M	31.05K	8.777M	0	42.33M	210.7M	36.44M	8.222M	1.126G
DFN	19.11M	47.77M	1.666K	38.77M	21.22M	2.555M	8.555M	3.222M	2.111M	110.1M	17.44M	968.2K	0	6.777M	103M	0	22.44M	52.22M	19M	475.3M
RedIRIS	25.88M	86.88M	1.728K	16.11M	916.9K	2.666M	1.111M	2.666M	1.222M	33.77M	5.555M	438.4K	0	1.666M	248M	8.666M	0	12M	1.777M	449.3M
JANET	71.88M	26.11M	0	15.77M	11.77M	1.555M	6.333M	1.666M	5.555M	91.33M	12.44M	351.0K	1.666K	9.777M	88.44M	104M	9.444M	0	2.222M	458.6M
PSNC	4.111M	8.111M	0	9.777M	0	2.111M	2.111M	1.444M	2.444M	1.888M	6.888M	336.1K	0	3M	3.111M	15M	3.111M	4.888M	0	68.33M

Part 5/6

- Introduction to IP flows
- IP flow monitoring in routers
- IP flow monitoring exporting standards
- An IP flow collection and analysis example:
GÉANT2
- List of tools for IP flows processing
- Advanced stuff

List of tools for IP flows processing

- <http://www.switch.ch/tf-tant/floma/software.html>
- Long list! – what to do when exploring it?
 - Try to understand the main application the tool targets
 - There is probably no tool good for all application (despite what they will claim..)
 - If freeware, try to understand if there's a user community behind the tool, and/or if somebody will help you in the installation/troubleshooting
 - Try to understand processing & disk space requirements, especially if you have unsampled Netflow data!

Part 6/6

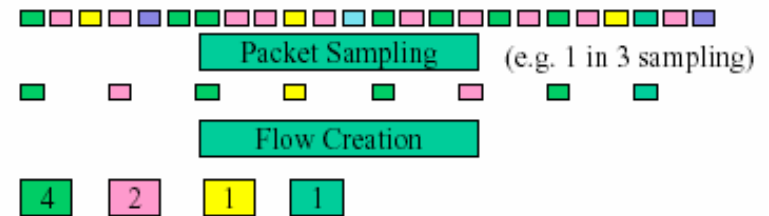
- Introduction to IP flows
- IP flow monitoring in routers
- IP flow monitoring exporting standards
- An IP flow collection and analysis example:
GÉANT2
- List of tools for IP flows processing
- **Advanced stuff**
 - **Sampling**
 - **PSAMP Working Group**
 - **Privacy considerations**

Sampling

- Most routers do deterministic 1:N or random 1:N sampling
 - As long as there are a lot of flows, these two types of sampling are equivalent

- Re-normalization:

- Packets: multiply by N
 - It's an “un-biased” estimator
- Bytes: multiply by N
 - It's correct as long as the sampled packet population well represents the bytes/pkt distribution
- Flows: multiplying by N is *wrong!*
 - No easy and universal formula (afaik)



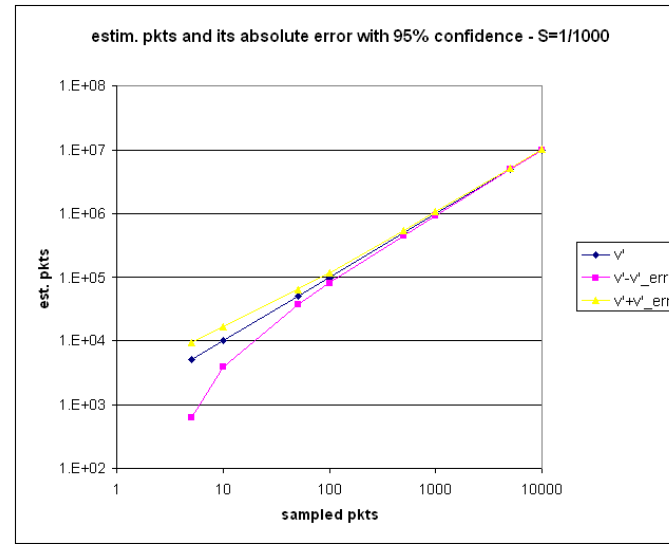
Sampling (cont.)

- The issue is to control the precision of the re-normalization
- Notation:
 - S=sampling rate (e.g. 1/1000)
 - H=true number of packets in a flow
 - h=sampled packets of a flow
 - N=true overall number of packets
 - n=number of overall sampled packets
 - **v'=h/S number of estimated packets in a flow (i.e, estimation of H)**
 - v=same as H (formulas more intuitive)
 - p=H/N true proportion of pkts of a flow in overall pkts
 - p'=h/n estimated proportion of pkts of a flow in overall pkts

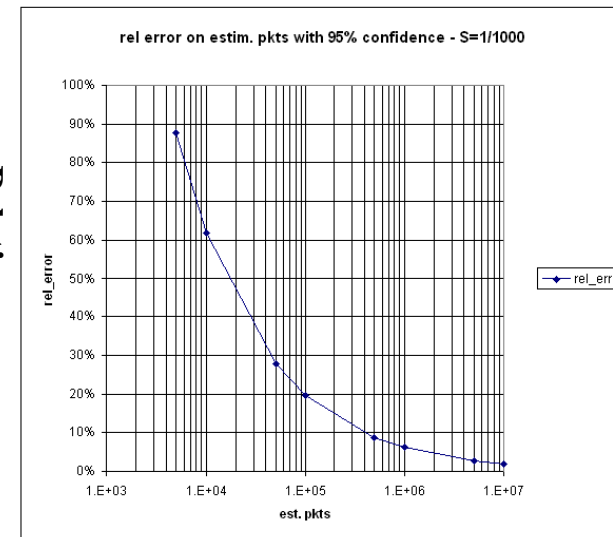
- Result: $v' - \epsilon_v < v < v' + \epsilon_v$, where

$$\epsilon_v = \frac{z_{1-\delta/2}}{S} \sqrt{h(1-p')}$$

- p' is unknown
 - Worst case assumption: p'=0



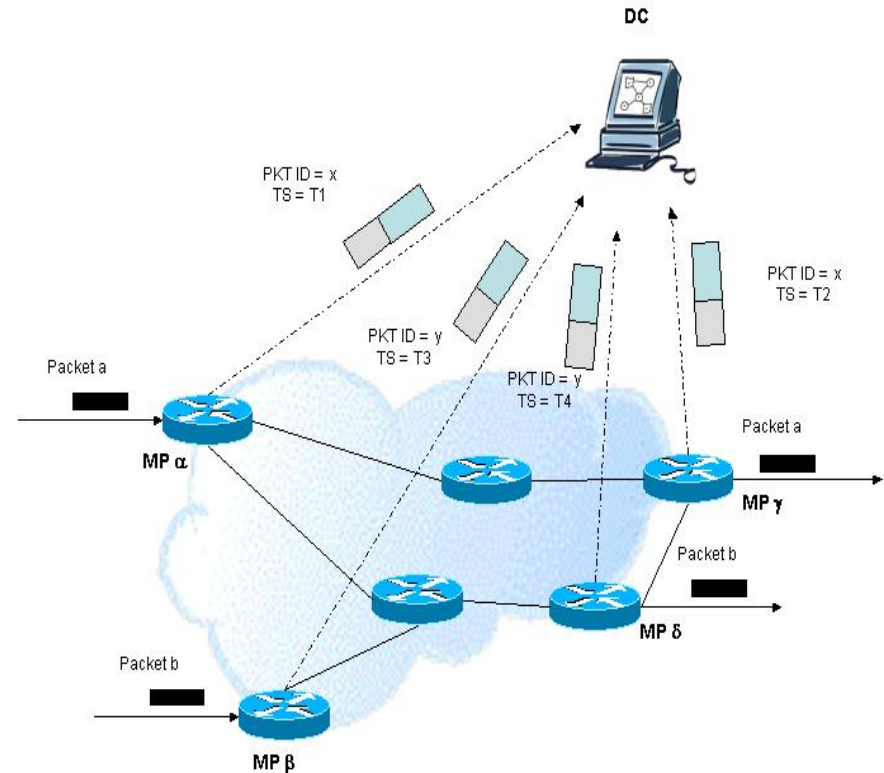
Absolute error



Relative error

PSAMP Working Group

- Focus on sampling packets and on transferring info to data collectors
- Target applications
 - traffic profiling, surveillance
 - monitoring network behaviour
 - Trajectory sampling (hash based sampling needed!)
- Will reuse/extend IPFIX for export
- <http://www.ietf.org/html.charters/psamp-charter.html>



Privacy Considerations

- un-anonymized Netflow data disclose packet headers
 - Who the end user is (provided access to DHCP logs or PTR records)
 - Who she contacts
 - What applications are used (those identifiable by ports)
- Is it a privacy breach? **YES!**
- But EU directives allow privacy breaches if
 - Motivated (e.g. NW operations, fraud prevention)
 - Commensurate to the potential threat
 - Described (what is collected/analysed/archived, for how long..)
- Research is generally harder to justify, but could be done
 - Precisely limiting the scope of the analysis
 - Putting constraints on disclosure, publications of IP addresses, etc.
 - Never dwelling in packet content

References

- <http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/netflsol/nfwhite.htm>
- <http://www.switch.ch/tf-tant/floma/software.html>
- <http://www.ietf.org/html.charters/ipfix-charter.html>

- <http://www.cisco.com/networkers/nw04/presos/docs/NMS-2032.pdf>
- http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/tflow_wp.pdf
- <http://www.splintered.net/sw/flow-tools/>
- <http://nfsen.sourceforge.net/>
- <http://software.uninett.no/stager/>
- www.ntop.org
- <http://silktools.sourceforge.net/>
- <http://www.caida.org/tools/utilities/flowscan/>
- Trajectory Sampling - <http://www.research.att.com/~duffield/papers/DG-TS-ToN.pdf>
- Archived video streaming: Andrew Cormack's presentation starts at minute 52:
http://www.terena.nl/conferences/tnc2005/programme/sessions/show.php?sess_id=93