La métrologie au GIP RENATER

François-Xavier Andreu
(francois-xavier.andreu@renater.fr)

PLAN

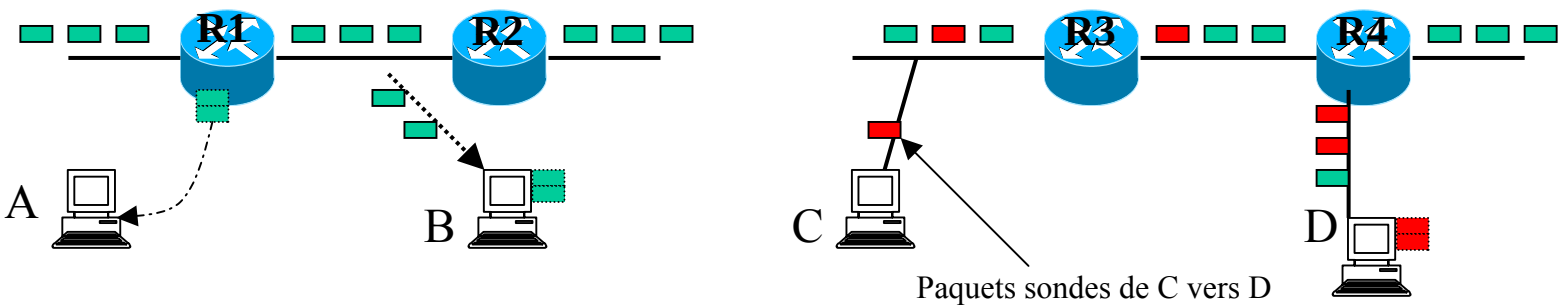
- Introduction, rappel sur la métrologie
- RENATER
- Les besoins élémentaires
- De nouvelles mesures... pour de nouveaux besoins
- Conclusion

PLAN

- Introduction, rappel sur la métrologie
- RENATER
- Les besoins élémentaires
- De nouvelles mesures... pour de nouveaux besoins
- Conclusion

Introduction, rappel sur la métrologie

- Les buts :
  - Supervision
  - Aide au diagnostique
  - Aide au dimensionnement
  - Détection des attaques (sécurité)
- La métrologie passive et active :

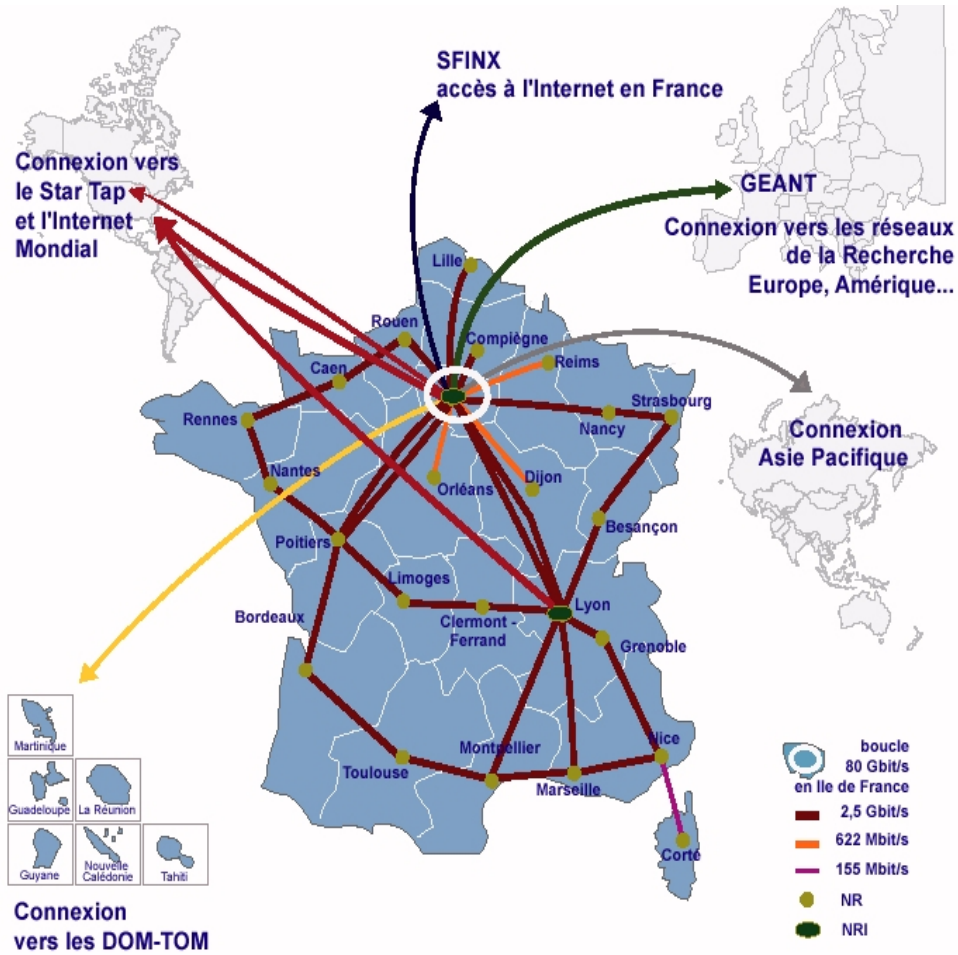


PLAN

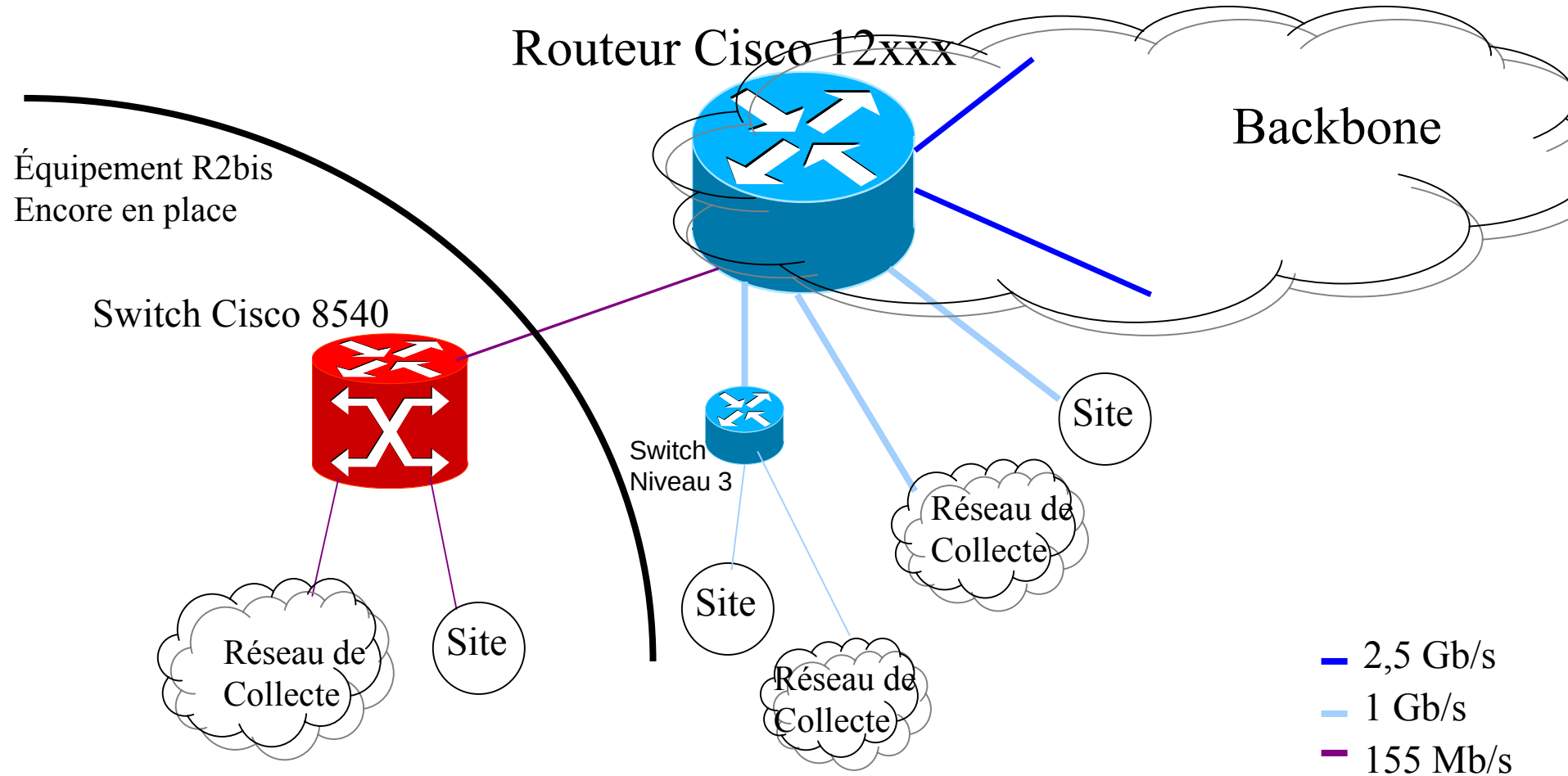
- Introduction, rappel sur la métrologie
- RENATER
- Les besoins élémentaires
- De nouvelles mesures... pour de nouveaux besoins
- Conclusion

RENATER-3

- Les services:
  - IPv4
  - IPv6
  - Allocation d'adresses IP (actuellement 4000 blocs en IPv4)
  - MPLS (service ATOM)
  - Multicast (v4,v6:www.m6bone.net): MBGP, MSDP
  - Les classes de services (cf article JRES2003)



Le "Nœud Régional"



PLAN

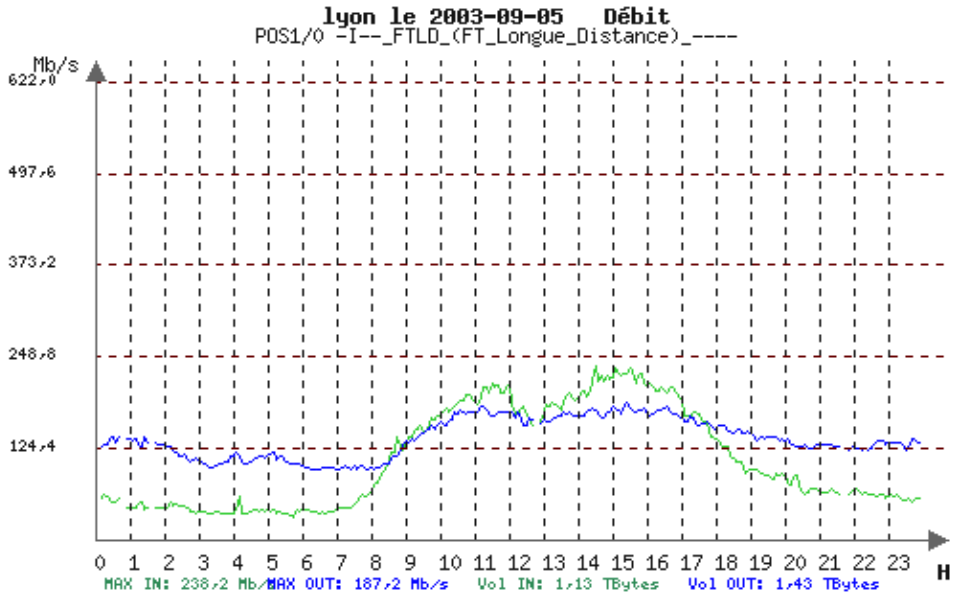
- Introduction, rappel sur la métrologie
- RENATER
- Les besoins élémentaires
  - Les MIBs et SNMP
  - NetFlow
- De nouvelles mesures... pour de nouveaux besoins
- Conclusion

Les MIBs et SNMP (1/3)

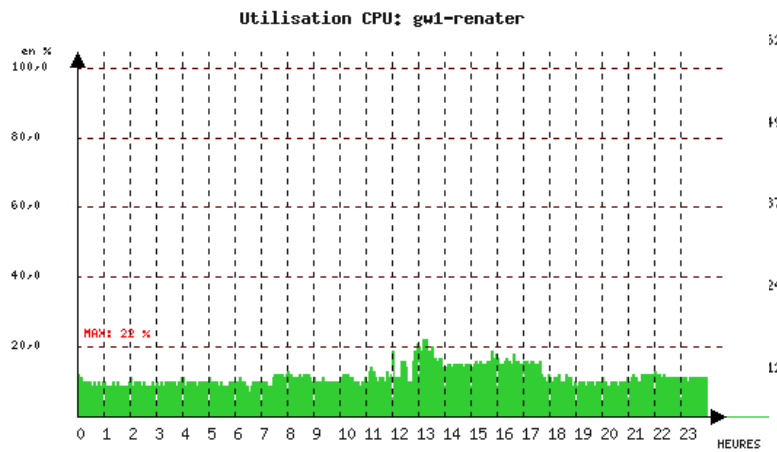
- MIB : bases de données sur les équipements accessibles par le protocole SNMP
- Logiciels libres très répandus : MRTG, Cricket, RTG...
- Caractéristiques de l'outil du GIP (RTG-like):
  - Écrit entièrement en C
  - Base de données MySQL
  - Intervalle d'interrogation paramétrable (possibilité < 1 min)
  - Interrogation asynchrone
    - 10 secondes pour l'ensemble des 80 équipements du réseau
  - Découverte des nouvelles interfaces
  - Minimum 4 oids relevés sur chacune des interfaces (1000*4*2 paquets UDP)
  - UCD-SNMP au début, aujourd'hui Net-SNMP version 5.0.7

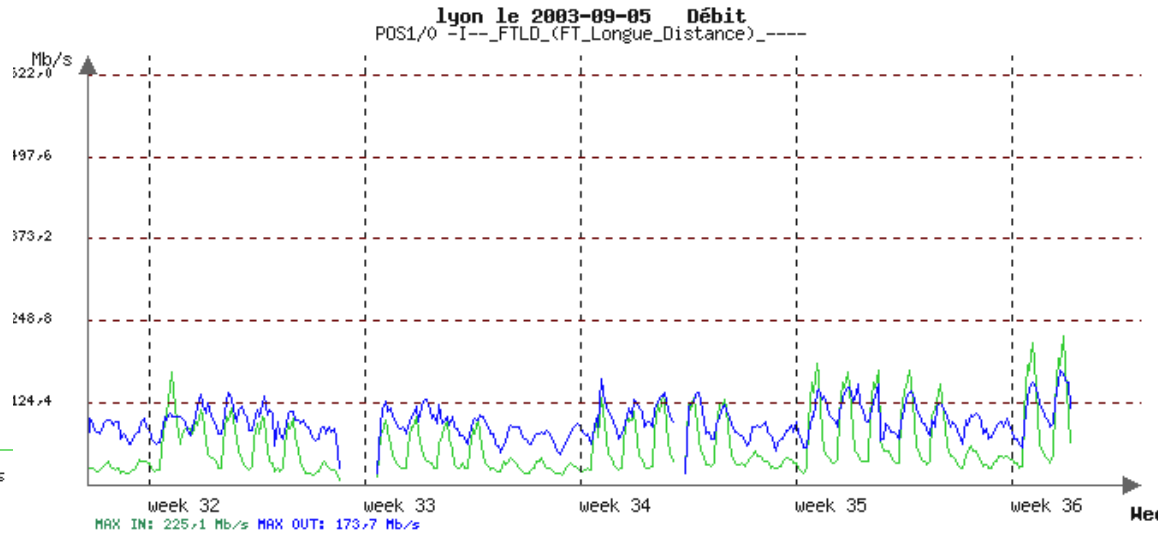
Les MIBs et SNMP (2/3)

- Débits des interfaces:
  - En Mb/s ou en paquets/s
  - Visualisation possible par tranche:
    - De 24 h (jusqu'aux 4 derniers jour
    - Hebdomadaire
    - Semestrielle
    - Annuelle



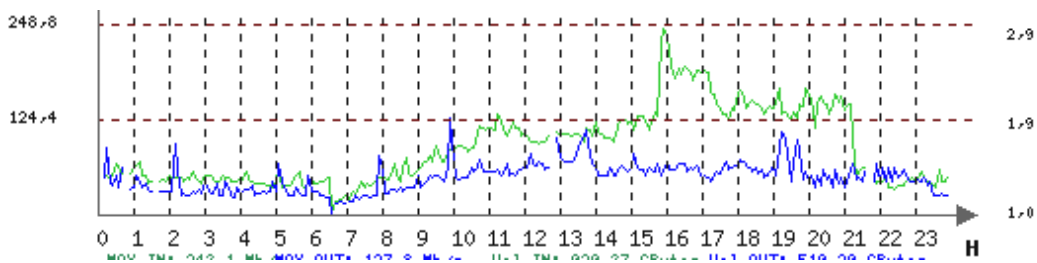
- Charge de la CPU:

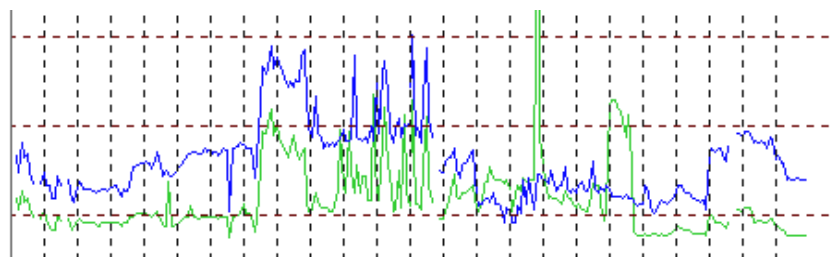




Les MIBs et SNMP (3/3)

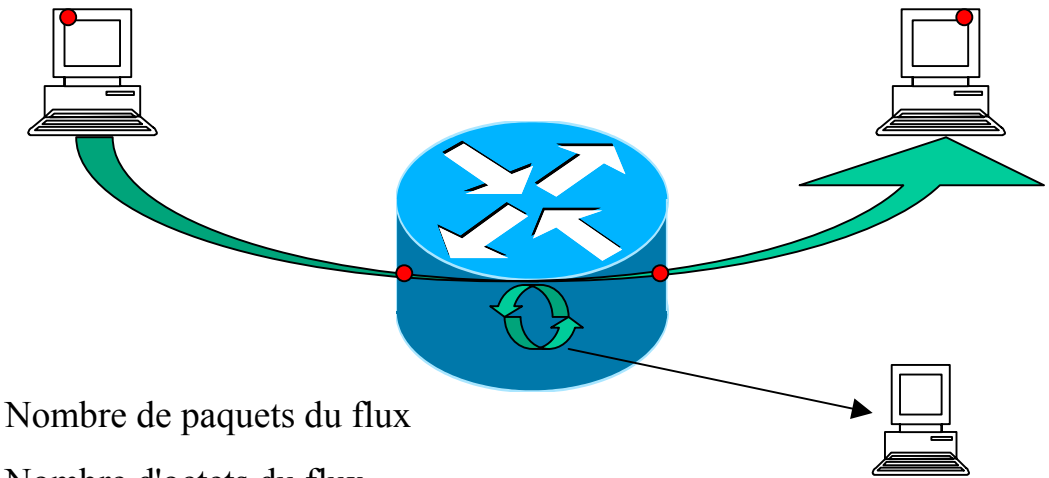
- MIB v6 (en cours)
- MIB MPLS
- Les limites:
  - Les courbes sont lissées (moyenne sur 5 minutes)
  - Vision du trafic IP (IPv4 avec IPv6)
  - Vision couche transport stricte
  - Difficultés de définir des alarmes sur le comportement d'un débit:





NetFlow

- Rappel:



Informations sur un flux:

- IP source
- IP destination
- port source
- port destination
- protocole
- type de service
- index de l'interface

+

Nombre de paquets du flux

Nombre d'octets du flux

temps: début et fin du flux

Index de sortie (interface)

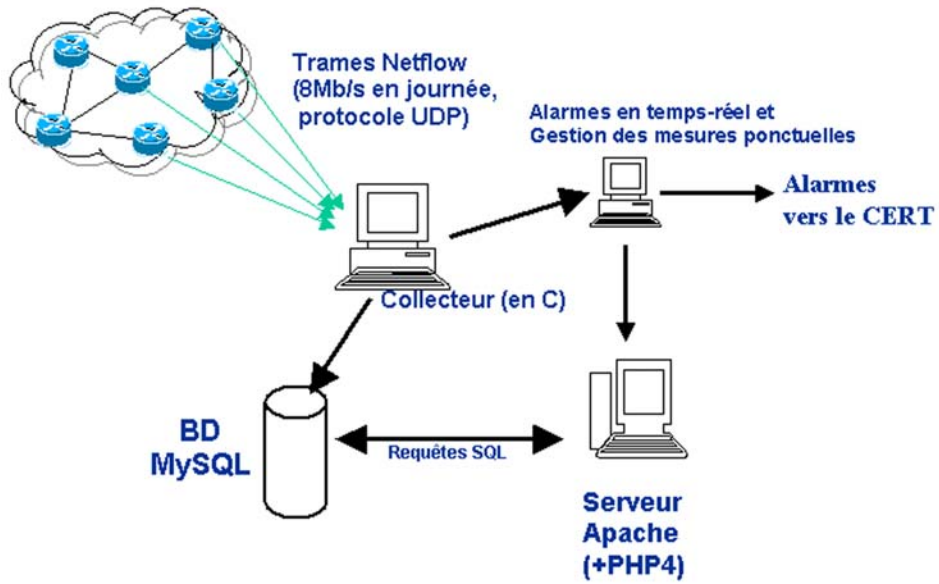
AS source et destination

Masque des IP source et destination

Cumul des flags TCP

NetFlow (suite)

- Architecture :



- Caractéristiques :
  - un débit de 8 Mb/s en journée, 3Mb/s la nuit vers le collecteur.
  - 6 millions de flux / 5 minutes en journée, un minimum de 2 millions la nuit (attention, 80% des équipements sont en mode échantillonnage).

NetFlow (suite)

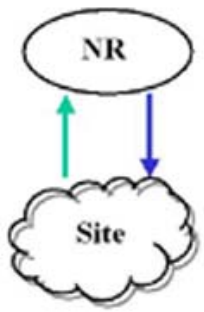
- Exemple d'une connexion sur un serveur web:

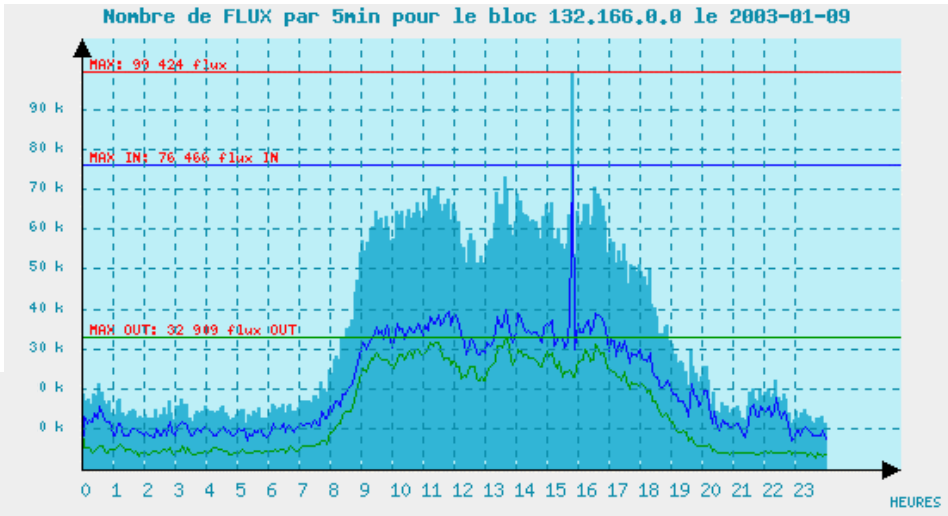
Adresse source	Adresse destination	Routeur	index d'entrée	index de sortie	Port source	Port dest.	Prot	Octets	Paquets	AS source	AS dest.
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1164	80	6	821	10	0	65037
194.199.8.10	193.49.159.141	193.51.179.66	14	16	80	1164	6	14456	14	65037	0
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1165	80	6	544	7	0	65037
194.199.8.10	193.49.159.141	193.51.179.66	14	16	80	1165	6	6582	7	65037	0
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1166	80	6	552	7	0	65037
194.199.8.10	193.49.159.141	193.51.179.66	14	16	80	1166	6	6641	7	65037	0
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1167	80	6	551	7	0	65037
194.199.8.10	193.49.159.141	193.51.179.66	14	16	80	1167	6	6895	8	65037	0
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1168	80	6	755	12	0	65037
194.199.8.10	193.49.159.141	193.51.179.66	14	16	80	1168	6	20203	17	65037	0
193.49.159.141	194.199.8.10	193.51.179.66	16	14	1169	80	6	460	5	0	65037

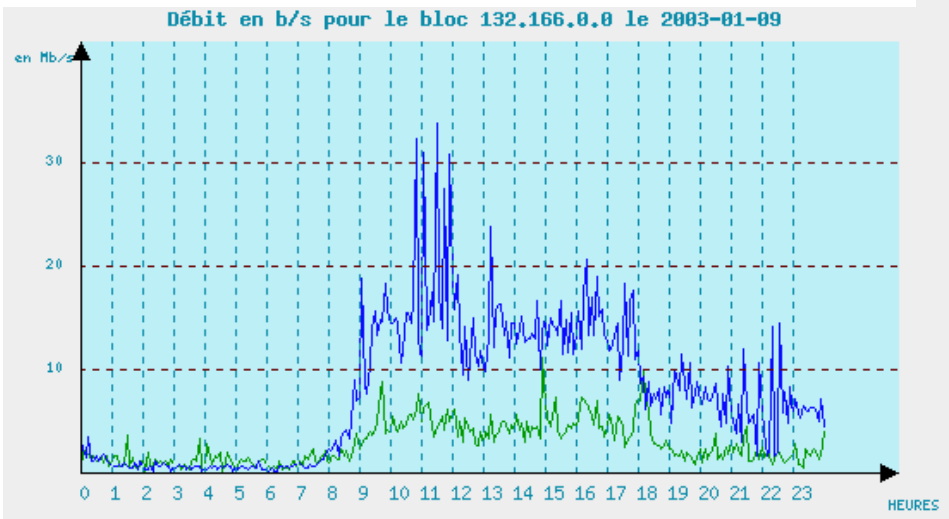
- Sauvegardes des informations :
  - Ponctuelle pour le trafic relatifs à :
    - Un routeur
    - Une interface
    - Une adresse IP
    - Un AS
  - Permanente pour :
    - Un bloc d'adresse (exemple: débit du trafic pour chaque bloc d'adresses)
    - Certains ports

NetFlow (suite)

- Débit d'un bloc d'adresse de RENATER :

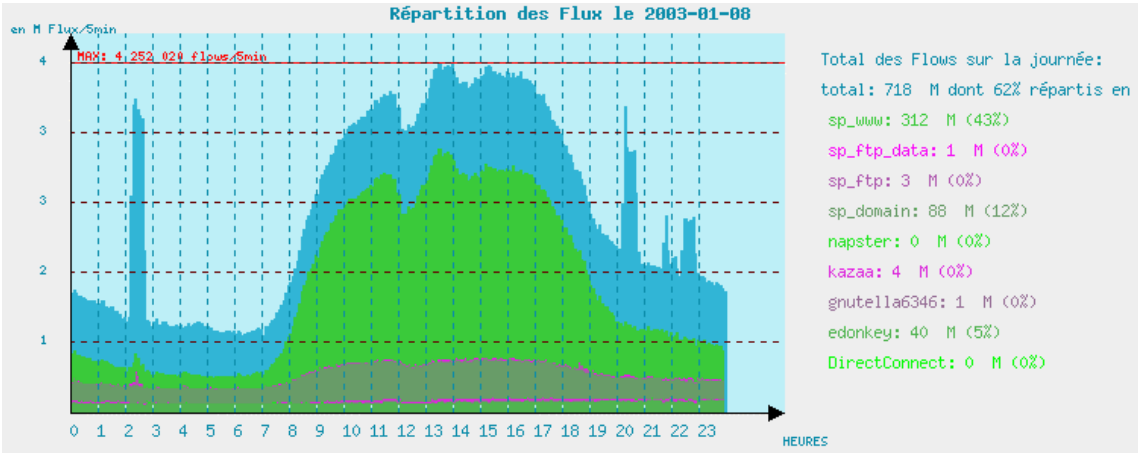






NetFlow (suite)

- Répartition du trafic suivant les ports :

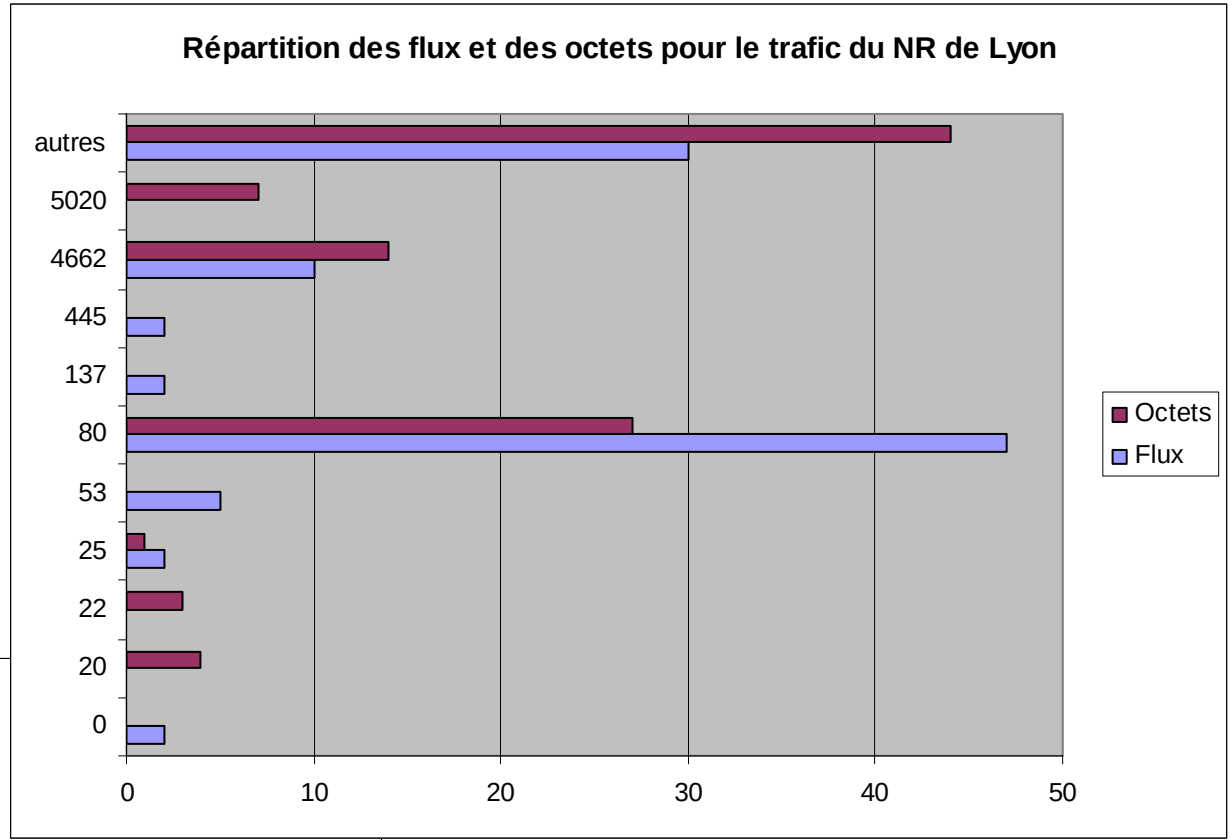


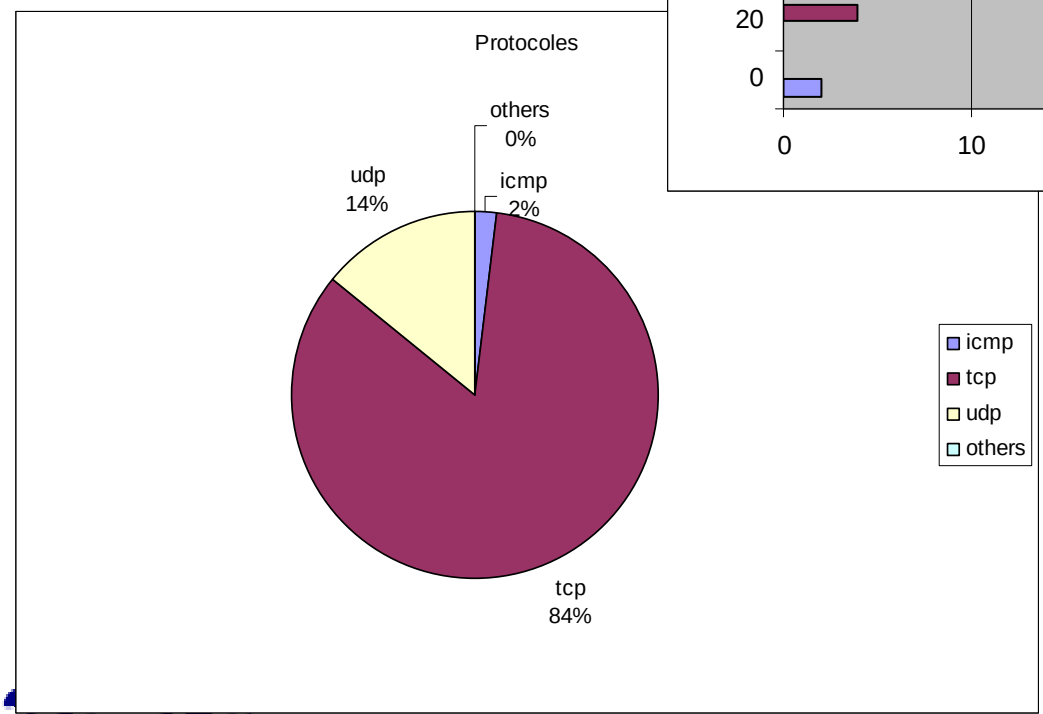
- Alarmes sur les problèmes de routages.
- Matrice de trafic.

- Capture des flux suivant des signatures particulières au niveau du collecteur (ports, nombre de paquets, tailles,…). Génération chaque jour d'un top 40 des adresses ayant un trafic "singulier" (P2P, ftp Warez). Ces informations sont passées au CERT-RENATER pour y être traitées. Depuis Juin 2002, le trafic ne respectant pas la charte RENATER à baisser de moitié.

NetFlow (suite)

Répartition par ports →

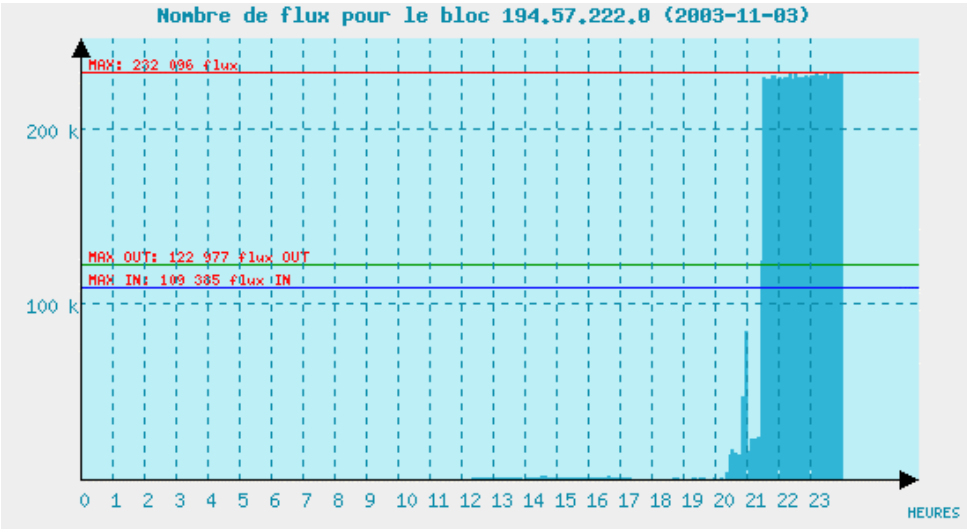




← Répartitions par protocoles

NetFlow (suite)

- Détection d'attaques :
  - Envoie de mail vers le CERT-RENATER lors de dépassement de seuil
  - Extrait des flux :

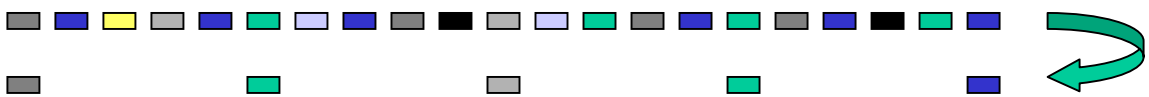


Adresse source	Adresse destination	Routeur	index d'entrée	index de sortie	Port source	Port dest.	Prot	Octets	Paquets	AS source	AS dest.
194.57.222.66	163.15.163.247	193.51.177.42	5	3	1445	80	6	40	1	1715	7539
194.57.222.211	163.15.163.247	193.51.177.42	5	3	1414	63	6	40	1	1715	7539
194.57.222.170	163.15.163.247	193.51.177.42	5	3	1191	53	6	40	1	1715	7539
194.57.222.190	163.15.163.247	193.51.177.42	5	3	1232	34	6	40	1	1715	7539
194.57.222.18	163.15.163.247	193.51.177.42	5	3	1610	25	6	40	1	1715	7539
194.57.222.10	163.15.163.247	193.51.177.42	5	3	1582	23	6	40	1	1715	7539
194.57.222.139	163.15.163.247	193.51.177.42	5	3	1103	1	6	40	1	1715	7539
194.57.222.203	163.15.163.247	193.51.177.42	5	3	1590	116	6	40	1	1715	7539
194.57.222.116	163.15.163.247	193.51.177.42	5	3	1877	113	6	40	1	1715	7539
194.57.222.34	163.15.163.247	193.51.177.42	5	3	1566	98	6	40	1	1715	7539
194.57.222.166	163.15.163.247	193.51.177.42	5	3	1122	90	6	40	1	1715	7539
194.57.222.1	163.15.163.247	193.51.177.42	5	3	1975	86	6	40	1	1715	7539
194.57.222.182	163.15.163.247	193.51.177.42	5	3	1248	82	6	40	1	1715	7539
194.57.222.163	163.15.163.247	193.51.177.42	5	3	1696	70	6	40	1	1715	7539
194.57.222.6	163.15.163.247	193.51.177.42	5	3	1270	62	6	40	1	1715	7539

NetFlow (suite)

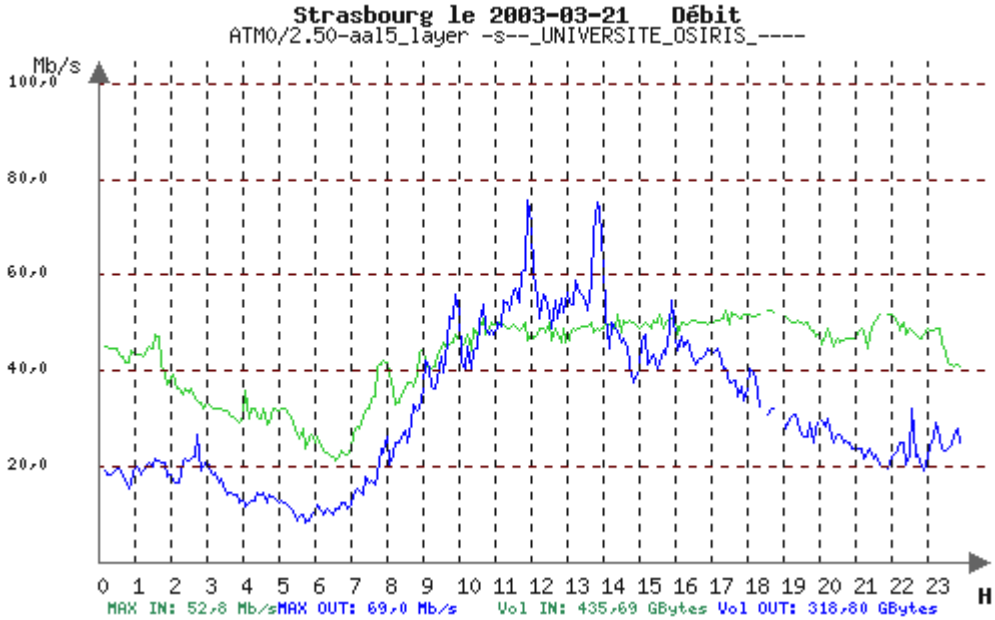
- Une limite :
  - Plus le trafic augmente, plus le traitement au niveau de l'équipement devient un problème.
- La solution:

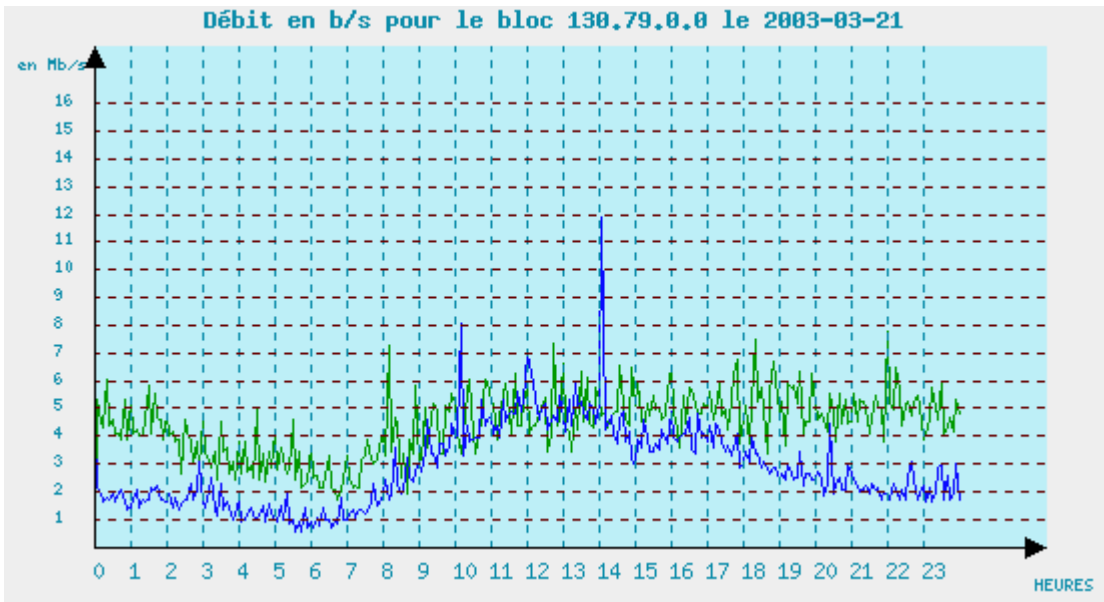
→ L'échantillonnage (Sampled Netflow):



Sur RENATER: 10% des paquets (~1/2 des flux) sauf BdC et DOM-TOM : 100%

Attention
à
l'interprétation





NetFlow (suite et fin)

- Nouveaux formats de transport (v9):
  - Utilisation de "templates"
  - Intégration de nouveaux protocoles:
    - IPv6
    - Multicast
    - MPLS
- Netflow Egress
- Différents modes d'échantillonnage :
  - Déterministe
  - Aléatoire

PLAN

- Introduction, rappel sur la métrologie
- RENATER
- Les besoins élémentaires
- De nouvelles mesures... pour de nouveaux besoins
  - METROPOLIS
  - Les Métriques
  - Exigences des applications
  - Ordre de grandeur sur RENATER
  - Problématiques
  - Solutions existantes
  - Première solution : re-spécification des besoins
- Conclusion

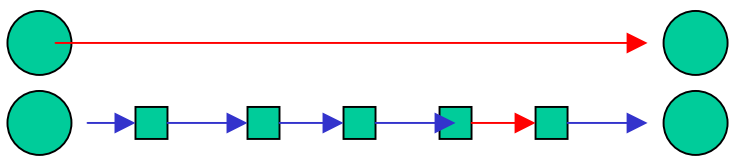
METROPOLIS

- Projet RNRT exploratoire (Métrologie pour l'Internet et les Services)
- Participants :

- Plate-forme de mesures actives et passives
- Thèmes abordés:
  - Classification et dimensionnement
  - Analyse du réseau
  - Méthodes pour la mesure et échantillonnage
  - Modélisation
  - Tarifications et SLAs

Les métriques (1/2)

- Standardisées par l'IETF :
  - groupe IPPM (IP Performance Groups)
  - délais, gigues, pertes, déséquencement...
- Pour chaque classe de service et type de protocole IP
- Notion de précision importante
- Mesures de bout en bout mais aussi en interne :



Les métriques (2/2)

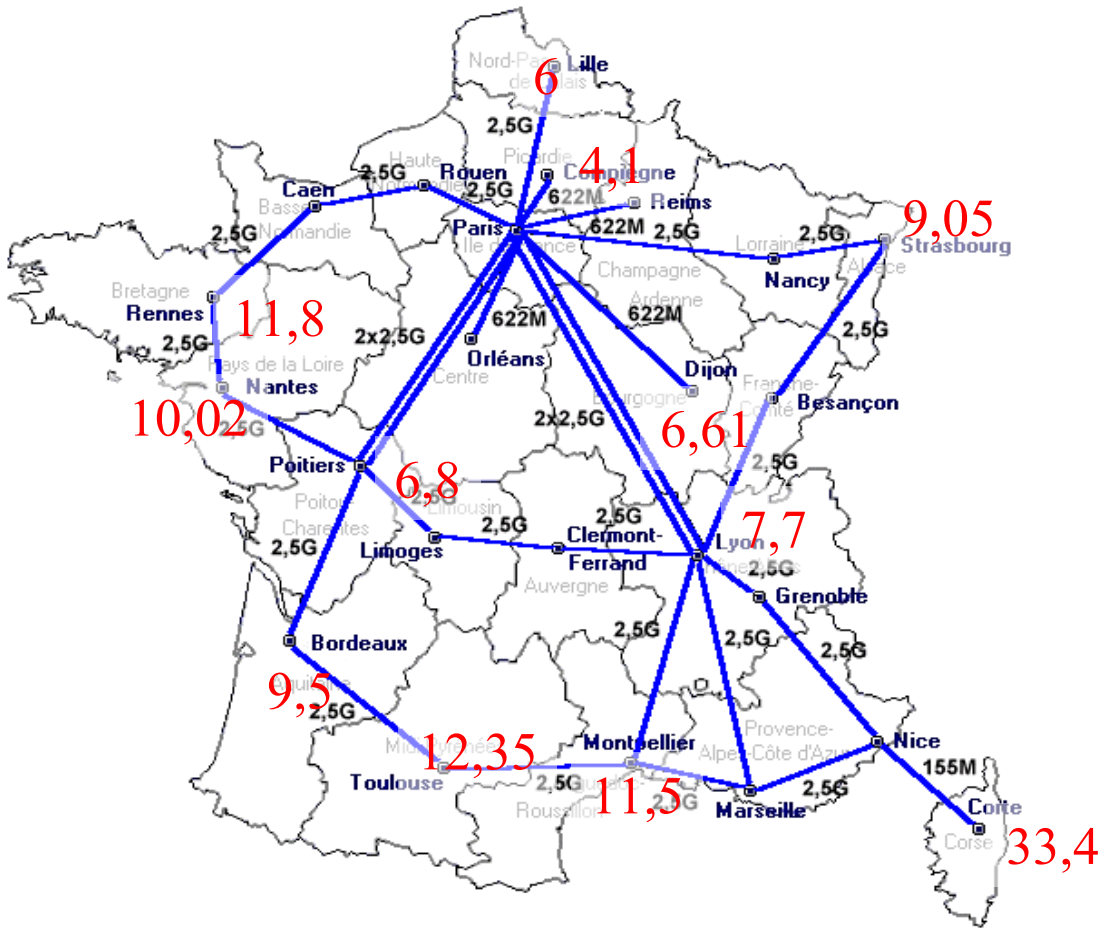
- Délai unidirectionnel (cf RFC 2679) :
  - importante pour les applications temps réels (VoIP, …)
  - permet la mesure de la qualité perçue par l'utilisateur (ex : qualité d'une application sur TCP dépend du délai d'arrivée des paquets et non des acquittements) → plus représentatif que le RTT
  - permet la surveillance du réseau et de l'efficacité de l'implémentation des classes de service en distinguant les trajets aller et retour
- Gigue (variation du délai unidirectionnel, cf RFC 3393) :
  - dû aux variations des files d'attente des routeurs et des trajets
  - importante dans les applications de type streaming (dimensionnement des buffers)
  - caractéristique de la dynamique et de la stabilité du réseau
- Taux de pertes unidirectionnel des paquets (cf RFC 2680) :
  - important pour tous les types d'application
  - dû à l'encombrement du réseau
- Déséquencement (cf draft-ietf-ippm-reordering) :
  - paquet déséquencé = paquet en retard (cf IPPM)
  - importante dans les applications de type streaming (paquet trop en retard = paquet perdu)
  - dû à l'existence de plusieurs chemins dans le réseau, à un protocole de retransmission sur erreur, …

Exigences des applications

- Applications temps réels (voix sur IP, visioconférence) : les plus exigeantes → délai < 150 ms, gigue < 50 ms, perte modéré, déséquencement faible
- Applications type streaming : existence d'un buffer → contraintes limitées sur le délai unidirectionnel et la gigue. Perte et déséquencement modérés
- Web, mail : délai modéré et perte faible, gigue et déséquencement sans grande importance
- Tansfert de données : délai, gigue et déséquencement sans grande importance, pertes faibles
- Applications interactives (telnet, jeux en ligne) et transactionnel : délais, déséquencement et pertes faibles

Ordres de grandeurs sur le backbone

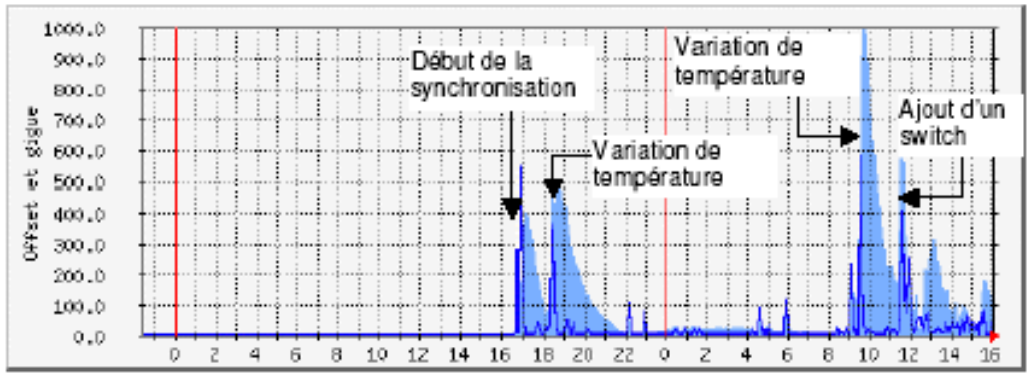
- Délais ≈ quelques ms
- Gigue ≈ ms
- Pertes << 10^{-3}
- Déséquencement très faib



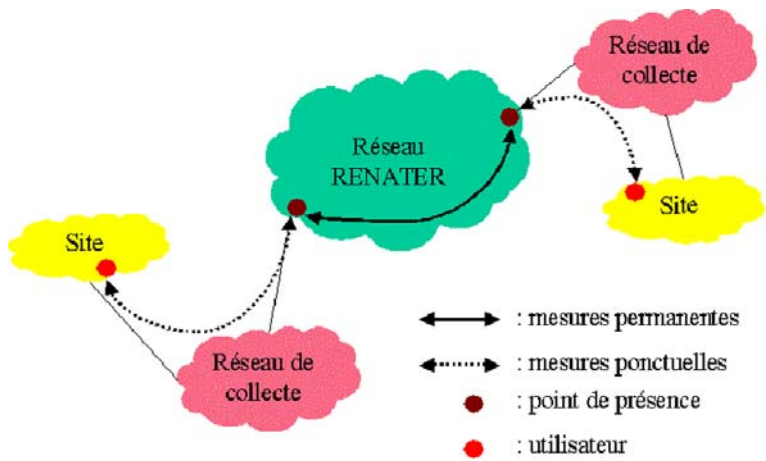
Problématiques

- Précision
  - délai ≈ quelques ms

    → précision ≈ 100 µs
- Synchronisation :
  - NTP :
    - En WAN > 1 ms
    - instabilité
  - GPS :
    - 10 µs
    - Problème de coût et d'installation
- Stabilité des OS
- Mesures de bout en bout :
  - Présence de sondes aux extrémités
  - Décomposition des mesures



Délais unidirectionnels entre deux stations A et B directement reliée, A est synchronisée sur B par NTP, B synchronisée sur son horloge locale.



Les solutions existantes

	SAA	RIPE TIM	Saturne	Rude/Crude	NIMI	QOS Metrix
Matériel	routeur	boîtier (pc/FreeBSD)	PC (free BSD)	PC (linux)	PC (divers OS)	boîtier
GPS intégré	non	oui	non	non	non	Selon modèle
Sécurité	Authentification sondes par MD5 possible	Aucune	Aucune	Paquets numérotés	Chiffrement	Intégrée
Réception des paquets sondes	Non précisée	Capture (pcap) + socket	Capture (bpf)	Socket		matérielle
Estampillage	Non précisé	Avant envoi sur socket	Kernel	Avant envoi sur socket	Dépend de l'outil utilisé	matériel
Paramétrage du DSCP des paquets sondes	oui	non	oui	oui		oui
Collecte	Distante, par snmp	Logs locaux envoyés au site central quotidiennement	Logs locaux, envoi au site central par rpc de chaque mesure	Logs locaux	Logs locaux	Vers site central en temps réel
Présentation des résultats	Non intégrée	Très complète, sur serveur web local et central (à Amsterdam)	Non intégrée	Non intégrée	Non intégrée	Site web central
Modification/ adaptation de l'outil	Impossible	Possible (mais nécessite demande à RIPE)	Possible	Possible	Possible (mais code source imposant)	Impossible
Licence	commerciale	commerciale	non définie actuellement	GPL	GPL	commerciale

Re-spécification des besoins et premières solutions

- Une représentation simplifiée de la qualité du backbone pour l'ensemble de la communauté
  → weathermap des délais sur le backbone (logiciel maison)
- Outils de diagnostique avec une vision de bout en bout
  → quelques sondes fixes sur le backbone et des sondes mobiles pour un usage ponctuel (faible incertitude de la mesure)
- Supervision des classes de services
  → encore quelques questions : permanente, très fine. Architecture précédente suffisante ?

Conclusion

- La métrologie au GIP RENATER se décompose en :
  - une évolution de l'existant
  - une réponse aux nouveaux besoins :
    - De l'exploitation (ex: supervision des classes de services)
    - Des utilisateurs (ex: mesures de bout en bout)