Outils pour quantifier et qualifier le trafic IP

yves Autran CNRS/UREC

La difficulté pour obtenir des chiffres détaillés concernant le trafic des datagrammes IP, entre un site et l'Internet, en utilisant les logiciels disponibles sur le marché, nous a conduit à envisager la création d'un outil logiciel capable de fournir des données permettant d'effectuer des statistiques ou plus exactement des comptages.

L'axe de recherche défini vise à obtenir de manière aisée et au moindre coût suffisamment de données sur les datagrammes IP pour pouvoir élaborer des courbes ou des tableaux, permettant de quantifier et de qualifier le flux des datagrammes.

Les données recueillies doivent permettre de connaître les informations relatives à la destinations (national, international), à l'identité de la station émettrice ou réceptrice, au prestataire de connectivité, et bien sûr au volume de l'échange de données. En outre, pour permettre d'évaluer les besoins infrastructurels, il est impératif de classer ces flux par type de protocole de niveau transport (tcp, udp, etc) et par service (numéro de port).

Bien évidemment cet outil conçu à l'origine pour analyser le trafic entre le site et l'Internet doit pouvoir être utilisé en n'importe quel point du réseau interne.

L'analyse des réponses possibles à un tel cahier des charges nous révèle trois voies. Deux d'entre elles s'appuient sur SNMP (**S**imple Network **M**angement **P**rotocol) et ont en commun d'être d'un coût de mise en service raisonnable et d'utiliser des outils de communication qui sont standardisés. La première concerne les objets de la MIB (**M**anagement **I**nformations **B**ase) spécifique au constructeur. La seconde utilise les objets de la MIB spécifique aux sondes RMON (**R**emote **MON**itoring). La troisième voie, plus difficile à mettre en oeuvre, consiste à développer un logiciel d'acquisition. Elle se révelera être la seule manière possible de parvenir à créer cet outil.

- *Exploitation de l'"accounting" du routeur d'entrée,*

  Cette solution s'appuie sur les objets de la MIB propriétaire du routeur Cisco dans le cas de notre étude. Mais les autres constructeurs de routeurs doivent aussi proposer des objets équivalents dans leurs MIB propriétaires. Les routeurs offrent en standard des possibilités de récupération d'informations concernant les adresses IP et le nombre d'octets et de paquets qui ont transités dans le routeur. Ces informations sont déja exploitées par nombre d'administrateurs de réseaux dans des langages tels que

PERL, shell ... avec des outils SNMP tels que Tricklet, le kit SNMP de Carnegie Mellon University ou celui du Massachuset Institute of Technologies pour récupérer ces données.

L'insuffisance de détail sur les informations collectées (comptabilité limitée au couple adresse source adresse destination et nombre d'octets et de paquets émis relatif à ce couple d'adresses IP) nous a conduit à rejeter cette approche.

- *Exploitation d'une sonde RMON,*

Les sondes RMON peuvent stocker en tampon de capture les parties d'une trame issues d'un masque de filtrage que l'on a défini dans le groupe de la Rmon MIB 'filter'.

Le principe consiste à récupérer la partie intéressante de chaque trame (Le champ type de la trame, les 20 octets de l'entête IP, les octets de numéro de port de l'entête de niveau 4 Udp/Tcp ...) pour l'analyser et en tirer les informations qui nous permettrons de procéder au classement.

Le temps important mis pour rapatrier les trames collectées, la pollution trés importante du réseau que cette opération génère et l'impossibilité de configurer la sonde pour effectuer un premier classement des trames capturées nous ont conduit, après plusieurs essais, à éliminer cette solution.

- *Création d'un outil de collecte des informations,*

Cet outil est composé d'un programme d'acquisition d'un logiciel d'exploitation et de modules d'extraction.

- Le programme d'acquisition.

Le programme d'acquisition, qui tourne en tâche de fond sur une machine de préférence dédiée, capture les trames Ethernet. Il en extrait le datagramme IP et le classe dans des tables dédiées aux protocoles étudiés (IP, TCP, UDP, ICMP). La table principale, qui sert de référence aux autres tables collecte les informations du protocole IP (protocole réseau). Elle contient les adresses d'origine et de destination qui sont contenues dans l'entête de chaque datagramme IP.

- Le programme d'exploitation.

Le programme d'exploitation récupère périodiquement ces tables par télécommunications. Il identifie les adresses contenues dans la table IP en interrogeant d'une part, la base européenne RIPE (**R**éseau **IP** **E**uropéen) pour obtenir le nom du pays, le fournisseur de connectivité, l'adresse de l'administrateur, et d'autre part le

service DNS (Domaine Name Service) pour obtenir le nom symbolique attaché à l'adresse numérique IP.

Ces opérations achevées les informations recueillies dans les tables spécifiques au protocole sont alors stockées dans le fichier de collecte.

Ce fichier sera exploité par des outils de mise en forme des données (présentation de tableaux, listes). Pour permettre à l'administrateur de construire des états à sa convenance, des programmes d'extraction de données dont le rôle est d'extraire des ensembles de données dans le fichier de collecte et de les présenter d'une manière qui soit exploitable par les outils les plus connus (PERL, shell etc ...) ont été ajoutés.