TABLE DES MATIERES

NOTATIONS

1.1 Minuscules latines

c_i	Classe
c_k	Classe
d	Nombre entier
e	Nombre entier
d	Nombre entier
$g(x_i,pa_i)$	Fonction générateur du réseau de Bayes
n	Nombre entier plus grand que e
p	Nombre premier au hasard
pa_i	Parent immédiat
q	Nombre premier au hasard
qi	Nombre possible d'instanciations pour pa_i
r_i	Nombre possible d'instanciations
t	Nombre entier
x_i	Variable caractéristique
X_j	Variable caractéristique

1.2 Majuscules latines

C	Message chiffré
M	Message claire
N	Nombre de cas dans la base de données
N_{ij}	Somme du N_{ijk} pour toutes valeurs de k
N_{ijk}	Nombre de cas pour qui x_i prend les valeurs x_{ik}
$P(X/x_i)$	Théorème de Bayes
X	Variable de classe

LISTE DES ABREVIATIONS

ACL	: Access Control List
ADSL	: Asymmetric Digital Subscriber Line
AFRINIC	: African Network Information Centre
ANSI	: American National Standards Institute
APIPA	: Automatic Private Internet Protocol Addressing
APNIC	: Asia Pacific Network Information Centre
ARIN	: American Registry for Internet Numbers
ARP	: Address Resolution Protocol
ATM	: Asynchronous Transfer Mode
BECN	: Backward Explicit Congestion Notification
BGP	: Border Gateway Protocol
BIOS	: Basic Input/Output System
CBAC	: Context Based Access Control
CCITT	: Comite Consultatif International Telegraphique et Telephonique
CCNA	: Cisco Certified Network Associate
CIDR	: Classless Inter-Domain Routing
CPU	: Central Processing Unit
C/R	: Command/Response
DARPA	: Defense Advanced Research Projects Agency
DCE	: Data Circuit terminating Equipment, Data Communication Equipment
DE	: Discard Eligibility
DHCP	: Dynamic Host Configuration Protocol
DLCI	: Data Link Connection Identifier
DNS	: Domain Name Service
DoS	: Deny of Service
DTE	: Data Terminal Equipment
EA	: Expanded Adress
EIGRP	: Enhanced Interior Gateway Routing Protocol
ESPA	: Ecole Supérieure Polytechnique d'Antananarivo
FECN	: Forward Explicit Congestion Notification

FR : Frame Relay
FTP : File Transfert Protocol
HDLC : High-level Data Link Control
HTTP : HyperText Transfer Protocol
HTTPS : HyperText Transfer Protocol Secure
IANA : Internet Assigned Numbers Authority
ICMP : Internet Control Message Protocol
IDS : Intrusion Detection System
IETF : Internet Engineering Task Force
IOS : Internetworks Operating System
IP : Internet Protocol
IPS : Intrusion Prevention System
ISDN : Integrated Services Digital Network
LACNIC : Latin America and Caribbean Network Information Centre
LAN : Local Area Networks
MAC : Media Access Control
MD5 : Message Digest 5
NAT : Network Address Translation
NVRAM : Non-Volatile Random Access Memory
OSI : International Standards Organization
OSPF : Open Shortest Path First
PDU : Protocol Data Unit
PVC : Permanent Virtual Circuit
RAM : Random Access Memory
RARP : Reverse Address Resolution Protocol
RIP : Routing Information Protocol
RIPE NCC : Réseaux IP Européens Network Coordination Centre
RIR : Regional Internet Registry
ROM : Read-Only Memory
RSA : Rivest, Shamir et Adleman
SMTP : Simple Mail Transfert Protocol
SNMP : Simple Network Management Protocol

SQL	: Structured Query Language
SSH	: Secure Shell
SSI	: Sécurité des Systèmes d'Information
SVC	: Switched Virtual Circuit
TACACS	: Terminal Access Controller Access-Control System
TCP	: Transmission Control Protocol
TFTP	: Trivial File Transfert
TTL	: Time To Live
UDP	: User Datagram Protocol
UIT	: Union Internationale des Telécommunications
VLSM	: Variable Lenght Subnet Mask
WAN	: Wide Area Networks
WWW	: World Wide Web

INTRODUCTION GENERALE

Les systèmes d'information font désormais partie intégrante du fonctionnement des administrations publiques, de l'activité des entreprises et du mode de vie des citoyens. Les services qu'ils assurent nous sont tout aussi indispensables que l'approvisionnement en eau ou en électricité.

Le développement global d'Internet a largement changé le monde et associe aux systèmes d'information une dimension incontournable au développement que ce soit économique, sociale et culturelle. En d'autres termes, la sécurité des systèmes d'information (SSI) est un enjeu à l'échelle mondial.

Les États-Unis ont parfaitement saisi, et ce depuis longtemps, tout l'intérêt stratégique et politique de la sécurisation et du contrôle de l'information. Pour cette raison, dans ce mémoire nous référons au politique de SSI relatif au routeur Cisco et au système de détection d'intrusion de l'initiative du DARPA (Defense Advanced Research Projects Agency) qui sont tous les deux des institutions américaines. Le principe de la SSI est de prendre connaissance des communications des autres tout en protégeant sa propre communication, capacité dans laquelle les États-Unis dominent.

Au niveau gouvernemental, le SSI s'agit d'un enjeu de souveraineté nationale. On parle en effet de la responsabilité de garantir la sécurité du système d'information national, la continuité de fonctionnement des institutions publique et des infrastructures vitales pour les activités socio-économiques du pays et la protection des entreprises et des citoyens.

Au niveau des entreprises, ces dernières doivent se protéger de la concurrence et de la malveillance que ce soit venant de l'intérieur ou de l'extérieur. Le système d'information irrigue l'ensemble de leur patrimoine (propriété intellectuelle et savoir-faire) et porte leur stratégie de développement. L'environnement lié aux technologies de l'information et de la communication est la cible de nombreuses menaces. L'ouverture des réseaux et leur complexité croissante associant des acteurs aux multiples profils, ont renforcé la vulnérabilité des systèmes d'information. Accéder à des données sensibles dans le but de les voler, de les modifier ou de nuire au bon fonctionnement des réseaux sont les principales menace.

Quelles formes prennent les attaques? De qui émanent-elles? Quelle est leur finalité?

Les outils nécessaires aux pirates sont aisément accessibles en ligne et il existe un échange constant d'information et de savoir-faire au sein de la communauté des pirates pour rendre ces

attaques de plus en plus efficaces. Les réseaux terroristes utilisent déjà largement Internet.

Certes, jusqu'à présent Madagascar n'a encore connu officiellement de cyber-attaque majeure motivée par des considérations politiques ou terroristes contre des systèmes d'information, mais rien ne permet d'exclure pour autant qu'une telle attaque ne se produise pas dans le futur.

L'exemple le plus spectaculaire porte sur la révélation, en juin 2005, des agissements d'une entreprise israélienne qui « louait un cheval de Troie » à ses clients ; une affaire qui a conduit à l'arrestation de plusieurs dirigeants d'entreprises à travers le monde. En s'adressant à cette société, un client demandait tout simplement à ce que le produit soit installé dans le système d'information de la cible, pour en extraire en toute impunité toutes les informations qu'il désirait. Cette exemple soulève à quelle point le système d'information doit être sécurisé.

Dans cet ouvrage, pour mieux aborder le concept de sécurisation de système d'information, le premier chapitre se focalisera sur les généralités sur les réseaux. Cette partie étalera les standards et protocoles existants, le crypto système et enfin les cyber-attaques. Le deuxième chapitre présente les routeurs Cisco. Le chapitre suivant abordera la sécurisation des routeurs Cisco puis on parlera du système de détection d'intrusion : le réseau Bayésiens et les bases de données du DARPA. Et le chapitre final se concentrera sur l'application IDS qu'on vient de développer sous java.

CHAPITRE 1
LES RESEAUX INFORMATIQUES ET LES ATTAQUES VISANT LES SYSTEMES D'INFORMATIONS

1.1 Introduction

Toute entreprise possède aujourd'hui un ou plusieurs systèmes de télécommunication qui véhiculent les différentes informations nécessaires à sa vie et à son développement. Ces systèmes sont organisés en réseaux, qu'on peut définir comme des ensembles d'équipements et de supports de transmission dont une des fonctions est de permettre le transfert d'informations. Nous sommes entrés dans l'ère de la communication où le volume et la diversité de ces informations se font de plus en plus grands. Aujourd'hui les progrès de l'informatique rendent possible le traitement d'informations de natures différentes sur le même ordinateur : séquences vidéos et sonores, présentation de documents. C'est le domaine du multimédia. Ces données transitent sur une plateforme suivant un modèle standard, le model OSI (International Standards Organization), suivent des protocoles. Nous parlerons de tous cela dans ce chapitre.

1.2 Le modèle OSI

Le modèle de référence de l'ISO fut introduit comme modèle pratique en 1984 par le comité de normalisation de l'ISO.

Modèle OSI

7	Application
6	Présentation
5	Session
4	Transport
3	Réseau
2	Liaison
1	Physique

Figure 1.01 : *Le modèle OSI*

Ce modèle comprend les sept couches suivantes :

- la couche application – activités spécialisées de réseau comme le transfert de fichiers et le courrier électronique ;

- la couche présentation – formatage de données, transcodage de caractère et cryptage ;
- la couche session – établissement de sessions entre un utilisateur et un nœud de réseau tel le login ;
- la couche transport – livraison des données de bout en bout, en mode sécurisé ou non ;
- la couche réseau – routage des PDU (Protocol Data Unit) à travers des réseaux multiples ; gestion de la congestion intermédiaire ;
- la couche liaison – formatage des données en trames et leur transmission sans erreur à travers un réseau physique ;
- la couche physique – transmission d'éléments binaires ou bits (binary digits) sur le support physique de communication.

Aux fonctionnalités présentes dans tout modèle de communication organisée en couches, l'OSI ajoute la méthode d'encapsulation de paquet qui permet de conserver l'intégrité nécessaire des PDU échangées entre entités homologues utilisant les services fournis par les entités des couches inférieures.

La PDU de chaque couche, à l'exception de celle de la couche physique, est composée de deux parties : l'en-tête et les données. L'en-tête contient des informations annexes utilisées uniquement par l'entité ou le module particulier. Les données, quant à elles, sont reçues pour traitement par la couche immédiatement supérieure. Rappelons que le modèle organisé en couches garantit à l'entité destinataire, la réception intacte d'une PDU telle que l'a envoyée l'entité émettrice. L'OSI se conforme à cette règle, faisant en sorte que l'entité émettrice qui a construit la PDU, la passe dans son intégralité (en-tête inclus) sous forme de données, à l'entité immédiatement inférieure. Quand la correspondante de l'entité inférieure à l'autre bout effectue le démultiplexage des données vers l'entité de la couche supérieure, celle-ci reçoit exactement ce qui lui est destiné. Ce procédé est valable pour toutes les couches sauf pour la couche application qui reçoit, en fait, les données finales en provenance du réseau. [1]

1.3 Le modèle Internet

Le modèle OSI, malgré sa définition assez exhaustive de la communication à couches, comporte quelques lacunes. Conçu à l'origine pour servir de cadre opératoire aux protocoles fonctionnant sur des réseaux locaux ou LAN (Local Area Networks), homogènes, il est peu adapté aux réseaux étendus ou WAN (Wide Area Networks). Si une fonction de routage est bel et bien spécifiée au niveau de la couche réseau du modèle OSI, sa description reste sommaire quant au rôle des

routeurs, sachant que ces derniers constituent les nœuds permettant de relier des réseaux mixtes de bout en bout. L'autre modèle le plus utilisé est le modèle Internet, alias TCP/IP (Transmission Control Protocol / Internet Protocol). À la différence du modèle OSI, le modèle Internet fut conçu pour servir de cadre opératoire aux protocoles fonctionnant sur des réseaux hétérogènes LAN et WAN. [1]

Le modèle Internet comprend quatre couches :

- la couche application – assure des activités spécialisées de réseau comme le terminal virtuel, le transfert de fichiers et le courrier électronique ;
- la couche transport – assure la livraison de données de bout en bout, sécurisées ou non ;
- la couche Internet – assure le routage de données à travers des réseaux hétérogènes et un contrôle de flux rudimentaire ;
- la couche d'accès réseau – assure le formatage de données en trames et leur acheminement sans erreur à travers un réseau physique ; c'est là que s'effectue la transmission de bits sur un support physique de communication.

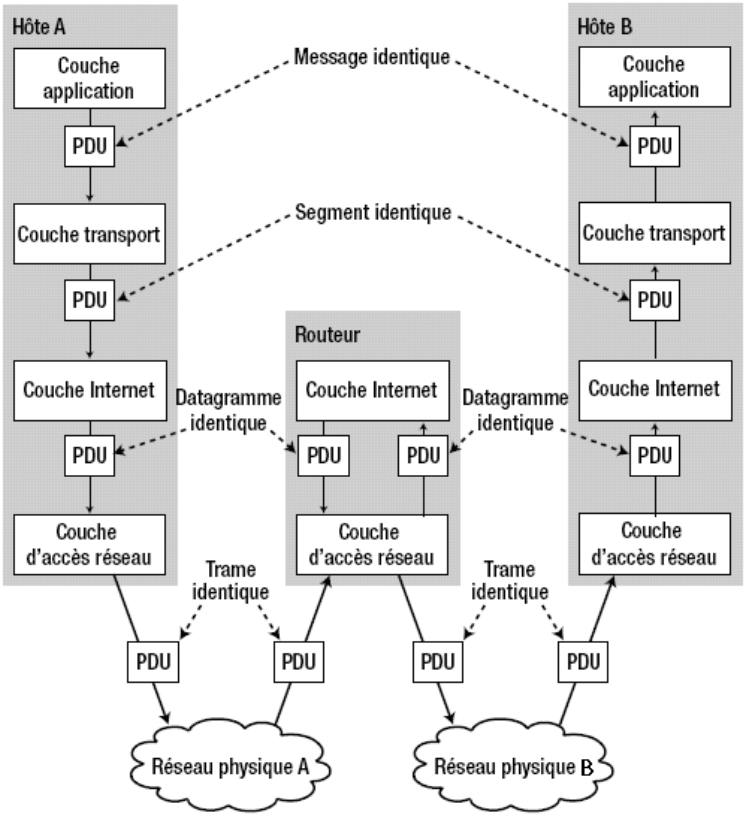


Figure 1.02 : *Deux types de pile dans le modèle Internet*

La fonctionnalité de ces couches est équivalente à celle de leurs homologues dans le modèle.
OSI (cf. Figure 1.03). Il faut cependant remarquer que la couche d'accès réseau regroupe les fonctionnalités de deux couches (liaison et physique). De même la couche application fusionne les couches session et présentation.
Le modèle Internet se différencie encore de celui de l'OSI pour ce qui est de la communication entre réseaux physiques divers, par l'introduction explicite du concept de routeur. Il possède deux types de piles à couches : l'une pour les nœuds terminaux ou « hôtes », dans la terminologie Internet, et l'autre pour les routeurs, autrefois appelés « Gateway ».
Pour tenir compte de ce nouveau schéma d'encapsulation, le modèle Internet distingue deux types de communication selon la couche concernée : le bout en bout ou hôte à hôte (host to host) et le proche en proche ou saut en saut (hop by hop). La communication de bout en bout suppose que les PDU du nœud émetteur sont expédiées vers le nœud récepteur sans se préoccuper du nombre de réseaux physiques intermédiaires à traverser. La communication de proche en proche n'intervient qu'entre deux nœuds situés sur le même réseau physique. Il est clair que la couche transport du modèle Internet assure la communication de bout en bout, tandis que la couche Internet assure la communication de proche en proche.
La comparaison entre le modèle OSI et le modèle Internet est la suivante :

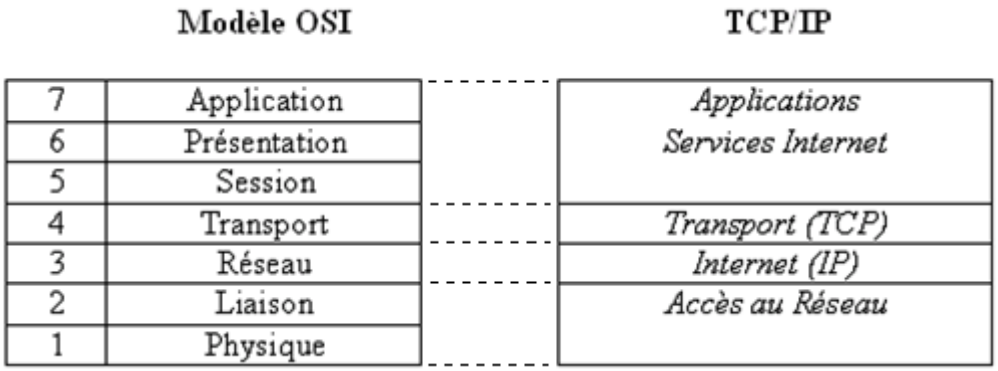


Figure 1.03 : *La correspondance OSI - Internet*

1.4 La suite de protocoles TCP/IP

TCP/IP, devenu standard de fait, est actuellement la famille de protocoles réseaux qui gère le routage la plus répandue sur les systèmes informatiques (Windows, Unix/Linux, Netware...). Plusieurs facteurs ont contribué à sa popularité à savoir sa maturité, son ouverture, son absence de propriétaire, sa richesse (il fournit un vaste ensemble de fonctionnalités), sa compatibilité

(différents systèmes d'exploitation et différentes architectures matérielles), et le développement important d'Internet. [2]

La famille de protocoles conçue autour des deux protocoles principaux, TCP (Transmission Control Protocol) et IP (Internet Protocol), est souvent appelée suite de protocoles TCP/IP. [1]

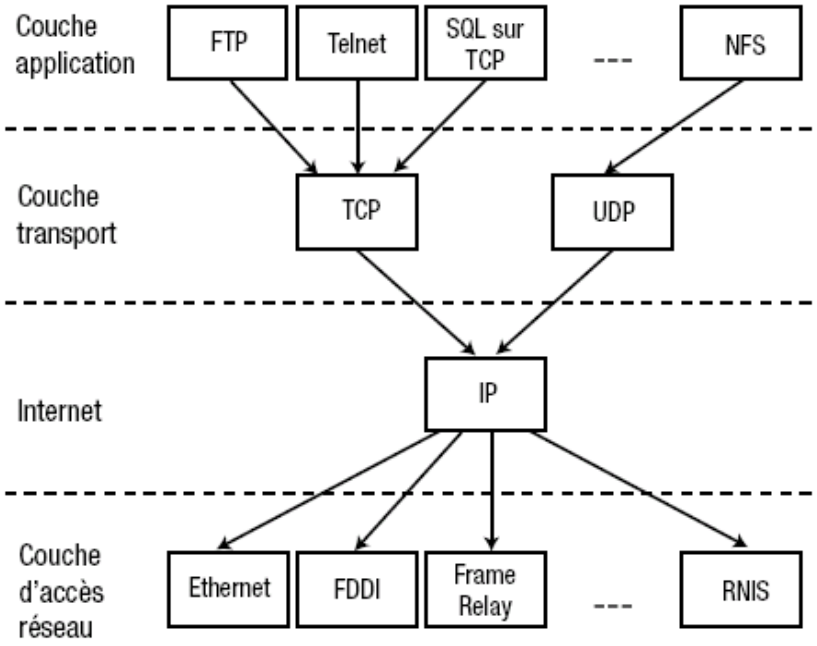


Figure 1.04 : *Série des protocoles TCP/IP*

1.4.1 Le protocole Internet IP

Dans le modèle Internet, le protocole IP exécute deux fonctions : le routage de datagrammes et un contrôle de congestion rudimentaire. La première fonction, le routage, est spécifique à IP. Aucun autre protocole, sur aucune autre couche du modèle Internet, ne peut effectuer le routage de paquets à travers des réseaux intermédiaires hétérogènes.

Contrairement au routage, l'autre fonction de IP qui est le contrôle de congestion est présent quasiment dans toutes les couches du modèle Internet. Celui de la couche Internet est qualifié de « rudimentaire » parce qu'il est très peu évolué, comparé à celui de TCP. La fonction principale d'IP restant tout de même la livraison des datagrammes de l'expéditeur au destinataire, le cas échéant, à travers un grand nombre de réseaux hétérogènes.

Le modèle Internet emploie des noms distincts pour désigner les PDU de la couche Internet et celles de la couche transport : datagramme dans le premier cas et segment dans le second. Le segment est encapsulé dans le datagramme, le pilote de la carte Ethernet les envoie sous forme de

trames sur le réseau physique.

C'est le datagramme qui contient les adresses IP source et destination ainsi que le champ Protocole indiquant quel protocole de couche supérieur recevra les données IP. Ce champ est utilisé pour le multiplexage/démultiplexage des données vers des protocoles de couche supérieure. Par exemple, la valeur du champ est 6 pour TCP, 17 pour UDP (User Datagram Protocol) et 7 pour ICMP (Internet Control Message Protocol). [1]

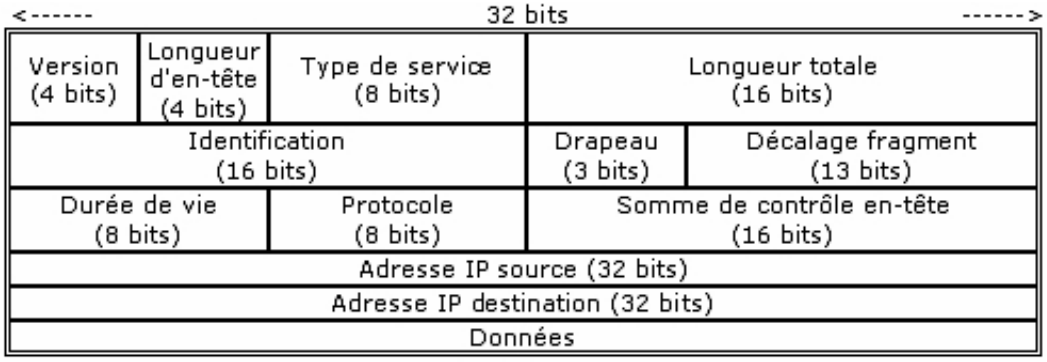


Figure 1.05 : *Datagramme IP*

Voici les protocoles essentiels dans TCP et UDP et leurs numéros de port respectif :

Numéro	Type	Description
7	TCP/UDP	ICMP (Internet Control Message Protocol)
20	TCP	FTP-data - File Transfert Protocol (flux de données)
21	TCP	FTP - File Transfert Protocol (le flux de contrôle pour le transfert de fichiers)
22	TCP	SSH (Secure Shell)
23	TCP	Telnet
25	TCP	SMTP (Simple Mail Transfert Protocol)
53	UDP	DNS (Domain Name Service)
65	TCP	TACACS (Terminal Access Controller Access-Control System) - Database Service
69	UDP	TFTP (Trivial File Transfert)
80	TCP	www-http (World Wide Web - HyperText Transfer Protocol)
110	TCP	Pop3 – Réception de courrier
118	TCP	SQLServ (Structured Query Language Services)
161	UDP	SNMP (Simple Network Management Protocol)
443	TCP	HTTPS (HyperText Transfer Protocol Secure)
546	UDP	DHCP (Dynamic Host Configuration Protocol)
1433	TCP	Microsoft SQL Server
3306	TCP	Mysql Server
8080	TCP	HTTP alternatif (webcache)

Tableau 1.01: *Ports TCP et UDP*

1.4.1.2 Les adresses logiques IPv4

Une adresse IP est un numéro d'identification qui est attribué de façon permanente ou provisoire à chaque appareil connecté à un réseau informatique utilisant l'Internet Protocol.

L'adresse IP est attribuée à chaque interface avec le réseau de tout matériel informatique (routeur, ordinateur, modem ADSL (Asymmetric Digital Subscriber Line), imprimante réseau, etc.) connecté à un réseau informatique utilisant l'Internet Protocol comme protocole de communication entre ses nœuds. Cette adresse est assignée soit individuellement par l'administrateur du réseau local dans le sous réseau correspondant, soit automatiquement via le protocole DHCP. Si l'hôte dispose de plusieurs interfaces, chacune dispose d'une adresse IP spécifique. Chaque paquet transmis par le protocole IP contient l'adresse IP de l'émetteur ainsi que l'adresse IP du destinataire. Les routeurs IP acheminent les paquets vers la destination de proche en proche.

Il existe des adresses IP de version 4 (sur 32 bits, soit 4 octets) et de version 6 (sur 128 bits, soit 16 octets). La version 4 est actuellement la plus utilisée : elle est généralement représentée en notation décimale avec quatre nombres compris entre 0 et 255, séparés par des points.

En pratique, il y a 3 classes d'adresse IP. Le but de la division des adresses IP en trois classes A, B et C est de faciliter la recherche d'un hôte sur le réseau. En effet, avec cette notation, il est possible de rechercher dans un premier temps le réseau que l'on désire atteindre puis de chercher un hôte sur celui-ci. Ainsi, l'attribution des adresses IP se fait selon la taille du réseau. [3]

Classe	Format	Première plage d'octets	Nombre de réseaux possibles	Nombre d'hôtes par réseau	Masque
A	Premier octet / Deuxième octet / Troisième octet / Quatrième octet : Réseau / Hôte / Hôte / Hôte	de 0 à 127	126	16 777 214	/8
B	Réseau / Réseau / Hôte / Hôte	de 128 à 191	16348	65 534	/16
C	Réseau / Réseau / Réseau / Hôte	de 192 à 223	2097152	254	/24

Tableau 1.02: *Les classes d'adresse IPv4 A, B et C*

D’après le RFC 5735, certaines adresses sont réservées à un usage particulier :

Bloc	Usage	Référence
0.0.0.0/8	Adresse réseau par défaut	RFC 1700
10.0.0.0/8	Adresses privées	RFC 1918
100.64.0.0/10	Espace partagé pour Carrier Grade NAT	RFC 6598
127.0.0.0/8	Adresse de bouclage (localhost)	RFC 1122
169.254.0.0/16	Adresses locales autoconfigurées (APIPA)	RFC 3927
172.16.0.0/12	Adresses privées	RFC 1918
192.0.0.0/24	Réservé par l'IETF	RFC 5736
192.0.2.0/24	Réseau de test TEST-NET-1	RFC 5737
192.88.99.0/24	6to4 anycast	RFC 3068
192.168.0.0/16	Adresses privées	RFC 1918
198.18.0.0/15	Tests de performance	RFC 2544
198.51.100.0/24	Réseau de test TEST-NET-2	RFC 5737
203.0.113.0/24	Réseau de test TEST-NET-3	RFC 5737
224.0.0.0/4	Multicast	RFC 5771
240.0.0.0/4	Réservé à un usage ultérieur non précisé	RFC 1112
255.255.255.255/32	Broadcast limité	RFC 919

Tableau 1.03: *Les différents usages des adresses IPv4*

L’IETF (Internet Engineering Task Force) est un groupe international informel, sans statut, sans membre et sans adhésion, ouvert à tout individu qui participe à l'élaboration de standards Internet. L'IETF produit la plupart des nouveaux standards d'Internet.

L’APIPA (Automatic Private Internet Protocol Addressing) ou IPv4LL est un processus qui permet à un système d'exploitation de s'attribuer automatiquement une adresse IP, lorsque le serveur DHCP est hors service ou injoignable.

Un Carrier Grade NAT est un NAT (Network Address Translation) à grande échelle utilisé par un fournisseur d'accès à Internet dans le but de diminuer la quantité d'adresses IPv4 nécessaires aux clients, et ainsi faire face à l'épuisement des adresses IPv4.

6to4 (parfois écrit « 6 to 4 ») est une méthode de transition entre IPv4 et IPv6 qui permet à un réseau IPv6 isolé de communiquer en IPv6 avec un autre réseau IPv6 à travers un réseau IPv4. 6to4 est utile quand deux hôtes souhaitent échanger des informations en IPv6 mais qu'une portion du réseau qui les sépare ne supporte qu’IPv4.

a) Notion d’adresse routable, privée et publique

Une adresse est dite « non routable » lorsqu'elle ne doit être utilisée que sur des réseaux strictement locaux.

Ces adresses vont de 10.0.0.0 à 10.255.255.255, de 172.16.0.0 à 172.31.255.255 et de 192.168.0.0 à 192.168.255.255.

Les adresses IPv4 sont dites publiques si elles sont enregistrées et routables sur Internet, elles sont donc uniques mondialement. À l'inverse, les adresses privées ne sont utilisables que dans un réseau local, et ne doivent être uniques que dans ce réseau.

La traduction d'adresse réseau NAT permet de faire correspondre une seule adresse externe publique visible sur Internet à toutes les adresses d'un réseau privé, et pallie ainsi l'épuisement des adresses IPv4. Les adresses privées ne peuvent pas être routées sur Internet. On dit qu'un routeur fait une traduction d'adresse réseau lorsqu'il fait correspondre les adresses IP internes non-uniques et souvent non routables d'un intranet à un ensemble d'adresses externes uniques et routables.

b) Le RIR (Regional Internet Registry)

L'IANA (Internet Assigned Numbers Authority) est en charge de la gestion de l'espace d'adressage IP, elle a segmentée l'espace d'adressage entre 256 blocs de taille /8 numéroté de 0/8 à 255/8. Chacun de ces blocs représente 16 millions d'adresses (24 bits). Puis l'IANA délègue l'administration de ces segments à 5 RIR (Regional Internet Registry) selon la cartographie suivante : [4]

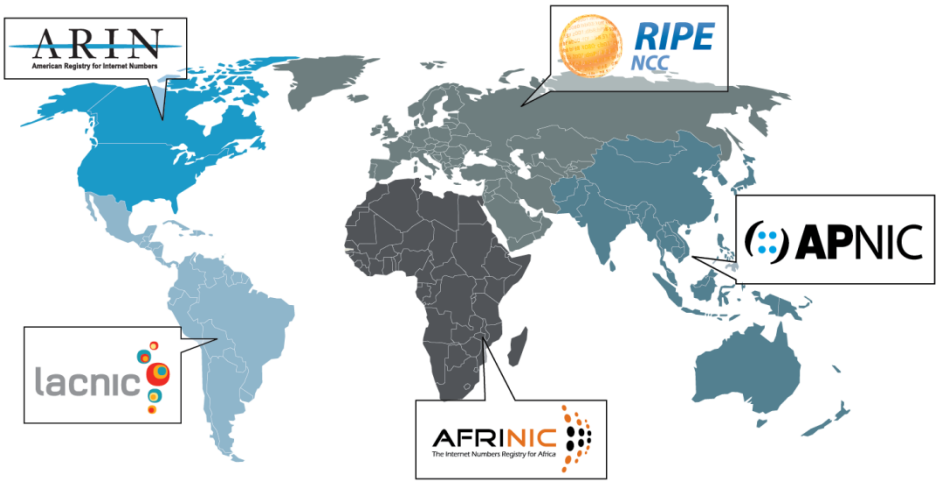


Figure 1.06 : *La carte Regional Internet Registry*

AFRINIC : African Network Information Centre

ARIN : American Registry for Internet Numbers

RIPE NCC : Réseaux IP Européens Network Coordination Centre

APNIC : Asia Pacific Network Information Centre

LACNIC : Latin America and Caribbean Network Information Centre

Madagascar est sous la juridiction de l'AFRINIC, ou plus précisément sous l'autorité du NIC-MG. L'IANA qui a pouvoir sur les adresses IP, les noms de domaines et tous les autres paramètres utilisés dans l'Internet, a délégué à l'ESPA (Ecole Supérieure Polytechnique d'Antananarivo) l'administration des services du NIC à Madagascar le 25 juillet 1995 (référence NIC-950627.146). L'ESPA s'est détachée de l'administration directe des services du NIC-MG pour des raisons d'ordre organisationnel interne à l'école. [5]

1.4.1.3 Les adresse logiques IPv6

La croissance du nombre d'utilisateurs et de serveurs d'Internet s'accompagne d'un épuisement des adresses IPv4, c'est-à-dire de la saturation progressive de la quantité d'adresses IPv4 publiques disponibles. La saturation menace la croissance du réseau internet. En février 2011, la réserve de blocs libres d'adresses publiques IPv4 de l'IANA est arrivée à épuisement. En effet, une adresse IPv4 comporte 32 bits, ce qui permet de créer jusqu'à 4 milliards (4 294 967 296 = 2^{32}) de numéros. [6]

Depuis la fin du XXe siècle, IPv6 est proposé comme solution pour faire face à la pénurie des adresses IPv4. Madagascar se trouve actuellement dans la première phase de transition (dite double pile ou dual stack), les ordinateurs disposent à la fois d'une adresse IPv4 et d'une adresse IPv6 et utilisent l'une ou l'autre adresse en fonction de la destination voulue. Ceci ne contribue pas à la diminution de la demande en adresse IPv4 mais permet aux ordinateurs qui ne disposent que d'une adresse IPv6 de continuer à accéder aux services disponibles sur Internet. Dans cet ouvrage, nous nous baserons seulement sur l'IPv4 car c'est encore l'attribution dominante à Madagascar.

1.4.2 Le protocole TCP (Transmission Control Protocol)

TCP est probablement le protocole IP de niveau supérieur le plus répandu. TCP fournit un service sécurisé de remise des paquets. TCP fournit un protocole fiable, orienté connexion, au-dessus d'IP (ou encapsulé à l'intérieur d'IP). TCP garantit l'ordre et la remise des paquets, il vérifie l'intégrité de l'en-tête des paquets et des données qu'ils contiennent. TCP est responsable de la retransmission des paquets altérés ou perdus par le réseau lors de leur transmission. Cette fiabilité fait de TCP/IP un protocole bien adapté pour la transmission de données basée sur la session, les applications client-serveur et les services critiques tels que le courrier électronique. [7]

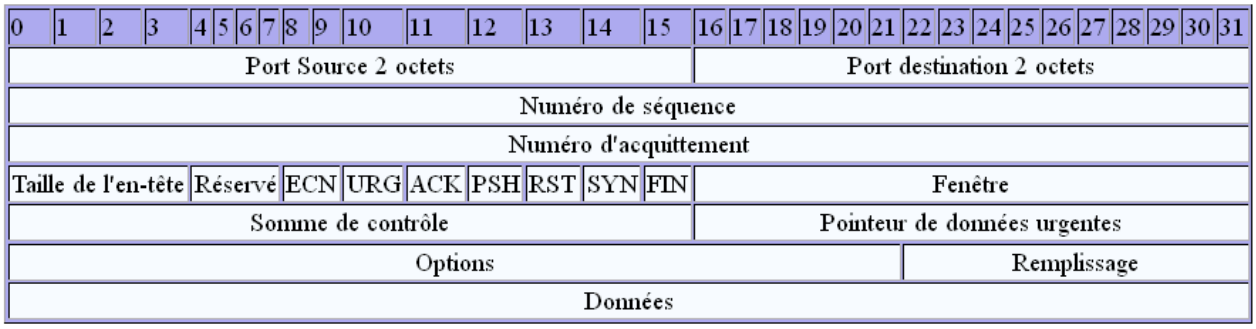


Figure 1.07 : *Format d'un segment TCP*

Port source : Numéro du port source

Port destination : Numéro du port destination

Numéro de séquence : Donne la position du segment dans le flux de données envoyées par l'émetteur

Numéro d'acquittement : Numéro de séquence du prochain octet attendu, c'est-à-dire le numéro de séquence du dernier octet reçu avec succès plus 1

Taille de l'en-tête : Longueur de l'en-tête en mots de 32 bits (les options font partie de l'en-tête)

Drapeaux :

- Réservé : réservé pour un usage futur
- ECN : signale la présence de congestion
- URG : Signale la présence de données urgentes
- ACK : signale que le paquet est un accusé de réception (acknowledgement)
- PSH : données à envoyer tout de suite (push)
- RST : rupture anormale de la connexion (reset)
- SYN : demande de synchronisation (SYN) ou établissement de connexion
- FIN : demande la FIN de la connexion

Fenêtre : Taille de fenêtre demandée, c'est-à-dire le nombre d'octets que le récepteur souhaite recevoir sans accusé de réception

Checksum : Somme de contrôle calculée sur l'ensemble de l'en-tête TCP et des données, mais aussi sur un pseudo en-tête (extrait de l'en-tête IP)

Pointeur de données urgentes : Position relative des dernières données urgentes

Options : facultatives

Remplissage : Zéros ajoutés pour aligner les champs suivants du paquet sur 32 bits, si nécessaire

L'établissement et la terminaison d'une connexion suivent le diagramme d'échange suivant :

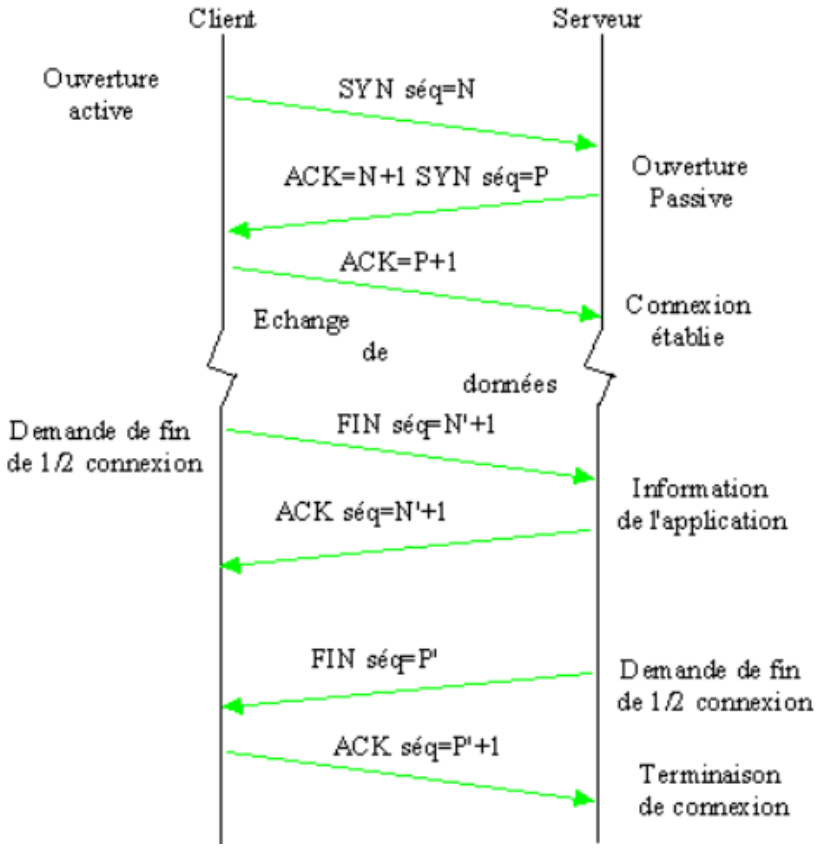


Figure 1.08 : *Les états des drapeaux d'un segment TCP pendant une communication*

Le client demande l'ouverture de la connexion en émettant un segment avec un drapeau "SYN" (bit SYN fixé à 1) spécifiant le numéro du port du serveur avec lequel il souhaite se connecter.
Il expédie aussi un numéro de séquence initial N. Cette phase est appelée ouverture active et consomme un numéro de séquence.
Le serveur répond en envoyant un segment avec les drapeaux "ACK" et "SYN" positionnés à 1 : ainsi, il acquitte le premier segment reçu avec une valeur ACK=N+1 et indique un numéro de séquence initial. Cette phase est appelée ouverture passive. Le client doit également acquitter ce deuxième segment en renvoyant un segment avec ACK=P+1 (pour le cas où 2 demandes de connexion auraient lieu en même temps, chacune dans un sens).

La terminaison d'une connexion peut être demandée par n'importe quelle extrémité et se compose de 2 "demi fermetures" (des flots de données pouvant s'écouler simultanément dans les deux sens). L'extrémité qui demande la fermeture (dans la figure le client) émet un segment où le drapeau FIN est positionné à 1 et où le numéro de séquence vaut N'. Le récepteur du segment l'acquitte en retournant un ACK=N'+1 et informe l'application de la demi fermeture vers l'extrémité l'ayant demandée). Dans l'autre sens, seuls des accusés de réception sont transmis.
Quand l'autre extrémité veut fermer la connexion, elle agit de même ce qui entraîne la terminaison complète de la connexion. [8]

1.4.3 Le protocole UDP (User Datagram Protocol)

UDP est un complément du protocole TCP qui offre un service de datagrammes sans connexion qui ne garantit ni la remise ni l'ordre des paquets délivrés. Les sommes de contrôle des données sont facultatives dans le protocole UDP. Ceci permet d'échanger des données sur des réseaux sans utiliser inutilement des ressources réseau ou du temps de traitement. Les messages (ou paquets UDP) sont transmis de manière autonome (sans garantie de livraison). Le protocole UDP prend également en charge l'envoi de données d'un unique expéditeur vers plusieurs destinataires. En guise d'exemple, TFTP (trivial FTP) s'appuie sur UDP, DHCP également, Windows utilise UDP pour les Broadcast en TCP-IP. [7]

1.4.4 Le protocole ICMP (Internet Control Message Protocol)

C'est un protocole de maintenance utilisé pour les tests et les diagnostics, qui véhiculent des messages de contrôle. Il permet à deux systèmes d'un réseau IP de partager des informations d'état et d'erreur. La commande « ping » utilise les paquets ICMP de demande d'écho et de réponse à un écho afin de déterminer si un système IP donné d'un réseau fonctionne. C'est pourquoi l'utilitaire « ping » est utilisé pour diagnostiquer les défaillances au niveau d'un réseau IP ou des routeurs. [7]

1.4.5 Le protocole RIP (Routing Information Protocol)

RIP est un protocole de routage dynamique qui permet l'échange d'informations de routage sur un inter-réseau. Chaque routeur fonctionnant avec RIP échange les identificateurs des réseaux qu'il peut atteindre, ainsi que la distance qui le sépare de ce réseau (nombre de sauts = nombre de routeurs à traverser). Ainsi chacun dispose de la liste des réseaux et peut proposer le meilleur chemin. [7]

1.4.6 Le protocole ARP (Address Resolution Protocol) et le protocole RARP (Reverse Address Resolution Protocol)

Le protocole ARP a un rôle phare parmi les protocoles de la couche Internet de la suite TCP/IP, car il permet de connaître l'adresse physique d'une carte réseau correspondant à une adresse IP en effectuant une diffusion du type *"qui est X.X.X.X ?"* puis en créant une table de correspondance entre les adresses logiques et les adresses physiques dans une mémoire cache, c'est pour cela qu'il s'appelle Protocole de résolution d'adresse.

Le protocole RARP est beaucoup moins utilisé, il signifie Protocole ARP inversé, il s'agit donc d'une sorte d'annuaire inversé des adresses logiques et physiques. En réalité le protocole RARP est essentiellement utilisé pour les stations de travail n'ayant pas de disque dur et souhaitant connaître leur adresse physique. Le protocole RARP permet à une station de connaître son adresse IP à partir d'une table de correspondance entre adresse MAC (adresse physique) et adresses IP hébergée par une passerelle (Gateway) située sur le même réseau local (LAN). [9]

1.4.7 La notion de sous-réseau

L'adresse du sous réseau est obtenue en appliquant l'opérateur *ET* binaire entre l'adresse IPv4 et le masque de sous réseau. L'adresse de l'hôte à l'intérieur du sous réseau est quant à elle obtenue en appliquant l'opérateur *ET* entre l'adresse IPv4 et le complément à un du masque (en anglais c'est le Wildcard Bits ou Wildcard Mask).

Prenons quelques exemples :

Adresse 192.168.1.2 et masque 255.255.255.0

→ 192.168.1.2 & 255.255.255.0 = 192.168.1.0 (adresse du réseau)

→ 192.168.1.2 & 0.0.0.255 = 0.0.0.2 (adresse de l'hôte)

Adresse 91.198.174.2/19

→ 91.198.174.2 & 255.255.224.0 = 91.198.160.0 (adresse du réseau)

→ 91.198.174.2 & 0.0.31.255 = 0.0.14.2 (adresse de l'hôte)

1.4.7.1 Subnetting ou structuration par sous-réseau

Le système de classe contribuait au gaspillage d'adresse, les besoins des entreprises étant souvent supérieur à ceux pouvant être satisfait par la classe C sans toutefois justifier l'attribution d'adresse de classe B.

L'adressage par classe est devenu tout à fait inadapté mais il fallait pourtant tenter de préserver les attributions déjà réalisées. L'idée consiste à emprunter un nombre de bit à définir dans le numéro d'hôte afin d'en faire une adresse de sous-réseau :

- 2 bits empruntés permettent de définir 4 sous-réseaux
- 3 bits empruntés permettent de définir 8 sous-réseaux et ainsi de suite…

Vu de l'extérieur, l'adresse est toujours valide, un datagramme destiné à l'une des machines de l'un des sous-réseaux est toujours transporté par Internet en mettant à profit l'adresse réseau, il n'y a pas d'impact sur Internet.

A l'intérieur par contre, il devient possible pour l'entreprise de mieux structurer son espace d'adressage et cela peut contribuer à éviter des demandes de bloc d'adressage supplémentaire dont l'objet ne serait pas de pourvoir un besoin d'adresse mais bien de permettre cette structuration.

C'est à l'administrateur qu'il revient de fixer la frontière entre l'adresse sous-réseau et l'adresse hôte selon les besoins de l'entreprise. Chaque bit emprunté supplémentaire multiplie par 2 le nombre de sous-réseau et divise par 2 le nombre potentiel d'hôte à l'intérieur d'un sous-réseau.

Le RFC interdit une adresse de sous-réseau dont tous les bits sont à '0' car il serait impossible de distinguer l'adresse de sous-réseau de celui du réseau globale. De la même façon, le RFC interdit une adresse de sous-réseau dont tous les bits sont à '1'. Cette fois-ci, c'est l'adresse de diffusion du sous-réseau qui chevauche celui du réseau global. [10]

1.4.7.2 L'adressage sans classe

a) Le VLSM (Variable Lenght Subnet Mask)

En adoptant VLSM, la structuration en sous-réseau prend une autre dimension. Il s'agit toujours d'un découpage en sous-réseau mais avec un masque de longueur variable. La longueur de préfix adopté doit se maintenir entre la longueur initiale + 1 et la longueur maximale. Les avantages sont au nombre de deux :

- Il devient possible de créer des sous-réseaux dont le nombre d'adresse hôte colle au plus près du besoin d'adresse du sous-réseau considéré. Le cas le plus évident est celui des liens point à point routeur qui nécessite deux adresses et à qui l'administrateur pourra affecter un sous-réseau avec un préfixe /30.
- En structurant le réseau de façon judicieuse, l'agrégation d'adresse est favorisée avec des avantages déterminant sur le volume de la table de routage, la consommation de ressource machine (temps passé à lire la table de routage, encombrement de la table de routage en

mémoire vive) ainsi que sur la bande passante consommée par le trafic d'acheminement (trafic ne transporte pas de donnée utile utilisateur mais les informations de topologie échangé entre processus de routage). [6]

L'administrateur place dans un tableau tous les sous-réseaux possibles du préfixe de l'adresse voulue (172.16.0.0/24 par exemple) avec les masques possibles allant de /25 à /30. Bien sure, un réseau de grande taille ne pourrait pas être représenté par un seul de ces tableaux. Un peu à la façon des cartes routières qui existent à différents échelles, l'administrateur qui aurait à gérer de grand espace à l'adresse devrait sans doute construire un tableau général puis des tableaux intermédiaires qui détailleraient les parties du préfixe initial.

L'idée est simple, il faut raisonner avec le nombre d'hôte à insérer dans la plage du sous-réseau. Par exemple si on cherche la plage convenable pour insérer 56 hôtes, 6 bits suffit pour les adresse hôtes. $2^6 = 64$, il y a une adresse pour déterminer le sous-réseau et un autre pour le broadcast de ce sous-réseau, ce qui laisse 62 adresses valides. C'est-à-dire que /26 est convenable pour insérer les 56 hôtes. Il faut donc enlever cette plage /26 du tableau pour qu'on ne puisse plus l'utiliser pour les autres besoins du même cahier de charge.

On peut donc raisonner mentalement que si on a besoin d'une liaison point à point entre deux routeurs adjacentes, le choix du masque est /30. Cela donnera un espace de découpage de 4 adresses dont 2 adresses valides pour attribuer aux deux routeurs.

b) Le CIDR (Classless Inter-Domain Routing)

Au début des années 90, suite à l'afflux des nouveaux utilisateurs d'Internet, surtout des entreprises, le système d'attribution des réseaux IP basé sur le système des classes commença à montrer ses limites car la taille des tables de routage se mit à gonfler exponentiellement. Un nouveau système de répartition des adresses en dehors des classes fut mis en place : le CIDR.

L'adressage CIDR est le système de gestion et d'allocation d'adresses IP le plus utilisé aujourd'hui. C'est la réponse à la question comment un système d'adressage par classes aurait-il pu supporter plus de 2 milliards d'internautes ? Sinon depuis les années quatre-vingt-dix, nous n'aurions plus d'adresses IP disponibles. [11]

CIDR	Bits disponibles	Masque de sous réseau	Nombre d'hôte par sous réseau
/1	31	128.0.0.0	2^{31}-2 = 2147483646
/2	30	192.0.0.0	2^{30}-2 = 1073741822
/3	29	224.0.0.0	2^{29}-2 = 536870910
/4	28	240.0.0.0	2^{28}-2 = 268435454
/5	27	248.0.0.0	2^{27}-2 = 134217726
/6	26	252.0.0.0	2^{26}-2 = 67108862
/7	25	254.0.0.0	2^{25}-2 = 33554430
/8	24	255.0.0.0	2^{24}-2 = 16777214
/9	23	255.128.0.0	2^{23}-2 = 8388606
/10	22	255.192.0.0	2^{22}-2 = 4194302
/11	21	255.224.0.0	2^{21}-2 = 2097150
/12	20	255.240.0.0	2^{20}-2 = 1048574
/13	19	255.248.0.0	2^{19}-2 = 524286
/14	18	255.252.0.0	2^{18}-2 = 262142
/15	17	255.254.0.0	2^{17}-2 = 131070
/16	16	255.255.0.0	2^{16}-2 = 65534
/17	15	255.255.128.0	2^{15}-2 = 32766
/18	14	255.255.192.0	2^{14}-2 = 16382
/19	13	255.255.224.0	2^{13}-2 = 8190
/20	12	255.255.240.0	2^{12}-2 = 4094
/21	11	255.255.248.0	2^{11}-2 = 2046
/22	10	255.255.252.0	2^{10}-2 = 1022
/23	9	255.255.254.0	2^{9}-2 = 510
/24	8	255.255.255.0	2^{8}-2 = 254
/25	7	255.255.255.128	2^{7}-2 = 126
/26	6	255.255.255.192	2^{6}-2 = 62
/27	5	255.255.255.224	2^{5}-2 = 30
/28	4	255.255.255.240	2^{4}-2 = 14
/29	3	255.255.255.248	2^{3}-2 = 6
/30	2	255.255.255.252	2^{2}-2 = 2
/31	1	255.255.255.254	
/32	0	255.255.255.255	

Tableau 1.04: *Le principe du CIDR*

Ce système, qui est régi par les RFC 1518 et 1519, a été conçu pour remplacer l'adressage par classes pour ces raisons évoquées précédemment. Le but de ce nouveau système s'articule autour de deux points :

- économiser les adresses IP.
- faciliter le routage.

Le CIDR donne la possibilité d'utiliser un seul réseau qui fusionne plusieurs sous-réseaux. Cette fusion de sous-réseaux, dite aussi *supernetting*, est l'essence même de CIDR. Cette technique est également appelée résumé de routes.

Pour implémenter un réseau fondé sur l'adressage CIDR, il faut utiliser un protocole qui puisse le supporter. Il en existe plusieurs, tels que BGP (Border Gateway Protocol) et OSPF (Open Shortest Path First). Si le protocole ne supporte pas ce type d'adressage, le routage échouera dans ce réseau. En général, les petits LAN et les réseaux domestiques n'implémentent pas l'adressage CIDR.

1.4.7.3 Masque Inverse ou complément à un du masque (Wildcard Mask ou Wildcard Bits)

Pour utiliser certains protocoles ou fonctionnalités des routeurs, on fait parfois appel aux « wildcard masks », afin d'identifier un sous réseau ou une plage d'adresses IP. C'est le cas pour l'utilisation des protocoles OSPF, EIGRP (Enhanced Interior Gateway Routing Protocol) ou encore pour les ACLs (Access Control List). Lors de l'application d'un « wildcard mask **»** il faut savoir que:

- Un bit avec une valeur de 0 vérifie la correspondance de l'adresse.
- Un bit avec une valeur de 1 ignore la valeur correspondante de l'adresse.

Donc, 0.0.0.255 correspond à un masque normal en /24 ou 255.255.255.0

0.0.255.255 correspond à un masque normal en /16 ou 255.255.0.0 [12]

Pour calculer rapidement le « wildcard mask » d'un sous réseau le plus simple est de faire une simple soustraction comme suit:

Exemple avec un masque /26 :

Code :

```
  255.255.255.255
- 255.255.255.192
------------------
  0 . 0 . 0 . 63
```

Autre exemple pour un masque en /19

Code :

```
  255.255.255.255
- 255.255.224. 0
----------------
  0 . 0 . 31 . 255
```

1.5 Le protocole X25 et le protocole Frame Relay

1.5.1 Les raisons de la migration vers le Frame Relay

Bien que considère actuellement comme un protocole d'ancienne génération, le X.25 a été une technologie de commutation de paquets très répandue car elle permettait d'obtenir une connexion très fiable sur des infrastructures câblées non fiables. Ce résultat était obtenu grâce à des contrôles de flux et d'erreur supplémentaires. Ces contrôles alourdissaient cependant le protocole.

Le protocole Frame Relay demande moins de temps de traitement que le X.25, du fait qu'il comporte moins de fonctionnalités. Par exemple, il ne fournit pas de correction d'erreur, car les réseaux étendus actuels permettent d'obtenir des connexions plus fiables que les anciens. Lorsqu'il détecte des erreurs, le nœud Frame Relay abandonne tout simplement les paquets sans notification. Toute correction d'erreur, telle que la retransmission des données, est à la charge des composants d'extrémité. La propagation des données d'une extrémité client à une autre est donc très rapide sur le réseau. [13]

Bien que les deux Frame Relay et X.25 utilisent le même protocole HDLC (High-level Data Link Control) de base, il existe plusieurs différences entre les deux. Certaines des différences importantes entre un réseau Frame Relay et X.25 et du réseau sont donnés dans le tableau 1.05

Caractéristique	**X.25**	**Frame Relay**
Protocole de trame de base utilisé	HDLC	HDLC
Vitesse typique (bande passante)	Faible	Haut
Des sessions interactives	Peine appropriée	Approprié
Connectivité LAN pour des transferts de fichiers rapides	Ne convient pas	Approprié
Complexité de protocole	Haut	Faible
Support de la voix	Pauvres	Bon
Correction d'erreur	Très bon	Non pris en charge
Commentaires	Une X.25 est un protocole très vieux, et largement mis en œuvre. X.25 met en œuvre une correction d'erreur de nœud à nœud, et très approprié pour circuits bruyants.	Frame Relay est largement mise en œuvre de nos jours. Frame Relay ne supporte aucune correction d'erreur de nœud à nœud (adapté avec l'évolution des canaux physiques actuels donc rend obsolète X.25)

Tableau 1.05: *Différence entre X.25 et Frame Relay*

1.5.2 La description du réseau Frame Relay

Lorsqu'on construit un réseau étendu, quel que soit le mode de transport choisi, deux sites sont toujours relies par un minimum de trois composants ou groupes de composants de base. Chaque site doit avoir son propre équipement (DTE) pour accéder au commutateur (DCE). Le troisième composant se trouve entre les deux, reliant les deux points d'accès.

La figure 1.09 représente la partie fournie par le réseau fédérateur Frame Relay.

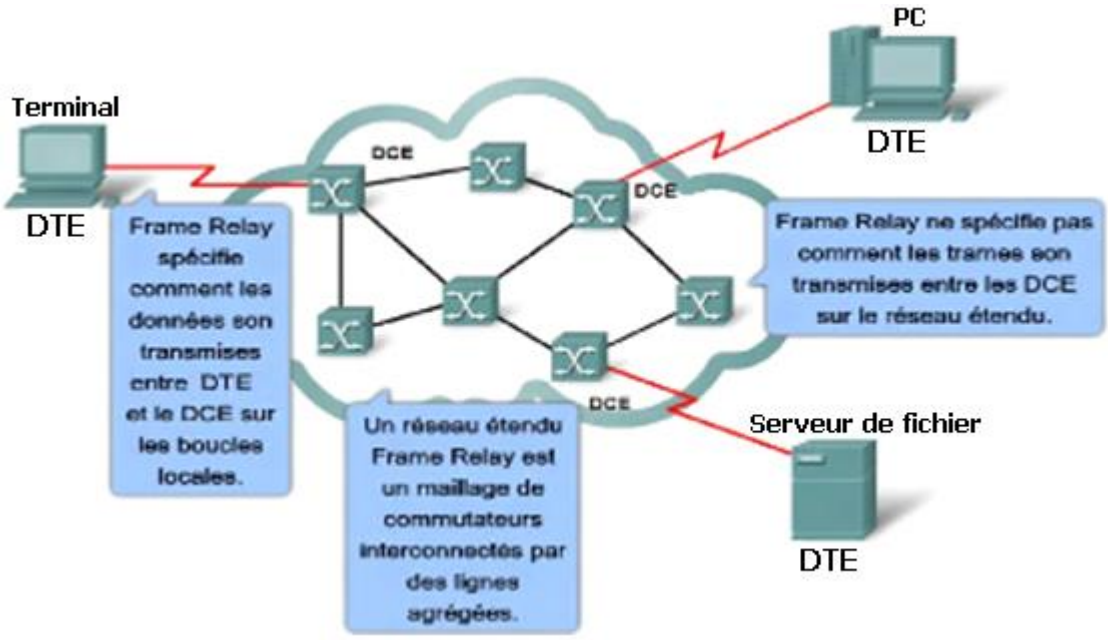


Figure 1.09 : ***Réseau étendu frame Relay***

- le DTE : (Data Terminal Equipment), c'est un équipement (généralement un routeur) de terminaison de réseau placé chez le client du fournisseur Frame Relay.
- le DCE : (Data Circuit terminating Equipment), c'est un équipement fournissant des services d'horloge et de commutation placé chez le fournisseur d'accès.

Le protocole Frame Relay intervient entre un périphérique d'utilisateur final, tel qu'un pont ou un routeur de réseau local, et un réseau. Le réseau proprement dit peut utiliser n'importe quelle méthode de transmission compatible avec la vitesse et l'efficacité requises par les applications Frame Relay. [13]

Le Frame Relay est un protocole à commutation de paquets situé au niveau de la couche de liaison (niveau 2) du modèle OSI, utilisé pour les échanges intersites (WAN). Il peut être vu :

- comme un successeur de X.25 : il a en effet remplacé ce protocole pour le raccordement des sites des entreprises aux infrastructures des opérateurs qui offrent des services réseau privé virtuel.
- comme une étape vers l'ATM (Asynchronous Transfer Mode) : il a souvent été présenté ainsi par les opérateurs ayant « voulu » la combinaison X.25 et l'ATM. Le Frame Relay est en effet issu d'une volonté américaine, de l'ANSI (American National Standards Institute) en particulier, X.25 n'ayant jamais été très populaire.
- comme faisant partie du ISDN (Integrated Services Digital Network) : c'est ainsi que l'UIT (Union Internationale des Telécommunications) l'a considéré et a défini des normes qui n'ont jamais été implémentées.

1.5.3 Le HDLC (High-level Data Link Control)

Le HDLC est un protocole de niveau 2 (couche de liaison) du Modèle OSI. Son but est de définir un mécanisme pour délimiter des trames de différents types, en ajoutant un contrôle d'erreur. Il est défini par l'Organisation internationale de normalisation sous la spécification ISO 3309 (cette norme a été révisée par: ISO/IEC 13239:2002). Les interfaces série des routeurs Cisco utilisent une version propriétaire de HDLC par défaut. [14]

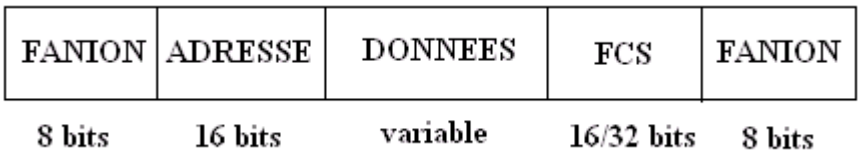


Figure 1.10 : *Trame HDLC*

Le format de trame HDLC contient les champs suivants :

- Le flag ou fanion (8 bits) est un délimiteur de trame pour la synchronisation. Pour HDLC, sa valeur est 011111102 ou 7Eh. Le fanion se trouve au début et à la fin du trame.
- L'adresse (8 bits) est celle du destinataire à qui est envoyée la trame.
- La commande ou contrôle (8 bits) permet de distinguer le type du trame : trame d'information (données), trame de supervision et trame non numérotée. Ce champ est souvent considéré comme appartenant au champ adresse. Par conséquent ce dernier sera compté 16 bits.

- Le champ donné de longueur variable contient les données à envoyer. Le nombre de bits à expédier n'a pas à être un multiple de 8 : comme ce champ n'a pas besoin d'être aligné du point de vue octet, il n'est pas nécessaire d'ajouter de bits de bourrage à la fin.
- Le Frame Check Sequence (FCS) est un code ajouté après les données pour détecter d'éventuelles erreurs de transmission. Il est codé habituellement sur 16 bits, mais après négociation entre les deux interlocuteurs, il peut être sur 32 bits. Cette séquence correspond au CRC calculé sur les champs adresse + commande + données.

Bien que Frame Relay utilise le protocole HDLC de base, examinons de plus près le champ Adresse car la configuration de ce champ est requise dans le paramétrage des routeurs Cisco.

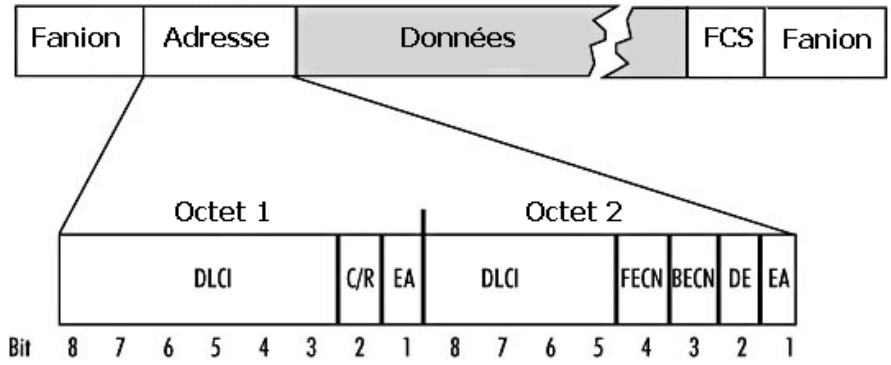


Figure 1.11 : *Le champ adresse de la trame Frame Relay*

Il s'agit d'un champ de 16 bits reparti comme suit :

- Le DLCI (Data Link Connection Identifier) est de 10 bits de large. DLCI identifie la connexion virtuelle entre le nœud d'extrémité (un appareil DTE) et le commutateur (un équipement DCE).
- Le C/R (Command/Response) indique si la trame est une commande ou une réponse. Ce bit ne peut être modifié par le réseau.
- Le FECN (Forward Explicit Congestion Notification) notifie qu'il y a une congestion du réseau FR. Il s'agit d'un domaine à un seul bit qui peut être réglé à 0 ou 1 par un interrupteur. Normalement, FECN est égal à zéro. Une valeur de 1 indique la congestion du réseau dans la direction de la source à la destination.
- Le BECN (Backward Explicit Congestion notification) est donc une notification mentionnant qu'il y avait une congestion. Il s'agit d'un domaine à un seul bit qui peut être réglé à 0 ou 1 par un interrupteur dans le réseau FR. Normalement, BECN est égal à

zéro. La valeur 1 indique que le réseau FR a connu congestion dans la direction de la source à la destination. En utilisant FECN et BECN, protocoles de couche supérieure peuvent contrôler la communication pour une utilisation efficace de réseau FR.

- Le DE (Discard Eligibility) signifie supprimer admissibilité. Il est défini par le DTE pour indiquer que la trame marquée peut être écartée en cas de congestion du réseau.
- L'EA (Expanded Adress) ou adresse étendu (EA) est utilisé pour indiquer que, s'il est mis à 1, l'octet actuel est déterminé à être le dernier octet de l'DLCI.

DLCI est un identifiant à valeur locale (à une interface) permettant d'acheminer un paquet jusqu'à sa destination sur le réseau de communication. Ce qui peut être fait par RARP. Par exemple, le même numéro peut être utilisé par plusieurs routeurs sans poser de problème de connectivité.
Le champ DLCI est localisé dans l'en-tête du relais de trames qui est l'adresse de destination de la trame correspondant à un circuit virtuel permanent (PVC ou Permanent Virtual Circuit). Le standard a été développé conjointement par l'ANSI et le CCITT (Comite Consultatif International Telegraphique et Telephonique) pour permettre l'existence de 1024 DLCI. Mais seulement les nombres de 16 à 1007 sont disponibles pour chaque utilisateur.
Le tableau 1.06 montre les valeurs que peut prendre le DLCI.

0	Etablissement de la liaison virtuelle (contrôle)
1 - 15	Réservés
16 - 1007	DLCI liaison virtuelle utilisateur (commutée ou permanente)
1008-1018	Réservés
1019-1022	Multicast
1023	Signalisation de congestion

Tableau 1.06: *Les valeurs DLCI*

1.5.4 Le SVC (Switched Virtual Circuit) et le PVC (Permanent Virtual Circuit)

PVC est une connexion permanente entre les nœuds terminaux (DTE) à l'intérieur d'un réseau Frame Relay. Le circuit virtuel est toujours disponible même si les données ne sont pas transmises. Ce type de connexion (PVC) est utilisé lorsqu'il est nécessaire de transférer des données cohérentes entre les nœuds d'extrémité.

A des circuits virtuels commutés SVC fournissent une connexion temporaire entre les nœuds d'extrémité (DTE) à travers un réseau de relais de trame.

1.6 Cryptosystème

1.6.1 Le MD5 (Message Digest 5)

1.6.1.1 Généralités sur le MD5

MD5 (Message Digest 5) est une fonction de hachage cryptographique qui calcule, à partir d'un fichier numérique, son empreinte numérique (en l'occurrence une séquence de 128 bits ou 32 caractères en notation hexadécimale) avec une probabilité très forte que deux fichiers différents donnent deux empreintes différentes.

Comme toute fonction de hachage cryptographique, MD5 peut aussi être utilisé pour calculer l'empreinte d'un mot de passe avec la présence d'un sel permettant de ralentir une attaque par force brute. Le salage (salting en anglais) consiste à ajouter une chaîne de caractères à l'information avant le hachage. Par exemple, dans un cadre cryptographique, au lieu de pratiquer le hachage sur le mot de passe seul, on peut le faire sur le résultat de la concaténation du mot de passe avec une autre chaîne de caractères pseudo-aléatoire, obtenue par un hachage de l'identifiant (login) concaténé avec le mot de passe.

Ainsi, plutôt que de stocker les mots de passe dans un fichier, ce sont leurs empreintes MD5 qui sont enregistrées, de sorte que quelqu'un qui lirait ce fichier ne pourrait pas découvrir les mots de passe. La commande *« enable secret »* des commutateurs et routeurs Cisco utilise le hachage MD5 pour stocker le mot de passe du mode privilégié dans le fichier de configuration de l'équipement.

La figure 1.12 représente le principe du chiffrement MD5.

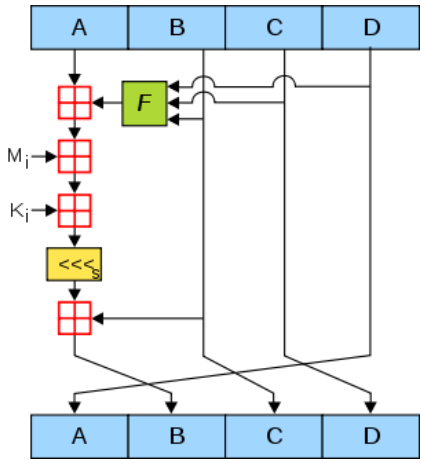


Figure 1.12 : *L'algorithme MD5*

[<<<s] est une rotation de s bits vers la gauche, *s* varie pour chaque opération.
[+] symbolise l'addition modulo 2^{32}.
La fonction F est égale à *[(B ET C) ou (non B ET D)]*

Voici l'empreinte (appelée abusivement signature) obtenue sur une phrase :
MD5("bonjour à tous") = d6aa97d33d459ea3670056e737c99a3d
En modifiant un caractère, cette empreinte change radicalement :
MD5("Bonjour à tous") = 5da8aa7126701c9840f99f8e9fa54976

Très concrètement, la vérification de l'empreinte ou somme de contrôle MD5 peut être réalisée de la façon suivante : lors du téléchargement d'un programme, il y a une série de caractères nommée "Signature MD5". Quand ce téléchargement est terminé, un utilitaire de calcul MD5 est lancé et le résultat est comparé avec la signature. Si les deux valeurs correspondent, on peut alors raisonnablement considérer que le fichier n'a pas été corrompu (volontairement ou non).

1.6.1.2 Mots de passe CISCO de Type 5 et de type 7

Les mots de passe Cisco de Type 5 sont hachés en utilisant l'algorithme MD5. Les mots de passes de niveaux 5 ne sont pas faciles à déchiffrer.
On nomme fonction de hachage (hash) une fonction particulière qui, à partir d'une donnée fournie en entrée, calcule une empreinte servant à identifier rapidement, bien qu'incomplètement, la donnée initiale.
Les mots de passes CISCO de type 7 sont les mots de passe qu'on peut trouver dans le fichier de configuration après avoir validé la commande "show running-config". On les trouve sous la forme de "enable password 7 062TyB56″. 062TyB56 est en fait un « Hash » du mot de passe que vous avez entré dans votre commande Cisco "enable password".
Le Hash de niveau 7 des équipements Cisco est basé sur l'algorithme propriétaire de Cisco qui est réversible. Il est très simple de retrouver le mot de passe d'origine. Du fait, il n'est pas conseillé d'utiliser le type 7. [15]

1.6.2 Algorithme RSA (Rivest, Shamir et Adleman)

Le chiffrement RSA (nommé par les initiales de ses trois inventeurs Ron Rivest, Adi Shamir et Leonard Adleman) est asymétrique : il utilise une paire de clés (des nombres entiers) composé d'une clé publique pour chiffrer et d'une clé privée pour déchiffrer des données confidentielles.

Les deux clés sont créées par une entité, souvent nommée par convention Alice, qui souhaite que lui soient envoyées des données confidentielles. Alice rend la clé publique accessible. Cette clé est utilisée par son (ses) correspondant(s), Bob, pour chiffrer les données qui lui sont envoyées. La clé privée est quant à elle réservée à Alice, et lui permet de déchiffrer ces données. La clé privée peut aussi être utilisée par Alice pour signer une donnée qu'elle envoie, la clé publique permettant à n'importe lequel de ses correspondants de vérifier la signature. [4]

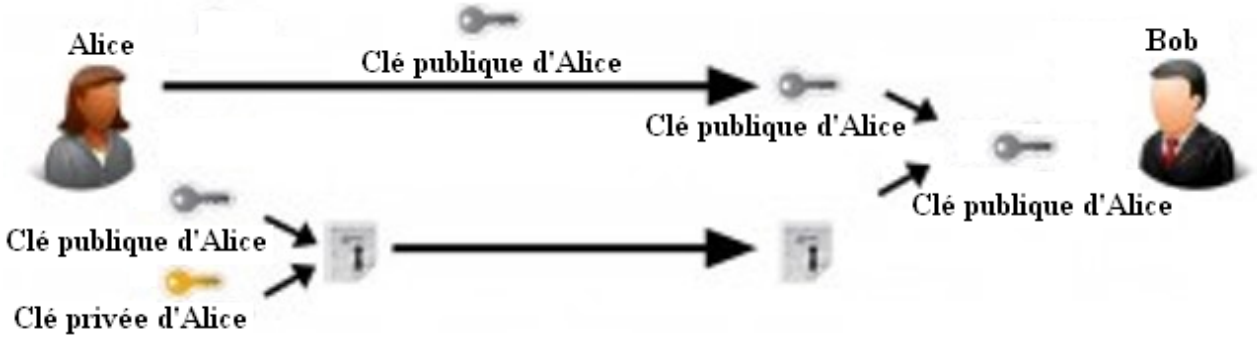


Figure 1.13 : *L'attribution des clés*

Une condition indispensable est qu'il soit « calculatoirement impossible » de déchiffrer à l'aide de la seule clé publique, en particulier de reconstituer la clé privée à partir de la clé publique, c'est-à-dire que les moyens de calcul disponibles et les méthodes connues au moment de l'échange (et le temps que le secret doit être conservé) ne le permettent pas.

Supposons qu'Alice veuille permettre à quiconque le désire de lui envoyer un message confidentiel qu'elle sera la seule à pouvoir lire. Pour cela on peut imaginer un annuaire dans lequel Alice écrira, avec son nom, deux nombres entiers que nous allons appeler n et e (il faut bien sûr que ces deux nombres aient des propriétés particulières, en particulier que n soit « grand »). Un message à destination d'Alice sera supposé être un nombre entier (d'une partie) de l'ensemble :

Figure 1.14 : *Scénario de fonctionnement du RSA*

$$M = \{0, 1, \dots , \quad n - 1\}$$

Pour faire parvenir le message M, il suffira à tout correspondant d'Alice de lui envoyer le cryptogramme :

$$C = M^e \text{modulo } n \qquad (1.01)$$

Si les nombres n et e ont été correctement calculés par Alice (et il faudrait ajouter si l'on prend soin d'éviter les M égaux à des nombres tels que 0, 1, n – 1 …), elle sera la seule à pouvoir déduire M de C.

Voyons maintenant comment Alice doit procéder. Tout d'abord, il lui faut choisir « au hasard » deux nombres premiers p et q grands (par exemple de 100 chiffres décimaux environ). Elle appelle n le produit pq. Ensuite, elle prend un nombre e strictement compris entre 1 et n–1 qui soit premier à p –1 et q–1. Le nombre e est appelé la clé publique d'Alice; le plus souvent, il n'y a pas de raison de le rendre aléatoire.

Du nombre e, Alice peut déduire un autre entier d, sa clé secrète tel que :

$$ed \equiv \text{modulo } [p - 1, q - 1] \qquad (1.02)$$

Ce qui signifie que

$$ed = 1 + t\,[p - 1, q - 1] \quad \text{pour un entier t} \qquad (1.03)$$

En effet, e est inversible modulo le plus petit commun multiple [p – 1, q – 1] de
p – 1 et q – 1 puisque premier à (p – 1) (q – 1).
Les nombres n et e sont rendus publics, le nombre d est gardé secret par Alice ainsi que la factorisation de n.
Supposons que Bob veuille envoyer un message M (que l'on peut supposer être un nombre tel que 1 < M < n – 1 qui sera identité à un élément de à Alice. Il lui adresse l'élément

$$C = M^e \text{ de } \mathbb{Z}/n\mathbb{Z} \tag{1.04}$$

Alice peut alors faire le calcul suivant modulo n :

$$C^d = M^{ed} = M^{1+t[p-1,q-1]} = M(M^{[p-1,q-1]})^t \equiv M \tag{1.05}$$

La dernière égalité se justifie bien en passant au produit d'anneau $\mathbb{Z}/p\mathbb{N} \times \mathbb{Z}/q\mathbb{N}$. En effet

$$M(M^{[p-1,q-1]})^t \equiv M \bmod p \tag{1.06}$$

Et aussi

$$M(M^{[p-1,q-1]})^t \equiv M \bmod q \tag{1.07}$$

La sécurité de ce procédé résulte du fait que l'on ne sait pas calculer d à partir de e et n sans connaître p et q, et que l'obtention de ces derniers (c'est-à-dire la factorisation de n) est difficile si n est « grand ». [4]

1.7 Les attaques visant les systèmes d'informations

De nombreux types d'attaques du réseau ont été identifiés. Ces attaques sont généralement classées en trois principales catégories : attaques dans le but de découvrir des informations, attaques par intrusion et attaques d'interruption de service. [16]

- La première catégorie d'attaque consiste à récolter des informations que les pirates utiliseront par la suite pour détruire les réseaux. Généralement, des outils logiciels tels que les "renifleurs de paquets" (sniffers) ou les scanneurs (scanners) sont utilisés pour analyser les ressources d'un réseau cible, d'un hôte ou d'une application et en exploiter les éventuelles faiblesses. Par exemple, il existe des logiciels spécialement conçus pour découvrir les mots de passe. Ces logiciels ont été créés à l'origine à l'intention des

administrateurs système afin de leur permettre de retrouver les mots de passe oubliés des employés ou de déterminer les mots de passe des employés ayant quitté la société sans communiquer cette information. Aux mains de pirates, ces logiciels peuvent se transformer en une arme redoutable.

- Les attaques par intrusion sont entreprises afin d'exploiter les faiblesses de certaines zones du réseau telles que les services d'authentification afin d'obtenir un accès aux comptes de messagerie électronique, aux bases de données et à d'autres informations confidentielles.
- Les attaques d'interruption de service saturent l'accès à une partie ou à l'intégralité d'un système. Elles s'exécutent généralement par l'envoi massif de données brouillées ou inexploitables à une machine connectée à un réseau d'entreprise ou à Internet, bloquant ainsi le trafic normal des données. Les attaques d'interruption de service distribué (DDoS, Distributed Denial of Service) qui consistent à saturer ainsi plusieurs machines ou hôtes sont encore plus nuisibles. Il existe encore des attaques par saturation appelées attaques en «buffer overflow» qui saturent la mémoire cache des CPU (Central Processing Unit) de n'importe quel élément du réseau. Ces attaques sont particulièrement dévastatrices car elles peuvent rendre un nœud de réseau totalement indisponible.

Les motivations des attaques peuvent être de différentes sortes :

- obtenir un accès au système ;
- voler des informations, tels que des secrets industriels ou des propriétés intellectuelles ;
- ramasser des informations personnelles sur un utilisateur ;
- récupérer des données bancaires ;
- s'informer sur l'organisation (entreprise de l'utilisateur, etc.) ;
- troubler le bon fonctionnement d'un service ;
- utiliser le système de l'utilisateur comme « rebond » pour une attaque ;
- utiliser les ressources du système de l'utilisateur, notamment lorsque le réseau sur lequel il est situé possède une bande passante élevée

Afin de contrer ces attaques, il est indispensable de connaître les principaux types d'attaques afin de mettre en œuvre des dispositions préventives.

1.7.1 Les attaques basées sur le protocole ARP

Il s'agit de malveillance qui a pour but d'attaquer tout réseau local utilisant le protocole de résolution d'adresse ARP, les cas les plus répandus étant les réseaux Ethernet et Wi-Fi.

On peut citer :

- L'ARP spoofing ou ARP poisoning : l'attaquant souhaite détourner des flux de communications transitant entre une machine cible et une passerelle. L'attaquant peut ensuite écouter, modifier ou encore bloquer les paquets réseaux. Il existe en fait deux types de paquet ARP : *is-at* et *who-has*. Dans l'ARP spoofing, l'attaquant va envoyer un paquet ARP qu'il aura lui même forgé :

  <adresse IP hôte ou passerelle cible> is-at <adresse_mac attaquant>

- Gratuitous ARP : l'attaquant émet une trame ARP en broadcast (à tout le réseau) dans laquelle il fait correspondre son adresse MAC (Media Access Control) à l'adresse IP de la passerelle. Le « gratuitous ARP » est initialement prévu pour que les équipements venant d'arriver sur le réseau s'annoncent (ce qui permet par exemple de détecter les IP dupliquées). Ce type de requête très utilisé par des équipements de réseau n'est pas mauvais en soi mais pourrait être détourné si les destinataires sont très mal protégés. Pour illustrer ce type d'attaque avec l'exemple ci-dessus, il faudrait remplacer l'argument « ip_destination » par l'adresse IP de broadcast (ex. : 192.168.1.255) ;
- Emission d'une requête ARP forgée : l'attaquant émet une requête en unicast vers la victime en spécifiant comme adresse IP émettrice, l'adresse IP qu'il veut usurper et en indiquant sa propre adresse MAC comme l'adresse MAC de l'émetteur. Ainsi, lorsque la victime reçoit la requête, elle enregistre la correspondance IP/MAC dans sa table ARP alors que celle-ci est erronée.

L'implémentation des pare-feux CBAC (Context Based Access Control) et/ou *« TCP established ACL »* au niveau du routeur Cisco permettent de contourner ces attaques basées sur le protocole ARP. [9]

1.7.2 L'attaque de l'homme du milieu

L'attaque de l'homme du milieu ou « man in the middle attack » est une attaque qui a pour but d'intercepter les communications entre deux parties, sans que ni l'une ni l'autre ne puisse se douter que le canal de communication entre elles a été compromis. L'attaquant doit d'abord être capable d'observer et d'intercepter les messages d'une victime à l'autre. [16]

L'implémentation SSH au niveau du routeur Cisco permet de surpasses ce type d'attaque. Dans le cadre de la cryptographie asymétrique, les deux personnes possèdent chacune leur clé publique

(qui sert à chiffrer) et leur clé privée (qui sert à déchiffrer). Ainsi, seules les clés publiques sont échangées, ce qui ne nécessite pas un canal sécurisé. Même si quelqu'un réussissait à intercepter et à lire ces clés publiques, elles ne lui seraient d'aucune utilité pour déchiffrer.
Une attaque par relais, connu en anglais sous le nom de « relay attack », est un type d'attaque informatique, similaire à l'attaque de l'homme du milieu et l'attaque par rejeu, dans lequel un attaquant ne fait que relayer mot pour mot un message d'un expéditeur vers un récepteur valide.

1.7.3 L'attaque par DoS (Deny of Service)

On appelle « attaque par déni de service » toutes les actions ayant pour résultat la mise hors-ligne d'un serveur. Techniquement, couper l'alimentation d'un serveur dans un but malfaisant peut-être considéré comme une attaque par déni de service. Dans les faits, les attaques par déni de service sont opérées en saturant un des éléments du serveur ciblé.
Une des attaques les plus courantes consistait à envoyer un paquet ICMP de plus de 65 535 octets. Au-dessus de cette limite, les piles IP ne savaient pas gérer le paquet proprement, ce qui entraine des erreurs de fragmentation UDP, ou encore les paquets TCP contenant des « flags » illégaux ou incompatibles.
Le principe de DoS (Deny Of Service) est d'avoir un effet de levier en utilisant plusieurs sources pour l'attaque. [16]

1.7.3.1 Le SYN Flood

Une attaque SYN Flood est une attaque visant à provoquer un déni de service en émettant un nombre important de demandes de synchronisation TCP incomplète avec un serveur. Quand un système (client) tente d'établir une connexion TCP vers un système offrant un service (serveur), le client et le serveur échangent une séquence de messages. Le système client commence par envoyer un message SYN au serveur. Le serveur reconnaît ensuite le message en envoyant un SYN-ACK message au client. Le client finit alors d'établir la connexion en répondant par un message ACK. La connexion entre le client et le serveur est alors ouverte, et le service de données spécifiques peut être échangé entre le client et le serveur. La figure 1.15 montre ce flux de messages.

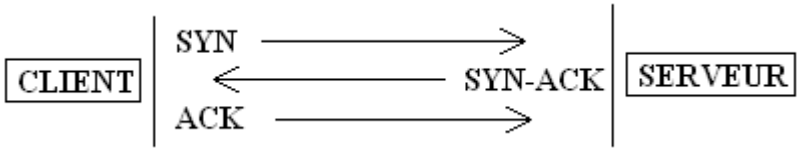


Figure 1.15 : *Scénario normal d'échange de demande de synchronisation TCP*

Créer des connexions semi-ouvertes s'accomplit facilement avec l'IP spoofing. Le système de l'agresseur envoie des messages SYN à la machine victime ; ceux-ci semblent être légitimes, mais font référence à un système client incapable de répondre au message SYN-ACK. Cela signifie que le message ACK final ne sera jamais envoyé au serveur victime.
Normalement il y a un délai d'attente associé à une connexion entrante, les semi-connexions ouvertes vont expirer et le serveur victime pourra gérer l'attaque. Toutefois, le système agresseur peut simplement continuer à envoyer des paquets IP falsifiés demandant de nouvelles connexions, plus rapides que le serveur victime.
Dans la plupart des cas, la victime aura des difficultés à accepter toute nouvelle connexion réseau entrante. Dans ces cas, l'attaque n'affecte pas les connexions entrantes, ni la possibilité d'établir des connexions réseau sortant. Toutefois, le système peut saturer la mémoire, ce qui provoque un crash rendant le système inopérant. [16]

1.7.3.2 L'UDP flooding

Ce déni de service exploite le mode non connecté du protocole UDP. Il crée un "UDP Packet Storm" (génération d'une grande quantité de paquets UDP) soit à destination d'une machine soit entre deux machines. Une telle attaque entre deux machines entraîne une congestion du réseau ainsi qu'une saturation des ressources des deux hôtes victimes. La congestion est plus importante du fait que le trafic UDP est prioritaire sur le trafic TCP. En effet, le protocole TCP possède un mécanisme de contrôle de congestion, dans le cas où l'acquittement d'un paquet arrive après un long délai, ce mécanisme adapte la fréquence d'émission des paquets TCP et le débit diminue. Le protocole UDP ne possède pas ce mécanisme. Au bout d'un certain temps, le trafic UDP occupe donc toute la bande passante, ne laissant qu'une infime partie au trafic TCP.
L'implémentation des ACL et du CBAC permet de réduire ce risque d'attaque SYN flood et UDP flooding. TCP established ACL et CBAC peuvent même lire en temps réel les états des fanions des paquets. [16]

1.7.3.3 Le Packet Fragment

Les dénis de service de type Packet Fragment utilisent des faiblesses dans l'implémentation de certaines piles TCP/IP au niveau de la défragmentation IP (ré-assemblage des fragments IP). Une attaque connue utilisant ce principe est Teardrop. L'offset de fragmentation du second fragment est inférieur à la taille du premier ainsi que l'offset plus la taille du second. Cela revient à dire que le deuxième fragment est contenu dans le premier (overlapping). Lors de la défragmentation,

certains systèmes ne gèrent pas cette exception et cela entraîne un déni de service.
Le système CBAC du routeur Cisco offre la possibilité de surveiller les offsets des trames qui transitent donc il permet d'éviter une telle attaque. [16]

1.7.3.4 Le Smurfing

Cette attaque utilise le protocole ICMP. Quand un ping (message écho ICMP) est envoyé à une adresse de broadcast (par exemple 10.255.255.255), celui-ci est démultiplié et envoyé à chacune des machines du réseau. Le principe de l'attaque est de truquer les paquets ICMP ECHO REQUEST envoyés en mettant comme adresse IP source celle de la cible. Le cracker envoie un flux continu de ping vers l'adresse de broadcast d'un réseau et toutes les machines répondent alors par un message ICMP ECHO REPLY en direction de la cible. Le flux est alors multiplié par le nombre d'hôtes composant le réseau. Dans ce cas tout le réseau cible subit le déni de service, car l'énorme quantité de trafic générée par cette attaque entraîne une congestion du réseau. [16]

1.7.4 Les utilisations détournées de TCP/IP

Lorsque l'on crée un outil, on n'a pas forcément conscience que celui-ci peut quelquefois être utilisé à des fins très différentes de celles prévues au départ. Ainsi, de même qu'un marteau sert initialement à enfoncer des clous mais peut également être utilisé pour provoquer des dommages physiques sur une personne, il existe des utilisations détournées de TCP/IP permettant de mener des "attaques" contre des machines ou des réseaux entiers. Nous nous intéresserons ici aux attaques que l'on peut mener via ou sur les couches Transport et IP. [16] Ces attaques sont de trois types:

- Reconnaissance : ces attaques ne sont pas "destructrices" au sens où elles empêchent une entité de fonctionner correctement, mais permettent d'acquérir des informations parfois cruciales pour mener une attaque de plus grande envergure plus tard. Il s'agit notamment :
    - des PORT SCANS (scanning de ports) qui permettent de savoir quels services sont actifs sur un poste donné. Certains scanners comme « nmap » ou « queso » permettent en plus de détecter le type de système d'exploitation surtout avec l'analyse de la structure des paquets (par exemple les paquets provenant de certaine variante d'UNIX ont des TTL, Time To Live, fixés à 64)

- des SWEEPS (balayage) qui permettent de détecter quelles machines sont actives sur le réseau, ce qui permet de dresser une carte de cibles potentielles. Ces balayages sont généralement effectués à l'aide de requêtes ICMP ("ping").
- solution "hybride" : il s'agit d'un balayage effectué sur un port particulier (par exemple, on contacte toutes les machines du réseau sur le port 25). L'attaquant cherche à déterminer quelles machines ont un certain type de service activé. Ces scans hybrides sont par conséquent utilisés couramment pour détecter les machines infectées par un trojan (en général le port est supérieur à 1024), ou bien si une faille existe sur le service visé.

- Déni de Service qu'on vient de décrire précédemment et qui engendre les dégâts suivants :
  - Consommation de bande passante : par inondation du réseau en utilisant la "force pure" (par exemple un réseau muni d'une connexion de type T1 - 1.5 Mbps - inondant un modem à 56 kbps) ou une méthode d'amplification (obtenu avec l'attaque smurfing)
  - Consommation de ressources : l'attaque vise à "brûler des cycles" sur la machine en faisant travailler le processeur pour rien, à occuper tout l'espace mémoire disponible ... en somme l'attaque vise les ressources du système plutôt que les ressources du réseau.
  - Failles dans les programmes : certains systèmes d'exploitation ne savent pas bien gérer les paquets inhabituels (non prévus dans les RFC) ou bien ne réagissent pas bien selon les stimuli. (obtenu avec es attaques de type teardrop)
  - Attaques sur DNS et Routage : pirater un cache de DNS ayant autorité pour une machine donnée peut empêcher quiconque souhaitant contacter cette machine d'y accéder, ou pire l'envoyer sur une machine-leurre. De telles attaques ont lieu souvent et peuvent servir par exemple à récupérer des mots de passe ou des numéros de carte de crédit si le site touché est un portail de e-business. Prenons un exemple typique sur internet :

    #1 le client souhaite se connecter au serveur web de microsoft, il interroge donc le serveur DNS de Microsoft pour connaitre l'IP associée à www.microsoft.com
    #2 l'attaquant a piraté le cache de DNS associé à microsoft, de façon à ce que l'adresse IP associée à www.microsoft.com soit en fait celle de www.hacker.com
    #3 le client est connecté à www.hacker.com en pensant être connecté à www.microsoft.com

De même, donner de fausses informations de routage à une machine peut l'empêcher d'accéder à des ressources, voire la bloquer complètement

- Détournement de connexion : par définition ces attaques ne concernent que TCP. Les attaquants utilisent une faille dans le modèle en couches de TCP/IP, à l'aide de la technique dite de prédiction de Sequence Number. Un attaquant se présente comme une machine autre (généralement une machine en laquelle la victime a "confiance", et qui a été mise hors service) et envoie sous cette identité un paquet de synchronisation de connexion (première étape du "3-way handshake" de TCP). La victime envoie son paquet de reconnaissance (SYN/ACK) avec son Sequence Number. Ce paquet ne peut être intercepté par l'attaquant en général, mais si celui-ci envoie un dernier paquet (ACK) avec son adresse réelle, en ayant deviné le SN de la victime correctement, alors la connexion sera établie par l'attaquant, et la victime croira avoir affaire à l'autre machine (celle en qui elle a "confiance"). Cette attaque permet de contourner les barrières comme les mots de passe: on attend que l'une des victimes ait établi une connexion complète et se soit authentifié avant de se faire passer pour lui.

1.8 Conclusion

Dans ce chapitre, on a commencé à donner un aperçu sur les réseaux et les plateformes sur lesquels les données transitent, on a même mis l'accent sur le model internet qui est le plus approprié pour représenter la fonction de routage et qui a été conçu pour servir de cadre opératoire aux protocoles fonctionnant sur des réseaux hétérogènes. TCP/IP est actuellement la famille de protocoles réseaux qui gère le routage la plus répandue sur les systèmes informatiques. Et dans la dernière partie, on a abordé les attaques visant les systèmes d'information que ce soit dans le but de découvrir des informations, d'attaques par intrusion et d'attaques d'interruption de service.

CHAPITRE 2
LES ROUTEURS CISCO ET LEUR SECURISATION

2.1 Introduction

On a vu précédemment que le rôle principal d'un routeur dans un WAN n'est pas seulement le routage mais principalement la compatibilité des connexions vers et entre les diverses normes physiques et de liaison de données d'un réseau WAN. Dans ce chapitre, on abordera les généralités sur les commandes de base des routeurs Cisco. Mais avant tous, sachons que Cisco, dont le siège social se trouve à San José en Californie, tire son nom et son logo de la ville où elle a été fondée, San Francisco et son fameux Golden Gate Bridge. La raison pourquoi on s'intéresse particulièrement à ce fabriquant est qu'il occupe plus de 70% du marché. La sécurisation à travers les routeurs Cisco est donc capitale dans le cadre du SSI. La deuxième partie de ce chapitre montre comment sécuriser l'accès à des réseaux particulier au moyen de firewalls Cisco, des Access listes et toutes les précautions à prendre relatif à la sécurisation du système d'information. Nous verrons comment configurer, déployer et maintenir ces équipements et mettre en place des filtrages de contenu et des Principe d'accès sécurisé AAA.

2.2 Les routeurs CISCO

2.2.1 L'architecture des routeurs Cisco

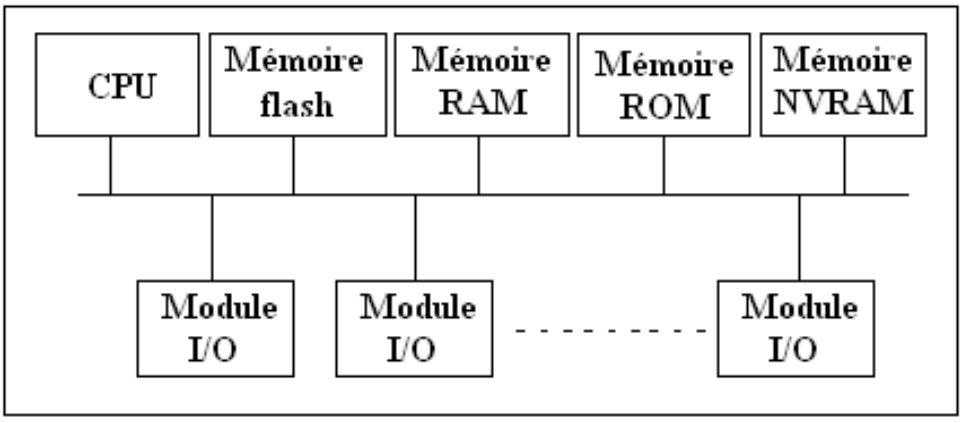


Figure 2.01 : *Architecture interne d'un routeur Cisco*

Les routeurs Cisco contiennent :

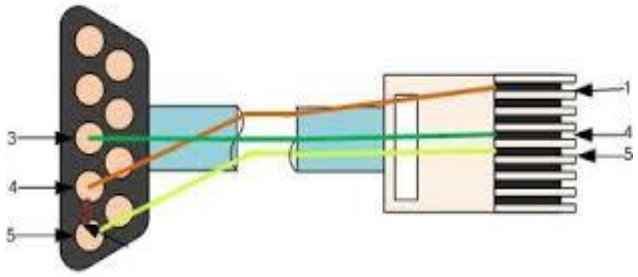
- Une carte mère qui est en général intégrée au châssis
- Une CPU qui est un microprocesseur Motorola avec un BIOS (Basic Input/Output System) spécial « Internetwork Operating System »
- Une mémoire NVRam (Non-Volatile Random Access Memory) pour Ram non Volatile et sur laquelle l'administrateur va stocker la configuration qu'il aura mise dans le routeur. Elle contient également la configuration de l'IOS (Internetworks Operating System)
- Une mémoire RAM (Random Access Memory) principale contenant le logiciel IOS, c'est dans laquelle tout sera exécuté un peu à la manière d'un simple ordinateur
- Une mémoire flash, également une mémoire non volatile sur laquelle on stocke la version courante de l'IOS du routeur
- Une mémoire ROM (Read-Only Memory) non volatile et qui, quant à elle, contient les instructions de démarrage (bootstrap) et est utilisée pour des opérations de maintenance difficiles de routages, ARP, etc.), mais aussi tous les buffers utilisés par les cartes d'entrée. [10]

2.2.2 L'IOS (Internetworks Operating System)

IOS est l'acronyme de "Internetworks Operating System" ou "Système d'exploitation pour l'interconnexion de réseaux". Ce système est administrable en lignes de commandes et est propres aux équipements de Cisco Systems. [10]

2.2.3 Le câble console Cisco

Quand les routeurs sont neuves ou quand on ne peut les accéder à distance (Telnet ou SSH), la connexion avec un câble console est la seule façon d'accéder à leur IOS. On utilise HyperTerminal (de Microsoft) pour effectuer les opérations nécessaires.



DB-9 Femelle RJ-45 Mâle

Figure 2.02 : *Modification du câble console Cisco*

2.2.4 Les commandes Cisco de base

Cette partie du chapitre 2 a pour but d'expliquer simplement les bases de la configuration et de l'administration de routeurs Cisco, des connaissances sur la fonction d'un routeur et les protocoles de routage nécessaires pour la sécurisation des routeurs Cisco. [10]

2.2.4.1 Les différents modes d'utilisateurs

L'IOS du routeur possède plusieurs privilèges d'utilisateur ou plus précisément d'administrateur :

- Mode Utilisateur: Permet de consulter toutes les informations liées au routeur sans pouvoir les modifier. Le shell est le suivant:

  Router >

- Utilisateur privilégié: Permet de visualiser l'état du routeur et d'importer/exporter des images d'IOS. Le shell est le suivant:

  Router #

- Mode de configuration globale: Permet d'utiliser les commandes de configuration générales du routeur. Le shell est le suivant:

  Router (config) #

- Mode de configuration d'interfaces: Permet d'utiliser des commandes de configuration des interfaces (Adresses IP, masque, etc.). Le shell est le suivant:

  Router (config-if) #

- Mode de configuration de ligne: Permet de configurer une ligne (exemple: accès au routeur par Telnet). Le shell est le suivant:

  Router (config-line) #

2.2.4.2 Résumé des commandes IOS de base

Si plusieurs commandes sont indiqués les unes en dessous des autres pour une même fonction, cela signifie qu'elles ont toute la même fonction et que l'une ou l'autre peut être utilisée au choix.

a) Passage entre les différents modes d'utilisateurs

- Utilisateur normal: Aucune commande à effectuer, c'est dans ce mode que commence une session.
- Utilisateur privilégié (à effectuer à partir du mode normal):

```
Router > enable
Router > en
```

- Mode de configuration globale (à effectuer à partir du mode Privilégié):

```
Router # configure terminal
Router # conf t
```

- Mode de configuration d'interface (à effectuer à partir du mode de configuration globale):

```
Router (config) # interface nom_interface
Router (config) # int nom_interface
```

- Mode de configuration de ligne (à effectuer à partir du mode de configuration globale):

```
Router (config) # line nom_de_la_ligne
```

b) Commandes d'information

Les commandes d'information permettent d'afficher les informations relatives au routeur. Elles commencent toutes avec le préfixe show ou sh. Elles sont, pour la plupart, à effectuer à partir du mode privilégié.

- Afficher le fichier de configuration courante du routeur:

```
show running-config
show run
sh run
```

- Afficher les informations sur la configuration matérielle du système et sur l'IOS:

```
show version
sh version
```

- Afficher les processus actifs:

```
show processes
```

- Afficher les protocoles configurés de couche 3 du modèle OSI:

```
show protocols
```

- Afficher les statistiques de mémoire du routeur:

```
show memory
```

- Afficher des informations et statistiques sur une interface:

```
show interfaces nom_interface
sh interfaces nom_interface
sh int nom_interface
```

- Afficher la table de routage IP:

```
sh ip route
```

c) Commandes d'enregistrement de la configuration courante

Ces commandes permettent de sauvegarder la configuration actuelle pour la réappliquer automatiquement en cas de redémarrage du routeur. Elles s'exécutent en mode Privilégié

- Sauvegarde avec demande de confirmation:

```
copy running-config startup-config
copy run start
```

- Sauvegarde sans demande de confirmation:

```
write
```

d) Commande d'annulation

Cette commande permet de revenir à la dernière configuration enregistrée, annulant toutes les modifications ayant été faites à la configuration depuis. Elle s'exécute en mode Privilégié.

```
copy startup-config running-config
copy start run
```

e) Annulation d'une commande particulière

Pour annuler une commande particulière, on utilisera le préfix no devant la commande précédemment exécutée. Par exemple, on annule la configuration d'une interface:

```
no ip address
```

f) Changer le nom du routeur

Le nom du routeur peut être modifié afin de permettre de les différencier sur le(s) réseau(x). La commande sera exécutée en mode de configuration globale.

```
host NouveauNom
```

Concrètement, un nom différent s'affichera lors de l'invite de commande (prompt) des sessions HyperTerminal ou Telnet.

- Avant le changement de nom :

```
Router >
```

- Après le changement de nom :

```
NouveauNom >
```

2.2.4.3 Activation des interfaces Ethernet du routeur

Pour faire communiquer les hôtes connectés au routeur. Admettons que le nom de l'interface reliée à un premier hôte est fa0/0 et celle reliée à un deuxième hôte est fa0/1 et qu'on est en mode de configuration globale. [10]

Voici les commandes à saisir:

- Interface fastEthernet 0/0:

```
Router > enable
Router # configure terminal
Router(config) # interface fastEthernet 0/0
Router (config-if) # ip address 192.168.1.1 255.255.255.0
Router (config-if) # no shutdown
Router (config-if) # exit
```

- Interface fastEthernet 0/1:

```
Router > enable
Router # configure terminal
Router(config) # interface fastEthernet 0/1
Router (config-if) # ip address 10.0.0.1 255.0.0.0
Router (config-if) no shutdown
Router (config-if) exit
```

Ces commandes sont liées à la configuration des interfaces du routeur. Elles sont, pour la plupart, à effectuer à partir du mode de configuration d'interface.

- Attribution d'une adresse IP à une interface:

```
ip address @IP masque
```

- Activation de l'interface:

```
no shutdown
```

2.2.4.4 Association d'un VLAN avec un module de commutation d'un routeur

Pour mieux aborder l'implémentation du SSH, prenons le cas du routeur Cisco 2811 munie d'un module switch 16 ports référencé NM-ESW-161. [10]

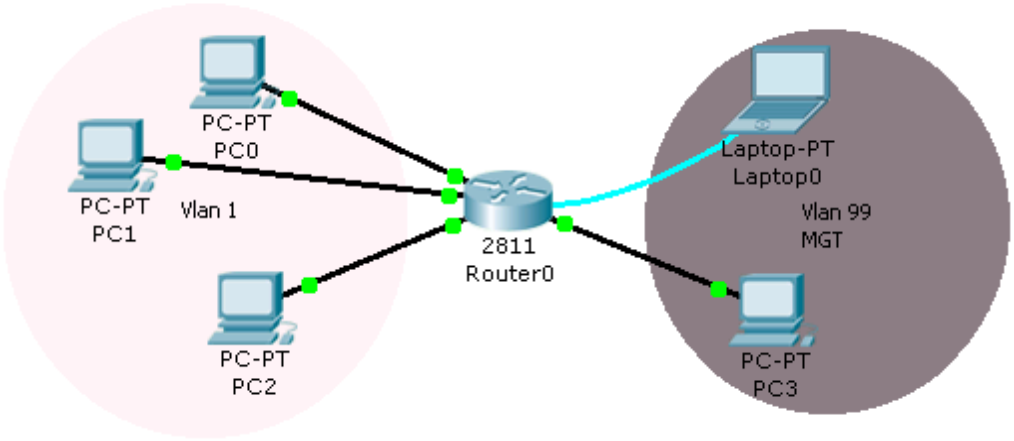


Figure 2.03 : *Architecture de plusieurs VLAN dans un routeur*

Tous d'abord créons un LAN virtuel (VLAN) que nous associeront ensuite avec un ou plusieurs port physique.

Comparons les procédés pour configurer un switch Cisco et un routeur Cisco

```
Switch > enable
Switch # configure terminal
Switch(config) # vlan 99
Switch(config-vlan) # name MGT
Switch(config-vlan) # end
Switch(config) #
```

Celui du routeur est

```
Router # vlan database
Router(vlan) # vlan 99 name MGT
Router(vlan) # exit
Router(config) #
```

Ici, le nom du LAN virtuel est par exemple MGT (acronyme du Management, c'est-à-dire un sous-réseau associé à l'administration du réseau global de l'entreprise). Associons maintenant ce VLAN avec le port physique 15 par exemple.

```
Router # configure terminal
Router(config) # interface fastEthernet 1/15
Router(config-if) # switchport mode access
Router(config-if)# switchport access vlan 99
Router(config-if) # no shutdown
Router(config-if)# end
```

Donnons un adresse au VLAN.

```
Router # configure terminal
Router(config) # interface vlan 99
Router(config-if) # ip address 192.168.99.2 255.255.255.0
Router(config-if) # no shutdown
Router(config-if) # end
```

2.2.4.5 Le protocole RIP (Routing Information Protocol)

Il suffit juste d'énumérer les réseaux directement rattachés au routeur.

```
Router # configure terminal
Router(config) # router rip
Router(config-router) # network 10.0.0.0
Router(config-router) # network 192.168.1.0
Router(config-router) # end
```

2.2.4.6 Route statique

Il s'agit d'enregistrer dans le routeur le réseau de destination en passant par l'interface du prochain pas. [10]

```
Router # configure terminal
Router(config) # ip route < adresse du réseau destination > < masque du réseau de destination > < interface du prochain pas >
```

2.2.4.7 Configuration du nuage Frame Relay

Le côté DCE de la connexion série est traité par le nuage Frame Relay. Le nuage simule l'interconnections de plusieurs commutateurs Frame Relay.

A l'intérieur de la configuration des ports séries, le DLCI est associé au routeur de destination. Par exemple, le DLCI = 103 du routeur R1 indiquant un parcours allant du Routeur R1 à R3 est associé au routeur R3, et ainsi de suite.

A l'intérieur de la configuration du Frame Relay, on associe le port série à l'entrée du nuage Frame Relay au routeur de destination à la sortie du nuage et on procède inversement pour obtenir un chemin aller-retour.

En ce qui concerne le routage statique du réseau Frame Relay, il faut tout d'abord changer l'encapsulation Ethernet en une encapsulation Frame Relay sur l'interface frontale après avoir attribué un adresse IP et un « no shutdown » à ce dernier.

Puis on définit la bande passante qui est exprimé en pourcentage. Et enfin on dresse le map (carte)

de parcours en spécifiant l'interface frontale de destination, le DLCI et le mot clé « broadcast ». Le tout est finalisé par « no shutdown ».

Le protocole de routage le plus pratique à implémenter dans ces cas là est le RIP. [10]

```
Router # configure terminal
Router(config) # interface serial 2/0
Router(config-if) # encapsulation frame-relay
Router(config-if) # bandwidth 64
Router(config-if) # frame-relay map IP 10.0.0.2 102 broadcast
Router(config-if) # end
```

2.3 La sécurisation des routeurs CISCO

2.3.1 Les raisons de la sécurisation

Internet a transformé et nettement amélioré les transactions commerciales, ce vaste réseau et les technologies qui lui correspondent ont ouvert la porte à un nombre croissant de menaces relatives à la sécurité contre lesquelles les entreprises doivent se prémunir. Bien que les attaques des réseaux soient généralement plus graves lorsqu'elles visent des sociétés qui stockent des données critiques, comme des dossiers confidentiels médicaux ou financiers, les conséquences de ces attaques sur une entreprise peuvent aller d'un léger désagrément à une paralysie complète de l'activité, des données importantes peuvent être perdues, la confidentialité peut être transgressée et plusieurs heures ou jours d'interruption du réseau peuvent s'en suivre. Maintenant, plus que jamais, il est impératif que les entreprises intègrent la sécurité au sein de l'architecture de leur réseau afin de limiter ces risques et de concrétiser le potentiel de croissance inhérent à l'environnement de réseau. [16]

2.3.2 Les acteurs des attaques

- Pirates informatiques (hackers)

  Ce terme générique s'applique aux passionnés d'informatique s'amusant à accéder aux ordinateurs et aux réseaux d'autres personnes. Le plus souvent, leur motif est le défi.

- Personnel non avisé

  Il arrive souvent que des employés, concentrés sur leurs activités professionnelles spécifiques outrepassent les règles de base de sécurité du réseau. Ils peuvent, par exemple, choisir des mots de passe simples à mémoriser afin de se connecter aisément au réseau. Ces mots de passe sont alors faciles à deviner ou à forcer par les pirates. Les employés peuvent involontairement être la source de failles dans la sécurité.

- Employés mécontents

  Ce problème est bien plus troublant que l'éventualité d'une erreur humaine endommageant le réseau : un employé mécontent peut vouloir nuire à l'entreprise. Les employés mécontents, souvent à la suite d'un licenciement ou d'une remontrance, peuvent intentionnellement supprimer des fichiers importants ou encore en accédant à des données confidentielles afin de fournir aux concurrents des informations qu'ils n'auraient pas pu obtenir d'une autre manière. [16]

2.3.3 Les sécurisations à effectuer sur les routeurs Cisco

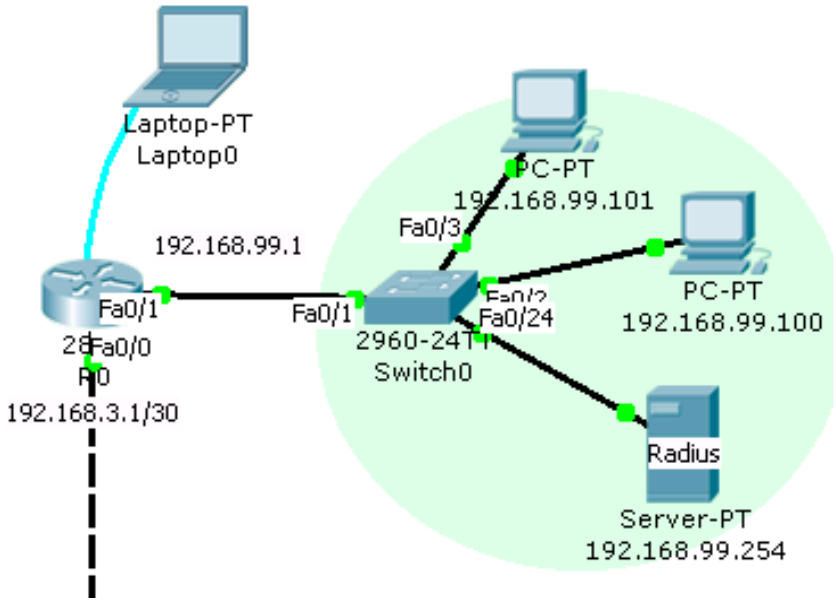


Figure 2.04 : *Architecture illustrant l'accès console et l'accès à distance*

2.3.3.1 Protection des ports physiques

a) Protection de l'accès local/physique (port console)

```
Router > enable(mode d'utilisateur privilégié)
Router # configure terminal(mode de configuration global)
Router(config) # line consol 0
Router(config-line) # password console@123456789!#
Router(config-line) # login
Router(config-line) # end
Router #
```

Il est aussi conseillé de définir un intervalle de temps dans laquelle l'utilisateur doit saisir le mot de passe exacte. Le but ici est d'empêcher les attaques par dictionnaire :

```
Router > enable
Router # configure terminal
Router(config) # line consol 0
Router(config-line) # exec-timeout 1
Router(config-line) # end
Router #
```

Une fois le délai dépassé, si l'utilisateur n'a pas trouvé le mot de passe exacte, la notification *«Password: timeout expired! Login invalid »* s'affiche. Si le processus se trouve déjà à l'intérieur d'une session utilisateur, une inactivité pendant 1 mn (dans l'exemple ci-dessus) déconnectera la liaison entre le PC et le routeur.

Les recommandations du CCNA Cisco concernant les mots de passe sont :

- la longueur minimale des caractères est égale à 10
- le mot de passe doit être constitué de lettre, de chiffre, de caractères spéciaux, de minuscule et de majuscule

- aucune relation avec les noms, les anniversaires, … il ne doit y avoir des relations qu'on peut facilement reconnaitre

Le manœuvre est identique pour la protection de l'accès à distance. [17]

b) Protection de l'accès à distance (Telnet)

```
Router > enable
Router # configure terminal
Router(config) # line vty 0 4
Router(config-line) # password vty@123456789!#
Router(config-line) # login
Router(config-line) # end
Router #
```

« 0 4 » signifie que le routeur peut supporter 5 session Telnet simultanées.

c) Protection de l'accès à distance par SSH (Secure SHell)

Telnet n'est pas sécurisé, il n'est pas chiffré. Les logiciels malveillants peuvent écouter la communication sur le réseau et peuvent même récupérer les informations confidentielles. C'est pourquoi on utilise SSH.

L'étape suivant consiste à désactiver Telnet et à active SSH s'il s'agit d'un Switch Cisco. Si c'est un routeur, les deux modes d'accès (Telnet et SSH) seront tous les deux valables. Avant toute chose, il faut s'assurer que le nom d'hôte du périphérique Cisco soit différent du nom d'hôte par défaut. Ce nom d'hôte sera utilisé pour le nom de la clé publique du chiffrement RSA.

```
Router # configure terminal
Router(config) # hostname Routeur1
Routeur1(config) # ip domain-name mpoina.com
Routeur1(config) # crypto key generate rsa
```

Le périphérique Cisco affichera le nom de la clé publique par une notification " *The name for the keys will be: Routeur1.mpoina.com*". Après cela il demandera le nombre de bit (entre 360 bits et 2048 bits) pour générer la clé privé RSA par la notification

How many bits in the modulus [512]: 1024

La recommandation de Cisco est 1024 bits, plus le nombre de bit est grand plus cela va prendre du temps. Puis après viennent les syntaxes suivantes

```
Routeur1(config) # ip ssh version 2
Routeur1(config) # end
```

Maintenant pour accéder à distance au switch Cisco, la commande « *telnet 192.168.99.2* » ne marche plus. A la place, nous utiliserons « *ssh –l admin 192.168.99.2* ». Pour les routeurs Cisco, les deux modes d'accès marchent. Il suffit donc de choisir entre le SSH ou le Telnet.
Pour plus de sécurité, il est recommandé d'utiliser des noms d'utilisateur autre qu'Admin qui est l'utilisateur par défaut. Pour ce faire, voir l'implémentation du protocole AAA plus bas.

Remarque : Si les routeurs Cisco prennent en charge la fonction commutation avec des modules additionnelles qu'on insère dans le boîtier du routeur, on peut associer un ou plusieurs de ces interfaces à un adresse IP, c'est-à-dire à un VLAN, pour pouvoir accéder à distance au routeur avec SSH ou Telnet. Cette démarche est détaillée dans la partie Association d'un VLAN avec un module de commutation d'un routeur du chapitre 2. [17]
Voici quelque module de commutation type :

- Le module HWIC-4ESW offre 4 ports de commutation RJ 45

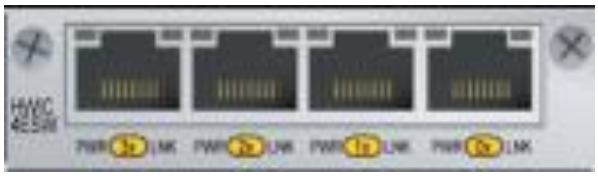


Figure 2.05 : *Module HWIC-4ESW offre 4 ports*

- Le module NM-ESW-161 offre 16 ports de commutation RJ45

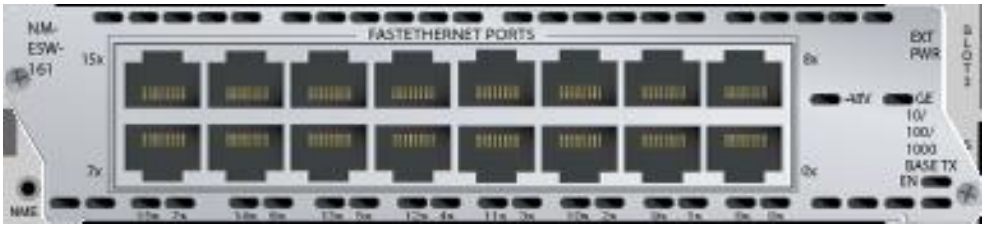


Figure 2.06 : *Module NM-ESW-161 offre 16 ports*

Pour terminer, il ne faut pas oublier de protéger l'accès vty par un mot de passe.

```
Router # configure terminal
Router(config) # line vty 0 15
Router(config-line) # password vty@123456789!#
Router(config-line) # end
```

2.3.3.2 Protection du mode d'utilisateur privilégié [17]

a) Protection de l'accès utilisateur privilégié par « enable password »

```
Router > enable
Router # configure terminal
Router(config) # enable password enable@123456789!#
Router(config) # end
```

b) Protection de l'accès utilisateur privilégié par « enable secret » (chiffrement MD 5)

```
Router > enable
Router # configure terminal
Router(config) # enable secret enable@123456789!#
Router(config) # end
```

Remarque : *enable password* n'est plus valable donc il faut l'enlever.

```
Router > enable
Router # configure terminal
Router(config) # no enable password
Router(config) # end
```

2.3.3.3 Chiffrement de tous les mots de passe (chiffrement type/level 7) [17]

```
Router > enable
Router # configure terminal
Router(config) # service password-encryption
Router(config) # end
```

2.3.3.4 Utilisation du protocole AAA (Authentification – Autorisation – Administration)

a) AAA intégré dans le routeur (stocké dans base de donnée du routeur)

L'implémentation du protocole AAA a pour but de récupérer dans une base de donnée les informations relatives à l'authentification d'un utilisateur. Cette base de données peut être stockée à l'intérieur même du routeur ou sur un serveur. Voyons en premier lieu le cas où elle se trouve à intérieur du routeur [17]. Pour un accès physique au port console, il faut d'abord entrer dans la ligne console, ensuite créer la session de l'utilisateur :

```
Router > enable
Router # configure terminal
Router(config) # line consol 0
Router(config-line) # login local
Router(config-line) # username mpoina secret mpoina@123456789!#
Router(config) # end
```

Pour un accès distant :

```
Router > enable
Router # configure terminal
Router(config) # line vty 0 4
Router(config-line) # login local
Router(config-line) # username mpoina secret mpoina@123456789!#
Router(config) # end
```

Remarque : le nom d'utilisateur par défaut « admin » est désactivé une fois qu'on a activé un profil d'utilisateur avec le protocole AAA, par conséquent on ne peut plus accéder avec le mot de passe générale de la line vty, on doit donc l'enlever. Le but est que tout individu qui veut gérer le routeur possède son propre compte. L'utilisation d'un mot de passe commun n'est pas conseillée. [17]

```
Router > enable
Router # configure terminal
Router(config) # line consol 0
Router(config-line) # no password
Router(config) # end
```

```
Router > enable
Router # configure terminal
Router(config) # line vty 0 4
Router(config-line) # no password
Router(config) # end
```

L’ajout ultérieur d’autre profile peut s’effectuer directement par :

```
Router > enable
Router # configure terminal
Router(config) # username jean secret jean@123456789!#
Router(config) # end
```

Pour plus de sécurité, on peut fixer la longueur minimale des mots de passe :

```
Router > enable
Router # configure terminal
Router(config) # security passwords min-length 10
Router(config) # end
```

Le cas échéant affichera la notification “*Password too short - must be at least 10 characters. Password not configured*” où 10 est la valeur de “*min-length*”

b) Serveur AAA Radius ou TACACS+

Dans le cas où il y a beaucoup de routeur et beaucoup de switch dans le réseau, il est plus approprié de centraliser les authentifications, c'est-à-dire utiliser un serveur Radius ou TACACS + qui se chargera d’authentifier les utilisateurs afin d’éviter de configurer un par un ces équipements CISCO. [17]

Les configurations à prendre en compte du côté du serveur Radius sont les suivants :

- le Nom du client est celui du routeur qui consultera les authentifications dans le serveur
- l’Adresse IP du client est celui du routeur aussi
- le Secret est le mot de passe pour crypter les différents échanges entre le routeur et le serveur
- le User setup contient la liste des noms d’utilisateur et les mots de passes associés pour l’authentification

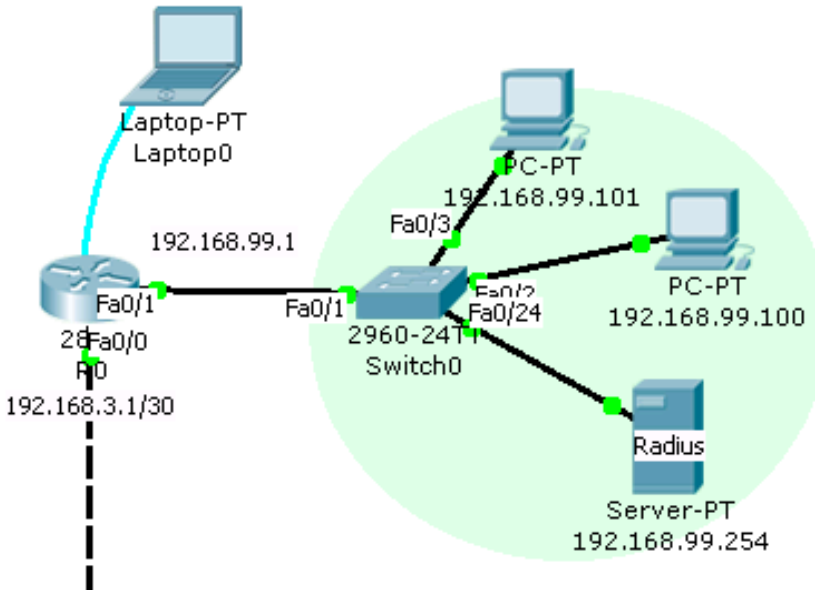


Figure 2.07 : *Architecture illustrant l'utilisation de serveur Radius*

Les lignes de commandes du côté du routeur sont les suivantes :

Router > enable

Router # configure terminal

Router(config) # aaa new-model

Router(config) # radius-server host 192.168.1.2 key secretrouteurserveur

Router(config) # aaa authentication login INSTANCE_AUTHENTIFICATION group radius

Router(config) # line vty 0 4

Router(config-line) # login authentication INSTANCE_AUTHENTIFICATION

Router(config-line) # end

2.3.4 Protection contre l'attaque par authentification

La syntaxe suivant permet de bloquer toute tentative excepté les comptes préétablis dans un « Acces List » (nous verrons les Access List plus tard).

```
Router > enable
Router # configure terminal
Router(config) # login block-for 180 attempts 5 within 60
Router(config) # login on-failure
Router(config) # login on-success
Router(config) # login quiet-mode
Router(config) # end
```

Cette syntaxe bloque pendant 180 secondes toutes tentatives d'authentification si dans un intervalle de 60 secondes il y a 5 mauvaises authentifications. La notification sera «... *Connection refused by remote host* ».

Les lignes de commandes « *login on-failure* » et « *login on-success* » permettent de capturer les authentifications que ce soit réussite ou échec.

Le terme « *login quiet-mode* » a pour but de bloquer toute tentative d'authentification en cas de connexion refusée par l'hôte, excepté les utilisateurs listés dans une liste d'accès. Cette ligne de commande n'est pas supportée par tous les routeurs. [16]

2.3.5 Utilisation du contrôle de liste d'accès ou Accès List Contrôle - ACL

2.3.5.1 Standard Access List

Les étapes à suivre :

- créer l'Access List
- appliquer cette Access List à une interface (entrante ou sortante)

L'Access List standard recevra un numéro entre 1 et 99, elle sera appliquée au routeur le plus proche de la destination. La destination ici est l'hôte qui fera l'objet de la restriction ou non. [17]

La figure qui suit explique la création de l'Access List standard:

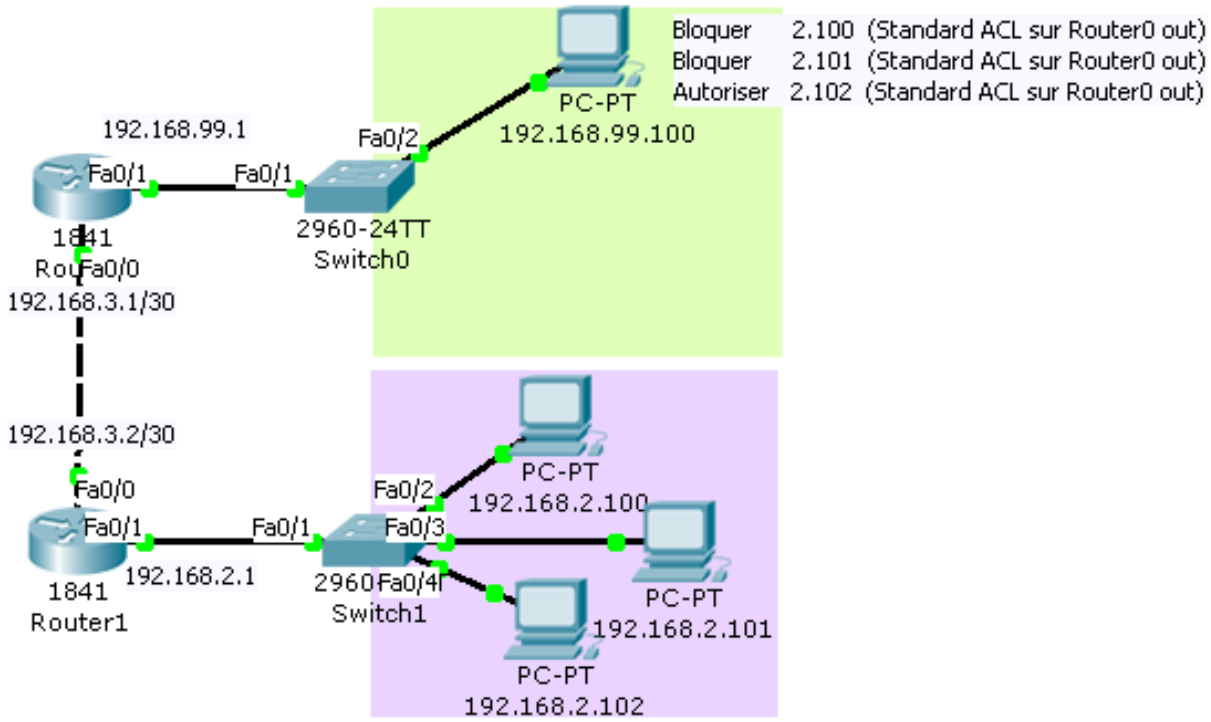


Figure 2.08 : *Architecture illustrant l'utilisation de l'ACL Standard*

Router > enable

Router # configure terminal

Router(config) # access-list 1 deny 192.168.2.100 0.0.0.0 ← 100 établie la liaison ?

Router(config) # access-list 1 deny 192.168.2.101 0.0.0.0 ← sinon, est-ce 101?

Router(config) # access-list 1 permit any ← autoriser les autres

Router(config) # access-list 1 deny any ← implicite et invisible

(Les déclarations générales en bas de la liste)

Router(config) # end

"1" est le numéro de l'Access List, « 192.168.2.100 » et « 192.168.2.101 » sont les hôtes qu'on va bloquer, « 0.0.0.0 » est le « wildcard bits » qui sert à bloquer un seul hôte. Si on veut bloquer le réseau 192.168.2.0 par exemple, le wildcard correspondant sera 0.0.0.255.

On peut remplacer la phrase

```
Router(config) # access-list 1 deny 192.168.2.101 0.0.0.0
```

Par

```
Router(config) # access-list 1 deny host 192.168.2.101
```

Les mots « permit any » permet d'autoriser tout autre terminal.

Appliquons maintenant cette Access List avec l'interface 0/1 sortant du routeur R0 (l'interface sortant est le plus proche de la destination) :

```
Router > enable
Router # configure terminal
Router(config) # interface fastEthernet 0/1
Router(config-if) # ip access-group 1 out
Router(config) # end
```

Les règles relatives à l'établissement d'un Access List :

- la lecture de la liste est procédurale, une fois qu'il y a une concordance, la lecture de la liste s'arrête.
- les déclarations spécifiques doivent être placées en haut de la liste
- les déclarations générales doivent être placées en bas de la liste
- à la fin de tout Access List se trouve un implicite (invisible mais existe) « deny any » c'est-à-dire refuser tous les autres terminaux.

Seul l'hôte 102 du réseau 2 peut accéder aux hôtes du réseau 99. L'Access List standard doit être appliquée à l'interface le plus proche de la destination. Il ne faut pas oublier de préciser les routes (statiques ou dynamique) pour la communication. [17]

2.3.5.2 Extended Access List

Ce type de liste d'accès est implémenté sur l'interface le plus proche de la source.

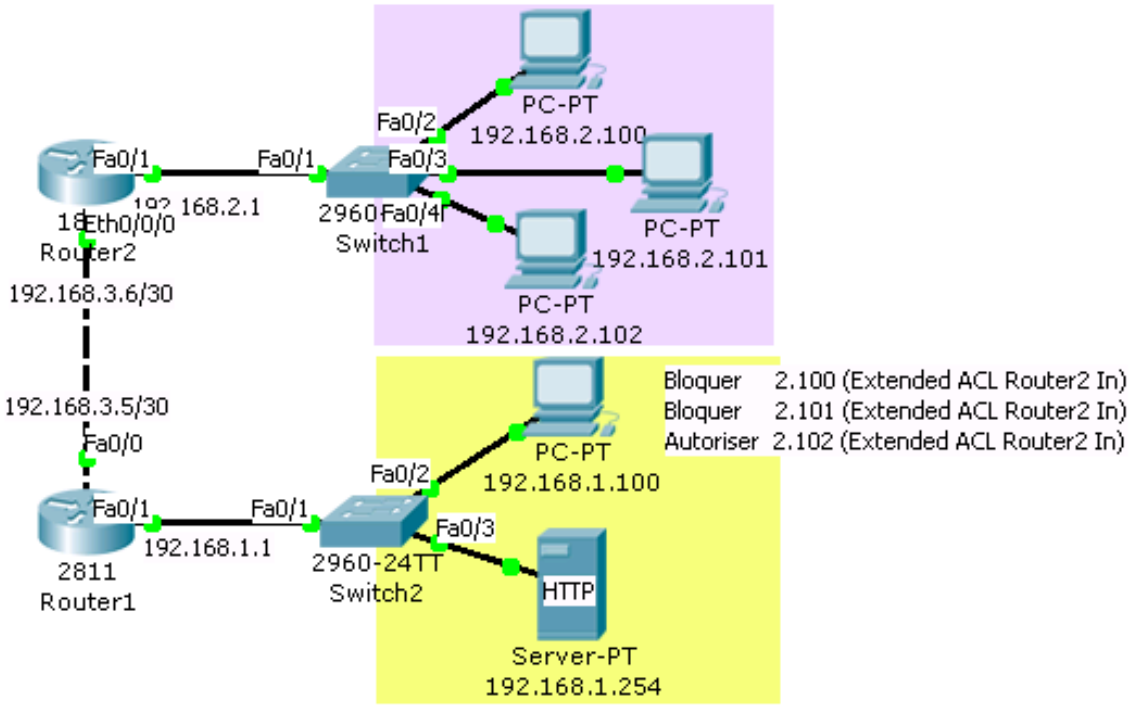


Figure 2.09 : *Architecture illustrant l'Extended ACL*

La liste d'accès étendue est de la forme :

> *Router(config) # access-list <numero de l'ACL> <permit ou deny> <tcp ou ip> <adresse de l'hôte ou du réseau source de la restriction> <wildcard bits> <adresse de l'hôte ou du réseau destination de la restriction> <wildcard bits>*

Pour le cas de la figure 3.06 :

> *Router(config) # access-list 100 deny ip host 192.168.2.100 192.168.1.0 0.0.0.255*
>
> *Router(config) # access-list 100 deny ip host 192.168.2.101 192.168.1.0 0.0.0.255*
>
> *Router(config) # access-list 100 permit ip any any*
>
> *Router(config) # access-list 100 deny ip any any* ← *implicite et invisible*

L'Extended ACL prend un numéro entre 100 et 199.

Il ne faut pas oublier d'appliquer cette liste à l'interface le plus proche de la source.

On associe souvent l'Extended ACL avec des services (protocoles) et non seulement avec des hôtes. Si on prend l'exemple des protocoles http et https, l'ACL devient :

```
Router(config) # access-list 199 permit tcp 192.168.2.0 0.0.0.255 192.168.1.254 0.0.0.0 eq 80
Router(config) # access-list 199 permit tcp 192.168.2.0 0.0.0.255 192.168.1.254 0.0.0.0 eq 443
Router(config) # access-list 199 permit ip 192.168.2.0 0.0.0.255 192.168.4.0 0.0.0.255
Router(config) # access-list 199 deny ip any any                    ← implicite et invisible
```

Le terme « *eq* » ordonne au routeur de n'accepter que les paquets provenant du numéro de port spécifié. [17]

2.3.5.3 Named Access List (ACL nommée)

Il s'agit de nommer les Access List que ce soit standard ou étendue. Pour le même cahier de charge type, on obtient la liste d'accès nommée suivante :

```
Router(config) # ip access-list extended WEB
Router(config-ext-nacl) # permit tcp 192.168.2.0 0.0.0.255 192.168.1.254 0.0.0.0 eq 80
Router(config-ext-nacl) # permit tcp 192.168.2.0 0.0.0.255 192.168.1.254 0.0.0.0 eq 443
Router(config-ext-nacl) # permit ip 192.168.2.0 0.0.0.255 192.168.4.0 0.0.0.255
```

WEB (World Wide Web) est le nom qu'on va donner à la liste d'accès. Apres il faut applique cette accès liste à l'interface la plus proche de la source c'est-à-dire un liste d'accès entrant. [17]

2.3.6 Les technologies pare-feu de Cisco

2.3.6.1 TCP Established ACL

Le cahier de charge type est le suivant :

- le réseau protégé ne doit être visible de l'extérieur
- le DMZ doit être accessible de l'extérieur
- les hôtes du réseau protégés doivent accéder aux serveurs extérieurs

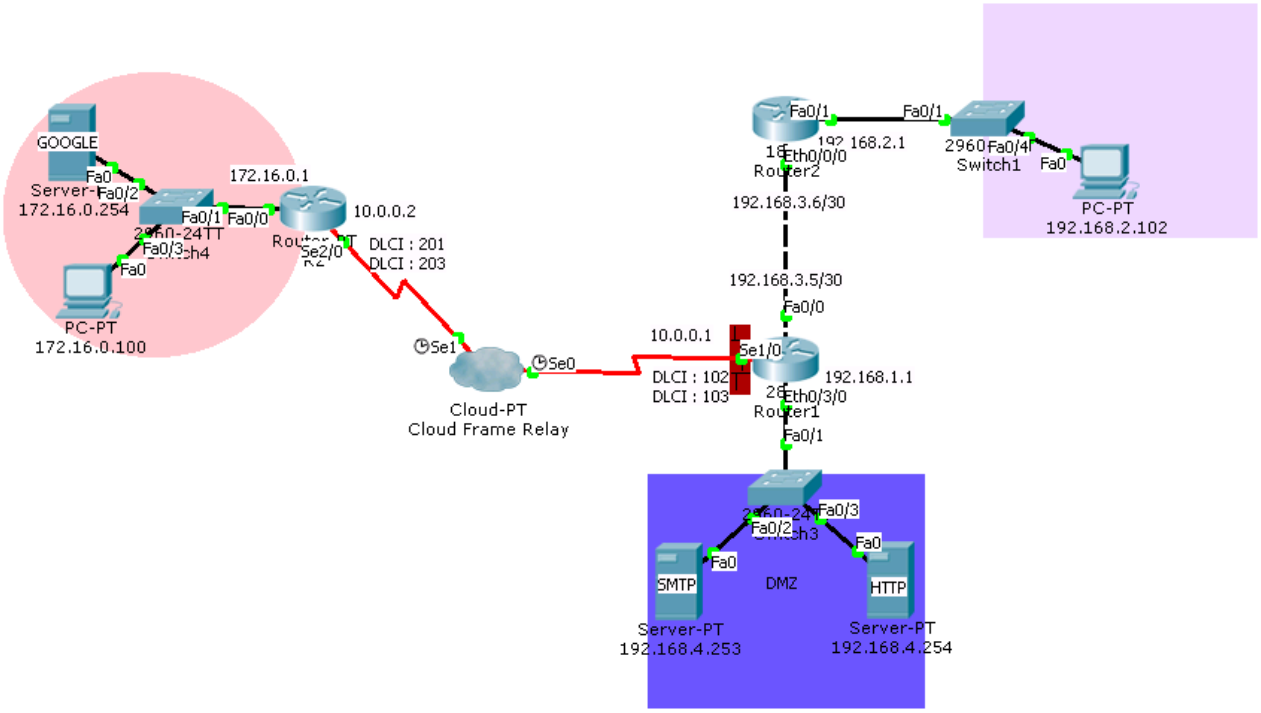


Figure 2.10 : *Architecture illustrant le TCP Established ACL*

La liste d'accès doit donc jouer le rôle de pare-feu :

```
Router(config) # access-list 199 permit tcp any host 192.168.4.254 eq 80
Router(config) # access-list 199 permit tcp any eq 80 192.168.2.0 0.0.0.255
Router(config) # access-list 199 deny ip any any
```

On remarque que le terme "eq" se trouve juste après l'hôte (ou le réseau) qui fournit le service.

Il est plus pratique d'utiliser un Named ACL :

```
Router(config) # ip access-list extended FIREWALL
Router(config-ext-nacl) # permit tcp any host 192.168.4.254 eq 80
Router(config-ext-nacl) # permit tcp any eq 80 192.168.2.0 0.0.0.255
```

Il ne faut pas oublier d'appliquer la liste d'accès à l'interface la plus proche de la source.

Concernant la 3è recommandation du cahier de charge, les requêtes peuvent sortir vers le nuage extérieur, mais les réponses ne peuvent pas entrer vers le réseau protégé. C'est pourquoi on raisonne à partir du point de vue de l'extérieur en attribuant la source au réseau extérieur.

Pour plus de sécurité, utilisons le mot clé «*established*» :

Revoyons l'architecture de la figure 3.07, quand l'hôte 102 du réseau 2 envoie une requête http, le serveur du réseau 3 envoie une réponse. Le tout se passe dans le controle de la liste d'accès. Supposons que l'hôte 172.16.0.100 est un pirate qui souhaite accéder au serveur stratégique du réseau protégé (serveur FTP par exemple), supposons qu'il connaît l'adresse du serveur FTP (il l'utilisera comme adresse de destination), qu'il connaît que le port http est autorisé par la liste d'accès (c'est une évidence et il l'utilisera comme port source) et le port de destination sera donc le port FTP (port 21). Le pirate n'a que créé un PDU de taille inférieur à 1400 et ce sera accepté à passer à travers le pare-feu.

Par contre, si on ajout le mot clé « *established* » à la liste d'accès, la liste d'accès contrôlera le flag de contrôle dans le champ TCP du segment du couche 4 (ACK, FIN, PSH, RST, SYN, URG) et saura que ce dernier et une réponse d'un requête provenant du réseau interne protégé. [17]

```
Router(config) # ip access-list extended FIREWALL
Router(config-ext-nacl) # permit tcp any host 192.168.4.254 eq 80
Router(config-ext-nacl) # permit tcp any eq 80 192.168.2.0 0.0.0.255 established
```

2.3.6.2 CBAC (Context-Based Access Control)

La technologie CBAC est une option disponible dans les routeurs Cisco permettant de spécifier quels trafics seront autorisés à passer à travers le routeur. Le CBAC inspecte les trafics (les services) spécifiés qui traversent le routeur. Cette inspection assurera que le trafic n'a pas subi d'intrusion extérieure et qu'il a été initialisé de l'intérieur du réseau. Des entrées dynamiques et temporaires autoriseront les trafics entrants venant de l'extérieur (venant de l'hôte destination). Le routeur lit donc les paquets tels que les contrôles de connexion TCP (les drapeaux ACK, SYN, RST, FIN), les numéros de séquences. La technologie CBAC est plutôt orientée protocole, elle inspecte les requêtes et les réponses DNS par exemple, les types de messages ICMP, le FTP, le TFTP. La CBAC peut même lire les translations d'adresse IP (NAT ou Network Address Tranlsation) et les translations d'adresse de port (PAT ou Port Address Tranlsation).

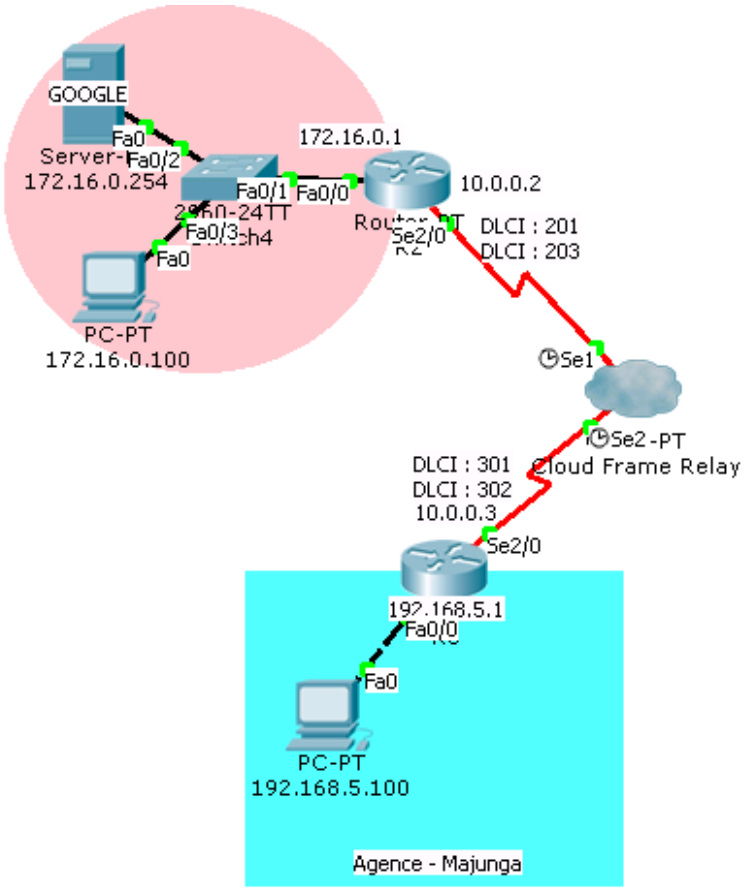


Figure 2.11 : *Architecture illustrant le CBAC*

Le principe est simple, il faut créer 2 ACL. L'un des ACL se trouve à l'intérieur du réseau et celui-ci s'assurera l'inspection ; l'autre ACL se trouve à l'extérieur, il bloquera par défaut tous ce qui n'est pas autorisé.

Voici les démarches proposés par Cisco pour implémenter CBAC : [17]

Tache 1: Bloquer les trafics venant de l'extérieur

- Démarche 1 : Vérifier les connectivités basiques

  Vérifier la connectivité des réseaux pour configurer le pare feu

  - A partir du poste 192.168.5.100, effectuer un ping vers le serveur 172 .16.0.254
  - A partir du poste 192.168.5.100, ouvrir le navigateur web pour atteindre la page web du serveur 172.16.0.254
  - A partir du serveur 172.16.0.254, effectuer un ping vers l'hôte 192.168.5.100.
- Démarche 2 : Configurer un ACL IP nommé sur le routeur R3 pour bloquer tout trafic généré par le réseau extérieur

  Utiliser la commande « ip access-list extended » pour créer l'ACL IP nommé

```
Router # configure terminal
Router(config) # ip access-list extended OUT-IN
Router(config-ext-nacl) # deny ip any any
```

- Démarche 3 : Appliquer l'ACL à l'interface Se2/0 du routeur R3

```
Router # configure terminal
Router(config) # int serial 2/0
Router(config-if) # ip access-group OUT-IN in
```

- Vérifier que les trafics entrants par l'interface Se2/0 sont bloqués

  A partir de l'hôte 192.168.5.100, effectuer un ping vers le serveur 172.16.0.254. Les réponses ICMP écho sont bloquées par l'ACL

Tache 2 : Créer une règle d'inspection CBAC

- Démarche 1 : Créer une règle d'inspection pour examiner les trafics ICMP, Telnet et http.

```
Router # configure terminal
Router(config) # ip inspect name IN-OUT-IN icmp
Router(config) # ip inspect name IN-OUT-IN telnet
Router(config) # ip inspect name IN-OUT-IN http
```

- Démarche 2 : Appliquer la règle d'inspection du trafic sortant à l'interface Se2/0 du routeur R3

```
Router # configure terminal
Router(config) # int serial 2/0
Router(config-if) # ip inspect IN-OUT-IN out
```

2.4 Conclusion

Cette partie parlant des routeurs CISCO nous a permis d'acquérir les connaissances et compétences nécessaires pour appréhender les concepts de routage et la mise en œuvre de routeurs Cisco. L'objectif était d'étudier les bases de la configuration et de l'administration de routeurs Cisco, des connaissances sur la fonction d'un routeur et les protocoles de routage nécessaires pour la sécurisation des routeurs Cisco. Du coup, nous avons même appris comment configurer le routeur Cisco face au réseau Frame Relay qui reste le réseau dominant dans la télécommunication. Nous avons aussi vu dans ce chapitre comment mettre en place des filtrages et les firewalls. Un point focal a été mis sur l'authentification des utilisateurs et des machines, méthode essentielle à la protection des SI. Ceci nécessite bien sûr l'étude des technologies de chiffrement. Certes le renforcement des routeurs Cisco est le premier démarche à prendre dans le cadre du SSI, mail il est primordiale de sécuriser l'accès physique des routeurs car tout accès physique au routeur rend obsolète tous ces sécurisations. C'est pourquoi on utilise toujours des armoires verrouillées pour placer les routeurs. L'accès dans la salle contenant les périphériques Cisco doit être strictement réservé aux personnels de confiances.

CHAPITRE 3
LES SYSTEMES DE DETECTION D'INTRUSION ET L'APPLICATION IDS

3.1 Introduction

L'évolution technologique du réseau de télécommunication favorise l'essor du nomadisme : les employés d'une entreprise pouvant travailler en dehors des locaux de l'entreprise et des plages horaires de celle-ci. De plus, la généralisation des liaisons hautes débit et la multiplication des accès distants (extranet, intranet, télétravail, cybercafé) laissent l'information de l'entreprise accessible à chaque instant, à partir de n'importe quel endroit, grâce aux réseaux virtuels privés. Parallèlement, les problèmes de sécurité, en particulier les intrusions par Internet, vont en s'amplifiant. Il est donc nécessaire de se protéger. Actuellement, les IDS (Intrusion Détection Systems) sont parmi les éléments incontournables des dispositifs de sécurité. La deuxième partie de ce chapitre détaillera l'utilisation de l'application Intrusion Detection System qu'on vient de réaliser. En d'autre terme, nous parlerons des paramètres d'entrée et des paramètres de sortie. Les outils java tels que les classes, les méthodes, la bibliothèque principale ont été déjà détaillés dans le chapitre précédent.

3.2 Les Systèmes de Détection d'Intrusion - IDS

3.2.1 Généralités sur les IDS

La détection d'intrusion a pour objectif de déceler toute violation de la politique de sécurité en vigueur sur un système informatique. Nous avons vu dans les chapitres précédents l'élaboration d'une politique de sécurité par la sécurisation des routeurs Cisco. Cependant, l'action préventive ne suffit pas, il faut lui associer une politique de détection d'intrusion.

La détection d'intrusion consiste en un processus de découverte et d'analyse de comportement hostile dirigé contre un réseau. Actuellement deux types de systèmes de détection d'intrusion existent et ont des champs d'action complémentaires : les systèmes orientés réseau (ou Network based Intrusion Detection software, soit NIDS) et les systèmes orientés poste (Host based IDS, ou HIDS).

Les systèmes orientés poste ont pour rôle de déterminer si un ordinateur donné est en train d'être attaqué ou si celui-ci a déjà été attaqué, résultant d'une compromission de la sécurité et de l'intégrité du système.

Les systèmes orientés réseau s'occupent de surveiller le trafic sur le réseau en confrontant les paquets détectés à un ensemble de signatures ou de règles. Si les règles sont violées ou si une

signature s'applique, le NIDS enregistre l'événement comme une attaque.

Il est aussi nécessaire de savoir ce que c'est qu'un système de prévention d'intrusion (ou IPS, Intrusion Prevention System) qui est un outil des spécialistes en sécurité des systèmes d'information, similaire aux IDS, permettant de prendre des mesures afin de diminuer les impacts d'une attaque. C'est un IDS actif, il détecte un balayage automatisé, l'IPS peut bloquer les ports automatiquement. Les IPS peuvent donc parer les attaques connues et inconnues. Comme les IDS, ils ne sont pas fiables à 100 % et risquent même en cas de faux positif de bloquer du trafic légitime. Dans cet ouvrage, nous ne parlerons davantage que de l'IDS.

Un système de détection d'intrusion est constitué classiquement de trois composants :

- Le capteur qui rassemble les informations sur l'évolution de l'état du système et qui fournit une séquence d'événements qui rendent compte de cette évolution.
- L'analyseur qui détermine si un sous-ensemble des événements fournis par le capteur est caractéristique d'une activité malveillante.
- Le manager qui réunit les alertes en provenance du capteur, les met en forme et les présente à l'opérateur. Il peut aussi avoir la responsabilité de la riposte appropriée.

La partie la plus importante est donc l'analyseur puisque c'est lui qui détecte les intrusions de manière automatique. Dans la pratique, les outils actuels ne sont pas configurés directement par les instances de sécurité.

L'analyseur doit repérer si la séquence d'événements fournie par le capteur peut laisser supposer une activité malveillante.

Les trois principales approches sont l'approche comportementale *(anomaly detection),* l'approche par scénarios *(misuse detection)* et l'approche hybride. [18]

3.2.1.1 L'approche comportementale

Cette approche repère une attaque en évaluant l'écart du système surveillé par rapport à un comportement normal préalablement défini.

L'approche comportementale consiste à considérer comme hostile tout ce qui n'est pas normal, au sens où on cherchera plutôt à bien modéliser ce qu'est un comportement normal sur le réseau pour pouvoir y opposer toute déviance, que l'on considérera comme étant une attaque ("si ce n'est pas normal, alors c'est dangereux"). Ce principe est présenté sur la figure 4.01.

Cette approche comprend donc deux phases :

- Extraction d'informations sur le milieu, afin de définir la "normalité",
- Etablir les limites de la "normalité", au-delà desquelles le comportement est nécessairement anormal.

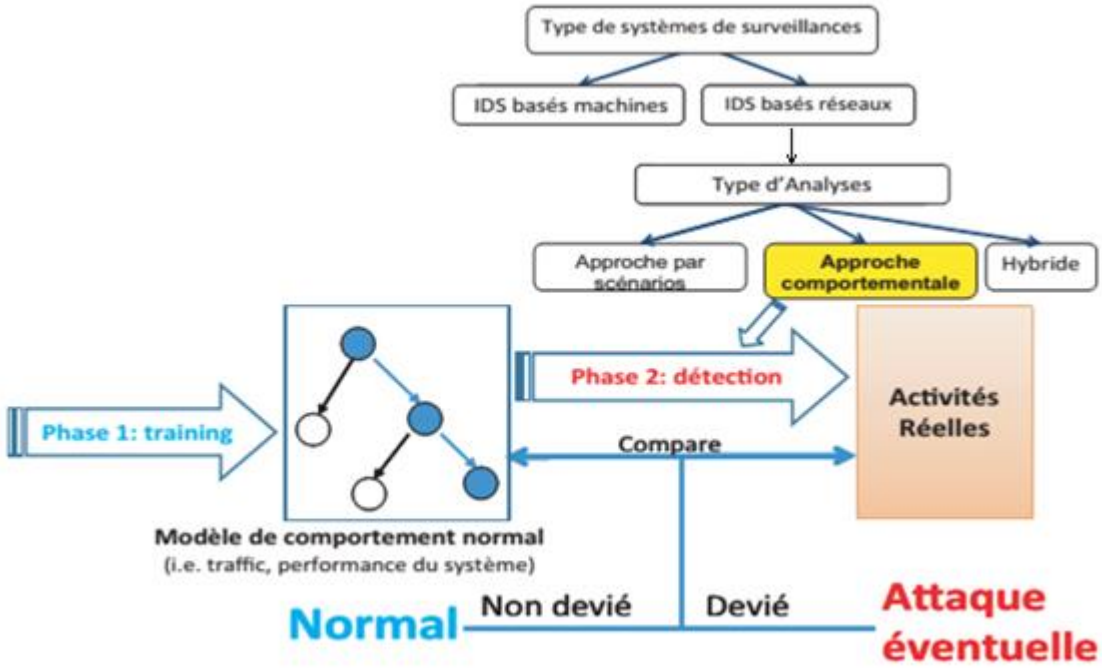


Figure 3.01 : *L'approche comportementale*

L'approche comportementale reviendra donc à repérer tout ce qui sortira du cadre de la normalité. Cette approche consiste à modéliser des comportements normaux pour détecter les comportements interdits.

Plusieurs méthodes ont été utilisées afin de construire ces comportements (profils). On trouve des méthodes statistiques, des approches qui se basent sur l'immunologie ou sur les réseaux de neurones, et les méthodes utilisant des graphes ou des réseaux Bayésiens.

La définition du comportement normal par apprentissage est complexe puisque les données apprises ont été recueillies antérieurement à l'IDS. En effet, la phase d'apprentissage requiert une base de données à la fois saine et achevé par rapport au comportement attendu des utilisateurs dans l'environnement réel. A partir de la figure 3.01, nous pouvons en déduire que la principale avantage est la possibilité de détecter de nouvelles attaques. Mais en conséquence, l'inconvénient majeur serait l'existence de fausse alertes (cf. Figure 3.03). [18]

3.2.1.2 L'approche par scénario

Cette approche utilise une base de signatures caractérisant les différentes attaques connues (ou scénarios) pour rechercher dans la séquence d'événements l'apparition d'un motif caractéristique d'une attaque.

L'approche par scenario considère comme normal tout ce qui n'est pas hostile : ici il est impératif de bien connaitre les attaques possibles et le mot d'ordre est plutôt " si ce n'est pas dangereux, alors c'est normal".

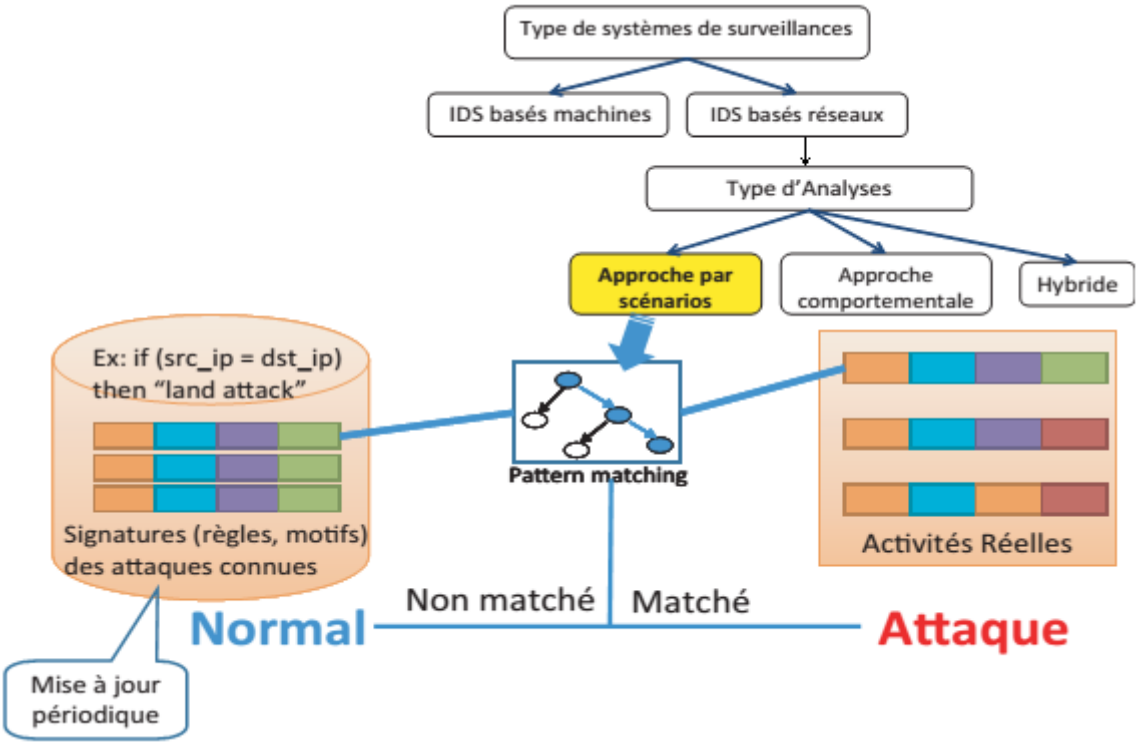


Figure 3.02 : *L'approche par scénario*

Si l'approche comportementale a l'avantage de pouvoir détecter des attaques pas forcément connues (et donc nécessairement non relevées par l'approche par mauvaise utilisation), l'approche par scénario assure cependant ce qui est relevé par le moteur d'analyse est une attaque avec une plus forte probabilité que pour l'approche précédente (l'approche par scénario amène nécessairement un grand nombre de fausses alertes). L'approche par scénario est d'autre part moins "lourde" à mettre en application, puisqu'elle ne nécessite pas de phase de renseignements sur le milieu avant d'être opérationnelle.

L'approche par scénario est actuellement la plus fréquente. Elle s'appuie sur une base de signatures d'attaque. Cependant, la difficulté vient de la définition des motifs. En effet, ceux-ci

doivent être suffisamment spécifiques pour pouvoir discriminer les différents types d'attaques, mais suffisamment générique pour pouvoir détecter les différentes variantes d'un même type d'attaque. Une signature trop générique conduira à l'augmentation du nombre de faux positifs, diminuant par la même la fiabilité.

La technique de détection par scénario nécessite en outre une maintenance active du système pour mettre à jour régulièrement la base des signatures.

En théorie, cette approche devrait produire peu de faux positifs (une connexion normale détectée comme étant une attaque) car le système utilise une connaissance a priori sur les attaques.

Les techniques de ce type restent toutefois faciles et rapides à mettre en œuvre, mais le problème de la fiabilité reste d'actualité concernant les fausses alertes.

A partir de cette figure 3.02, nous pouvons en déduire que la principale avantage est le nombre réduit de false positive. Mais en conséquent, l'inconvénient majeur serai la non détection de nouvelles attaques. [18]

3.2.1.3 L'approche hybride

L'approche hybride qui résulte de la fusion des deux approches précédentes utilise simultanément une base de signature caractérisant les attaques connues et une base de comportements normaux de la part des utilisateurs. [18]

3.2.2 Développements actuels en détection d'intrusion **[19]**

Pour implémenter ces approches, différentes méthodes sont actuellement développées et utilisées :

- Système expert: il s'agit d'un système corrélé à une base de connaissances établie par des experts humains, offrant des possibilités de résoudre des problèmes ou au moins de fournir une aide à la décision. D'une façon générale, le système utilise la base de connaissances sous forme de règles, dont l'activation provoque la détection d'une attaque ou éventuellement l'appel de nouvelles règles (structure de type "si règle n°1 alors règle n°2..."). Il est ici crucial de pouvoir assurer la maintenance de la base de connaissances, et il est à noter que la qualité de cette base dépend grandement de l'expert humain qui la définit.
- Langage de spécifications : cette méthode consiste en des déclarations d'événements en cours, complétées par un ensemble de règles, de la forme "motif -> action" (si motif est détecté, alors le système provoque action). Une fois établies en fonction de ce qui est observé (les déclarations gardent une trace sous forme d'arguments des spécificités de

l'événement), les déclarations d'événements sont confrontées aux motifs Le langage défini permet de rendre plus complexe les motifs utilisés par composition séquentielle, possibilité de tenir compte de contraintes temporelles, etc.

- Système à scénarios : le système compare ce qu'il observe à un ensemble de scénarios prédéfinis (par exemple des attaques se déroulant en plusieurs étapes). L'analyse consiste en la détection de l'étape du scénario en cours, puis à faire l'hypothèse de la prochaine étape que l'on devrait rencontrer si ce scénario est effectivement en cours d'application. Le système tente alors de détecter cette prochaine étape dans les données qu'il possède. Au fur et à mesure que les données sont analysées, le système tient à jour les probabilités d'occurrence de chaque scénario. Ici encore, la définition des scénarios est une étape cruciale.
- Analyse par automates : les attaques sont considérées comme des suites de transitions d'états du système surveillé. Les états dans le motif d'attaque correspondent aux états du système et sont associés à des tests logiques qui doivent être validés avant de passer à l'état suivant. Les états successifs sont reliés entre eux par des arcs correspondants aux conditions requises pour changer d'état.
- Analyse par graphe : il existe certaines attaques sur les réseaux tels que les worms (programmes se propageant de façon autonome de machine en machine, et utilisant chaque machine contaminée comme "base de lancement" et pour une tâche définie par l'attaquant : calcul parallèle, ou bien tout simplement désactivation de la machine) dont l'activité est facilement représentable par un graphe de structure caractéristique (structure en arbre ou en éventail). L'idée de l'analyse par graphe est donc de construire un graphe représentant les activités et les machines sur le réseau à protéger : si le graphe présente une structure similaire à celle d'une attaque, alors le réseau est probablement soumis à une activité hostile. Le type d'attaques que cette méthode permet de détecter est cependant limité (sweeps, worms).
- Intelligence artificielle : l'utilisation d'outils capables de s'auto-configurer tels que les réseaux neuronaux permet de faciliter le travail des opérateurs humains, notamment dans le domaine de la classification. Cette méthode est donc particulièrement indiquée si on a choisi l'approche comportementale. De plus ce système est assez flexible par rapport aux modèles utilisant des signatures, ce qui permet de détecter de légères variations dans les attaques.

3.2.3 Les problèmes de la fiabilité des IDS

L'analyseur doit détecter de manière automatique les intrusions. Dans la pratique, les outils actuels ne sont pas configurés directement par les instances de sécurité. Ainsi, s'ils détectent certaines intrusions, ils détectent aussi des tentatives d'intrusions infructueuses, ce qui n'est pas souhaitable.

En outre, la relative naïveté des algorithmes de détection conduit à un nombre élevé d'alertes, dont une proportion significative est en fait constituée de fausses alertes (faux positifs). Enfin, certaines intrusions peuvent ne pas être détectées (faux négatifs). [18]

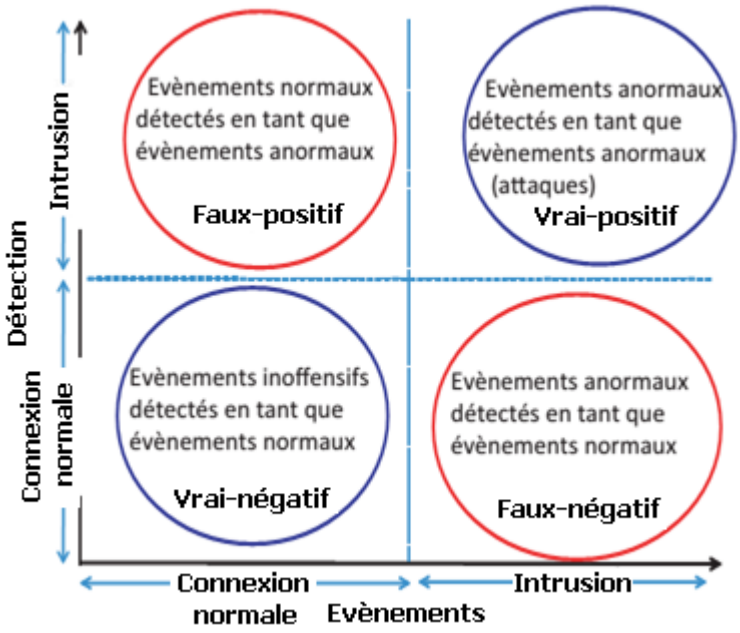


Figure 3.03 : *Les problèmes de la fiabilité des IDS*

- Le faux-positif signifie que l'événement est détecté en tant qu'intrusion alors que c'est une connexion normale
- Le faux-négatif signifie que l'événement est détecté en tant que connexion normale alors que c'est une intrusion
- Le vrai-positif signifie qu'une intrusion est vraiment détectée en tant qu'intrusion
- Le vrai-négatif signifie qu'une connexion normale est vraiment détectée en tant que connexion normale

Le terme naïf sera explicité davantage dans le paragraphe 3.2.4

3.2.4 Le réseau Bayésiens

Les problèmes de classification constituent une famille de problèmes auxquels il est possible d'appliquer des méthodes d'apprentissage supervisé, c'est-à-dire où l'on dispose d'une base d'exemples correctement identifiés. Le but consiste, à partir de ces exemples, de construire un modèle capable de prédire avec un bon degré de confiance à quelle classe (connexion normale, intrusion) appartient un individu à partir de la seule connaissance de certaines de ses caractéristiques (la durée de la connexion, le type du protocole, le service, le flag, le nombre de faux fragment, le nombre de tentatives d'ouverture échouées…)

Nous allons étudier ici sur la base d'un exemple concret un type de modèle de classification en particulier : le classifieur Bayésien naïf.

Sa caractéristique principale est qu'il émet une hypothèse forte a priori inadaptée aux cas pratiques : l'indépendance des caractéristiques étudiées.

En termes simples, un classifieur bayésien naïf suppose que l'existence d'une caractéristique déterminant l'appartenance à une classe est indépendante de l'existence d'autres caractéristiques.

Prenons un exemple dans la vie courante pour mieux expliquer, un animal sera considéré avec un bon degré de confiance comme un canard s'il a des ailes, un bec et qu'il cancane comme un canard, en ignorant totalement la possibilité que ces caractéristiques puissent être corrélées.

D'une part, le modèle naïf se comporte d'autant mieux que le nombre de classes parmi lesquelles choisir est faible. Ceci est dû en bonne partie au fait que le degré de certitude avec laquelle ils prennent leur décision peut se permettre d'être médiocre tout en restant suffisant, par un effet de seuil.

D'autre part, plus la répartition de la covariance des caractéristiques est distribuée équitablement entre les différentes classes, plus leurs contributions respectives à une mauvaise classification auront tendance à s'annuler dans l'estimation finale de la probabilité d'appartenance à une classe donnée : en d'autres termes, ils sont relativement peu sensibles au bruit dès lors que la contribution de quelques caractéristiques à une classe « pèse » suffisamment lourd.

Toutefois, les classifieurs bayésiens naïfs conservent l'avantage avec des échantillons de taille réduite sur lesquels ils convergent plus rapidement pour estimer les paramètres nécessaires à la classification. [20]

3.2.4.1 Le modèle bayésien naïf

Le modèle probabiliste pour un tel classifieur est le modèle conditionnel *P(X | x1,…,xn)*, où *X* est la variable « de classe » (celle qui indique si un individu appartient à une classe donnée) conditionnée par plusieurs variables caractéristiques x_i (la durée de la connexion, le type du protocole, le service, le flag, le nombre de faux fragment, le nombre de tentatives d'ouverture échouées…). [20]

La notation suivante énonce le théorème de Bayes:

$$P(X/x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n/X)P(X)}{P(x_1, \dots, x_n)} \quad (3.01)$$

Ici, le dénominateur correspond à la répartition des caractéristiques au sein de la population des individus, il est donc indépendant de la variable de classe *X* elle-même.

Le numérateur quant à lui peut s'écrire, par application de la formule de Bayes (c'est-à-dire la définition de la probabilité conditionnelle, soit *P(X,Y) = P(X|Y)P(Y))* :

$$P(X/x_1, \dots, x_n) = P(X, x_1, \dots, x_n) = P(X)P(x_1/X)P(x_2, \dots, x_n/X, x_1) \quad (3.02)$$

On poursuit alors en développant le dernier membre :

$$P(x_2, \dots, x_n/X, x_1) = P(x_2/X, x_1)P(x_3, \dots, x_n/X, x_1, x_2) \quad (3.03)$$

On itère ensuite de la sorte jusqu'à avoir développé les *n* caractéristiques. C'est ici qu'intervient l'hypothèse d'indépendance des caractéristiques (ou hypothèse naïve) : si les x_i sont indépendants deux à deux, alors

$$P(x_i/X, x_j) = P(x_i / X) \quad (3.04)$$

Ce qui fait que les facteurs conditionnés par X, d'une part, et des x_i d'autre part, peuvent être récrits sous la forme de probabilités conditionnées par X seulement, de sorte que l'expression se réduit en fin de compte à la forme suivante :

$$P(X/x_1, \dots, x_n) = \frac{P(X)\Pi_{i=1}^{n} P(x_i / X)}{P(x_1, \dots, x_n)} \quad (3.05)$$

Plus clairement :

- le membre de gauche représente la probabilité qu'un individu *X* appartienne à une classe donnée (connexion normale ou intrusion) sachant qu'il possède les caractéristiques *x1,…,xn*
- le membre de droite représente la probabilité de distribution de la classe dans la population, pondérée par l'occurrence de chaque caractéristique pour cette classe et une constante (le dénominateur)

A titre de remarque, ce dénominateur est appelé évidence (du faux-ami anglais evidence), que l'on peut estimer à partir d'un ensemble d'entraînement constitué d'un échantillon d'individus aux classes connues et supposé représentatif. En pratique cependant, il ne dépend que des caractéristiques pour lesquels on cherche à déterminer la loi de probabilité d'appartenance à une classe : dans le calcul du maximum de l'expression, il agit donc comme une constante comprise entre 0 et 1 que l'on peut donc tout simplement choisir d'omettre de calculer.

3.2.4.2 Estimation des paramètres du modèle

Les paramètres (distribution des classes et des caractéristiques) peuvent être estimés par les fréquences des classes par rapport aux caractéristiques sur la base d'exemples d'individus dont les classes sont connues (notre ensemble d'entraînement). Il est nécessaire de choisir a priori une distribution pour les classes et pour les caractéristiques [20]. Elles dépendront en pratique du type de valeurs rencontrées :

- pour des caractéristiques binaires, on choisira une loi de Bernoulli
- pour des caractéristiques à valeurs discrètes on choisira une loi multinomiale
- pour des caractéristiques à valeurs continues on choisira une loi normale

L'approximation fréquentielle reviendra à calculer les valeurs suivantes :

$$P(X = c_i) = \frac{nombre\ d'individus\ de\ classe\ c_i}{nombre\ total\ d'individus\ dans\ l'echantillon} \tag{3.06}$$

$$P(x_i/X = c_i) = \frac{nombre\ d'individus\ de\ classe\ c_i\ ayant\ la\ carasteristique\ x_i}{nombre\ d'individus\ de\ la\ classe\ c_i\ dans\ l'échantillon} \tag{3.07}$$

3.2.4.3 Construction d'un classifieur à partir d'un échantillon

Nous savons désormais calculer les paramètres de notre modèle. Pour fabriquer un classifieur à proprement parler, il reste à trouver une règle de décision qui puisse les exploiter. La règle la plus simple consiste à évaluer la probabilité de chaque classe *c* compte tenu des caractéristiques de l'individu et de choisir celle qui maximise l'expression du numérateur de la formule 3.05, soit

$$P(X = c)\prod_{i=i}^{n} P(x_i \,/\, X) \tag{3.08}$$

Cette règle de décision est appelée règle du maximum a posteriori. [20]

3.2.4.4 Application du réseau Bayésien pour l'IDS

Dans cette ouvrage, nous manipulerons un échantillon de base de données donc nous opterons pour l'approche Bayésien ou plus précisément l'algorithme K2. L'algorithme d'apprentissage K2 montre une haute performance dans beaucoup de travaux de recherche. Le principe de l'algorithme K2 est de définir la base de données des variables : x1,… xn, et de construire un Graphe Orienté. Les variables constituent un réseau de nœud. Les arcs représentent les relations causes à effet entre les variables.

L'algorithme K2 utilisé dans l'apprentissage a besoin de :

- un ordre entre les variables
- un nombre de parent des nœuds (u)

L'algorithme K2 procède en commençant par un nœud unique (le premier variable dans l'ordre définie qui est dans notre cas le *protocol_type*), puis ajoute de façon incrémental une connexion avec d'autres nœuds qui peut augmenter la probabilité résultante de la structure, ceci est calculé en utilisant la fonction g. [21]

$$g(x_i, pa_i(x_i)) = \prod_{j=1}^{qi} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \coprod_{k=1}^{ri} N_{ijk}! \tag{3.09}$$

Pour chaque variable x_i :

- r_i est le nombre possible d'instanciations
- N est le nombre de cas dans la base de données

- *qi* est le nombre possible d'instanciations pour pa_i
- N_{ijk} est le nombre de cas pour qui *xi* prend les valeurs x_{ik} (avec pa_i instancié à la j-ème instanciation du pa_i dans la base de données).
- N_{ij} est la somme du N_{ijk} pour toutes valeurs de *k.*

Concernant le système de l'algorithme K2 :

- Les variables d'entrées : un ensemble de variable *x1,…, xn*, un ordre donné parmi eux, une limite seuil *u* du nombre de parents pour un nœud, une base de données sur *x1,…, xn*.
- Les variables de sortie : un graphe orienté acyclique régie par la portion de program dans la suivante

/ ********** Debut du code **********/

For i := 1 to n do

$pa_i(x_i) = \emptyset$; *Flag :=true ;*

$P_{old} := g(x_b, pa_i(x_i))$;

While Flag and $|pa_i(xi)| < u$ *do*

Soit z un nœud dans l'ensemble du prédecesseurs de x_i *qui n'appartient pas à* $pa_i(x_i)$ *qui maximise* $g(x_i pa_i(x_i) \cup \{z\})$;

$P_{new} := g(x_i pa_i(x_i) \cup \{z\})$;

If $P_{new} > P_{old}$ *then*

$P_{new} = P_{old;}$

$pa_i(x_i) = pa_i(x_i) \cup \{z\}$

Else Flag := false;

/ ********** Fin du code **********/

L'ordre des variables du réseau est comme suit:

Protocole_type, service, land, wrong_fragment, num_failed_logins, logged_in, root_shell, is_guest_login, attack_type.

La limite supérieure des nœuds parents u = 4 ce qui est suffisant pour représenter ces 9 variables. Le réseau Bayésiens résultante est représenté par la figure 3.04.

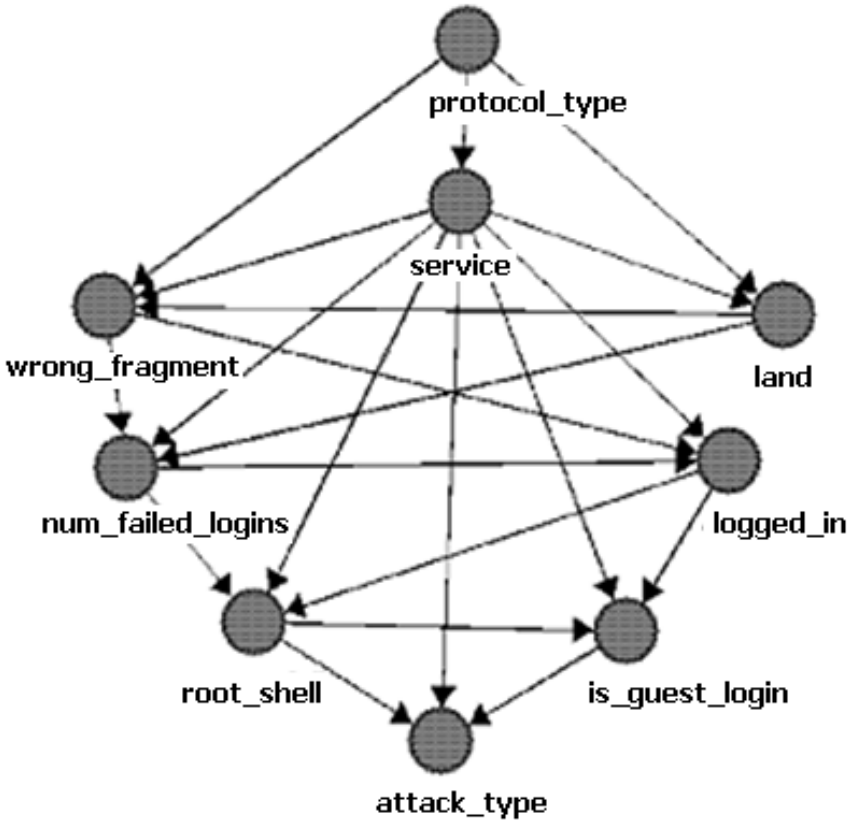


Figure 3.04 : *Lé réseau Bayésien pour l'application IDS*

3.2.5 Présentation des données KDD – Cup 1999

Les données utilisées pour nos expérimentations sont des données réelles issues de la base KDD – Cup 1999. Elles ont été préparées et contrôlées par le laboratoire MIT (Michigan Institute of Technology) Lincoln pour le programme d'évaluation de détection d'intrusion DARPA (Defense Advanced Research Projects Agency) en 1998. Ces données ont aussi été utilisées pour le concours de détection d'intrusion de KDD 1999. Chaque connexion est étiquetée en tant que connexion normale ou attaque.

Ces données sont décrites au moyen de différents attributs explicatifs. Pour une meilleure compréhension, ceux-ci ont été classifiées en cinq types d'attributs.

Les attributs d'une même machine décrivent seulement les connexions faites durant les deux dernières secondes et ayant le même destinataire que la connexion courante.
Les attributs de même service décrivent seulement les connexions faites durant les deux dernières secondes et ayant le même service que la connexion courante.
Les attributs d'une même machine et de même service définissent les critères du trafic de connexion faits en une fenêtre de deux secondes. On trouve aussi des attributs de connexion TCP individuelles.
Enfin, il existe des attributs qui indiquent un comportement anormal dans les données, ainsi que le nombre de tentative d'ouverture échouées. Il s'agit des attributs de contenu. [18]
Les données KDD – Cup 1999 sont construites à partir des données collectées par le programme d'évaluation de détection d'intrusion de DARPA en 1998. Ces données qui correspondent à environ quatre gigaoctet de données binaires TCPdump compressées, contiennent sept semaines de trafic de réseau. Ceci à été transformé en environ cinq millions de connexions. Les données d'apprentissage KDD – Cup 99 possèdent 4 900 000 connexions étiquetées normale ou attaque. Chaque connexion contient 41 variables descriptives, mais pour la réalisation de ce projet de détection d'intrusion, les neufs premiers variables sont suffisants.

Nom d'attributs	Description	Type
Attributs du trafic de connexion pendant une fenêtre de deux secondes		
Count	nombre de connexion à la même machine que la connexion courante durant les deux dernières secondes	Continu
Attributs des connexions de même machine		
Serror_rate	nombre de connexion qui ont des erreurs de SYN	Continu
Rerror_rate	nombre de connexion qui ont des erreurs de REJ	Continu
Same_srv_rate	nombre de connexion au même service	Continu
diff_srv_rate	nombre de connexions au service différents	Continu
srv_count	nombre de connexions au même service que le raccordement courant dans les dernières deux secondes	Continu
Attributs des connexions de même service		
srv_serror_rate	nombre de connexions qui ont des erreurs de SYN	Continu
srv_rerror_rate	nombre de connexions qui ont des erreurs de REJ	Continu
srv_diff_host_rate	nombre de connexions aux différents hosts machines	Continu
Attributs des connexions TCP individuelles		
durée	longueur (nombre de secondes) de la connexion	Continu
protocol_type	type de protocole, par exemple TCP, UDP	Discret
service	service de réseau pour la destination, par exemple, HTTP, Telnet, …	Discret
src_bytes	nombre de bytes de données de la source à la destination	Continu
flag	statut normal ou erreur de la connexion	Discret
land	1 si la connexion est from/to même host/port; 0 autrement	Discret
wrong_fragment	nombre de "faux" fragments	Continu
urgent	nombre de paquets urgents	Continu
Attributs de contenu		
hot	Nombre de hot indicateurs	Continu
num_failed_logins	nombre de tentatives d'ouvertures échouées	Continu
logged_in	1 si entré avec succès; 0 autrement	Discret
num_compromised	le nombre de conditions compromises	Continu
root_shell	1 si root shell est obtenue; 0 autrement	Discret
su_attempted	1 si la commande su root racine a été essayée; 0 autrement	Discret
num_root	nombre d'accès root	Continu
nom_file_creations	nombre d'opérations de création de dossier	Continu
num_shells	nombre de shell sollicités	Continu
num_access_files	nombre d'opérations sur des dossiers de contrôle d'accès	Continu
num_outbound_cmds	nombre de commandes venant d'une session FTP	Continu
is_hot_login	1 si l'ouverture appartient à la hot liste; 0 autrement	Discret
is_guest_login	1 si l'ouverture est un guest login; 0 autrement	Discret

Tableau 3.01: *Les différents attributs des données KDD – Kup 1999*

3.2.6 Programmation java avec WEKA

Weka (acronyme pour Waikato Environment for Knowledge Analysis, en français : « Environnement Waikato pour l'analyse de connaissances ») est une suite populaire de logiciels et de bibliothèque d'apprentissage automatique développée à l'université de Waikato, Nouvelle-Zélande. L'espace de travail Weka contient une collection d'outils de visualisation et d'algorithmes pour l'analyse des données et la modélisation prédictive.

Utiliser Weka nous permet par exemple de définir un protocole ou un programme JAVA générique d'apprentissage où il nous suffira de changer une ligne dans le programme pour pouvoir utiliser un classifieur à la place d'un autre. [22]

Pour pouvoir utiliser les algorithmes dans nos programmes, il nous faut :

- Pouvoir définir ou charger à partir d'un fichier un ensemble d'exemples d'apprentissage.
- Connaître les quelques méthodes qui permettent de définir, initialiser et utiliser un classifieur.
- Connaître la méthode qui permet d'utiliser un classifieur pour trouver la classe d'un nouvel exemple.
- Afficher, lire, interpréter les classifications obtenues.

Le cadre dans lequel nous travaillons, et les algorithmes que nous étudions et utilisons se basent sur le schéma de fonctionnement suivant :

- On dispose d'un ensemble d'exemples (instances), chaque exemple étant défini par :
    - Sa description : c'est un ensemble de valeurs définissant cet exemple (la durée de la connexion, le type du protocole, le service, le flag, le nombre de faux fragment, le nombre de tentatives d'ouverture échouées…)
    - La classe qu'on lui a associée (connexion normale, intrusion)
- On fournit cet ensemble d'exemples à un programme qui va générer un classifieur. Un classifieur est un programme qui, quand on lui fournira un exemple, essaiera de deviner sa classe. Dit autrement, le programme prédit la classe d'un exemple à partir de sa description. Dit encore autrement, le programme cherche la relation qui lie la description à la classe.

3.2.6.1 Définir un ensemble d'apprentissage

Un ensemble d'apprentissage est défini dans Weka par la classe Instances (au pluriel). [22]

```
import weka.core.Instances;
```

Un objet de cette classe contient :

- Une description de la structure des exemples (liste des attributs, type de chaque attribut, indice de l'attribut qui sert de classe)
- La liste des exemples

3.2.6.2 Charger un ensemble d'exemples

On peut charger un ensemble d'exemples à partir d'un fichier d'extension ".arff". [22]

```
import java.io.File;
import weka.core.converters.ArffLoader;
import weka.core.Instances;

public class DataSource {
        public static Instances read(String filePath) throws Exception {
                ArffLoader arffLdr = new ArffLoader();
                arffLdr.setSource(new File(filePath));
                Instances dataSet = arffLdr.getDataSet();
                return dataSet;
                                                        }
                        }
```

3.2.6.3 Construire et utiliser un classifieur

Les arbres de décision correspondent à la classe weka.classifiers.trees.J48, et sont une sous-classe de weka.classifiers.Classifier.
Une fois que le classifieur est construit, on peut l'utiliser pour classer de nouveaux exemples.
La méthode *public double classifyInstance(Instance instance)* de la classe Classifier retourne un réel qui correspond à la classe attribuée par le classifieur à l'exemple passé en paramètre.
Dans le cas où la classe est discrète, comme par exemple quand le classifieur est un arbre de décision, il faut encore reconvertir ce réel pour retrouver la valeur nominale de l'attribut classe. Les concepteurs de Weka ont préféré coder en interne la valeur de chaque attribut par un réel, mais dans le cas des attributs nominaux, ce réel est en fait un entier, qui est l'indice de la valeur de l'attribut. [22]

```
Instances     train = DataSource.read(ptrain);              // donnees pour l'apprentissage
Instances     test = DataSource.read(ptest);                // donnees à tester
train.setClassIndex(train.numAttributes() - 1);
test.setClassIndex(test.numAttributes() - 1);

if (!train.equalHeaders(test))
throw new IllegalArgumentException("Les données ne sont pas compatibles");

J48 j48 = new J48();                                        // instancier le classifieur
j48.setUnpruned(true);

FilteredClassifier fc = new FilteredClassifier();
fc.setFilter(rm);
fc.setClassifier(j48);
fc.buildClassifier(train);// construire le classifieur avec donnees train
int nbInstance=0, nbIntrusion=0;

for (int i = 0; i < test.numInstances(); i++) {
double pred = fc.classifyInstance(test.instance(i));
jta.append("\nConnexion " + (i+1) +":");
jta.append(" Evènement : "+ test.classAttribute().value((int) test.instance(i).classValue()));
jta.append(", Prédiction : "+ test.classAttribute().value((int) pred));
if                                                          ((test.classAttribute().value((int)
pred).intern()=="intrusion")|test.classAttribute().value((int)
pred).intern()=="anomaly"){nbIntrusion++;};
nbInstance++;
jta.append("\n_____________________________________________\n");
} // fin du boucle for
jta.append("\n\nNombre d'instance  : "+nbInstance+"\n");
```

3.2.7 La base de données ARFF

Le format d'entrée par défaut de Weka est l'ARFF (Attribute Relation File Format). Tous ce qu'il faut savoir sur ce type de fichier sont : [22]

- Les commentaires sont précédés de %
- La définition du nom de l'ensemble de données avec @relation (Le nom doit être aussi compréhensible que possible)
- La définition des caractéristiques avec @attribute
    - Attributs nominaux suivis des valeurs entre accolades
      @attribute 'protocol_type' {'tcp','udp', 'icmp'}
    - I Attributs numériques avec real
      @attribute 'wrong_fragment' real
    - I Attributs chaines avec string, les valeurs doivent être entre doubles guillemets
      @attribute unTexte string
    - Attributs dates avec date (yyyy-MM-dd-THH :mm :ss)
      @attribute uneDate date
- @data signale le début des instances

3.3 L'exploitation de l'application IDS

Le principal but de notre réalisation est de prédire si les connexions testées sont des connexions normales ou des intrusions.

Après avoir lancé l'application, l'écran d'accueil (cf. Annexe Figure A1.01) s'affiche pendant 10 secondes, le temps pour que l'utilisateur puisse lire les informations concernant l'application.

Puis la fenêtre principale s'affiche.

La fenêtre est subdivisée en 2 entités :

- La partie du gauche concerne l'ensemble d'apprentissage et le classifieur naïf bayésien. Un JTextArea affiche le contenu de l'ensemble d'apprentissage (les 41 attributs de chaque instance). C'est à partir de cette base de données qu'est établi le classifieur naïf bayésien, le JTextArea en bas à gauche.

- La partie de la droite de l'écran concerne les connexions à tester. Le programme prédit la classe des exemples à partir de leurs descriptions. Dit encore autrement, le programme cherche la relation qui lie la description à la classe.

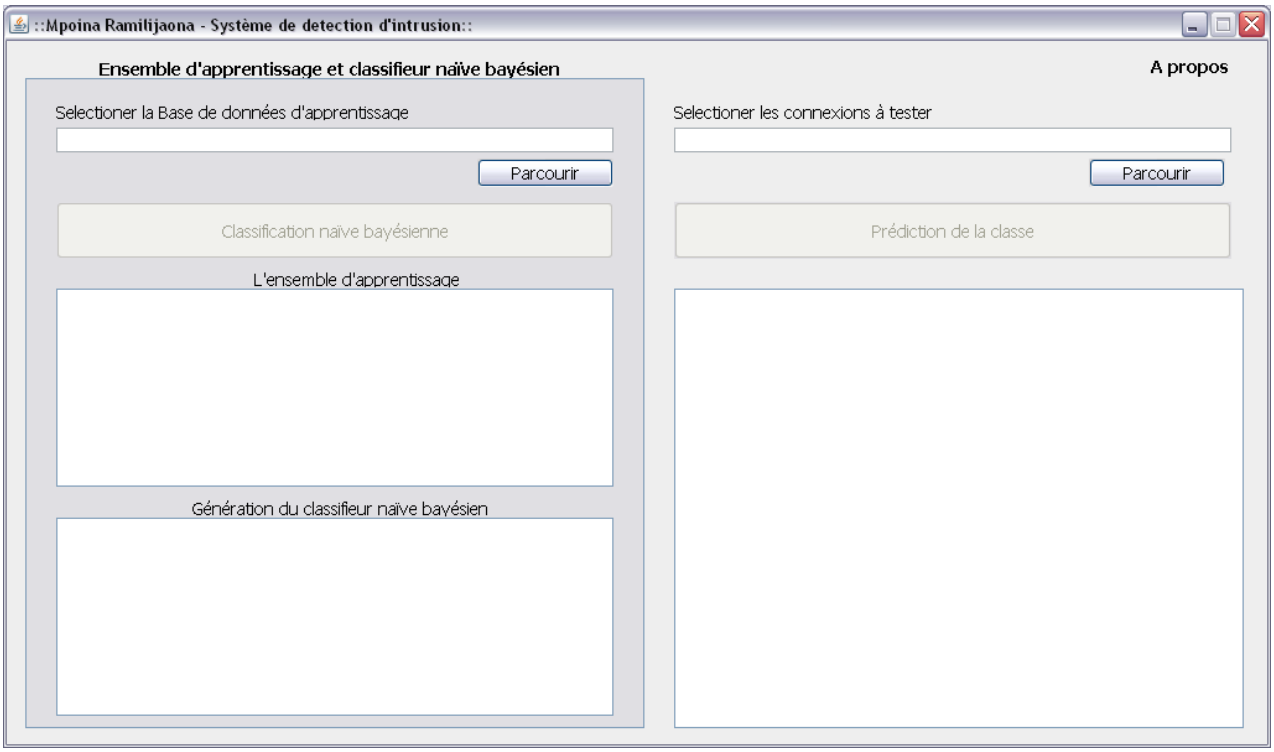


Figure 3.05 : *La fenêtre principale initialement vide*

La zone de texte invitant à « Sélectionner la base de données d'apprentissage » affiche l'emplacement du fichier ARFF du KDD avec lequel le système va générer le classifieur naïf bayésien. Le bouton « Parcourir » permet de sélectionner ce fichier ARFF.

Rappelons que les fichiers ARFF contiennent les données d'apprentissages préparées et contrôlées par le laboratoire MIT pour le programme d'évaluation d'intrusion DARPA.

En tout, il y a moins de 5 millions de signatures de connexion étiquettes comme intrusion ou saine.

Les classifieurs bayésiens naïfs conservent l'avantage avec des échantillons de taille réduite pour estimer les paramètres nécessaires à la classification. Seule une partie de ces 5 millions d'instances est donc suffisant pour entrainer le système. Nous utiliserons par exemple 20% de ces instances.

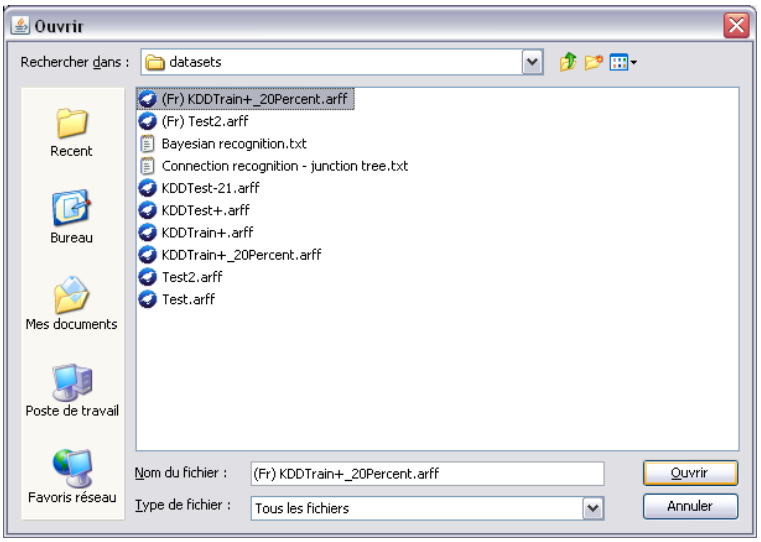


Figure 3.06 : *La sélection du fichier ARFF*

Apres avoir sélectionné un fichier ARFF, l'emplacement de ce dernier est affiché dans le zone de texte « Sélectionner la base de données d'apprentissage ». La deuxième étape consiste à générer le classifieur naïf bayésien.

Une fois avoir cliqué le bouton « Classification naïve bayésienne », le système génère les classifieurs « connexion normale » et « intrusion » en suivant les processus énoncé dans ce chapitre.

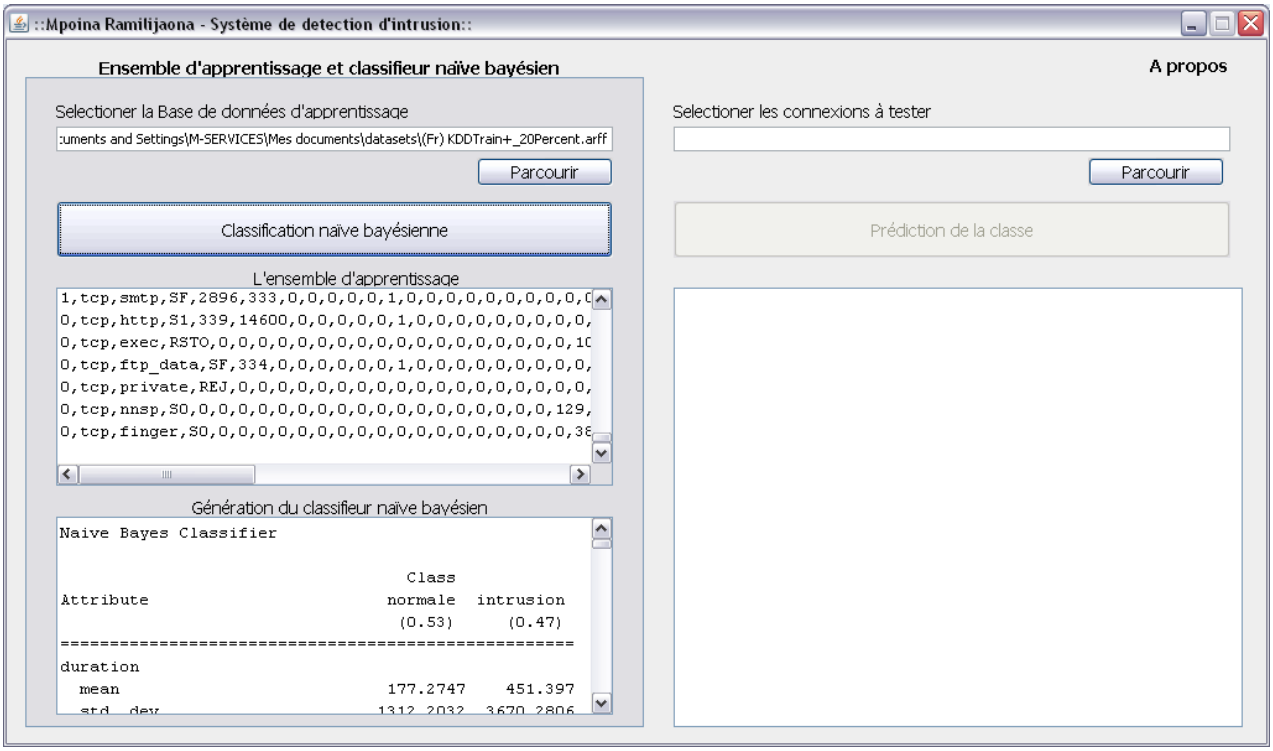


Figure 3.07 : *Génération du classifieur naïf bayésien*

Le JTextArea « L'ensemble d'apprentissage » affiche les instances (connexions) étiquetés contenues dans le KDD. Rappelons que seul les neuf-premier de ces attributs est suffisant pour dresser le réseau bayésien (cf. Figure 4.04).

Le JTextArea « Génération du classifieur naïf bayésien » affiche les classifieurs crées (cf. 3.2.4.3 et 3.2.6.3)

La zone de texte invitant à « Sélectionner les connexions à tester» affiche l'emplacement du fichier ARFF contenant les signatures des connexions à tester. Le bouton « Parcourir » permet de sélectionner ce fichier ARFF.

Le bouton « Prédiction de la classe » effectue le processus de décision du système suivant la règle de décision du maximum à postériori (cf. 3.2.7). C'est la raison d'être de ce système de détection d'intrusion.

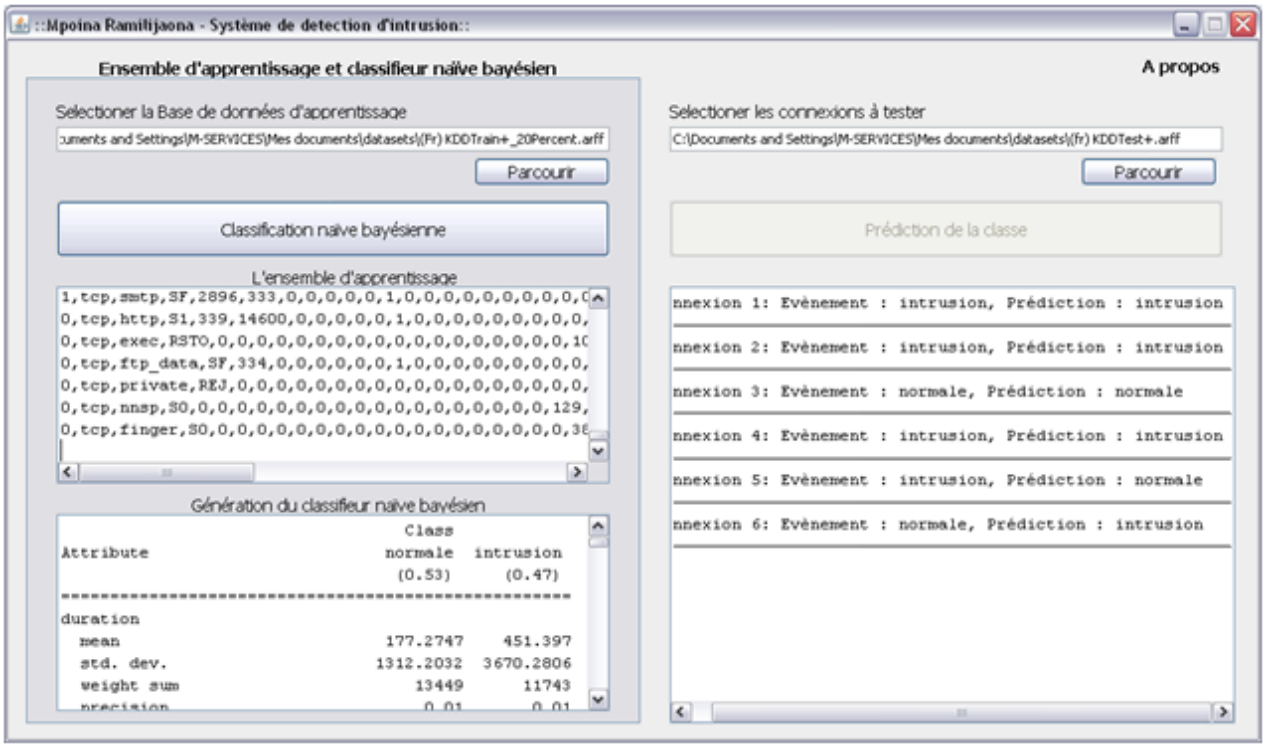


Figure 3.08 : *Résultat d'une détection d'intrusion sur un échantillon de connexion*

L'analyseur du système détecte de manière automatique les intrusions. Dans notre cas, le système à trouvé des fausses alertes sur l'échantillon (cf. paragraphe 3.2.3)

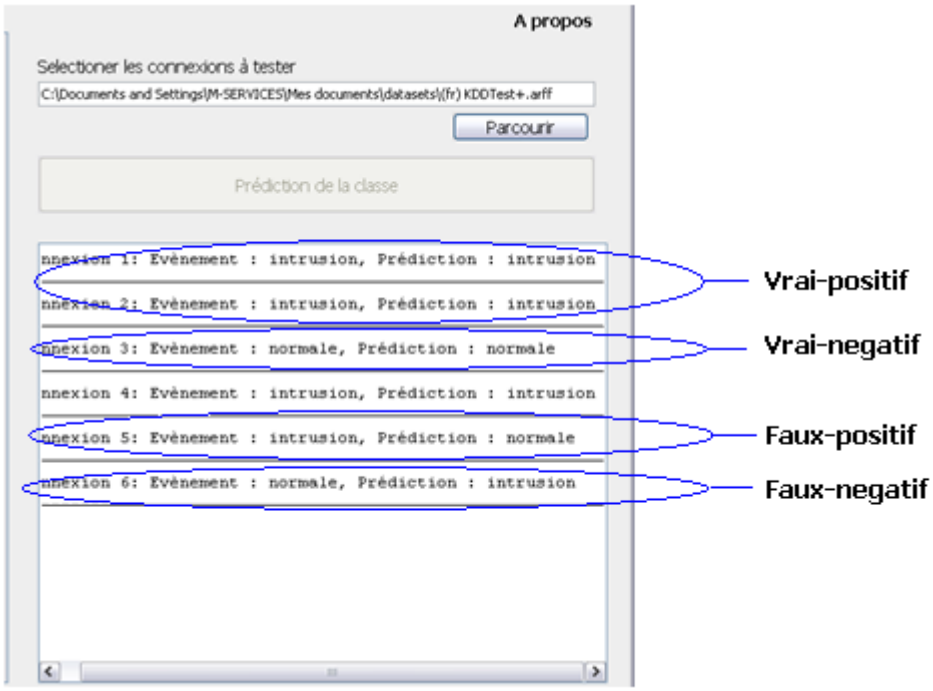


Figure 3.09 : *Résultat de l'analyse sur un échantillon de connexion*

3.4 Conclusion

Les Information Détection Systems sont parmi les éléments incontournables des dispositifs de sécurité. Avec la base de données KDD 1999 et grâce aux réseaux Bayésiens, le système construit un modèle capable de prédire avec un bon degré de confiance la classe normale/intrusion d'une connexion. L'implémentation sous java de l'IDS est possible grâce à la bibliothèque d'apprentissage automatique Weka. Dans ce chapitre, avons fait la présentation détaillé de l'application Intrusion Detection System ainsi que l'analyse d'un échantillon de signature de quelques connexion. L'objectif était d'analyser de façon automatique la classe d'appartenance d'une connexion (normale ou intrusion) en consultant sa signature. Cette prédiction est basée sur le réseau bayésien.

CONCLUSION GENERALE

Sécuriser le système d'information par renforcement des routeurs Cisco et par réalisation d'une application IDS était l'objectif principal de notre travail. En effet, l'accomplissement de cette tâche a nécessité l'approfondissement de la connaissance en réseau IP & sécurité, en curriculum Cisco surtout la CCNA Security, en cyber-attaque visant les systèmes d'informations, en Java particulièrement le data mining, et principalement en mathématique statistique.

La généralité sur les réseaux informatiques domine le début de cet ouvrage : les différentes modèles qu'il faut savoir et les protocoles essentiels comme le TCP/IP, X25 et Frame Relay. Comme il s'agit de sécurisation de Système d'information, il est important d'étaler les différentes attaques possibles ainsi que leurs comportements pour pouvoir les réprimer avec le système CISCO et le Système de Détection d'Intrusion.

Le durcissement des routeurs Cisco débute la première étape de la sécurisation du Système d'Information, à commencer par l'architecture et les commandes de base de ces routeurs, pour ensuite entamer à leur sécurisation. En effet, toutes les protections qu'on puisse implémenter sur un périphérique Cisco comme les Access List et les Pare-feu ont y été évoqués.

La dernière partie de l'ouvrage soulève les bases nécessaires à la réalisation d'une application de détection d'intrusion, dans lesquelles le réseau de Bayes et les bases de données d'apprentissage du DARPA y figure. L'application Java du Network Intrusion Detection System a été présentée. Pour ce qui est de l'avenir de ce logiciel, de nombreuses améliorations peuvent être apportés comme par exemple l'adoption du réseau de neurone au lieu du réseau de Bayes, l'interface graphique, et ainsi de suite. L'application IDS qui d'être développé se limite au cerveau proprement dit des IDS mais n'effectue pas ni la translation des connexions réseau en leur signatures (en amont de ce noyau IDS), ni la prise de décision ou les alarmes en cas d'intrusion (en aval de ce noyau IDS).

Sécuriser le Systèmes d'Informations revient à déterminer à l'avance les comportements de toutes les attaques possibles, après cela interviennent les mesures de sécurité au niveau de l'interface entre le Système d'Information lui-même et le réseau internet.

Certes les mesures préventives listées dans cet ouvrage agissent en permanence à travers les flux qui transitent dans le système, mais les attaques évoluent de jour en jour ce qui nécessite la mise à jour de la connaissance sur ces attaques, voir même l'apparition de nouveaux menaces.

Les Systèmes d'Informations sont des patrimoines propres aux entités de toute taille allant d'un seule poste d'ordinateur à un réseau d'une firme internationale. L'enjeu sur leurs sécurités demeure primordial nécessitant en permanence la sécurisation des Systèmes d'informations basique comme le renforcement des routeurs CISCO et la mise en place d'un Système de Détection d'Intrusion.

ANNEXES

ANNEXE 1 : Les autres fenêtres de l'application IDS

Figure A1.01 : *L'écran d'accueil*

L'écran d'accueil s'affiche pendant 10 secondes, le temps pour que l'utilisateur puisse lire les informations pertinentes.

Et pour finir, la fenêtre sur les informations générales :

Figure A1.02 : *A propos du logiciel*

ANNEXE 2 : Téléchargement des fichiers d'entrainement KDD

Nombreux sont les pages qui hébergent les fichiers d'entrainement KDD en format ARFF, en voici un exemple :

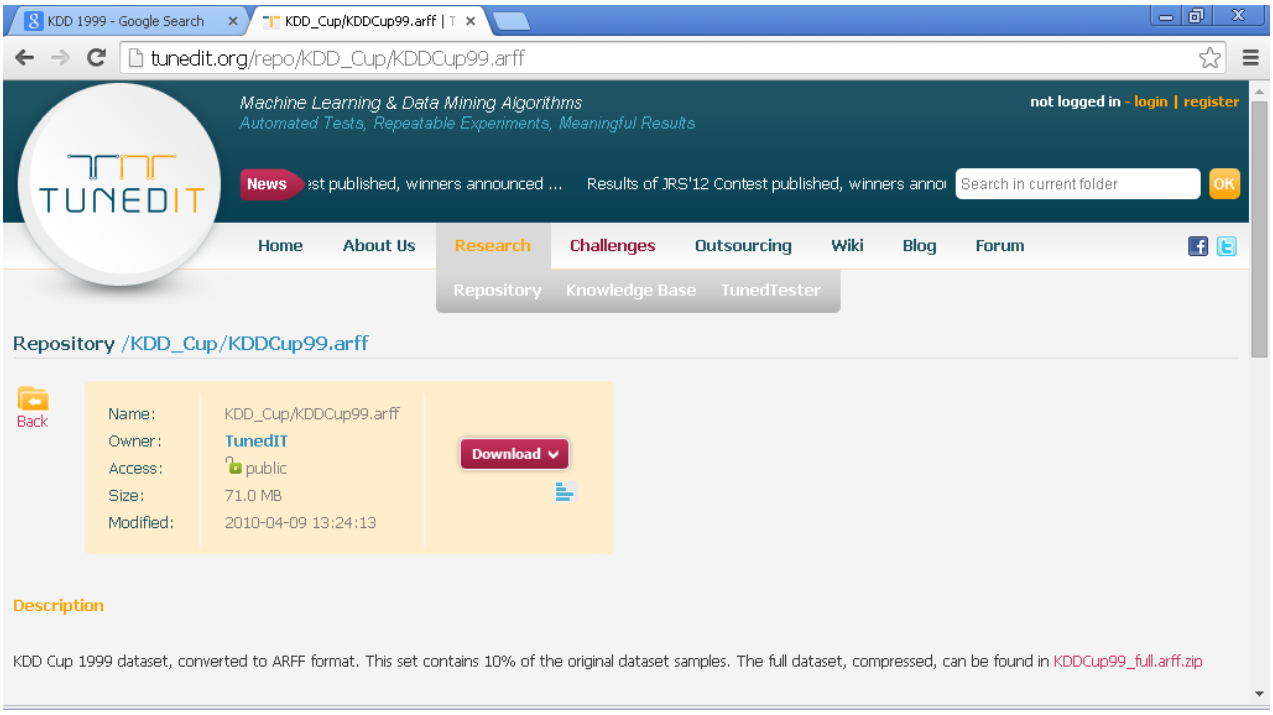


Figure A2.01 : *Exemple d'un site web hébergeant le fichier KDD*

BIBLIOGRAPHIE

[1] I. Rudenko, *"Cisco Routers for IP Routing*", The Coriolis Group LLC, 2000

[2] Huin, *"Rapport de DEA: pattern matching et detection d'intrusion »,* taz.newffr.com, 2014

[3] L.E. Randriarijaona, *"Réseau IP & sécurité"*, Cours I4 Dépt TCO, AU 2011-2012

[4] P Ayers, C Matthews, B Yates, "*Regional Internet Registry*", wikipedia.com, 2014

[5] N.R. Razafindrakoto, R.R. Andriandrova, *" L'association NIC-MG"*, slideplayer.fr, 2005

[6] N. VanHaute, *"Le protocole IPv6"*, commentcamarche.net, 2014

[7] L'équipe Freeduc-Sup, *"La suite de protocoles TCP / IP"*, linux-france.org, 2004

[8] Y. Duchemin, *"TCP / IP"*, yann.duchemin.free.fr, 2000

[9] N. VanHaute, *"Le protocole ARP"*, commentcamarche.net, 2014

[10] A. Vaucamps, D. Dadi, "Cisco : Installer et configure un routeur", fr.scribd.com, 2013

[11] N. VanHaute, *"Le CIDR"*, commentcamarche.net, 2014

[12] Association Linux Online, *"Masque et masque inverse"*, linuxtricks.fr, 2014

[13] A. Amine, *"Mise en œuvre d'un cœur réseau IP/MPLS"*, memoireonline.com, 2011

[14] P. Ayers, C Matthews, B Yates, *"High-level Data Link Control*", wikipedia.com, 2014

[15] S. Fontaine, *"Décryptez votre Hash Cisco 7"*, authsecu.com, 2014

[16] Cisco, *"Sécurisation des entreprises"*, cisco.com, 2001

[17] D. Alberghetti, *"CCNA Security"*, danscourses.com, 2014

[18] E. Bahri, N. Harbi, D.M. Farid, *"Application: Detection d'intrusion"*, waset.org, 2010

[19] H. Loria, *"Pattern matching et detection d'intrusion"*, madchat.fr, 2014

[20] A. Gaudelas, *"Classer ses dépenses à l'aide de la classification bayésienne naïve"*, blog.octo.com, 2011

[21] S. Meharouoech, *"Optimisation de la fiabilité et la pertinence des Systèmes de Détection et Prévention d'Intrusions"*, supcom.mincom.tn, 2010

[22] M. Sharina, K. Jindal, A. Kumar, *"Intrusion Detection System using Bayesian Approach for wireless Network"*, research.ijcaonline.org, 2012

FICHE DE RENSEIGNEMENTS

Nom : RAMILIJAONA

Prénom : ‘Mpoina

Adresse de l’auteur : Lot F I115 Ambohimahitsy
Ambohimangakely
Antananarivo 103 – Madagascar
Tel : +261 32 61 889 71
E-Mail: mpoina@gmail.com
Skype: mpoina.ramilijaona
B.P : 8466 Antananarivo 101

Titre du mémoire :

« SECURISATION DES SYSTEMES D’INFORMATION : RENFORCEMENT DES ROUTEURS CISCO ET REALISATION D’UN SYSTEME DE DETECTION D’INTRUSION »

Nombre de pages : 107

Nombre de tableaux : 7

Nombre de figures : 38

Mots clés :

SSI, CISCO, IDS, CCNA Security, Data protection

Directeur de mémoire :

Nom : RATSIMBAZAFY

Prénoms : Andriamanga

Grade : Maître de conférences

Tel : +261 33 75 638 84

E-mail : ndriamanga@gmail.com

RESUME

« Le système d'information représente un patrimoine essentiel d'une organisation, qu'il convient de protéger. La sécurité informatique consiste à garantir que les ressources matérielles ou logicielles d'une organisation sont uniquement utilisées dans le cadre prévu ». La sécurité informatique est un défi d'ensemble qui concerne une chaîne d'éléments : les infrastructures matérielles de traitement ou de communication, les logiciels (systèmes d'exploitation ou applicatifs), les données, le comportement des utilisateurs. Les conséquences d'éventuelles attaques peuvent concerner la vie privée d'une ou plusieurs personnes, notamment par la diffusion d'informations confidentielles comme ses coordonnées bancaires, sa situation patrimoniale, ses codes confidentiels. Deux principes essentiels sur la sécurisation d'un système d'information peuvent être réalisés : par la sécurisation des routeurs et par l'implémentation d'une application de détection d'intrusion.

ABSTRACT

"Information system is a key asset of organizations, which should be protected. IT security is to ensure that the hardware and software resources of an organization are used only in the space provided." Computer security is a challenge that affects all string elements: physical infrastructure processing or communications, software (operating systems or applications), data and user behavior. Drawback of possible attacks can affect the privacy of one or more persons, including the dissemination of confidential information such as bank details, financial situation, secret codes. Two fundamental principles of the information system security can be applied: secure routers and implement intrusion detection software.