orange™ Business Services

Les cahiers de la Sécurité

La sécurité de la téléphonie sur IP en entreprise

Cédric Baillet

sommaire

en synthèse

L'intégration des services de communication au sein des infrastructures réseaux est un phénomène qui a commencé il y a bien longtemps avec par exemple les services de messagerie mail ou encore de messagerie instantanée. Aujourd'hui, ce phénomène continue à prendre de l'ampleur avec l'arrivée de la téléphonie sur IP (ToIP), de la messagerie unifiée ou encore des applications CTI. A l'origine de cet engouement, deux grands facteurs : la réduction des coûts et la flexibilité de la solution qui permet d'intégrer facilement de nouveaux services.

Bien qu'ayant le même but à la base, la téléphonie sur IP et la téléphonie traditionnelle utilisent des architectures et des technologies très différentes. En effet, au sein d'un réseau informatique classique, la voix et la signalisation sont désormais considérées comme des données et transitent par le LAN, le WAN voire dans certains cas Internet, tandis que dans une architecture de téléphonie classique, chaque ligne a un circuit physique privé ainsi qu'une infrastructure dédiée à cette application.

La téléphonie sur IP est digitalisée puis encodée avant d'être diffusée sur le réseau. Elle est désormais considérée comme de la donnée et se comporte comme telle, mais possède en même temps des contraintes particulières imposant des règles sur la qualité de service et la disponibilité. Créer une architecture robuste respectant ces contraintes tout en étant sécurisée n'est pas une tâche facile.

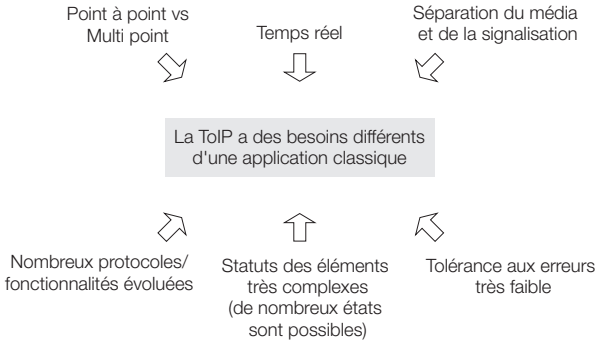
Comme toute nouvelle technologie, la téléphonie sur IP introduit de nouveaux risques et de nouvelles attaques. Héritant à la fois des propriétés de la téléphonie traditionnelle et des réseaux de données, la téléphonie sur IP est donc susceptible de connaître les problématiques sécurité de ces deux mondes. Ainsi, les attaques rencontrées dans la téléphonie traditionnelle comme la manipulation des protocoles de signalisation ou encore le "phreaking" (détournement du système de téléphonie pour éviter la facturation) trouvent un pendant en téléphonie sur IP. Le but de ces attaques lui ne change pas : la fraude et la manipulation de données. Par ailleurs, la sécurité des réseaux informatiques, plus complexes, offre un panel d'attaques beaucoup plus large.

Ainsi, tous les éléments du réseau, du niveau physique jusqu'à la couche applicative, font partie intégrante de la sécurité de la téléphonie sur IP et offrent des possibilités d'attaques importantes et potentiellement inhabituelles en téléphonie classique comme les dénis de service.

les particularités de la téléphonie sur IP

L'une des erreurs souvent commises lorsque l'on considère la sécurité de la ToIP est l'analogie entre cette technologie et les applications couramment utilisées.
En effet, les solutions de sécurité classiques ne prennent généralement pas en compte les particularités de la ToIP comme la nécessité de tendre vers le "temps réel" pour avoir une conversation compréhensible. Ceci demande des mécanismes de qualité de service stricts comme un temps de latence de maximum 150 ms ou encore une perte de paquets inférieure à 3%.

Les particularités de la ToIP



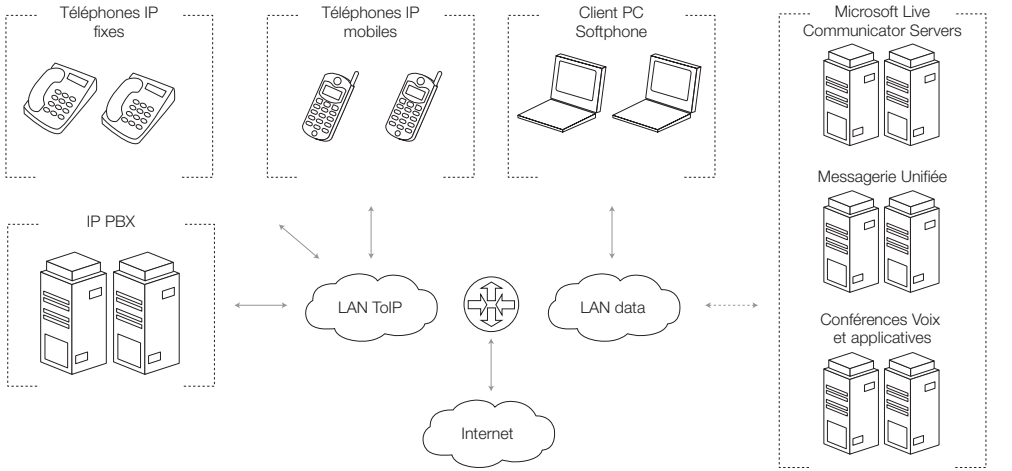
Ainsi, en prenant en compte ce type de facteurs, nous pouvons nous apercevoir que la ToIP est une application sensible, susceptible d'être très facilement perturbée et qui par conséquent pourrait être une cible de choix pour toutes les attaques considérées comme "déni de service"(DoS). On pourra par ailleurs envisager des déclinaisons adaptées à la ToIP (comme l'envoi massif de paquets de signalisations forgés par l'attaquant) et impactant de nombreux périphériques et applicatifs ToIP (les téléphones IP pouvant ainsi se retrouver figés et les serveurs voir leurs fonctionnements perturbés par une attaque n'occupant même pas 1Mb/sec). Notons qu'il est désormais plus intéressant d'avoir un très fort taux d'émission de paquets plutôt qu'une occupation importante de la bande passante pour perturber ces nouvelles applications. En effet, cela imposera à la victime un temps de traitement des paquets émis par l'attaquant beaucoup plus important et par conséquent une possible dégradation de la qualité de service en dessous des niveaux tolérables.

les architectures ToIP

généralités

En simplifiant, nous pouvons considérer que la ToIP doit établir et gérer les sessions de communication lors de la transmission de données “voix” sur des réseaux IP.
Les technologies ToIP pourront par ailleurs être adaptées pour supporter des transmissions de données avec différents formats, comme la vidéo. Ainsi, il est nécessaire de maintenir une transmission fiable durant toute une conversation, puis de mettre fin à cette dernière lorsque l’un des interlocuteurs le décide.

Exemple d'architecture ToIP



Deux grandes familles de protocoles sont utilisées en ToIP pour réaliser ces actions : les protocoles portant sur la signalisation et ceux portant sur le transport de la voix.

Aujourd’hui, deux grands types d’architecture ont émergé comme standards et sont largement déployés : H323 et SIP.

H323

H323 est un standard de l'UIT (Union Internationale des Télécommunications) pour la transmission de la voix et de la vidéo sur des réseaux "par paquets". H323 est aujourd'hui un "Framework" englobant un grand nombre de protocoles de signalisation tels que H225, H245, etc. et RTP (Real time Transmission Protocol) comme protocole de transmission des données. Chacun de ces protocoles a un rôle spécifique dans l'établissement des appels et est basé pour la grande majorité sur un adressage dynamique des ports.

Un réseau H323 est constitué d'un grand nombre de terminaux utilisateurs, de passerelles, éventuellement d'un "gatekeeper" (résolution des adresses et contrôle de la bande passante), de MCU (Multipoint Control Unit) pour gérer les conférences, et de serveurs pour sauvegarder les configurations des terminaux utilisateurs.

SIP (Session Initiation Protocol)

SIP est un standard de l'IETF (Internet Engineering Task Force) permettant l'initialisation d'une session de communication bi-directionnelle. Aujourd'hui, SIP a donné son nom à l'ensemble de l'architecture ToIP reposant sur ses fonctionnalités de signalisation. Par ailleurs, il est important de noter que SIP n'est pas spécifique à la ToIP et peut être utilisé dans toutes les technologies utilisant la notion de session. SIP est basé sur du texte à l'instar de http et peut être transmis à l'aide d'UDP, TCP ou SCTP. UDP permet de réduire la bande passante utilisée tout en améliorant la vitesse et l'efficacité de la transmission, tandis que TCP permet d'apporter des contrôles supplémentaires et l'implémentation de la sécurité via SSL/TLS2 par exemple.

L'architecture d'une solution SIP est différente de celle évoquée précédemment pour H323. Cette dernière est composée de terminaux utilisateurs (aussi appelés User Agent), de "proxy" et/ou de serveurs de redirection permettant la retransmission des messages, de serveurs de localisation pour trouver les points de connexion des utilisateurs et enfin d'un "registrar" enregistrant le profil d'un utilisateur lors de sa connexion sur le réseau.

les risques d'attaques

la reconnaissance

Il s'agit tout simplement de mener une enquête sur le réseau et le système téléphonique installé pour les connaître le mieux possible et de trouver des points de vulnérabilité ou encore des informations directement en relation avec un "bug" déjà référencé. Des méthodes comme le "scan" de ports, de plages d'adresses IP ou de numéros de téléphone, la reconnaissance de système via le "fingerprinting" peuvent être utilisées.

Une reconnaissance positive sur un réseau permet de connaître son plan d'adressage IP, les serveurs opérationnels avec les versions installées, les protocoles utilisés dans l'entreprise, les versions d'IOS, etc.

Une fois ces informations réunies, une attaque peut être lancée sur un point bien particulier comme un périphérique réseau, un service Windows ou Unix…

le déni de service (aussi connu sous le nom DoS)

Le DoS est une attaque sur un système informatique, un réseau, qui peut provoquer une interruption du service initialement rendu à l'utilisateur à cause d'une réduction de la bande passante du système, ou de l'utilisation de toutes les ressources système.

Ce type d'attaque peut être réalisé de nombreuses façons. On remarque cependant trois schémas de base :

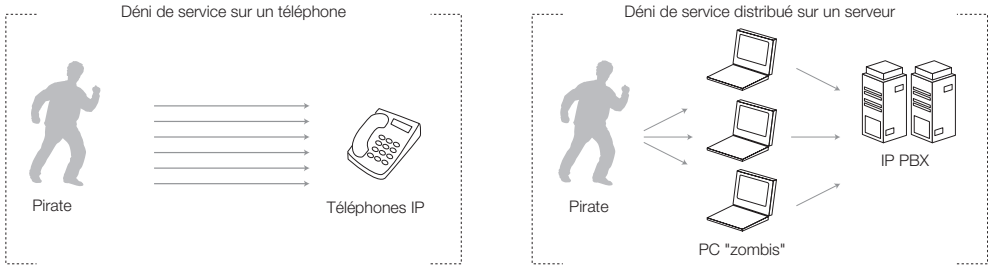
> la réduction des ressources informatiques comme la bande passante, l'espace disque ou la puissance CPU,

> la perturbation des tables de routage,

> la perturbation d'éléments physiques du réseau.

La figure ci-dessous schématise le déroulement de deux types de dénis de service.

Les dénis de services



Le premier cas décrit un déni de service simple sur un téléphone IP : l'attaquant envoie un volume important de messages aléatoires ne pouvant être compris par un téléphone. Celui-ci mettra de plus en plus de temps à traiter ces messages jusqu'au moment où il épuisera ses ressources CPU et ne pourra plus effectuer ses tâches nominales (téléphoner …). Un autre exemple pourrait être tout simplement l'épuisement de la bande passante de son port Ethernet.

Le second cas nous montre le déroulement d'un déni de service distribué sur un serveur. La technique reste identique dans ses fondamentaux mais repose sur une couche intermédiaire : les PC "Zombis". Il s'agit de PC dont l'attaquant a pris le contrôle et qui vont relayer son attaque pour que l'impact soit beaucoup plus important.

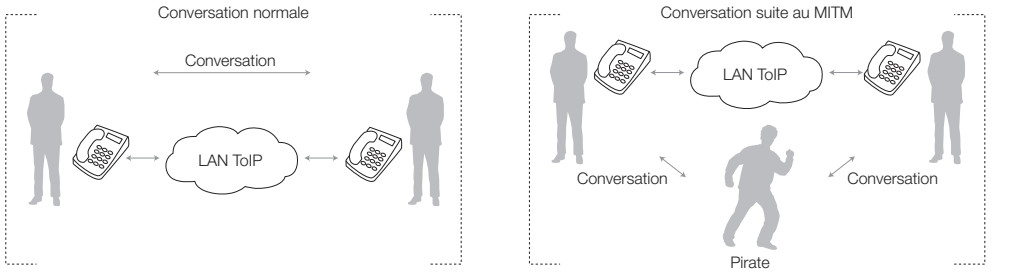
les écoutes

Il s'agit d'intercepter et d'écouter une conversation sans que les interlocuteurs principaux ne soient conscients de ce qui est en train de se produire. En ToIP, cela concerne plus généralement une famille d'attaques permettant à une personne malintentionnée d'intercepter une conversation et de l'enregistrer. Considérant que des numéros de cartes de crédit pour les particuliers ou encore des éléments confidentiels de certains contrats dans les entreprises peuvent être énoncés lors d'une conversation téléphonique, l'impact de ce type d'attaque devient évident.

Man In The Middle (MITM) :

MITM est une attaque au cours de laquelle la personne malveillante a la capacité de lire, insérer ou modifier les messages échangés entre les deux parties, et cela sans qu'elles n'en soient conscientes. Ce type d'attaque peut notamment être utilisé pour réaliser des écoutes ou encore des dénis de service.

Man In The Middle



le "fuzzing"

Le "fuzzing" est une méthode permettant de tester les logiciels et de mettre en évidence des disfonctionnements potentiels. Une des méthodes couramment utilisées consiste à entrer des informations tronquées dans le système pour voir comment ce dernier réagit. En cas de comportement non prévu, il suffit alors d'implémenter un "patch" pour corriger le problème.

Cette méthode peut cependant être détournée par des personnes malintentionnées pour trouver des faiblesses dans un système, ToIP dans notre cas. Ce type d'attaque peut éventuellement permettre en cas de succès de rentrer dans le système avec des droits d'administrateurs, de causer des délais c'est-à-dire d'allonger le temps de réponse, ou tout simplement de mettre le système hors service.

les fraudes téléphoniques

Ce type d'attaque consiste à contourner le système téléphonique mis en place pour pouvoir passer des appels non autorisés.

Ce type d'abus peut venir :

> soit d'une mauvaise configuration du système détectée par l'utilisateur malintentionné qui va en profiter,

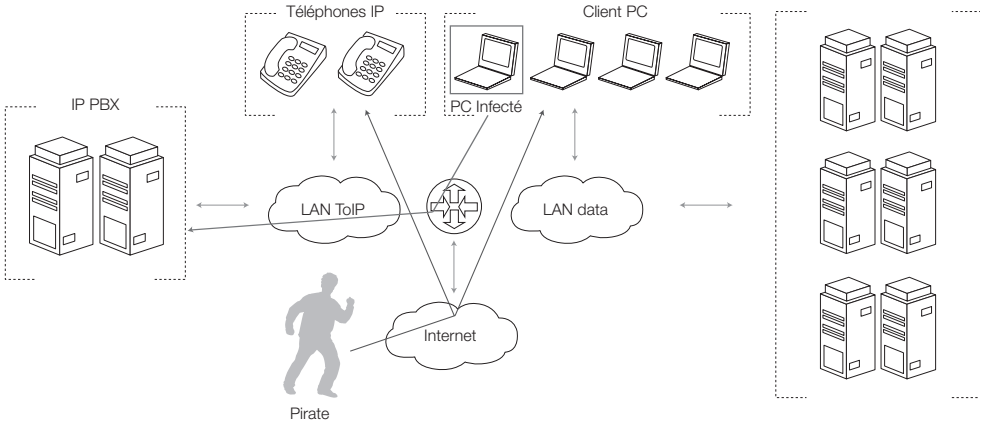
> soit d'une personne malintentionnée qui va utiliser une des méthodes évoquées plus haut.

le ToIP SPAM ou SPIT (Spam over Internet Telephony)

Le SPIT correspond à des messages téléphoniques non sollicités qui peuvent être envoyés à un utilisateur comme à l'ensemble de l'entreprise. En plus du dérangement de l'utilisateur, ce type de message peut provoquer une surcharge du système tant au niveau du réseau que sur les différents serveurs.

Les auteurs de ce type d'attaques sont souvent difficiles à tracer sur Internet. Ils peuvent donc envoyer des messages qui ne sont pas seulement publicitaires mais qui peuvent aussi être exploités pour organiser des actions frauduleuses, pour utiliser des ressources non autorisées ou encore pour récupérer des informations confidentielles.

SPIT



exemples et statistiques

quelques exemples concrets des risques encourus

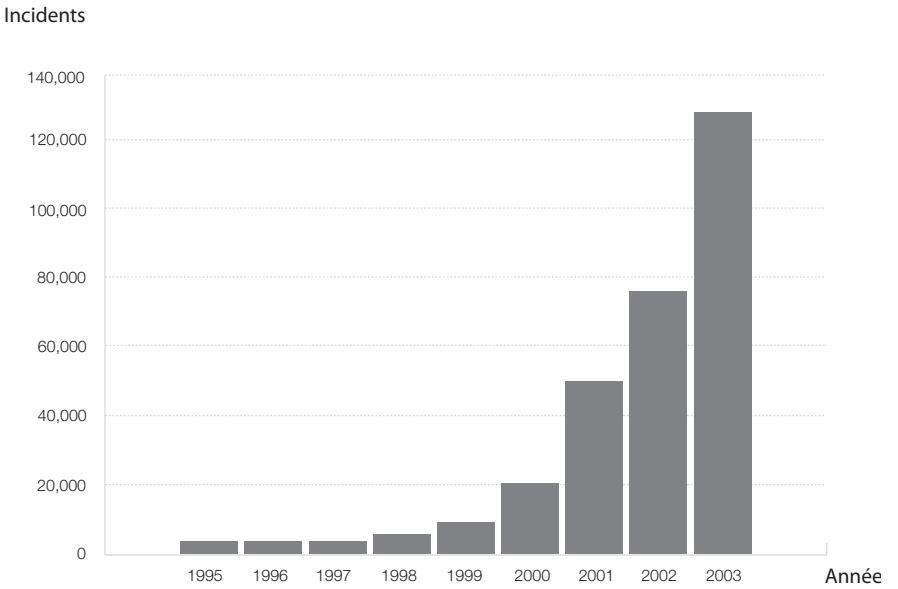
Type de risque	Menaces
Interruption du Service et des données ToIP	Envoi massif de paquets réseaux pour perturber le flux du protocole de signalisation Voix
	Envoi massif de paquets réseaux pour perturber le flux Voix
	Envoi massif de paquets TCP/UDP/ICMP à destination du téléphone IP pour fragiliser sa pile IP
	Exploitation des failles et des faiblesses d'implémentation des protocoles Voix (signalisation ou média)
	Exploitation des failles ou faiblesses d'implémentation de l'OS supportant l'IP PBX
	Déni de Service sur l'accès sans fil
	Déni de Service sur les services réseaux
	Attaque sur les systèmes d'authentification utilisateurs présents dans une solution de ToIP
	Injection de paquets Voix
	Modification des paquets Voix
	Modification de la qualité de service
	Modification des VLAN
Vol des services et données ToIP	Utilisation des techniques de "social engineering" pour contourner les limitations imposées par l'administrateur
	Connexion d'un téléphone IP frauduleux
	Attaque de type "cache poisoning"
	Détournement des appels ToIP
	Vol de données des applications ToIP (ex taxation)
	Ecoute des conversations
	Interception des protocoles de signalisations
	Fraude téléphonique (on retrouve le même type de faiblesse que sur un système classique)
	Utilisation détournée de la messagerie ou accès aux données utilisateurs

Un descriptif plus complet est disponible sur le site http://toipsa.com

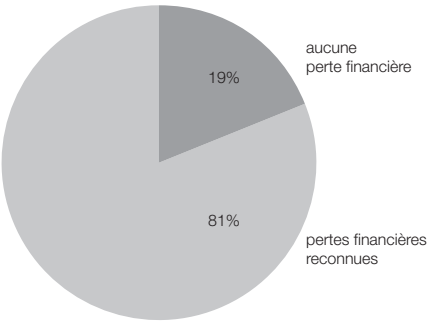
Téléchargez le pdf sur THREAT ANATOMY dans la rubrique "Activities"

quelques statistiques

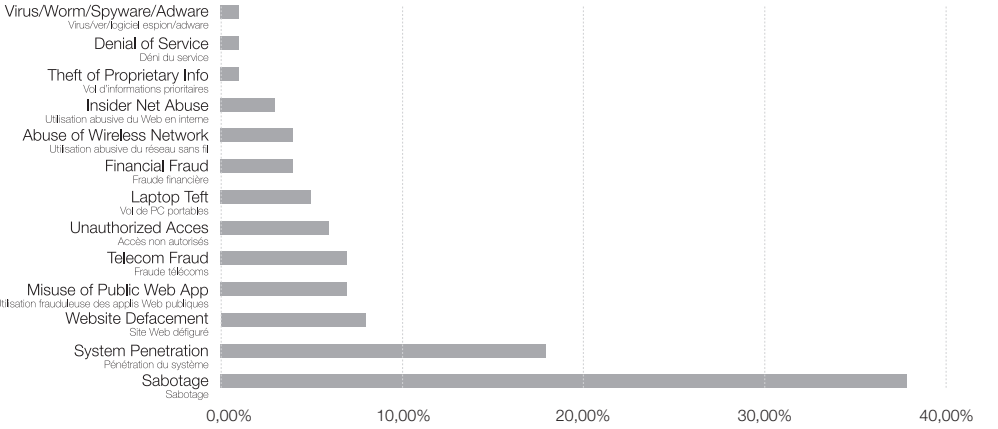
L'augmentation du nombre d'alertes sécurité (source : www.cert.org)



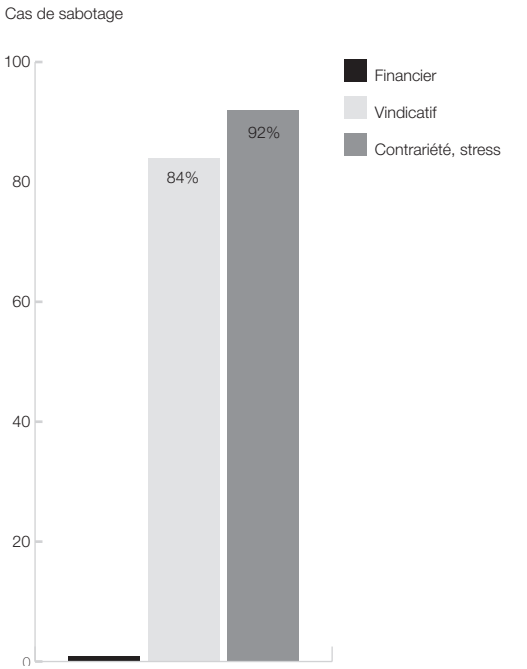
Les risques financiers lors d'une attaque (source : www.cert.org)



Répartition des pertes financières dues à un incident sécurité (source CSI/FBI)



Les motifs poussant un employé "à la faute" (source : www.cert.org)



les solutions existantes pour sécuriser une solution de ToIP

la sécurité physique

La sécurité physique est une partie essentielle de tout environnement sécurisé.
Elle doit permettre la limitation des accès aux bâtiments et équipements (ainsi qu'à toutes les informations qu'ils contiennent) évitant ainsi les intrusions inopportunes, le vandalisme, les catastrophes naturelles, et les dommages accidentels (pic d'intensité électrique, température trop élevée…).

A moins que le trafic Voix ne soit chiffré sur le réseau, toute personne ayant un accès physique au réseau d'une société peut potentiellement se connecter à tout moment et intercepter des communications. Les lignes téléphoniques classiques peuvent certes subir le même type d'attaque, mais les prises LAN présentes dans la plupart des bureaux permettent non seulement un accès beaucoup plus simple au réseau, mais évitent aussi et surtout au pirate de se faire remarquer. En effet, qui soupçonnerait une personne en train de travailler dans un bureau de mener une attaque sur le système de téléphonie ?
Même avec le chiffrement des communications mis en place, un accès physique aux serveurs Voix ou aux passerelles peut permettre à un attaquant d'observer le trafic (qui appelle qui ? à quelle fréquence ? etc.). Une politique de contrôle d'accès pour restreindre l'accès aux composants du réseau de ToIP via des badgeuses, serrures, service de sécurité, etc., permettra d'établir un premier périmètre sécurisé.

Lors de la mise en place d'un système de ToIP, l'alimentation électrique doit être étudiée en détail pour éviter toute interruption de service due à une coupure de courant.
Deux possibilités peuvent être utilisées pour alimenter le poste IP :

> brancher le téléphone sur le secteur via un transformateur,

> utiliser le protocole PoE (Power over Ethernet – alimentation électrique du poste via le réseau informatique).

L'utilisation du PoE se répand de plus en plus mais soulève de nouvelles questions autour des commutateurs assurant cette fonction :

> Chaque poste téléphonique nécessite 15,4W pour fonctionner. Ainsi, plus le nombre de postes est élevé, plus le commutateur est obligé de délivrer une puissance électrique importante. La chaleur dégagée par le commutateur sera donc directement proportionnelle au nombre de postes à alimenter. Une attention toute particulière doit donc être apportée à la climatisation des locaux.

> L'alimentation électrique des commutateurs doit absolument être redondée pour éviter des coupures du service téléphonique lors de problème sur le réseau électrique. On peut envisager la mise en place d'UPS (attention à la capacité) ou encore de double alimentation électrique pour les commutateurs via deux circuits distincts.

Il est important d'inclure les budgets nécessaires pour gérer les questions d'alimentation électrique et de climatisation pour obtenir le vrai coût d'une solution de ToIP.

Il est important d'inclure les budgets nécessaires pour gérer les questions d'alimentation électrique et de climatisation pour obtenir le vrai coût d'une solution de ToIP.

la sécurisation des serveurs

L'ensemble des serveurs participant à une solution de ToIP doit respecter une procédure de mise en place standard et être sécurisé avant toute connexion au réseau.

Une seule équipe au sein de l'entreprise doit être en charge de la rédaction des procédures d'installation et de sécurisation des serveurs et cela quel que soit le type de système (Windows, Linux, Unix propriétaire, etc.).

La sécurisation des serveurs comprend notamment :

> la suppression des comptes inutiles,

> la vérification du bon niveau de droit des différents comptes,

> la suppression des services inutiles,

> la suppression des logiciels ou modules inutiles,

> le bon niveau de correction par rapport aux publications des éditeurs/constructeurs.

Par ailleurs, nous recommandons un audit régulier des serveurs en production par la même équipe. Celle-ci vérifiera le bon fonctionnement des serveurs et s'assurera que les utilisateurs ne détournent pas les serveurs de leurs fonctionnalités initiales, provoquant alors une baisse du niveau de sécurité de l'entreprise.

La ToIP repose sur un grand nombre de services fournis par le réseau pour fonctionner correctement (diffusion de la configuration, supervision de la solution, localisation des utilisateurs, …). On retrouve notamment les classiques suivants : DNS, DHCP, LDAP, RADIUS, HTTP, HTTPS, SNMP, SSH, TELNET, NTP, TFTP, ainsi que la gestion dynamique de la qualité de service. Idéalement, les services utilisés par l'infrastructure de ToIP devraient être dédiés et les serveurs suivre les conditions de sécurisation citées plus haut.

la supervision

Généralités

Les outils permettant la supervision des réseaux doivent normalement pouvoir être adaptés pour superviser l'ensemble de l'infrastructure convergente téléphonie sur IP et Data. C'est l'un des grands avantages de l'unification des infrastructures. Néanmoins, si la plupart de ces outils fonctionnent sur l'environnement convergé complet, une mise à jour devrait sans doute être nécessaire pour les adapter aux outils de la ToIP et à leurs particularités.

Par ailleurs, il est recommandé de séparer le trafic généré par les solutions de supervision du reste des applications - un mode "out of band" sur un segment réseau dédié est fortement recommandé.

Les solutions des constructeurs sont naturellement performantes et parfaitement adaptées à leurs produits. On notera cependant que le ticket d'entrée reste élevé. Une solution alternative à base de produits "open source" comme MRTG ou BigBrother pourrait alors être envisagée. Attention cependant, il sera nécessaire dans ce cas là :

> de posséder les MIBs adaptées aux produits que l'on souhaite superviser,

> de posséder les compétences pour adapter les produits "open source" aux solutions de téléphonie sur IP,

> de pouvoir libérer le temps nécessaire à la mise en place de ces solutions de supervision.

Par ailleurs, les moyens de surveillance active du réseau et de l'ensemble de ses périphériques ne fournissent pas seulement un niveau de défense supplémentaire mais aussi des moyens de récupérer des informations sur le déroulement d'événements non prévus dans un fonctionnement nominal.

On comprend dans la surveillance active non pas la supervision du réseau évoqué plus haut, mais la détection d'intrusions réseau, la détection d'intrusions système et les remontées d'alarmes sur erreur des journaux système et logs réseau (via SNMP par exemple).

Les tests d'intrusions et de vulnérabilités permettent de valider les niveaux de contrôle de la sécurité évoqués ci-dessus.

Syslog et SNMP

Syslog

La fonctionnalité Syslog permet d'avoir une technique pour gérer les échanges de notification au travers d'un réseau IP entre un client et un serveur. Les messages échangés ne sont par chiffrés par défaut puisqu'il s'agit d'un protocole très simple ; il est donc nécessaire de restreindre ce type d'application à un réseau interne ou protégé.

Certaines alternatives existent et offrent des fonctionnalités de sécurité renforcées. Leur utilisation est recommandée pour éviter des fuites d'informations sur le type d'architecture ToIP mis en place ainsi que ses faiblesses potentielles.

SNMP

SNMP signifie "Simple Network Management Protocol" (protocole simple de gestion de réseau). Il s'agit d'un protocole qui permet aux administrateurs réseau de gérer les équipements du réseau et de diagnostiquer les problèmes de réseau.

Le système de gestion de réseau est basé sur deux éléments principaux : un superviseur et des agents. Le superviseur est la console qui permet à l'administrateur réseau d'exécuter des requêtes de management. Les agents sont des entités qui se trouvent au niveau de chaque interface connectant l'équipement managé au réseau et permettant de récupérer des informations sur différents objets.

Commutateurs, routeurs et serveurs sont des exemples d'équipements contenant des objets pouvant être supervisés. Ces objets peuvent être des informations sur les matériels, des paramètres de configuration, des statistiques de performance et autres objets qui sont directement liés au comportement en cours de l'équipement en question. Ces objets sont classés dans une sorte de base de données appelée MIB ("Management Information Base"). SNMP permet le dialogue entre le superviseur et les agents afin de recueillir les objets souhaités dans la MIB.

L'architecture de gestion du réseau proposée par le protocole SNMP est donc basée sur trois principaux éléments :

> les équipements managés (managed devices) sont des éléments du réseau (ponts, hubs, routeurs ou serveurs), contenant des "objets de gestion" (managed objects) pouvant être des informations sur le matériel, des éléments de configuration ou des informations statistiques,

> les agents, c'est-à-dire une application de gestion de réseau résidant dans un périphérique et chargé de transmettre les données locales de gestion du périphérique au format SNMP,

> les systèmes de management de réseau (network management systems notés NMS), c'est-à-dire une console au travers de laquelle les administrateurs peuvent réaliser des tâches d'administration.

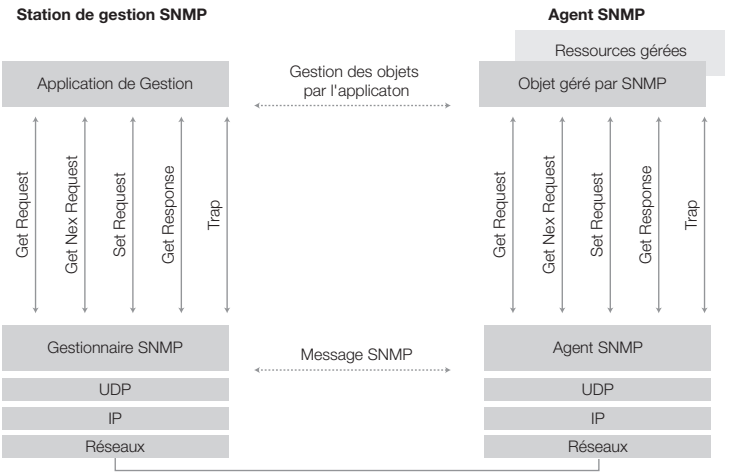
Concrètement, dans le cadre d'un réseau, SNMP est utilisé :

> pour administrer les équipements,

> pour surveiller le comportement des équipements.

Une requête SNMP est un datagramme UDP usuellement à destination du port 161.
Les schémas de sécurité dépendent des versions de SNMP (v1,v2 ou v3).
Dans les versions 1 et 2, une requête SNMP contient un nom appelé communauté, utilisé comme un mot de passe. Il y a un nom de communauté différent pour obtenir les droits en lecture et pour obtenir les droits en écriture. Dans bien des cas, les lacunes de sécurité que comportent les versions 1 et 2 de SNMP limitent l'utilisation de SNMP à la lecture des informations. Un grand nombre de logiciels libres et payants utilisent SNMP pour interroger régulièrement les équipements et produire des graphes rendant compte de l'évolution des réseaux ou des systèmes informatiques.

Le protocole SNMP définit aussi un concept de “trap”. Une fois défini, si un certain événement se produit, comme par exemple le dépassement d'un seuil, l'agent envoie un paquet UDP à un serveur. Ce processus d'alerte est utilisé dans les cas où il est possible de définir simplement un seuil d'alerte.



Le fonctionnement du protocole SNMP

NIDS et HIDS

Les systèmes de détection d’intrusion réseau ou NIDS (Network-based Intrusion Detection Systems) ont pour but d’alerter les administrateurs de la solution en cas de trafic anormal ou jugé malicieux.

Un trafic qualifié de malicieux peut correspondre à la propagation d’un ver ou d’un exploit connu (programme développé spécifiquement pour exploiter une faiblesse clairement identifiée dans un logiciel donné), tandis qu’un trafic anormal fait plutôt référence à une utilisation détournée du réseau (par rapport à son but premier) et aux règles de sécurité définies (une connexion en Peer-to-Peer par exemple).

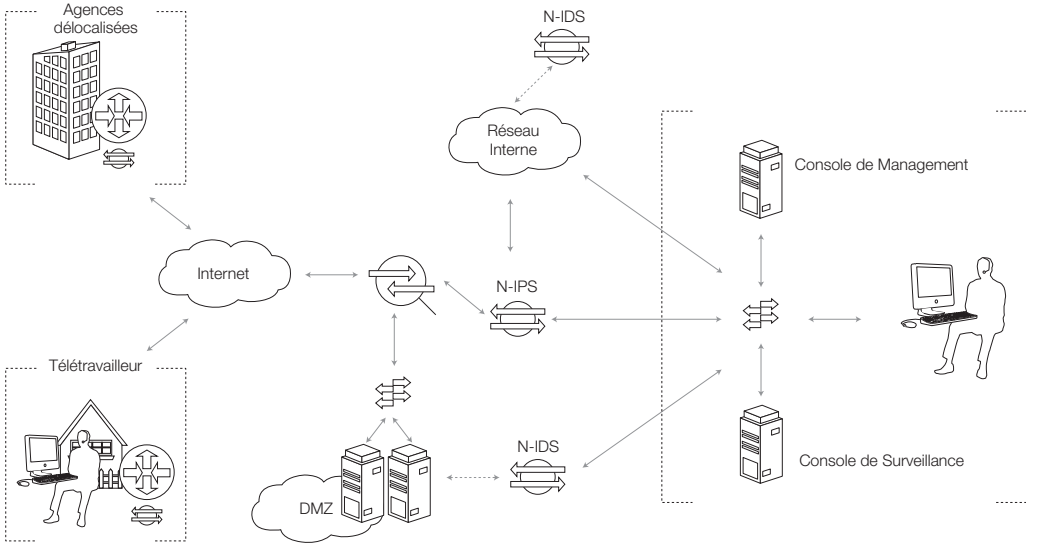
La détection d’intrusion système ou HIDS (Host-based Intrusion Detection System) fait référence à une famille de logiciels collectant des informations sur les serveurs ou encore les postes utilisateurs pour les analyser. Ce type d’action permet d’avoir une vue détaillée sur les différentes activités et d’identifier les processus ou les utilisateurs ayant des activités non autorisées.

La mise en place de ces deux types de systèmes est conseillée.

L’utilisation de sondes NIDS aux points clés du réseau – accès internet, DMZ, etc. - permet de superviser une partie importante du trafic aux points sensibles de l’infrastructure.
La gestion courante de ces sondes est réalisée autant que possible via des liens sécurisés (chiffrés). On notera que pour une efficacité maximum, et donc un minimum d’erreur de qualification dans les incidents (“false positive”), des mises à jour régulières et une configuration précise et détaillée seront nécessaires .

Les sondes d’intrusion système doivent être mises en place sur l’ensemble des serveurs participant aux infrastructures ToIP (DNS, DHCP, TFTP, RADIUS, NTP, LDAP…) et suivre un processus de supervision en temps réel au sein de l’entreprise.

Un exemple de déploiement de sonde N-IDS sur le réseau



N-IDS versus N-IPS

N-IDS pour Network-based Intrusion Detection System — Terme traditionnellement utilisé pour désigner des solutions travaillant en mode « promiscuité » (comme les sniffers) et capables de remonter des alertes lors d'intrusions sur le réseau.

N-IPS pour Network-based Intrusion Prevention System — Terme commun pour désigner les solutions situées en passerelle ou en coupure sur un segment du réseau, forçant le trafic à circuler par le système et proposant des contre-mesures sur le réel trafic analysé.

Tests d'intrusions et de vulnérabilités

Un test de vulnérabilités est un test d'identification de failles connues. Le résultat de ce test est un tableau synthétique dressant la liste des équipements concernés pour chaque faille trouvée.

Les tests d'intrusions consistent à éprouver les moyens de protection d'un système d'information en essayant de s'introduire dans le système en situation réelle. Les résultats du test de vulnérabilités pourront être exploités pour atteindre ce but.

On distingue généralement deux méthodes distinctes :

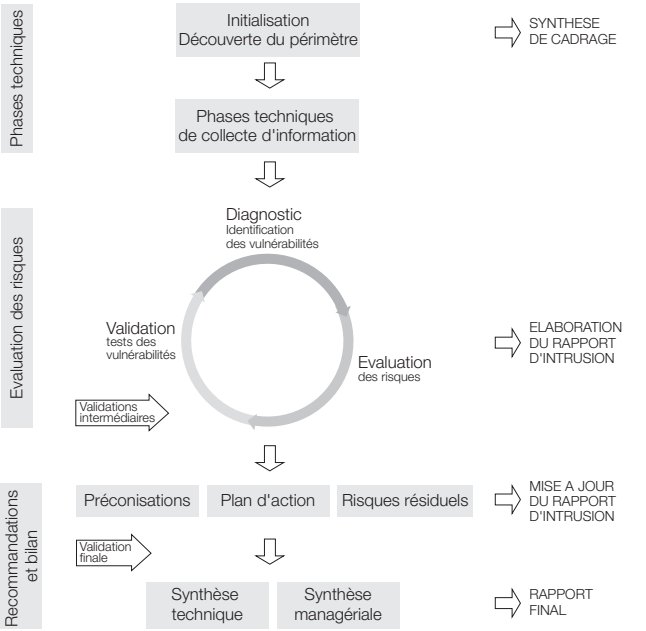
> la méthode dite "boîte noire" consistant à essayer d'infiltrer le réseau sans aucune connaissance du système, afin de réaliser un test en situation réelle,

> la méthode dite "boîte blanche" consistant à tenter de s'introduire dans le système en ayant connaissance de l'ensemble du système, afin d'éprouver au maximum la sécurité du réseau.

Une telle démarche doit nécessairement être réalisée avec l'accord du plus haut niveau de la hiérarchie de l'entreprise, car elle peut aboutir à des dégâts éventuels d'autant que ces méthodes sont interdites par la loi sans l'autorisation du propriétaire du système.

Un test d'intrusions, lorsqu'il met en évidence une faille, est un bon moyen de sensibiliser les acteurs d'un projet. A contrario, il ne permet pas de garantir la sécurité du système, puisque des vulnérabilités peuvent avoir échappé aux testeurs. Les audits de sécurité permettent d'obtenir un bien meilleur niveau de confiance dans la sécurité d'un système : en effet, les aspects organisationnels et humains sont pris en compte et la sécurité est analysée de l'intérieur.

Le déroulement d'un test d'intrusion



l'authentification des utilisateurs

Généralités

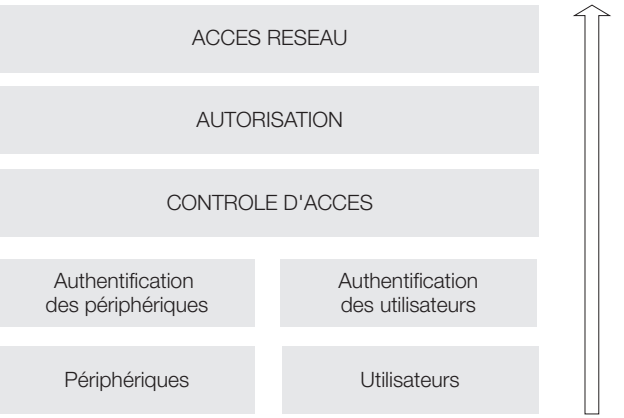
L'une des méthodes les plus importantes pour anticiper une attaque sur un système de téléphonie est de déterminer clairement l'identité des périphériques ou des personnes participant à la conversation. On parlera d'authentification.

L'authentification est généralement basée sur un secret partagé par les différentes parties (vous êtes authentifié(e) si vous connaissez le secret) ou sur un système de clés publiques et de certificats (vous êtes authentifié(e) si vous possédez la clé correcte). On parlera de PKI (voir le point p.32 pour plus de détails).

Si l'authentification établit l'identité des différents membres participant à l'échange, l'autorisation permettra d'établir le degré d'accès aux systèmes que vous pouvez posséder. Ces deux paramètres sont importants et doivent toujours être associés.

La figure ci-dessous montre un système d'authentification/autorisation. Bien qu'il ne respecte pas parfaitement les règles de l'ISO, il illustre bien que les utilisateurs et les périphériques doivent être authentifiés, souvent via des processus différents. Quoi qu'il en soit, l'authentification peut être séparée de l'autorisation sans poser de problème de sécurité supplémentaire.

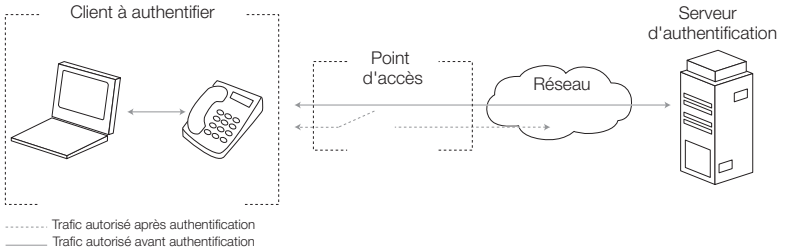
Système d'authentification/autorisation schématique



802.1X

Le protocole 802.1X permet de restreindre l'accès au LAN en empêchant les clients non autorisés de se connecter. En effet, ces derniers devront d'abord être authentifiés, puis autorisés par un serveur d'authentification, un RADIUS par exemple, avant que le port du commutateur ne soit ouvert et que le réseau ne soit accessible. Le protocole EAP (Extensible Authentication Protocol) est un "framework" utilisé dans la phase d'authentification d'un client. Il fonctionne avec un grand nombre de méthodes d'authentification et n'est donc pas restrictif.

L'authentification 802.1



La plupart des périphériques récents ayant implémenté EAP sont constitués de deux composants principaux : une authentification externe et une authentification interne, séparées par un membre de la famille EAP (PEAP ou EAP-MSCHAPv2). L'authentification externe définit la méthode utilisée pour établir un circuit chiffré entre le client et le serveur d'authentification (un tunnel TLS par exemple). Une fois ce circuit établi, le processus d'authentification interne transmettra les références de l'utilisateur au serveur d'authentification.

Il est recommandé d'utiliser le protocole 802.1X pour l'authentification des périphériques et des utilisateurs que ce soit sur des réseaux câblés ou sans fil (WIFI).

En fonction des environnements, les associations suivantes doivent être mises en place (attention à l'évolution rapide de ces technologies) :

> EAP TLS - PKI

> EAP-PEAP – Clients Windows

> EAP-TTLS – Clients autres

On notera qu'en fonction du type de serveur RADIUS utilisé, des paramètres supplémentaires comme les VLAN pourront être assignés dynamiquement au client authentifié.

Les PKI (Public Key Infrastructure)

Une PKI (Public Key Infrastructure) est un système de gestion des clés publiques qui permet de gérer des listes importantes de clés publiques et d'en assurer la fiabilité, pour des entités généralement dans un réseau.

Une PKI (Public Key Infrastructure) est un système de gestion des clés publiques qui permet de gérer des listes importantes de clés publiques et d'en assurer la fiabilité, pour des entités généralement dans un réseau.

La PKI offre un cadre global permettant d'installer des éléments de sécurité tels que la confidentialité, l'authentification, l'intégrité et la non-répudiation tant au sein de l'entreprise que lors d'échanges d'information avec l'extérieur.
Une infrastructure PKI fournit donc quatre services principaux :

> fabrication de bi-clés,

> certification de clés publiques et publication de certificats,

> révocation de certificats,

> gestion de la fonction de certification.

Signature :

Dans la signature nous avons une bi-clé : une clé (privée) pour la création de signature et une clé (publique) pour la vérification de signature. Voici les différentes étapes de la signature d'un message :

> à l'aide de la clé privée de signature de l'expéditeur, une empreinte de taille fixe connue sous le nom "message digest" est générée par hachage en utilisant l'algorithme SHA-1 ou MD5, le plus utilisé étant SHA-1. Cette empreinte est ensuite cryptée avec cette clé privée de signature,

> on joint au message l'empreinte et le certificat contenant la clé publique de signature,

> le destinataire vérifie la validité du certificat et sa non révocation dans l'annuaire,

> le destinataire transforme l'empreinte avec la clé publique de signature ainsi validée. Cette opération permet de s'assurer de l'identité de l'expéditeur,

> ensuite le destinataire génère une empreinte à partir du message reçu en utilisant le même algorithme de hachage. Si les deux empreintes sont identiques, cela signifie que le message n'a pas été modifié. Donc la signature vérifie bien l'intégrité du message ainsi que l'identité de l'expéditeur.

Exemples d'algorithmes de signature : RSA,DSA

Définitions

> Confidentialité : les informations échangées deviennent illisibles, cette confidentialité est assurée par le chiffrement.

> Authentification : identification de l'origine de l'information.

> Non-répudiation : l'émetteur des données ne pourra pas nier être à l'origine du message.

> Intégrité : fonction permettant d'assurer que l'information n'a pas subi de modification .

Chiffrement

Il y a deux types de chiffrement possible :

> Chiffrement à clé secrète (symétrique) : L'émetteur utilise une clé pour chiffrer le message et le destinataire utilise la même clé (le même algorithme mais en sens inverse) pour déchiffrer le message.

> Chiffrement à clé publique (asymétrique) : Un message chiffré avec une clé publique donnée ne peut être déchiffré qu'avec la clé privée correspondante. Par exemple si A souhaite envoyer un message chiffré à B, il le chiffrera en utilisant la clé publique de B (qui peut être publiée dans un annuaire). La seule personne qui déchiffre le message est le détenteur de la clé privée de B.

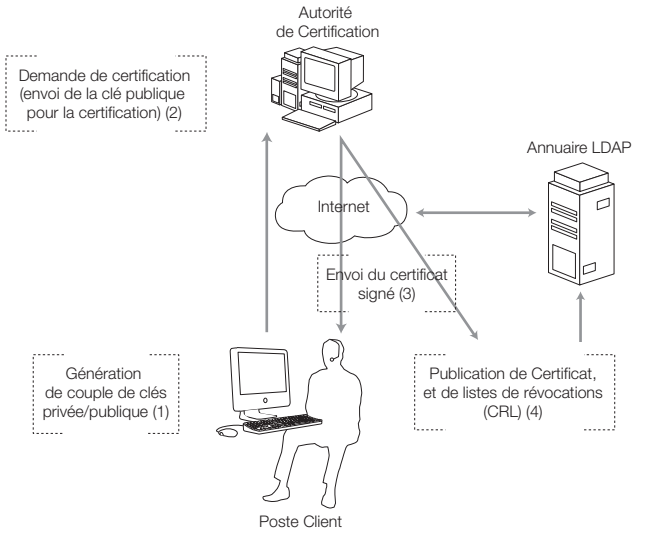
Exemples d'algorithmes de chiffrement :

> Symétrique : DES, AES.

> Asymétrique : RSA.

Dans une infrastructure à clés publiques, pour obtenir un certificat numérique, l'utilisateur fait une demande auprès d'une autorité d'enregistrement. Celle-ci génère un couple de clés (clé publique, clé privée), envoie la clé privée au client, applique une procédure et les critères définis par l'autorité de certification - qui certifie la clé publique - et appose sa signature sur le certificat, parfois fabriqué par un opérateur de certification.

Fonctionnement d'une PKI



Sur les solutions ToIP, une PKI donne le moyen de garantir l'identité des utilisateurs et des périphériques, tout en travaillant avec des protocoles comme le 802.1X. C'est donc un moyen permettant de fédérer l'ensemble des identifications de la solution tout en restant suffisamment ouvert pour travailler avec tous les grands protocoles du marché.

la séparation et la sécurisation des flux

Une fois numérisé, le trafic voix n'est plus identifiable comme tel et se confond avec les flux data sur le réseau. Il devient ainsi victime des problèmes rencontrés couramment en data. Si l'on réussit à séparer les flux avant d'en arriver à cette situation, un traitement particulier peut alors être mis en place et permettre de réduire autant que possible ce type d'inconvénients (broadcast, congestions, DoS, …). La séparation des flux voix et data peut être réalisée via l'utilisation de techniques comme les VLAN, la mise en place de qualité de service ou encore de filtrage.

La séparation logique des flux voix et data à l'aide de VLAN permet d'éviter que les incidents rencontrés sur l'un des flux ne puissent perturber l'autre.

VLAN

La séparation logique des flux voix et data à l'aide de VLAN est une mesure fortement recommandée. Elle doit permettre d'éviter que les incidents rencontrés sur l'un des flux ne puissent perturber l'autre. Les VLAN ou réseaux locaux virtuels, peuvent être représentés comme une séparation logique d'un même réseau physique. Cette opération se fait au niveau 2 du modèle OSI. On notera cependant qu'un VLAN est souvent configuré pour correspondre directement à un subnet IP bien identifié, préparant ainsi le travail à effectuer sur la couche supérieure.

Dans un environnement commuté complet, les VLAN vont créer une séparation logique des domaines de broadcast et de collisions – des problèmes dus à ces deux éléments sont souvent rencontrés dans des LAN trop importants ou lorsqu'on utilise encore des hubs. De plus, la réduction des domaines de broadcast permet de réduire le trafic sur le réseau, libérant ainsi plus de bande passante pour les applications utiles et réduisant les temps de traitement sur les périphériques réseau.

On peut utiliser cette technique pour organiser les postes utilisateurs selon leurs situations physiques dans les bâtiments, le service auquel appartient l'utilisateur, la vitesse de connexion, ou tout autre critère ayant du sens dans l'entreprise. Un renforcement de la sécurité peut être réalisé en mettant en place un filtrage inter-VLAN, n'autorisant que les utilisateurs d'un VLAN à y accéder. Le risque de DoS peut ainsi être réduit.

Qualité de service

Un réseau de données travaille par défaut en mode "Best Effort". Concrètement, cela signifie que l'ensemble du trafic possède la même priorité et le même nombre de chances d'arrivée sans être supprimé. Ainsi, si un encombrement de réseau survient, le trafic à supprimer pour rétablir le service sera sélectionné au hasard. Ceci n'est pas acceptable sur un trafic contenant de la voix : une dégradation importante de la qualité ne peut être tolérée par l'utilisateur. Il est donc nécessaire de mettre en place un système de priorité donnée au trafic voix sur les trafics moins sensibles.

Appliquée aux réseaux à commutation classique, la qualité de service ou QoS désigne donc l'aptitude à pouvoir garantir un niveau acceptable de perte de paquets, défini contractuellement, pour un usage donné (voix, vidéo, web, ftp, etc.).

La Qualité de Service (QoS) garantit un niveau acceptable de perte de paquets, pour donner par exemple, la priorité au trafic voix sur les trafics moins sensibles comme la data.

Les principaux critères permettant d'apprécier la qualité de service sont les suivants :

> débit (en anglais “bandwidth”), parfois appelé bande passante par abus de langage, il définit le volume maximal d'informations (bits) par unité de temps,

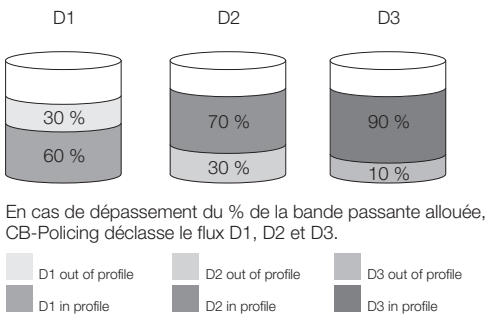
> gigue (en anglais jitter) : elle représente la fluctuation du signal numérique dans le temps ou en phase,

> latence, délai ou temps de réponse (en anglais “delay”) : elle caractérise le retard entre l'émission et la réception d'un paquet,

> perte de paquets (en anglais “packet loss”) : elle correspond à la non-délivrance d'un paquet de données, la plupart du temps due à un encombrement du réseau,

> déséquencement (en anglais “desequencing”) : il s'agit d'une modification de l'ordre d'arrivée des paquets.

Un exemple de QoS : Le Class Base Policing



En cas de dépassement du % de la bande passante allouée, CB-Policing déclasse le flux D1, D2 et D3.

Firewalls

Un pare-feu (appelé aussi coupe-feu, garde-barrière ou “firewall” en anglais), est un système permettant de protéger un ordinateur ou un réseau d'ordinateurs des intrusions provenant d'un réseau tiers (notamment Internet). Le pare-feu filtre les paquets de données échangés avec le réseau. Il s'agit ainsi d'une passerelle filtrante comportant au minimum les interfaces réseau suivantes :

> une interface pour le réseau à protéger (réseau interne),

> une interface pour le réseau externe.

Le système firewall est un système logiciel, reposant parfois sur un matériel réseau dédié, constituant un intermédiaire entre le réseau local (ou la machine locale) et un ou plusieurs réseau(x) externe(s). Il est possible de mettre un système pare-feu sur n'importe quelle machine et avec n'importe quel système pourvu que :

> la machine soit suffisamment puissante pour traiter le trafic,

> le système soit sécurisé,

> aucun autre service que le service de filtrage de paquets ne fonctionne sur le serveur.

Dans le cas où le système pare-feu est fourni dans une boîte noire “clé en main”, on utilise le terme ”d'appliance”.

Fonctionnement d'un système pare-feu

Un système pare-feu contient un ensemble de règles prédéfinies permettant :

> d'autoriser la connexion (“allow”),

> de bloquer la connexion (“deny”),

> de rejeter la demande de connexion sans avertir l'émetteur (“drop”).

L'ensemble de ces règles permet de mettre en œuvre une méthode de filtrage dépendant de la politique de sécurité adoptée par l'entité.

On distingue habituellement deux types de politiques de sécurité permettant :

> soit d'autoriser uniquement les communications ayant été explicitement autorisées : "Tout ce qui n'est pas explicitement autorisé est interdit",

> soit d'empêcher les échanges qui ont été explicitement interdits.

La première méthode est sans nul doute la plus sûre, mais elle impose toutefois une définition précise et contraignante des besoins en communication.

Le filtrage simple de paquets

Un système pare-feu fonctionne sur le principe du filtrage simple de paquets (en anglais “stateless packet filtering”). Il analyse les en-têtes de chaque paquet de données (datagramme) échangé entre une machine du réseau interne et une machine extérieure.

Ainsi, les paquets de données échangés entre une machine du réseau extérieur et une machine du réseau interne transitent par le pare-feu et possèdent les en-têtes suivants, systématiquement analysés par le firewall :

> adresse IP de la machine émettrice (source),

> adresse IP de la machine réceptrice (destination),

> type de paquet (TCP, UDP, etc.),

> numéro de port (rappel : un port est un numéro associé à un service ou une application réseau).

Les adresses IP contenues dans les paquets permettent d'identifier la machine émettrice et la machine cible, tandis que le type de paquet et le numéro de port donnent une indication sur le type de service utilisé.

Les ports reconnus (dont le numéro est compris entre 0 et 1023) sont associés à des services courants (les ports 25 et 110 sont par exemple associés au courrier électronique, et le port 80 au Web). La plupart des dispositifs pare-feu sont au minimum configurés de manière à filtrer les communications selon le port utilisé. Il est généralement conseillé de bloquer tous les ports qui ne sont pas indispensables (selon la politique de sécurité retenue).

Le port 23 est par exemple, souvent bloqué par défaut par les dispositifs pare-feu car il correspond au protocole Telnet, permettant d'émuler un accès par terminal à une machine distante de manière à pouvoir exécuter des commandes à distance. Les données changées par Telnet ne sont pas chiffrées, ce qui signifie qu'un individu est susceptible d'écouter le réseau et de voler les éventuels mots de passe circulant en clair. Les administrateurs lui préfèrent généralement le protocole SSH, réputé sûr et fournissant les mêmes fonctionnalités que Telnet.

Le filtrage dynamique

Le filtrage simple de paquets ne s'attache qu'à examiner les paquets IP indépendamment les uns des autres, ce qui correspond au niveau 3 du modèle OSI. Or, la plupart des connexions reposent sur le protocole TCP, qui gère la notion de session, afin d'assurer le bon déroulement des échanges. D'autre part, de nombreux services (le FTP par exemple) initient une connexion sur un port statique, mais ouvrent dynamiquement (c'est-à-dire de manière aléatoire) un port afin d'établir une session entre la machine faisant office de serveur et la machine cliente.

Ainsi, il est impossible avec un filtrage simple de paquets de prévoir les ports à laisser passer ou à interdire. Pour y remédier, le système de filtrage dynamique de paquets est basé sur l'inspection des couches 3 et 4 du modèle OSI, permettant d'effectuer un suivi des transactions entre le client et le serveur. Le terme anglo-saxon est “stateful inspection” ou “stateful packet filtering” (le filtrage de paquets avec état).

Un dispositif de pare-feu du type “stateful inspection” est ainsi capable d'assurer un suivi des échanges, c'est-à-dire de tenir compte de l'état des anciens paquets pour appliquer les règles de filtrage. De cette manière, à partir du moment où une machine autorisée initialise une connexion à une machine située de l'autre côté du pare-feu, l'ensemble des paquets transitant dans le cadre de cette connexion sera implicitement accepté par le pare-feu.

Si le filtrage dynamique est plus performant que le filtrage de paquets basique, il ne protège pas pour autant de l'exploitation des failles applicatives, liées aux vulnérabilités des applications. Or ces vulnérabilités représentent la part la plus importante des risques en termes de sécurité.

Le filtrage applicatif

Le filtrage applicatif permet, comme son nom l'indique, de filtrer les communications application par application. Le filtrage applicatif opère donc au niveau 7 (couche application) du modèle OSI, contrairement au filtrage de paquets simple (niveau 3). Le filtrage applicatif suppose donc une bonne connaissance des protocoles utilisés par chaque application.

Un firewall effectuant un filtrage applicatif est appelé généralement “passerelle applicative” (ou “proxy”), car il sert de relais entre deux réseaux en s'interposant et en effectuant une validation fine du contenu des paquets échangés. Le proxy représente donc un intermédiaire entre les machines du réseau interne et le réseau externe, subissant les attaques à leur place. De plus, le filtrage applicatif permet la destruction des en-têtes précédant le message applicatif, ce qui permet de fournir un niveau de sécurité supplémentaire.

Il s'agit d'un dispositif performant, assurant une bonne protection du réseau, pour peu qu'il soit correctement administré. En contrepartie, une analyse fine des données applicatives requiert une grande puissance de calcul et se traduit donc souvent par un ralentissement des communications, chaque paquet devant être finement analysé.

Par ailleurs, le proxy doit nécessairement être en mesure d'interpréter une vaste gamme de protocoles et de connaitre les failles afférentes pour être efficace.

Enfin, un tel système peut potentiellement comporter une vulnérabilité dans la mesure où il interprète les requêtes qui transitent par son biais. Ainsi, il est recommandé de dissocier le pare-feu (dynamique ou non) du proxy, afin de limiter les risques de compromission.

Les limites des firewalls

Un système pare-feu n'offre bien évidemment pas une sécurité absolue, bien au contraire. Les firewalls n'offrent une protection que dans la mesure où l'ensemble des communications vers l'extérieur passe systématiquement par leur intermédiaire et qu'ils sont correctement configurés. Ainsi, les accès au réseau extérieur par contournement du firewall sont autant de failles de sécurité. C'est notamment le cas des connexions effectuées à partir du réseau interne à l'aide d'un modem ou de tout moyen de connexion échappant au contrôle du pare-feu.

De la même manière, l'introduction de supports de stockage provenant de l'extérieur sur des machines internes au réseau ou bien d'ordinateurs portables peut porter fortement préjudice à la politique de sécurité globale.

Enfin pour garantir un niveau de protection maximal, il est nécessaire d'administrer le pare-feu et notamment de surveiller son journal d'activité afin d'être en mesure de détecter les tentatives d'intrusion et les anomalies. Par ailleurs, il est recommandé d'effectuer une veille de sécurité (en s'abonnant aux alertes de sécurité des CERT par exemple) afin de modifier le paramétrage de son dispositif en fonction de la publication des alertes.

Le tableau ci-dessous donne une première liste de protocoles couramment utilisés dans les technologies ToIP. On apportera une attention toute particulière à la gestion des ports dynamiques par les firewalls pour éviter d’ouvrir des plages de ports extrêmement importantes.

Ports couramment utilisés en ToIP

Services	Ports
Skinny	TCP 2000-2002
TFTP	UDP 69
MGCP	UDP 2427
Backhaul (MGCP)	UDP 2428
Tapi/Jtapi	TCP 2748
HTTP	TCP 8080/80
SSL	TCP 443
SCCP	TCP 3224
RTP	16384-32767
SNMP	UDP 161
SNMP Trap	UDP 162
DNS	UDP 53
NTP	UDP 123
LDAP	TCP 389
H.323RAS	TCP 1719
H.323 H.225	TCP 1720
H.323 H.245	TCP 11000-11999
H.323 Gatekeeper Discovery	UDP 1718
SIP	TCP 5060
SIP/TLS	TCP 5061

Translation d’adresses IP (NAT)

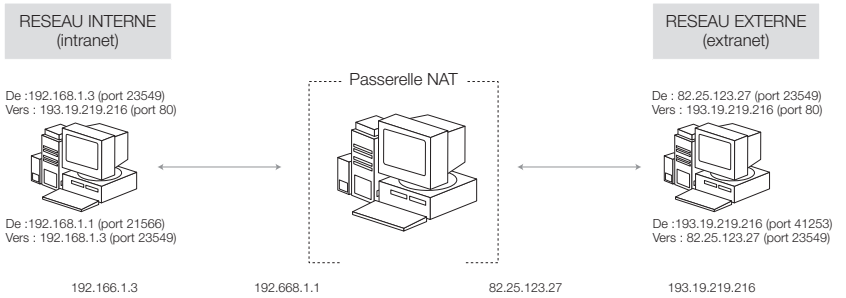
Principe du NAT

Le mécanisme de translation d'adresses (en anglais Network Address Translation noté NAT) a été mis au point afin de répondre à la pénurie d'adresses IP avec le protocole IPv4 (le protocole IPv6 répondra à terme à ce problème).

En effet, en adressage IPv4, le nombre d'adresses IP routables (donc uniques sur la planète) n'est pas suffisant pour permettre à toutes les machines de se connecter à Internet.

Le principe du NAT consiste à utiliser une adresse IP routable (ou un nombre limité d'adresses IP) pour connecter l'ensemble des machines du réseau en réalisant, au niveau de la passerelle de connexion à Internet, une translation (littéralement une “traduction”) entre l'adresse interne (non routable) de la machine souhaitant se connecter et l'adresse IP de la passerelle.

Exemple de NAT simple vers un serveur web



D'autre part, le mécanisme de translation d'adresses permet de sécuriser le réseau interne car il masque complètement l'adressage interne. Ainsi, pour un observateur externe au réseau, toutes les requêtes semblent provenir de la même adresse IP.

Translation statique

Le principe du NAT statique consiste à associer une adresse IP publique à une adresse IP privée interne au réseau. Le routeur (ou plus exactement la passerelle) permet donc d'associer à une adresse IP privée (par exemple 192.168.0.1) une adresse IP publique routable sur Internet et de faire la traduction, dans un sens comme dans l'autre, en modifiant l'adresse dans le paquet IP.

La translation d'adresses statiques permet ainsi de connecter des machines du réseau interne à Internet de manière transparente mais ne résout pas le problème de la pénurie d'adresses dans la mesure où n adresses IP routables sont nécessaires pour connecter n machines du réseau interne.

Translation dynamique

Le NAT dynamique permet de partager une adresse IP routable (ou un nombre réduit d'adresses IP routables) entre plusieurs machines en adressage privé. Ainsi, toutes les machines du réseau interne possèdent virtuellement, vu de l'extérieur, la même adresse IP. C'est la raison pour laquelle le terme de “mascarade IP” (en anglais IP “masquerading”) est parfois utilisé pour désigner le mécanisme de translation d'adresse dynamique.

Afin de pouvoir “multiplexer” (partager) les différentes adresses IP sur une ou plusieurs adresses IP routables, le NAT dynamique utilise le mécanisme de translation de port (PAT - “Port Address Translation”), c'est-à-dire l'affectation d'un port source différent à chaque requête de manière à pouvoir maintenir une correspondance entre les requêtes provenant du réseau interne et les réponses des machines sur Internet, toutes adressées à l'adresse IP du routeur.

Les translations d’adresses et la ToIP

Les translations d’adresses restent une difficulté importante dans les solutions implémentant la ToIP et le chiffrement. En effet, le chiffrement d’un paquet cache l’ensemble des données d’origine pour ne traiter que le transport du paquet au travers d’un réseau. Les informations portant sur un traitement particulier du routage et de la signalisation ne peuvent donc pas être exploitées et la qualité de service notamment peut en être fortement impactée. On privilégiera donc des périphériques “ToIP aware” tout au long de la chaîne de communication pour contourner cette limitation.

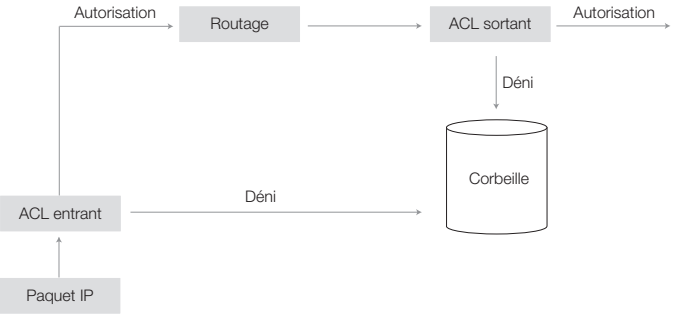
Access Control Lists

Les ACL (Access Control Lists) permettent de filtrer des paquets suivant des critères définis par l'utilisateur. Sur des paquets IP, il est ainsi possible de filtrer les paquets entrants ou sortant d'un routeur en fonction de l'IP source, de l'IP destination, des ports ...

Il existe 2 types d'ACL :

> Standard :
uniquement sur les IP sources,

> Etendue :
sur quasiment tous les champs des en-têtes IP, TCP et UDP.



Il est possible de résumer le fonctionnement des ACL de la façon suivante :

> le paquet est vérifié par rapport au 1er critère défini,

> s'il vérifie le critère, l'action définie est appliquée,

> sinon le paquet est comparé successivement par rapport aux ACL suivantes,

> s'il ne satisfait aucun critère, l'action déni en anglais (“deny”) est appliquée.

Exemple d’ACL:

```
interface Ethernet0
  ip address 172.16.1.1 255.255.255.0
  ip acces-group 1 out
!
access-list 1 deny 172.16.3.10 0.0.0.0
access-list 1 permit 0.0.0.0. 255.255.255.255
```


access-list 1 deny 172.16.3.10 0.0.0.0

> refuse les paquets d'IP source 172.16.3.10,

> le masque (également appelé “wildcard mask”) signifie ici que tous les bits de l'adresse IP sont significatifs

access-list 1 permit 0.0.0.0 255.255.255.255

> tous les paquets IP sont autorisés,

> le masque 255.255.255.255 signifie qu'aucun bit n'est significatif.

La création, la mise à jour, le “troubleshooting” nécessitent beaucoup de temps et de rigueur dans la syntaxe. Il est donc conseillé de :

> créer les ACL à l'aide d'un éditeur de texte et de faire un copier/coller dans la configuration du routeur,

> placer les extended ACL au plus près de la source du paquet pour le détruire le plus vite possible,

> placer les ACL standards au plus près de la destination au risque de détruire un paquet trop tôt,

> placer la règle la plus spécifique en premier,

> désactiver l’ACL sur l'interface concerné, avant de faire le moindre changement sur celle-ci.

Par ailleurs, on notera qu’une ACL :

> impose une utilisation beaucoup plus intensive des ressources du routeur où elle est implémentée (elle doit examiner chaque paquet pour vérifier s’il correspond ou non aux règles configurées),

> ne supporte en aucun cas la gestion des ports dynamiques,

> ne remonte pas dans les couches applicatives.

Chiffrement

RTP (Real-time Transport Protocol) encode, transporte et décode la voix. La qualité du flux étant essentielle dans le domaine de la ToIP, tant sur le plan de la vitesse que sur la qualité, le compromis sécurité/utilisation est donc essentiel : débit, qualité de la voix, temps d'établissement des communications, etc.

Certaines solutions classiques ne sont donc pas viables et le meilleur compromis se détermine au cas par cas selon le nombre de clients, les débits souhaités, le niveau de sécurité requis, la vitesse du média utilisé, les types de données…

Les contraintes d'intégrité, de confidentialité et d'authenticité ne tenaient pas une place de choix dans les premières solutions et les protocoles clés ne disposaient d'aucune protection fiable (SIP, RTP, RTCP ...).

Plusieurs protocoles sont apparus avec notamment les équivalents chiffrés de RTP et RTCP : SRTP et SRTCP (respectivement Secure Real-time Transport Protocol et Secure Real-time Transport Control Protocol). Les paquets SRTP se différencient des paquets RTP par 3 champs :

> un champ additionnel pour le type d'algorithme utilisé (“Authentication tag”),

> un deuxième contenant différentes informations sur la clé (“Master Key Identifier” -MKI),

> et bien sûr un payload chiffré.

SRTP et SRTCP ont été développés dans un souci de performance, le but étant de sécuriser au maximum les échanges à moindre coût en minimisant la surcharge liée au chiffrement du “payload”.

```
…
Synchronization Source (SSRC) identifier
Contributing Source (CSRC)
RTP extension
PAYLOAD (chiffré)
MKI
Authentication tag
```


Comme le montre la figure ci-dessus, seul le “payload” est chiffré dans un paquet SRTP, ce qui ne permet donc pas d'assurer à 100% l'intégrité des paquets transmis : l'en-tête du paquet SRTP peut être modifié, tout comme les champs optionnels MKI ou “Authentication tag”.

Quand SRTP est utilisé conjointement avec SIP, le déroulement chronologique des transactions est le suivant :

> appelant : établissement de la communication (sonnerie) avec un paquet.

SIP INVITE :

> appelé :
accord du tiers distant avec une réponse 200 (OK),

> appelant :
paquet SIP de type ACK pour confirmer la réception du paquet précédent et établir la session SIP.

Le schéma d'établissement ressemble étrangement à une ouverture de session TCP… Une fois la communication établie, si les deux participants se sont préalablement mis d'accord sur l'algorithme de chiffrement utilisé, la communication sécurisée est établie.

Dans le cas contraire, si une négociation des clés est nécessaire, un protocole de gestion des clés s'impose sur le même principe qu’IKE (pour plus de détails sur ce protocole, voir p.49) pour IPSec par exemple.

C'est dans ce but que MiKEY (“Multimédia Internet Keyring”) a été développé : il s'agit d'un protocole récent encore en test et très peu implémenté. MiKEY est encapsulé dans les paquets SIP et permet d'utiliser :

> un secret commun (PSK pour “Pre-Shared Key”), généralement sous forme de mot de passe,

> des protocoles Diffie-Hellman d’échange de clés,

> une PKI.

Cette dernière alternative n'a pas encore été implémentée et ses performances globales sont très controversées à l’heure actuelle.

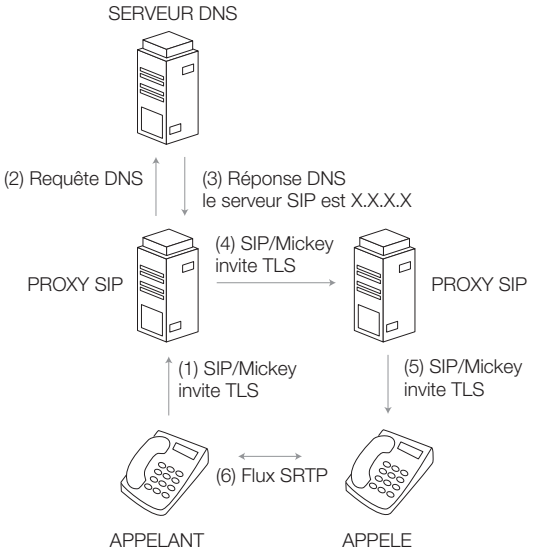
MiKEY, tout comme SRTP/SRTCP, tente de minimiser les coûts et les impacts de la protection. Il doit assurer une sécurité optimale des transactions de clés sans affecter de façon significative la rapidité des échanges.

MiKEY s'encapsule dans SIP avec le champ a=key-mgmy qui permet d'assurer l'authentification qui permettra de faire transiter un flux SRTP par la suite avec des algorithmes et des clés adéquats.

MiKEY repose sur SIP : toute attaque sur ce dernier remet donc en cause la sécurité.

Il est indispensable de veiller à l'intégrité et l'authenticité des paquets SIP. La nécessité pour certains intermédiaires d'accéder, voire de modifier des portions de paquets (proxies, registrars, redirecteurs...) rendent la tâche délicate. Minisip(entre autre) propose l'utilisation de TLS pour limiter ces risques.

Mise en place d'un flux sécurisé



La ToIP faisant appel à de nombreux processus (SIP, DNS, SRTP, voire SRTCP pour le contrôle du flux SRTP : CODECs, timing, etc.), il est indispensable de sécuriser chaque étape de la communication. Les nombreuses contraintes imposées par cette technologie ne facilitent pas la tâche : il est aussi parfois nécessaire de traverser des pare-feux, de fonctionner avec des translations d'adresses (NAT) et certains choix sont alors limités.

Tunnel IPsec

Les différentes solutions de “tunneling” présentent toutes des avantages et des inconvénients pour la ToIP avec ses nombreuses exigences. IPSec, qui travaille sur la couche réseau, permet d'assurer une plus grande fiabilité des informations. Notons par exemple que le problème des en-têtes SRTP modifiables n'est plus un souci ici.

Cependant, le coût de cette solution est parfois considérable, tant sur le plan des ressources matérielles que sur le trafic réseau. IKE (Internet Key Exchange) permet alors de remplacer MiKEY et d'assurer la gestion des clés pour l'ensemble des communications ToIP.

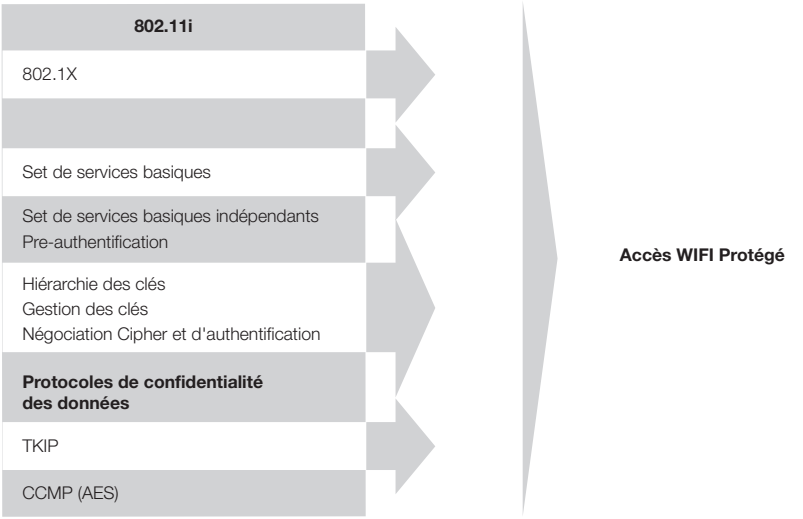
La surcharge engendrée par IPSec peut être minimisée en configurant le tunnel pour traiter uniquement les flux de téléphonie sur IP (pour des machines/protocoles fixés). Un atout intéressant est la possibilité d'utiliser la totalité des softphones disponibles puisqu'ils n'ont plus à gérer la sécurité des échanges (via SRTP/MiKEY...).

UDP limite les types de tunnels utilisables, notamment SSL ou SSH, même s'il reste possible d'utiliser vtun (ou un équivalent) pour faire de l'UDP over TCP, mais les performances deviendraient rapidement médiocres !

Pour résumer : les tunnels simplifient le déploiement de la ToIP sécurisée, mais ne peuvent pas être employés sur de larges infrastructures.

Et la ToIP sur WIFI ?

Concernant les réseaux WIFI, la norme sécurisée 802.11i récemment homologuée par l'IEEE augmente la sécurité du WIFI en introduisant l'algorithme AES (Advanced Encryption Standard). L'AES est une technique de chiffrement fort à clé symétrique, où la clé est la même pour le chiffrement et le déchiffrement. Les longueurs de clé de chiffrement utilisées sont 128, 192 ou 256 bits. L'AES, créé en 1998, est peu utilisé par le grand public. Plusieurs gouvernements l'ont homologué pour usage administratif, y compris pour le chiffrement des données les plus sensibles.



Un exemple de framework sécurisé basé sur 802.11i (v2 au firmware de l'IP Phone 7920 du constructeur Cisco).

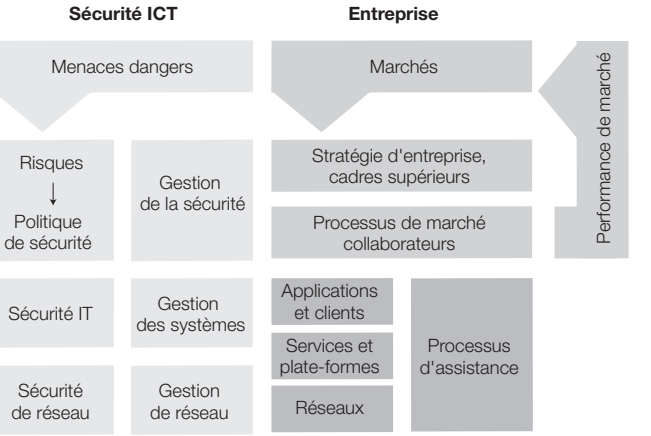
Un projet ToIP est l'occasion de sensibiliser les collaborateurs à la sécurité de l'ensemble de l'infrastructure de l'entreprise. En effet, qui aujourd'hui se soucie de la sécurité de ses e-mails ?

La sécurité comme une mission complète

La sécurité ne se limite pas uniquement à de simples mesures techniques, elle nécessite plutôt un concept intégral (voir schéma ci-dessous). Il s'agit ici d'un "framework" uniforme qui, entre autre, octroie une responsabilité de sécurité claire, stratégique et opérative. Les menaces doivent être évaluées en fonction de l'entreprise et les risques mesurables doivent être définis. On en déduit des concepts de sécurité pour des cas normaux et des situations d'urgence, qui sont mis en oeuvre sur le plan technique à l'aide de solutions de sécurité. Si les collaborateurs ne sont pas impliqués, ces solutions ne sont cependant qu'à moitié efficaces. De cette façon, la sécurité pénètre finalement les processus d'entreprise.

Un haut niveau de sécurité peut être atteint aujourd'hui en impliquant aussi les composants de sécurité déjà utilisés pour le réseau de données. Les fabricants et les comités compétents travaillent aussi à améliorer les technologies (matérielles, logicielles) et les protocoles les plus souvent utilisés.

Les risques et les frais qu'ils peuvent entraîner ne sont pas une raison pour renoncer aux avantages de la ToIP : un coût avantageux sur le long terme, des gains de flexibilité et de productivité grâce à des applications de communication intégrées. En outre, les efforts de sécurité se répercutent sur l'ensemble de l'infrastructure : l'idée d'être espionné lors d'une discussion effraie et sensibilise les collaborateurs sûrement plus que le souci de la sécurité de leurs e-mails.



politique de sécurité

Lors de la configuration d'un réseau, qu'il s'agisse d'un réseau local (LAN), d'un réseau local virtuel (VLAN), ou d'un réseau étendu (WAN), il est important de définir dès le début les politiques de sécurité. Les politiques de sécurité sont des règles programmées et stockées dans un dispositif de sécurité, destinées à contrôler des aspects comme les droits d'accès. Ces politiques de sécurité sont, bien sûr, également des règlements écrits régissant le fonctionnement d'une société. De plus, les sociétés doivent désigner le responsable de l'application et de la gestion de ces politiques et déterminer le mode d'information des employés à propos des règles et des protections.

Une politique de sécurité repose essentiellement sur quatre piliers :

> décrire clairement votre modèle métier.

Il serait absurde de concevoir ou de déployer une solution de sécurité qui ne serait pas fondée sur la nature de vos objectifs métier. Identifiez clairement vos objectifs métier en y incluant le type de services et d'accès qui vous sont nécessaires pour les atteindre.

A mesure que vous ajoutez de nouveaux systèmes ou de nouveaux services à votre réseau, vous introduisez de nouveaux risques. Les procédures régulières d'administration du réseau doivent comprendre une évaluation régulière et complète des faiblesses du système.

> identifier en détail les risques associés.

Si vous prévoyez d'héberger un segment de services au public (encore appelé zone démilitarisée ou DMZ) et d'offrir des activités de commerce électronique, vous devez comprendre toutes les manières dont les pirates chercheront à exploiter vos systèmes et vos services. Quels sont les risques si la page Web est saccagée, si un pirate s'introduit sur un serveur ou si une base de données de clients est attaquée ? De quelle manière ces attaques sont-elles menées ? Les pirates contournent-ils le pare-feu en se cachant dans le trafic Web autorisé ou cherchent-ils à exploiter des vulnérabilités dans des systèmes d'exploitation mal mis à jour ? Ces questions doivent être soigneusement examinées et comprises avant de pouvoir passer à l'étape suivante. A mesure que vous ajoutez de nouveaux systèmes ou de nouveaux services à votre réseau, vous introduisez de nouveaux risques. Les procédures régulières d'administration du réseau doivent comprendre une évaluation régulière et complète des faiblesses du système.

La mise en œuvre des solutions de sécurité repose sur les trois "P" : les Personnes, les Produits et les Procédures.

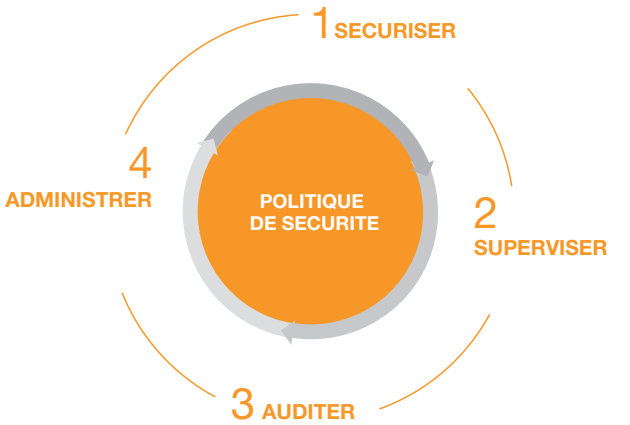
> adopter une démarche systématique de limitation de ces risques.

Tout dans un réseau peut constituer une cible, qu'il s'agisse des routeurs, des commutateurs, des hôtes, des applications, des réseaux et des systèmes d'exploitation. Pour être efficace, une politique de sécurité doit tenir compte de chacune de ces composantes. La mise en oeuvre des solutions de sécurité repose sur les trois "P" : les Personnes, les Produits et les Procédures. Vous devez disposer de techniciens compétents pour mettre en oeuvre votre politique, vous devez utiliser des outils spécifiquement conçus pour supporter votre stratégie e-business et vous devez associer à tout cela une administration système et une politique d'analyse efficaces.

> garder à l'esprit que la sécurité est un processus.

Une politique de sécurité n'est pas une solution "gravée dans le marbre". La sécurité exige des études, des analyses et des améliorations régulières pour offrir le niveau de protection dont votre entreprise a besoin. Vous pouvez également envisager l'acquisition d'un utilitaire d'évaluation des vulnérabilités ou encore signer un contrat avec un partenaire extérieur de contrôle de la sécurité afin de vérifier votre politique, vos procédures et votre mise en œuvre et, dans certains cas, vous décharger de certaines tâches à forte composante en main-d'oeuvre comme la surveillance.

Le cycle vertueux de la sécurité



SECURISER
> identifier et authentifier
> filtrer
> chiffrement et VPN

SUPERVISER
> détection d'intrusions /contre mesures
> détection en mode "content base"/contre mesures
> supervision du personnel

AUDITER
> valider la politique de sécurité
> rechercher les vulnérabilités
> audit d'application/patch

ADMINISTRATION
> analyse des logs et génération de rapports
> gestion des configurations
> sécurisation des modes d'administration

conclusion

Les sociétés déployant ou envisageant de déployer des solutions de téléphonie sur leurs réseaux auront un effort à fournir encore plus important que par le passé pour en sécuriser l'ensemble. La sécurité coûtera plus cher et demandera un personnel qualifié. Par ailleurs, le développement juridique de ce domaine traité via Sarbanes-Oxley, GBLA et CALEA pour les Etats-Unis ou le DEPC en Europe transforme les tentatives de pénétrations et autres actes malveillants en actes criminels. L'air du temps est donc à la professionnalisation poussée de ce domaine, les conséquences pouvant être extrêmement lourdes en cas de faute.

La convergence de la voix et de donnée sur les mêmes infrastructures physiques doit permettre d'alléger les charges d'exploitation et de s'ouvrir à de nouveaux applicatifs utilisant une nouvelle palette de possibilités. Cependant, avec la mise en place de cette nouvelle architecture, la voix devient aussi vulnérable que les applications classiques aux failles de sécurité sur le réseau, posant ainsi une nouvelle problématique : si les risques deviennent les mêmes, la tolérance aux pannes des utilisateurs en matière de téléphonie reste extrêmement faible en comparaison des applications data. Un niveau de service identique à la téléphonie traditionnelle est attendu.

La plupart des sociétés possèdent aujourd'hui des périmètres de sécurité bien établis. Cependant, en observant en détail les solutions mises en place, il apparaît rapidement que tout n'est pas coordonné et que certaines faiblesses persistent. Idéalement, une entreprise devrait posséder plusieurs niveaux de défense coordonnés sur l'ensemble d'un périmètre de sécurité défini, confrontant ainsi un pirate ayant réussi à passer un premier niveau avec un deuxième, puis un troisième. De cette manière, un accès rapide aux informations sensibles de l'entreprise doit pouvoir être évité.

Par ailleurs, le postulat indiquant que tout utilisateur authentifié et autorisé est supposé digne de confiance et peut posséder un accès important aux ressources du réseau n'est pas suffisamment remis en question au travers du nouveau masque de lecture imposé par la convergence des applications et ressources. Ce manque de granularité peut créer des risques ouvrant des ouvertures à toute personne malintentionnée ou encore permettre la propagation de vers ou de virus sur le réseau.

Enfin, n'oublions pas que si les risques d'attaques sont souvent associés à l'image du "hacker" indépendant, ayant une grande maîtrise des technologies, une motivation à toute épreuve et utilisant Internet pour accéder aux infrastructures de l'entreprise, la réalité est tout autre. Les risques et les pertes financières pour l'entreprise sont souvent causés par les utilisateurs eux-mêmes.

Nous n'insisterons jamais assez sur ce point, d'autant que la convergence des applications démultiplie les risques de pertes financières à la suite d'une interruption du service. Les points clés suivants, nécessaires à la sécurisation du réseau, doivent absolument être étudiés :

> le renforcement des politiques de sécurité,

> la communication autour des politiques de sécurité mises en place,

> la mise en place d'une sécurité physique rigoureuse,

> la vérification des accréditations des utilisateurs,

> la supervision des installations (Logs, firewall, IDS/IPS, etc.),

> la séparation des flux,

> le renforcement logique des serveurs,

> la mise en place de chiffrement dès que cela est envisageable.

Orange Business Services offre aux entreprises des solutions et services pour les aider à gérer leurs enjeux de sécurité. Qu'il s'agisse de les conseiller sur les solutions à mettre en œuvre ou de leur apporter le soutien nécessaire pour les mettre à la disposition des utilisateurs, les entreprises bénéficient d'un accompagnement personnalisé.

En outre, Orange Business Services qui maîtrise l'ensemble des technologies touchant à la sécurité, propose un haut niveau de sécurité embarquée dans l'ensemble de ses solutions et services.

glossaire

Algorithme RSA

RSA (de ses concepteurs Rivest, Shamir et Adleman) est un algorithme asymétrique de cryptographie à clé publique, très utilisé dans le commerce électronique, et plus généralement pour échanger des données confidentielles sur Internet.

Call hijacking

Détournement d'appel

DoS Attack

Littéralement, Attaque par Déni de Service. Le but est de rendre un service (la téléphonie par exemple) totalement inutilisable pour perturber l'activité d'une société.

EAP

Le protocole EAP (Extensible Authentication Protocol) est une norme IETF (Internet Engineering Task Force) décrite dans le document HYPERLINK RFC 3748 ("http://www.ietf.org/rfc/rfc3748.txt"), qui définit une infrastructure permettant aux clients d'accès réseau et aux serveurs d'authentification d'héberger des modules pour les méthodes et technologies d'authentification actuelles et futures.

Microsoft® Windows® utilise EAP pour authentifier l'accès réseau pour les connexions PPP (Point-to-Point Protocol) -accès distant et réseau privé virtuel- et pour l'accès réseau basé sur IEEE 802.1X aux commutateurs Ethernet et points d'accès sans fil.

Softphone

Un Softphone est un logiciel que l'on utilise pour faire de la téléphonie sur Internet depuis son ordinateur. Les interfaces de ces Softphones sont souvent simples d'utilisation et très omplètes puisque toutes les fonctionnalités qui existent sur des téléphones classiques existent aussi sur les Softphones.

UDP (User Datagram Protocol)

Un des principaux protocoles de télécommunication utilisé par Internet. Il fait partie de la couche transport de la pile de protocole TCP/IP. Le rôle de ce protocole est de permettre la transmission de paquets (aussi appelés datagrammes) de manière très simple entre deux entités, chacune étant définie par une adresse IP et un numéro de port. Ce protocole qui fonctionne en mode non-connecté est souvent décrit comme étant un protocole non-fiable par opposition au protocole TCP qui fonctionne en mode connecté (ouverture de connexion, transfert de données, fermeture de connexion).

La téléphonie sur IP coexiste aujourd'hui dans deux environnements très différents : Internet et les réseaux d'entreprises. Si dans le premier cas la ToIP est difficile à sécuriser, le réseau d'entreprise en revanche, beaucoup moins ouvert sur l'extérieur et bien mieux maîtrisé, permet l'implémentation de solutions à même de renforcer la sécurité de la téléphonie sur IP de façon significative.

Il est indéniable que le changement de média véhiculant la voix comporte des risques que l'on peut retrouver sur les applications informatiques traditionnelles. Pour autant, la tolérance aux pannes devra être nécessairement diminuée, compte tenu du niveau de service attendu de la téléphonie traditionnelle.

Pour sécuriser une solution de ToIP en entreprise, il sera donc nécessaire :

> d'implémenter des mécanismes de sécurité renforcés par rapport au monde informatique,

> de prendre en compte le caractère temps réel de cette nouvelle application par l'implémentation de mécanismes de qualité de service sur l'ensemble de l'environnement,

> d'intégrer cette nouvelle application dans les différentes politiques de sécurité de l'entreprise,

> de ne pas oublier les leçons du passé (les défauts de configuration des IP PBX et des PBX classiques en matière de téléphonie restant identiques).

Dans ce troisième cahier de la Sécurité, notre expert présente de façon détaillée et pédagogique les risques d'attaques de solutions ToIP en entreprise puis propose une analyse éclairée des différentes solutions existantes pour sécuriser une solution de ToIP, lesquelles ne sont pas que techniques.

Les cahiers de la Sécurité déjà parus :

OTP, une solution d'authentification forte

PKI, une solution de cryptage

France Télécom - 6 place d'Alleray 75505 Paris Cedex 15 – SA au capital de 10 426 692 520 euros - 380 129 866 RCS Paris - document non contractuel - code EAN 3699690005017
www.uniteam.fr - 05/2007

Téléchargez les cahiers de la sécurité sur www.orange-business.com

Business Services

orange™