MISC N° 16

Sécurité de la Voix sur IP

Pierre BETOUIN - École supérieure d'informatique, d'électronique et d'automatique (ESIEA)
pierre.betouin@security-labs.org

Patrick CHAMBET - Consultant Senior, EdelWeb - Groupe ON-X
http://www.edelweb.fr - http://www.chambet.com - patrick@chambet.com

Nicolas FISCHBACH, Senior Manager, Network Engineering Security, COLT Telecom
nico@securite.org - http://www.securite.org

Introduction

La voix sur IP est déjà, ou va devenir, un projet stratégique en 2005 pour bon nombre d'entreprises et d'opérateurs. Pour l'utilisateur elle permet uniquement de téléphoner à moindre coût via l'Internet, la voix restant une forme de communication bien plus conviviale que le courrier électronique ou les formes de messageries instantanées. Pour l'entreprise et les opérateurs ce facteur « coût de la communication » est important, mais le déploiement de réseaux privés virtuels MPLS, l'introduction de la qualité de service dans les réseaux (QoS), la convergence voix-données (CTI), les divers projets de consolidation des deux dernières années, l'arrivée des auto-commutateurs IP, la disponibilité de postes téléphoniques intégrant des fonctionnalités de plus en plus avancées sont des facteurs tout au moins aussi déterminants. Des études récentes montrent que la sécurité de la VoIP est un élément clé pour les décideurs, mais les déploiements observés ont malheureusement tendance à montrer le contraire.

Après avoir présenté les principaux protocoles liés à la voix sur IP et avoir présenté les rudiments de la sécurisation des différents équipements nous allons détailler différentes attaques et quels « ajouts » sécurité peuvent limiter leur impact. Pour finir nous discuterons de l'interception de trafic et présenterons deux évolutions récentes de la VoIP.

Les différents protocoles

Nous allons nous intéresser tout particulièrement aux solutions de voix sur IP qui reposent sur les protocoles SIP [2] (signalisation et contrôle) et RTP [12] (transport de la voix). Il nous semble important de rappeler qu'une solution complète repose sur un large nombre de protocoles et que H.323, qui n'est ici que cité, est encore présent dans bon nombre de déploiements.

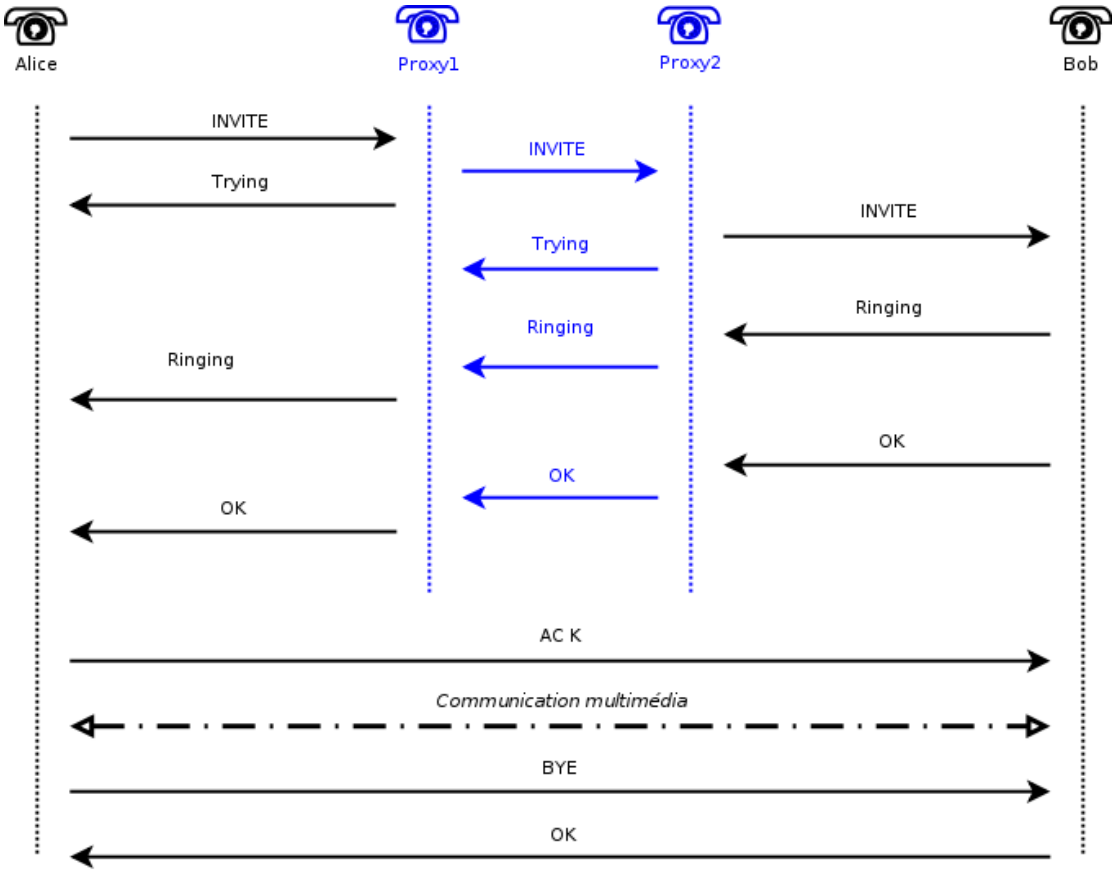
SIP

SIP (Session Initiation Protocol) est le standard IETF pour la signalisation (établissement, terminaison, redirection, relayage, etc) de communications multimédias interactives. Ce protocole est de nos jours celui qui est déployé couramment. Le format est proche d'une adresse de messagerie: sip:nico@securite.org avec une syntaxe proche de celle d'HTTP.

SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions) est une extension de SIP dont le but est de supporter les messageries instantanées.

Le projet PROTOS [16] s'est intéressé à SIP, et comme pour SNMP, a trouvé un nombre important d'implémentations qui ne passaient pas le "batch" de tests sans se planter ou redémarrer. Cela concerne aussi bien les téléphones que les relais SIP et les équipements de sécurité.

Le "SIP proxy" joue le rôle de relais et fait suivre une requête SIP au prochain serveur. Le "SIP redirect server" renvoie une réponse au client contenant le prochain serveur à contacter. Le "SIP registrar" enregistre le lien entre l'adresse IP et l'adresse SIP. Comme pour IPsec, des solutions alternatives ont dû être développées pour gérer les contraintes introduites par la traduction d'adresses (NAT).



Exemple de session SIP

Méthode	***Principe***
INVITE	Débute une communication SIP. Création du *Call-ID* et éventuellement du *tag* sur le champ *From*.

Méthode	*Principe*
REGISTER	Enregistre un utilisateur auprès d'un noeud VoIP (non utilisé en communication directe entre 2 clients).
ACK	Confirmation de l'établissement de la session SIP. *Call-ID* identique à celui du paquet INVITE associé. Ajout éventuel du *tag* sur le champ *To* du paquet INVITE relatif.
BYE	Met fin à la communication. Même *Call-ID* que les paquets précédents. Idem si utilisation des *tags*.
CANCEL	Annule un SIP INVITE (appel). Même *Call-ID* et éventuellement *tag* du champ *From* du SIP INVITE.

Principales méthodes SIP

H.323

H.323 [14] a été le premier protocole développé pour permettre des communications multimédias. SIP est son "concurrent". H.323 est relativement complexe et SIP tente de simplifier les échanges en utilisant une sémantique proche de HTTP.
H.235 [13] définit certains mécanismes de sécurité (authentification et chiffrement).

Les protocoles secondaires

Le service DNS est utilisé pour fournir des services d'annuaire et de localisation. TFTP et HTTP sont utilisés par les téléphones et différents autres éléments pour télécharger leur configuration. ENUM permet de lier des adresses SIP via DNS aux numéros de téléphone au format E.164.

Sécurité des équipements

Les offres commerciales de VoIP nécessitent des serveurs et des applicatifs pour fonctionner. Nous allons prendre l'exemple d'une infrastructure Cisco et détailler les besoins de sécurité sur les différents équipements.

Serveurs Windows

Les applicatifs Cisco sont hébergés sur des serveurs Windows. Ces serveurs sont bien souvent installés par défaut et laissés en l'état. Il est donc nécessaire de commencer par sécuriser l'OS de ces serveurs (voir MISC No 2 et [9]).
Les points suivants sont particulièrement à surveiller, notamment sur les serveurs pré-Windows 2003 :

- filtrer les ports accessibles sur les serveurs depuis le réseau des utilisateurs, au niveau des routeurs ;
- désinstaller ou stopper les services inutiles ;
- mettre à jour le serveur avec les derniers correctifs de sécurité ;
- interdire les connexions NetBIOS anonymes ;
- appliquer des permissions d'accès strictes sur le système de fichiers et la base de registre ;
- activer l'audit de sécurité et configurer des journaux de grande taille ;
- installer IIS (nécessaire pour le Call Manager par exemple) de manière sécurisée (voir MISC No 1 et [10]).

L'application Cisco Call Manager utilisant également une base de données SQL Server, il est très important de sécuriser particulièrement celui-ci. Pour cela, les recommandations principales sont les suivantes :

- appliquer les derniers correctifs de sécurité de SQL Server ;
- faire tourner les services de SQL Server sous des comptes spécifiques, non administrateurs ;
- attribuer un mot de passe robuste au compte 'sa' ;
- supprimer les bases par défaut ;
- si possible (en fonction des applications), créer des utilisateurs spéciaux, voire nominatifs et leur attribuer des mots de passe non triviaux ;
- appliquer des permissions d'accès sur les objets de SQL Server (tables, procédures stockées, etc…) en fonction des utilisateurs ;
- supprimer les procédures stockées par défaut si elles ne sont pas utilisées.

A noter que, sans sécurisation particulière, il est possible d'effectuer un transfert de zone DNS du domaine interne utilisé par défaut (avvid.com). On y trouve également la liste des services et des ports utilisés par les applicatifs de gestion VoIP :

```
 [[123.123.123.123]]
 avvid.com.    SOA     unity01.avvid.com admin. (26 900 600 86400 3600)
 avvid.com.    A       123.123.123.123
 avvid.com.    NS      unity01.avvid.com
 _kerberos._tcp.default-first-site-name._sites.dc._msdcs
               SRV     priority=0, unity01.avvid.com
```

[…]

Il convient donc d'interdire le transfert de zone au niveau des serveurs DNS internes.

*Serveurs *nix*

En ce qui concerne la sécurisation des serveurs Unix, il est recommandé d'appliquer les recommandations générales suivantes, en fonction des applicatifs VoIP hébergés :

- minimisation du système (désinstallation des démons et serveurs inutiles) ;
- appliquer les derniers correctifs de sécurité de l'OS ;
- créer des comptes spécifiques et leur attribuer des mots de passe robustes ;
- chrooter (mise en "cage") les applicatifs et les faire tourner sous des comptes non privilégiés ;
- appliquer les derniers correctifs de sécurité des applicatifs ;
- activer les logs (journalisation) au maximum.

Des informations sur la sécurisation d'Asterisk par exemple, un PBX logiciel tournant sur Unix, se trouvent ici : [17].

Téléphones

Le téléphone peut se présenter sous deux formes : soit un téléphone "classique", soit un téléphone "logiciel" qui s'exécute sur l'ordinateur de l'utilisateur. En terminologie SIP c'est un UA (User Agent).

Au niveau des téléphones IP, l'accès à la partie protégée par HTTP Basic Authentication se fait en clair :

```
http://IP-Phone/CGI/
```

Si le mot de passe est le même sur tous les téléphones, n'importe qui peut exécuter des commandes de configuration sur l'ensemble des téléphones.

Il est bien sûr recommandé de mettre à jour le *firmware* des téléphones IP, et de désactiver les interfaces de gestion du type HTTP et telnet des téléphones IP (cela peut être fait depuis l'application Cisco Call Manager, que nous étudierons plus loin).

Signalons qu'il est possible d'effectuer un déni de service sur un téléphone IP par l'intermédiaire de l'application Cisco Call Manager 3.x. En effet, celle-ci gère le profil des utilisateurs, qui sont téléchargés sur les téléphones IP lorsque les utilisateurs se loguent dessus. En dépassant largement la limite de 99 services téléphoniques IP par téléphone (en développant par exemple un script qui crée plus de 1000 services dans l'application Call Manager), le profil de l'utilisateur devient inutilisable sur un téléphone IP : on ne peut plus ni se loguer sur un téléphone ni se déloguer, et celui-ci devient inutilisable (plus de tonalité). Il est nécessaire de l'éteindre et de le rallumer, après avoir corrigé le profil utilisateur dans le Call Manager.

Autres équipements

Certains constructeurs proposent des équipements propriétaires pour gérer la VoIP. Cisco propose par exemple les passerelles voix-données VG 200 et VG 248, fonctionnant sous IOS, le système d'exploitation des routeurs Cisco.
Il est fréquent de rencontrer le service telnet ouvert sur de tels équipements. Il est donc fortement recommandé de ne pas laisser l'interface d'administration des équipements accessible depuis l'ensemble du réseau local.
De plus, étant donné les possibilités d'interception des connexions, même à travers le switch (cf. plus loin dans cet article), il est fortement déconseillé d'utiliser un protocole d'administration non chiffré.

Nous vous recommandons d'activer les fonctionnalités de sécurité présentes dans IOS et CatOS (voir MISC 1 ou [15]).

Applicatifs

Les applications de gestion du réseau VoIP sont bien souvent des applications Web, et, comme telles, elles sont sensibles aux attaques classiques contre ce type d'applications (voir Linux Mag HS No 12 et [11]). Il est donc nécessaire de prendre un soin particulier à leur sécurisation, d'autant plus que certaines d'entre elles ont besoin d'être accessibles depuis une grande partie du réseau local, voire depuis l'Internet !

Le Call Manager, par exemple, fournit des fonctions de base de gestion des appels, des utilisateurs et des configurations, mais également des fonctionnalités avancées comme la conférence, les boîtes vocales, etc. Il peut être vu comme un IP PBX. A ne pas confondre avec des PBX traditionnels (pas de VoIP) que l'on peut administrer à distance via une connexion TCP/IP (qui remplace la connexion locale ou de télé-maintenance via un modem). L'application Unity, quant à elle, est l'applicatif permettant d'avoir une messagerie vocale VoIP.

L'authentification lors des accès au Call Manager 3.x se fait par l'intermédiaire de formulaires HTML :

```
http://CallMgr/CCMUser/logon.asp
http://CallMgr/CCMAdmin/logon.asp
http://CallMgr/CCMCIP/authenticate.asp
http://CallMgr/ma/desktop/maLogin.jsp
```

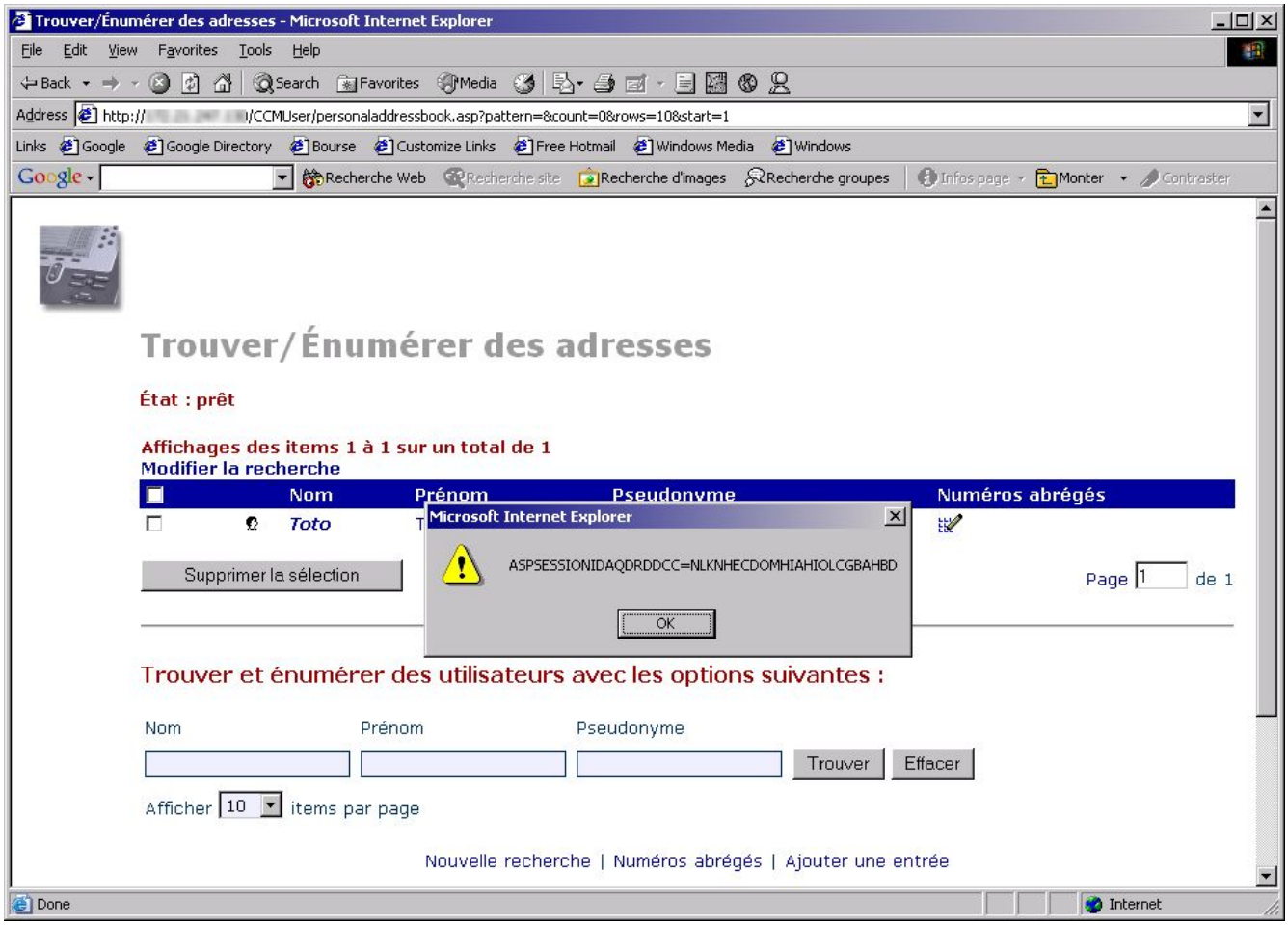
De même, l'authentification pour l'accès à l'application Cisco IP Manager Assistant (IPMA) se fait à l'aide d'un formulaire HTML et d'une applet Java.

Dans tous les cas (accès utilisateurs et accès administrateur), le trafic réseau transite en clair. Il est facile, par une attaque ARP sur les commutateurs (voir plus loin), de capturer le trafic afin de récupérer le mot de passe de l'administrateur lorsqu'il se connecte. Il est donc indispensable de forcer l'utilisation de SSL pour accéder aux pages HTML d'authentification applicative.

D'autre part, les données saisies par l'utilisateur dans les formulaires HTML de l'application ne sont contrôlées que côté client (par des scripts JavaScript), ce qui est l'erreur classique : elles ne sont pas contrôlées à nouveau côté serveur. Il est donc facile pour un attaquant d'outrepasser les contrôles afin d'entrer des données dangereuses dans l'application (code HTML, scripts, injection SQL, etc…).

Des applications directes de ce type de vulnérabilité sont par exemple les suivantes :

- Déni de service sur les téléphones IP (voir plus haut) ;
- Cross Site Scripting (XSS) : voir figure ci-dessous.



L'exemple ci-dessus montre qu'il est possible de récupérer le cookie d'authentification d'un utilisateur. Ce cookie est suffisant pour accéder à l'application à sa place.

D'autres recommandations, plus classiques, sont également à prendre en compte lors de l'installation des applications de gestion VoIP :

- configurer le Call Manager, IIS et/ou TomCat afin de ne pas afficher des messages d'erreur détaillés pouvant fournir des informations précieuses à un attaquant ;
- supprimer les répertoires par défaut et tous les répertoires et fichiers inutiles ;
- installer un filtre d'URLs (de type URLScan) afin d'interdire les attaques classiques sur les URLs et les caractères spéciaux ;
- etc (voir [11]).

Il faut également savoir que l'accès à l'annuaire d'entreprise est en libre accès par défaut, aussi bien à partir des téléphones IP que de tout navigateur Web : il est possible de faire des recherches et de consulter l'ensemble de l'annuaire à l'aide de requêtes de la forme :

```
http://CallMgr/CCMCIP/xmldirectorylist.asp?l=A&f=&start=1
```

Le résultat donne les informations suivantes :

```
<CiscoIPPhoneDirectory>
<Prompt>Records 1 to 32 of 1254</Prompt>
…
<DirectoryEntry>
<Name>Glloq, Alain</Name>
<Telephone>12345</Telephone>
</DirectoryEntry>
[...]
</CiscoIPPhoneDirectory>
```

De même, un accès LDAP anonyme aux serveurs applicatifs révèle le contenu de l'annuaire :

```
ld = ldap_open("123.123.123.123", 389);
Expanding base 'DC=avvid,DC=com'...
Result <0>: (null)
Matched DNs:
Getting 1 entries:
>> Dn: DC=avvid; DC=com;
       1> masteredBy: CN=NTDS Settings,CN=UNITY01,CN=Servers,CN=Default-
First-Site-Name, CN=Sites,CN=Configuration,DC=avvid,DC=com;
[…]
       1> fSMORoleOwner: CN=NTDS Settings,CN=UNITY01,CN=Servers,CN=Default-
First-Site-Name, CN=Sites,CN=Configuration,DC=avvid,DC=com;
       1> gPLink: [LDAP://CN={31B2F340-016D-11D2-945F-
00C04FB984F9},CN=Policies,CN=System, DC=avvid,DC=com;0];
[…]
```

Attaques locales

Les attaques étudiées ici sont des attaques locales, au niveau du LAN principalement. Elles ont été testées sur des équipements Cisco IP Phone et leur infrastructure (équipements, serveurs et applicatifs).

Des vulnérabilités graves ont été mises en lumière et exploitées, allant d'un simple déni de service sur les équipements (téléphones et serveurs) au détournement de flux audio conduisant à l'écoute de n'importe quelle conversation sur le réseau, en passant par la récupération d'informations confidentielles et la facturation d'appels à d'autres personnes.

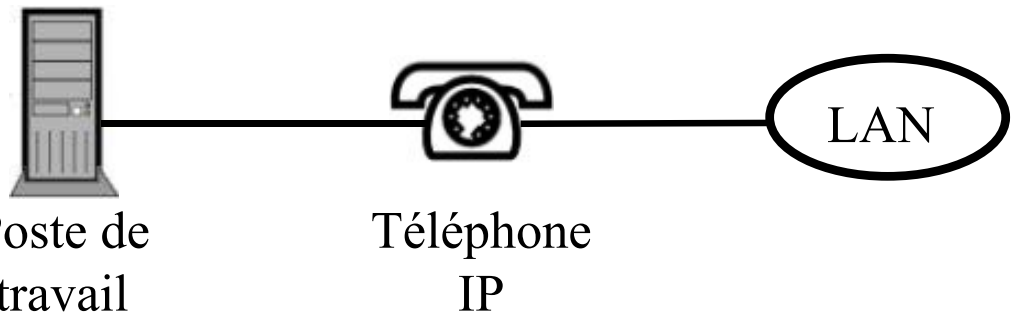
Utilisation frauduleuse des postes téléphoniques

Bien que la configuration locale des téléphones IP puisse être modifiée par l'utilisateur (en utilisant le code de déverrouillage '**#' par exemple sur les Cisco IP Phones), il est difficile d'effectuer une attaque à partir des postes téléphoniques IP. En effet, la plupart des paramètres ne sont pas modifiables localement. Seule une reconfiguration des interfaces du téléphone est possible.

Étanchéité des réseaux voix / données

L'IP Phone de Cisco, par exemple, intègre deux interfaces réseau et un switch interne, séparant le flux voix et le flux données dans deux VLANs différents. Dans l'utilisation normale d'un IP Phone Cisco, l'ordinateur de l'utilisateur est connecté sur

la prise de données de son IP Phone, celui-ci étant lui-même connecté au réseau local de l'entreprise :



Tentatives de franchissement données vers voix

Dans la configuration ci-dessus, le franchissement données vers voix est impossible, car l'IP Phone procède à un filtrage du VLAN voix. Cependant, pour accéder au flux voix, il suffit de brancher un ordinateur directement sur la prise murale, avec des outils appropriés pour se placer sur le VLAN voix (par exemple noyau Linux avec extension protocole 802.1Q).

Tentatives de franchissement voix vers données

Le franchissement voix vers données n'a pas véritablement de sens dans la mesure où le flux de données n'est pas protégé par un VLAN spécifique. On accède donc aux mêmes informations, que l'on soit branché d'un côté ou de l'autre de l'IP Phone.

Bien sûr, le franchissement voix vers données n'est pas possible depuis le seul poste téléphonique.

Attaque des flux VoIP

Le flux voix VoIP utilise deux types de canaux de communication, l'un de contrôle, l'autre de données.

Le canal de contrôle utilise le protocole Cisco Skinny sur les ports TCP 2000-2002. Chaque téléphone IP établit une connexion TCP avec le serveur hébergeant le Call Manager 3.x sur le port 2000. Cette connexion est gardée active grâce à l'émission de « KeepAlive » toutes les 30 secondes.

Les caractéristiques importantes d'un point de vue sécurité sont les suivantes :

- le canal de contrôle est connecté en permanence ;
- la connexion est initialisée par le poste IP-phone ;
- les actions Skinny sont transmises en clair.

Le canal de données sert à faire transiter les données audio des communications. Il utilise un dérivé du protocole RTSP : il s'agit de paquets UDP envoyés sur des ports dynamiquement négociés à travers le canal de contrôle.

Une communication classique comporte deux flux voix simultanés, ayant un débit de 50 paquets par seconde chacun. Chaque paquet est composé d'un en-tête de 54 octets (en-tête UDP + entête RTSP) puis de 160 octets de données audio encodées en u-law 8bits, 8kHz.

Les caractéristiques importantes d'un point de vue sécurité sont les suivantes :

- les ports UDP sont négociés dynamiquement au travers du canal de contrôle ;
- les paquets sont numérotés séquentiellement dans l'entête RTSP ;
- le flux audio est compressé mais non chiffré ;
- l'absence de flux est autorisée et est synonyme de silence dans la communication téléphonique.

Insertion de paquets dans un flux VoIP

Si on est en mesure d'intercepter le flux d'un téléphone IP, il est possible d'insérer des paquets dans le canal de contrôle ou dans le canal de données, avec des résultats très différents.
Les attaques suivantes ont été réellement effectuées au cours de tests d'intrusion.

Insertion d'un paquet remplaçant un autre paquet de même taille dans le canal de contrôle

Lorsqu'il est possible de prédire l'émission d'un évènement particulier dans le canal de contrôle par l'utilisateur d'un téléphone IP, il suffit d'envoyer, avec une très légère anticipation (quelques centièmes de seconde), un autre paquet qui sera pris en compte à la place du paquet prédit.

La substitution d'un paquet conduit aux résultats suivants selon les cas :

- simuler l'appui d'une touche du pavé numérique du téléphone, pour modifier à la volée le numéro composé par l'utilisateur du téléphone (ou bien l'enregistrer afin d'espionner les numéros composés), ou d'effectuer des choix à sa place lors de la consultation d'un serveur vocal, par exemple ;
- envoyer un faux paquet de type `OpenReceivedChannelAck`, juste avant l'émission du véritable paquet, contenant un mauvais numéro de port UDP pour spécifier le canal de donnée : l'utilisateur légitime n'entendra alors que du silence dans son combiné.

Insertion d'un paquet quelconque dans le canal de contrôle

Dans certains cas, on ne peut de prédire avec certitude le prochain paquet émis sur le canal de contrôle par le Call Manager et reçu par le téléphone phone IP (ou vice versa).
Cependant, il est toujours possible d'envoyer un paquet qui sera pris en compte par le téléphone IP (un signal de sonnerie occupée par exemple). Mais dans ce cas, les numéros de séquence TCP étant erronés, le téléphone plantera quelques secondes plus tard, lors de l'émission du paquet suivant.

Déni de service généralisé

Nous avons vu qu'il suffit d'envoyer un paquet TCP de taille quelconque, mais doté d'un numéro de séquence correctement choisi, à un téléphone pour le rendre inutilisable jusqu'à son extinction manuelle. Il est donc facile d'effectuer un déni de service généralisé sur l'ensemble du réseau VoIP.

Le scénario suivant pourrait par exemple être mis en œuvre :

1- écoute passive des *broadcasts* émis par les téléphones IP au moment de leur démarrage ;
2- identification d'un téléphone IP en train de booter ;
3- écoute passive de ce téléphone, à l'aide de ARP spoofing, le temps de récupérer la séquence TCP (30 secondes suffisent en général) ;
4- envoi d'un paquet TCP au téléphone IP, mettant celui-ci hors service.

Le réseau téléphonique entier serait alors hors service, pendant tout le temps que durerait l'attaque : à l'instant ou un téléphone serait redémarré manuellement, l'attaque le mettrait à nouveau hors service de manière quasi immédiate.
La localisation de l'origine d'une attaque de ce type risquerait d'être longue et difficile.

Nous verrons les parades à ce type d'attaque au paragraphe sur les recommandations.

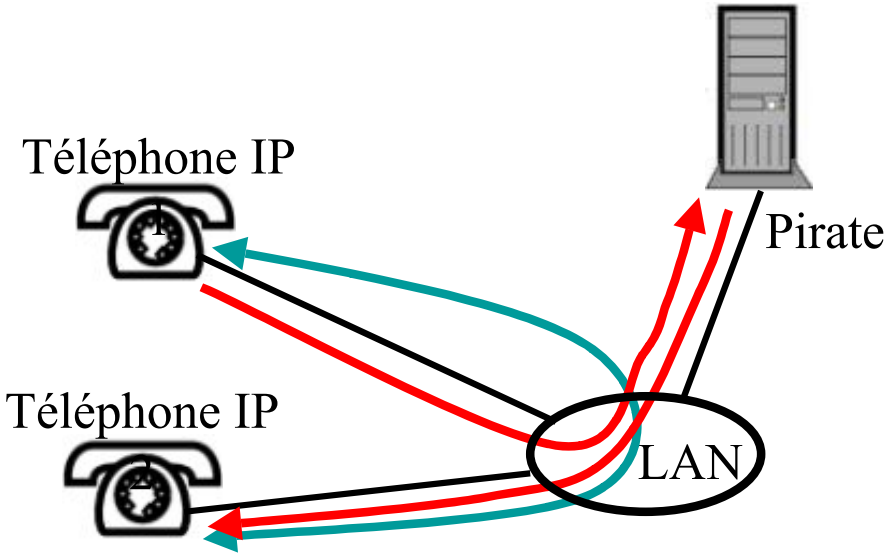
Détournement du trafic VoIP

En supposant que le LAN est commuté, il est nécessaire de recourir à la technique d'ARP spoofing pour intercepter le trafic associé à un téléphone IP. Cette technique consiste à envoyer sur le réseau Ethernet des paquets ARP forgés annonçant que l'adresse MAC possédant l'adresse IP cible est celle de notre carte réseau, et non celle de la machine cible réelle. Le résultat est que les paquets IP sont envoyés vers notre machine, et non plus vers la machine légitime. Nous pourrons alors rerouter le trafic reçu vers la machine cible, ou bien le faire disparaître.

Dans le cas de la VoIP, le but est :

- de se faire passer pour le serveur Call Manager, la passerelle voix-données VG200 ou un autre téléphone IP distant vis-à-vis du téléphone IP à écouter ;
- de se faire passer pour le téléphone IP à écouter vis-à-vis du Call Manager, de la passerelle VG200 ou du téléphone IP distant.

La figure ci-dessous illustre ce mécanisme :



Trajet du flux voix sans (vert) et avec (rouge) détournement de trafic

Dans cette configuration, tout flux reçu ou émis par le téléphone IP de la victime passe par l'ordinateur de l'attaquant. Celui-ci peut alors enregistrer la conversation, la modifier à la volée ou la renvoyer vers d'autres postes.

Il est tout à fait possible d'écouter un numéro de téléphone particulier. En fonction de ce numéro, on localise le téléphone IP correspondant en se connectant successivement à tous les téléphones pour accéder à son interface web et récupérer le numéro. Une recherche complète sur un réseau de taille importante ne dure que quelques minutes.

Une fois le téléphone IP cible identifié (par son adresse IP), on se met en écoute et on attend l'établissement d'une communication. A ce moment, l'ARP spoofing est lancé afin d'intercepter le flux voix. Ce flux étant récupéré de manière transparente, le processus est indétectable par les utilisateurs espionnés. L'assemblage des paquets audio ne pose en général pas de problème pour reconstituer la conversation d'origine. Bien souvent, il suffit d'additionner le signal circulant dans un sens et dans l'autre pour reconstituer la « bande son » de la conversation entre les deux protagonistes, qu'ils soient tous deux situés dans l'entreprise ou que l'un d'entre eux seulement s'y trouve.
L'écoute téléphonique est donc possible à large échelle sur un réseau local VoIP. Il suffit de se connecter à une prise réseau pour écouter l'intégralité des communications émises ou reçues.

Attaques « génériques »

Anonymat

Dans la majorité des échanges informatiques, les différentes adresses des correspondants (Mails, IPs, FQDN...) ne sont pas des facteurs déterminants, au contraire du contenu des transactions.

Pour la VoIP, les adresses « directes » sont au premier plan, au même titre que les numéros de téléphones RTC, d'où l'apparition de la fonction « numéro caché » et de la liste rouge !
L'anonymat des correspondants n'est pas forcément simple à mettre en œuvre car il devient difficile de filtrer le trafic et de cohabiter avec des règles strictes de filtrage.
Les relais SIP permettent d'assurer l'anonymat des utilisateurs, dans une certaine mesure (car limités à des plages d'adresses la plupart du temps), en se référant au *Call-ID* qui doit être présent dans tous les paquets de la même transactions pour qu'ils soient recevables. En revanche, les dénis de service (par envoi continu de SIP INVITE par exemple) sont alors plus difficiles à contrer.
L'anonymat ne peut donc pas être assuré dans le cas d'une communication VoIP directe sans *proxy* entre 2 clients.

Spoofing

Dans un paquet SIP, les identifiants d'une communication lorsqu'elle est établie sont le *Call-ID*, et les *tags* des champs *From* et *To* s'ils sont utilisés. Les 3 principaux types de *spoofing* sur le protocole SIP s'effectuent avec des paquets SIP INVITE, BYE et CANCEL (cf. *Figure 5*).

Le premier est trivial à mettre en oeuvre en LAN comme sur Internet en forgeant un paquet SIP INVITE (avec des champs *Call-ID*, *From*, *To*, et *Cseq* adéquats). L'hôte distant sonnera mais la conversation ne pourra s'établir que si le *tag* du champ *To* est identique à celui retourné par l'appelé lorsqu'il accepte l'appel. Malheureusement, de nombreuses implémentations de SIP n'utilisent pas les *tags* et sont donc vulnérables à ces attaques, même lorsque l'attaquant n'est pas en mesure de renifler le trafic.

Dans le paquet SIP le champ *From* est suivi d'un *tag*. Lors de la réception du SIP INVITE, l'appelé ajoutera également son *tag* aléatoire et toutes les communications suivantes devront inclure ces 2 *tags* respectifs sur *From* et *To*.
La probabilité de réussite pour falsifier les 2 paquets suivants est quasiment nulle sans écoute du trafic si l'aléa du *Call-ID* et des *tags* sont corrects.

Dans le cas d'un LAN où il est possible d'écouter le trafic réseau, et donc de récupérer les *Call-IDs*, *From*, *To*, et les *tags* s'ils sont utilisés, tous les types d'attaques sont envisageables, du MITM au DoS en passant par une écoute passive temps réel des conversations ou une modification à la volée des paquets. Seule la voix pourrait alors faire défaut... !

Et sur Internet ?

Comme nous l'avons vu, bon nombre d'attaques se font grâce à de l'ARP spoofing ou du DNS spoofing. L'attaque ARP n'est pas très réaliste sur l'Internet et c'est pourquoi l'interception de la communication se fera plutôt en utilisant d'autres moyens.

La proximité (au sens réseau) du pirate et de la source ou de la destination est un facteur clé pour faciliter l'attaque et l'écoute.

L'attaque la plus commune reste le déni de service (mutualisé) à l'encontre du lien d'accès à l'Internet ou de la passerelle VoIP-PSTN/LAN. Toujours aussi simple à mettre en œuvre et complexe à filtrer et à tracer.

En ce moment, de petits malins recherchent activement ce genre de plates-formes VoIP-PSTN connectées à l'Internet et mal configurées : cela leur permet de terminer gratuitement leur communication VoIP via l'Internet sur le réseau téléphonique commuté dans différents pays…

"Ajouts" et recommandations sécurité

RTP (Real Time Transport Protocol) encode, transporte et décode la voix. La qualité du flux étant essentielle dans le domaine de la VoIP, 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 clefs ne disposaient d'aucune protection fiable (SIP, RTP, RTCP, ...). Les problèmes s'accentuaient avec l'utilisation massive d'UDP pour accélérer les échanges, entre autres avec les problèmes évidents de *spoofing*.

La VoIP permet de remplacer la totalité des lignes RTC classiques (avec l'utilisation de PABX-IP), et c'est alors que la sécurité de l'architecture est rentrée dans le cahier des charges.
Plusieurs protocoles sont apparus avec notamment les équivalents chiffrés de RTP et RTCP : SRTP et SRTCP [1] (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 : *Authentification tag* ;
- un deuxième contenant différentes informations sur la clef : *Master Key Identifier (MKI)* ;
- et bien sûr un *payload* chiffré.

SRTP et SRTCP ont été développés dans un souci de performances, 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) identifiers
RTP extension
PAYLOAD (CHIFFRÉ)
MKI
Authentification tag

Paquet SRTP

Comme le montre la figure, 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 pourrait être modifié, voire les champs optionnels *MKI* ou *Authentification tag*.

Quand SRTP est utilisé conjointement avec SIP [2], le déroulement chronologique des transactions est le suivant :

- Appelant : établissement de la communication (*ringing*) 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, et si les deux participants se sont préalablement mis d'accord sur l'algorithme de chiffrement utilisé, alors la communication sécurisée est établie.

Dans le cas contraire, si une négociation des clefs est nécessaire, un protocole de gestion des clefs s'impose sur le même principe, par exemple, qu'IKE pour IPSec.

C'est dans ce but que MiKEY [3] (*Multimédia Internet Keyring*) a été développé : il s'agit d'un protocole récent à l'état de *draft* et encore très rarement 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 clefs sans affecter de façon significative la rapidité des échanges.

Le projet *minisip* [4], bien qu'encore trop jeune pour être utilisé en production, est l'un des plus avancés dans ce domaine à l'heure actuelle. Il s'agit d'un client implémentant MiKEY et SRTP avec au choix une authentification *PSK* ou Diffie-Hellman. Il supporte également l'utilisation de TLS pour sécuriser les échanges SIP.

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 clefs adéquats :

```
INVITE sip:lolo@rstack.org;user=phone SIP/2.0
From: <sip:mp@domain.org;user=phone>;tag=1017556314
To: <sip:lolo@rstack.org;user=phone>
Call-ID: 1664131130@rstack.org
CSeq: 101 INVITE
Contact: <sip:mp@rstack.org:5060;user=phone;transport=UDP>;expires=900
Content-Type: application/sdp
```

```
INVITE sip:lolo@rstack.org;user=phone SIP/2.0
Via: SIP/2.0/UDP 192.168.0.1:5060
[…]
m=audio 32945 RTP/AVP 0 97
a=key-mgmt:mikey  AQAFgAAAAsAxNbayXB+WngBEH (…) jPANqgLOmmfPvB+/f56vA==
a=rtpmap:0 PCMU/8000/1
a=rtpmap:97 iLBC
```

Paquet SIP INVITE + MiKEY

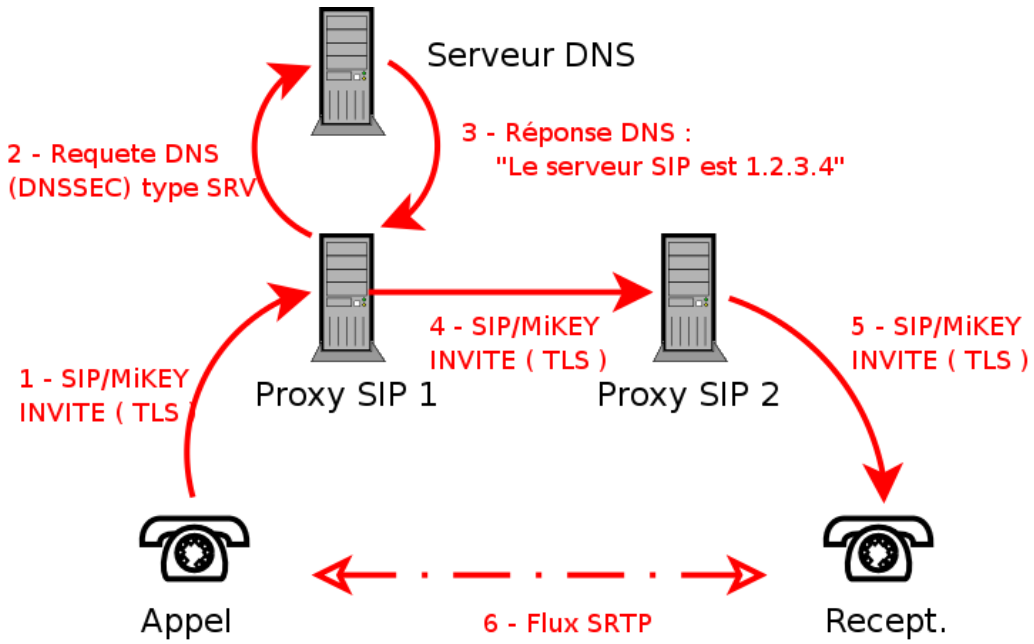
En utilisant des relais SIP, lorsque mp@rstack.org cherche à contacter lolo@rstack.org, le relais sortant effectue une requête DNS de type SRV pour connaître l'IP du serveur SIP distant, par exemple celle de sip.domaine.org.

Cette requête est une cible idéale pour un pirate, par exemple via une attaque de *spoofing* sur le *DNS ID* ou de *cache poisoning*[5] afin de renvoyer l'adresse d'un *proxy* SIP qu'il contrôle. L'utilisation de DNSSEC pourrait donc être envisagée à ce niveau.

Comme nous l'avons vu, 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* propose l'utilisation de TLS pour limiter ces risques.

La *figure 3* représente une solution sécurisée d'architecture de VoIP, mais il en existe bien sûr beaucoup d'autres. *Minisip* permet de tester cette solution dans son intégralité. Pour plus de détails, vous pouvez consulter [6].



Solution sécurisée, également celle retenue par le projet minisip...

La VoIP 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 VoIP 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 clefs pour l'ensemble des communications VoIP.
La surcharge engendrée par IPSec peut être minimisée en configurant le tunnel pour traiter uniquement les flux de voix sur IP (pour des machines/protocoles fixés).Un atout intéressant est la possibilité d'utiliser la totalité des *soft phones* 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 VoIP sécurisée, mais ne peuvent pas être employés sur de larges infrastructures ou sur des *soft phones* peu puissants.

Filtrage réseau

Comme nous l'avons vu plus haut, les serveurs de gestion VoIP (surtout installés par défaut) ont un nombre de ports ouverts par défaut très conséquent.
Il est donc fortement recommandé de filtrer les ports accessibles sur les serveurs depuis le réseau des utilisateurs, au niveau des routeurs. Pour cela, il peut être utile de placer les serveurs sur un sous-réseau dédié.

Il convient de lister les services et les ports associés qui doivent être pris en considération lors de l'implémentation d'une politique de filtrage sur un réseau VoIP. Certains protocoles sont propriétaires (ex : Cisco Skinny) et d'autres ne font pas bon ménage avec du filtrage sans état (non-*stateful*) ou si des relais applicatifs, qui savent décoder le protocole, ne sont pas présents (ALGs – Application Level Gateways). Attention, bon nombre de pare-feux se limitent à gérer l'ouverture de ports en fonction des communications et n'inspectent pas les flux (au niveau protocolaire – ie. AGLs). De plus cet élément additionnel risque d'introduire un délai ainsi qu'une gigue, c'est pourquoi ils sont absents dans bien des déploiements.

IDS

Il n'est pas très courant de trouver des outils de détection d'intrusion pour des solutions de voix sur IP. La quantité de faux positifs dus à l'observation du flux RTP pourrait être plus que conséquente. Bien qu'elle permette de détecter des dénis de service par exemple, la détection d'intrusion se ramène souvent à de la détection de fraude.

Recommandations

Les principales recommandations permettant de se protéger contre les attaques précédentes sont les suivantes:

- verrouiller les adresses MAC / adresses IP par port dans les commutateurs traversés par du flux VoIP (cela peut être relativement simple lorsque les téléphones IP ne se déplacent jamais) ;
- utiliser IPSec entre les équipements VoIP ;
- mettre à jour les applicatifs et les *firmwares* des équipements VoIP ;
- activer le chiffrement des communications VoIP ;
- utiliser l'authentification des équipements VoIP ;
- activer au mieux les fonctions de sécurité offertes par les équipements de l'infrastructure VoIP utilisée et qui ne seraient pas activées par défaut.

En ce qui concerne l'architecture Cisco, la nouvelle version du Call Manager (4.x) introduit des améliorations notables en matière de sécurité :

- signature du code des *firmwares* et vérification lors du téléchargement sur les IP Phones ;
- authentification réciproque des IP Phones et des serveurs Call Manager par certificats (un certificat unique est installé d'origine dans chaque IP Phone, et une Certificate Trust List ou CTL est utilisée) ;
- authentification de la signalisation (TLS est utilisé afin de vérifier que les paquets de signalisation n'ont pas été modifiés) ;
- chiffrement des communications (à l'aide du protocole SRTP, qui utilise des clés asymétriques pour chiffrer et authentifier les paquets de données audio) ;
- possibilité de désactiver certaines fonctionnalités des IP Phones (désactivation du port de connexion au PC, désactivation du broadcast de l'adresse ARP du téléphone au démarrage, désactivation de l'accès au VLAN voix, désactivation de l'accès à certaines pages de configuration et de statistiques).

Il est donc fortement recommandé de migrer une architecture CCM 3.x en version 4.x.

Interception "légale" de trafic

L'interception légale de trafic est une fonction/interface qui est présente dans la majorité des réseaux téléphoniques (fixes et cellulaires). Pour l'instant cette obligation légale est encore au stade de discussions dans beaucoup de pays et les autorités de régulations ont des idées assez divergentes : voir [18] et [19].

Une anecdote assez intéressante : un nombre conséquent de personnes pensent que les agences de renseignement doivent avoir de plus en plus de mal à intercepter les communications vu le nombre de médias et la quantité d'informations. Un commentaire d'un haut responsable d'une telle agence suggère tout le contraire : en effet, « avant », il fallait numériser toutes ces informations pour devoir les traiter, aujourd'hui cette étape conséquente se fait aux extrémités ce qui leur simplifie largement la tâche…

Les réseaux de téléphonie “classiques”

A titre de comparaison, nous allons discuter, sans entrer dans les détails, de la sécurité des réseaux téléphoniques plus classiques.

PSTN/POTS

Le réseau téléphonique filiaire (PSTN/POTS – Public Switched Telephone Network/Plain Old Telephone System) transporte voix et données. A la différence d’un réseau VoIP, où les équipements qui “gèrent” les communications sont souvent dans le même réseau et accessibles, cela n’est pas le cas dans le RTC (SS7 et le “switch” par exemple).

Il est bien connu que la majorité des téléphones sans fil analogiques peuvent être écoutés, mais que la distance est limitée à quelques centaines de mètres. Les téléphones sans fil du type DECT peuvent chiffrer la communication.

GSM

Le réseau cellulaire (GSM – Global System for Mobile Communications), à la différence d’un réseau sans fil de type 802.11, ne permet pas d’écouter facilement une communication. En revanche, les communications ne sont chiffrées qu’entre le téléphone de l’utilisateur et la station à laquelle il est rattaché. Ce n’est pas un chiffrement de bout en bout entre l’appelant et l’appelé, mais des solutions existent qui emploient par exemple le mode données pour transporter de la voix chiffrée.

Un changement majeur est apparu ces dernières années avec le déploiement des réseaux de troisième génération (GPRS et UTMS) : les téléphones ne sont plus des grille-pains mais sont livrés avec un système d’exploitation, Java, une pile TCP/IP, etc… De plus, ils sont connectés en permanence au réseau et disposent d’une adresse IP.

Skype

Skype [7] est un logiciel récent développé par l’équipe de *KaZaA* qui innove dans la VoIP en proposant un système reposant sur le principe du *P2P* (*peer-to-peer*) via le réseau *FastTrack*.
Les clients s’interrogent pour connaître les « réseaux de contacts » et ainsi communiquer directement entre eux. *Skype* ne demande aucune configuration particulière sur les équipements du réseau puisqu’il nécessite uniquement des connections TCP sortantes vers les ports 80 ou 443, par exemple. Il gère également les relais HTTP classiques.

En mode *PC-to-PC*, la communication est chiffrée en AES (*Advanced Encryption Standard*) 256 bits entre les clients. Reste à déterminer le niveau de confiance que l’on peut accorder à un algorithme dont on ne maîtrise pas l’implémentation : le code source n’est pas disponible ! Les clefs AES sont négociées en utilisant RSA (1536 bits à 2048 bits). *SkypeOut*, qui permet de contacter le réseau RTC, ne chiffre pas les échanges : il aurait pu être possible cependant de le faire jusqu’au PABX-IP distant…

Skype semble offrir une belle perspective à la VoIP sur Internet pour un usage personnel. Les contraintes de sécurité ont été prises en compte, ce qui est souvent rare

dans ce type de projets. En revanche, « l'opacité » de *Skype* sur les protocoles qu'il utilise ou sur son fonctionnement le rend difficile à utiliser dans un contexte professionnel, surtout si la confidentialité est au centre des préoccupations.

D'autres projets, comme SIPshare [18], s'intéressent par exemple à l'utilisation de SIP dans un réseau P2P.

VoWLAN

VoWLAN, comme son nom l'indique, n'est rien d'autre que de la VoIP appliquée sur des réseaux sans fil. Elle offre encore de nouvelles perspectives, notamment avec l'utilisation de PDAs ou de téléphones SIP compatibles WLAN. La qualité peut alors largement dépasser celle des téléphones portables cellulaires et il est possible de couvrir de grandes distances avec plusieurs AP en *roaming*. Le protocole IAPP (*Inter Access Point Protocol*) assure la normalisation de cette technologie pour passer d'AP en AP de façon transparente.

En revanche, l'utilisation du Wifi pose certains problèmes, notamment sur les questions de débit et de sécurité... Des CODECs appropriés sont nécessaires (PCM, GSM, …) pour assurer une qualité minimale sur une bande passante réduite, en fonctions de différents critères, dont la compression, les délais (établissement, latence...), et la qualité (écoute, écho, pertes, ...).
La sécurité du réseau Wifi doit être assurée au préalable (cf. [8]) tout en garantissant un débit minimum afin d'éviter les surcharges liées à la sécurisation du réseau sans fil (WEP/WPA, IPSec, tunnels, etc.).

La VoWLAN présente tous les enjeux de sécurité de la VoIP auxquels viennent s'ajouter ceux des réseaux sans fil.

Conclusion

Nous avons couvert le sujet de la voix sur IP d'un point de vue technique et nous pensons qu'aujourd'hui une solution de voix sur IP peut-être sécurisée à un niveau acceptable.

Un projet de voix sur IP est complexe, car il n'existe pas de solution générique, et une étude au cas par cas s'impose avant la mise en oeuvre de cette technologie. Le facteur sécurité doit être pris en compte avant même la phase de conception en posant les bonnes questions aux vendeurs que vous êtes en train de sélectionner.

La convergence des deux réseaux (informatique et téléphonie) change la donne pour la partie voix. Il convient de définir clairement quel département est responsable de la sécurité des parties et de l'ensemble, ce qui bien trop souvent n'est pas fait.

Références

- [1] RFC 3711 : SRTP et SRTCP
- [2] RFC 3261 : SIP
- [3] RFC (*draft*) MiKEY - draft-ietf-msec-mikey-08.txt
- [4] minisip- *Soft phone* SIP/MiKEY - http://www.minisip.org
- [5] « Failles intrinsèques du protocole DNS » - http://betouin.security-labs.org/Article_DNS_PBT.pdf
- [6] « Secure Mobile Voice over IP » - ftp://ftp.it.kth.se/Reports/DEGREE-PROJECT-REPORTS/030626-Israel_Abad_Caballero-final-report.pdf
- [7] Skype - http://www.skype.com
- [8] “La sécurité des réseaux 802.11 : quoi de neuf depuis un an ?” - MISC n°12
- [9] Sécurisation de Windows NT/2000/XP/2003 - http://www.chambet.com/publications/sec-win2k/
- [10] Sécurisation d’IIS - http://www.chambet.com/publications/iis-security/
- [11] Vulnérabilités et sécurisation des applications Web - http://www.chambet.com/publications/sec-web-apps/
- [12] RFC 1189 : RTP
- [13] H.235 : http://www.javvin.com/protocolH235.html
- [14] H.323 : http://www.h323forum.org/papers/
- [15] Sécurisation des routeurs et des commutateurs Cisco - http://www.miscmag.com/articles/index.php3?page=214
- [16] PROTOS SIP - http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/
- [17] http://www.voip-info.org/wiki-Asterisk+security
- [18] CALEA - http://www.fcc.gov/calea/
- [19] ETSI and LI - http://www.ripe.net/ripe/meetings/ripe-48/presentations/ripe48-eof-etsi.pdf
- [20] SIPshare - http://www.research.earthlink.net/p2p/