VoIP 1

Etude et implementation avec SIP

Essentiel

Campus Booster ID : 338 - 385
Version 1.0

Sommaire

1. La VoIP

1.1. Description

Parfois appelée VoIP (Voice over IP) ou ToIP (Telephony over IP), la transmission de la voix sur les réseaux informatiques est le résultat entre les besoins permanents en communication de notre société et la démocratisation de ces réseaux informatiques, ces derniers offrant un support de plus en plus fiable pour le transport des données, et des offres de connexion à Internet toujours plus accessibles et attractives.

Il y a une différence notable entre la VoIP et la ToIP. La VoIP concerne le transport de flux en temps-réel, principalement la voix, sur les réseaux de données. La VoIP devient la ToIP lorsqu'elle est raccordée au réseau téléphonique classique. La ToIP est donc un sous-ensemble de la VoIP, avec des restrictions imposées par le monde de la téléphonie classique, comme notamment l'obligation de respecter la nomenclature des numéros d'appel.

L'objectif de la VoIP est donc de remplacer, au moins en partie, la téléphonie classique souvent onéreuse, principalement pour les communications internationales, en utilisant les réseaux informatiques déployés de part le monde.

Les réseaux informatiques étant construits pour le transport des données, la voix est donc numérisée, via un codec, puis encapsulée dans un paquet avant d'être transportée. Les codecs utilisés pour la téléphonie sur IP ne dépendent pas vraiment du protocole utilisé, mais plutôt de l'implémentation sur les logiciels et les équipements réseau.

Le transport de la voix sous forme numérique peut être intégral ou plus généralement partiel. Dans ce dernier cas, les communications téléphoniques internationales seraient par exemple réduites au coût d'une communication locale. Des sociétés comme Skype ont largement contribuées à démontrer cet intérêt réel en termes d'économie.

1.2. Historique

La numérisation de la voix existe depuis très longtemps, et les échanges vocaux par le biais des ordinateurs ne sont pas rares depuis l'apparition de la messagerie instantanée par exemple.

Il y avait une trop grande barrière entre le monde informatique et le monde de la téléphonie. La VoIP est donc réellement apparue au moment où l'on a commencé à faire le lien entre les moyens de communication informatique et la téléphonie classique, conjointement à l'élaboration et à la ratification de normes et protocoles spécifiques.

La VoIP existe depuis plusieurs années au sein des entreprises, en fonction des moyens technologiques et financiers existants. Un bon exemple est celui de la société Cisco Systems, qui a mis en place dès le début leur propre offre de produits et services en usage interne. Certains fournisseurs exercent d'ailleurs leur métier autour du service voix sur IP depuis maintenant plus de 10 ans.

Comme toutes les nouvelles technologies, la VoIP a gagné en popularité lorsque les prix des produits associés ont grandement baissé. L'offre des ISP (Internet Service Provider) incluant le service de téléphonie a aussi grandement contribué à démarginaliser la VoIP aux yeux du grand public.

1.3. Comparatif avec la téléphonie classique

1.3.1. Avantages

La montée en puissance de la téléphonie sur IP est flagrante. Ceci est principalement dû aux avantages que cette technologie apporte par rapport à la téléphonie classique.

Parmi ces avantages, on peut citer les points suivants :
- Architecture unique
- Economies
- Services ajoutés
- Mobilité

Architecture unique : L'un des objectifs de la téléphonie sur IP est l'intégration du réseau de téléphonie au réseau de données, pour former un seul et unique réseau. De plus, il est possible de n'utiliser qu'une seule liaison vers un opérateur pour le transport des données (connexion à Internet), alors qu'auparavant une liaison téléphonique classique était obligatoire.

Economies : Un autre avantage est le coût des communications. En effet, une bonne infrastructure VoIP offre des coûts de communication inférieurs, voir même nuls dans certains cas, qu'elles soient locales ou internationales. La possibilité de ne pas souscrire à un abonnement téléphonique classique permet aussi de faire des économies substantielles.

Services ajoutés : L'avantage ayant le plus grand impact sur les bénéfices de la VoIP est l'offre presque illimitée de services pouvant être greffés. On peut rapidement lister :
- **Voice Mail** : Répondeur téléphonique synchronisé avec la messagerie électronique, pour recevoir les messages vocaux sous la forme d'emails.
- **Click-to-Dial** : Lancement d'appels téléphoniques vers un destinataire via le client de messagerie (exemple : Microsoft Outlook).
- **Gestion de présence** : Redirection automatique vers le terminal le plus proche de l'utilisateur.
- **Synchronisation des contacts** : Centralisation des adresses postales, emails et numéros de téléphone autour d'un unique annuaire.

Mobilité : Malheureusement, les téléphones classiques de bureau ne peuvent être qu'à un seul emplacement physique. Les téléphones IP peuvent suivre l'utilisateur quelque soit le lieu, la seule réelle restriction étant un accès au réseau de données. De plus, il est possible d'avoir plusieurs terminaux IP pour un même utilisateur (un téléphone IP au bureau et un softphone avec accès VPN pour les déplacements par exemple), un protocole s'occupant alors de la gestion de présence de l'utilisateur et de la redirection des appels vers le bon terminal.

1.3.2. Inconvénients

Certains avantages peuvent apparaître comme des inconvénients, en fonction du contexte. On peut donc citer les inconvénients suivants :
- Architecture unique
- Coût de la VoIP
- Qualité et fiabilité

Architecture unique : Mutualiser des réseaux peut provoquer des problèmes qui n'existaient pas avant la fusion. En effet, la VoIP est une application qui vient se poser par-dessus une infrastructure réseau standard comme n'importe quelle autre application. Certains détails doivent alors être pris en compte (comme la QoS, la sécurité des transmissions, la disponibilité, et la résistance aux attaques réseaux, etc.), afin d'assurer le service de téléphonie.

Coût de la VoIP : Malheureusement, la VoIP a un coup principalement lié à l'infrastructure et aux équipements. C'est pourquoi certaines entreprises ne seraient pas obligatoirement gagnantes avec un passage vers la VoIP. En général, les entreprises optent pour un passage progressif vers la VoIP pour remplacer, à terme, les classiques téléphones et PABX. Ce passage progressif se fait via l'utilisation d'adaptateurs, principalement des FXS et FXO.

Par conséquent, et sauf exceptions, seules les entreprises qui démarrent optent pour une solution IP intégrale, vu que le réseau de données peut être prévu pour l'usage de la VoIP dès le départ. Les autres entreprises préfèreront plutôt opter pour une transition progressive.

Qualité et fiabilité : La téléphonie sur IP utilisant les réseaux de données, y compris Internet, pour faire transiter les flux, les appels peuvent alors subir quelques désagréments (perte de paquets, délais, etc.) nuisant à la qualité générale de la communication.

Par exemple, il est généralement reconnu qu'un délai inférieur à 150 ms est requis pour une qualité optimale. Or, il est malheureusement courant sur certaines liaisons de dépasser ce délai (délai moyen de 500 ms observé sur une liaison satellite).

De plus, la VoIP étant une application transitant sur le réseau, elle est donc tout aussi sensible que les autres applications par rapport aux problèmes pouvant survenir sur ce réseau, comme les dénis de services (DoS et DDoS) ou plus simplement la congestion.

1.4. Les acteurs de la VoIP

Le marché de la téléphonie sur IP est très vaste. De nombreuses entreprises ont investi dans ce marché, en proposant leurs solutions.

Les grands noms actuels dans le monde de la VoIP sont :

- Alcatel
- Audiocode
- Cirpack
- Cisco
- Linksys
- Quintum
- RAD

De plus, de multiples opérateurs offrent des services liés à la VoIP. Dresser la liste de ces opérateurs serait impossible, compte tenu de leur nombre. Par contre, on peut les différencier en plusieurs catégories :

- **ISP proposant des services VoIP** : Tous les opérateurs historiques (ou presque) proposent maintenant des services VoIP à leurs clients. On peut citer par exemple France Telecom, ou bien tous les ISP pour particuliers fournissant le service VoIP au travers de leur « box » (Free, Neuf Télécom, AOL, etc.).

- **Opérateurs dédiés à la VoIP pour les entreprises** : Ses fournisseurs, parfois assez récents sur le marché de la VoIP, se sont dédiés à l'offre de solutions aux entreprises (NetCentrex, Verizon, etc.).

- **Opérateurs dédiés à la VoIP pour les particuliers** : Ces opérateurs (Skype, Vonage, VoIP Buster, etc.) ne fournissent en général que l'accès à une infrastructure VoIP, les flux transitant ainsi au travers d'une connexion Internet classique. Ces solutions sont donc limitées en termes de fonctionnalités et de qualité de service. Elles sont par conséquent réservées aux particuliers.

1.5. Le futur : Everything over IP

La VoIP va pouvoir tirer partie des capacités et fonctionnalités presque illimitées, en termes de services ajoutés à la simple transmission de la voix, que propose les TIC. Ces services n'ont en effet pour limite que l'imagination des développeurs d'applications.

Le futur de la VoIP s'oriente clairement vers la mutualisation et la multiplication des services offerts (plates-formes IP Centrex), mais aussi vers la mobilité avec les systèmes 3G/Wi-Fi (smartphones « dual mode »).

Les solutions IP Centrex commencent à être largement disponibles, au travers de multiples solutions. L'IP Centrex correspond globalement à l'offre de services centralisés autour d'une même plate-forme, et hébergés chez l'opérateur.

Les protocoles actuellement en place permettent une grande souplesse dans leur utilisation. Il est donc théoriquement possible de faire passer n'importe quel type de flux temps-réel, la limitation se trouvant concrètement dans les fonctionnalités des terminaux et des plates-formes de services.

De plus, les nouveaux smartphones possédant une interface Wi-Fi permettent de mettre en place des solutions de communications « dual mode ». Ce système fournit le meilleur moyen de communication par rapport à l'environnement disponible. Si un hot-spot Wi-Fi est disponible, alors un client VoIP peut être utilisé. Sinon, le téléphone peut basculer automatiquement sur le réseau GSM.

2. Protocoles liés à la VoIP

Beaucoup de protocoles sont, ou furent, utilisés pour la transmission de flux temps-réel sur les réseaux IP. Il est possible de distinguer deux grandes familles de protocoles :

- **Les protocoles de signalisation** : Utilisés pour l'établissement et le contrôle des appels. Ces protocoles peuvent être ensuite redécoupés suivant leur mode de fonctionnement :
  - Client/Serveur (asymétrique)
  - Peer-to-Peer (symétrique)

Protocole	Client/Serveur	Peer-to-Peer
Cisco SCCP	✓	
MGCP	✓	
SIP		✓
H.323		✓

- **Les protocoles de transport des flux** : Permettant de transit des flux sur le réseau IP.

2.1. Protocoles de signalisation

2.1.1. Protocole SIP

SIP (Session Initiation Protocol) est un protocole de la couche application du modèle OSI. Il a été spécifié par le groupe de travail MMUSIC (Multiparty Multimedia Session Control) de l'IETF (Internet Engineering Task Force) en mars 1999. La ligne de conduite était alors de concevoir un protocole de signalisation facile à implémenter, évolutif et flexible. En juin 2002, une nouvelle normalisation, la RFC 3261, est publiée. Elle constitue aujourd'hui le recueil des spécifications fondamentales du protocole SIPv2.

SIP a pour fonction d'établir, modifier et terminer des sessions multimédia avec un ou plusieurs participants, indépendamment des protocoles de la couche transport et sans dépendance sur le type de session qui est établie. Un participant peut aussi être invité dans une session préétablie. De même, une donnée pourra être rajoutée ou supprimée d'une session existante.

Par session, on entend un ensemble d'appelants et d'appelés qui communiquent entre eux. Les conférences multimédias et les appels téléphoniques via Internet en sont des exemples.

Toutefois, SIP n'est pas le seul protocole nécessaire aux équipements de communication. En effet, son but est de rendre la communication possible, la communication en elle-même doit être effectuée par d'autres moyens. Ce qui implique que, pour obtenir une plateforme multimédia complète, SIP doit être combiné avec d'autres protocoles.

Typiquement, ceci implique, selon la RFC 3261, les protocoles suivants :

- **RTP (Real-time Transport Protocol)** : Pour assurer le transport des flux en temps réel. Il encode et divise les données en paquets, puis les transporte à travers le réseau IP.
- **SDP (Session Description Protocol)** : Pour la description des paramètres des sessions multimédia.
- **RTSP (Real-Time Streaming Protocol)** : Pour contrôler la livraison des flux en streaming.
- **MGCP (Media Gateway Control Protocol)** : Pour les passerelles de contrôle au réseau téléphonique commuté public (PSTN).

RTP et SDP sont les protocoles le plus souvent employés avec le protocole SIP.

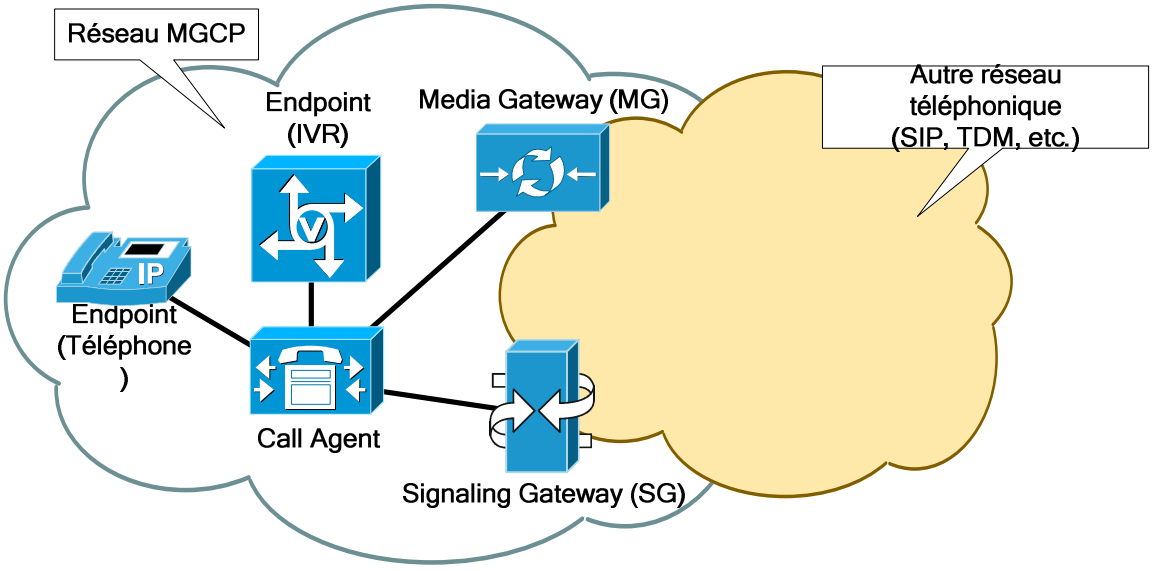
SIP est basé sur le protocole HTTP, lequel peut être également considéré comme un protocole de signalisation dans la mesure où il permet de demander à un serveur une ressource précise. SIP profite de la valeur éprouvée du protocole sans doute le plus utilisé et reconnu à travers le globe.

SIP peut être transporté dans des segments TCP ou UDP. Le numéro port par défaut est le 5060, sauf pour SIP sur TLS (Transport Layer Security) qui utilise le port 5061.

2.1.2. Protocole MGCP

MGCP (Media Gateway Control Protocol) est un protocole de signalisation spécifié en janvier 2003 par la RFC 3435 et a pour base la RFC 3015. Cette dernière a été conjointement développée par le groupe de travail MEGACO de l'IETF et ITU-T (International Telecommunication Union – Telecom standardization).

MGCP est généralement encapsulé dans des segments UDP, sur le port 2427. Il utilise SDP pour la description du média et RTP pour le transport des flux.



Architecture MGCP

MGCP est une architecture client/serveur composée des éléments suivants :

- **Call Agent** : C'est le softswitch du réseau VoIP avec MGCP. Sa principale mission est la coordination des MG et SG, en leur indiquant qui doit accomplir les fonctions requises. C'est le Call Agent qui reçoit les notifications et qui dirige les fonctions de l'infrastructure VoIP.

- **Media Gateway (MG)** : Au moins une MG doit être présente dans l'architecture MGCP. Elle s'occupe principalement de la conversion des flux entre circuits (TDM) et paquets (IP). Plus généralement, elle s'occupe du traitement des flux.

- **Signaling Gateway (SG)** : Au moins une SG doit être présente dans l'architecture MGCP s'il y a une connexion avec un autre réseau de téléphonie. Elle permet la conversion des informations de signalisation (appels, etc.) depuis et vers un autre réseau.

- **Point d'extrémité (Endpoint)** : C'est une source de signal. Cela peut être un téléphone, un serveur de conférence ou un serveur vocal interactif (IVR) par exemple.

Il peut y avoir plusieurs Call Agents dans une même infrastructure, garantissant ainsi une disponibilité par redondance et une répartition de la charge de gestion des appels.

Les différentes fonctions du réseau VoIP sont ensuite déléguées à une ou plusieurs passerelles MGCP (MG et/ou SG).

Il est possible de mutualiser plusieurs fonctions dans un même dispositif réseau. Il n'est donc pas rare de trouver des serveurs intégrant un Call Agent, une MG, une SG, et un point d'extrémité (serveur de conférence).

2.1.3. Suite de protocoles H.323

H.323 est une suite de protocoles élaborée par l'ITU-T définissant des standards pour les communications multimédias. La première publication date de 1996, et la version actuelle (version 6) date de Juin 2006.

Les protocoles décrits sont regroupés en sept catégories :

- **Contrôle d'appel et signalisation**
  - H.225.0 : Protocoles de signalisation d'appel et mise en paquets des flux média (utilise une partie de la signalisation Q.931)
  - H.225.0/RAS : Enregistrement, admission et statut (Registration, Admission and Status)
  - H.245 : Protocole de contrôle pour communication multimédia

- **Codecs audio**
  - G.711
  - G.722
  - G.723.1
  - G.728
  - G.729

- **Codecs vidéo**
  - H.261
  - H.263

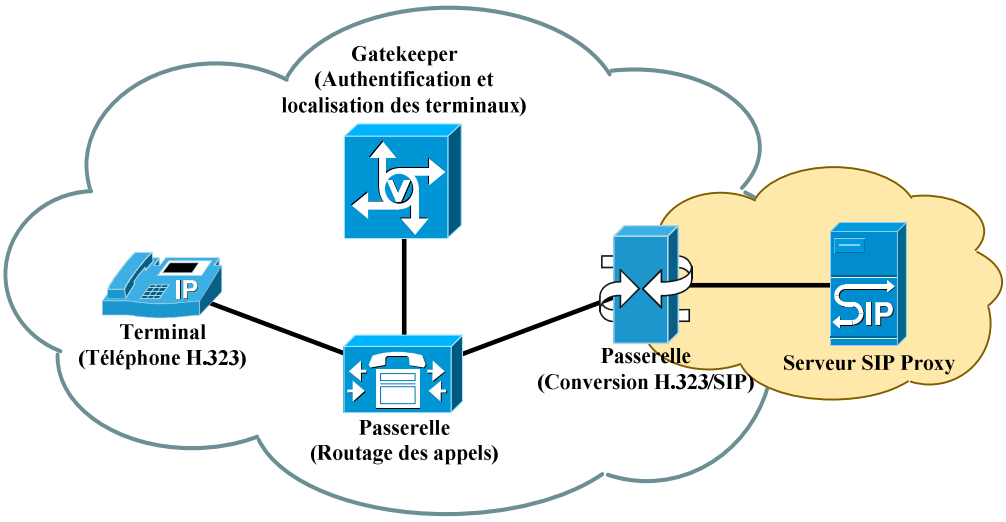
- **Transmission de données**
  - T.120 : Suite de protocoles pour la transmission de données (utilisé par des applications de travail collaboratif en temps-réel)

- **Transport sur le média**
  - RTP
  - RTCP

- **Sécurité**
  - H.235 : Sécurité et cryptage pour les terminaux multimédia (séries H)

- **Services supplémentaires**
  - H.450.1 : Fonctions génériques pour les services supplémentaires
  - H.450.2 : Transfert d'appel
  - H.450.3 : Déviation d'appel
  - H.450.4 : Mise en attente d'appel
  - H.450.5 : Parquage et récupération d'appel
  - H.450.6 : Mise en attente d'appel
  - H.450.7 : Indication de message en attente
  - H.450.8 : Services d'identification de noms
  - H.450.9 : Services de complétion d'appel pour les réseaux H.323

Les communications H.323 ne nécessitent pas obligatoirement de point central. Nous sommes dans un mode de fonctionnement Peer-to-Peer. Un terminal H.323 peut donc directement communiquer avec un autre terminal H.323 sans passer par un serveur.



Architecture H.323

L'architecture du standard H.323 est composée de :

- **Terminaux** : Décrit le dispositif d'extrémité de chaque lien. Il fournit deux méthodes de communication en temps réel avec un autre terminal H.323, une passerelle ou un MCU. Cette communication se compose d'une combinaison de dialogues, de données et/ou de vidéos.

- **Passerelles** : Elles établissent la connexion entre terminaux H.323, de même qu'avec les terminaux de réseaux utilisant d'autres protocoles tels que le réseau téléphonique commuté classique, SIP ou encore MGCP.

- **Gatekeepers** : Fournit des mécanismes d'enregistrement et d'authentification des terminaux, permet le contrôle de la bande passante, assure la translation entre numéro de téléphone et adresse IP, mais aussi le transfert et le renvoi d'appel par exemple.

- **MCUs (Multipoint Control Unit)** : Ils établissent les conférences et sont composés de :
    - Multipoint Control mandaté qui assure la signalisation d'appels et le contrôle de conférence.
    - Multipoint Processor qui fournit la commutation et le mixage des flux. Occasionnellement, il assure le transcodage des flux audio et vidéo reçus.

2.1.4. Protocole Cisco SCCP

SCCP (Skinny Client Control Protocol) est un protocole propriétaire Cisco. Il est utilisé entre la plate-forme Cisco CallManager et les téléphones IP de l'équipementier.

Le protocole s'architecture autour d'un serveur, qui peut être sous la forme d'un cluster, qui va centraliser l'intégralité du traitement, simplifiant ainsi les fonctions devant être gérées par les terminaux (téléphones et passerelles IP/TDM).

Les fonctions centralisées vont du simple traitement des appels, au maintien à jour des versions logicielles des terminaux, en passant par la fourniture de multiples services.

Les messages sont transportés dans des segments TCP sur le port 2000.

2.1.5. Comparatif des différentes solutions

H.323 était de loin le protocole le plus populaire, car il était employé couramment dans les communications multimédias. De plus, le protocole est en place depuis plusieurs années, son utilisation éprouvée et sa grande maturité en faisaient une solution idéale. Ainsi, des investissements lourds ont été consentis à la conception de grands réseaux H.323. De nombreux produits conçus par de grands noms de l'informatique tels que Cisco, Microsoft, IBM, et Intel, utilisent ce standard.

Toutefois, SIP est très en vogue ces dernières années et sa côte de popularité est exponentielle. Son habilité à combiner aisément la voix et les services IP est son principal atout. L'établissement de l'appel est plus rapide qu'avec H.323, grâce à la séparation des champs d'entête du corps du message qui est traité plus facilement et dont le temps de transition sur le réseau diminue. SIP s'est donc aujourd'hui imposé comme le standard *de facto* pour la VoIP/ToIP.

2.2.Protocoles de transport des flux

2.2.1. Protocole RTP

RTP, pour Real-Time Transport Protocol, a été conçu en janvier 1996. La dernière version est décrite dans la RFC3550 de juillet 2003.

RTP fournit des fonctions de transport bout-à-bout appropriées aux applications transmettant des données en temps réel, telles que les flux audio, vidéo ou encore simulation de données, à travers des services de diffusion multicast ou unicast. RTP fonctionne conjointement avec RTCP, qui se charge de la QoS et du transport des informations en rapport avec les participants dans une session en cours.

Les données RTP sont typiquement transportées dans des segments UDP, les applications utilisant RTP étant généralement peu sensibles aux pertes de paquets mais plus sensibles aux délais. Il n'y a pas de numéro de port par défaut, excepté un détail : le flux RTP utilise un numéro de port pair et le flux RTCP associé le numéro impair suivant (port RTCP = port RTP + 1).

Les services proposés par RTP sont :
- Identification de type de charge utile
- Numérotation de séquence
- Horodatage
- Surveillance de la livraison

Les mécanismes de délivrance opportune et autres garanties de QoS ne sont pas assurés par RTP mais par des protocoles de couche inférieure. Par conséquent, RTP part sur le principe que le réseau est fiable.

Diverses applications sont faites de l'usage de RTP. En voici quelques-unes :
- Conférences multimédias à plusieurs participants
- Stockage de données en continue
- Simulation distribuée interactive
- Badge actif
- Applications de contrôle et mesures

2.2.2. Protocole RTCP

RTCP (Real-Time Transport Control Protocol) fait partie intégrante de la RFC 3550 définissant aussi RTP.

Ce protocole fournit différents services périodiques de contrôle pour les flux RTP hors bande (out-of-band), vu que les paquets RTP et RTCP sont distincts.

RTCP fournit quatre fonctions principales :

- **Feedback sur la qualité de la transmission** : Ceci est effectué par les rapports d'émetteur (SR) et de récepteur (RR).

- **Transport de l'identifiant de la source de flux RTP (CNAME)** : Utile par exemple lorsqu'un récepteur doit associer 2 flux RTP différents pour une même session (flux voix et vidéo d'une visioconférence).

- **Paquets RTCP envoyés par tous les participants** : Cette fonction permet à chaque participant d'observer le nombre de participants dans une session multimédia. Cela permet aussi de calculer la fréquence d'émission des rapports (SR et/ou RR) pour s'adapter à n'importe quel nombre de participants.

- **Informations de contrôle minimal de session (optionnel)** : Ces informations fournissent des détails sur l'arrivée ou le départ de participants dans une conférence. Il est donc par exemple possible de maintenir en temps réel une liste des participants dans une conférence.

L'utilisation de RTCP n'est pas obligatoire mais elle est fortement recommandée pour toute session, surtout pour celles fonctionnant dans un environnement multicast (plusieurs participants).

2.2.3. Protocole SRTP

Le protocole SRTP, pour Secure RTP, est une évolution de la RFC 3550 résolvant les problématiques d'authentification, de confidentialité et d'intégrité des flux transportés. Il est défini conjointement à SRTCP (Secure RTCP) dans la RFC 3711 de mars 2004.

La confidentialité est assurée par l'algorithme AES, qui peut être implémenté suivant différents modes de chiffrage (cipher).

L'authentification et l'intégrité, ainsi que le Replay Protection sont assurés par l'algorithme HMAC-SHA1 (empreinte sur 160 bits).

Les fonctions de RTP et RTCP sont maintenues dans ces versions sécurisées.

2.3.Codecs

Un codec est un algorithme de compression/décompression utilisé pour représenter un signal, généralement audio ou vidéo, dans un environnement digital.

Il existe une multitude de codecs, chacun ayant des différences principalement au niveau de la qualité du signal compressé, représentée par un indice MOS (Mean Opinion Score), et de la charge de calcul pour traiter les signaux.

Ces codecs ont plusieurs critères similaires :

- Débit binaire
- Taille des échantillons (typiquement 20 ou 30 ms)
- Latence induite par l'échantillonnage (égale à la taille de l'échantillon)
- Temps de compression/décompression du signal (latence variable en fonction de l'algorithme)
- Nombre de trames par secondes

Voici un tableau regroupant les codecs les plus courants dans le monde de la VoIP, leur débit binaire, et la bande passante utilisée sur un réseau Ethernet (en partant sur la base de 40 octets pour la somme des en-têtes UDP, IP et Ethernet) :

Codec	Codec Bit Rate	Bande passante sur un réseau Ethernet
G.711	64 Kbps	87,2 Kbps
G.723.1	5,3 Kbps	20,8 Kbps
G.723.1	6,4 Kbps	21,9 Kbps
G.726	24 Kbps	47,2 Kbps
G.726	32 Kbps	55,2 Kbps
G.728	16 Kbps	31,5 Kbps
G.729	8 Kbps	31,2 Kbps

Bande passante utilisée par différents codecs

On peut ainsi estimer le nombre maximum d'appels simultanés que pourrait gérer une liaison.

Un calculateur de bande passante est disponible à cette adresse : http://www.bandcalc.com

```
Les débits du tableau précédent s'entendent pour un flux. Une
communication téléphonique comprend en général deux flux, l'émission
et la réception, il faut donc prendre soin de dimensionner les
liaisons en conséquence.
```

3. Protocole SIP

3.1. Définitions

- **Dialogue** : Echange entre deux User Agent pendant une période donnée. Un dialogue est un ensemble de transactions.
- **Initiateur** : L'entité qui initie une session avec une requête INVITE.
- **Invitation** : Requête INVITE.
- **Invité** : Le récepteur d'une requête INVITE.
- **Message** : Demande ou réponse échangée entre éléments SIP.
- **Méthode** : Désigne le type de requête transmis à un serveur. Par exemple, les requêtes INVITE et BYE.
- **UAC (User Agent Client)** : Un UAC est une entité logique qui remplit le rôle de client d'une application client/serveur. C'est lui qui envoie des requêtes et reçoit des réponses.
- **UAS (User Agent Server)** : Un UAS est une entité logique qui remplit le rôle de serveur d'une application client/serveur. C'est lui qui reçoit des requêtes et transmet les réponses.
- **URI (Uniform Resource Identifier)** : Une URI identifie une entité en employant une syntaxe, proche de celle utilisée pour les emails, de la forme « sip:identifiant@domaine » (par exemple sip:john@sip.labo-voip.com).
- **Proxy Server** : Entité intermédiaire à la fois client et serveur qui fournit un service de routage aux clients qui souhaitent joindre d'autres clients. Par conséquent, le serveur Proxy effectue des requêtes au nom d'autres clients.
- **Redirect Server** : UAS qui redirige vers un ensemble d'URIs alternatifs en générant des réponses 3xx aux requêtes qu'il reçoit.
- **Registrar Server** : Serveur qui accepte les requêtes REGISTER qu'il reçoit et stocke les informations. Il est utilisé pour l'identification et/ou l'authentification des utilisateurs.
- **Requête** : Envoyé d'un client à un serveur, ce message SIP permet d'invoquer une opération particulière.
- **Réponse** : Envoyé d'un serveur à un client, ce message SIP indique le statut d'une requête envoyée précédemment par le client au serveur.
- **Session** : Flux multimédia échangé entre un ensemble d'émetteurs et de récepteurs.
- **Transaction** : Se compose de tous les messages échangés entre un client et un serveur, de la première requête à la réponse finale.

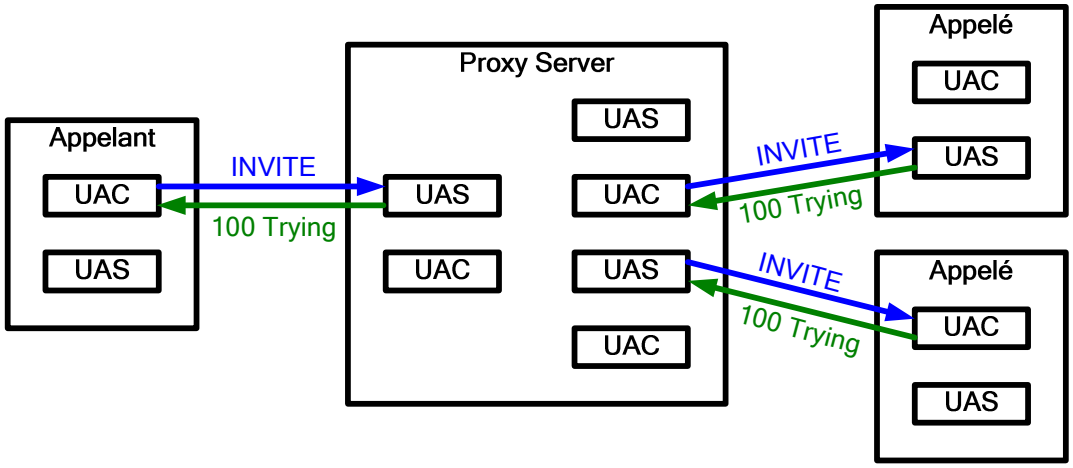
- **Stateful Proxy** : Maintient l'état lors de transactions entre client et serveur.
- **Stateless Proxy** : Transmet chaque requête et réponse qu'il reçoit sans maintien d'état de la transaction.

3.2. Architecture

3.2.1. User Agents

Ce sont des entités logicielles ou physiques qui utilisent SIP pour trouver une autre entité de destination.

Les User Agents peuvent être (liste non exhaustive) :
- Softphones (applications logicielles)
- Téléphones IP (fixes ou Wi-Fi)
- Smartphones et PDAs
- IPBX
- Passerelles IP/TDM



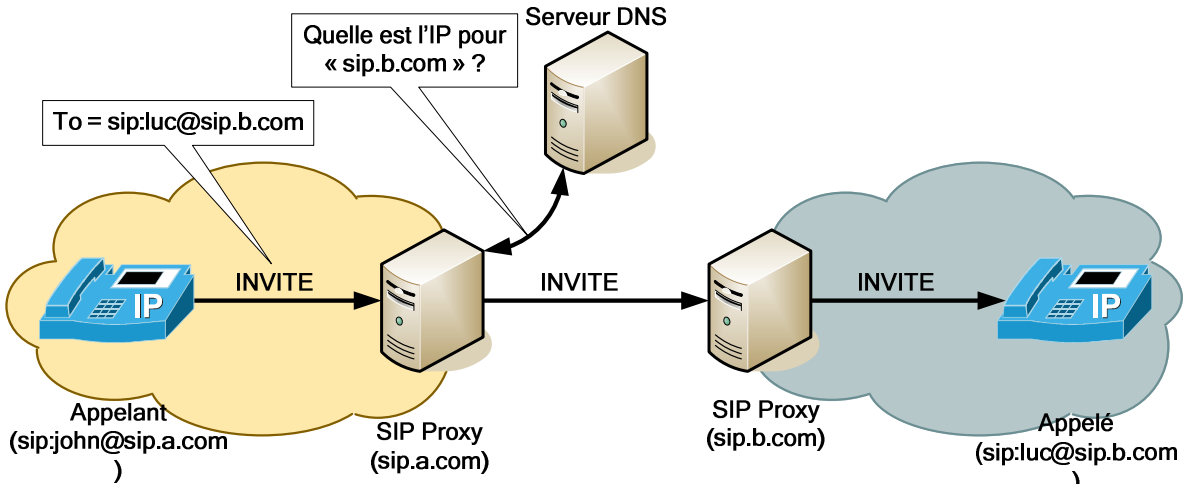
Exemple simple d'architecture SIP basique

Chaque User Agent comprend un UAS et un UAC. Ce sont des entités logiques qui permettent pour l'un, d'envoyer des réponses, recevoir des requêtes et pour l'autre d'envoyer des requêtes, recevoir des réponses. Il est à noter que l'état client ou serveur ne dure que la durée d'une transaction. Ainsi, un User Agent est à tour de rôle client et serveur.

3.2.2. Proxy server

Pièce importante de l'architecture SIP, il fournit un service de routage des messages envoyés par un client, en tenant compte de certaines fonctions importantes comme :
- La localisation actuelle de l'appelé
- La comptabilité (pour facturation)
- Etc.

Les messages peuvent traverser un ensemble de serveurs Proxy, jusqu'à atteindre celui qui connaît la localisation de l'appelé.



Transfert entre les serveurs Proxy pour joindre la destination

Il existe deux types de serveur Proxy abordés ci-après :

- **Stateless Server** : Simple et plus rapide que le Stateful Server, il transmet les messages indépendamment des autres sans tenir compte de l'état des transactions. De ce fait, le Stateless Server ne fournit pas de mécanismes de retransmission de messages. Toutefois, il est utilisé pour le partage de charge, la translation de messages et le routage. Globalement, un Stateless Proxy ne se charge que de transférer les messages qu'il reçoit. Il ne génèrera donc pas ses propres messages de réponse temporaire par exemple.

- **Stateful Server** : Contrairement au Stateless Server, il maintient l'état de la transaction de la première requête à la réponse finale. Cette particularité inclut un temps de traitement supplémentaire et rend le serveur moins rapide, mais permet d'avoir des fonctions très avantageuses :
  - Le forking en est un exemple, il permet de redistribuer une requête vers plusieurs destinations (initiation d'une session avec plusieurs destinataires).
  - La retransmission de messages, car il connaît le contenu de la transaction.
  - La localisation des utilisateurs, il est ainsi possible de renvoyer un appel vers le mobile d'un utilisateur, alors que l'appel était initialement transmis vers le téléphone du bureau.
  - La comptabilité.
  - L'aide à la translation NAT.

En général, les utilisateurs du réseau de VoIP/ToIP utilisent les noms de domaines de l'entreprise pour la partie réseau de l'URI. Le ou les serveurs SIP Proxy sont alors identifiés via des entrées DNS de type SRV, à l'instar des serveurs de messagerie identifiés par des entrées de type MX. Cela permet d'avoir une URI unique, quel que soit le SIP Proxy utilisé par l'entreprise.

Les entrées DNS de type SRV sont écrites en suivant le formalisme de la RFC 2782 :

```
{_Service._Protocol} SRV {Priorité} {Port} {Cible}
```

Les entrées pour un domaine DNS d'une entreprise pourraient donc ressembler à ceci :

```
_sip._udp SRV 0 5060 sip.a.com
_sip._udp SRV 1 5060 backupsip.a.com
```

Ainsi, toutes les URI des utilisateurs de ce réseau pourraient être formatées comme suit :

Sans entrée DNS de type SRV	Avec entrée DNS de type SRV
utilisateur@sip.a.com	utilisateur@a.com

3.2.3. Registrar Server

C'est un serveur qui fournit un moyen de localiser les utilisateurs. Pour cela, les utilisateurs s'enregistrent en envoyant des requêtes d'enregistrement au serveur (REGISTER). Ce dernier extrait les informations permettant de localiser l'utilisateur, telles que l'adresse IP, le numéro de port et le nom d'utilisateur. Puis, il stocke sur une base de données ces informations.

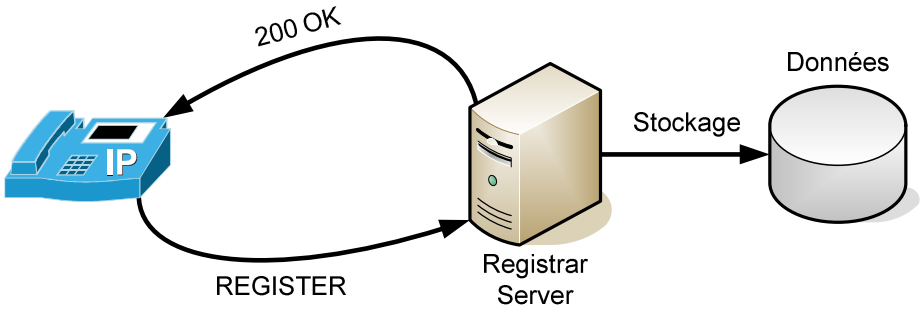


Schéma simple d'enregistrement auprès d'un serveur SIP Registrar

Le serveur Registrar peur accomplir une simple identification, processus minimum pour localiser les utilisateurs sur le réseau IP. Il est aussi possible de mettre en place une authentification, pour contrôler les utilisateurs qui se connectent et utilisent le réseau VoIP.

Il est possible d'identifier ou d'authentifier l'appelant et/ou l'appelé.

3.2.4. Redirect Server

Le serveur de redirection permet d'obtenir une liste des locations courantes d'un utilisateur particulier. La base de données créée par un Registrar Server est la source d'informations utilisée par le serveur de redirection pour dresser cette liste, qui est transmise par une réponse de la classe 3xx. De cette façon l'appelant possède une liste des destinations possibles de l'appelé.

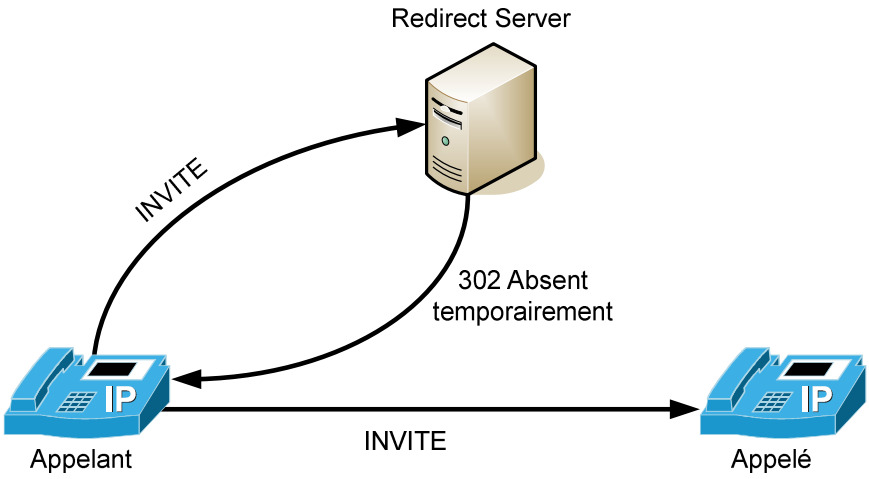


Schéma simple de redirection par un serveur SIP Redirect

3.2.5. Autres types de serveurs SIP

Il existe une multitude de serveurs SIP, chacun répondant à une fonctionnalité précise. On peut citer les serveurs suivants (liste non exhaustive) :

- **Serveur de conférences** : Ce serveur va procéder au mélange des flux RTP en provenance des différents intervenants, et s'occupe de toutes les fonctions relatives à la gestion de ces conférences.

- **Serveur de Voice Mail** : Il centralise les fonctions de messagerie vocale. Les messages vocaux peuvent ensuite être gérés via une interface vocale, une interface HTML, ou bien au travers de mail, lorsqu'un lien est créé entre ce serveur de Voice Mail et le serveur de messagerie de l'entreprise (Microsoft Exchange par exemple).

- **Serveur IVR (Interactive Voice Response)** : Les serveurs vocaux interactifs permettent la création de menus vocaux pour prétraiter les appels. Ce type de service est particulièrement utilisé par les supports techniques.

Tous les serveurs SIP peuvent être des entités réseaux séparées, ou bien fusionnées dans une unique machine. Il est aussi possible de multiplier certains serveurs, pour différentes raisons allant de la redondance à la répartition de charge.

3.3. Méthodes SIP

3.3.1. Messages SIP

Les communications SIP se font au moyen d'une série de messages qui peuvent être de deux natures :

- **Requêtes** : Permet d'invoquer une opération particulière.
- **Réponses** : Permet d'informer l'initiateur d'une requête que cette dernière a bien été reçue, traitée, voir aussi du résultat obtenu après traitement.

Chaque message est composé d'une première ligne qui indique le type de message, de l'en-tête du message (en-tête SIP) et optionnellement du corps du message. Les deux derniers sont séparés par une ligne vide.

Le corps du message peut être de plusieurs types. Le plus courant est un message SDP inclus dans une requête INVITE.

La grande malléabilité du protocole SIP provient entre autres de la liberté de créer des requêtes et/ou réponses personnalisées. Il est donc possible de créer des services supplémentaires.

3.3.2. En-tête SIP

L'en-tête SIP est écrit sous la forme d'une succession de champs, dont voici les principaux :

Champs	Description
Via	Indique le chemin emprunté par le message (typiquement l'adresse de l'UAC qui vient d'envoyer le message)
From	Indique l'initiateur du message
To	Indique le destinataire du message
Contact	Fournit la ou les URIs pour joindre l'appelant pour des communications futures
Call-ID	Identifiant unique permettant de distinguer une communication
CSeq	(Command Sequence) Identifiant unique de transaction au sein d'une même session
Content-Type	Indique le type de média du corps du message envoyé
User-Agent	Chaîne de caractères stipulant le terminal utilisé pour envoyer ce message
Content-Length	Indique la taille du corps du message

Voici un exemple de message INVITE envoyé :

```
INVITE sip:luc@sip.b.com SIP/2.0
Via: SIP/2.0/UDP 10.1.16.170:5060;rport;branch=C4BF7BAD282A1EA948DFA
From: John <sip:john@sip.a.com>;tag=3580587940
To: <sip:luc@sip.b.com>
Contact: <sip:john@10.1.16.170:5060>
Call-ID: FC9C664C-8134-47F2-877B-2ACBF60DB1B9@10.1.16.170
CSeq: 47647 INVITE
Max-Forwards: 70
Content-Type: application/sdp
User-Agent: X-Lite release 1105x
Content-Length: 254
```

3.3.3. En-tête SDP

Le message SDP, alias le corps du message SIP, contient plusieurs champs répartis en trois catégories :

- Description de la session
- Description temporelle
- Description du média

Il existe 20 champs différents répartis dans les trois catégories ci-dessus. Il est inutile de toutes les présenter, par contre, connaître les principales peut s'avérer utile :

Champs	Signification	Description
v	Version	Version du protocole SDP (v=0)
o	Origin	Fournit des informations sur l'origine de la session (<username> <session id> <version> <network type> <address type> <address>)
c	Connection Data	Indique les données de la connexion (<network type> <address type> <connection address>)
t	Times	Fournit les informations de temps de la session (<start time> <stop time>)
m	Media Announcements	Spécifie des détails du transport du ou des flux sur le réseau, le dernier paramètre indiquant le ou les codecs utilisés (décrits par les champs « a=rtpmap ») (<media> <port> <transport> <fmt list>)
a	Attributes	Différents attributs de session, servant ici principalement à énumérer les différents codecs pouvant être utilisés pour la communication (rtpmap:<payload type> <encoding name>/<clock rate>)

```
RTP/AVP = Real-Time Transport Protocol using the Audio/Video profile carried over UDP
```

Voici un exemple d'en-tête SDP envoyé dans un message INVITE :

```
v=0
o=john 16742548 16742652 IN IP4 10.1.16.170
s=X-Lite
c=IN IP4 10.1.16.170
t=0 0
m=audio 8000 RTP/AVP 3 98 97 101
a=rtpmap:3 gsm/8000
a=rtpmap:98 iLBC/8000
a=rtpmap:97 speex/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
a=sendrecv
```

3.3.4. Requêtes SIP

Il existe plusieurs types de requêtes SIP. Néanmoins, les plus importantes sont décrites ci-après :

- **INVITE** : Permet d'initier une session multimédia.
- **REGISTER** : Contient les informations de la localisation courante d'un utilisateur, l'adresse IP et le numéro de port. Cette requête est à destination d'un serveur Registrar.
- **BYE** : Permet de mettre fin à une session établie.

- **ACK** : Accuse réception de la réponse finale à une requête INVITE. La durée de l'établissement de la session en utilisant une méthode en trois étapes est aléatoire. En effet, elle dépend du temps que prendra l'appelé à accepter ou rejeter l'appel. Alors, l'appelé renvoie périodiquement la réponse jusqu'à la réception de l'accusé de réception.

- **CANCEL** : Annule la session en cours d'établissement. Par exemple, lorsque l'appelé prend trop de temps à donner une réponse.

3.3.5. Réponses SIP

Les réponses sont identifiées par un code définit par la version 2 du protocole SIP. Le code consiste en une valeur allant de 100 à 699, ces dernières étant classées en 6 catégories de réponses :

- **Réponse prévisionnelle 1xx** : Le traitement d'une requête peut être plus ou moins long. Aussi les réponses 1xx permettent d'informer l'émetteur que la requête a bien été reçue et est en cours de traitement. Cela permet à l'initiateur d'arrêter la retransmission de la requête. Le chiffre 100 (TRYING) est utilisé lors de requêtes INVITE, et le chiffre 180 pour signaler une sonnerie en cours (RINGING).

- **Réponse finale positive 2xx** : Indique qu'une requête à été traitée et acceptée. 200 (OK) est la réponse positive à une requête INVITE par exemple.

- **Redirection 3xx** : Quand un serveur Proxy ne peut satisfaire un appel, il redirige l'appelant vers un service alternatif qui pourra établir l'appel. Ce service peut être un autre serveur Proxy ou la nouvelle localisation de l'appelé.

- **Réponse finale négative 4xx (erreur client)** : Indique qu'une requête ne peut être traitée ou que la requête à une mauvaise syntaxe et que le problème vient de l'appelant.

- **Réponse finale négative 5xx (erreur serveur)** : Indique que le serveur ne peut traiter la requête bien quelle soit valide. L'appelant retransmettra la requête par la suite.

- **Réponse finale négative 6xx (échec global)** : Indique que la requête ne peut être traitée par aucun serveur. Généralement, l'appelé décline sa participation à une session par une réponse 603.

```
La première ligne contient un message dans le langage humain
exprimant la raison de la réponse transmise par le User Agent de
l'utilisateur.
```

3.3.6. Liste des réponses SIP prédéfinies

Il existe plusieurs réponses prédéfinies. Les codes et leurs significations sont présentés dans ce tableau :

Code (Message)	Signification
100	Trying
180	Ringing
181	Call Is Being Forwarded
182	Queued
183	Session Progress
200	OK
202	Accepted
300	Multiple Choices
301	Moved Permanently
302	Moved Temporarily
305	Use Proxy
380	Alternative Service
400	Bad Request
401	Unauthorized
402	Payment Required
403	Forbidden
404	Not Found
405	Method Not Allowed
406	Not Acceptable
407	Proxy Authentication Required
408	Request Timeout
410	Gone
412	Conditional Request Failed
413	Request Entity Too Large
414	Request-URI Too Long
415	Unsupported Media Type
416	Unsupported URI Scheme
420	Bad Extension
421	Extension Required

Code (Message)	Signification
422	Session Interval Too Small
423	Interval Too Brief
429	Provide Referrer Identity
480	Temporarily Unavailable
481	Call/Transaction Does Not Exist
482	Loop Detected
483	Too Many Hops
484	Address Incomplete
485	Ambiguous
486	Busy Here
487	Request Terminated
488	Not Acceptable Here
489	Bad Event
491	Request Pending
493	Undecipherable
494	Security Agreement Required
500	Server Internal Error
501	Not Implemented
502	Bad Gateway
503	Service Unavailable
504	Server Time-out
505	Version Not Supported
513	Message Too Large
580	Precondition Failure
600	Busy Everywhere
603	Decline
604	Does Not Exist Anywhere
606	Not Acceptable

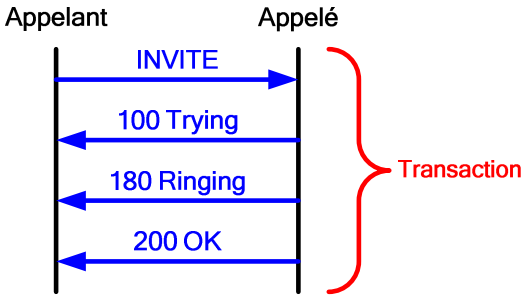
3.4. Echanges SIP

3.4.1. Transactions SIP

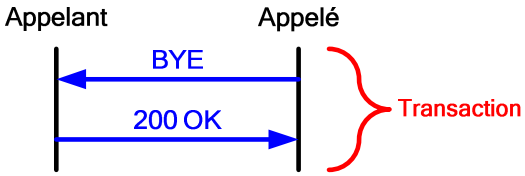
SIP est un protocole transactionnel, ce qui implique qu'une requête et toutes les réponses associées soient regroupées en transactions.

Les transactions sont facilement identifiables, car tous les messages SIP utiliseront le même numéro de séquence (CSeq).

Toutefois, une particularité est à noter avec les ACK. En effet, l'ACK n'est pas pris en compte lors d'une réponse finale positive à une requête car, bien qu'il n'y ait qu'une requête, plusieurs participants peuvent répondre positivement à cette dernière. Par contre, l'ACK sera pris en compte lors d'une réponse finale négative.



Transaction d'établissement d'un appel



Transaction de cloture d'un appel

3.4.2. Dialogues SIP

Un dialogue SIP est un échange de transaction entre deux User Agents dans le temps. Par ailleurs, il facilite l'ordonnancement et le routage de messages entre points d'extrémités SIP.

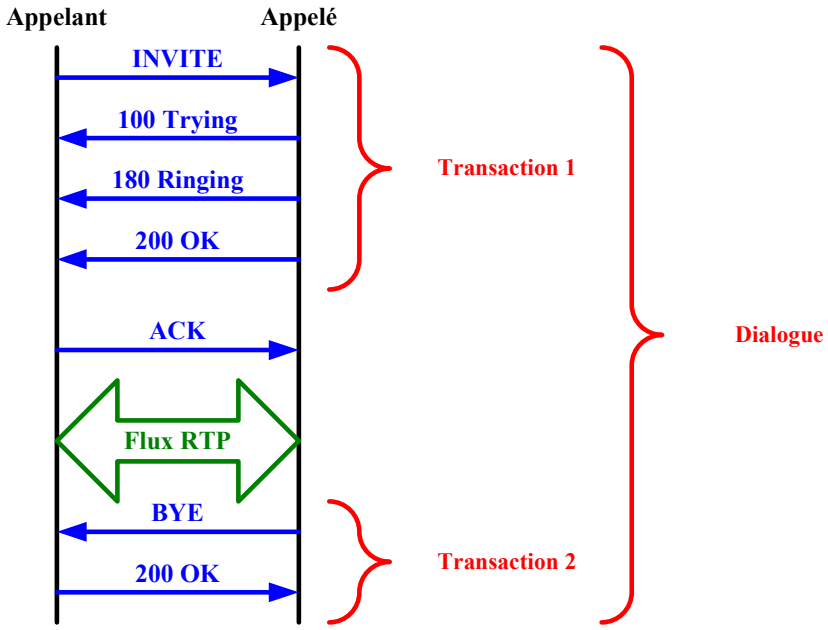
D'un point de vue pragmatique, un dialogue est une succession logique de transactions.

Les champs suivant d'un message SIP permettent d'identifier un dialogue :

- **Call Id** : Identifie un appel composé d'un ou plusieurs dialogues. Permet ainsi de distinguer les dialogues entre eux.
- **From** : Identifie le dialogue côté appelant.
- **To** : A l'inverse identifie le dialogue côté appelé.
- **CSeq** : Ordonne les messages au sein du dialogue et permet d'identifier une transaction.

En effet, un dialogue, et donc aussi les transactions correspondantes, est composé des messages qui partagent les mêmes paramètres d'identification. L'identification de dialogue permet à deux User Agent de poursuivre leur relation sans recours à un serveur Proxy une fois que tous deux connaissent leur localisation.

De plus, certains messages établissent un dialogue, d'autre pas. Le meilleur exemple est la requête BYE, qui à lieu dans le dialogue préétabli par une requête INVITE.

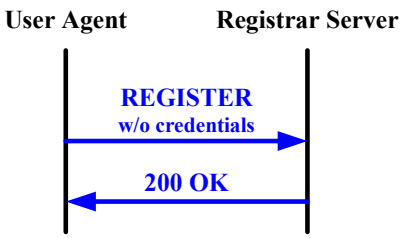


Exemple d'un dialogue découpé en deux transactions

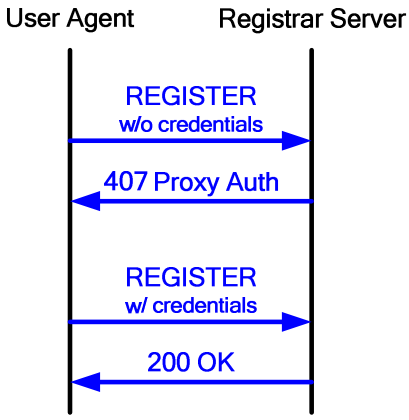
3.5. Transactions typiques

3.5.1. Enregistrement

L'enregistrement d'un client auprès d'un serveur SIP Registrar s'effectue via une requête REGISTER. Le serveur peut être configuré pour une simple identification, pour obtenir les informations de localisation du client, ou pour l'authentification, de manière à s'assurer de l'identité du client.



Enregistrement avec identification



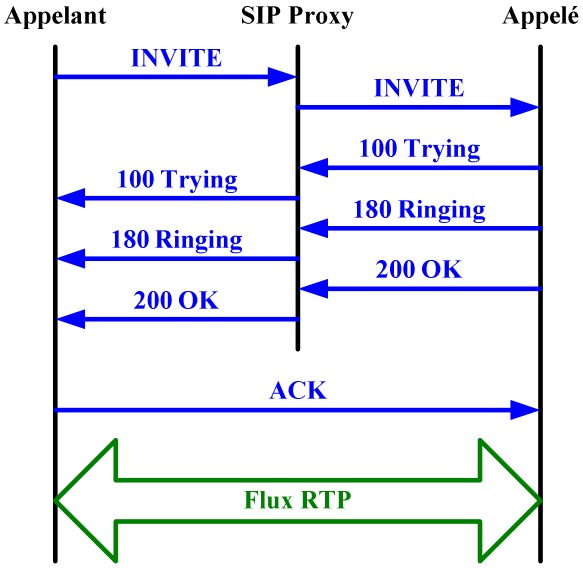
Enregistrement avec authentification

3.5.2. Invitation

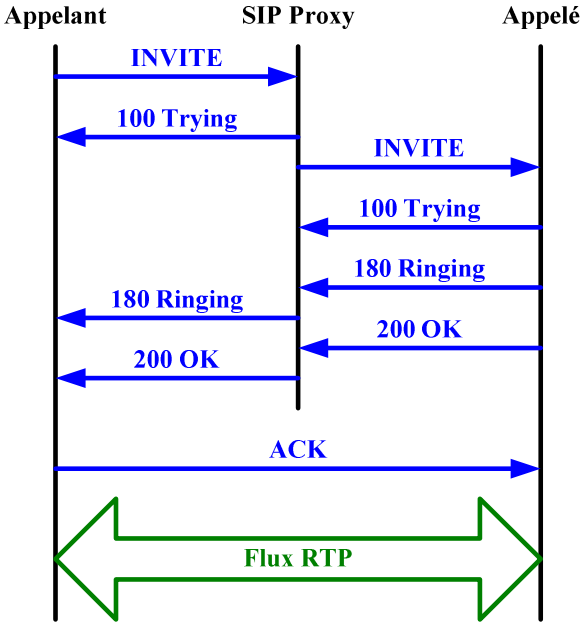
La manière dont sont traitées les requêtes INVITE dépend du type de serveur Proxy utilisé. Un Stateless Proxy ne fera que rediriger les messages reçus vers une destination, alors qu'un Stateful Proxy sera capable de garder l'état des transactions en cours, et par exemple de générer ses propres réponses.

Un Stateful Proxy peut activer le Record Routing, de manière à forcer le passage de tous les messages suivants par le Proxy. Pour cela, le Proxy ajoute un champ « Record-Route » dans l'en-tête SIP dans les requêtes envoyées.

Chaque Proxy utilisant le Record Routing ajoutera son champ « Record-Route ».



Etablissement d'un appel via un Stateless Proxy

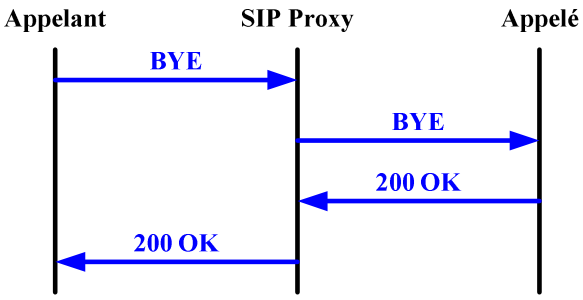


Etablissement d'un appel via un Stateful Proxy (sans Record Routing)

3.5.3. Terminaison de session

La terminaison de l'appel est directement effectuée entre les deux User Agents, sauf si le Record Routing est activé sur le Proxy. Dans ce cas, tous les messages passeront par le Proxy.

Terminaison de session (sans Record Routing)



Terminaison de session (avec Record Routing)

4. Equipements

4.1. Côté abonné

4.1.1. Téléphones IP

Un téléphone IP est un terminal téléphonique qui se connecte à un équipement réseau au lieu d'une prise téléphonique standard.

Ainsi toute communication téléphonique ne circule non plus sur une ligne téléphonique standard mais sur un réseau de données.

Il existe cependant deux types de téléphones IP :

- Les téléphones IP fixes
- Les téléphones IP portables (utilisant les réseaux sans fil)

Cisco IP Phone 7970 / G et Zyxel P2000W

4.1.2. Logiciels de téléphonie IP

Les logiciels de téléphonie IP, appelés aussi « Softphones », permettent de téléphoner via un ordinateur muni d'un casque et d'un microphone comme s'il s'agissait d'un téléphone physique et disposent des mêmes fonctionnalités.

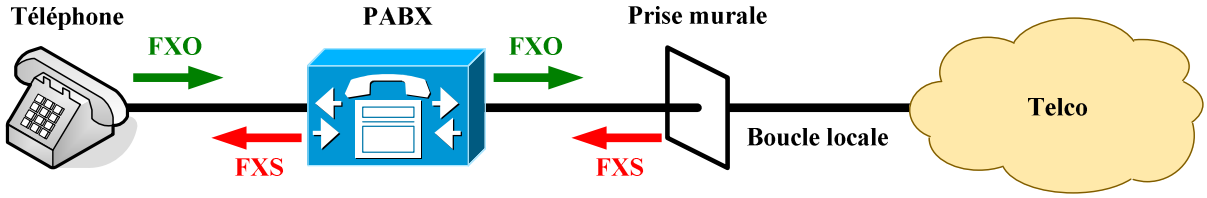
Il existe de nombreux logiciels de téléphonie IP, parmi les plus connus on retrouve : Skype, MSN Messenger, Counterpath eyeBeam, et bien d'autres.

Les fonctionnalités et les protocoles supportés dépendent du logiciel. Il faut donc choisir le logiciel en fonction de la plate-forme utilisée.

CounterPath eyeBeam

4.1.3. ATA (Analog Telephone Adapter)

Les ports FXS (Foreign eXchange Subscriber) et FXO (Foreign eXchange Office) sont les interfaces d'un réseau téléphonique analogique. Le port FXO est l'interface descendante (allant du téléphone vers le PABX par exemple), alors que le port FXS est l'interface remontante (l'interface du PABX allant vers un téléphone).



Le port FXS fournit à l'abonné les services de tonalité et d'alimentation électrique. C'est le port qui va vers l'abonné.

Le port FXO fournit principalement le service de fermeture de la boucle en indiquant si le combiné est raccroché ou pas (on-hook/off-hook). C'est le port qui va vers l'opérateur.

Cisco ATA 186

Linksys PAP2

Afin de permettre un basculement progressif vers une infrastructure VoIP, des adaptateurs de type FXS sont apparus pour interconnecter des téléphones analogiques sur un IPBX, ou inversement de relier une infrastructure téléphonique classique (téléphones et PABX) vers un réseau IP.

Les ports FXO sont généralement situés sur les passerelles IP/TDM.

4.2. Côté opérateur

4.2.1. PABX et IPBX

Un PABX (Private Automatic Branch eXchange) ou PBX (Private Branch eXchange) est un autocommutateur téléphonique. Cet équipement permet l'interconnexion des plusieurs terminaux téléphoniques analogiques d'une entreprise. Il propose de multiples services comme le transfert d'appels et la musique d'attente.

Il existe aussi des PABX virtuels, aussi appelés IP PBX ou IPBX (Intranet Private Branch eXchange), qui sont globalement les équivalents des PABX traditionnels mais pour un usage dédié à la VoIP. Ces IPBX sont utilisés entre autres par les plates-formes IP Centrex.

4.2.2. Passerelle IP/TDM

La grande difficulté avec la VoIP est l'interconnexion avec le réseau téléphonique classique (souvent appelé TDM, pour Time Division Multiplexing).

Des passerelles existent donc pour permettre cette interconnexion IP vers TDM. Des versions logicielles (Asterisk) et des versions matérielles (Cisco AS5x00 ou CIRPACK par exemple) sont disponibles.

Cisco AS5400

CIRPACK MultiNode B

Elles permettent ainsi de transformer les flux IP vers les flux téléphoniques classiques et les rendre cohérents réciproquement.

4.2.3. Serveurs SIP

Les serveurs SIP permettent de centraliser les requêtes des différents téléphones IP afin de pouvoir établir les communications demandées par les utilisateurs. Ils sont d'ailleurs à la base de solutions IP Centrex.

Ces serveurs servent de centrale d'appels (SIP Proxy) et d'annuaire téléphonique (SIP Registar) puisqu'ils ont généralement la liste de toutes les entités reliées au réseau téléphonique IP de l'entreprise.

Ainsi, il va être possible pour un utilisateur, via un URI, de contacter facilement n'importe quelle personne de son entreprise, quel que soit l'emplacement physique de cette personne.

Il existe de nombreux serveurs SIP différents. Dans la catégorie OpenSource gratuit, on retrouve par exemple SER et Asterisk, dont l'utilisation est assez répandue.

5. Infrastructure du réseau supportant la VoIP

5.1. Interconnexion entre la VoIP et la téléphonie classique

Comme vu précédemment, l'interconnexion entre un réseau VoIP et un réseau téléphonique classique se fait par l'intermédiaire d'une passerelle IP/TDM, qui peut être un équipement dédié ou une fonctionnalité incluse dans l'IPBX.

Ces passerelles sont généralement reliées à un opérateur télécom au travers d'un trunk (de technologie variable, allant du simple accès PSTN à une liaison E1 par exemple). Ce trunk n'est autre qu'un tuyau permettant de faire passer un ou plusieurs flux voix vers cet opérateur.

L'avantage des réseaux VoIP est aussi la totale liberté offerte quant au choix des identifiants d'appels. En effet, ces identifiants ne sont pas limités à des numéros de téléphone. Ils peuvent donc être des chaînes alphanumériques presque quelconques. Une interconnexion vers un réseau de téléphonie classique impose donc une restriction de taille, à savoir l'utilisation de numéros de téléphone compatibles et utilisables sur ce dernier.

Par conséquent, l'intégralité des réseaux VoIP interconnectés au réseau téléphonique classique nécessite l'obtention, généralement sous la forme d'un abonnement, de numéros de téléphones auprès d'un opérateur spécialisé.

5.2. Infrastructure LAN

5.2.1. QoS et VLANs

Toutes les technologies utilisant le transport de la voix sur IP nécessitent un minimum de bande passante pour assurer les communications audio établies.

Les réseaux au sein desquels les transferts de données sont trop importants augmentent la latence et diminuent la disponibilité des équipements réseaux, ce qui affaiblit grandement la qualité des communications audio, voir même de les rendre impossibles (pertes de paquets, congestion, délais trop importants, etc.).

C'est pourquoi il devient nécessaire, pour ne pas dire primordial, de différencier les flux audio des flux de données au sein d'un réseau. Pour ce faire, il faut mettre en place des solutions de priorité entre les trafics ainsi que des quotas d'utilisation de la bande passante en fonction de la nature des données.

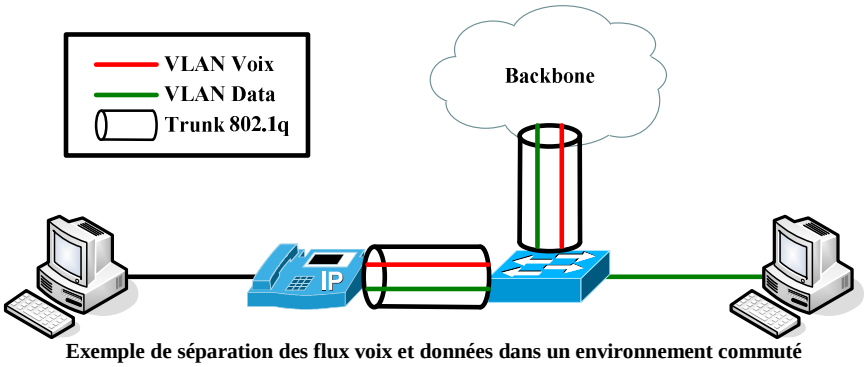
De cette manière, une bande passante minimum pourra être garantie pour les flux audio au sein d'un réseau pour éviter, entre autre, l'indisponibilité du réseau téléphonique IP (élément critique pour certaines entreprises).

Dans cette optique, il faut différencier clairement les flux au sein de l'entreprise, et leur accorder la priorité et la bande passante minimale nécessaire.

Voici un exemple de priorité qui pourrait être appliqué :

Flux	Priorité	Exemple(s)
Convergence du réseau	1	Trames BPDU, mises à jour de routage
Signalisation des appels	2	Paquets SIP
Flux voix	3	Paquets RTP
Données prioritaires	4	Réplication Active Directory Authentification (Kerberos, LDAP, RADIUS)
Données non prioritaires	5	Pages Web, téléchargements

En plus de la QoS, il faut mettre en place une étanchéité entre les trafics voix et données. Cela suppose la création d'au moins deux VLANs : un VLAN Data et un VLAN Voix.



Exemple de séparation des flux voix et données dans un environnement commuté

Le VLAN Voix est utilisé pour toute communication en provenance ou à destination des équipements utilisant des technologies de voix sur IP (signalisation et flux RTP), tandis que le VLAN Data est utilisé pour tout autre type de trafic.

Cela permet de faire une séparation complète des flux et de définir plusieurs classes de service, chacune regroupant des flux de même type (ici tout ce qui est flux audio, et tout le reste). Des priorités de trafics, des quotas de bande passante et des limitations du nombre de requêtes pourront être appliqués sur ces différentes classes de services.

La méthode de priorité généralement retenue pour une implémentation de la VoIP est le LLQ (Low Latency Queuing), vu que cette dernière favorise les trafics sensibles à la latence, comme c'est le cas pour les flux voix.

En ce qui concerne les classes de service, les quotas de bande passante et tout autre caractéristique de la QoS, ils doivent être configurés sur les équipements en prenant en compte l'architecture du réseau et les besoins des utilisateurs sur celui-ci.

En conclusion, les recommandations sont de mettre en place :

- Priorité entre les trafics (LLQ)
- Réservation de bande passante (Traffic Shaping)
- Séparation des trafics avec des VLANs

Le maître-mot ici est donc la disponibilité.

5.2.2. Sécurité

L'infrastructure permettant le bon fonctionnement de toute la téléphonie IP ainsi que tout autre type de communication audio doit absolument rester fiable quoi qu'il arrive.

En effet, cette infrastructure doit être protégée de tout type d'attaque pour éviter l'obtention d'informations confidentielles (espionnage industriel), le détournement de communications téléphoniques ou encore la perte de fonctionnalité de toute cette infrastructure (élément critique dans la majorité des grandes entreprises).

Les éléments à sécuriser au sein d'une telle infrastructure sont nombreux et concernent autant les équipements réseaux, que les terminaux téléphoniques ou encore les systèmes d'exploitation hébergeant aussi bien un serveur SIP qu'un simple softphone.

Pour commencer, il faut penser à sécuriser l'accès aux terminaux et aux différents équipements utilisant la voix sur IP en demandant à ceux-ci ainsi qu'aux utilisateurs de s'authentifier via un mot de passe ou un certificat numérique.

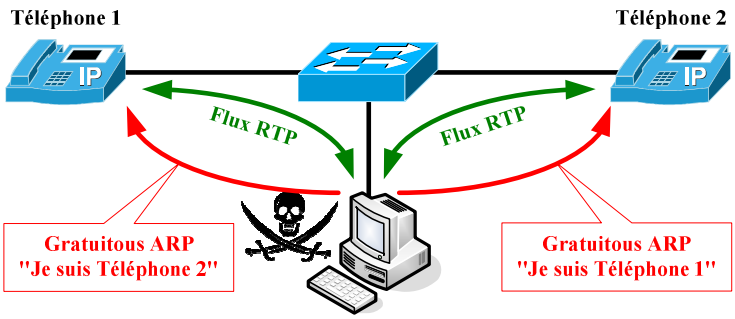
Cependant, cela ne suffit pas puisque une attaque de type « bruteforce » (attaque consistant à essayer tous les mots de passe possibles jusqu'à ce que le bon soit trouvé), ainsi que le vol ou la récupération d'un certificat numérique sont toujours possibles.

Il faut de plus garantir une étanchéité des informations, s'assurer qu'une information circulant dans le VLAN Voix ne puisse pas être accessible depuis un utilisateur d'un autre VLAN. Cependant, il est toujours possible de brancher un ordinateur à la place d'un téléphone IP pour usurper son identité et accéder à ce VLAN Voix.

Cette étanchéité ne s'étend pas qu'aux informations mais aussi aux équipements en charge des communications audio, afin d'éviter toute tentative d'exploitation d'une vulnérabilité ou tout simplement éviter des attaques de type dénis de service (DoS ou DDoS).

En effet, la mise hors service d'un équipement critique de cette infrastructure pourrait entrainer un disfonctionnement total de tous les services de téléphonie sur IP.

Du fait que la majorité des flux audio utilisent le protocole UDP, il est facilement possible de détourner des conversations téléphoniques, usurper une identité, enregistrer ou brouiller une conversation.



Exemple d'attaque de type « man in the middle » avec des Gratuitous ARP

Des attaques de type « man in the middle » utilisant principalement des requêtes ARP formatées d'une manière spécifique (Gratuitous ARP) ou encore l'usurpation d'adresse IP (IP Spoofing) sont utilisées pour détourner ces communications.

Il y aura malheureusement toujours moyen pour une personne mal intentionnée, malgré les sécurités pouvant être mises en place, de pouvoir trouver une faille et accéder aux flux.

C'est pourquoi il faut aussi mettre en place un système de cryptage des données, qui certes va augmenter légèrement la latence des équipements, mais pourra permettre une bien meilleure protection et confidentialité des informations.

Il est ainsi possible d'utiliser le protocole TLS pour sécuriser les requêtes du protocole SIP, ou encore l'utilisation des protocoles SRTP et SRTCP à la place de leurs homologues non sécurisés.

Sécuriser une infrastructure LAN est vraiment une tâche difficile pour un administrateur réseau, aucune mesure ne pourra jamais garantir une sécurité parfaite à 100%, il est seulement possible de sécuriser au mieux une infrastructure pour limiter les risques.

En résumé, les recommandations pouvant être appliquées sont les suivantes :

- **Authentification** : Empêcher une personne non autorisée d'utiliser le service, via l'implémentation de mots de passe ou de certificat numérique (PKI).

- **Etanchéité** : Utilisation de VLANs et filtrage des accès aux applications et équipements critiques.

- **Cryptage** : Utilisation des mécanismes disponibles pour protéger les protocoles de signalisation (TLS pour SIP par exemple) et de transport des flux (SRTP au lieu de RTP)

5.2.3. VoIP et les réseaux sans fil

La voix sur IP a évolué jusqu'à présent au sein d'infrastructures filaires et commence à s'étendre au niveau des réseaux sans fil (VoWiFi, pour Voice over Wi-Fi).

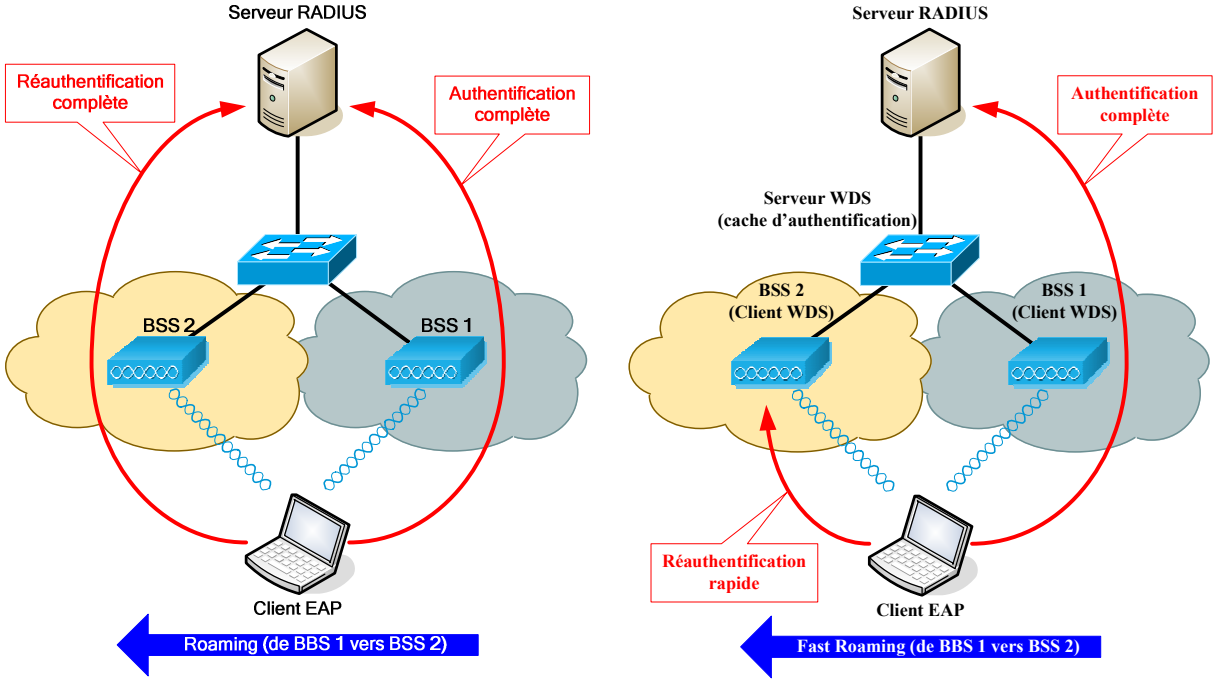
Cette nouvelle tendance apporte tout d'abord un avantage financier. En effet, il n'y a plus qu'à maintenir une seule infrastructure radio ce qui limite grandement le câblage physique des bâtiments.

La VoWiFi a pour vocation l'interconnexion des équipements mobiles tels que les SmartPhones, les PDA ou encore les ordinateurs portables pour leur permettre d'établir des conversations téléphoniques.

Cependant, la VoWiFi est encore sujette à de nombreux problèmes. En effet, il est difficile de pouvoir mettre en place de la QoS sur des points d'accès Wi-Fi et de pouvoir gérer correctement les problèmes engendrés par la concurrence d'accès au niveau des utilisateurs. En effet, nous n'avons que très peu de contrôle sur les ondes et plus particulièrement sur les perturbations pouvant survenir.

De plus, pour qu'une conversation soit considérée de bonne qualité, il ne faut pas dépasser un délai de plus de 150ms ce qui est un véritable problème en termes de mobilité, puisqu'un utilisateur est souvent amené à se connecter dynamiquement à des bornes différentes (principe du roaming) tout en maintenant ses connexions actives.

La latence apportée par le roaming est un réel problème dans la VoWiFi puisque cela dégrade la qualité de la conversation et peut mener à la coupure des communications, si un système d'authentification centralisé est utilisé (méthodes EAP).



C'est pourquoi l'utilisation du Fast Roaming est fortement recommandée au sein des infrastructures Wi-Fi car cette méthode permet à un utilisateur de se réauthentifier plus rapidement sur la nouvelle borne lorsqu'il est amené à se déplacer (infrastructure WDS par exemple).

5.3. Infrastructure WAN

5.3.1. QoS

Lorsqu'une QoS (Quality of Service) est mise en place au sein d'une infrastructure LAN, celle-ci peut être maitrisée jusqu'au routeur frontière.

Au-delà de cette limite, seul un fournisseur d'accès à Internet (FAI ou ISP) a la capacité, bien que souvent partielle, de pouvoir étendre cette QoS au niveau d'un réseau WAN.

Une entreprise est donc contrainte aux solutions que peuvent lui fournir un ISP dans le but de répondre à ses besoins, et doit s'en remettre aux capacités de cet ISP à garantir une qualité de service suffisante pour pouvoir acheminer correctement ses flux à travers un réseau WAN.

Cependant, tous les ISP n'ont pas les mêmes capacités de gestion pour maintenir une qualité de service correspondant aux besoins d'une entreprise. Les opérateurs classiques fournissent généralement une bande passante. Ce sont plutôt les opérateurs Voix qui fournissent un service liant bande passante symétrique et qualité de service pour la ToIP.

Attention, une fois sur le réseau Internet, aucun mécanisme ne permettra de garantir la délivrance optimale des flux multimédias. Il est donc bien souvent impossible de pouvoir assurer une QoS de bout en bout si les réseaux en question ne sont pas gérés par le même fournisseur de services.

Pour pouvoir mettre en place une QoS à travers une infrastructure WAN et ainsi permettre le bon transport de la voix sur IP, il est donc préférable de louer chez un ISP une liaison symétrique dédiée à la voix disposant d'une bande passante suffisante et permettant de garantir correctement la priorité des flux.

Des opérateurs se sont spécialisés dans le transport des flux voix. Ces derniers proposent une qualité de service adaptée aux besoins en VoIP des entreprises. Des offres complètes, incluant l'attribution de numéros de téléphone, existent aussi.

5.3.2. Sécurité

Dans une infrastructure WAN, un ISP doit s'assurer que les informations transmises par ses clients ne puissent pas être détournées ou dérobées par une personne mal intentionnée.

C'est pourquoi il doit mettre en place des solutions permettant l'étanchéité des données circulant à travers ses réseaux ainsi que le cryptage de ces données afin de garantir une confidentialité des informations qui y sont contenues.

Pour ce faire, un ISP utilise généralement des connexions VPN entre les sites afin de pouvoir créer un tunnel dans lequel transiteront toutes les informations de manière cryptée.

De plus, pour permettre de contrôler l'accès à ces informations et accroître la disponibilité de ses services, une entreprise a généralement recours à l'utilisation de deux liaisons différentes : une pour la voix et une autre pour les données.

En revanche l'utilisation de firewall doit être implémentée en prenant en compte les spécificités de la VoIP, sous peine de causer de nombreux problèmes.

Par exemple, suivant la façon dont le firewall a été configuré, les appels devant traverser le firewall pourraient être bloqués. Il faut donc utiliser des firewalls actifs capables de comprendre les protocoles de VoIP et leur interaction, de manière à garantir le passage opportun des flux. Ces firewalls devront donc être capables d'interpréter correctement les informations contenues dans les messages SIP pour permettre les échanges ultérieurs ainsi que les flux RTP associés.

5.3.3. NAT/PAT

Le NAT permet la translation d'adresses privées en adresses publiques pour que des utilisateurs d'un LAN puissent communiquer sur Internet.

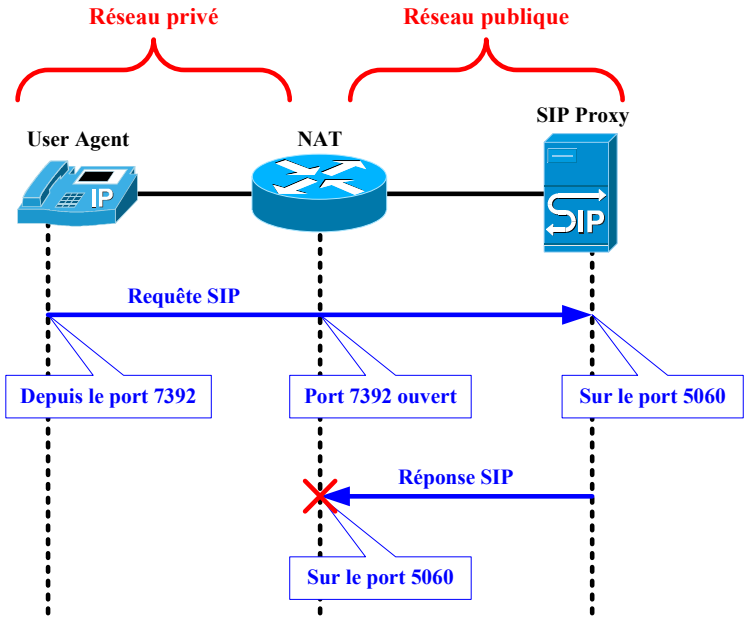
Les requêtes provenant de l'extérieur du réseau pour initialiser une conversation téléphonique avec un téléphone IP du LAN (appels entrants) sont bien souvent bloquées au niveau du routeur gérant la connexion à Internet.

De plus, la translation d'adresse s'effectue uniquement au niveau du paquet IP et non au niveau des en-têtes SIP par exemple. Les messages SIP et les flux RTP sont donc touchés par la problématique du NAT, comme le montre les schémas ci-dessous.

Ici, une requête SIP sur UDP est envoyée par un téléphone à destination d'un serveur Proxy situé derrière un NAT.

La requête passera correctement, et la réponse tentera d'atteindre le téléphone sur le numéro de port par défaut de SIP.

Ce numéro de port n'étant pas ouvert sur le NAT, la réponse ne pourra pas aboutir.

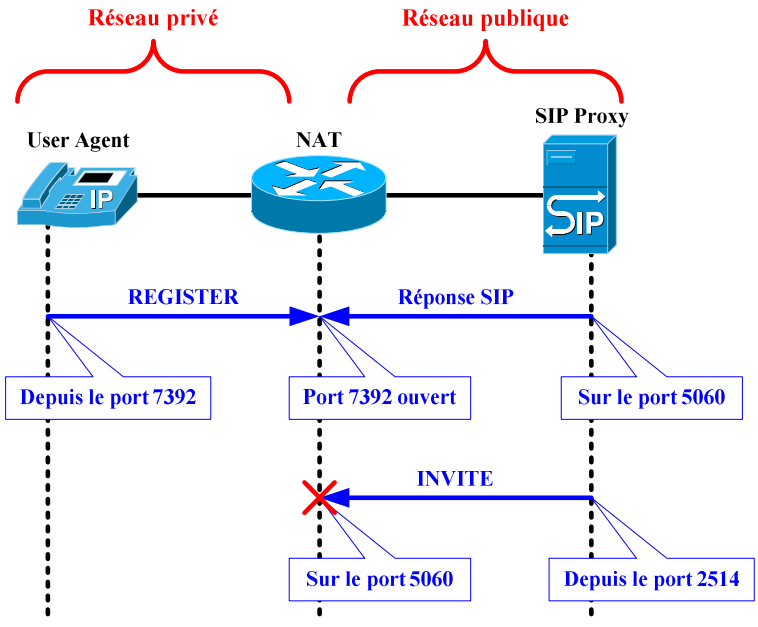


Signalisation SIP sur UDP au travers d'un NAT

Ce contexte présente une requête SIP sur TCP envoyée par un téléphone, à destination d'un serveur Proxy situé derrière un NAT.

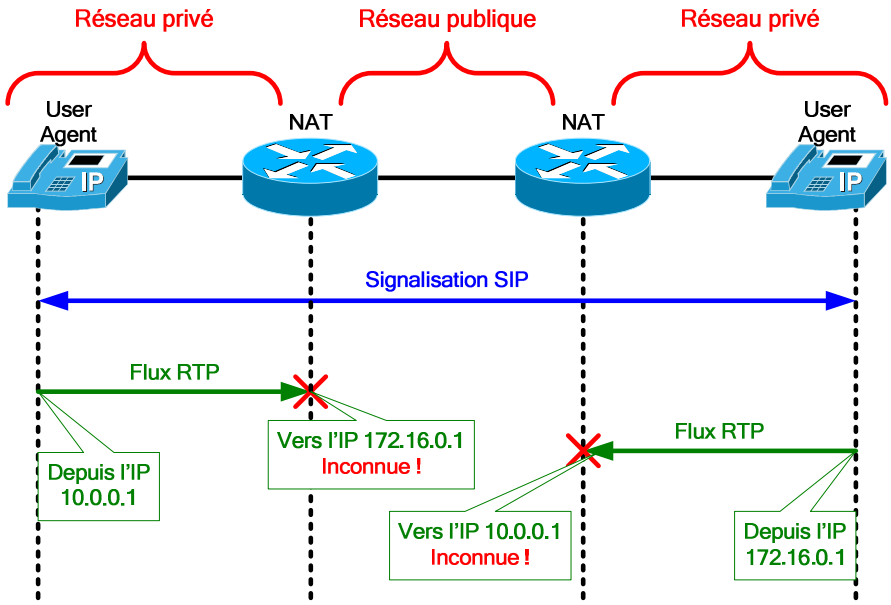
La connexion TCP fait que les réponses à une requête pourraient passer sur la même connexion. Il n'y aurait donc pas de problème.

Néanmoins, si une requête venait à être initiée depuis le réseau public, par l'intermédiaire du serveur Proxy, cette requête se retrouverait bloquée par le NAT, car le numéro de port par défaut de SIP ne serait pas ouvert.



Signalisation SIP sur TCP au travers d'un NAT

Dans le contexte idyllique où la signalisation SIP pourrait fonctionner, le problème de la traversée d'un NAT se pose aussi pour les flux RTP, qui vont utiliser les adresses IP et numéros de port UDP indiqués par SDP. Les téléphones tenteraient donc d'initier un flux RTP sur des adresses invalides sur le réseau public.



Flux RTP sur UDP au travers d'un NAT

Rien n'est actuellement ratifié, mais plusieurs solutions existent :

- Signalisation SIP
  - Réponse symétrique
  - Réutilisation des connexions
- Flux RTP
  - RTP symétrique
  - STUN (Simple Traversal of UDP through NAT), défini par la RFC 3489
  - TURN (Traversal Using Relay NAT)
  - ICE (Interactive Connectivity Establishment)

Ces solutions sont expliquées dans le draft IETF suivant : http://tools.ietf.org/html/draft-ietf-sipping-nat-scenarios-05 (Juin 2006).

5.3.4. Fiabilité et disponibilité des liaisons WAN

Contrairement à une infrastructure LAN où tous les aspects peuvent être maitrisés par l'administrateur réseau, il faut espérer qu'un ISP tienne ses engagements de qualités en termes de délai et de fiabilité de ses liaisons WAN.

Pour entretenir une communication VoIP de bonne qualité, il faut que le délai reste inférieur ou égal à 150 ms, restriction particulièrement difficile si l'on passe par des liaisons satellites par exemple.

Il faut donc s'assurer que l'ISP puisse fournir une liaison adaptée aux besoins de son client afin de garantir des délais convenables et une disponibilité continue de cette liaison.

Pour connaître la bande passante dont peut avoir besoin une entreprise au niveau de la VoIP, il suffit d'analyser la bande passante que génère une communication et de regarder combien de flux peuvent ainsi passer en même temps sur une liaison.

Il faut aussi prévoir un chemin de secours pour les flux, en cas de panne de la liaison WAN.

5.3.5. Implémentation sur différents médias et technologies WAN

Chaque liaison WAN à sa propre capacité en termes de bande passante, certaines sont symétriques, d'autres asymétriques (capacité d'émission différente de celle de réception).

Parmi les technologies WAN existantes, celles qui sont les plus adaptées à la VoIP sont celles qui offrent à l'utilisateur le moins de latence possible et une bande passante suffisante aussi bien au niveau de l'émission que de la réception.

C'est pourquoi, les liaisons symétriques à haut débit telles que les LS, SDSL ou PRI (E1, T2, etc.) sont fortement recommandées pour l'utilisation d'une infrastructure VoIP.

Certaines liaisons ne répondent pas aux besoins de la transmission de la voix sur IP. Les deux problèmes les plus fréquemment rencontrés sont :

- **La bande passante** : Avec liaisons asymétriques comme le RTC, le RNIS BRI (Numéris) ou encore l'ADSL.
- **La latence réseau** : Fortement visible sur des liaisons satellite par exemple.