Voix sur IP - VOIP

1 - Introduction

La voix sur IP (Voice over IP) est une technologie de communication vocale en pleine émergence. Elle fait partie d'un tournant dans le monde de la communication. En effet, la convergence du triple play (voix, données et vidéo) fait partie des enjeux principaux des acteurs de la télécommunication aujourd'hui. Plus récemment l'Internet s'est étendu partiellement dans l'Intranet de chaque organisation, voyant le trafic total basé sur un transport réseau de paquets IP surpasser le trafic traditionnel du réseau voix (réseau à commutation de circuits). Il devenait clair que dans le sillage de cette avancée technologique, les opérateurs, entreprises ou organisations et fournisseurs devaient, pour bénéficier de l'avantage du transport unique IP, introduire de nouveaux services voix et vidéo. Ce fût en 1996 la naissance de la première version voix sur IP appelée H323. Issu de l'organisation de standardisation européenne ITU-T sur la base de la signalisation voix RNIS (Q931), ce standard a maintenant donné suite à de nombreuses évolutions, quelques nouveaux standards

prenant d’autres orientations technologiques.

Pour être plus précis et néanmoins schématique, le signal numérique obtenu par numérisation de la voix est découpé en paquets qui sont transmis sur un réseau IP vers une application qui se chargera de la transformation inverse (des paquets vers la voix). Au lieu de disposer à la fois d'un réseau informatique et d'un réseau téléphonique commuté (RTC), l'entreprise peux donc, grâce à la VoIP, tout fusionner sur un même réseau. Ca par du fait que la téléphonie devient de la "data". Les nouvelles capacités des réseaux à haut débit devraient permettre de transférer de manière fiable des données en temps réel. Ainsi, les applications de vidéo ou audioconférence ou de téléphonie vont envahir le monde IP qui, jusqu’alors, ne pouvait raisonnablement pas supporter ce genre d’applications (temps de réponse important, jigue-jitter, Cos-Qos…). Jusque vers le milieu des années 90, les organismes de normalisation ont tenté de transmettre les données de manière toujours plus efficace sur des réseaux conçus pour la téléphonie. A partir de cette date, il y a eu changement. C'est sur les réseaux de données, que l'on s'est évertué à convoyer la parole. Il a donc fallu développer des algorithmes de codage audio plus tolérants et introduire des mécanismes de contrôle de la qualité de service dans les réseaux de données. Faire basculer différents types de données sur un même réseau permet en plus, de simplifier son administration.

Comme toute innovation technologique qui se respecte, la VoIP doit non seulement simplifier le travail mais aussi faire économiser de l'argent. Les entreprises dépensent énormément en communications téléphoniques, or le prix des communications de la Toip (Téléphonie sur Ip) est dérisoire en comparaison. En particulier, plus les interlocuteurs sont éloignés, plus la différence de prix est intéressante. De plus, la téléphonie sur IP utilise jusqu'à dix fois moins de bande passante que la téléphonie traditionnelle. Ceci apportant de grand intérêt pour la voix sur réseau privée. Il semblerait que les entreprises après avoir émis un certain nombre de doutes sur la qualité de services soient désormais convaincues de la plus grande maturité technologique des solutions proposées sur le marché. Qu’il s’agisse d’entreprises mono-site ou multisites, les sondages montrent que le phénomène de migration vers les systèmes de téléphonie sur IP en entreprise est actuellement

engagé.

Les premières technologies de VoIP imaginées étaient propriétaires et donc très différentes les unes des autres. Pourtant, un système qui est censé mettre des gens et des systèmes en relation exige une certaine dose de standardisation. C'est pourquoi sont apparus des protocoles standards, comme le H323 ou le SIP.

2 - Le Réseau Téléphonique Commuté

Mais qu'est-ce que le RTC? Le RTC est tout simplement le réseau téléphonique que nous utilisons dans notre vie de tous les jours et qui nous donne accès à de multiple fonction. En effet outre le fait de pouvoir téléphoner, le RTC nous permet d'utiliser de multiples services tel que la transmission et réception de fax, l'utilisation d'un minitel, accéder à Internet etc… Il représente donc l'un des protocole de discussion utilisé sur la paire de cuivre boucle locale.

2.1 - Histoire de la téléphonie

Du premier télégraphe de Chappe en 1790 au RTC actuelle, l'histoire des communications à connu de grands moments et de grandes avancés dû à l'ingéniosité de certains et aux progrès technologique et électronique. Nous retiendrons quelques grandes dates tel que :

1837 Premier télégraphe électrique inventé par Samuel Morse

1889 Almon B. Strowger (USA) invente le premier « sélecteur » automatique et donne ainsi naissance à la commutation téléphonique automatique

1938 Alec Reeves (Français) dépose le brevet des futurs systèmes à modulation par impulsion et codage (MIC) : quantification et échantillonnage du signal à intervalles réguliers, puis codage sous forme binaire.

1962 Les premiers systèmes de transmission multiplex de type MIC apparaissent aux Etats-Unis ils permettent une liaison à 24 voies entre centraux téléphonique, à la même époque en France on installe des MIC à 32 voies.

1970 Un nouveau pas est franchi dans le domaine de la commutation électronique avec la mise en service en France, par le CNET, des premiers centraux téléphoniques publics en commutation électronique temporelle.

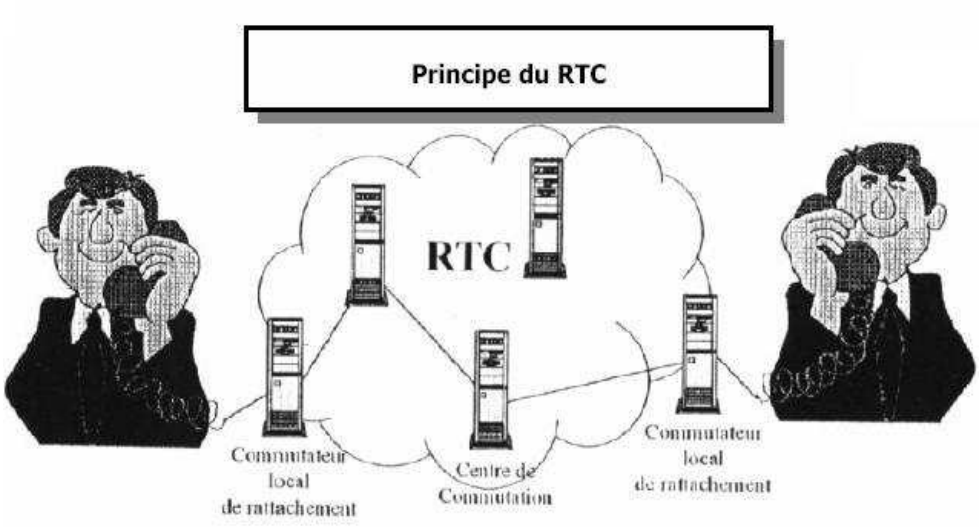
1979 Lancement du minitel en France

1987 Le RNIS est mis en service en France.

1990 De nouveaux concepts apparaissent tel que la commutation temporelle asynchrone (ATM) et la hiérarchie numérique synchrone.

2.2 - Principe du Rtc

Le réseau téléphonique public (RTPC, Réseau Téléphonique Public Commuté ou simplement RTC) a essentiellement pour objet le transfert de la voix. Le transport des données n'y est autorisé, en France, que depuis 1964. Utilisant le principe de la commutation de circuits, il met en relation deux abonnés à travers une liaison dédiée pendant tout l'échange.



On distingue deux grandes parties dans ce réseau :

Le réseau capillaire ou de distribution, c'est le raccordement depuis chez l'abonné à un point d'entrée du réseau. Cette partie du réseau est analogique.

Le réseau de transit, effectue pour sa part le transport des communications entre les nœuds de transit concentrateurs / commutateurs). Cette portion du réseau est actuellement numérique.

La numérisation offre plusieurs avantages. Puisqu'il ne s'agit que de 0 et de 1, la qualité du signal est préservée, quelle que soit la distance entre les convertisseurs (analogique numérique et numérique analogique). Ce n'est pas le cas des communications analogiques où le signal est pollué à chaque manipulation.

La gestion générale du réseau discerne trois fonctions :

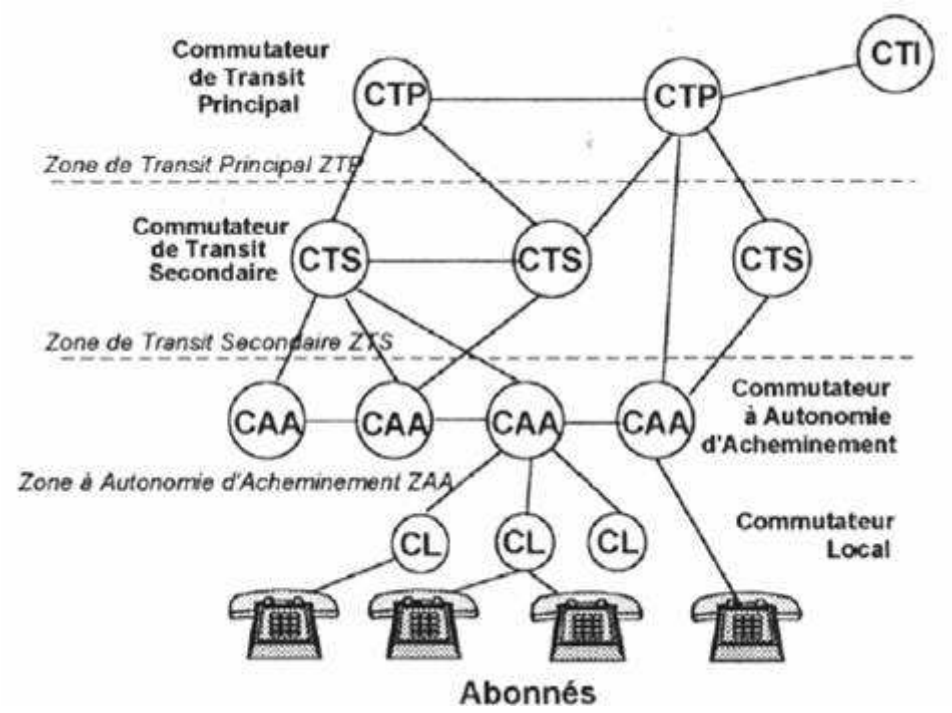
La distribution, celle-ci comprend essentiellement la liaison d'abonné ou boucle locale (paire métallique torsadée) qui relie l'installation de l'abonné au centre de transmission de rattachement. Cette ligne assure la transmission de la voix (fréquence vocale de 300 à 3 400 Hz), de la numérotation (10 Hz pouf la numérotation décimale -au cadran- et 697 à 1633 Hz pour la numérotation fréquentielle) et de la signalisation générale (boucle de courant, fréquences supra vocales)

La commutation, c'est la fonction essentielle du réseau, elle consiste à mettre en relation deux abonnés, maintenir la liaison pendant tout l'échange et libérer les ressources à la fin de celui-ci. C'est le réseau qui détermine les paramètres de taxation et impute le coût de la communication à l'appelant

La transmission, c'est la partie support de télécommunication du réseau, cette fonction est remplie soit par un système filaire cuivre (en voie de disparition), de la fibre optique ou des faisceaux hertziens. Aujourd'hui, le réseau est pratiquement intégralement numérisé, seule la liaison d'abonné reste analogique.

2.3 - Architecture du réseau

Le réseau téléphonique commuté a une organisation hiérarchique à trois niveaux. Il est structuré en zones correspondant à un niveau de concentration.



On distingue :

Zone à Autonomie d'Acheminement (ZAA), cette zone, la plus basse de la hiérarchie, comporte un ou plusieurs Commutateurs à Autonomie d'Acheminement (CAA) qui eux-mêmes desservent des Commutateurs Locaux (CL). Les commutateurs locaux ne sont que de simples concentrateurs de lignes auxquels sont raccordés les abonnés finals. La ZAA (Zone à Autonomie d'Acheminement) est un réseau étoilé, elle constitue le réseau de desserte;

Zone de Transit Secondaire (ZTS), cette zone comporte des Commutateurs de Transit Secondaires (CTS). Il n'y a pas d'abonnés reliés aux CTS (Commutateurs de Transit Secondaires). Ils assurent le brassage des circuits lorsqu'un CAA (Commutateur à Autonomie d'Acheminement) ne peut atteindre le CAA destinataire directement (réseau imparfaitement maillé);

Zone de Transit Principal (ZTP), cette zone assure la commutation des liaisons longues distances. Chaque ZTP (Zone de Transit Principal) comprend un Commutateur de Transit Principal (CTP), L'un des commutateurs de transit principal (CTP) est relié au commutateur international de transit.

3 - Les enjeux de la téléphonie sur Ip

Dans cette partie, nous allons voir pourquoi la téléphonie IP est devenue importante pour les entreprises. L'enjeu est de réussir à faire converger le réseau de donnée IP et le réseau téléphonique actuel. Voici les principales motivations pour déployer la téléphonie sur IP (Source Sage Research 2003, sondage auprès de 100 décisionnaires IT).

Motivations	Pourcentage
Réduction de coûts	75 %
Nécessité de standardiser l'équipement	66 %
Hausse de la productivité des employés	65 %
Autres bénéfices de productivité	64 %
Hausse du volume d'appels à traiter	46 %
Autres facteurs	50 %

La téléphonie sur IP exploite un réseau de données IP pour offrir des communications vocales à l'ensemble de l'entreprise sur un réseau unique voix et données. Cette convergence des services de communication données, voix, et vidéo sur un réseau unique, s'accompagne des avantages liés à la réduction des coûts d'investissement, à la simplification des procédures d'assistance et de configuration, et à l'intégration accrue de filiales et de sites distants aux installations du réseau d'entreprise.

Les coûts généraux de l'infrastructure de réseau sont réduits. Le déploiement d'un unique réseau convergé voix et données sur tous les sites permet de réaliser des économies sur les investissements productifs, l'ordre d'idée en 2004-2005 atteint les 50% si l'on prend on compte les communications inter-site. De plus, comme le téléphone et le PC partagent le même câble Ethernet, les frais de câblage sont réduits. Les frais d'administration du réseau sont également minimisés. Il est ainsi

possible de réaliser des économies à court et à long terme sur de nombreux postes : administration d'un seul réseau, fournisseur d'accès unique, unique contrat de maintenance, câblage commun, gratuité des communications interurbaines, réduction de la complexité de l'intégration d'applications. Enfin, la migration de la solution actuelle vers la Téléphonie sur IP s'effectue en douceur. Les solutions de téléphonie sur Ip sont conçues pour dégager une stratégie de migration à faible risque à partir de l'infrastructure existante.

Le scénario vers lequel va s'orienter la téléphonie sur Ip dépend beaucoup de l'évolution du réseau lui-même. En effet, si Internet reste à peu près dans sa configuration actuelle où il est essentiellement dimensionné en fonction d'une qualité de service moyenne pour la transmission des données, il est fort probable que la téléphonie sur Ip restera un marché réservé au réseau de type Frame, Mpls. Les seules exceptions seraient alors les cas d'interconnexion de PBX d'entreprises, commerce électronique, applications nouvelles associant la voix pour une véritable utilisation multimédia d'Internet. En effet, ce qui ralenti considérablement l'explosion de ce secteur est le fait qu'il y ait encore trop peu de déploiements opérationnels en France et même dans le monde. De nombreuses entreprises connaissent la téléphonie sur IP, mais toutes en sont au même stade : le test. De plus, il faut savoir que la plupart des déploiements opérationnels de téléphonie sur IP ont été réalisés pour des universités, or, les universités n'ayant pas les mêmes exigences qu'une entreprise, ces déploiements ne sont pas réellement pris en compte.

Les applications et les services Ip intégrés améliorent la productivité et le soin de la clientèle. Les bénéfices récurrents seront apportés par les gains de productivité liés à l'utilisation de nouveaux services et de nouveaux applicatifs tels que la messagerie unifiée qui permettent de libérer, selon les spécificités des métiers, entre 25 et 40 minutes de temps de travail par collaborateur, les assistants personnels qui permettent au collaborateur de personnaliser sur l'Intranet toutes les fonctions avancées de renvoi d'appel en fonction de son agenda propre ou partagé et les applications « d'eLearning », qu'il convient de faire apparaître dans une démarche de démonstration de retour sur l'investissement à court et moyen terme. De plus, les fonctions simplifiées de création, de déplacement et de modification réduisent le

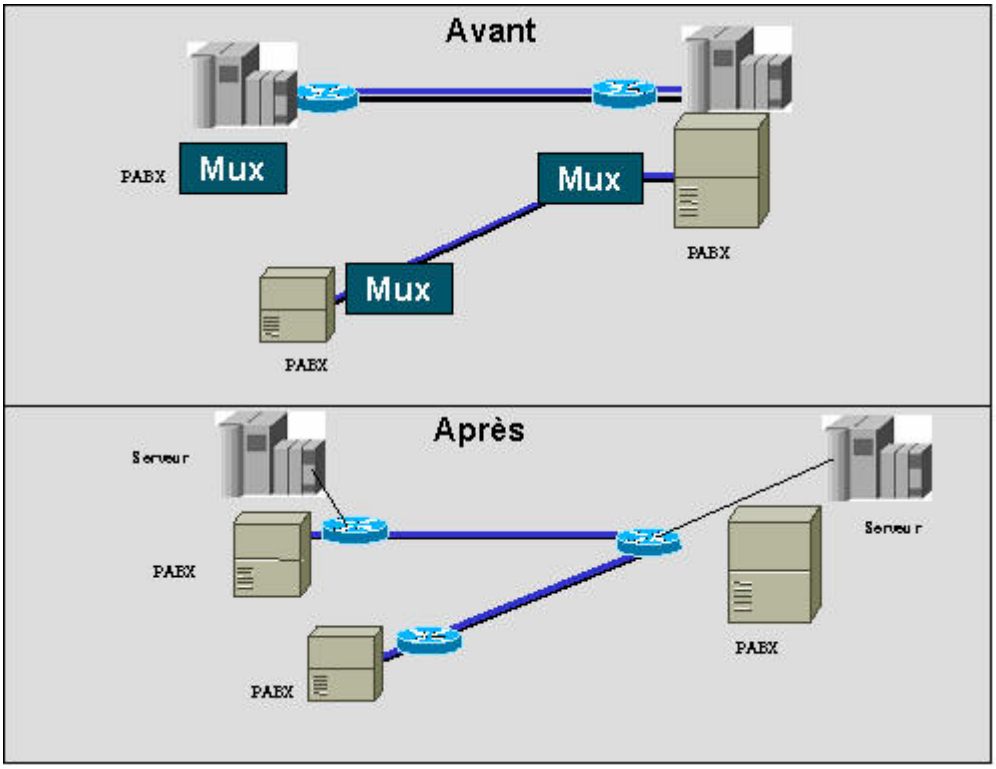
temps nécessaire pour ajouter de nouveaux utilisateurs au réseau. Le déploiement de nouveaux services est accéléré. L'utilisation d'une infrastructure IP commune et d'interfaces standard ouvertes permet de développer et de déployer très rapidement des applications innovantes. Enfin, les utilisateurs accèdent à tous les services du réseau partout où ils peuvent s'y connecter notamment à travers l'extention mobility (substitution de postes).

4 - Les avantages

La VoIP offre de nombreuses nouvelles possibilités aux opérateurs et utilisateurs qui bénéficient d'un réseau basé sur Ip. Les avantages les plus marqués sont les suivants.

4.1 - Réduction des coûts

En déplaçant le trafic voix Rtc vers le réseau privé WAN/IP les entreprises peuvent réduire sensiblement certains coûts de communications. Réductions importantes mises en évidence pour des communications internationales, ces réductions deviennent encore plus intéressantes dans la mutualisation voix/données du réseau IP inter-sites (WAN). Dans ce dernier cas, le gain est directement proportionnel au nombre de sites distants.



4.2 - Standards ouverts et interopérabilité multi-fournisseurs

Trop souvent par le passé les utilisateurs étaient prisonniers d'un choix technologique antérieur. La VoIP a maintenant prouvé tant au niveau des réseaux opérateurs que des réseaux d'entreprises que les choix et les évolutions deviennent moins dépendants de l'existant.

Contrairement à nos convictions du début, nous savons maintenant que le monde VoIP ne sera pas uniquement H323, mais un usage multi-protocoles selon les besoins de services nécessaires. Par exemple, H323 fonctionne en mode " peer to peer " alors que MGCP fonctionne en mode centralisé. Ces différences de conception offrent immédiatement une différence dans l'exploitation des terminaisons considérées.

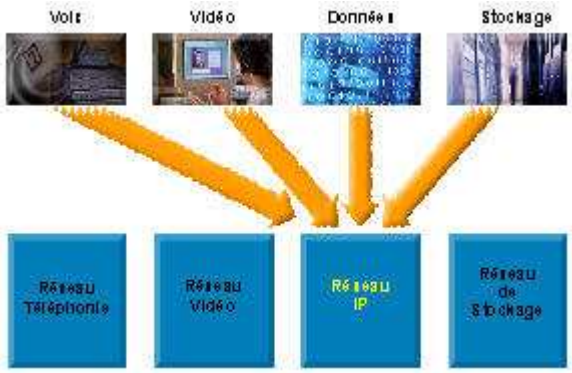
4.3 - Choix d'un service opéré

Les services opérateurs ouvrent les alternatives VoIP. Non seulement l'entreprise peut opérer son réseau privé VoIP en extension du réseau RTC opérateur, mais l'opérateur lui-même ouvre de nouveaux services de transport VoIP qui simplifient le

nombre d'accès locaux à un site et réduit les coûts induits. Le plus souvent les entreprises opérant des réseaux multi-sites louent une liaison privée pour la voix et une pour la donnée, en conservant les connexions RTC d'accès local. Les nouvelles offres VoIP opérateurs permettent outre les accès RTC locaux, de souscrire uniquement le média VoIP inter-sites.

4.4 - Un réseau voix, vidéo et données (triple play)

En positionnant la voix comme une application supplémentaire du réseau IP, l'entreprise ne va pas uniquement substituer un transport opérateur RTC à un transport IP, mais simplifier la gestion des trois réseaux (voix, données et vidéo) par ce seul transport. Une simplification de gestion, mais également une mutualisation des efforts financiers vers un seul outil. Concentrer cet effort permet de bénéficier d'un réseau de meilleure qualité, plus facilement évolutif et plus disponible, pourvu que la bande passante du réseau concentrant la voix, la vidéo et les données soit dimensionnée en conséquence.



4.5 - Un service PABX distribué ou centralisé

Les PABX en réseau bénéficient de services centralisés tel que la messagerie vocale, la taxation, etc… Cette même centralisation continue à être assurée sur un réseau VoIP sans limitation du nombre de canaux. A l'inverse, un certain nombre de services sont parfois souhaités dans un mode de décentralisation. C'est le cas du centre d'appels où le besoin est une centralisation du numéro d'appel (ex : numéro vert), et une décentralisation des agents du centre d'appel. Difficile à effectuer en téléphonie traditionnelle sans l'utilisation d'un réseau IP pour le déport de la gestion des ACD distants. Il est ainsi très facile de constituer un centre d'appel ou centre de

contacts (multi canaux/multi-médias) virtuel qui possède une centralisation de supervision et d'informations.

Il convient pour en assurer une bonne utilisation de dimensionner convenablement le lien réseau. L'utilisation de la VoIP met en commun un média qui peut à la fois offrir à un moment précis une bande passante maximum à la donnée, et dans une autre période une bande passante maximum à la voix, garantissant toujours la priorité à celle-ci.

4.6 - Evolution vers un réseau de téléphonie sur Ip

La téléphonie sur IP repose totalement sur un transport VoIP. La mise en œuvre de la VoIP offre là une première brique de migration vers la téléphonie sur IP.

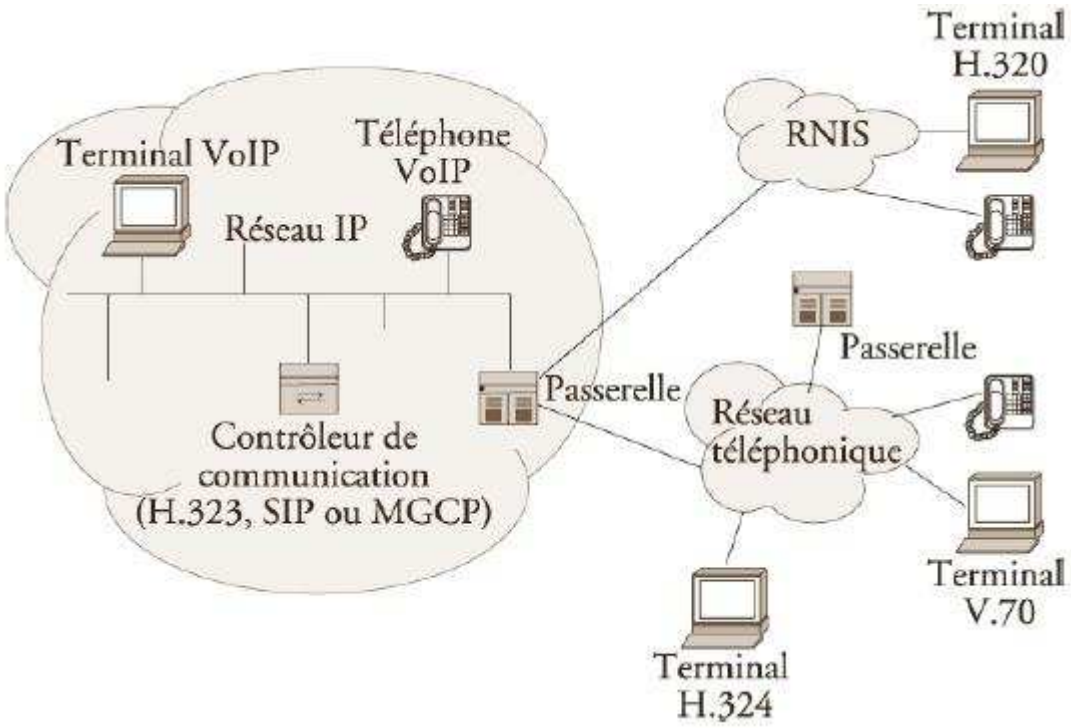
4.7 - Intégration des services vidéo

La VoIP intègre une gestion de la voix mais également une gestion de la vidéo. Si nous excluons la configuration des " multicasts " sur les composants du réseau, le réseau VoIP peut accueillir des applications vidéo de type vidéo conférence, vidéo surveillance, e-learning, vidéo on demand,…, pour l'ensemble des utilisateurs à un coût d'infrastructure réseau supplémentaire minime.

5 - L'Architecture Voip

5.1 - Les schémas

Voici le schéma générale de l'utilisation de la Voip en entreprise :



La VoIP étant une nouvelle technologie de communication, elle n'a pas encore de standard unique. En effet, chaque constructeur apporte ses normes et ses fonctionnalités à ses solutions. Il existe tout de même des références en la matière. Je vais décrire les trois principales que sont H.323, SIP et MGCP/MEGACO. Tous les acteurs de ce marché utilisent comme base pour leur produit une ou plusieurs de ces trois architectures. Il existe donc plusieurs approches pour offrir des services de téléphonie et de visiophonie sur des réseaux IP. Certaines placent l'intelligence dans le réseau alors que d'autres préfèrent une approche peer to peer avec l'intelligence répartie à la périphérie (terminal de téléphonie IP, passerelle avec le réseau téléphonique commuté...). Chacune a ses avantages et ses inconvénients.

Le schéma ci-dessus, décrit de façon générale la topologie d'un réseau de téléphonie IP. Elle comprend toujours des terminaux, un serveur de communication et une passerelle vers les autres réseaux. Chaque norme a ensuite ses propres caractéristiques pour garantir une plus ou moins grande qualité de service. L'intelligence du réseau est aussi déportée soit sur les terminaux, soit sur les passerelles/Gatekeeper (contrôleur de commutation). On retrouve les éléments communs suivants :

Le routeur : Il permet d'aiguiller les données et le routage des paquets entre deux réseaux. Certains routeurs, comme les Cisco 2600, permettent de simuler un gatekeeper grâce à l'ajout de cartes spécialisées supportant les protocoles VoIP.

La passerelle : il s'agit d'une interface entre le réseau commuté et le réseau IP.

Le PABX : C'est le commutateur du réseau téléphonique classique. Il permet de faire le lien entre la passerelle ou le routeur et le réseau RTC. Une mise à jour du PABX est aussi nécessaire. Si tout le réseau devient IP, il n'y a plus besoin de ce matériel.

Les Terminaux : Des PC ou des téléphones VoIP.

Pour transmettre les paquets, on utilise RTP, standardisé en 1996. Il est un protocole adapté aux applications présentant des propriétés temps réel. Il permet ainsi de reconstituer la base de temps des flux (horodatage des paquets : possibilité de re-synchronisation des flux par le récepteur), de détecter les pertes de paquets et en informer la source, et d'identifier le contenu des données pour leurs associer un transport sécurisé. En revanche, ce n'est pas "la solution" qui permettrait d'obtenir des transmissions temps réel sur IP. En effet, il ne procure pas de réservation de ressources sur le réseau (pas d'action sur le réseau de type RSVP, diffserv, Policeur), de fiabilisation des échanges (pas de retransmission automatique, pas de régulation automatique du débit) et de garantie dans le délai de livraison (seules les couches de niveau inférieur le peuvent) et dans la continuité du flux temps réel. Bien qu'autonome, RTP peut être complété par RTCP. Ce dernier apporte un retour d'informations sur la transmission et sur les éléments destinataires. Ce protocole de contrôle permet de renvoyer à la source des informations sur les récepteurs et ainsi lui permettre, par exemple, d'adapter un type de codage ou encore de modifier le débit des données.

5.2 - Gateway et Gatekeeper

Pour commencer je vais parler d'un des éléments clefs d'un réseau VoIP, la passerelle et leurs « Gatekeepers » associés. Les passerelles ou gateways en téléphonie IP sont des ordinateurs qui fournissent une interface où se fait la convergence entre les réseaux téléphoniques commutés (RTC) et les réseaux basés

sur la commutation de paquets TCP/IP. C'est une partie essentielle de l'architecture du réseau de téléphonie IP. Le gatekeeper est l'élément qui fournit de l'intelligence à la passerelle. Comme nous l'avons déjà fait remarqué, nous pouvons séparer les parties matérielles et logicielles d'une passerelle. Le gatekeeper est le compagnon logiciel de la gateway.

Une gateway permet aux terminaux d'opérer en environnements hétérogènes. Ces environnements peuvent être très différents, utilisant diverses technologies tels que le Numéris, la téléphonie commutée ou la téléphonie IP. Les gateways doivent aussi être compatible avec les terminaux téléphoniques analogiques. La gateway fournit la possibilité d'établir une connexion entre un terminal analogique et un terminal multimédia (un PC en général). Beaucoup de sociétés fournissent des passerelles mais cela ne signifie pas qu'elles fournissent le même service. Les gateways (partie physique) et les gatekeepers (partie logicielle) font l'objet de deux sections séparées pour bien cerner la différence. Certaines sociétés vendent un produit " gateway ", mais en réalité, elles incorporent une autre gateway du marché avec leur gatekeeper pour proposer une solution commerciale. La plus-value ne se fait pas sur la gateway mais sur le gatekeeper car c'est sur celui-ci qu'on peut faire la différence.

Un gatekeeper deux services principaux : la gestion des permission et la résolution d'adresses. La gatekeeper est aussi responsable de la sécurité. Quand un client veut émettre un appel, il doit le faire au travers du gatekeeper. C'est alors que celui-ci fournit une résolution d'adresse du client de destination. Dans le cas où il y a plusieurs gateways sur le réseau, il peut rediriger l'appel vers un autre couple gateway/gatekeeper qui essaiera à son tour de router l'appel. Pendant la résolution d'adresse, le gatekeeper peut aussi attribuer une certaine quantité de bande passante pour l'appel. Il peut agir comme un administrateur de la bande passant disponible sur le réseau. Le gatekeeper répond aux aspects suivant de la téléphonie IP :

Le routage des appels : en effet, le gatekeeper est responsable de la fonction de routage. Non seulement, il doit tester si l'appel est permis et faire la résolution d'adresse mais il doit aussi rediriger l'appel vers le bon client ou la bonne passerelle.

Administration de la bande passante : le gatekeeper alloue une certaine quantité de bande passant pour un appel et sélectionne les codecs à utiliser. Il agit en tant que régulateur de la bande passante pour prémunir le réseau contre les goulots d'étranglement (bottle-neck).

Tolérance aux fautes, sécurité : le gatekeeper est aussi responsable de la sécurité dans un réseau de téléphonie IP. Il doit gérer les redondances des passerelles afin de faire aboutir tout appel. Il connaît à tout moment l'état de chaque passerelle et route les appels vers les passerelles accessibles et qui ont des ports libres.

Gestion des différentes gateways : dans un réseau de téléphonie IP, il peut y avoir beaucoup de gateways. Le gatekeeper, de par ses fonctionnalités de routage et de sécurité, doit gérer ces gateways pour faire en sorte que tout appel atteigne sa destination avec la meilleure qualité de service possible.

Ainsi, le gatekeeper peut remplacer le classique PABX. En effet, il est capable de router les appels entrant et de les rediriger vers leur destination ou une autre passerelle. Mais il peut gérer bien d'autres fonctions telles que la conférence ou le double appel. Il n'existe pas les mêmes contraintes avec un gatekeeper qu'avec un PABX. En effet, ce dernier est constitué par du logiciel et l'opérateur peut implémenter autant de services qu'il le désire. Alors qu'avec un PABX, l'évolutivité est limitée par le matériel propriétaire de chaque constructeur, avec le gatekeeper, l'amélioration des services d'un réseau de téléphonie IP n'a pas de limites. Le grand bénéfice du développement d'un gros gatekeeper est de remplacer le PABX classique. En effet, chaque PABX utilise son propre protocole pour communiquer avec les postes clients, ce qui entraîne un surcoût. Avec le couple gateway/gatekeeper, ce problème n'existe pas. Il utilise des infrastructures qui existent, le LAN et des protocoles tel qu'IP.

6 - Standards VoIP

6.1 - Protocole H323

6.1.1 - Introduction

Avec le développement du multimédia sur les réseaux, il est devenu nécessaire de créer des protocoles qui supportent ces nouvelles fonctionnalités, telles que la visioconférence : l'envoi de son et de vidéo avec un soucis de données temps réel. Le protocole H.323 est l'un d'eux. Il permet de faire de la visioconférence sur des réseaux IP.

H.323 est un protocole de communication englobant un ensemble de normes utilisés pour l'envoi de données audio et vidéo sur Internet. Il existe depuis 1996 et a été initié par l'ITU (International Communication Union), un groupe international de téléphonie qui développe des standards de communication. Concrètement, il est utilisé dans des programmes tels que Microsoft Netmeeting, ou encore dans des équipements tels que les routeurs Cisco. Il existe un projet OpenH.323 qui développe un client H.323 en logiciel libre afin que les utilisateurs et les petites entreprises puissent avoir accès à ce protocole sans avoir à débourser beaucoup d'argent.

6.1.2 - Fonctionnement

Le protocole H.323 est utilisé pour l'interactivité en temps réel, notamment la visioconférence (signalisation, enregistrement, contrôle d'admission, transport et encodage). C'est le leader du marché pour la téléphonie Ip. Il s'inspire du protocole H.320 qui proposait une solution pour la visioconférence sur un réseau numérique à intégration de service (Rnis ou Isdn en anglais), comme par exemple le service numéris proposé par France Telecom. Le protocole H.323 est une adaptation de H.320 pour les réseaux Ip. A l'heure actuelle, la visioconférence sur liaison Rnis est toujours la technique la plus déployée. Elle existe depuis 1990. Les réseaux utilisés sont à commutation de circuits. Ils permettent ainsi de garantir une Qualité de Service (QoS) aux utilisateurs (pas de risque de coupure du son ou de l'image). Aujourd'hui, c'est encore un avantage indiscutable. Par contre, comme pour le téléphone, la facturation est fonction du débit utilisé, du temps de communication et de la distance entre les appels.

H.323 définit plusieurs éléments de réseaux :

Les terminaux - Dans un contexte de téléphonie sur IP, deux types de terminaux H.323 sont Aujourd'hui disponibles. Un poste téléphonique IP raccordés directement au réseau Ethernet de l'entreprise. Un PC multimédia sur lequel est installé une application compatible H.323.

Les passerelles (GW: Gateway) - Elles assurent l'interconnexion entre un réseau Ip et le réseau téléphonique, ce dernier pouvant être soit le réseau téléphonique public, soit un Pabx d'entreprise. Elles assurent la correspondance de la signalisation et des signaux de contrôle et la cohésion entre les médias. Pour ce faire, elles implémentent les fonctions suivantes de transcodage audio (compression, décompression), de modulation, démodulation (pour les fax), de suppression d'échos, de suppression des silences et de contrôle d'appels. Les passerelles sont le plus souvent basées sur des serveurs informatiques standards (Windows NT, Linux) équipés d'interfaces particuliers pour la téléphonie (interfaces analogiques, accès de base ou accès primaire RNIS, interface E1, etc.) et d'interfaces réseau, par exemple de type Ethernet. La fonctionnalité de passerelle peut toutefois être intégrée directement dans le routeur ainsi que dans les Pbx eux-mêmes.

Les portiers (GK: Gatekeeper) - Ils sont des éléments optionnels dans une solution H.323. Ils ont pour rôle de réaliser la traduction d'adresse (numéro de téléphone - adresse Ip) et la gestion des autorisations. Cette dernière permet de donner ou non la permission d'effectuer un appel, de limiter la bande passante si besoin et de gérer le trafic sur le Lan. Les "gardes-barrière" permettent également de gérer les téléphones classiques et la signalisation permettant de router les appels afin d'offrir des services supplémentaires. Il peuvent enfin offrir des services d'annuaires.

Les unités de contrôle multipoint (MCU, Multipoint Control Unit) - Référence au protocole T.120 qui permet aux clients de se connecter aux sessions de conférence de données. Les unités de contrôle multipoint peuvent communiquer entre elles pour échanger des informations de conférence.

Dans un contexte de téléphonie sur IP, la signalisation a pour objectif de réaliser les fonctions suivantes :

Recherche et traduction d'adresses - Sur la base du numéro de téléphone du destinataire, il s'agit de trouver son adresse IP (appel téléphone . PC) ou l'adresse IP de la passerelle desservant le destinataire. Cette fonction est prise en charge par le Gatekeeper. Elle est effectuée soit localement soit par requête vers un annuaire centralisé.

Contrôle d'appel - L'équipement terminal (« endpoint » = terminal H.323 ou passerelle) situé à l'origine de l'appel établit une connexion avec l'équipement de destination et échange avec lui les informations nécessaires à l'établissement de l'appel. Dans le cas d'une passerelle, cette fonction implique également de supporter la signalisation propre à l'équipement téléphonique à laquelle elle est raccordée (signalisation analogique, Q.931, etc.) et de traduire cette signalisation dans le format défini dans H.323. Le contrôle d'appel est pris en charge soit par les équipements terminaux soit par le Gatekeeper. Dans ce cas, tous les messages de signalisation sont routés via le Gatekeeper, ce dernier jouant alors un rôle similaire à celui d'un PBX.
Services supplémentaires : déviation, transfert d'appel, conférence, etc.

Trois protocoles de signalisation sont spécifiés dans le cadre de H.323 à savoir :

RAS (Registration, Admission and Status) - Ce protocole est utilisé pour communiquer avec un Gatekeeper. Il sert notamment aux équipements terminaux pour découvrir l'existence d'un Gatekeeper et s'enregistrer auprès de ce dernier ainsi que pour les demandes de traduction d'adresses. La signalisation RAS utilise des messages H.225.0 6 transmis sur un protocole de transport non fiable (Udp, par exemple).

Q.931 - H.323 utilise une version simplifiée de la signalisation RNIS Q.931 pour l'établissement et le contrôle d'appels téléphoniques sur Ip. Cette version simplifiée est également spécifiée dans la norme H.225.0.

H.245 : ce protocole est utilisé pour l'échange de capacités entre deux équipements terminaux. Par exemple, il est utilisé par ces derniers pour s'accorder sur le type de codec à activer. Il peut également servir à mesurer le retard aller-retour (Round Trip Delay) d'une communication.

Une communication H.323 se déroule en cinq phases :

Établissement d’appel

Échange de capacité et réservation éventuelle de la bande passante à travers le protocole RSVP (Ressource reSerVation Protocol)

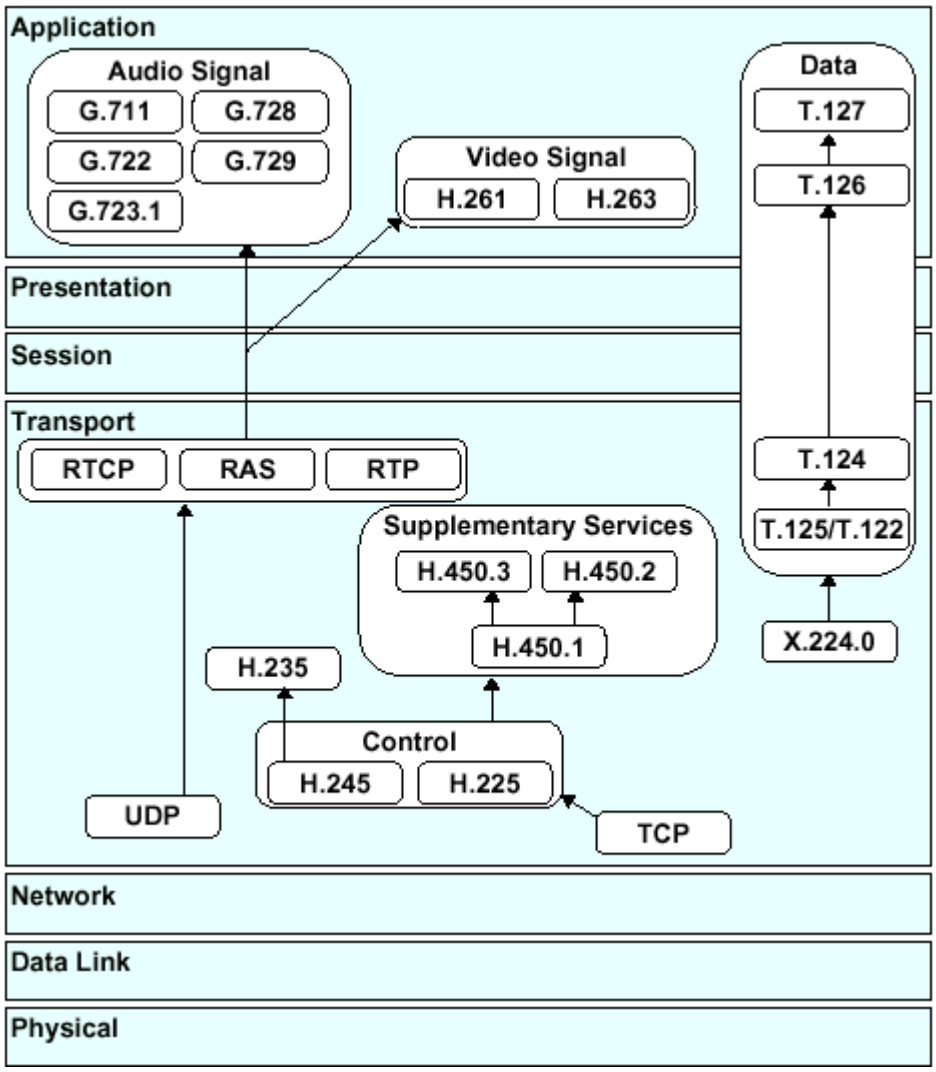
Établissement de la communication audio-visuelle

Invocation éventuelle de services en phase d’appel (par exemple, transfert d’appel, changement de bande passante, etc.)

Libération de l’appel.

6.1.3 - H323 dans le modèle Osi

Les différents protocoles sont représentés ci-dessous dans le modèle OSI :



6.1.4 - La visioconférence sur Ip

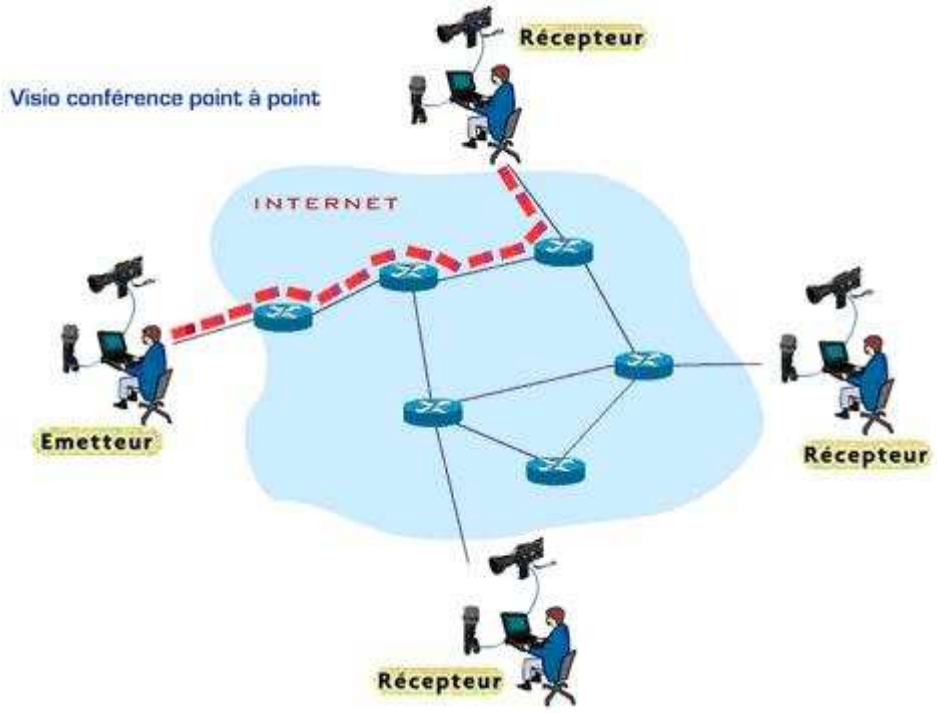
Tout d'abord, au niveau économique, la visioconférence sur Ip s'avère moins coûteuse que celle sur liaison RNIS car d'un côté, l'équipement d'un PC est relativement peu cher : ce système ne nécessite pas l'installation de prises RNIS spéciales. D'autre part, une liaison Rnis a un coût calculé selon la longueur de l'appel, le débit, et la distance. Alors que dans une liaison IP, le prix est forfaitaire selon le débit. En fin de compte, la visioconférence par Ip s'avère souvent moins onéreuse que par liaison Rnis.

Ensuite, qualitativement parlant, la visioconférence sur Ip peut utiliser des débits supérieurs et ainsi avoir une image et un son meilleurs qu'avec une liaison Rnis. En

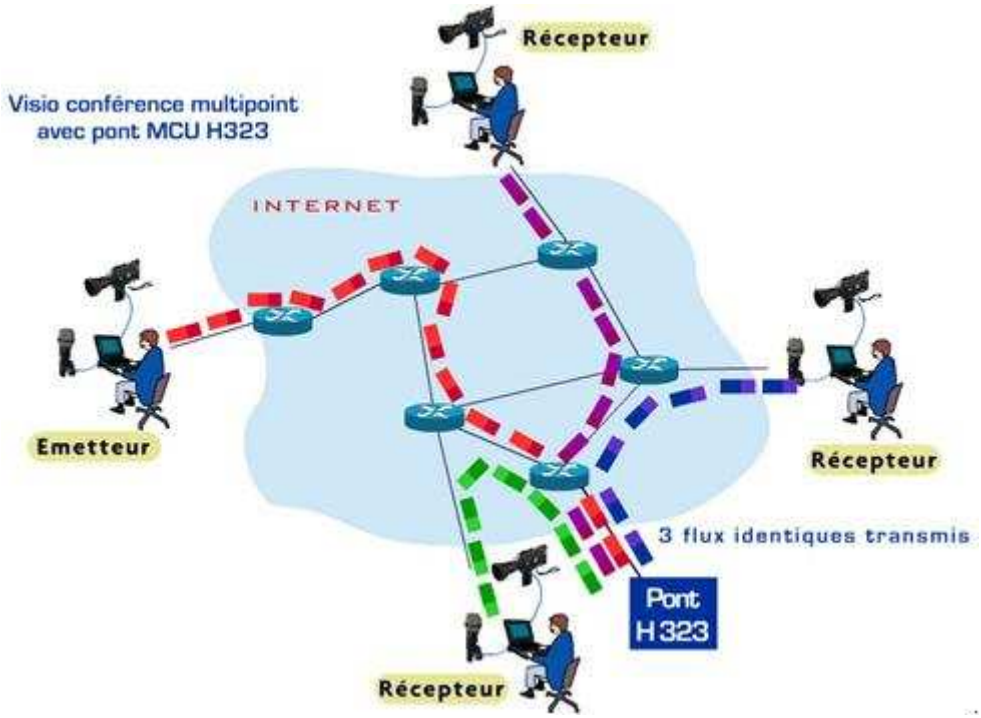
effet, la visioconférence sur Numeris utilise des débits allant de 128Kb/s à 384Kb/s, alors qu'en mutualisant certaines liaisons Ip, on peut obtenir des lignes haut débit allant jusqu'à plusieurs Mb/s. Malheureusement, le problème majeur de la visioconférence sur Ip est l'absence d'une Qualité de Service (QoS) sur les réseaux Ip. C'est également ce qui fait l'avantage des réseaux Rnis. Cependant, avec l'évolution des réseaux Ip, on sait désormais qu'il est possible qu'on puisse disposer d'une QoS sur ceux-ci tel que Rsvp, Diffserv, gestion de file d'attente. On pourrait donc avoir des flux avec priorité sur ces réseaux.

En dehors du protocole H.323, il existe des normes de visioconférence sur Ip ayant des possibilités analogues à H.323 telles que Ip multicast, qui est particulièrement adaptés au téléenseignement et à la diffusion de séminaires et conférences car il permet la connexion de plusieurs dizaines de sites voire plus. Il existe également le système Vrvs qui est utilisé dans certaines communautés scientifiques, notamment la physique, en raison de sa convivialité. Il intègre Ip multicast et H.323.

Pour pouvoir suivre une visioconférence, il faut bien entendu le matériel adéquat. Ce peut être un matériel dédié contenant tout ce qu'il faut : moniteur, micro, et caméra vidéo. Ou alors, un ensemble matériel et logiciel sur un poste de travail normal (PC, etc.). Si la visioconférence ne compte que deux interlocuteurs, alors la liaison est point à point comme illustré sur le schéma ci-dessous :



Dans le cas où il y a plus de deux interlocuteurs, la visioconférence nécessite l'utilisation d'un pont multipoint comme illustré sur le schéma ci-dessous :



Pour se connecter entre eux, les interlocuteurs sont identifiés par un numéro ou une adresse E.164. Elle est composée de numéros et est structurée comme un numéro

de téléphone. En particulier, un numéro de téléphone est une adresse E.164. « E.164 » est le nom de la norme qui définit ces adresses.

Pour router un appel H.323 dans le réseau, il est nécessaire d'avoir un « GateKeeper ». C'est un élément logiciel qui fonctionne dans un PC, ou encore dans un pont multipoint ou dans un routeur IP (Exemple dans les routeurs Cisco). En fonction de l'adresse destinataire contenue dans l'appel H.323, les différents GateKeeper vont établir la communication entre émetteur et destinateur et mettre en place le routage.

Par ailleurs, le protocole H.323 intègre la norme T.120 qui permet le partage d'applications. On peut, par exemple, afficher des documents sur les postes de travail des autres interlocuteurs.

6.1.5 - Avantages et inconvénients

Les réseaux IP sont à commutation de paquets, les flux de données transitent en commun sur une même liaison. La visioconférence IP mise sur une disponibilité de ces liaisons. Les débits des réseaux IP doivent donc être adaptés en fonction du trafic afin d'éviter tout risque de coupure du son et de la vidéo. Tous les sites n'ont pas le même débit. Plus le débit sera élevé et plus le risque de coupure sera faible. Par ailleurs, tant que la Qualité de Service n'existera pas dans les réseaux IP, la fiabilité des visioconférences sur les lignes à faible débit sera basse.

A l'heure actuelle, la compatibilité entre les différentes normes de visioconférence est assez faible. La visioconférence H.323 et H.320 sont compatibles mais elles nécessitent l'emploi de passerelles H.320/H.323.
En ce qui concerne les différentes normes pour la visioconférence sur Ip, H.323 et Ip Multicast ne sont, en règle générale, pas compatibles, sauf dans le cadre de VRVS qui permet un certain degré d'interopérabilité, mais ne gère pas la norme T.120.

Voici les principaux bénéfices qu'apporte la norme H.323 sont les suivants :

Codec standards : H.323 établit des standards pour la compression et la décompression des flux audio et vidéo. Ceci assure que des équipements provenant de fabricants différents ont une base commune de dialogue.

Interopérabilité : Les utilisateurs veulent pouvoir dialoguer sans avoir à se soucier de la compatibilité du terminal destinataire. En plus d'assurer que le destinataire est en mesure de décompresser l'information, H.323 établit des méthodes communes d'établissement et de contrôle d'appel.

Indépendance vis à vis du réseau : H.323 est conçu pour fonctionner sur tout type d'architecture réseau. Comme les technologies évoluent et les techniques de gestion de la bande passante s'améliorent, les solutions basées sur H.323 seront capables de bénéficier de ces améliorations futures.

Indépendance vis à vis des plates-formes et des applications : H.323 n'est lié à aucun équipement ou système d'exploitation.

Support multipoint : H.323 supporte des conférences entre trois points terminaux ou plus sans nécessiter la présence d'une unité de contrôle spécialisée.

Gestion de la bande passante : Le trafic audio et vidéo est un grand consommateur de ressources réseau. Afin d'éviter que ces flux ne congestionnent le réseau, H.323 permet une gestion de la bande passante à disposition. En particulier, le gestionnaire du réseau peut limiter le nombre simultané de connexions H.323 sur son réseau ou limiter la largeur de bande à disposition de chaque connexion. De telles limites permettent de garantir que le trafic important ne soit pas interrompu.

Support multicast : H.323 supporte le multicast dans les conférences multipoint. Multicast envoie chaque paquet vers un sous-ensemble des destinataires sans réplication, permettant une utilisation optimale du réseau.

A l'heure actuelle, le standard de fait pour les systèmes de téléphonie sur IP est la norme H.323 de l'UIT. Indispensable pour permettre un minimum d'interopérabilité entre équipements de fournisseurs différents, ce standard présente toutefois les inconvénients suivants :

Protocole complexe, créé initialement pour les conférences multimédia et qui incorpore des mécanismes superflus dans un contexte purement téléphonique. Ceci a notamment des incidences au niveau des terminaux H.323 (téléphones IP, par exemple) qui nécessitent de ce fait une capacité mémoire et de traitement non sans incidence au niveau de leur coût.

Comprend de nombreuses options susceptibles d'être implémentées de façon différentes par les constructeurs et donc de poser des problèmes d'interopérabilité ou de plus petit dénominateur commun (dans le choix du codec, par exemple) ; D'autre part, comme le seul codec obligatoire est le codec G.711 (64 Kps) et que le support des autres codecs plus efficaces est optionnel, l'interopérabilité entre produits provenant de constructeurs différents ne signifie pas qu'ils feront un usage optimal de la bande passante. En effet, dans le cas où les codecs à bas débits sont différents, le transport de la voix se fera à 64 Kbps, ce qui, en terme de bande passante, ne présente guère d'avantages par rapport à un système téléphonique classique.

6.1.6 - Comparaison avec Sip

Sip est un autre protocole pour l'interactivité en temps réel. Il a été développé par l'IETF et s'inspire du protocole Http alors que H.323 s'inspire de la téléphonie. Sip est plus modulaire et peut fonctionner avec d'autres protocoles. Il est donc plus souple que H.323.

6.1.7 - Conclusion

Le protocole H.323 est une des normes envisageables pour la visioconférence sur Ip. Cependant, elle est pour l'instant surtout employé par des programmes propriétaires (Microsoft, etc.). La documentation est difficile d'accès car l'ITU fait payer les droits d'accès aux derniers développements de cette technologie, en dehors des efforts faits par le projet OpenH.323 pour rendre cette technologie accessible à tous. Cet ensemble de normes ne s'avèrent pas toujours compatibles avec d'autres protocoles à cause de son développement inspiré de la téléphonie, ce qui peut rendre son utilisation un peu "rigide".

6.2 - Protocole SIP

6.2.1 - Introduction

Le protocole Sip (Session Initiation Protocole) a été initié par le groupe MMUSIC (Multiparty Multimedia Session Control) et désormais repris et maintenu par le groupe SIP de l'IETF donnant la Rfc 3261 rendant obsolète la Rfc 2543. Sip est un protocole de signalisation appartenant à la couche application du modèle Osi. Son rôle est d'ouvrir, modifier et libérer les sessions. L'ouverture de ces sessions permet de réaliser de l'audio ou vidéoconférence, de l'enseignement à distance, de la voix (téléphonie) et de la diffusion multimédia sur Ip essentiellement. Un utilisateur peut se connecter avec les utilisateurs d'une session déjà ouverte. Pour ouvrir une session, un utilisateur émet une invitation transportant un descripteur de session permettant aux utilisateurs souhaitant communiquer de s'accorder sur la compatibilité de leur média, Sip permet donc de relier des stations mobiles en transmettant ou redirigeant les requêtes vers la position courante de la station appelée. Enfin, SIP possède l'avantage de ne pas être attaché à un médium particulier et est sensé être indépendant du protocole de transport des couches basses.

6.2.2 - Fonctionnement

Sip intervient aux différentes phases de l'appel :

Localisation du terminal correspondant,

Analyse du profil et des ressources du destinataire,

Négociation du type de média (voix, vidéo, données…) et des paramètres de communication,

Disponibilité du correspondant, détermine si le poste appelé souhaite communiquer, et autorise l'appelant à le contacter.

Etablissement et suivi de l'appel, avertit les parties appelant et appelé de la demande d'ouverture de session, gestion du transfert et de la fermeture des appels.

Gestion de fonctions évoluées : cryptage, retour d'erreurs, …

Avec Sip, les utilisateurs qui ouvrent une session peuvent communiquer en mode point à point, en mode diffusif ou dans un mode combinant ceux-ci. Sip permet donc l'ouverture de sessions en mode :

Point-à-point - Communication entre 2 machines, on parle d'unicast.

Diffusif - Plusieurs utilisateurs en multicast, via une unité de contrôle M.C.U (Multipoint Control Unit)

Combinatoire - Plusieurs utilisateurs pleinement interconnectés en multicast via un réseau à maillage complet de connexions.

Voici les différents éléments intervenant dans l'ouverture de session :

Suivant nature des échanges, choix des protocoles les mieux adaptés (Rsvp, Rtp, Rtcp, Sap, Sdp).

Détermination du nombre de sessions, comme par exemple, pour véhiculer de la vidéo, 2 sessions doivent être ouvertes (l'une pour l'image et l'autre pour la vidéo).

Chaque utilisateur et sa machine est identifié par une adresse que l'on nomme Url Sip et qui se présente comme une Url Mailto.

Requête Uri permettant de localiser le proxy server auquel est rattaché la machine de l'appelé.

Requête Sip, une fois le client (machine appelante) connecté à un serveur Sip distant, il peut lui adresser une ou plusieurs requêtes Sip et recevoir une ou plusieurs réponses de ce serveur. Les réponses contiennent certains champs identiques à ceux des requêtes, tels que : Call-ID, Cseq, To et From.

Les échanges entre un terminal appelant et un terminal appelé se font par l'intermédiaire de requêtes :

Invite - Cette requête indique que l’application (ou utilisateur) correspondante à l’Url Sip spécifié est invité à participer à une session. Le corps du message décrit cette session (par ex : média supportés par l’appelant). En cas de réponse favorable, l’invité doit spécifier les médias qu’il supporte.

Ack - Cette requête permet de confirmer que le terminal appelant a bien reçu une réponse définitive à une requête Invite.

Options - Un proxy server en mesure de contacter l'UAS (terminal) appelé, doit répondre à une requête Options en précisant ses capacités à contacter le même terminal.

Bye - Cette requête est utilisée par le terminal de l’appelé à fin de signaler qu’il souhaite mettre un terme à la session.

Cancel - Cette requête est envoyée par un terminal ou un proxy server à fin d’annuler une requête non validée par une réponse finale comme, par exemple, si une machine ayant été invitée à participer à une session, et ayant accepté l’invitation ne reçoit pas de requête Ack, alors elle émet une requête Cancel.

Register - cette méthode est utilisée par le client pour enregistrer l’adresse listée dans l’URL TO par le serveur auquel il est relié.

Une réponse à une requête est caractérisée, par un code et un motif, appelés code d’état et raison phrase respectivement. Un code d’état est un entier codé sur 3 bits indiquant un résultat à l’issue de la réception d’une requête. Ce résultat est précisé par une phrase, textbased (UTF-8), expliquant le motif du refus ou de l’acceptation de la requête. Le code d’état est donc destiné à l’automate gérant l’établissement des sessions Sip et les motifs aux programmeurs. Il existe 6 classes de réponses et donc de codes d’état, représentées par le premier bit :

1xx = Information - La requête a été reçue et continue à être traitée

2xx = Succès - L’action a été reçue avec succès, comprise et acceptée

3xx = Redirection - Une autre action doit être menée afin de valider la requête

4xx = Erreur du client - La requête contient une syntaxe éronnée ou ne peut pas être traitée par ce serveur

5xx = Erreur du serveur - Le serveur n’a pas réussi à traiter une requête apparemment correcte

6xx = Echec général - La requête ne peut être traitée par aucun serveur

Dans un système Sip on trouve deux types de composantes, les users agents (UAS, UAC) et un réseau de serveurs :

L’UAS (User Agent Server) - Il représente l’agent de la partie appelée. C’est une application de type serveur qui contacte l’utilisateur lorsqu’une requête Sip est reçue. Et elle renvoie une réponse au nom de l’utilisateur.

L’U.A.C (User Agent Client) - Il représente l’agent de la partie appelante. C’est une application de type client qui initie les requêtes.

Le relais mandataire ou PS (Proxy Server), auquel est relié un terminal fixe ou mobile, agit à la fois comme un client et comme un comme serveur. Un tel serveur peut interpréter et modifier les messages qu’il reçoit avant de les retransmettre :

Le RS (Redirect Server) - Il réalise simplement une association (mapping) d’adresses vers une ou plusieurs nouvelles adresses. (lorsqu’un client appelle un terminal mobile - redirection vers le PS le plus proche - ou en mode multicast - le message émis est redirigé vers toutes les sorties auxquelles sont reliés les destinataires). Notons qu’un Redirect Server est consulté par l'Uac comme un simple serveur et ne peut émettre de requêtes contrairement au Ps.

Le LS (Location Server) - Il fournit la position courante des utilisateurs dont la communication traverse les Rs et PS auxquels il est rattaché. Cette fonction est assurée par le service de localisation.

Le RG (Registrar) - C'est un serveur qui accepte les requêtes Register et offre également un service de localisation comme le LS. Chaque PS ou RS est généralement relié à un Registrar.

6.2.3 - Sécurité et Authentification

Les messages Sip peuvent contenir des données confidentielles, en effet le protocole Sip possède 3 mécanismes de cryptage :

Cryptage de bout en bout du Corps du message Sip et de certains champs d'en-tête sensibles aux attaques.

Cryptage au saut par saut (hop by hop) à fin d'empêcher des pirates de savoir qui appelle qui.

Cryptage au saut par saut du champ d'en-tête Via pour dissimuler la route qu'a emprunté la requête.

De plus, à fin d'empêcher à tout intrus de modifier et retransmettre des requêtes ou réponses Sip, des mécanismes d'intégrité et d'authentification des messages sont mis en place. Et pour des messages Sip transmis de bout en bout, des clés publiques et signatures sont utilisées par Sip et stockées dans les champs d'en-tête Autorisation.

Une autre attaque connue avec Tcp ou Udp est le « deny of service », lorsqu'un Proxy Server intrus renvoie une réponse de code 6xx au client (signifiant un échec général, la requête ne peut être traitée). Le client peut ignorer cette réponse. Si il ne l'ignore pas et émet une requête vers le serveur "régulier" auquel il était relié avant la réponse du serveur "intrus", la requête aura de fortes chances d'atteindre le serveur intrus et non son vrai destinataire.

6.2.4 - Comparaison avec H323

Voici les avantages du protocole H.323 :

Il existe de nombreux produits (plus de 30) utilisant ce standard adopté par de grandes entreprises telles Cisco, IBM, Intel, Microsoft, Netscape, etc.

Les cinq principaux logiciels de visioconférence Picturel 550, Proshare 500, Trinicon 500, Smartstation et Cruiser 150 utilisent sur Ip la norme H.323.

Un niveau d’interopérabilité très élevé, ce qui permet à plusieurs utilisateurs d'échanger des données audio et vidéo sans faire attention aux types de média qu'ils utilisent.

Voici les avantages du protocole Sip :

Sip est un protocole plus rapide. La séparation entre ses champs d’en-tête et son corps du message facilite le traitement des messages et diminue leur temps de transition dans le réseau.

Nombre des en-têtes est limité (36 au maximum et en pratique, moins d'une dizaine d'en-têtes sont utilisées simultanément), ce qui allège l'écriture et la lecture des requêtes et réponses.

Sip est un protocole indépendant de la couche transport. Il peut aussi bien s’utiliser avec Tcp que Udp.

De plus, il sépare les flux de données de ceux la signalisation, ce qui rend plus souple l'évolution "en direct" d'une communication (arrivée d'un nouveau participant, changement de paramètres…).

	SIP	H323
Nombre échanges pour établir la connexion	1,5 aller-retour	6 à 7 aller-retour
Maintenance du code protocolaire	Simple par sa nature textuelle à l'exemple de Http	Complexe et nécessitant un compilateur
Evolution du protocole	Protocole ouvert à de nouvelles fonctions	Ajout d'extensions propriétaires sans concertation entre vendeurs
Fonction de conférence	Distribuée	Centralisée par l'unité MC
Fonction de téléservices	Oui, par défaut	H.323 v2 + H.450

Détection d'un appel en boucle	Oui	Inexistante sur la version 1 un appel routé sur l'appelant provoque une infinité de requêtes
Signalisation multicast	Oui, par défaut	Non

6.2.5 - Conclusion

La simplicité, la rapidité et la légèreté d'utilisation, tout en étant très complet, du protocole Sip sont autant d'arguments qui pourraient permettre à Sip de convaincre les investisseurs. De plus, ses avancées en matière de sécurité des messages sont un atout important par rapport à ses concurrents.

6.3 - Transport Rtp & Rtcp

6.3.1 - Introduction

Rtp est un protocole qui a été développé par l'IETF afin de facilité le transport temps réel de bout en bout des flots données audio et vidéo sur les réseaux Ip, c'est à dire sur les réseaux de paquets. Rtp est un protocole qui se situe au niveau de l'application et qui utilise les protocoles sous-jacents de transport Tcp ou Udp. Mais l'utilisation de Rtp se fait généralement au-dessus de Udp ce qui permet d'atteindre plus facilement le temps réel. Les applications temps réels comme la parole numérique ou la visio-conférence constitue un véritable problème pour Internet. Qui dit application temps réel, dit présence d'une certaine qualité de service (QoS) que Rtp ne garantie pas du fait qu'il fonctionne au niveau Applicatif. De plus Rtp est un protocole qui se trouve dans un environnement multipoint, donc on peut dire que Rtp possède à sa charge, la gestion du temps réel, mais aussi l'administration de la session multipoint.

Rtp et Rtcp sont définis, depuis juillet 2003, par la Rfc 3550 rendant obsolète la version précédente Rfc 1889.

6.3.2 - Les fonctions de Rtp

Le protocole Rtp, Real Time Transport Protocol, standardisé en 1996, a pour but d'organiser les paquets à l'entrée du réseau et de les contrôler à la sortie. Ceci de façon à reformer les flux avec ses caractéristiques de départ. Rtp est géré au niveau de l'application donc ne nécessite pas l'implémentation d'un Kernel ou de librairies. Comme nous l'avons dit dans l'introduction, Rtp est un protocole de bout en bout. Rtp est volontairement incomplet et malléable pour s'adapter aux besoins des applications. Il sera intégré dans le noyau de l'application. Rtp laisse la responsabilité du contrôle aux équipements d'extrémité.

Rtp, est un protocole adapté aux applications présentant des propriétés temps réel. Il permet ainsi de :

Reconstituer la base de temps des flux (horodatage des paquets : possibilité de resynchronisation des flux par le récepteur)

Mettre en place un séquencement des paquets par une numérotation et ce afin de permettre ainsi la détection des paquets perdus. Ceci est un point primordial dans la reconstitution des données. Mais il faut savoir quand même que la perte d'un paquet n'est pas un gros problème si les paquets ne sont pas perdus en trop grands nombre. Cependant il est très important de savoir quel est le paquet qui a été perdu afin de pouvoir pallier à cette perte. Et ce par le remplacement par un paquet qui se compose d'une synthèse des paquets précédent et suivant.

Identifier le contenu des données pour leurs associer un transport sécurisé.

L'identification de la source c'est à dire l'identification de l'expéditeur du paquet. Dans un multicast l'identité de la source doit être connue et déterminée.

Transporter les applications audio et vidéo dans des trames (avec des dimensions qui sont dépendantes des codecs qui effectuent la numérisation). Ces trames sont incluses dans des paquets afin d'être transportées et doivent de ce fait être récupérées facilement au moment de la phase de dépaquétisation afin que l'application soit décodée correctement.

En revanche, ce n'est pas "la solution" qui permettrait d'obtenir des transmissions temps réel sur IP. En effet, il ne procure pas de :

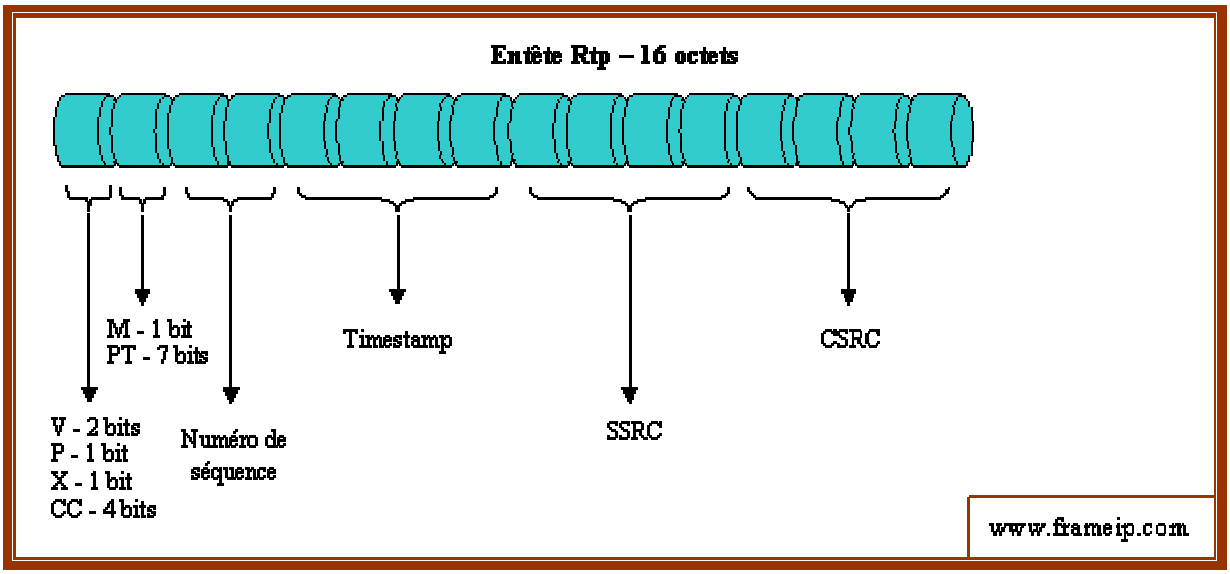
Réservation de ressources sur le réseau (pas d'action sur le réseau, cf. RSVP);

Fiabilité des échanges (pas de retransmission automatique, pas de régulation automatique du débit);

Garantie dans le délai de livraison (seules les couches de niveau inférieur le peuvent) et dans la continuité du flux temps réel.

6.3.3 - Entête Rtp

L'entête d'un paquet Rtp est obligatoirement constitué de 16 octets. Cette entête précède le "payload" qui représente les données utiles.



6.3.3.1 - V

Ce champ, codé sur 2 bits, permet d'indiquer la version de Rtp. Actuellement, V=2.

6.3.3.2 - P

Ce bit indique, si il est à 1, que les données possèdent une partie de bourrage.

6.3.3.3 - X

Ce bit spécifie, si il est à 1, que l'entête est suivie d'une entête supplémentaire.

6.3.3.4 - CC

Ce champ, codé sur 4 bits, représente le nombre de CSRC qui suit l'entête.

6.3.3.5 - M

Ce bit, lorsqu'il est à 1, définie que l'interprétation de la Marque est par un profil d'application.

6.3.3.6 - PT

Basé sur 7 bits, ce champ identifie le type du payload (audio, vidéo, image, texte, html, etc.).

6.3.3.7 - Numéro de séquence

Ce champ, d'une taille de 2 octets, représente le numéro d'ordre d'émission des paquets. Sa valeur initiale est aléatoire et il s'incrémente de 1 à chaque paquet envoyé, il peut servir à détecter des paquets perdus.

6.3.3.8 - Timestamp

Ce champ horodatage, de 4 octets, représente l'horloge système ou l'horloge d'échantillonnage de l'émetteur. Elle doit être monotone et linéaire pour assurer la synchronisation des flux.

6.3.3.9 - SSRC

Basé sur 4 octets, ce champ identifie de manière unique la source de synchronisation, sa valeur est choisie de manières aléatoire par l'application.

6.3.3.10 - SSRC

Ce champ, sur 4 octets, identifie les sources de contribution. La liste des participants ayant leur contribution (audio, vidéo) aux donnée du paquet.

6.3.4 - Les fonctions de Rtcp

Le protocole Rtcp est fondé sur la transmission périodique de paquets de contrôle à tous les participants d'une session. C'est le protocole Udp (par exemple) qui permet le multiplexage des paquets de données Rtp et des paquets de contrôle Rtcp. Le protocole Rtp utilise le protocole Rtcp, Real-time Transport Control Protocol, qui transporte les informations supplémentaires suivantes pour la gestion de la session :

Les récepteurs utilisent Rtcp pour renvoyer vers les émetteurs un rapport sur la QoS. Ces rapports comprennent le nombre de paquets perdus, le paramètre indiquant la variance d'une distribution (plus communément appelé la gigue : c'est à dire les paquets qui arrivent régulièrement ou irrégulièrement) et le délai aller-retour. Ces informations permettent à la source de s'adapter, par exemple, de modifier le niveau de compression pour maintenir une QoS.

Une synchronisation supplémentaire entre les médias. Les applications multimédias sont souvent transportées par des flots distincts. Par exemple, la voix, l'image ou même des applications numérisées sur plusieurs niveaux hiérarchiques peuvent voir les flots gérées suivre des chemins différents.

L'identification car en effet, les paquets Rtcp contiennent des informations d'adresses, comme l'adresse d'un message électronique, un numéro de téléphone ou le nom d'un participant à une conférence téléphonique.

Le contrôle de la session, car Rtcp permet aux participants d'indiquer leur départ d'une conférence téléphonique (paquet Bye de Rtcp) ou simplement de fournir une indication sur leur comportement.

Le protocole Rtcp demande aux participants de la session d'envoyer périodiquement les informations citées ci-dessus. La périodicité est calculée en fonction du nombre de participants de l'application. On peut dire que les paquets Rtp ne transportent que les données des utilisateurs. Tandis que les paquets Rtcp ne transportent en temps réel, que de la supervision. On peut détailler les paquets de supervision en 5 types:

200 : rapport de l'émetteur

201 : rapport du récepteur

202 : description de la source

203 : au revoir

204 : application spécifique

Ces différents paquets de supervision fournissent aux nœuds du réseau les instructions nécessaires à un meilleur contrôle des applications temps réel.

6.3.5 - Entête Rtcp

Ce protocole défini cinq paquets de contrôle :

200 - SR (Sender Report) : Ce rapport regroupe des statistiques concernant la transmission (pourcentage de perte, nombre cumulé de paquets perdus, variation de délai (gigue), …Ces rapports sont issus d'émetteurs actifs d'une session.

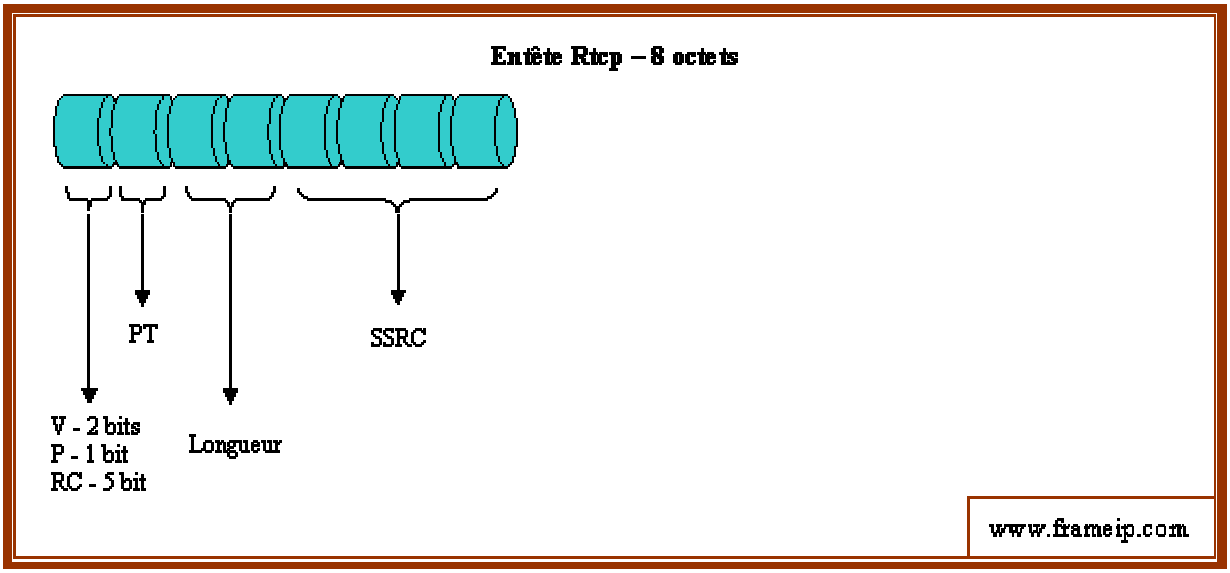
201 - RR (Receiver Report) : Ensemble de statistiques portant sur la communication entre les participants. Ces rapports sont issus des récepteurs d'une session.

202 - SDES (Source Description) : Carte de visite de la source (nom, e-mail, localisation).

203 - BYE : Message de fin de participation à une session.

204 - APP : Fonctions spécifiques à une application.

Voici l'en-tête commun à tous les paquets Rtcp.



6.3.5.1 - V

Ce champ, codé sur 2 bits, permet d'indiquer la version de Rtp, qui est la même que dans les paquets Rtcp. Actuellement, V=2.

6.3.5.2 - P

Ce bit indique, si il est à 1, que les données possèdent une partie de bourrage.

6.3.5.3 - RC

Ce champ, basé sur 5 bits, indique le nombre de blocs de rapport de réception contenus en ce paquet. Une valeur de zéro est valide.

6.3.5.4 - PT

Ce champ, codé sur 1 octet, est fixé à 200 pour identifier ce datagramme Rtcp comme SR.

6.3.5.5 - Longueur

Ce champ de 2 octets, représente la longueur de ce paquet Rtcp incluant l'entête et le bourrage.

6.3.5.6 - SSRC

Basé sur 4 octets, ce champ, représente l'identification de la source pour le créeateur de ce paquet SR.

6.3.6 - Conclusion

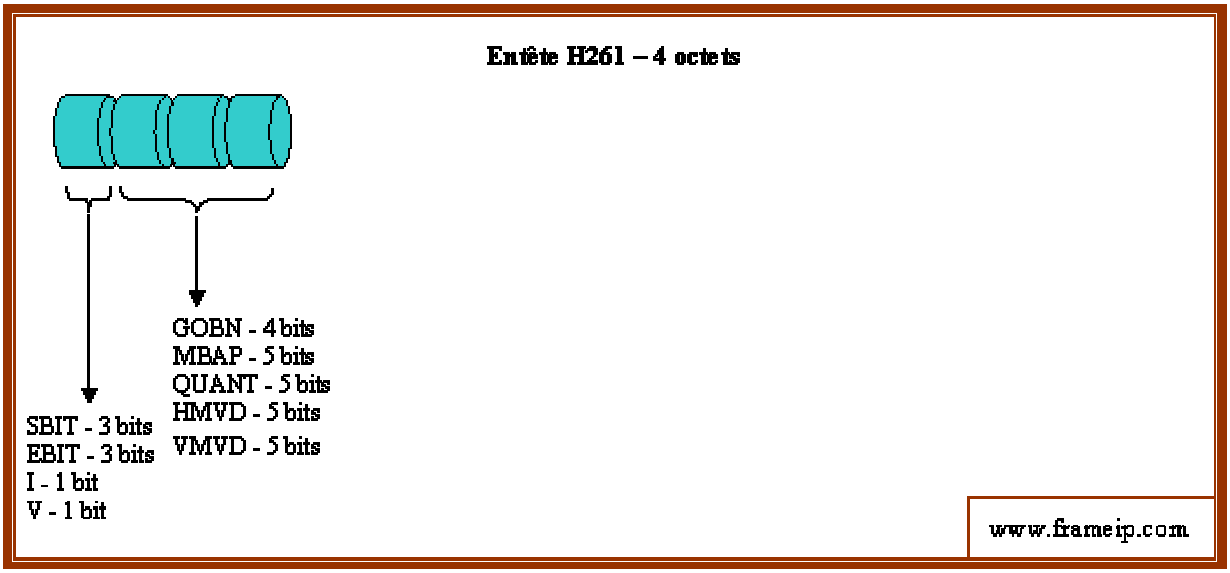
Rtp nécessite le protocole de transport Udp, (en-tête 8 octets), qui fournira les numéros de port source et destination nécessaire à la couche application. Pour l'instant le protocole Rtp se trouve au dessus de Udp, tandis que dans le futur, on aura une indépendance vis à vis des couches réseaux.

En résumant, ces deux protocoles sont adaptés pour la transmission de données temps réel. Cependant, ils fonctionnent en stratégie bout à bout et donc ne peuvent contrôler l'élément principal de la communication : le réseau.

Ces protocoles sont principalement utilisés en visioconférence où les participants sont tour à tour émetteurs ou récepteurs. Pour le transport de la voix, ils permettent une transmission correcte sur des réseaux bien ciblés. C'est-à-dire, des réseaux qui implémentent une qualité de service adaptée. Des réseaux bien dimensionnés (bande passante, déterminisme des couches sous-jacentes, Cos, ...) peuvent aussi se servir de cette solution.

6.4 - H261

Le protocole H.261 est décrit dans la RFC 2032, cette norme décrit le transport d'un flux vidéo sur Rtp. Le format de l'en-tête est le suivant :



SBIT (Start Bit) - Basé sur 3 bits, ce champ représente le nombre de bits de poids forts à ignorer dans le premier octet de données.

EBIT (End Bit) - Basé sur 3 bits, ce champ représente le nombre de bits de poids faible à ignorer dans le dernier octet de données.

I (Intra-frame encoded data flag) - Basé sur 1 bit, ce flag doit être mis à 1 si il contient seulement des intra-frame codé.

V (Motion Vector) - Basé sur 1 bit, ce flag indique si le Motion Vector est utilisé ou pas.

GOBN (GOB number) - Basé sur 4 bits, ce champ code le nombre de GOB actif au début du paquet. Placez à 0 si le paquet commence par un en-tête de GOB.

MBAP (Macroblock Address Predictor) - Basé sur 5 bits, ce champ code le prédicteur d'adresse de Macroblock. Placez à 0 si le paquet commence par un en-tête de GOB.

QUANT (Quantizer) - Basé sur 5 bits, ce champ représente la valeur actif avant le début de ce paquet.

HMVD (Horizontal Motion Vector Data) - Basé sur 5 bits, ce champ doit être à 0 si le flag V est à 0 ou si le paquet commence avec une entête Gob.

VMVD (Vertical Motion Vector Data) - Basé sur 5 bits, ce champ doit être à 0 si le flag V est à 0 ou si le paquet commence avec une entête Gob.

6.5 - Audio

Le transport de la voix sur un réseau IP nécessite au préalable tout ou une partie des étapes suivantes :

Numérisation : dans le cas où les signaux téléphoniques à transmettre sont sous forme analogique, ces derniers doivent d'abord être convertis sous forme numérique suivant le format PCM (Pulse Code Modulation) à 64 Kbps. Si l'interface téléphonique est numérique (accès RNIS, par exemple), cette fonction est omise.

Compression : le signal numérique PCM à 64 Kbps est compressé selon l'un des formats de codec (compression / décompression) (Tableau 3-3) puis inséré dans des paquets IP. La fonction de codec est le plus souvent réalisée par un DSP (Digital Signal Processor). Selon la bande passante à disposition, le signal voix peut également être transporté dans son format originel à 64 Kbps.

Décompression : côté réception, les informations reçues sont décompressées .il est nécessaire pour cela d'utiliser le même codec que pour la compression- puis reconverties dans le format approprié pour le destinataire (analogique, PCM 64Kbps, etc.).

L'objectif d'un codec est d'obtenir une bonne qualité de voix avec un débit et un délai de compression le plus faibles possibles. Le coût du DSP est lié à la complexité du codec utilisé. Le Tableau ci-dessous présente les caractéristiques des principaux codecs standards de l'UIT. Les codecs les plus souvent mis en oeuvre dans les solutions VoIP sont G.711, G.729 et G.723.1.

La qualité d'un codec est mesurée de façon subjective en laboratoire par une population test de personnes. Ces dernières écoutent tout un ensemble de

conversations compressées selon les différents codecs à tester et les évaluent qualitativement selon la table suivante :

Tableau : Echelle utilisé pour l'évaluation de la qualité de voix

Qualité de la parole	Score
Excellente	5
Bonne	4
Correcte	3
Pauvre	2
Insuffisante	1

Sur la base des données numériques des appréciations, une opinion moyenne de la qualité d'écoute (Mean Opinion Score . MOS) est ensuite calculée pour chaque codec. Les résultats obtenus pour les principaux codecs sont résumés dans le tableau ci-dessous :

Tableau : Score MOSdes différents codecs

Codec VoIP	Débit (Kbps)	Score MOS
G.711 (PCM)	64	4.1
G.726	32	3.85
G.729	8	3.92
G.723.1	6.4	3.9
G.723.1	5.3	3.65
GSM	13	3.5
G.729 x2		3.27
G.729 x3		2.68

G.729 x GSM 3.17

Deux observations principales peuvent être tirées du Tableau 3-5 :

La qualité de la voix obtenue par les codecs G.729 et G.723.1 (à 6.4Kbps) est très proche de celle du service téléphonique actuel, et ce pour des débits entre 8 et 10 fois inférieurs. Ces deux codecs présentent une meilleure qualité que celle des réseaux téléphoniques cellulaires (GSM).

Le cumul, dans une même communication, d'opérations de compression/décompression conduit à une rapide dégradation de la qualité. Les solutions mises en oeuvre doivent éviter des configurations en tandem dans lesquelles un PBX reçoit un appel d'un poste distant à travers une liaison VoIP et le redirige vers une autre liaison semblable.

Offrant une qualité de voix très proche, les codecs G.729 et G.723.1 se distinguent essentiellement par la bande passante qu'ils requièrent et par le retard que chacun introduit dans la transmission. Le choix d'un équipement implémentant l'un ou l'autre de ces codecs devra donc être fait selon la situation, en fonction notamment de la bande passante à disposition et du retard cumulé maximum estimé pour chaque liaison (selon les standards de l'UIT, le retard aller (« one-way delay ») devrait être inférieur à 150 ms). Le facteur du jitter est primordiale pour une bonne écoute de la Voip.

7 - Problème & QoS

7.1 - Latence

La maîtrise du délai de transmission est un élément essentiel pour bénéficier d'un véritable mode conversationnel et minimiser la perception d'écho (similaire aux désagréments causés par les conversations par satellites, désormais largement remplacés par les câbles pour ce type d'usage).

Or la durée de traversée d'un réseau IP dépend de nombreux facteurs:

Le débit de transmission sur chaque lien

Le nombre d’éléments réseaux traversés

Le temps de traversée de chaque élément, qui est lui même fonction de la puissance et la charge de ce dernier, du temps de mise en file d'attente des paquets, et du temps d'accès en sortie de l’élément

Le délai de propagation de l'information, qui est non négligeable si on communique à l'opposé de la terre. Une transmission par fibre optique, à l'opposé de la terre, dure environ 70 ms.

Noter que le temps de transport de l'information n'est pas le seul facteur responsable de la durée totale de traitement de la parole. Le temps de codage et la mise en paquet de la voix contribuent aussi de manière importante à ce délai.

Il est important de rappeler que sur les réseaux IP actuels (sans mécanismes de garantie de qualité de service), chaque paquet IP « fait sont chemin » indépendamment des paquets qui le précèdent ou le suivent: c'est ce qu'on appelle grossièrement le « Best effort » pour signifier que le réseau ne contrôle rien. Ce fonctionnement est fondamentalement différent de celui du réseau téléphonique où un circuit est établi pendant toute la durée de la communication.

Les chiffres suivants (tirés de la recommandation UIT-T G114) sont donnés à titre indicatif pour préciser les classes de qualité et d'interactivité en fonction du retard de transmission dans une conversation téléphonique. Ces chiffres concernent le délai total de traitement, et pas uniquement le temps de transmission de l'information sur le réseau.

Classe n° Délai par sens Commentaires

Classe n°	Délai par sens	Commentaires

1	0 à 150 ms	Acceptable pour la plupart des conversations
2	150 à 300 ms	Acceptable pour des communications faiblement interactives
3	300 à 700 ms	Devient pratiquement une communication half duplex
4	Au delà de 700 ms	Inutilisable sans une bonne pratique de la conversation half duplex

En conclusion, on considère généralement que la limite supérieure "acceptable" , pour une communication téléphonique, se situe entre 150 et 200 ms par sens de transmission (en considérant à la fois le traitement de la voix et le délai d'acheminement).

7.2 - Perte de paquets

Lorsque les buffers des différents élément réseaux IP sont congestionnés, ils « libèrent » automatiquement de la bande passante en se débarrassant d'une certaine proportion des paquets entrant, en fonction de seuils prédéfinis. Cela permet également d'envoyer un signal implicite aux terminaux TCP qui diminuent d'autant leur débit au vu des acquittements négatifs émis par le destinataire qui ne reçoit plus les paquets. Malheureusement, pour les paquets de voix, qui sont véhiculés au dessus d'UDP, aucun mécanisme de contrôle de flux ou de retransmission des paquets perdus n'est offert au niveau du transport. D'où l'importance des protocoles RTP et RTCP qui permettent de déterminer le taux de perte de paquet, et d'agir en conséquence au niveau applicatif.

Si aucun mécanisme performant de récupération des paquets perdus n'est mis en place (cas le plus fréquent dans les équipements actuels), alors la perte de paquet IP se traduit par des ruptures au niveau de la conversation et une impression de hachure de la parole. Cette dégradation est bien sûr accentuée si chaque paquet contient un long temps de parole (plusieurs trames de voix de paquet). Par ailleurs, les codeurs à très faible débit sont généralement plus sensibles à la perte d'information, et mettent plus de temps à « reconstruire » un codage fidèle.

Enfin connaître le pourcentage de perte de paquets sur une liaison n'est pas suffisant pour déterminer la qualité de la voix que l'on peut espérer, mais cela donne une bonne approximation. En effet, un autre facteur essentiel intervient; il s'agit du modèle de répartition de cette perte de paquets, qui peut être soit « régulièrement » répartie, soit répartie de manière corrélée, c'est à dire avec des pics de perte lors des phases de congestion, suivies de phases moins dégradées en terme de QoS.

7.3 - Gigue

La gigue est la variance statistique du délai de transmission. En d'autres termes, elle mesure la variation temporelle entre le moment où deux paquets auraient dû arriver et le moment de leur arrivée effective. Cette irrégularité d'arrivée des paquets est due à de multiples raisons dont: l'encapsulation des paquets IP dans les protocoles supportés, la charge du réseau à un instant donné, la variation des chemins empruntés dans le réseau, etc…

Pour compenser la gigue, on utilise généralement des mémoires tampon (buffer de gigue) qui permettent de lisser l'irrégularité des paquets. Malheureusement ces paquets présentent l'inconvénient de rallonger d'autant le temps de traversée global du système. Leur taille doit donc être soigneusement définie, et si possible adaptée de manière dynamique aux conditions du réseau.

La dégradation de la qualité de service due à la présence de gigue, se traduit en fait, par une combinaison des deux facteurs cités précédemment: le délai et la perte de paquets; puisque d'une part on introduit un délai supplémentaire de traitement (buffer de gigue) lorsque l'on décide d'attendre les paquets qui arrivent en retard, et que d'autre part on finit tout de même par perte certains paquets lorsque ceux-ci ont un retard qui dépasse le délai maximum autorisé par le buffer.

8 - Etat du marché

On compte une bonne vingtaine de firmes sur le marché. Les principaux sont Cisco, Clarent, Avaya, Alcatel, Nortel Network, Siemens, Ténovis, 3COM … Ce qu’il faut souligner, c’est le fait qu’il y ait peu de concurrents car comme je l’ai dit

précédemment, la téléphonie sur Ip est un marché très jeune et très novateur. D'ailleurs, le fait que la téléphonie sur IP soit un marché chevauchant 2 secteurs qui se rapprochent et étaient complètement différent auparavant, la téléphonie et l'informatique, nous assistons ici à une concurrence ayant des origines différentes. En effet, nous retrouvons le géant de l'équipement réseaux Cisco en concurrence avec des entreprises de téléphonies tel que Alcatel ou Siemens. Mais Cisco et Clarent arrivent largement en tête, sur un marché qui de 259 millions de dollars cette année pourrait atteindre 2,89 milliards en 2006. La téléphonie sur IP propose 3 types de terminaux différents : Les hardphones qui sont des téléphones physiques IP, les softphones qui sont des logiciels permettant de téléphoner sur IP au travers d'un PC et les téléphones IP Wi-fi qui sont des téléphones sans-fil IP. Mais la plupart des concurrents proposent ces 3 produits qui sont plutôt homogènes. Un softphone Cisco et un Softphone Siemens sont quasi-identiques. Seule l'interface graphique les distingue. Pour le client, le produit des 2 concurrents est identique dans la mesure où il apporte les mêmes services.

9 - Conclusion

Actuellement, il est évident que la téléphonie IP va continuer de se développer dans les prochaines années. Le marché de la téléphonie IP est très jeune mais se développe à une vitesse fulgurante. C'est aujourd'hui que les entreprises doivent investir dans la téléphonie IP si elles veulent y jouer un rôle majeur.

Le fait est que IP est maintenant un protocole très répandu, qui a fait ses preuves et que beaucoup d'entreprises disposent avantage de la téléphonie IP, car elle demande un investissement relativement faible pour son déploiement. La téléphonie IP ouvre la voie de la convergence voix/données et celle de l'explosion de nouveaux services tels que les CTI.

Maintenant que la normalisation a atteint une certaine maturité, il n'est plus dangereux de miser sur le standard H323 qui a été accepté par l'ensemble de la communauté.

La téléphonie IP est une bonne solution en matière d'intégration, de fiabilité,

d’évolutivité et de coût. Elle fera partie intégrante des Intranets d’entreprises dans les années à venir et apparaîtra aussi dans la téléphonie publique pour permettre des communications à bas coût.

Enfin, le développement de cette technologie représente-t-il un risque ou une opportunité pour les opérateurs traditionnels ? La réponse n’est pas tranchée. D’un coté, une stagnation des communications classiques; d’un autre coté l’utilisation massive d’Internet va augmenter le trafic et développer de nouveaux services que pourront développer les opérateurs. Bientôt nous téléphonerons tous sur IP...

On peut ainsi vraisemblablement penser que le protocole IP deviendra un jour un standard unique permettant l'interopérabilité des réseaux mondialisés. C'est pourquoi l'intégration de la voix sur IP n'est qu'une étape vers EoIP : Everything over IP.