

Voice over IP

par Luigi Rizzo 

C'est notamment en raison du phénomène Internet que l'IP est en passe de s'affirmer rapidement comme le protocole dominant les réseaux et la voix peut désormais n'être plus considérée que comme l'une des nombreuses applications possibles, permettant ainsi une grande simplification des infrastructures et la réalisation d'un véritable réseau intégré.

Sur les réseaux de télécommunication, le volume global du trafic de données croît actuellement à un rythme beaucoup plus rapide que le trafic vocal, en raison notamment du succès phénoménal de l'Internet, d'Intranet et d'Extranet. Depuis quelques années déjà, les données ont dépassé la voix et du même coup justifié la conception de réseaux optimisés destinés à la circulation des données (c'est à dire à commutation de paquet), plutôt qu'à la voix (commutation de circuit). Cette tendance débouche sur la nécessité d'exploiter des technologies de transport de la voix sur des réseaux de données, comme le Voice over IP, plutôt que le Voice over Frame Relay ou le Voice over ATM. On ne peut cependant pas nier que l'IP est en passe de s'affirmer comme le protocole unique de référence : ce qui fait de la VoIP la technologie la plus prometteuse et la plus stratégique pour le futur. Un autre aspect à considérer est le contexte compétitif global qui s'est progressivement créée, tant entre fournisseurs de services de circulation de la voix et des données qu'entre les clients. La conséquence est un effort considérable pour la recherche de technologies permettant de réduire les coûts en réalisant, par exemple, une infrastructure de réseau unique pour la fourniture de services de n'importe quel type. La dernière importante motivation de VoIP est technologique : l'amélioration rapide du rapport prix/prestations des technologies électroniques

pour le traitement numérique des signaux (les Digital Signal Processors) permet d'utiliser la compression et la Voice Activity Detection à grande échelle, réduisant de beaucoup (théoriquement de huit fois) la bande nécessaire au trafic. En ce qui concerne le marché, on peut constater que VoIP a décroché ses premières applications dans le milieu professionnel, où il était facile de justifier de simples réalisations par une réduction des coûts ; par la suite, les fournisseurs alternatifs de services de données (et en particulier les Internet Service Provider) ont vu en VoIP la possibilité d'offrir des services nouveaux et plus profitables. Enfin les nouveaux fournisseurs de services, dotés de réseaux en fibre optique de grande capacité, voient en VoIP la meilleure technologie pour construire un réseau téléphonique à faible coût sur une infrastructure IP.

La gestion numérique de la voix

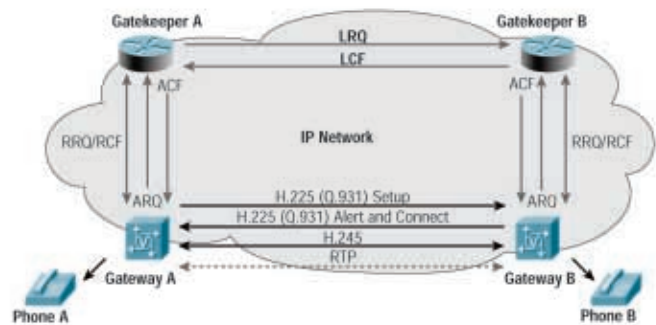
Il existe plusieurs organisations opérant dans le développement de standard pour VoIP, mais avant d'en parler, il est utile de rappeler les principes de base de la gestion numérique de la voix en téléphonie. La voix est numérisée à travers un échantillonnage à 8 KHz et un codage (non linéaire) sur 8 bits, qui produit un flux de 64 Kbps standardisé sous le sigle G.711, connu comme Pulse Code Modulation (PCM), et qui est à la

base des réseaux téléphoniques numériques. Le flux numérique peut être comprimé en éliminant la corrélation entre les différents échantillons numériques afin de réduire la bande nécessaire pour supporter la session vocale : cette opération respecte elle aussi certains algorithmes standard qui permettent d'obtenir différents taux de compression et donc différents niveaux de qualité de la voix. Les applications les plus sophistiquées permettent d'obtenir des résultats encore meilleurs : il est notamment possible d'activer une fonction dénommée Voice Activity Detection (VAD) qui permet de ne pas transmettre de données pendant les différentes phases de pause généralement présentes dans une conversation et donc d'entraîner un intéressant gain de bande supplémentaire.

Certains paramètres essentiels pour l'obtention d'une bonne qualité de transport de la voix sur réseaux IP, auxquels aucune ressource de transmission n'est préalablement allouée, comme dans les réseaux à commutation de circuit, nécessitent une attention particulière en phase de projet. Le plus important est constitué par le retard end-to-end, c'est-à-dire le temps compris entre l'émission du signal vocal et sa réception à l'autre bout de la ligne. Si ce retard est trop élevé, il est alors impossible de converser normalement : la valeur maximum conseillée est de 150 ms. Pour garantir le respect de cette valeur, il est nécessaire que le backbone IP soit conçu de manière appropriée en termes de transmission et que soient mis en œuvre des mécanismes sophistiqués de gestion des congestions, toujours possibles sur un réseau multiservices. La VoIP de haute qualité est donc rendue possible par un réseau IP avec QoS (Quality of Service), c'est à dire en mesure d'offrir des garanties sur les temps maximum de traversée du réseau. L'alternative à la gestion de la QoS, peu applicable, est de disposer d'un backbone largement surdimensionné sur lequel aucune congestion ne peut se vérifier et où les retards de transmission et d'acheminement sont minimales. L'Internet n'appartenant à aucune de ces catégories, il est impossible d'obtenir une qualité très élevée avec Internet, surtout en utilisant des PC en guise de téléphone. Tout change par contre lorsque l'on utilise exclusivement le réseau d'un unique ISP (ou de plusieurs ISP associés), où il est possible d'améliorer la QoS et de maintenir un niveau qualitatif élevé. Utiliser un réseau IP pour transporter la voix présente des criticités mais offre également des avantages substantiels : d'abord la bande est mieux exploitée car elle n'est plus allouée de manière rigide sur les différents circuits ; de plus l'IP est " media indépendant " et peut donc opérer en end-to-end sur n'importe quelle infrastructure de transmission, y compris sans fil.

Pour réaliser un réseau téléphonique, outre les protocoles de gestion du signal vocal en soi, il est nécessaire de tenir compte également des protocoles de signalisation et de contrôle, qui permettent d'instaurer les connexions. Le standard le plus diffus à l'heure actuelle pour les applications en milieu professionnel est sans conteste le H.323, qui couvre la commu-

nication en temps réel sur réseau à paquet et est donc bien plus vaste que le seul transport de la voix. Le standard H.323 regroupe un ensemble de recommandations émanant de la ITU (International Telecommunications Union) et comprend plusieurs composants, parmi lesquels le H.225 (spécifiant les messages pour le call control, en termes de signalisation, d'enregistrement, d'accès et de paquets / synchronisation des flux audio et vidéo) et le H.245 (spécifiant les messages pour l'ouverture et la fermeture des canaux de transmission des flux). Le H.323 exploite à son tour un protocole RTP (Real-Time Transport Protocol) défini par la IETF (Internet Engineering Task Force) et permet une interaction entre des produits multivendor car il est supporté par les principales entreprises du secteur.



Réseau H.323 [Source : Cisco]

Le Network Pabx

En milieu professionnel, les applications de VoIP les plus diffuses à l'heure actuelle naissent dans l'objectif de réduire les coûts téléphoniques en acheminant le trafic téléphonique existant entre deux sièges de l'entreprise vers le réseau de données au lieu du réseau de téléphone public. Dans ce contexte, les standards téléphoniques peuvent être reliés directement aux routeurs exploitant la VoIP par interfaces analogiques ou numériques. L'application est généralement constituée du fameux Network Pabx, où le standard téléphonique traditionnel relié aux téléphones est remplacé par une infrastructure constituée de téléphones VoIP (reliés directement via Ethernet) et/ou par des PC multimédiaux dotés d'application VoIP, par des appareils à connexion LAN et par un serveur doté d'un logiciel supportant toutes les fonctions avancées d'un standard ou d'un call center. L'infrastructure voix/données de l'entreprise est donc entièrement intégrée et les énormes capacités de commutation des switch du réseau local sont exploitées au mieux, assurant ainsi la fonction de call control à travers un logiciel fonctionnant sur un serveur standard.

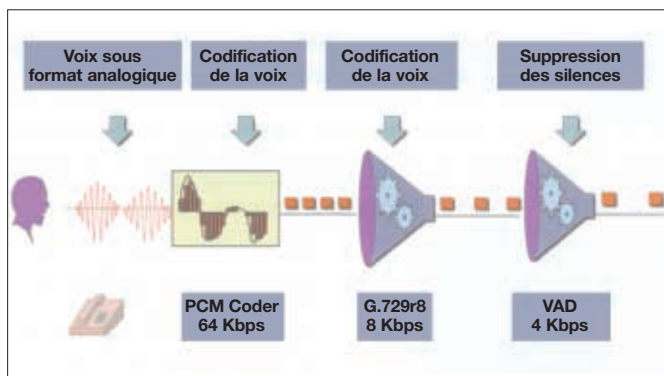
VOIX: LA TECHNOLOGIE

Pour que la voix puisse être transportée sur des réseaux de données, il est nécessaire que le signal analogique soit numérisé et segmenté en paquets pouvant être envoyés séparément avant d'être rassemblés selon la séquence originale. Les réseaux utilisables pour ce type de transport peuvent être l'intranet de l'entreprise, l'Internet lui-même, du moment que les critères de qualité end-to-end nécessaires sont respectés.

La voix, comme tout ce que nous entendons, a naturellement un format analogique et les premiers systèmes de téléphonies avaient les mêmes caractéristiques. Les signaux analogiques sont souvent représentés sous forme d'ondes sinusoïdales parfaitement régulières, mais la voix et les autres signaux sonores contiennent plusieurs fréquences et ont une structure généralement plus complexe.

Tandis que les êtres humains sont bien équipés pour la communication analogique, la transmission analogique n'est pas particulièrement efficace : lorsque les signaux analogiques faiblissent en raison de dispersions au cours de la transmission, il devient difficile de séparer la structure analogique complexe de la voix de la structure casuelle du bruit de transmission.

Amplifier les signaux analogiques revient nécessairement à amplifier aussi le bruit, et les connexions analogiques finissent par devenir trop bruyantes.



La codification de la voix (Source: Cisco)

Le format de transmission numérique le plus courant, appelé Pulse Code Modulation (PCM), convertit la voix en format numérique en échantillonnant les signaux-voix 8000 fois par seconde et en convertissant chacun des échantillons en un code.

La fréquence d'échantillonnage de 8000 fois par seconde (125 microsecondes entre chaque échantillonnage) a été choisie parce que pratiquement tous les signaux compréhensibles par les êtres humains circulent à des fréquences inférieures à 4000 Hz : échantillonner les ondes vocales toutes les 125 microsecondes suffit à relever les fréquences inférieures à 4000 Hz.

On ne peut penser à transporter la voix sur un réseau IP sans se préoccuper en premier lieu de la Quality of Service (QoS) du réseau intéressé.

La Voice over IP nécessite en effet de respecter des critères bien plus rigoureux que ceux nécessaires à de simples transferts de fichiers ou de e-mail. La QoS est un concept qui doit intéresser le réseau IP dans son ensemble, des gateway d'accès à l'architecture backbone.

Comme toutes les choses importantes, la transmission de la voix sur réseau IP a son prix. Les concepteurs de réseaux connaissent parfaitement les exigences de Quality of Service (QoS) d'applications critiques comme les processus transactionnels en ligne : eh bien, la Voice over IP exige des caractéristiques encore plus contraignantes. Si le réseau n'est pas conçu de manière appropriée pour remplir ces critères end-to-end, la QoS s'en ressent et la clarté du signal vocal est compromise. Mais quels sont les facteurs qui influencent cet aspect ?

Latence, codec, écho, pertes de paquets et Voice Activity Detection sont les facteurs déterminants de la qualité de la voix. Une qualité satisfaisante dépend d'un bon canal, d'une faible latence et d'un bon contrôle de l'écho. Un bon canal est caractérisé par une faible perte de paquets, par un codec de haute qualité et d'un bon algorithme de Voice Activity Detection (VAD, détection de l'activité vocale), qui a pour but d'éliminer les moments de silence contenus dans une conversation (qui représentent souvent jusqu'à 60% de la durée globale d'un appel téléphonique). La faible latence est nécessaire pour une communication vocale interactive. Le contrôle de l'écho fait en sorte que le locuteur n'interprète pas l'écho comme le début de la réponse de l'autre personne.

Pour un réseau Voice over IP de haute qualité, l'un des facteurs essentiels est la capacité de réduction et de contrôle du delay (retard). La latence des communications vocales est définie généralement comme le retard dans une direction entre la bouche de la personne qui parle et l'oreille de celle qui écoute. Le delay, ou latence, d'une communication téléphonique normale est de l'ordre de 50 ms. Le GSM a un delay supérieur à 200 ms ; les communications via satellite atteignent en moyenne plus de 500 ms. La latence maximum acceptable pour un réseau de données visant à une qualité vocale de type "carrier class", c'est-à-dire comparable à celle des opérateurs de téléphonie traditionnels, est de 150/200 ms. Parmi les éléments qui contribuent à la latence, on trouve les compressions codec (qui varient de 1 ms pour le G.711 à 20 ms pour le G.729), la paquetsisation et le framing, qui comportent environ 10 ms ; le buffer nécessaire pour compenser les variations de delay (jitter) peut ajouter 40 ms supplémentaires, mais 80% de la qualité en VoIP dépend généralement de la conception du réseau dans son ensemble.

Voix et données sur W-LAN

À la convergence de la voix et des données sur IP, Symbol Technologies ajoute un lecteur intégré et une liaison sans fil au réseau de l'entreprise : il s'agit du Netvision Data Phone, un dispositif portable léger et innovant qui permet à l'opérateur d'effectuer et de recevoir des appels téléphoniques, de saisir des informations ou de les acquérir par lecture de code barres, de communiquer en mode sans fil, d'accéder à des applications sur le serveur de l'entreprise. Grâce au partenariat avec des entreprises leader du secteur, comme Ericsson et Nortel, et au support fourni par celles-ci, les terminaux Netvision Data Phone peuvent être intégrés au réseau téléphonique de l'entreprise à travers une gateway spécifique. L'éventail des applications d'un instrument aussi polyvalent va de la grande distribution aux magasins de dépôt, en passant par les milieux hospitaliers : l'union entre un puissant dispositif de collecte de données et un téléphone – relié au réseau de l'entreprise à travers l'infrastructure Wi-Fi – donne les moyens de résoudre même les situations les plus critiques. Le récent accord selon lequel Psion Teklogix pourra revendre la gamme de téléphones sans fil NetLink de Spec-

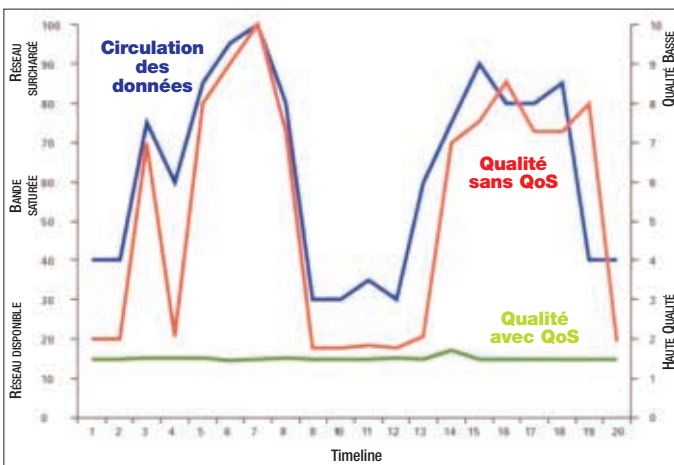


NetVision Data Phone de Symbol Technologies

traLink est également significatif. Les téléphones de SpectraLink garantissent la robustesse généralement requise dans les milieux d'utilisation des terminaux Psion Teklogix et en partagent l'infrastructure Wi-Fi. De plus, le protocole SVP (SpectraLink Voice Priority) sera

installé sur les access point de Psion Teklogix afin de garantir la Quality of Service (QoS) du signal vocal. L'annonce de SpectraLink de compter LXE parmi les partenaires pouvant revendre ses téléphones sans fil remonte même au mois d'août dernier: la connaissance du marché de la logistique de LXE constitue un excellent point de départ pour l'intégration de solutions innovantes comme celle que représentent les téléphones NetLink. Mais d'autres événements importants prouvent que la "Voice over Wi-Fi" a le vent en poupe: dès le début de l'été, Cisco lancera sur le marché un téléphone portable Wi-Fi; SpectraLink a annoncé un nouveau téléphone (toujours Wi-Fi) qui coûtera pratiquement la moitié

du modèle précédent; mais le véritable tournant sera celui résultant des développements sur lesquels Motorola, Proxim et Avaya sont en train de se pencher: des téléphones et des systèmes en mesure de gérer le roaming entre les réseaux Wi-Fi et les réseaux de téléphones cellulaires publics. □



L'importance de la Quality of Service (Source : Cisco)

Le retard produit dans une infrastructure de réseau est décomposable normalement en accès up-link, backbone et liaison d'accès down-link et dépendant du retard de propagation dû aux caractéristiques physiques des médias (cuivre, fibre, sans fil), que du "retard de réseau" causé par l'ensemble des appareils de réseau traitant les paquets vocaux le long de leur parcours. Le delay peut engendrer des interruptions du rythme ou de la cadence d'un appel, et lorsqu'il est élevé, il peut transformer la conversation en véritable communication via CB, contraignant les interlocuteurs à adopter une structure formelle du genre "je parle/tu parles" pour éviter que les interlocuteurs ne commencent à parler au même moment en profitant d'une pause apparente due, en réalité, au retard. Etant donné que le réseau IP constitue un facteur crucial de la qualité de la VoIP, il est facile de se rendre compte de l'importance d'un bon projet de réseau et d'instruments technologiques pouvant garantir une continuité effective de la Quality of Service.