

La Qualité de Service le la Voix sur IP

Principes et Assurance

Introduction

La généralisation des infrastructures IP dans les entreprises s'accompagne du développement de techniques d'amélioration ou de garantie de la qualité des services proposés. Jadis réservé au transport des données, IP est désormais utilisé pour le transport de la voix et on assiste à un réel mouvement dans ce sens sur le marché. Cette tendance se vérifie aussi bien dans les grandes entreprises qui sont nombreuses à franchir le pas que parmi les opérateurs qui offrent des services de voix sur IP.

Si les techniques de mise en œuvre d'IP permettent théoriquement de transporter la voix avec un bon niveau de qualité, encore faut-il que les choix d'architecture, de dimensionnement et de paramétrage permettent d'atteindre l'un des objectifs visé qui est le partage des ressources d'infrastructures entre les données et la voix. En effet des choix inadaptés en terme d'ingénierie de l'infrastructure et du trafic rendront impossible la cohabitation voix/données.

Ce document s'adresse en priorité aux équipes en charge de cette problématique dans les entreprises ou à leurs partenaires. Il décrit comment déterminer si une infrastructure IP écoulant un certain trafic de données va pouvoir supporter un flux additionnel de type voix sur IP. Il présente pour cela comment les principaux paramètres de performance et de qualité des flux IP vont influencer sur la qualité de la voix. Il rappelle sommairement comment les systèmes 5View Applications et 5View VoIP Stream permettent d'apprécier de manière précise la structure des flux applicatifs IP de l'entreprise afin d'évaluer leur impact présent et futur sur l'infrastructure, et de mesurer la qualité actuelle de la VoIP grâce à des simulations de trafic suivant différentes hypothèse d'implémentation de la voix.

Les paramètres de la voix sur IP

Les aspects déterminants pour la qualité de la voix sur un réseau sont le traitement de la voix, la clarté, le délai de bout en bout et l'écho. Ils dépendent des différents composants de la chaîne de transmission, de leur paramétrage, de l'architecture générale de la chaîne, et dans le cas de la VoIP des flux concurrents. Ces aspects sont les suivants :

- Traitement de la voix : lors de l'émission du signal, la voix est traitée, c'est à dire codée et éventuellement compressée, avant d'être transmise
- La clarté est la mesure de fidélité de la voix reçue par rapport à la voix émise
- Le délai de bout en bout est le temps de propagation de la voix à travers le réseau de l'émetteur vers le récepteur
- L'écho est le son émis par l'émetteur qui lui revient

La problématique de qualité de la voix sur IP est particulière car la voix attend de son transporteur autre chose que les données. La transmission de données classique (fichiers, messages, transactions ...) ne supporte aucune perte en ligne sous peine de graves conséquences pour l'interprétation et l'utilisation de ces données par l'équipement récepteur, mais elle supporte en revanche une dérive importante en terme de durée d'acheminement.

Peu importe qu'un paquet arrive avec 100 ms de retard. Le comportement attendu pour la voix est exactement inverse : 1% ou 2% de perte de données de voix en ligne ne sont pas trop gênants pour la qualité du service de VoIP, mais en revanche une variation fréquente de 100 ms sur le délai de transit est catastrophique et rend le service inutilisable.

La voix attend donc du transport IP l'inverse de ce qu'exigent les données. Et cette formulation n'est qu'un raccourci car en fait le transport de la voix exige beaucoup plus : il bénéficiera évidemment de l'intégrité exigée pour le transport des données laquelle est garantie par les réseaux modernes – bien qu'il puisse s'en affranchir dans une certaine limite – mais exigera beaucoup plus au niveau des autres paramètres, notamment en ce qui concerne la stabilité du réseau dans le temps.

Nous présenterons à la suite les principaux paramètres influents en VoIP, dans l'ordre les échantillonnages (codecs), le délai de transit, la gigue de phase et les pertes de données.

Les différents échantillonnages

Le paramètre d'échantillonnage ou codec (pour compression / décompression) est structurant en VoIP. Le codec détermine à quelle vitesse la voix est échantillonnée et dimensionne par la même le flux de données numériques que va générer la transformation d'un échantillon temporel de voix analogique. Les codecs sont répertoriés par leur nom à l'ITU. Les codecs les plus utilisés et leurs vitesses d'échantillonnage sont les suivants :

G.711	64 kbps
G.726	32 kbps
G.726	24 kbps
G.728	16 kbps
G.729	8 kbps
G.723.1 MPMLQ	6.3 kbps
G.723.1 ACELP	5.3 kbps

Le choix du codec est un compromis entre la qualité de service souhaitée et la capacité de l'infrastructure IP à délivrer une bande passante et des paramètres de QoS qui vont impacter cette qualité. Le paramètre le plus déterminant auquel on s'intéresse pour commencer est la bande passante que l'on met en regard du nombre de communications simultanées à écouler. Le tableau suivant permet d'effectuer rapidement le bilan de bande passante en fonction du codec choisi :

Codec	Echantillonnage (codec)			Calcul de bande passante nécessaire									
	Débit (kbps)	Interv. échant. (ms)	Délai échant. (ms)	Volume de données de voix par échant. de codec (octet)	Volume des données de voix dans RTP (octets)	Durée des données de voix dans RTP (ms)	Nb de paq. par sec.	Bande passante IP/UDP/RTP (kbps)	Bande passante ethernet avec IP/UDP/RTP (kbps)	Bande passante IP/UDP/cRTP (kbps)	Bande passante ethernet avec IP/UDP/cRTP (kbps)	Bande passante RTP/IP pour 10 canaux (kbps)	Bande passante RTP/IP pour 32 canaux (kbps)
G.711	64	20	1	160	160	20	50	80,0	87,2	65,6	72,8	800	2560
G.726	32	20	1	80	80	20	50	48,0	55,2	33,6	40,8	480	1536
G.726	24	20	1	60	60	20	50	40,0	47,2	25,6	32,8	400	1280
G.728	16	20	25	40	60	30	33	26,7	31,5	17,1	21,9	267	853
G.729	8	20	25	20	20	20	50	24,0	31,2	9,6	16,8	240	768
G.723.1	6,3	30	67.5	24	24	30	33	17,1	21,9	7,5	12,3	171	546
G.723.1	5,3	30	67.5	20	20	30	33	16,0	20,8	6,4	11,2	160	512

Le choix du codec G.711 permet de bénéficier à réseau constant de la meilleure qualité de service, tandis que les compressions G.726, G.728, G.729 et G.723 apportent avec elles des diminutions initiales de la QoS, immédiatement reflétées dans le score MOS de mesure de la qualité que nous étudierons plus avant.

Le délai de transit

Le délai de transit (ou *end-to-end delay* dans la dénomination anglo-saxonne) est un des paramètres critiques influençant fortement la QoS d'un service de voix sur IP. C'est le temps que va mettre en moyenne un paquet IP contenant un échantillon de voix pour traverser l'infrastructure entre deux interlocuteurs. Ce temps de transit comporte quatre composantes :

- Le délai d'échantillonnage
- Le délai de propagation
- Le délai de transport
- Le délai des buffers de gigue

Le délai d'échantillonnage est la durée de numérisation de la voix à l'émission puis de conversion en signal voix à la réception. Ce temps dépend du type de codec choisi et varie de quelques millisecondes avec le codec G.711 (échantillonnage 64 kbps) à plus de 50 ms en G.723 (échantillonnage 6,3 ou 5,3 kbps). C'est une des raisons pour laquelle le choix du codec impacte le score MOS d'appréciation de la clarté de la voix, indépendamment des autres caractéristiques de l'infrastructure.

Le délai de propagation est la durée de transmission en ligne des données numérisées. Cette durée est normalement très faible par rapport aux autres composantes du délai de transit, de l'ordre de quelques millisecondes.

Le délai de transport est la durée passée à traverser les routeurs, les commutateurs et les autres composants du réseau et de l'infrastructure de téléphonie IP. L'ordre de grandeur est de plusieurs dizaines de millisecondes, voir centaines de millisecondes.

Le délai des buffers de gigue est le retard introduit à la réception en vue de lisser la variation de temps de transit, et donc de réduire la gigue de phase. L'ordre de grandeur est de 50 ms. Les éléments d'infrastructure, notamment les routeurs, peuvent également mettre en œuvre des buffers de gigue.

La gigue de phase

La variation de temps de transit, ou gigue de phase, est la conséquence du fait que tous les paquets contenant des échantillons de voix ne vont pas traverser le réseau à la même vitesse. Cela crée une déformation de la voix ou un hachage.

La gigue de phase est indépendante du délai de transit. Le délai peut être court et la gigue importante ou inversement. La gigue est une conséquence de congestions passagères sur le

réseau, ce dernier ne pouvant plus transporter les données de manière constante dans le temps. La valeur de la gigue va de quelques ms à quelques dizaines de ms.

La perte de données

La transmission de la voix par paquets s'appuie sur le protocole RTP (real-time transport protocol). Ce dernier permet de transmettre sur IP les paquets de voix en reconstituant les informations même si la couche de transport change l'ordre des paquets. Il utilise pour cela des numéros de séquence et s'appuie sur UDP.

Les contraintes temps réel de délai de transit évoquées plus haut rendent inutile la retransmission des paquets perdus : même retransmis un datagramme RTP arriverait bien trop tard pour être d'une quelconque utilité dans le processus de reconstitution de la voix. En voix sur IP on ne retransmet donc pas les données perdues. Ces pertes de données VoIP sont dues aux congestions sur le réseau, qui entraînent des rejets de paquets tout au long du réseau, ou à une gigue excessive qui va provoquer des rejets de paquet dans les buffers de gigue du récepteur, ceux-ci ne pouvant pas accueillir tous les paquets arrivés en retard.

Une perte de données régulière mais faible est moins gênante en voix sur IP que des pics de perte de paquets espacés mais élevés. En effet l'écoute humaine s'habitue à une qualité moyenne mais constante et en revanche supportera peu de soudaines dégradations de la QoS.

Le taux de perte en VoIP est typiquement de quelques pourcents ou dixièmes de pourcent.

Prédiction de la qualité de service, E-model et score MOS

Un des intérêts de réaliser des simulations de trafic en avant projet de déploiement de voix sur IP est de pouvoir calculer une valeur estimée de score MOS.

La méthode MOS est une mesure subjective de la QoS voix par des opérateurs humains. Elle n'a pas pour objet de fournir des données objectives mais d'obtenir une appréciation de la clarté de la voix reçue grâce à une enquête auprès d'un panel d'utilisateurs ou d'opérateurs représentatifs. Cette méthode est définie par la spécification ITU P800 (MOS Mean Opinion Score, pour « note moyenne d'appréciation »). Son application est coûteuse et délicate, mais elle reste le « juge de paix » en dernier recours lors de problèmes de plainte des utilisateurs du service de téléphonie. Elle s'applique aussi bien à la téléphonie traditionnelle qu'à la ToIP.

Calcul du facteur R de l'E-model

L'ETSI a développé un modèle de calcul de la qualité de transport de la voix de bout en bout, de la bouche de l'émetteur à l'oreille du récepteur, connu sous le nom de E-model (référence ETSI : ETR 250). Ce modèle a été standardisé par l'ITU sous la référence G.107. Le principe de l'E-model consiste à calculer une grandeur unique R en fonction des paramètres que nous avons précédemment décrits dans ce document. La formule simplifiée du calcul de R est la suivante :

$$R = R_0 - I_s - I_d - I_e + A$$

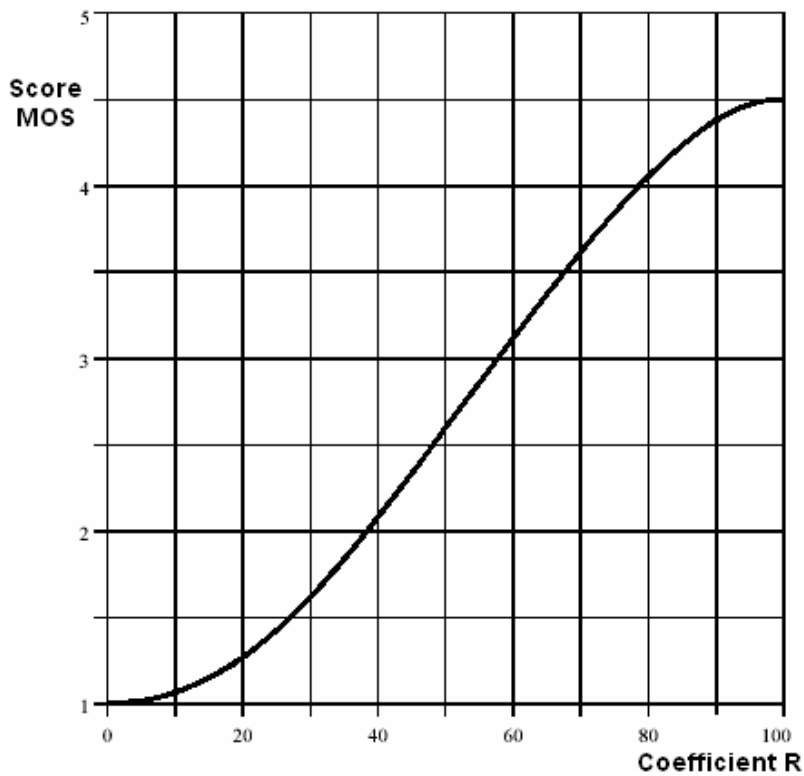
Le principe de la formule est de partir d'un certain capital de QoS, égal à R_0 , et de lui imputer les dommages causés par les différents aspects de la transmission.

R_0 : coefficient initial signal / bruit, « capital initial de QoS », égal à 94,3 en VoIP.
 I_s : coefficient de dommages simultanés avec l'émission de la voix (bruit de fond ...)
 I_d : coefficient de dommages dus au délai de transmission et de transport
 I_e : coefficient de dommage de distorsion causés par les équipements
A : coefficient d'amélioration

Le facteur R ainsi calculé de 0 à 100 permet de déduire directement un coefficient MOS de 0 à 5. Dans la réalité les valeurs de R oscillent entre 50 et 93,2, soit la limite basse acceptable pour le récepteur, et la limite haute liée aux possibilités techniques de transformation de la voix humaine en signal. La moyenne de la valeur R sur les services en production se situe dans la fourchette 70 à 80.

Conversion R / MOS

Le graphique suivant présente la courbe $MOS = f(R)$. Le facteur MOS est la perception humaine tirée d'un protocole de test précis. L'inflexion vers le haut indique qu'à partir d'un certain niveau de qualité, l'augmentation de celle-ci, reflétée par le facteur R, est moins perçue par l'utilisateur. On observe le même phénomène en bas de la courbe, une diminution de R étant moins perçue par l'utilisateur quand la qualité est déjà très dégradée.



Courbe de conversion R / MOS

Le tableau suivant présente la correspondance entre les valeurs de R et la qualité de la voix transmise (nous avons conservé la terminologie anglo-saxonne de la recommandation ITU-T) :

Coefficient R	Qualité de transmission de la voix
90-100	best
80-90	high
70-80	medium
60-70	low
0-60	very poor

Coefficients de la formule de calcul de R

- R_0 est la valeur que l'on obtiendrait si la transmission était parfaite. C'est le « capital initial de QoS ». Comme le simple fait de numériser la voix à l'émission pour la reconvertir en signal analogique à la réception provoque une dégradation, la recommandation attribuée à R_0 a une valeur par défaut de 94,3, correspondant à une valeur MOS de 4,5.

- I_s intègre les dommages qui sont simultanés à l'émission de la voix, dus notamment aux conditions d'émission.

Remarque : I_s et R_0 ne diffèrent pas entre la téléphonie classique et la ToIP.

- I_d intègre le délai de transit comprenant toutes les composantes citées plus haut sauf le délai d'échantillonnage.

- I_e intègre la probabilité qu'un paquet soit retransmis sur le réseau, ainsi que les facteurs de distorsion introduits par le codage de la voix.

Le codec introduit donc une composante dans I_e , dont les valeurs sont données ci-après :

Codec	Débit codec	Coefficient I_e	Facteur R « intrinsèque »
G.711 PCM	64 kbps	0	94,3
G.726 ADPCM	32 kbps	7	87,3
G.726 ADPCM	24 kbps	25	69,3
G.728 LD-CELP	16 kbps	7	87,3
G.729 CS-ACELP	8 kbps	10	84,3
G.723.1 MP-MLQ	6.3 kbps	15	79,3
G.723.1 ACELP	5.3 kbps	19	75,3

En cas de transcodages multiples les coefficients I_e se cumulent, ce qui rend cette opération extrêmement coûteuse en terme de bilan de QoS.

Le taux de perte de données influera d'autant plus sur I_e que l'on aura choisi un codec lent.

- A est un coefficient de prise en compte de facteurs d'amélioration du réseau. Le principe est de considérer que le fait d'avoir une facilité d'accès au service de téléphonie permet de supporter quelques désagréments : par exemple $A = 10$ pour les mobiles. On accepte sur un téléphone portable des imperfections que l'on ne tolérerait pas en téléphonie fixe.

E-model et simulation de trafic

En introduisant dans la formule de calcul du coefficient R les valeurs mesurées en simulation de trafic VoIP on obtient une très bonne approximation du score MOS tel que l'on peut le mesurer en situation réelle.

S'assurer de la qualité de la voix sur IP

Afin de s'assurer de la qualité de voix avant déploiement sur un environnement IP existant, on est amené à travailler dans quatre directions :

- Analyser le trafic voix existant à basculer en VoIP et déterminer le niveau de qualité recherché
- L'analyse de l'architecture et du paramétrage de l'infrastructure IP existante
- La réalisation d'une expertise des flux sur la chaîne de transmission à emprunter par la VoIP suivie d'une simulation de flux de voix
- L'étude du paramétrage des futurs équipements de ToIP

La liste des questions et vérifications à traiter comprend entre autres les points suivants :

- Déterminer le nombre de canaux de voix à établir et le niveau de QoS voix recherché
- inventorier les éléments de la chaîne de transmission appelée à transporter les flux VoIP
- connaître les éléments de paramétrage de QoS IP et autres sur la chaîne de transmission existante
- connaître le taux d'utilisation des éléments précédents
- expertiser sur une période significative la structure des flux IP concurrents des futurs flux VoIP sur la chaîne considérée
- répertorier les scénari de dégradation de la QoS IP à partir de l'expertise des flux réalisée
- connaître les éléments possibles de paramétrage des éléments de ToIP sur la future chaîne de transmission
- procéder à une simulation de trafic VoIP de bout en bout sur la chaîne de transmission avec différentes hypothèses d'appels simultanés et éventuellement différentes hypothèses de codec
- mesurer les paramètres critiques de délai de transit, de gigue de phase et de perte de données, et en déduire le niveau de QoS estimé sur la base des hypothèses établies

- formuler des recommandation de paramétrage des éléments de ToIP

Les recommandations pour assurer la qualité de la voix sur IP

Les valeurs cibles

Nous avons présenté plus haut les principaux paramètres déterminants pour le niveau de qualité d'un service de voix sur IP. Pour s'assurer qu'une infrastructure IP peut recevoir un tel service il faut vérifier l'obtention régulière dans le temps de valeurs générant un bon niveau d'écoute humaine (niveau **Bon**). Dans le cadre d'un étude préliminaire visant à qualifier un réseau IP, il faudra tenir compte des composants de ToIP non encore présents sur le réseau mais qui auront un impact sur le délai de transit dès l'infrastructure complètement déployée, tels que les passerelles et les postes téléphoniques IP. Des mesures ou simulations qui n'intégreraient pas des éléments à venir seraient incomplètes.

Les spécifications G.114 et G.131 de l'ITU-T fournissent la valeur recommandée pour le délai de transit de la voix de bout en bout, à 150 ms.

Attention : ladite recommandation est valable pour le transport de la voix analogique ou non compressée G.711 à 64 kbps. En VoIP si l'on choisit un codec plus lent la valeur de délai de transit de 150 ms ne garantira pas la même QoS qu'en G.711.

Le tableau ci-après présente les seuils de valeurs pour les paramètres critiques et les conséquences constatées pour le niveau de service de VoIP en codec G.711 64 kbps :

	Bon	Moyen	Mauvais
Délai de transit	D < 150 ms	150 ms < D < 400 ms	400 ms < D
Gigue de phase	G < 20 ms	20 ms < G < 50 ms	50 ms < G
Perte de données	P < 1%	1% < P < 3%	3% < P

Choix d'architecture et de configuration IP

L'architecture réseau devra offrir une bande passante suffisante pour le nombre de conversations de voix que l'on souhaite satisfaire simultanément. Contrairement à ce qui se produit avec les données où les utilisateurs peuvent tolérer un certain niveau de dégradation des temps de réponse, le manque de bande passante interrompt immédiatement le service de voix sur IP. Il n'est donc pas possible d'économiser la bande passante en la maintenant en de ça du bilan technique imposé par le nombre de communications à assurer et par leur besoin unitaire en débit.

Comme l'exigence de la VoIP est importante en ce qui concerne le délai de transit et la stabilité des paramètres de QoS IP, on affectera à la voix une classe de service correspondante, c'est à dire de type *Gold* pour reprendre la terminologie en vigueur chez les fournisseurs de service VPN IP.

Sur les réseaux IP avec des débit moyens ou faibles (artères de débit inférieur à 1 Mbps) on activera avec profit la segmentation des paquets IP. Cela permet de diminuer la taille moyenne des paquets de données (non voix) sur le réseau et de réduire les délais d'attente des petits paquets de voix.

Toujours dans le cas des réseaux avec des débits faibles ou moyens, lorsque qu'un projet de voix sur IP se trouve en situation difficile par manque de bande passante et qu'il n'est pas possible d'augmenter cette dernière, on peut avoir recours à la compression des 40 octets d'entête RTP (mécanisme dit cRTP) pour arriver à moins de 5 octets d'entête. Cette solution à l'inconvénient d'augmenter le délai de transit, et elle est donc réservée à des réseaux rapides à l'échelle des besoins de la voix sur IP.

Choix d'échantillonnage

Sauf à être contraint par un besoin impératif de gain de bande passante, on préférera le codec G.711 qui assure la numérisation la plus rapide de la voix ainsi que la plus grande clarté pour le récepteur.

Le choix de codec lents doit se faire en connaissance de cause. S'il est viable sur les réseaux téléphoniques traditionnels, il s'avère généralement très pénalisant en VoIP.

Mise en œuvre du mode PLC

Le mode PLC (pour Packet Loss Concealment) permet d'améliorer la clarté, en atténuant l'effet négatif sur le rendu de la voix causé par la perte aléatoire de quelques paquets. Il consiste à remplacer la portion du signal de voix manquant par une extrapolation de la portion précédente de l'information, ce qui revient à lisser un signal incomplet. Le PLC améliore la qualité du signal dégradé suite à des pertes aléatoires de paquets, mais n'est d'aucune efficacité en cas de pic de perte de données dû à une congestion du réseau. Le PLC ne consommant quasiment aucune ressource dans le bilan global du délai de transit on mettra en œuvre ce mécanisme pour améliorer la régularité de la clarté.

Utilisation de 5View Applications et 5View VoIP Stream

Le système 5View VoIP Stream d'Accellent permet de simuler un trafic téléphonique VoIP basé sur RTP. Il permet de simuler 100 communications simultanées et de produire les mesures nécessaires à la qualification d'une infrastructure IP appelée à transporter de la voix. Il permet également de mesurer automatiquement la note MOS à intervalles réguliers.

5View VoIP Stream est généralement utilisé conjointement avec le système 5View Application qui permet d'analyser le comportement des applications en réseau en même temps que l'on simule du trafic de voix. Il est de la sorte possible de mesurer l'impact futur du trafic de voix sur les applications métiers ou autres empruntant le même réseau IP.

Les outils et services d'Accellent permettent de réaliser aisément l'étape d'expertise des flux qui comprend l'analyse de l'existant et la projection d'un trafic de voix. L'avantage essentiel apporté par l'approche d'Accellent est de vérifier de manière précise l'interaction entre le futur service de voix et le comportement du référentiel applicatif en place.

Pour toute information complémentaire, contactez info@accellent-group.com ou www.accellent-group.com.

accellent

Le Descartes I
29, Promenade Michel Simon
93160 Noisy le Grand
Télécopie : 01 48 15 54 89
E-mail : info@accellent-group.com
www.accellent-group.com