**UNIVERSITE KASDI MERBAH, OUARGLA**

***Faculté des Nouvelles Technologies de l'Information et de la Communication - FNTIC***

***Département de l'Informatique et des Technologies de l'information***

N° d'Ordre : .....
N° de Série : .....

# Mémoire

En vue de l'obtention du diplôme de

## Magistère en Informatique

(**Option :** Technologies de l'Information et de la Communication)

**Présenté par :**

**Abdellatif CHERADID**

**THEME**

# *Etude des approches adaptatives liées à la QoE dans le cadre des applications de Téléphonie sur IP*

Soutenu le : / / devant le jury :

| | | |
|---|---|---|
| **Dr. Said Mohamed SAID** | **Maître de Conférences Habilité, Université de Ouargla** | **Président** |
| **Pr. Brahim BELATTAR** | **Professeur, Université de Batna** | **Examinateur** |
| **Dr. Ahmed KORICHI** | **Maître de Conférences Habilité, Université de Ouargla** | **Examinateur** |
| **Dr. Driss KORICHI** | **Maître de Conférences Habilité, Université de Ouargla** | **Rapporteur** |

**Année Universitaire 2013-2014**

# *Remerciements*

*Louange et remerciement à Allah tout puissant qui m'a donné la volonté, l'aide, la patience et le courage pour accomplir ce modeste travail.*

*Mes sincères remerciements au Dr. Dris KORICHI, Maître de Conférences Habilité à l'université de Ouargla, qui a proposé le sujet et accepté de m'encadrer. Je le remercie d'avoir compris et soutenu mes efforts pour mener à bien ce travail de thèse.*

*Je tiens à remercier chaleureusement les membres du jury : Monsieur Pr. Brahim BELATTAR, Professeur à l'université de Batna, et Monsieur Dr. Ahmed KORICHI, Maître de Conférences Habilité à l'université de Ouargla, pour avoir accepté de juger ce travail. Je remercie également Monsieur Dr. Said Mohamed SAID, Maître de Conférences Habilité à l'université de Ouargla, pour m'avoir fait l'honneur de présider mon jury de thèse.*

*Je remercie particulièrement Mr.Med Salah HADJI pour toute son aide. Aussi, je tiens à remercier mon ami et mon frère Bachir MESSAID ; je le remercie pour le temps qu'il m'a accordé dans les moments les plus durs.*

*Je remercie également les enseignants de notre promotion pour la qualité de leur enseignement et la disponibilité dont ils ont fait preuve à notre égard.*

*Abdellatif CHERADID*

# ملخص

تعتبر في العصر الحديث خدمة أو خدمات الصوت عبر بروتوكول الإتصال أي بي **IP** من الخدمات الواسعة الاستعمال في المجال الشخصي الاحترافي على حد السواء. كما أنها سوف تمثل الخدمات الأساسية التي سترتكز عليها شبكات إتصال الجيل القادم. لذا ومن المنطقي قياس وتحسين كل الوسائط والمعاملات التي تتعلق بهذا النوع من الخدمات وهذا لتأمين نوعية خدمة جيدة. هذا المصطلح الذي يعرف على أنه مجموعة من الخدمات الواجب توفيرها من طرف الشبكة أثناء النقل. أيضا لتأمين نوعية خبرة جيدة وثابتة، وهنا نعني بنوعية الخبرة، نوعية الخدمة حسب الإدراك الشخصي لمستخدمي خدمات الشبكة مع وظائف الشبكة مثل إدارة الموارد والتحكم بالولوج ومراقبة التدفق المعلوماتي لدعم المستخدمين لشبكات الجيل القادم للوسائط المتعددة .

في هذا العمل نقترح آلية جديدة للتوجيه بإستعمال نوعية الخبرة في الشبكات التي تعتمد خدمة الصوت عبر البروتوكول **IP**. هته الآلية تعتمد على مبدأ التحسين والتوجيه باستعمال فكرة "مستعمرات النمل" وباستعمال درجة تقييم نوعية الصوت باستخدام طريقة التقييم الموضوعي المسماة النموذج E-Model والمقترحة من طرف الاتحاد الدولي للاتصالات (ITU G.107)

**كلمات دالة: نوعية الخدمة، نوعية التجربة، طريقة قياس جودة الصوت E-Model , التحسين باستعمال طرق "مستعمرات النمل".**

# *Résumé*

Les services de la voix sur IP (VoIP) sont actuellement présents et utilisés dans nos activités quotidiennes, personnelles et professionnelles. Ils sont certainement parmi les principaux services dans des réseaux de prochaine génération (NGN). Par conséquent et dans ce contexte, la qualité de la livraison pour les besoins des services de la VoIP doivent être mesurés et optimisés pour assurer la Qualité du Service (QoS). Ceci permet aux opérateurs de garder et d'attirer de nouveaux clients. La QoS représente l'ensemble le service prérequis à remplir par le réseau lors du transport d'un flux. En revanche, la nouvelle notion de « qualité de l'expérience » (QoE) représente la perception subjective des utilisateurs des services du réseau avec des fonctions du réseau telles que le contrôle d'admission, la gestion des ressources, le routage, le contrôle de flux, …etc. Ceci permis un soutien aux utilisateurs dans les futurs systèmes de réseaux multimédias.

Dans ce travail, nous proposons un mécanisme de routage conduit par la QoE des utilisateurs finaux. Il s'agit d'un protocole de routage pour les application Voix sur IP basé sur le principe d'optimisation et de routage par colonie de fourmis et conduit par le degré d'évaluation de la qualité de la voix en utilisant la méthode d'évaluation objective E-Model proposée par l'Union Internationale des Télécommunications (ITU-T Recommandation G.107).



***Mots-clés*** : Qualité de service (QoS), Qualité d'expérience (QoE), E-model, Colonie de fourmis

# *ABSTRACT*

Voice over IP services are currently present in our personal and professional activities and will be the principal services in the Next Generation Networks (NGN). Consequently and in order to keep and attract new customers, the quality of the delivery so that the needs of VoIP services are measured and optimized to ensure the quality of the service “QoS”. QoS represents the whole of pre-necessary service to fill by the network during transport of a flow. The quality of experience “QoE” means the subjective user’s perception of the networks services with functions of the network, such as the control of admission, stock management, the routing, the control of flow, etc… support users in future multimedia systems networks.

In this work, we propose a routing mechanism driven by the end-user QoE; It's a routing protocol for IP voice applications based on the principle of optimization and routing ant colony and driven by the degree of evaluation of voice quality using the E-Model objective evaluation method proposed by the International Telecommunication Union (ITU-T Recommendations G.107).



***KEYWORDS***: *Quality of Service (QoS), Quality of Experience (QoE), E_model, Ant Colony optimization.*

# Liste des Figures

# Liste des Tables

# • *Intoduction Générale*

# Introduction Générale

Les services de la Voix sur IP (VoIP) sont maintenant offerts par différents prestataires de service et souscrits par un grand nombre d'utilisateurs fixes et mobiles dans les systèmes de gestion des réseaux multimédias. La VoIP apporte des gains pour les fournisseurs de services ainsi pour les clients. Pour les fournisseurs de services, les coûts opérationnels sont réduits en raison de la distribution des différents services, voix et données, qui utilisent la même infrastructure réseau avec des ressources partagées. Pour les clients, la VoIP présente les caractéristiques d'un réseau public traditionnel de téléphonie commuté (RTC), comme la messagerie vocale et la conférence vocale, d'une manière omniprésente ainsi que le soutien au niveau qualité et bien évidemment un prix abordable.

# Problématique

Contrairement au réseau téléphonique public, un réseau IP n'est pas une connexion « Point à Point ». Le « chemin » entre émetteur et récepteur est déterminé « à la volée ». Les paquets de données peuvent donc subir des retards ou devoir être transmis à nouveau. Si cela a peu d'importance pour le courrier électronique, qui est lu une fois tous les paquets assemblés, cela peut avoir des conséquences désastreuses quand il s'agit d'une conversation téléphonique en temps réel. Par conséquence, l'amélioration de la qualité des services comme perçue par les utilisateurs, désignée généralement par le nom de la qualité de l'expérience (QoE), a un grand effet pour les fournisseurs de services avec l'objectif de minimiser le taux de désabonnement tout en conservant leur avantage concurrentiel. Le terme QoE a été introduit, en combinant la perception utilisateur, l'expérience et les espérances ainsi que les paramètres liés à la qualité de service.

Le sujet de ce mémoire s'inscrit dans cette problématique. En d'autres termes, il s'agit d'étudier le routage orienté qualité de service et particulièrement celui orienté délai dans les réseaux. Nous nous sommes intéressés plus particulièrement à l'étude du routage adaptatif qui est par nature orienté qualité de service. Le calcul de la table de routage dans ce type de routage est en effet basé sur des agents explorateurs chargés de rassembler l'information sur l'état des liens et procéder par la suite à des mises à jour au niveau des nœuds du réseau. Le routage adaptatif est dit aussi probabiliste. Il se base sur une fonction de renforcement qui permet la mise à jour des entrées de la table de routage et ainsi la construction des routes optimales.

## Plan de la mémoire

Dans le premier chapitre, nous présentons une étude bibliographique de l'état de l'art sur la Voix sur IP en se focalisant sur les principes de ce type de service ainsi les protocoles standards utilisés, la sécurité et les applications VoIP (softphone).

Le deuxième chapitre sera consacré aux notions de base liées à la qualité de service dans les réseaux informatiques, ainsi que la notion de la qualité perçue de l'utilisateur d'un service réseau et les déférents types d'applications qui demandent la perception utilisateur. Nous présentons également divers méthodes d'évaluation de la qualité de la voix.

En se basant sur les techniques proposées dans la littérature, nous abordons ensuite une étude sur les approches de routage adaptatif les plus représentatives. Nous nous sommes intéressés à leurs différents modes de fonctionnement.

Le dernier chapitre est consacré au protocole que nous proposons. Après avoir présenté nos motivations, nous décrivons l'ensemble des modules qui définissent le fonctionnement de ce dernier, et nous abordons par la suite l'implémentation et l'analyse des performances de notre protocole. La validation et l'évaluation des performances de l'algorithme proposé ont été réalisées sur différentes topologies en utilisant l'outil de simulation Network Simulator 2(NS-2).

La conclusion générale et les perspectives possibles à notre travail sont présentées dans la conclusion générale.

# Chapitre I

- *VoIP et ToIP*
  *Principes et Standards*

# I. VoIP et ToIP : Principes et Standards

## I.1. Introduction

La Voix sur IP, ou «VoIP» pour Voice over IP, est une technique qui permet de communiquer par la voix sur des réseaux IP que ce soit des réseaux privés ou sur Internet. Cette technologie est notamment utilisée pour supporter le service de téléphonie sur IP («ToIP» pour Telephony over Internet Protocol).

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, opérateurs, entreprises, 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, 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 peut donc, grâce à la VoIP, tout fusionner sur un même réseau. La téléphonie devient alors 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, 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,...). Jusqu'au 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 les différents types de données sur un même réseau permet en plus, de simplifier son administration.

Comme toute innovation technologique, 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 avec la téléphonie classique. En particulier, lorsque les interlocuteurs sont éloignés, ou la différence de prix devienne importante. De plus, la téléphonie sur IP utilise jusqu'à dix fois moins de bande passante que la téléphonie

traditionnelle. Ceci apporte de grand intérêt pour la voix sur réseau privé. Les entreprises après avoir émis un certain nombre de doutes sur la qualité de services, sont désormais convaincues de la plus grande maturité technologique des solutions proposées sur le marché. Qu'il s'agisse d'entreprises mono-site ou multi-sites, les sondages montrent que le phénomène de migration vers les systèmes de téléphonie sur IP en entreprise est actuellement engagé.

## I.2. Réseau Téléphonique Commuté (RTC)

Le Réseau Téléphonique Commuté (ou RTC) est le réseau du téléphone (fixe et mobile), dans lequel un poste d'abonné est relié à un central téléphonique par une paire de fils alimentée en batterie centrale (la boucle locale). Les centraux sont eux-mêmes reliés entre eux par des liens offrant un débit de 2Mb/s: ce sont les Blocs Primaires Numériques (BPN)[ Seb, 2004].

Dans le cas d'un réseau construit par un opérateur public, on parle parfois de Réseau Téléphonique Commuté Public (RTCP) ou PSTN, de l'anglais (Public Switched Telepon Netwrok)

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

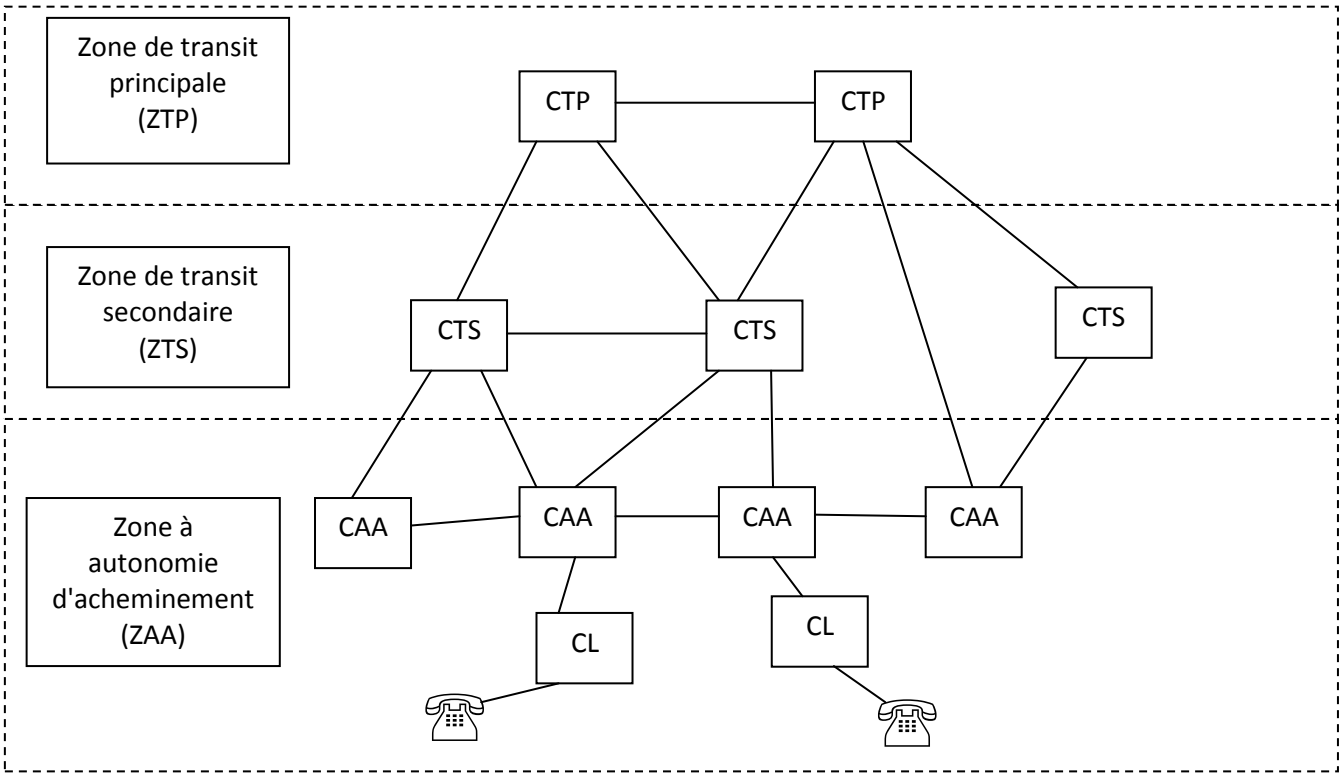



*Fig I-1 : Organisation hiérarchique d'un réseau téléphonique commute (RTC)*

## I.3. Historique de la voix sur IP

Y.HADDADE à chronologie l'histoire de la voix sur IP comme suit [Haddade 2006] :

En décembre 1966, S. Saito et F. Itakura publient un rapport pour le laboratoire de communications électriques du NTT de Tokyo. Ils décrivent une approche statistique pour coder la voix. En parallèle, la même année, Glen Culler propose son système temps réel à l'université de Santa Barbara en Californie. Ce système permet le traitement de signaux temps-réel sur les terminaux des étudiants. On peut dire que c'est le premier système de traitement du signal numérique temps-réel (DSP: Digital Signal Processing) dans une salle de classe.

En 1969, le premier réseau nommée ARPANET, avec quatre terminaux fait son apparition. Les premiers efforts dans le but de développer la transmission de paquets de données contenant des échantillons de voix sur le réseau ARPANET furent initiés en 1972 par Bob Kahn. Pour explorer les possibilités de transmission de paquets audio sur ARPANET, Kahn forme le groupe NSC (Network Secure Communications).

En 1974, le protocole NVP (Network Voice Protocol) est présenté par Danny Cohen et son équipe. Il est expliqué clairement comment la parole en temps-réel peut être transmise sur le réseau ARPANET.

En 1995, la compagnie " Vocaltec" lance le premier logiciel de téléphonie par internet. Ce logiciel était conçu à la base pour permettre une communication entre deux PC équipés chacun d'une carte son, microphone et écouteur. Vocaltec connut un premier succès avec le produit "Internet Phone" puis en 1996 avec IPO. Ce fut le "Skype" du milieu des années 1990. L'inconvénient majeur à cette époque fut le manque d'infrastructure large bande permettant les transmissions haut débit, nécessaire à une bonne qualité audio. En conséquence, la téléphonie classique restait encore l'unique service fiable de téléphonie de qualité. Cependant ce fut une étape décisive pour cette nouvelle technologie, dans la mesure où ce fut le premier téléphone sur IP qui n'ait jamais existé.

La VoIP continua à évoluer au fil des années, au point qu'en 1998, des compagnies proposaient déjà des services de téléphonie de PC à téléphone fixe. Les services de téléphone à téléphone apparurent peu de temps après, même s'ils requerraient souvent un ordinateur pour établir la connexion. L'apparition des réseaux large bande permit d'améliorer la clarté des appels et de réduire les délais.

L'envolée de la VoIP arriva lorsque les grands fabricants tels que Cisco Systems et Nortel proposèrent des équipements VoIP permettant la commutation. En d'autres termes il n'était plus nécessaire de passer par un ordinateur pour transmettre les communications sous forme de paquets de données sur le réseau internet.

Depuis 2000, la VoIP a connu une croissance fulgurante. Tandis que les entreprises font transiter systématiquement leurs communications par la VoIP pour économiser sur les longues distances et les coûts d'infrastructures, les services de VoIP se développent

chez les particuliers. Les bénéfices prévus pour la fin 2008 dépassent les 8.5 milliards de dollars.

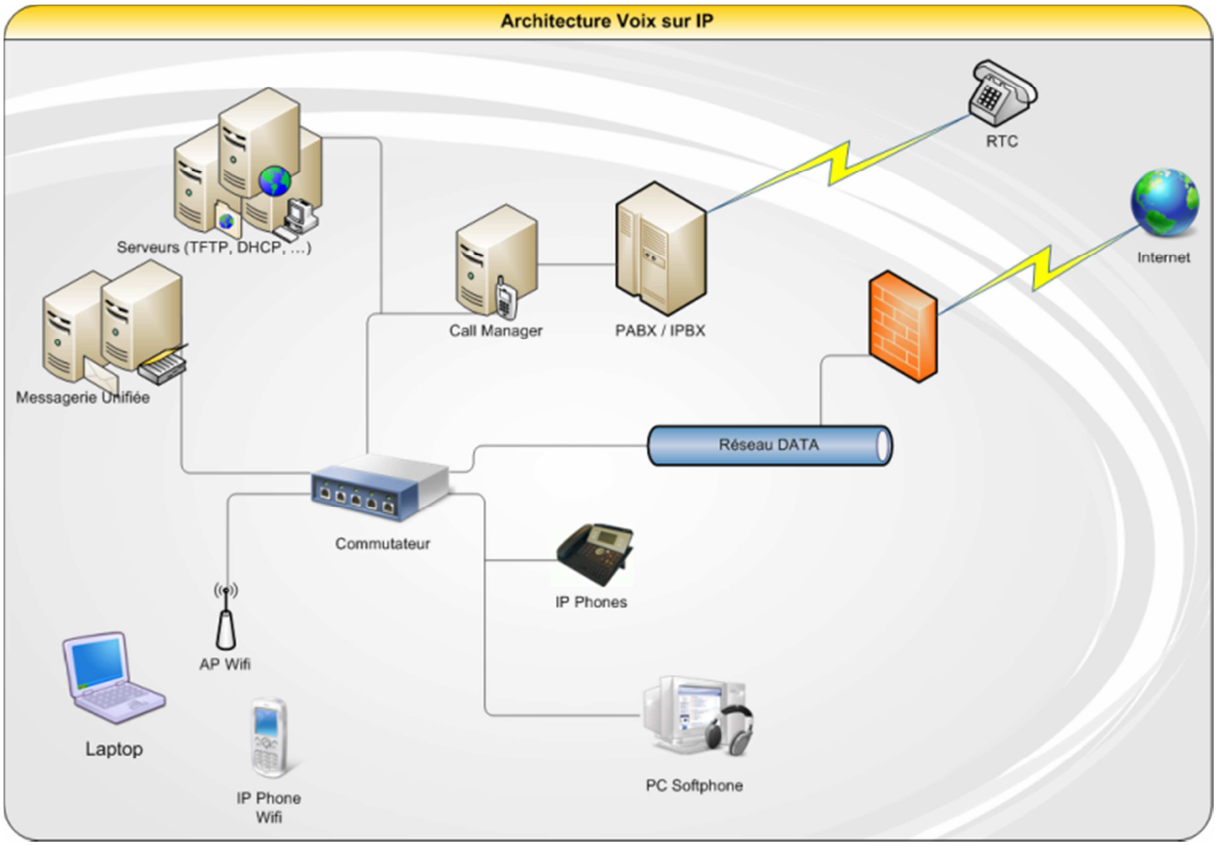



*Fig I-2 : Architecture générale d'un réseau VoIP[Guillet, 2010]*

## I.4. Codage de la voix

Le transport de la voix sur un réseau IP nécessite au préalable tout ou une partie des étapes suivantes [Mellouk, 2009]:

*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) 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 les plus faibles possibles. Traditionnellement on distingue 3 classes de techniques de codage de la voix, le codage forme d'onde, codage paramétrique et codage hybride

***Les codeurs de forme d'onde*** essayent de coder la forme exacte de la forme d'onde de son articulé, sans considérer la nature de la production de la parole et de la perception du langage humain. Ces codeurs appropriés mieux au codage de débit binaire élevé depuis des baisses d'exécution rapidement avec la diminution de débit binaire.

***Le codage paramétrique***, modélise le processus de fabrication du son et extrait les paramètres convenables, qui sont transmis au décodeur. Les voici quisont habitués pour reconstruire une forme d'onde qui est souvent très différente de celle du signal original mais qui produit un son qui est semblable ou près de l'original. Cette technique de codage fonctionne bien pour de bas débits binaires. Augmentant le débit binaire normalement ne traduit pas une meilleure qualité, puisqu'elle est limitée par le modèle choisi.

***Le codage hybride*** combine la force d'un codeur de forme d'onde avec celle d'un codeur paramétrique. Comme un codeur paramétrique, il se fonde sur un modèle de production de la parole. Comme dans des codeurs de forme d'onde, une tentative est faite d'assortir le signal original avec le signal décodé dans le domaine de temps. Cette technique domine les codeurs de débit binaire moyen.

Les codeurs de la parole opèrent habituellement en simples blocs, qui doivent être accumulés avant que le traitement commence. La taille des blocs d'entrée varie avec le codeur de voix utilisé. Par exemple, le codeur G.729 segmente le discours d'entrée dans 80 échantillons (10ms) et assigne 80 bits par bloc codé, menant à un débit binaire de 8 Kbps.

| *Classe de codage* | *Technique de codage* | *Standard* | *Débit binaire* |
|---|---|---|---|
| Forme d'onde | PCM | G.711 | 64 kbps |
| | ADPCM | G.726 | 16-40 kbps |
| Paramétrique | LPC | FS 1015 | 2.4 kbps |
| Hybride | CELP | FS 1016 | 4.8 kbps |
| | LD-CELP | G.728 | 16 kbps |
| | CS-ACELP | G.729 | 8 kbps |

*Tableau I-1 : Les standards de codage de la voix [Mellouk, 2009]*

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 le tableau suivant :

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

*Tableau I-2 : Echelle utilisée pour l'évaluation de la qualité de voix [ITU, 2005]*

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 :

| 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.723.2 | 5.3 | 3.65 |
| GSM | 13 | 3.5 |

*Tableau I-3 : Score MOS des différents codecs [Ouakil et pujolle, 2008]*

## I.5. Les Standards Protocoles de la VOIP

Les principaux protocoles utilisés pour l'établissement des connexions en Voix sur IP sont :

### I.5.1. Le Protocole H.323

Le protocole H.323 regroupe un ensemble de protocoles de communication de la voix, de l'image et de données sur IP. C'est un protocole développé par l'UIT-T. Il est dérivé du protocole H.320 utilisé sur RNIS [Ouakil et pujolle, 2008].

Plus qu'un protocole, H.323 ressemble davantage à une association de plusieurs protocoles différents et qui peuvent être regroupés en trois catégories : la signalisation, la négociation de codec, et le transport de l'information.

Les messages de signalisation sont ceux que l'on envoie pour demander d'être mis en relation avec une autre personne, qui indiquent que la ligne est occupée, que le téléphone sonne… Cela comprend aussi les messages que l'on envoie pour signaler que tel téléphone est connecté au réseau et peut être joint de telle manière. En H.323, la signalisation s'appuie sur le protocole RAS ((en)Registration Admission Status) pour l'enregistrement et l'authentification, et le protocole Q.931 pour l'initialisation et le contrôle d'appel.

La négociation est utilisée pour se mettre d'accord sur la façon de coder les informations qu'on va s'échanger. Il est important que les téléphones (ou systèmes) parlent un langage commun s'ils veulent se comprendre. Il serait aussi préférable, s'ils ont plusieurs alternatives de langages qu'ils utilisent le plus adapté. Il peut s'agir du codec le moins gourmand en bande passante ou de celui qui offre la meilleure qualité. Le protocole utilisé pour la négociation de codec est le H.245

Le transport de l'information s'appuie sur le protocole RTP qui transporte la voix, la vidéo ou les données numérisées par les codecs. On peut aussi utiliser les messages RTCP pour faire du contrôle de qualité, voire demander de renégocier les codecs si, par exemple, la bande passante diminue.

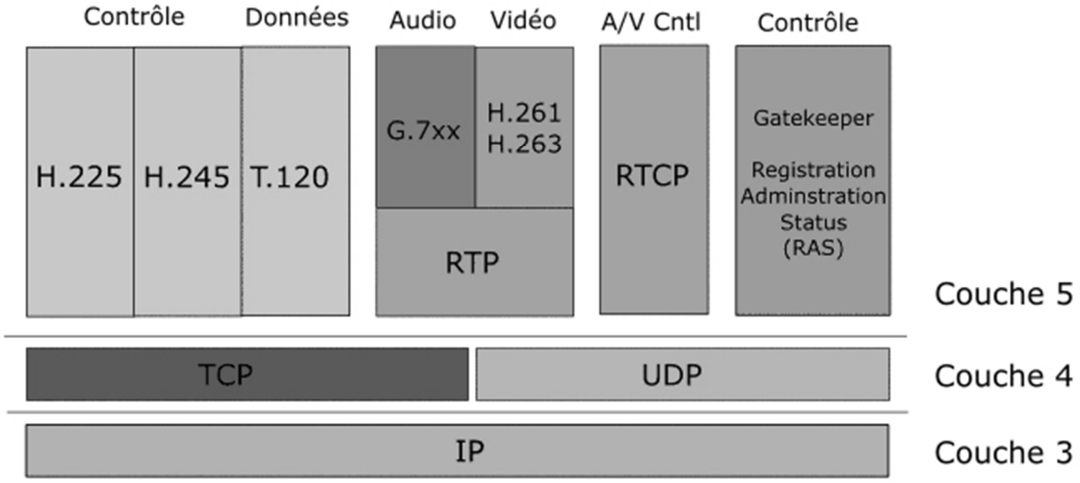



*Fig I-3 : La Pile protocolaire H323 [Ouakil et pujolle, 2008]*

## I.5.2. Le Protocole SIP (Session Initiation Protocol)

*Session Initiation Protocol* (SIP) est un protocole standard ouvert de gestion de sessions souvent utilisé dans les télécommunications multimédia (son, image, etc.). Il est depuis 2007 le plus courant pour la téléphonie par internet (laVoIP).

SIP n'est pas seulement destiné à la VoIP mais aussi à de nombreuses autres applications telles que la visiophonie, la messagerie instantanée, la réalité virtuelle ou même les jeux vidéo

Session Initiation Protocol (dont l'abréviation est SIP) est un protocole normalisé et standardisé par l'IETF (décrit par le RFC 3261 qui rend obsolète le RFC 2543, et est complété par le RFC 3265) qui a été conçu pour établir, modifier et terminer des sessions multimédia. Il se charge de l'authentification et de la localisation des multiples participants. Il se charge également de la négociation sur les types de média utilisables par les différents participants en encapsulant des messages SDP (Session Description Protocol). SIP ne transporte pas les données échangées durant la session comme la voix ou la vidéo. SIP étant indépendant de la transmission des données, tout type de données et de protocoles peut être utilisé pour cet échange. Cependant le protocole RTP (Real-time Transport Protocol) assure le plus souvent les sessions audio et vidéo. SIP remplace progressivement H.323 [Ouakil et pujolle, 2008].

SIP a été étendu afin de supporter de nombreux services tels que la présence, la messagerie instantanée (similaire au service SMS dans les réseaux mobiles), le transfert d'appel, la conférence, les services complémentaires de téléphonie, etc.

SIP a été retenu par le 3GPP pour l'architecture IMS (IP Multimedia Subsystem) comme protocole pour le contrôle de session et de service. Il remplacera à terme les protocoles ISUP (utilisé pour le contrôle d'appel dans le Réseau Téléphonique Commuté) et INAP (utilisé pour le contrôle de service dans l'architecture Réseau Intelligent)

Le protocole SIP n'est qu'un protocole de signalisation. Une fois la session établie, les participants de la session s'échangent directement leur trafic audio/vidéo à travers le protocole RTP (Real-Time Transport Protocol).

Par ailleurs, SIP n'est pas un protocole de réservation de ressource, il ne peut donc pas assurer la QoS. Il s'agit d'un protocole de contrôle d'appel et non de contrôle du média. SIP n'est pas non plus un protocole de transfert de fichier tel que HTTP, utilisé afin de transporter de grands volumes de données. Il a été conçu pour transmettre des messages de signalisation courts afin d'établir, maintenir et libérer des sessions multimédia. Des messages courts non relatifs à un appel peuvent néanmoins être transportés par SIP à la manière des SMS [Ouakil et pujolle, 2008].

SIP définit deux types d'entités: les ***clients*** et les ***serveurs***.

Le serveur proxy (Proxy server), Le serveur de redirection (Redirect server), L'agent utilisateur (UA, User Agent) et L'enregistreur (Registrar)

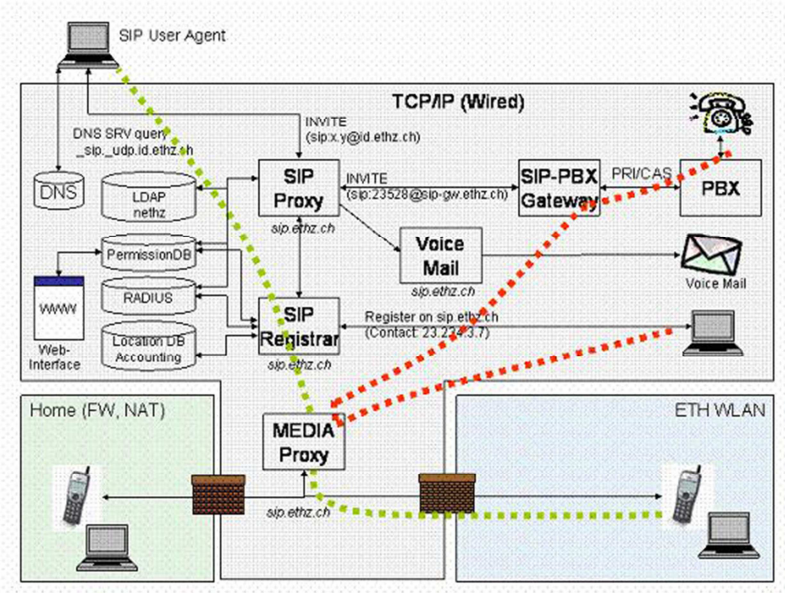



*Fig I-4 : SIP – Présentation Technique[Guillet, 2010]*

### I.5.3. Le Protocole MGCP (Media Gateway Control Protocol)

MGCP est un protocole de signalisation et de contrôle d'appel utilisé dans la Voix sur IP(VoIP) qui généralement interagit avec le réseau téléphonique public commuté (RTC).

MGCP utilise la Session Description Protocol (SDP) pour la spécification et la négociation des flux média à transmettre dans une session d'appel et le Real-time Transport Protocol (RTP) pour l'encadrement des flux multimédia.

MGCP définit les entités suivantes :

• Le ***Call Agent***, qui sert à piloter et administrer les passerelles de manière centralisée.

• Les ***passerelles***, qui maintiennent la connectivité entre réseaux de natures différentes.

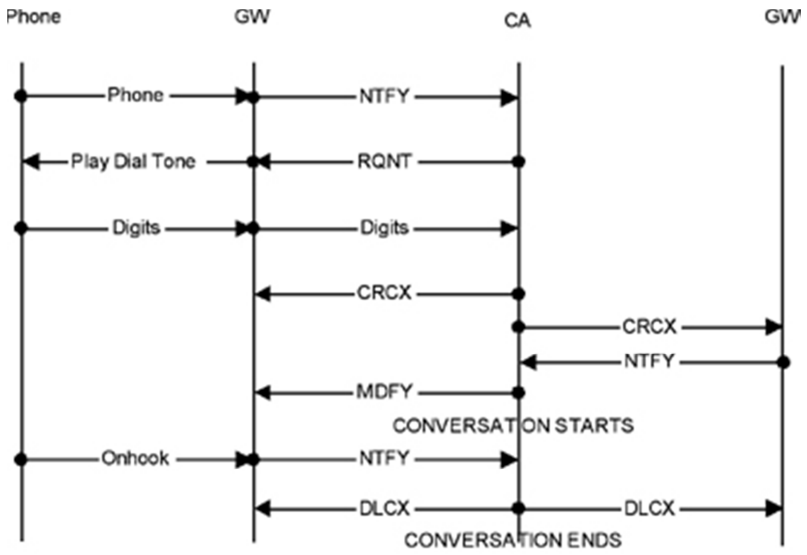



*Fig I-5 : diagramme de flux d'appels mgcp[Ouakil et pujolle, 2008]*

## I.6. Les Protocoles de Transport de la Voix

***RTP (Real-time Transport Protocol)*** [Ouakil et pujolle, 2008] est un protocole qui a été développé par l'IETF afin de faciliter 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 l'UDP ce qui permet d'atteindre plus facilement le temps réel. Les applications temps réel comme la parole numérique ou la visioconférence constituent un véritable problème pour Internet. Qui dit application temps réel, dit présence d'une certaine qualité de service (QoS) que RTP ne garantit 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.

L'entête d'un paquet RTP est obligatoirement constitué de 16 octets. Cet en-tête précède le "payload" qui représente les données utiles.

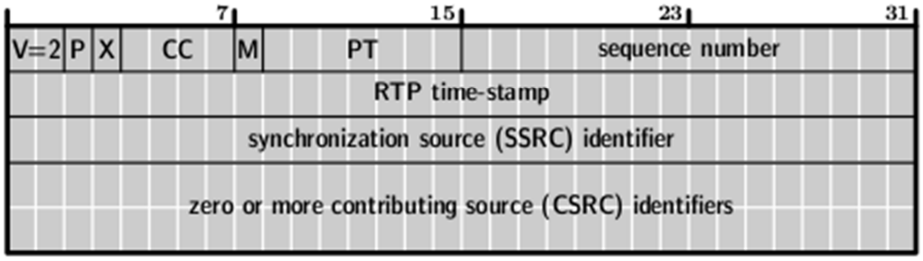



*Fig I-6 : en-tête d'un paquet RTP[Schadle, 2006]*

Le protocole RTP utilise le protocole ***RTCP, Real-time Transport Control Protocol***, [Ouakil et pujolle, 2008] 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 avoir 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.

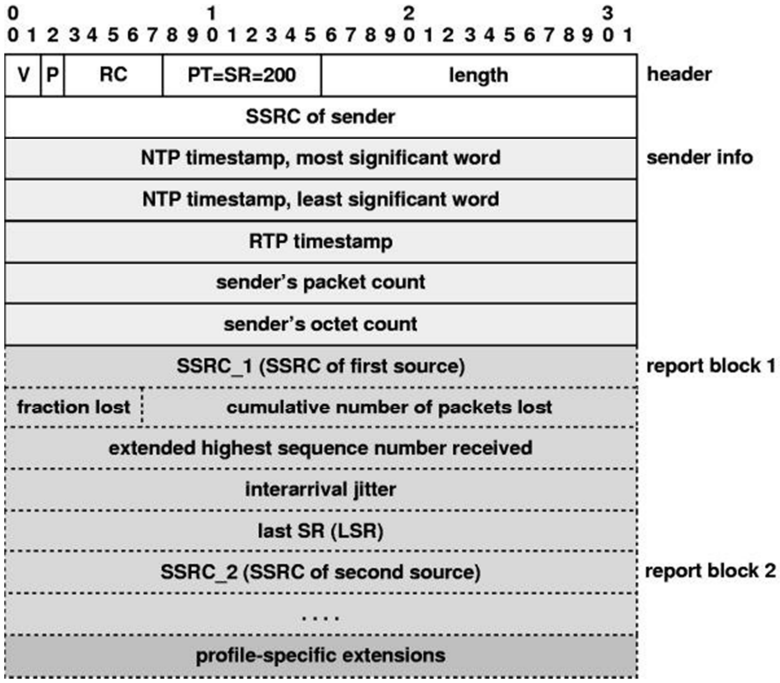



*Fig I-7 : en-tête d'un paquet RTCP[Schadle, 2006]*

## I.7. VoIP et Sécurité

La technologie de ToIP est apparue il y a plus de dix ans. Elle a subi de multiples standardisations internationales, qui, sans la mettre à l'abri des évolutions permanentes, inhérentes aux technologies réseau, la rendent désormais suffisamment mature pour envisager un déploiement à grande échelle. À condition toutefois de maîtriser la sécurité et son intégration au monde du sans-fil.

Les vulnérabilités dont les attaques peuvent tiré parti peuvent avoir cinq origines [Porter et Gough, 2007] :

• les protocoles ;

• les logiciels ;

• le système d'exploitation ;

• l'infrastructure physique ;

• l'erreur humaine.

Chacune d'elles est une source potentielle de faille, qu'il convient d'étudier avec précaution dans la mise en place d'une solution de ToIP.

Une attaque peut avoir trois objectifs :

• Acquisition de service. L'objectif d'une telle attaque est de s'approprier des droits et fonctionnalités qui n'ont pas véritablement été attribués à l'attaquant.

• Interception de service. Cette attaque compromet la confidentialité du service et vise à en analyser ou modifier le contenu.

• Interruption de service. L'objectif est purement de nuire au bon déroulement du service en cherchant à le mettre hors d'usage.

Les attaques de sécurité réseau sont divisées en attaques passives et actives. Ces deux classes sont elles-mêmes divisées en sous-classes. Les conséquences de ces attaques sont le coût légal et de recouvrement et la perte d'informations propriétaires, d'image ou de services réseau.

Une attaque est dite passive lorsqu'un individu non autorisé obtient un accès à une ressource sans modifier son contenu. Les attaques passives peuvent être des écoutes ou des analyses de trafic, parfois appelées analyses de flot de trafic.

Une attaque est dite active lorsqu'un parti non autorisé apporte des modifications aux messages et flux de données ou de fichiers. Il est possible de détecter ce type d'attaque.

Les attaques actives peuvent prendre la forme d'un des quatre types suivants, seul ou en combinaison (Mascarade, Rejet, Modification de message ou Déni de service)

La sécurité dans les réseaux téléphoniques classiques est particulièrement forte. La disponibilité du réseau y atteint les 5 « neuf », c'est-à-dire que le système marche 99,999 % du temps. Avec la téléphonie sur IP, il faut introduire des éléments de sécurité supplémentaires puisque le support est partagé et qu'une écoute peut se produire.

Classiquement, la sécurité s'appuie sur cinq services de base : ***l'identification, l'authentification, la confidentialité, l'intégrité des données et la non-répudiation.***

Schématiquement, les infrastructures de sécurité des réseaux peuvent être classées en cinq catégories [Ouakil et pujolle, 2008] :

• ***Chiffrement*** au niveau physique. Dans la cryptographie optique (PMD), le saut de fréquences pseudo-aléatoire ou le chiffrement du flux d'octets (une méthode couramment déployée par les banques), les clés sont distribuées manuellement.

• ***Confidentialité, intégrité de données, signature de trames MAC.*** La distribution des clés est réalisée dans un plan particulier, décrit par la norme IEEE 802.1x. Dans ce cas, on introduit la notion de contrôle d'accès au réseau LAN. C'est une notion juridique importante, dont le rôle est d'interdire le transport des informations à des individus non authentifiés, et donc potentiellement dangereux.

• ***Confidentialité, intégrité des données, signature des paquets IP ou TCP.*** C'est typiquement la technologie IPsec en mode tunnel. Un paquet IP chiffré et signé est encapsulé dans un paquet IP non protégé. En effet, le routage à travers Internet implique l'analyse de l'en-tête IP par les passerelles traversées. IPsec crée un tunnel sécurisé entre le réseau d'accès et le domaine du fournisseur de services. On peut déployer une gestion manuelle des clés ou des protocoles de distribution automatisés, tels que ISAKMP (Internet Security Association and Key Management Protocol). La philosophie de ce protocole s'appuie sur la libre utilisation du réseau d'accès, qui ne va pas sans soulever des problèmes juridiques. Par exemple, si des utilisateurs mal intentionnés protègent leurs échanges téléphoniques, il est impossible aux réseaux traversés de détecter leur complicité dans le transport d'informations illégales.

• ***Insertion d'une couche de sécurité additive.*** Le protocole SSL (Secure Sockets Layer) fondé sur un chiffrement asymétrique assure la protection d'applications telles que la navigation Web ou la téléphonie IP. SSL réalise généralement une simple authentification entre serveur et client et négocie un secret partagé (Master Secret), à partir duquel sont dérivées les clés de chiffrement utilisées par l'algorithme de chiffrement négocié entre les deux parties. Une fois le tunnel sécurisé établi, le client s'authentifie à l'aide d'un login et d'un mot de passe. Il obtient alors une identité temporaire associée à un simple cookie.

### I.7.1. Les équipements de sécurité

***Pare-feu*** : Bon nombre de pare-feux se limitent à gérer l'ouverture de ports en fonction des communications et n'inspectent pas les flux (au niveau protocolaire). De plus cet élément additionnel risque d'introduire un délai ainsi qu'une gigue, c'est pourquoi ils sont absents dans bien des déploiements [Fischbach, 2004].

***IDS***(***Intrusion Détection System***): Il n'est pas très courant de trouver des outils de détection d'intrusion pour des solutions de voix sur IP. La quantité de faux positifs dus `a l'observation du flux RTP pourrait être plus que conséquente. Bien qu'elle permette de détecter des dénis de service par exemple, la détection d'intrusion se ramène souvent à de la détection de fraude [Fischbach, 2004].

### I.7.2. Les types d'attaques

***Déni de service:***Les attaques par déni de service se retrouvent sous plusieurs formes. Les plus classiques sont celles qui visent à utiliser toute la bande passante disponible ou abuser de problèmes intrinsèques à TCP/IP.

Dans le cadre d'une solution VoIP bien des éléments peuvent être attaquées : letéléphone, le réseau, le système d'exploitation, l'application, etc. Autant un dénide service sur l'Internet peut être filtré avec des mécanismes et des techniquesplus ou moins avancées, autant celui à l'encontre d'une communication seradifficile à traiter et aura un impact direct sur les possibilités de communications.

Par exemple un nombre trop important de messages SIP INVITE ou de simples messages ICMP peuvent créer une situation de déni de service [Baudoin et Karle 2004].

***Interception :*** L'interception d'une communication peut être l'œuvre d'un " criminel informatique " ou des autorités.Les techniques bien connues d'écoute de réseau (" sniffing " ou de l'homme du milieu (" man in the middle " s'appliquent à l'interception. En revanche, contrairement à une attaque MITM contre un protocole comme telnet ou http o`u il est possible de modifier le contenu à la volée, cela s'avère plus contraignant et plus facilement détectable avec de la voix encodée. Le protocole RTP transporte la voix encodée, sans aucun chiffrement, ce qui rend l'interception relativement triviale [Fischbach, 2004].

***Call-ID :*** Le service de présentation du numéro de l'appelant est devenu,pour bon nombre d'abonnés le facteur principal de prise ou de rejet d'appel, tout particulièrement depuis l'essor des téléphones portables.

Il est également relativement simple de manipuler l'identifiant de l'appelant.L'impact n'est pas très important, sauf dans le cas ou le CLID est utilisé pour authentifier l'appelant et autoriser l'accès à une ressource (boîte vocale, appels internationaux ou numéros spéciaux, etc...) [Fischbach, 2004].

***Non-répudiation et fraude :*** La fraude la plus connue est l'accès gratuit ou à un coût réduit à des services à valeur ajoutée ou des appels internationaux [Fischbach, 2004].

***La compromission de serveurs :*** Les serveurs jouent un rôle important dans une solution de voix sur IP, et même s'il n'est pas forcément possible d'intercepter un appel si un serveur est compromis, il est souvent possible de récupérer des CDRs (Call Detail Records) qui contiennent toutes les traces des appels effectués. En revanche la compromission d'une passerelle entre le réseau VoIP et le réseau téléphonique classique permet d'écouter de manière transparente les appels, même s'ils sont chiffrés du côté VoIP (SRTP) [Fischbach, 2004].

### I.7.3. Architecture de Sécurité

***Les systèmes et le réseau:*** En fonction des options de déploiement choisies, il est probable que la majorité des systèmes doivent être accessibles depuis " partout ". Il convient donc de sécuriser ces éléments comme tout serveur, et dans la mesure du possible de mettre en place une solution avec des pare-feux. La séparation entre le trafic voix et la signalisation peut se faire au niveau du réseau à l'aide de VLANs.

***Déni de service:*** Le mécanisme le plus important est la qualité de service (QoS) Sans déploiement de bout-en-bout, et tout particulièrement sur les réseaux locaux ou ce n'est pas très commun, il est possible pour n'importe qui de générer un déni de service. La bande passante disponible est également un facteur, mais la QoS et la gestion des files d'attentes est le point clé.

Les différents éléments qui composent une architecture de type SIP, comme le relais, sont livrés par certains vendeurs avec des mécanismes de détection de déni de service (comme l'envoi massif de messages INVITE).

Des solutions de filtrage de contenu ou d'analyse anti-virus commencent à être disponibles et utilisent SIP pour permettre l'inspection, un peu à l'image de protocoles comme IPSEC.

***Interception:*** La seule solution pour limiter l'interception est l'usage de mécanismes cryptographiques pour les flux de signalisation et de données (voix encodée). Ce chiffrement devrait être de bout en bout et il est important d'en évaluer les impacts. Le plus important étant celui sur la qualité de la communication à cause du délai supplémentaire introduit. Comme dans toute solution cryptographique, l'authentification forte des parties impliquées dans les échanges est un élément, sinon les attaques classiques de MITM peuvent s'appliquer.

La version " S " de SIP, SIPS, reposant sur TLS (Transport Layer Security) sur TCP permet de chiffrer les échanges SIP qui contiennent, par exemple, le nom d'utilisateur, le mot de passe ainsi que le numéro appelé.

Une nouvelle version du protocole RTP, S-RTP intègre des mécanismes de chiffrement.

***Non-répudiation et fraude:*** La fraude peut être limitée par la configuration correcte de la passerelle VoIP vers RTC (Réseau Téléphonique Commuté) ainsi qu'une gestion des droits/classes de service par utilisateur. Les passerelles doivent être configurées pour éviter qu'un utilisateur se connecte directement sans passer par le relais SIP.

## I.8. Mobile VoIP

VoIP mobile (Voice over Internet Protocol Mobile), wVoIP (Wireless VoIP) ou VoWLAN (Voice over WLAN) est une extension de la technologie VoIP.La VoIP mobile est plus riche que la VoIP au travers de réseaux Wi-Fi. N'importe quel réseau IP sans fil, comme les réseaux UMTS, HSDPA, voire le WiMAX, peut véhiculer la VoIP et les services qui l'accompagnent. Plusieurs technologies sont utilisées en matière de téléphonie par Internet sur mobile. Il faut d'abord distinguer trois types de connexion.

- les connexions GSM classiques
- les connexions 3G+
- les connexions WiFi

Pour les connexions wifi et 3G+ les possibilités diffèrent par ailleurs suivant les systèmes d'exploitation (problèmes de compatibilité des logiciels de téléphonie mobile par IP). O

### I.8.1. VoIP mobile sur connexion GSM

Ces systèmes fonctionnent généralement par un système de numéro d'accès.*Mobivox* qui fonctionne sur n'importe quel poste fixe ou mobile. *ManiVoip de manifone*utilise le compte SIP d'une *Freebox* ou d'un opérateur SIP tiers depuis son téléphone mobile.

### I.8.2. VoIP mobile sur connexion 3G

Concerne les propriétaires de téléphones disposant d'un accès Internet via un abonnement 3G ou assimilé (UMTS, GPRS, etc). Le callback consiste à donner son numéro et celui de son destinataire, l'opérateur VoIP prenant en charge la mise en correspondance des deux personnes (jajah,betamax,…). Certains éditeurs de logiciels de téléphonie proposent une version pour téléphone mobile (skype,gizmo,truphone). Notons par ailleurs l'existence de softphones compatibles SIP pour mobiles (VOYP,BeWip,Gosip,siphone…).

### I.8.3. VoIP mobile sur WiFi

La téléphonie par Internet via WiFi est parfois désignée sous le terme VoWiFi.

Trois types de téléphones doivent être distingués:

- les téléphones mobiles WiFiautonomes

- les téléphones mobiles bi-modes GSM+ / WiFi
- les téléphones mobiles bi-modes GSM+ / WiFi avec autoroaming

### I.8.4. VoIP et WiMax

WiMax-Mobile, ou Wi-Mobile, ou encore Universal WiMax, correspond à la norme IEEE 802.16e. Son objectif est de concurrencer les normes de réseau de mobiles 3G, comme l'UMTS ou le CDMA 2000[Ouakil et pujolle, 2008].

WiMax-Mobile s'adresse aux réseaux métropolitains sans fil mais adaptés aux connexions d'utilisateurs mobiles. Le standard WiMax-Mobile a été finalisé en décembre 2005, et les premiers équipements sont arrivés sur le marché dans le courant de l'année 2007. Le retard important par rapport aux normes déjà en place dans le monde des mobiles, comme l'UMTS, ne pourra être comblé que par un coût très inférieur de ces nouveaux équipements.

Le réseau WiMax-Mobile est doté de cinq classes de services pour la téléphonie

.

- UGS (Unsolicited Grant Service), dévolu à la téléphonie grâce à une forte garantie de la qualité de service.
- rtPS (Real-time Packet Service), qui correspond à des applications ayant de fortes contraintes temporelles, mais avec des débits qui peuvent être variables.
- nrtPS (Non-real-time Packet Service), qui correspond à des applications sans contraintes temporelles mais avec des contraintes de débit.
- BE (best-effort), qui ne possède aucune garantie.
- La cinquième classe s'appelle ertPS (enhanced real-time Packet Service), complète UGS (Unsolited Grand Service), qui ne propose qu'un flux constant, de telle sorte que si le flux téléphonique est compressé et donne un débit variable, une perte de bande passante importante peut avoir lieu.

## I.9. VoIP et ADSL

Avec l'invention de l'ADSL, il a suffi de menus aménagements du réseau téléphonique pour que la même paire de fils de cuivre puisse acheminer de très hauts débits informatiques. Aujourd'hui, la boucle est bouclée avec les services de VoIP par ADSL, qui utilisent toujours le même couple de fils pour acheminer de la voix, sous forme de paquets de données.

Cette technologie nouvelle permet simplement d'utiliser sa connexion Internet pour ses communications téléphoniques. Nous parlons bien sûr de téléphone et non pas de logiciels type MSN qui imposent à ceux qui l'utilisent maintes contraintes. Non, nous parlons d'un véritable combiné téléphonique qui vous permet de recevoir des appels ou d'appeler tout numéro de téléphone (national, mobile, international, numéros spéciaux, etc.), tout comme si vous utilisiez votre ligne classique.

Pour avoir accès à votre ligne VoIP, il faut brancher votre combiné téléphonique sur un terminal spécifique (c'est à dire le *box* fourni par votre fournisseur d'accès). La box qui vous est remise est en fait un modem ADSL amélioré possédant la fonction appropriée pour la téléphonie sur IP. On les connait aujourd'hui sous le nom de xxxBOX (exemple : la livebox d'Orange, la NeufBox, ou encore AssilaBox de EEPAD en Algérie).

## I.10. Les Softphones

Un softphone (anglicisme) est un type de logiciel utilisé pour faire de la téléphonie par Internet depuis un ordinateur plutôt qu'un téléphone. Les communications peuvent se faire au moyen d'un microphone et d'un casque ou de haut-parleurs reliés à la carte son, mais il existe aussi un type de périphérique dédié à cette tâche, semblable à un téléphone et se branchant sur un port USB. Les appels sont généralement gratuits de softphone à softphone et payant de softphone vers poste fixe ou mobile.

### I.10.1. Structure générale des applications VoIP

La transmission de VoIP dans un réseau à commutation de paquets commence par la digitalisation de la parole et forme des paquets de voix discrets. La figure suivante montre la structure générale d'une application VoIP etmet en évidence la séquence d'opérations à exécuter à l'extrémité de l'expéditeur aussi bien qu'à l'extrémité du récepteur [Mellouk, 2009].

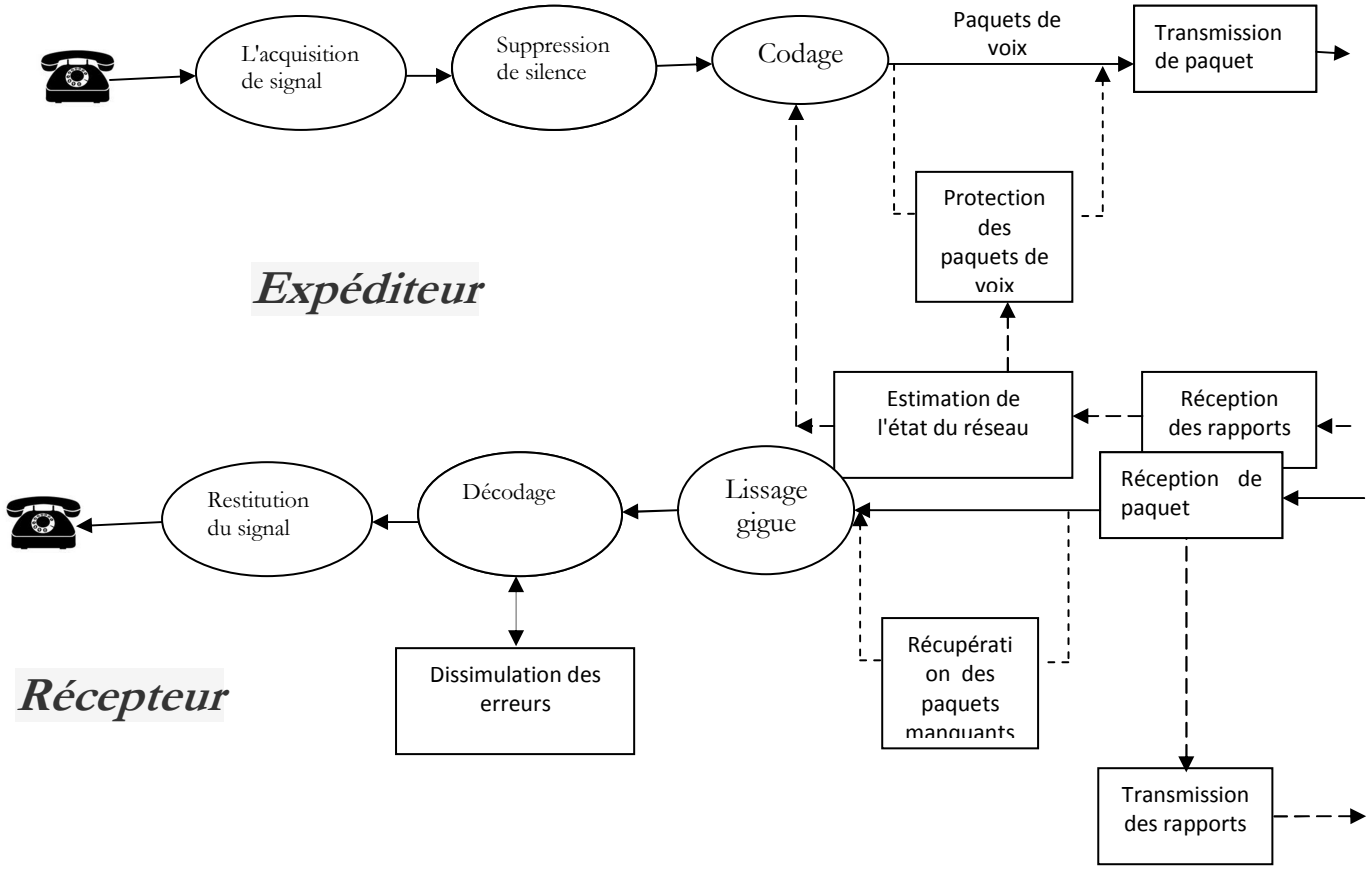



*Fig I-8 : structure générale d'un SoftPhone[Mellouk, 2009]*

Les opérations suivantes effectuées sur le côté d'expéditeur, serviront à préparer les paquets de voix pour être envoyées sur le réseau

- L'acquisition de signal est effectuée par un équipement spécifique, composé d'un microphone pour capter le signal de voix analogique, et un convertisseur analogique-numérique pour l'échantillonnage et la quantification du signal. l'échantillonnage est le processus d'encodage d'un signal continu sous la forme numérique discrète en lisant son niveau à intervalles de temps espacés précisément alors que la quantification est le processus qui s'occupe de convertir les échantillons obtenus de taille en nombre fini de valeurs discrètes. La qualité de la voix digitalisée dépendra de la fréquence d'échantillonnage et de la précision de la quantification (le nombre de bits employé pour représenter l'amplitude d'échantillons), ces deux paramètres agissant sur la déformation de signal. Un taux de 8000 échantillons par seconde avec une représentation de 8 bits est satisfaisant pour une application de VoIP
- La suppression de silence est une façon très efficace pour diminuer la largeur de bande passante de réseau consommée par une session voip. La parole humaine consiste des talkspurts (segment de parole qui se trouve entre deux silences) et lacunes de silence, également connues sous le nom modèles on-off. une méthode de détection d'activité vocale est utilisée pour détecter destalkspurts et les intervalles de silence. En conséquence, la transmission de paquets de voix est arrêtée pendant les périodes de silence.
- codage de la parole vise à produire de la qualité la plus élevée possible audiscours avec le plus bas possible débit. Traditionnellement on distingue 3 classes de techniques de codage de la voix, le codage forme d'onde, codage paramétrique et codage hybride
- des paquets de voix se composent d'un ou plusieurs blocs codés par parole. la transmission de paquet est accomplie utilisant le protocole RTP au-dessus du protocole IP. La RTP s'exécute au-dessus de l'UDP et fournit des services de livraison bout à bout pour des applications temps réel.
- la protection du paquet de voix vise à améliorer la robustesse de reconnaissance de la parole contre des pertes de paquet dans le réseau utilisantla méthode arrangement FEC (ForwardError Correction). ça consiste à insérer des informations redondantes au paquet de voix sur le côté expéditeur. ceci permet au récepteur de récupérer le contenu des paquets perdus.

Les opérations suivantes effectuées à l'extrémité du récepteur, serviront à reconstituer le signal de voix des paquets entrants:

- lors de la réception de paquets, l'en-tête UDP / RTP est vérifié pour détecter les pertes de paquets, les paquets doublons, les paquets corrompus, et le non séquencement des paquets, le récepteur peut calculer le délai inter-paquet entre le paquet juste reçu et le paquet qui le précède, pour estimer la gigue, le délai de

transmission peut également être estimé si les horloges de l'expéditeur et le récepteur sont synchronisés.

- Lissage gigue (jittersmoothing) est nécessaire parce que le récepteur ne recevra pas typiquement des paquets de voix à intervalles réguliers en raison de la gigue, même si l'expéditeur génère des paquets de voix à intervalles réguliers. lissage consiste d'organiser les paquets entrants dans une mémoire tampon de lecture assez longue pour permettre le plus lent des paquets pour arriver à temps pour être joué dans la séquence correcte.
- décodage de la voie effectue l'opération inverse décodage de la parole effectuée à côté de l'émetteur, il régénère la bit Stream représentant les échantillons de voix quantifiée. Décodage peut inclure certaines techniques de dissimulation d'erreur afin de réduire les effets de la perte des parquets.
- la restitution du signal convertit les échantillons numériques en signal de voix analogique qui est introduit dans le haut-parleur.
- La transmission des rapports fournit la rétroaction pour aider l'expéditeur à traiter la performance du réseau imprévisible. Rapports envoyés périodiquement utilisant le protocole de RTCP qui est les contre-parties de RTP. Elles incluent le nombre de perte des paquets de voix, le délai aller-retour et la gigue. Une telle rétroaction permet à l'expéditeur de s'adapter aux états de réseau et de maintenir un certain niveau de qualité de voix, basé sur le de débit binaire ou les mécanismes d'adaptation FEC.

Les applications temps réel de voix ont une limite supérieure sur le délai de bout en bout tolérable (de l'acquisition du signal à l'extrémité de l'expéditeur a la restitution du signal à l'extrémité du récepteur). Parmi les opérations décrites ci-dessus (par exemple acquisition de signal, codage et décodage de la parole, transmission et réception de paquet, etc.), présenter plus le délai.

### I.10.2. Skype

Skype est l'un des premiers logiciels grand public à avoir permis la jonction entre la téléphonie du monde Internet et celle du monde RTC. C'est sans doute là la clé de son succès. Grâce à une qualité d'écoute excellente, une facilité d'utilisation ne nécessitant généralement aucune configuration (y compris dans les infrastructures réseau déployant des pare-feu), une mobilité accrue, une gamme de services complémentaires et un prix incomparablement moins cher que la téléphonie traditionnelle, Skype s'est répandu de manière virale.

Skype a été lancé le 29 août 2003 à l'initiative de Niklas Zennström, un Suédois de 36 ans, et Janus Friis, un Danois de 26 ans, tous deux experts en technologies de peer-to-peer puisqu'ils avaient fait frémir l'industrie de l'entertainment au début des années 2000 avec le logiciel Kazaa, qu'ils avaient conçu.

Skype fonctionne selon un mode décentralisé et une architecture peer-to-peer (P2P), au sein du réseau skype il existe deux types de nœud : Les clients ordinaires et ce que l'on appellera les "super node". Les clients ordinaires sont, comme leur nom l'indique des applications Skype transmettant des flux de voix. Les "Super nodes" permettent eux de

transmettre les flux entre les différents utilisateurs. Tous les clients Skype qui possèdent les capacités requises (CPU, mémoire et bande passante suffisante) peuvent devenir des Super nodes. Ceux-ci agissent bien entendu également comme des clients ordinaires en permettant de transmettre leurs propres flux. Une autre entité importante du réseau skype est le serveur de login. Tous les clients doivent s'authentifier auprès de celui-ci pour s'assurer que les identifiants soient uniques sur le réseau. Ce serveur est la seule entité centralisée du protocole Skype. [Baset et Schulzrinne, 2006]

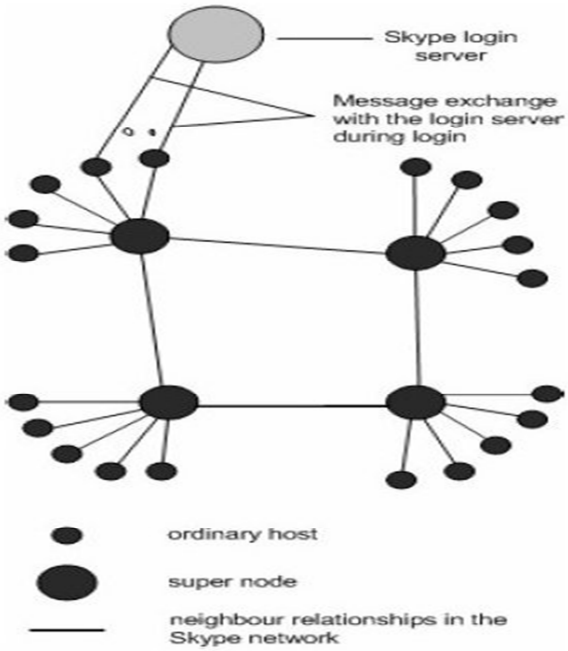



*Fig I-9: architecture peer-to-peer (P2P) de Skype [Baset et Schulzrinne, 2006]*

Une condition essentielle de la réussite de la ToIP est la possibilité de traverser les pare-feu. Les communications de ce type exploitent des ports dynamiques, qui ne sont généralement pas ouverts par ces pare-feu. Par ailleurs, le réseau sur lequel se trouve l'utilisateur peut mettre en œuvre un mécanisme de NAT (Network Address Translation), ou translation d'adresse réseau, qui donne à l'utilisateur une adresse IP non routable sur Internet. Pour ces deux raisons, la communication directe entre correspondants est impossible. Skype a trouvé la parade en exploitant différentes techniques. L'une d'elles consiste en l'utilisation de ports standards, qui sont étrangers à la téléphonie sur IP, mais qui présentent l'avantage d'être le plus souvent ouverts par les pare-feu. C'est le cas du port 80, associé généralement au Web pour le protocole HTTP.

Skype permet en outre d'utiliser des ressources situées à l'extérieur de la zone protégée par le pare-feu. Cette ressource peut être un utilisateur parmi d'autres, choisi pour accomplir cette tâche selon un algorithme propriétaire. Les flux IP de Skype suivent ainsi un chemin détourné lorsque le chemin direct est impossible. Ce sont de tels chemins qu'empruntent les communications entre utilisateurs de Skype, lesquels se prêtent à la fonctionnalité de routage sans en avoir conscience et pour les besoins d'autres clients.

### I.10.3. Windows Live Messenger

Windows Live Messenger (WLM, anciennement MSN Messenger) est un logiciel client propriétaire lié à un service propriétaire de messagerie instantanée (utilisable gratuitement) pour Windows XP, des consoles de jeux et des téléphones portables, et produit par Microsoft. Il offre les services de VoIP et de visioconférence depuis sa version 8.0.

Le système protocolaire utilisé dans WLM est propriétaire. Appelé MSNP (Mobile Status Notification Protocol), il fonctionne sur une couche de transport TCP avec le port 1863.

Publié sous forme de draft en 1999 dans sa version 2, MSNP a beaucoup évolué et est à sa version 14. Ce système est totalement fermé, et seuls des analyses des flux entrants et sortants ou des partenariats technologiques permettent d'envisager des interactions.

L'architecture générale de WLM est centralisée. Les serveurs Microsoft assurent à eux seuls l'enregistrement de tous les abonnés, avec leurs statuts et leurs pseudonymes. En changeant d'ordinateur, ces derniers conservent à la fois leur liste de contacts et le dernier pseudonyme qu'ils ont choisi.

Toutes les communications textuelles sont également centralisées et transitent par les serveurs Microsoft. Les utilisateurs ne sont donc pas mis en relation directement, sauf s'ils décident d'initier un transfert de fichier ou une communication audio ou vidéo, auquel cas les flux de données sont trop volumineux pour transiter par les serveurs Microsoft et ces derniers se contentent de faire connaître aux correspondants leur adresse IP respective.

### I.10.4. Yahoo! Messenger

Yahoo! Messenger est un système propriétaire de messagerie instantanée, de VoIP et de visioconférence créé par la société Yahoo!. Yahoo! Messenger nécessite un enregistrement préalable auprès de Yahoo pour l'ouverture d'un compte. Yahoo! Messenger, , est très proche de la messagerie de Microsoft.

Parmi les facettes remarquables du client Yahoo! Messenger, le service de téléphonie est exploitable pour appeler aussi bien des utilisateurs équipés du logiciel (téléphonie IP pure) que des téléphones standards.

### I.10.5. Asterisk

Asterisk est un autocommutateur téléphonique privé (PABX) open source pour systèmes UNIX, Mac OS et Windows. Il permet, entre autres, la messagerie vocale, les files d'attente, les agents d'appels, les musiques d'attente et les mises en garde d'appels, la distribution des appels. Il est possible également d'ajouter l'utilisation des conférences par le biais de l'installation de modules supplémentaires et la recompilation des binaires.

Asterisk implémente les protocoles H.320, H.323 et SIP, ainsi qu'un protocole spécifique nommé IAX (Inter-AsteriskeXchange). Ce protocole IAX permet la communication entre deux serveurs Asterisk ainsi qu'entre client et serveur Asterisk. Asterisk peut également jouer le rôle de registrar et passerelle avec les réseaux publics (RTC, GSM, etc.) Asterisk est extensible par des scripts ou des modules en langage Perl, C, Python, PHP, et Ruby. Asterisk est publié sous licence GPL et licence propriétaire

Asterisk comprend un nombre très élevé de fonctions permettant l'intégration complète pour répondre à la majorité des besoins en téléphonie. Il permet de remplacer totalement, par le biais de cartes FXO/FXS, un PABX propriétaire, et d'y adjoindre des fonctionnalités de VoIP pour le transformer en PBX IP. Il permet également de fonctionner totalement en VoIP, par le biais de téléphones SIP ou IAX du marché. Enfin, des fonctionnalités de routage d'appel, menu vocal et boites vocales—entre autres—le placent au niveau des PBX les plus complexes. Au sein des grandes installations d'Asterisk, il est courant de déployer les fonctionnalités sur plusieurs serveurs. Une unité centrale ou plus seront dédiées au traitement des appels et seront épaulées par des serveurs auxiliaires traitant les tâches secondaires (comme une base de données, les boîtes vocales, les conférences).

Des modules tiers permettent de visualiser ou paramétrer le PBX via une interface Flash ou via un client léger.

Enfin, notez qu'une distribution particulière d'Asterisk, *@home, est dédiée au PBX léger sur un réseau domestique.

### I.10.6. Jabber

Jabber est une plate-forme libre développée pour assurer l'interopérabilité des logiciels de messagerie instantanée et fédérer les réseaux de messagerie autour de normes communes. Les utilisateurs peuvent communiquer entre eux indépendamment de leur client logiciel et de leur serveur de traitement. Chacun d'eux reste libre de choisir son serveur, qui lui fournit une adresse afin de s'identifier, et d'utiliser le client qu'il souhaite parmi la gamme des logiciels compatibles disponibles. La clé de voûte de cette plate-forme est d'offrir un protocole libre et standardisé

Jabber repose sur le protocole XMPP (eXtensible Messaging and Presence Protocol), qui a été conçu au sein de la communauté libre Jabber avant d'être soumis à l'IETF pour standardisation et évolution.

Le standard XMPP est un protocole extensible de messagerie et de présence conçu par la communauté Jabber pour formaliser l'échange de flux de messages à contraintes temps réel.

Jabber repose sur un modèle de type client-serveur distinguant les trois types d'entités logiques suivants (Clients Jabber, Serveurs Jabber et les Passerelles)

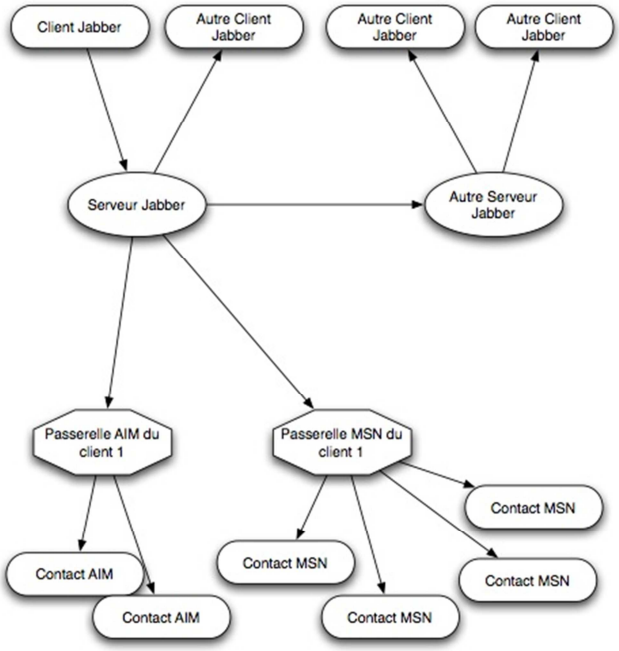



*Fig I-10: modèle de type client-serveur de jabber[Ouakil et pujolle, 2008]*

### I.10.7. Google talk

Google Talk est un logiciel propriétaire et service de messagerie instantanée et de voix sur IP basé sur Jabber développé par la société Google sorti en version bêta le 24 août2005. Google Talk représente à la fois le service ainsi que le logiciel propriétaire client permettant de se connecter à ce service.

Google Talk utilise le protocole standard et ouvert Jabber/XMPP tout en encourageant à utiliser son propre client afin de se connecter au service Google Talk.

De par l'utilisation d'un protocole ouvert, on peut aussi vous connecter au service de messagerie instantanée Google Talk à partir d'un client Jabber standard. La voix sur IP n'est toutefois pas encore complètement développée. La communication inter-serveur Jabber (« s2s », pour « server to server ») a été implémentée dans le service de Google Talk le 17 janvier2006, Google Talk est donc ouvert à l'ensemble du réseau Jabber public mondial (réseau dit « fédéré »).

## I.11. Conclusion

La téléphonie sur IP va inéluctablement remplacer la téléphonie numérique classique. Les enjeux sont considérables puisque l'ensemble des entreprises aura adopté cette technologie dans les années à venir. Ainsi que la Voip ouvre la porte à des technologies d'avenir liées à la convergence voix et données et aux services qui y seront associés. Mais avant de décider de passer à la ToIP il est important de réfléchir aux cinq problèmes clés suivants :

*La disponibilité, la sécurité, le contrôle, la gestion et la qualité de service*

# Chapitre II

- *VoIP et Qualité de Service*

# II. VoIP et Qualité de Service

## II.1. Introduction

Le terme QoS (acronyme de « Quality of Service », en français « Qualité de Service ») désigne la capacité à fournir un service (notamment un support de communication) conforme à des exigences en matière de temps de réponse et de bande passante.

Appliquée aux réseaux à commutation de paquets (réseaux basés sur l'utilisation de routeurs) la QoS désigne l'aptitude à pouvoir garantir un niveau acceptable de perte de paquets, défini contractuellement, pour un usage donné (voix sur IP, vidéo-conférence, etc.).

*Les principaux critères permettant d'apprécier la qualité de service sont les suivants :*

***Débit*** (en anglais ***bandwidth***), parfois appelé bande passante par abus de langage, il définit le volume maximal d'information (bits) par unité de temps.

***Gigue*** (en anglais ***jitter***) : elle représente la fluctuation du signal numérique, dans le temps ou en phase.

***Latence***, ***délai*** ou ***temps de réponse*** (en anglais ***delay***) : elle caractérise le retard entre l'émission et la réception d'un paquet.

***Perte de paquet*** (en anglais ***packet loss***): elle correspond à la non-délivrance d'un paquet de données, la plupart du temps dûe à un encombrement du réseau.

***Déséquencement*** (en anglais ***desequencing***) : il s'agit d'une modification de l'ordre d'arrivée des paquets.

Le terme « *niveau de service* » (en anglais *Service level*) définit le niveau d'exigence pour la capacité d'un réseau à fournir un service point à point ou de bout en bout avec un trafic donné. On définit généralement trois niveaux de QoS :

***Meilleur effort*** (en anglais ***best effort***), ne fournissant aucune différenciation entre plusieurs flux réseaux et ne permettant aucune garantie. Ce niveau de service est ainsi parfois appelé lack of QoS.

***Service différencié*** (en anglais ***differenciated service*** ou ***soft QoS***), permettant de définir des niveaux de priorité aux différents flux réseau sans toutefois fournir une garantie stricte.

***Service garanti*** (en anglais ***guaranteed service*** ou ***hard QoS***), consistant à réserver des ressources réseau pour certains types de flux. Le principal mécanisme utilisé pour obtenir un tel niveau de service est RSVP (Resource reSerVation Protocol, traduisez Protocole de réservation de ressources).

En effet, contrairement aux réseaux à commutation de circuits, tels que les réseaux téléphonique commuté, où un circuit de communication est dédié pendant toute la durée de la communication, il est impossible sur internet de prédire le chemin emprunté par les différents paquets.

Au-delà de la qualité de la voix qui doit rester respectable après avoir été numérisée et compressée et remodelée en sens inverse, il importe de garantir un niveau de qualité sur toute la transmission des paquets de voix sur le réseau IP. Pour cela Plusieurs solutions peuvent être mises en œuvre selon deux grandes directions : contrôle de la qualité de service niveau applicatif (AQoS) et contrôle de la qualité de service niveau réseau(NQoS). Dans le premier cas, l'application s'adapte au réseau. Si le réseau est chargé, l'adaptation s'effectue en diminuant le débit du flux. Dans le second cas, c'est le réseau qui s'adapte à l'application. Il faut en ce cas demander au réseau un SLA (Service Level Agreement), qui, compte tenu du débit et des contraintes temporelles, assure la qualité de service demandée.

## II.2. Le Contrôle QoS niveau applications

Le contrôle QoS niveau applicatif se rapporte aux équipements inclus dans une application qui préservent la qualité de son utilisation prévue, AQoS se rapporte à deux domaines principaux [Oneil, 2002]:

- La signalisation
- La manipulation terminale des flux de médias

Et Ce qui suit est une liste courte des services obtenus par l'intermédiaire de la signalisation et des médias manipulation :
• Découverte/enregistrement de portier
• Enregistrement de point final
• Admission d'appel
• Échange de possibilités
• Négociation de largeur de bande
• Perte, frousse et latence relatives de réseau d'AQoS
• Dégagement d'appel.

## II.3. Le contrôle QoS niveau réseau

En effet, la voix est très sensible aux retards, et le principe de best effort propre à IP ne garantit pas non plus l'ordonnancement des paquets à la réception de la même façon qu'à l'émission. De plus, un paquet perdu serait une véritable catastrophe pour la VoIP, car la retransmission du paquet que pourrait prendre en charge TCP ne ferait qu'augmenter le retard.

Les paquets doivent donc faire face au cours de la transmission sur le réseau à différents facteurs comme **l'écho**, la **latence** (latency) ou encore la **gigue** (jitter), qui peuvent provoquer de graves dommages à la qualité de la voix qui sera reçue. Ce sont donc autant de facteurs qui doivent être combattus.

En revanche, c'est le **délai** dans le transport de la voix, au quel l'on fait référence par les termes de latence et de gigue (la gigue étant la variation du délai, rendant le signal audio reçu inintelligible pour l'oreille humaine), qui nécessitent l'implémentation de nouveaux mécanismes d'allocation de priorités sur le trajet emprunté par les paquets. Et c'est cette notion de **priorité** qui est recherchée dans les principaux protocoles de qualité de service tels que **RSVP**, et les mécanismes de qualité de service comme **IntServ**, **MPLS** ou **DiffServ**, de sorte à aiguiller chaque type de trafic différent sur le chemin qui l'avantage le mieux [Mellouk, 2009].

Le but n'est pas ici de détailler chacun de ces protocoles, mais de savoir qu'ils existent et de pouvoir les différencier en en connaissant les principaux atouts et le mode de fonctionnement global. On distingue principalement trois différentes philosophies de QoS sur IP

## II.4. Analyse de délai de Bout-en-Bout

Il ya de nombreuses sources de délai dans les applications VoIP[Mellouk, 2009]. Certains de ces facteurs peuvent être gérés et certains ne peuvent pas. Les dépréciations causées par les retards comprennent d'écho et les chevauchements du parleur. L'écho est provoqué par l'isolement acoustique entre le haut-parleur et le microphone du dispositif d'un utilisateur. À mesure que le délai augmente, l'écho devient plus apparent et ennuyeux... Echo devient un problème important lorsque le délai d'aller-retour devient supérieur à 50 millisecondes. Ce problème est abordé utilisant les téléphones qui comportent l'annulation d'écho. Le chevauchement du parleur se produit quand une partie coupe la parole de l'autre partie en raison de long délai d'attente (c.-à-d. quand les participants parlent simultanément). Ce problème devient bien plus significatif si le délai d'un sens unique devient plus grand que 250millisecondes. Par conséquent, il est important de tenir compte de toutes les sources de délai correctement pour vérifier que la transmission vocale de haute qualité de bout en bout est possible. Sources de délai comprennent délai de codage et le délai de découpage en paquets de l'expéditeur, le délai de réseau, le délai de gigue, délai de réassemblage des paquets et le délai de décompression côté récepteur.

### II.4.1. Délai de codage et décodage

Les codeurs de parole ont un certain délai fixe [ITU, 1999]. Ce délai varie en fonction du codeur utilisé. Plus le taux de compression est élevé, plus la complexité du processus de compression est élevée, et plus le délai de codage est élevé. Le délai de codage est lui-même divisé en trois délais : délai de *framing size*, délai de *look-ahead* et délai de *data processing* .[Mellouk, 2009]

- **Le délai de Framing size $D_{FS}$** est l'intervalle de temps de discours du codeur utilise dans le processus d'encodage. En effet, le codeur a besoin de recueillir une certaine quantité d'échantillons avant de commencer l'algorithme de codage
- **Délai de Look-ahead $D_{LH}$** , se produit quand le codeur doit rassembler quelques blocs simples pour fournir des orientations en codant le bloc actuel. L'idée de la

lecture anticipée est de tirer profit de la corrélation étroite existant entre les blocs successifs d'échantillons.

- **Délai Data processing $D_{DP}$** est le temps pris pour exécuter l'algorithme de codage sur un bloc d'échantillons. Parce que différents codeurs travaillent de différentes façons, ce délai varie en fonction du codeur utilisé et la vitesse du processeur.

| Technique de codage | débit | $D_{FS}$ | $D_{LA}$ | $D_{DP}$ |
|---|---|---|---|---|
| G.711 | 64 kbps | 0.125 ms | 0 ms | 0.75-1.0 ms |
| G.726 | 16-40 kbps | 0.125 ms | 0 ms | 1-2 ms |
| FS 1015 | 2.5 kbps | 22.5 ms | 90 ms | 25-30 ms |
| FS 1016 | 4.8 kbps | 30 ms | 7.5 ms | 12-15 ms |
| G.728 | 16 kbps | 0.625 ms | 0 ms | 3-5 ms |
| G.729 | 8 kbps | 10 ms | 5 ms | 2.5-10 ms |

*Tableau II-1 : Les délais pour les codeurs standards de la voix. [Guillet, 2010]*

À l'extrémité du récepteur, le temps de décodage, $C_R$, est approximativement 10% des $D_{DP}$ pour chaque bloc simple, vu les codecs standard qui sont connus pour leur bonne qualité [Vleeschauwer et al., 2000]. Cependant, le temps de décodage est proportionnel au nombre des blocs par paquet parce que un paquet peut contenir plusieurs blocs.

### II.4.2. Délai de paquetage

Le délai de Paquetage est le temps nécessaire pour remplir un paquet avec des données de voix codées. Ce délai ($P_E$) est une fonction de la longueur d'un bloc simple exigée par le codeur de voix et le nombre des blocs codés (n) placé dans un simple paquet. Plus la longueur de paquet est grande, plus le temps est exigé. L'utilisation des paquets plus courtes peut diminue ce délai, mais ceci diminuera l'efficacité de réseau parce que plus de paquets doivent être envoyés, chacune avec l'information d'en-tête presque redondante.

*RFC 1890* spécifie que la période de Paquetage de défaut ne dépasse pas 20 ms. Pour G.711, ceci signifie que 160 blocs simples seront accumulés et puis transmis dans un seul paquet de voix. De la même manière, la documentation de Cisco nous conseille d'essayer d'obtenir un délai de Paquetage qui ne doit pas dépassé 30 ms [Cisco, 2000]. Bien que chaque simple bloc éprouve le délai de codage et le délai de Paquetage.

Sur la réception d'un paquet entrant, le délai de De-Paquetage est le temps nécessaire pour vérifier et enlever l'en-tête de RTP, puis tamponner le simple bloc. Un tel délai est négligeable.

### II.4.3. Le délai de réseau

Le délai de réseau, $D_N$, est le temps pris par un paquet de voix pour traverser un chemin de réseau multi-saut particulier pour atteindre sa destination. Pour chaque saut, les divers composants de réseau incluent le délai d'enchainement, le délai de propagation, le délai de file d'attente, et le délai de changement de chemins ou de routage. Le délai d'enchainement est en fonction de la taille du paquet et de la capacité du média physique fondamental. Le délai de propagation dépend de la distance entre les dispositifs du réseau et la vitesse du signal dans le moyen employé pour les relier ensemble. Le délai de routage dépend de la vitesse du traitement inhérent des dispositifs du réseau que les paquets traversent (par exemple des routeurs). Le délai de la queue est directement lié aux longueurs de la file d'attente dans les dispositifs du réseau et la charge du trafic courant. C'est le composant de délai le plus variable. Une autre cause de variation de délai de réseau est que les paquets sont conduits indépendamment, c.-à-d. les paquets envoyés de la même source à la même destination pourraient prendre différents chemins de réseau. En conséquence, ils pourraient éprouver différents troncs de réseau, différents dispositifs de réseau et situation différente du trafic.[ Trad, 2002]

### II.4.4. Compensation de délai de Gigue

Gigue est la variation dans le temps d'arrivée des paquets présentés par les délais variables dans le réseau. Éviter la gigue exige que les paquets doivent être tamponnés et tenus assez longtemps pour luis permettre les plus lents d'arriver à temps et d'être joué dans l'ordre correct (chaque période de Paquetage, par exemple chaque 20ms). Le délai de compensation de gigue, $D_{JC}$, est le temps d'entreposage du paquet de voix dans le buffer de de-gigue si on pourrait dire.

### II.4.5. Calcul de délai de bout-en-bout

Considérant un codec de la voix dont les caractéristiques ($D_{FR}$, $D_{LA}$, $D_{DP}$) sont connues, le délai de bout-en-bout pour le $i^{eme}$ paquet de voix, dénoté par D (i), peut être calcul utilisant l'équation suivante.

$$D(i)=\underbrace{D_{FS}+D_{LA}+D_{DP}+n*P_E}_{\text{délai de codage}} + \underbrace{D_N(i)+d_{jc}(i)}_{\text{délai de réseau+temporisation}} + \underbrace{n*0.1*D_{DP}}_{\text{délai de décodage}}$$

Ou n est le nombre de blocs simples codés placés dans le paquet. Le délai de codage et de décodage sont fixe. D (i), devrait être constant pour chaque paquet de telle sorte que les blocs simples doit-être diffusés d'une manière synchronisée, le délai de playout (réseau + temporisation) peut être fixé en ajustant le délai de compensation de gigue sur le délai de réseau.[ Trad, 2002]

## II.5. Conditions de qualité du service pour VoIP

Puisque les réseaux IP fournissent par défaut un service de meilleur-effort et ne fournissent pas des garanties pour la livraison de paquet, la conversation de VoIP de bonne qualité dépendent de maintenir des contraintes strictes pour le délai, la perte de paquet, et la gigue. Le long délai affecte l'interactivité normale de conversation, et cause l'hésitation et le talkover. La perte de paquet peut causer des lacunes ou objets façonnés qui ont comme conséquence une parole déformée ou même incompréhensible. Des niveaux bas de gigue sont absorbés par temporisateur de De-gigue. Cependant, les niveaux élevés de la gigue peuvent causer un paquet qui va réussi à arriver à sa destination, mais qui doit être rejeté puisque il a été trop retardé [Mellouk, 2009].

### II.5.1. Contrainte de délai

Le délai de bout-à-bout peut aller jusqu'à 150ms et il est généralement considérés acceptable pour des applications de voix de haute qualité. La recommandation G.114 [ITU, 1996] de l'IUT-T fournit les lignes directrices concernant la limite de délai pour des applications de voix dans l'écho est adéquatement contrôlé (voir le tableau II.2). Le délai de bout-en-bout peut être maintenu à une valeur constante parce que l'expéditeur produit des paquets de voix à intervalles réguliers et ceux-ci doivent être joués dans le coté du récepteur à mêmes intervalles.

| N° de classe | Délai d'une voie | Commentaires |
|---|---|---|
| 1 | 0-150 ms | Acceptable pour la plupart des applications utilisateurs |
| 2 | 150-400 ms | Acceptable, à condition que les administrations sont conscients de l'impact du temps de transmission sur la qualité de transmission des demandes des utilisateurs |
| 3 | >400 ms | Inacceptable pour les besoins généraux de la planification des réseaux, cependant, il est reconnu que dans certains cas exceptionnels cette limite sera dépassée |

*Tableau II-2 : Les spécifications des délais de bout-en-bout pour le codeur G.114 [Mellouk, 2009]*

### II.5.2. Contrainte de perte de paquets

Les paquets de voix ont à traverser un réseau IP, qui est un processus non fiable. Les paquets peuvent être abandonnés en raison de la congestion de réseau ou de corruption des données. Par ailleurs, pour les trafics en temps réel tels que la voix, la retransmission des paquets perdus à la couche transport n'est pas pratique en raison du délai d'addition. En outre, les paquets qui arrivent en retard à cause de délai sont rejetés au niveau du récepteur et sont donc perdus.

VoIP est sensible à la perte de paquet. La dégradation de la qualité perçue de la parole provoquée par la perte de paquet dépend de la distribution de perte de paquet, de la taille

et de l'activité du signal du son elle-même. Elle dépend également du codec de la voix utilisé. Le codec PCM G.711. Par exemple, n'a aucune dépendance d'échantillon-à-échantillon ou de paquet-à-paquet. Une perte de paquet de 20 ms cause exactement la dégradation de la parole de 20 ms sur le signal audio. Cependant, pour le codeur de la voix G.729, par exemple, la dégradation sur cet intervalle de voix, elle affectera également les paquets adjacents de voix de sorte que la dégradation réelle de signal audio ait pu être plus de 20 ms. En bref, une perte de paquet de 5% dans le codec G.711 n'est pas égale à la perte de paquet de 5% dans le codec G.729 en termes de dégradation perçue de vitesse. Quand le codec G.711 est adopté, la qualité de voix est inacceptable si le pourcentage des pertes est plus de 10%. La perte de paquet ne doit pas dépasser 5% pour des applications de voix de haute qualité. Ainsi plus de compression de codec peuvent tolérer même moins de perte de paquet. Cisco indique que le codec du défaut G.729 exige de la perte de paquet d'être plus moins de 1% pour éviter des affaiblissements audibles [Mellouk, 2009].

Naturellement, la dégradation perçue de qualité de voix provoquée par la perte de paquet dépend également de la façon dont le paquet perdu est caché ou manipulé. Un paquet perdu, par exemple, peut être remplacé par intervalle silencieux, ou par le bloc précédent, ou être remplacé par l'interpolation entre les blocs adjacents, ou même ignoré, causant ainsi un glissement de voix. Dans chaque cas, la qualité de voix perçue sera différente pour la même quantité de perte de paquet avec le même codec.

### II.5.3. Contrainte gigue

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.

Comme indiqué plus tôt, la stratégie générale en faisant face à la gigue est de tenir les paquets entrants dans un buffer de De-gigue de sorte qu'ils puissent être joués dehors d'une façon temps-régulier. Ce buffer doit être assez large pour permettre aux paquets les plus lents d'arriver à temps. Plus la gigue est grande, le plus long certains des paquets sera tenu dans le buffer, qui présente le retard additionnel. L'algorithme de mise en mémoire tampon de playout peut être fixe ou adaptatif. Si il est fixe, le délai de p²layout, c.-à-d. DN (i)+djc (i) dans la formule de calcul de délai indiqué précédemment, sera être les mêmes pour tous les paquets de voix, par conséquent le délai de bout à bout D (i) sera maintenu à une valeur constante pour tous les paquets [Cho et Un, 1994]. Un tel algorithme peut fonctionner efficacement dans les réseaux avec la basse gigue mais pas pour l'Internet parce que les paquets peuvent éprouver, parfois, la grande variation du délai. [Sanghi et al., 1993].

## II.6. La Qualité Perçue et Qualité d'Expérience

### II.6.1. Introduction

La qualité d'expérience ou la qualité perçue est l'idée qu'on se fait par rapport aux sensations et à l'opinion qu'on en a. Dans toute relation d'échange les produits se doivent s'offrir à la fois des services réels mais aussi de les exprimer dès le premier regard. Leurs dimensions matérielles et immatérielles participent donc à la perception de leurs qualités. Il s'agira des qualités de base toujours nécessaires, des services offerts et des signes de qualité qui sont perçus par un client. La qualité perçue est une notion qui n'est pas uniquement une propriété du produit/service mais qui dépend aussi du celui qui va l'apprécier.

D'un point de vue historique, le sigle QoS est apparu en premier dans les années 90 pour désigner un ensemble de techniques réseaux garantissant l'acheminement des trafics sensibles comme la voix ou les applications "transactionnelles". Depuis, le sigle QoS a été mis à toutes les "sauces" pour mettre en avant l'amélioration des performances apportées par des produits matériels et/ou logiciels. Le sigle QoE est quant à lui apparu beaucoup plus récemment. Il surfe sur une vague de popularité croissante car il touche directement aux utilisateurs. Présentée comme une approche globale de la qualité (mesure de bout en bout), la QoE consiste à mesurer la performance d'utilisation d'un service directement auprès de l'utilisateur. La QoE s'appuie sur une mesure subjective du service fourni, tandis que la QoS est basée sur la mesure objective des composants qui constituent le service.

Le concept de QoE en ingénierie est également connu comme la qualité de service perçue (PQoS), dans le sens ou la qualité de service comme elle est finalement perçue par l'utilisateur final. Prenons l'exemple d'un service de téléphonie. L'approche QoE définit la qualité d'écoute d'une communication voix selon l'indice MOS (Mean Opinion Score) attribué par un échantillon d'utilisateurs et qui varie de 0 (très mauvais) à 5 (excellent). Tandis que, l'approche QoS consiste à mesurer les paramètres techniques du réseau (délai réseau, gigue, perte de paquets) et des Codecs qui influent sur la qualité de la voix.

L'évaluation de la PQoS c'est des procédures objectives et subjectives chaque fois ayant lieu après le processus d'encodage (poste-encodage d'évaluation)

### II.6.2. Les types d'applications

Pour déterminer l'influence exacte que les conditions dans le réseau ont sur une application, on doit effectuer des expériences pour cette application particulière. Donc au lieu de considérer chaque application séparément, on doit diviser les applications réseau en classes. Cette division est utile avant tout dans l'étape de définition des métriques.

On peut distinguer les classes d'application suivantes [Beuran, 2004] :

### II.6.2.1. Applications unidirectionnelles

Cette classe concerne les applications qui sont de manière générale unidirectionnelles. Telles que les applications de type transfert de données par le protocole TCP (FTP pour le transfert des fichiers ou HTTP pour l'accès WEB).

Pour FTP le temps d'achèvement du transfert puisse être l'un des paramètres utilisés pour juger la qualité perçue. L'effet des paramètres liés au délai particulièrement n'est pas important pour le temps d'achèvement du transfert, par contre l'influence de la perte de paquets est assez significative sur le temps d'achèvement ainsi sur l'efficience d'utilisation de la bande passante du réseau. Pour une telle application le débit utile moyen est une métrique importante de la qualité perçue par l'utilisateur

Pour HTTP les requêtes ont d'habitude une taille plus petite que les réponses, le délai semble être plus important du point de vue de l'utilisateur étant donnée l'interactivité des applications Web. Même en l'absence de contraintes de temps réel, le temps de réponse d'un serveur Web peut être considéré comme un paramètre important dans ce cas pour la qualité perçue.

### II.6.2.2. Applications unidirectionnelles avec contraintes temporelles

L'exemple typique dans cette classe est le streming en continu. Une telle application plus généralement multimédia (MPEG-2, MPEG-4), dépend essentiellement de la perte de paquets. Le délai et sa variation sont importants dans la mesure où ils doivent être compensés par l'utilisation d'une mémoire tampon. La diffusion de vidéo en continu a des exigences particulièrement élevées sur la bande passante, comme précisé dans le tableau

| ***Application*** | ***Nécessaire de bande passante (Mbps)*** |
|---|---|
| Vidéo dans une fenêtre | <1.5 |
| Qualité VHS (plein écran) | 1-2 |
| Diffusion NTSC | 2-3 |
| Diffusion PAL | 4-6 |
| PAL professionnel | 8-10 |
| Diffusion HDTV | 12-20 |
| HDTV professionnel | 32-40 |
| NTSC brut | 168 |
| PAL brut | 216 |
| HDTV brut | 1000-1500 |

*Tableau II-3 : La bande passante requise pour la diffusion vidéo en continu [Beuran, 2004]*

### II.6.2.3. Applications bidirectionnelles

Les applications bidirectionnelles, généralement ne sont pas des applications temps réel, mais imposent des demandes d'interactivité chaque fois qu'un utilisateur est directement impliqué. Donc sont celles dans les deux directions de communication sont également significatives. Telles que Les systèmes de fichiers réseau (NFS), Les systèmes des noms des domaines (DNS), l'accès à distance (Telnet)

### II.6.2.4. Applications bidirectionnelles avec contraintes temporelles

Les applications bidirectionnelles avec contraintes temporelles sont les plus exigeants en termes de contraintes imposées aux réseaux. Généralement font usage des flux RTP/UDP on peut distinguer deux catégories en fonction de la bande passante demandé

- Application avec un usage réduit de bande passante (e.g. la téléphonie IP)
- Applications avec un usage plus élevé de bande passante (e.g. la visioconférence)

La téléphonie IP (nommée aussi Voice over IP, VoIP) est l'une des applications qui commence à être de plus en plus utilisée, stimulant les efforts de fourniture de QoS. Son interactivité est telle que les utilisateurs percevront d'infimes variations des paramètres QoS. On dit que la perte est plus importante que la gigue qui est à son tour plus importante que le délai.

La visioconférence (Video Teleconferencing, VTC) est une application de nature similaire qui envoie aussi des données vidéo. Plusieurs standards existent dans ce cas aussi : H.323, H.322, H.261, H.263 etc. Les demandes de bande passante sont plus grandes que pour VoIP.

Ainsi on peut classifier les applications en fonction de la manière d'utilisation des ressources du réseau et leur adaptabilité

### II.6.2.5. Applications élastiques

Les applications élastiques sont celles dont la base est constituée de protocoles adaptatifs du type TCP. Ce protocole essaye d'occuper une partie aussi large qu'il peut gérer de la bande passante disponible. Le taux de transmission est adapté aux conditions réseau courantes en utilisant un mécanisme de contrôle et évitement de la congestion. Il s'agit donc des applications dont le trafic est régulé.

### II.6.2.6. Applications inélastiques

Les applications inélastiques sont, par exemple, celles qui sont fondées sur des protocoles de diffusion en continu de type UDP. Ce type de protocole utilise une bande passante fixe et n'a pas de mécanismes de correction d'erreurs intrinsèques, donc le trafic

est non-régulé. Il ne peut pas s'adapter aux conditions courantes dans le réseau, donc ses exigences QoS sont strictes.

## II.6.3. Evaluation de La Qualité Perçue pour VoIP

Ce n'est pas facile d'évaluer la qualité perçue de la voix. La meilleure manière de l'évaluer est d'avoir de vraies personnes qui font l'évaluation parce que la qualité est un concept très subjectif. UIT-T a défini plusieurs standards qui permettent une évaluation de la qualité de la communication vocale.

Les réseaux de télécommunication modernes fournissent un large ensemble de services vocaux utilisant beaucoup de systèmes de transmission. Le déploiement rapide des technologies numériques en particulier a conduit à un besoin accru d'évaluation des caractéristiques de transmission des nouveaux équipements de communication. UIT-T a défini plusieurs standards qui permettent une évaluation de la qualité de la communication vocale.

### II.6.3.1. Méthodes subjectives

#### II.6.3.1.1. Score moyen d'opinion (Mean Opinion Score, MOS)

En 1996, UIT-T a défini la méthodologie de déterminer le degré de satisfaction concernant une certaine ligne téléphonique, sous le nom de score moyen d'opinion (Mean Opinion Score, MOS) (la norme P-800). Cette méthode peut s'appliquer à toute forme possible de dégradation : perte de paquets, distorsions de circuit, erreurs de transmission, distorsions environnementales, écho, distorsions de codage etc. La procédure d'évaluation est basée sur des tests subjectifs dans lesquels la qualité est notée par un « panel » d'expérimentateurs. Les valeurs suivantes sont assignées en fonction de la qualité perçue de la connexion [ITU, 1996]:

| | |
|---|---|
| ***Excellent*** | **5** |
| ***Bien*** | 4 |
| ***Acceptable*** | 3 |
| ***Pauvre*** | 2 |
| ***Mauvais*** | 1 |

*Tableau II-4 : Les valeurs assignées pour la qualité perçue du MOS [ITU, 2005]*

***Remarque*** : (CMOS et DMOS) Le CMOS est le résultat de l'analyse par catégories de comparaisons CCR (Comparison Category Rating) dans laquelle on fournit à un groupe d'auditeurs des signaux par paires. L'auditeur compare les deux signaux de chaque paire en termes de qualité en précisant lequel est le meilleur et évalue la différence selon une échelle de notation bien définie. Quant au DMOS, il résulte de l'analyse par catégories de dégradations DCR (Degradation Category Rating) dans laquelle on fournit à un groupe d'auditeurs des paires de signaux pour comparer cette fois-ci la qualité en terme de

dégradation. Contrairement au CMOS, les auditeurs savent a priori que la qualité du second signal est moins bonne que celle du premier. Ils doivent donc indiquer à quel point le second est justement moins bon suivant l'échelle de DMOS [Keagy, 2000]

| Score CMOS | Qualité du second comparé au premier |
|---|---|
| 3 | Bien meilleure |
| 2 | Meilleure |
| 1 | Légèrement meilleure |
| 0 | a peu près équivalente |
| -1 | Un peu moins bonne |
| -2 | Moins bonne |
| -3 | Nettement médiocre |

*Tableau II-5 : Echelle CMOS [Keagy, 2000]*

| Score DMOS | Qualité DMOS |
|---|---|
| 5 | Dépourvu de dégradation |
| 4 | Dégradation audible mais pas gênante |
| 3 | Dégradation un peu gênante |
| 2 | Dégradation gênante |
| 1 | Dégradation très gênante |

*Tableau II-6 : Echelle DMOS [Keagy, 2000]*

### II.6.3.2. Méthodes objectives

Les mesures objectives de qualité des signaux vocaux les plus communément utilisées sont citées et classées dans le tableau suivant.

| Mesures temporelles | Mesures Fréquentielles | Mesures perceptuelles |
|---|---|---|
| SNR | IS | BSD |
| segSNR | CD | MBSD |
| | WSS | PSQM |
| | LLR | PESQ |

*Tableau II-7 : Classification de critères d'évaluation objective [Mellouk, 2009]*

Les critères temporels et fréquentiels se basent essentiellement sur l'évaluation de la qualité en termes de comparaison de distorsion de formes entre signal de référence et signal débrutie, sans tenir compte de l'aspect perceptif. Certes, c'est une condition nécessaire mais non suffisante dans la mesure où deux signaux pratiquement de même forme peuvent être perçus différemment [Wang, 1992], d'où l'intérêt d'introduire le facteur psycho-acoustique pour tout système ayant pour objectif de conserver la qualité de la parole. Diverses mesures objectives perceptuelles sont élaborées conduisant à de bonnes corrélations avec la perception humaine. Elles sont essentiellement dédiées au codage de la parole, mais trouvent leur application en débruitage de la parole.

#### II.6.3.2.1. Mesure de la qualité perceptuelle vocale (Perceptual Speech Quality Measure, PSQM)

Mesure de la qualité perceptuelle vocale est une méthode objective standardisé par UIT-T en 1998 concernant la mesure de la qualité des codeurs vocaux dans la bande téléphonique. Cette méthode a été initialement développée par la société KPN aux Pays-Bas.

PSQM utilise des représentations psychophysiques qui sont aussi près que possible des représentations internes humaines pour les signaux vocaux. Une valeur zéro de PSQM signifie qu'aucune dégradation n'est présente, et une valeur de 6,5 signifie un canal totalement inutilisable. Les valeurs PSQM peuvent être transposées sur une échelle de type MOS afin de pouvoir obtenir une estimation de la qualité subjective [Beerends et Stemerdink, 1992].

#### II.6.3.2.2. Le modèle E

Le modèle E (E-model 2000), est un modèle de calcul à utiliser dans la planification de transmission [ITU, 2008]. C'est un modèle d'appréciation de la qualité de transmission qui peut être utilisé pour garantir que les utilisateurs seront satisfaits par la performance de transmission de bout en bout. Le modèle intègre les facteurs de dégradation qui affectent l'équipement de transmission, incluant le délai et les codeurs bas débit. Ces dégradations sont calculées en fonction d'une série de paramètres d'entrée pour lesquels des valeurs par défaut et des gammes permises sont précisées. Ils doivent être utilisés si la situation de dégradation correspondante apparaît. Le score MOS équivalent (de type conversation), sur une échelle de 1 à 4,5, peut être obtenu du facteur R, qui est le résultat du modèle E, par les formules suivantes :

$$\begin{cases} MOS = 1, & R < 0 \\ MOS = 1 + 0{,}035R + R(R-60)(100-R)\cdot 7\cdot 10^{-7}, & 0 \le R \le 100. \\ MOS = 4.5, & R > 100 \end{cases}$$

Le graphique de la dépendance entre MOS et facteur R est donné ci-dessous. A noter que le score maximum qu'on peut obtenir dans ce cas est 4,5, le score moyen qui résulte typiquement de la suite de tests subjectifs pour une qualité excellente, car il est connu que les notes d'expérimentateurs varient entre 4 et 5 dans ces conditions.

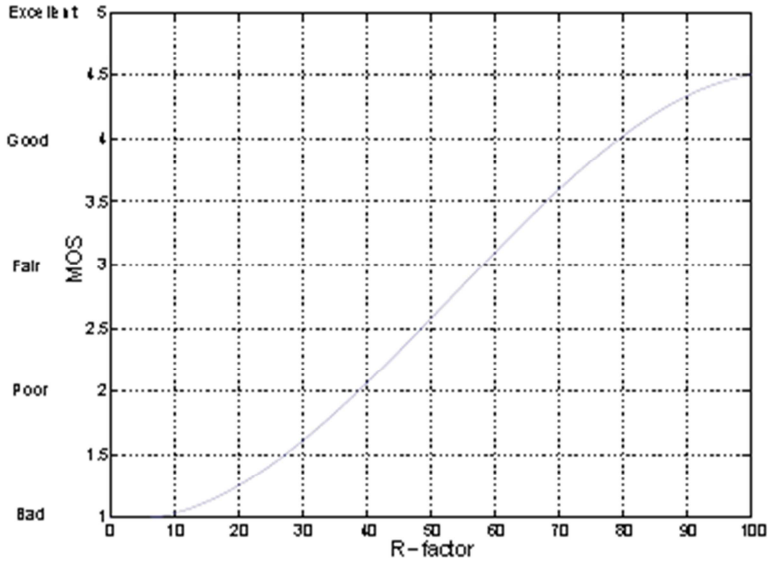



*Figure I-1 : MOS en fonction de R [ITU, 2005]*

| ***Facteur R*** | ***MOS*** | ***Satisfaction des utilisateurs*** |
|---|---|---|
| 90 | 4.35 | Très satisfaits |
| 80 | 4.03 | satisfaits |
| 70 | 3.60 | Certains satisfaits |
| 60 | 3.10 | Beaucoup non satisfaits |
| 50 | 2.58 | Presque tous non satisfaits |

*Tableau II-8 : MOS en fonction des valeurs R du model E [ITU, 2005]*

### II.6.3.2.3. Système d'analyse/mesure perceptuelle (Perceptual Analysis/Measurement System, PAMS)

PAMS est une méthode développée par British Telecom pour déterminer la qualité vocale anticipée pour un système de transmission à travers n'importe quel réseau, y compris ceux où apparaissent des pertes de paquets ou une variation de délai. Le processus de calcul de PAMS utilise un modèle qui combine une description mathématique des propriétés psychophysiques de l'ouïe humaine avec une technique d'analyse qui prend en compte la subjectivité des erreurs dans la perspective de la perception humaine. Le processus PAMS compare deux versions d'un même signal, original et dégradé, et détermine des prédictions MOS, sur une échelle de 1 à 5, pour la qualité de l'écoute et l'effort d'écoute.

Des tests subjectifs extensifs ont été menés avec des sujets humains sur une large variété de technologies de transport de la voix pour valider le modèle PAMS. L'algorithme a été vérifié dans une large gamme de conditions, incluant les codeurs à bas débit. Il en est résulté que le score calculé par PAMS est proche à 0,5 près du MOS typique déterminé par des tests subjectifs contrôlés de laboratoire [Rix, 2002].

#### II.6.3.2.4. Évaluation perceptuelle de la qualité vocale (Perceptual Evaluation of Speech Quality, PESQ)

PESQ est une méthode objective de prédiction de la qualité subjective pour la téléphonie avec bande passante réduite et codeurs vocaux. PESQ combine les meilleurs aspects de PSQM et PAMS. Comparé au PSQM, PESQ prend en compte, en plus, le filtrage, le délai variable, les distorsions de codage et les erreurs de canal.

Le processus clé de PESQ, comme pour les méthodes apparentées, est la transformation des signaux original et dégradé en représentations psychophysiques proches de celles des signaux auditifs du système auditif humain.

Suite à des tests menés de façon extensive, PESQ s'avère fournir des précisions acceptables dans les types d'applications suivants : l'évaluation et la sélection des codeurs, tests dans des réseaux actifs (avec connexions numériques ou analogiques pour la communication VoIP), tests des réseaux émulés et prototype.

Le score PESQ est produit sur une échelle similaire au MOS, avec des valeurs situées entre -0,5 et 4,5. Les valeurs habituelles sont entre 1,0 et 4,5, les scores MOS obtenus dans des expériences subjectives sur la qualité de l'écoute.

La relation entre les scores PESQ et la qualité audio est la suivante :

- Des scores PESQ entre 3 et 4,5 désignent une qualité perçue acceptable (avec 3,8 comme seuil de la qualité dans les systèmes téléphoniques traditionnels) – on va se référer à ce niveau comme qualité « bonne »
- Des valeurs entre 2 et 3 indiquent qu'un effort est nécessaire pour la compréhension du parler – on va se référer à ceci comme qualité « basse »;
- Scores inférieurs à 2 signifient que la dégradation a rendu la communication très difficile ou même impossible, par conséquent la qualité est « inacceptable ».

#### II.6.3.2.5. Signal to Noise Ratio (SNR) ET Segmental SNR

Rapport signal sur bruit est une spécification qui mesure le niveau du signal audio par rapport au niveau de bruit présent dans le signal. Signal aux spécifications bruit sont courantes dans de nombreux composants, y compris les amplificateurs, lecteurs phonographiques, lecteurs CD/DVD, magnétoscopes et autres. Le bruit est décrit comme le souffle, comme dans une platine cassette, ou tout simplement en général un bruit de fond électronique qui se trouve dans toutes les composantes. Comme son nom l'indique, rapport signal sur bruit est un rapport de signal sur bruit et il est exprimé en décibels. Rapport signal sur bruit est abrégé S/N Ratio ou (SNR), un nombre plus élevé signifie une meilleure spécification. S/N ratio de 70db est beaucoup plus souhaitable que le ratio S/N de 50dB. Un S/N de 100db est considéré comme excellent.

IL est bien connu que les traditionnels SNR est pas une mesure exacte dans l'évaluation de la performance des codeurs audio. Cela est, depuis le SNR ne prend pas en

compte les effets de masquage. Segmental SNR est déterminé par le calcul du SNR pour chaque frame établissement d'une moyenne de lui au-dessus du signal entier

### II.6.3.2.6. Measuring Normalizing Blocks (MNB)

Est une méthode d'estimation objective pour la qualité vocale perçue fournir des estimations fiables de qualité pour 4kHz. Il est basé sur un modèle de la perception auditive humaine. La sortie de l'algorithme MNB est la distance auditive (échelle de 0 - ∞) et une cartographie logistique de la distance auditive (échelle de 1-0). Comme une méthode d'essai objective BMN offre plusieurs avantages. Il est rapide, peu coûteuse, facilement reproductible et les résultats sont absolus [Voran, 2002]

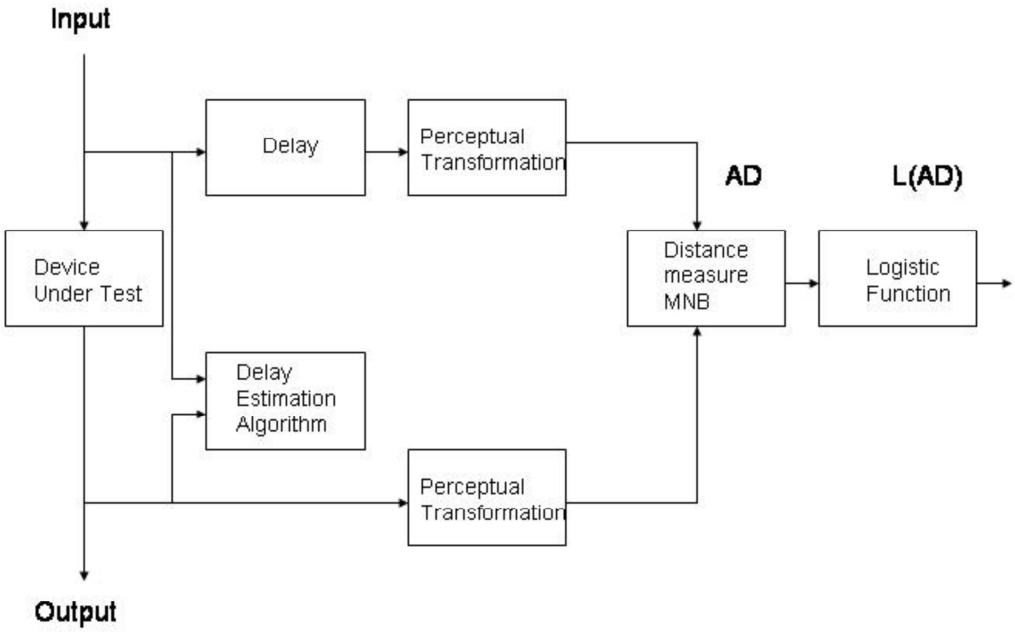



*Figure II-2 : schéma de l'estimation objective pour la distance auditif.[ Voran, 2002]*

### II.6.3.3. Méthodes hybrides

#### II.6.3.3.1. PSQA (Evaluation Pseudo-subjective de la Qualité des Flux VoIP : une approche basée sur les RNN

Cette approche qui se base sur les réseaux de neurones aléatoires est une hybride entre l'évaluation subjective et l'évaluation objective. L'idée de l'approche est de faire évaluer subjectivement plusieurs échantillons dégradés et utiliser ensuite ces résultats, en lui apprenant la relation entre les paramètres à l'origine de dégradation des échantillons et la qualité perçue. L'avantage de cette méthode par rapport aux autres méthodes objectives c'est qu'elle ne base pas sur la comparaison de la qualité du signal émis avec ce le signal reçu et c'est le principe de la plupart des méthodes objectives et donc elles ne sont pas adaptées à une utilisation en temps réel. Finalement l'auteur a démontré que cette méthode a des performances comparable à celles des méthodes les plus performantes, et même meilleures dans les cas où l'on rencontre des conditions de réseau semblables à celles d'internet. De plus cette méthode permet l'évaluation de la qualité d'un flux reçu sans accéder à la version original. Il faut noter aussi que cette approche permet de considérer n'importe quel ensemble de paramètres (bien sûr qui peuvent avoir un effet sur la qualité perçu. Pour VoIP on peut choisir le codec, le taux de perte dans le réseau, le délai moyen de bout en bout, …), ainsi que cette approche peut être appliquée aussi bien à la VoIP, qu'à l'audio hi-fi et à la vidéo. La figure suivante présente une comparaison entre les valeurs de la MOS réel et les valeurs estimées par le RNN la figure montre que l'erreur absolue maximale est 0.58 [Varlela, 2006]

## II.7. Conclusion

La téléphonie par paquets est une application complexe, pour laquelle une excellente qualité de service est nécessaire. Elle peut être obtenue par différentes méthodes, suivant que le réseau est commuté ou routé.

Si le réseau est commuté, les chemins utilisés pour le transport des paquets peuvent être dimensionnés par des techniques d'ingénierie de trafic. Si le réseau est routé, des solutions comme DiffServ permettent de classifier le trafic et de surdimensionné le réseau par rapport au trafic prioritaire constitué essentiellement des paquets de parole.

Ainsi le protocole RTP/RTCP est une solution encore très utilisée aujourd'hui pour compresser plus ou moins le flot téléphonique en fonction de l'état du réseau. L'application s'adapte ainsi au réseau, alors que les solutions plus modernes visent à ce que le réseau s'adapte à l'application.

Chacune de ces solutions est à la recherche pour améliorer la qualité du service, sans prendre en compte la satisfaction du client. Un grand nombre de méthodes d'évaluation de la satisfaction clients utilisateurs du service VoIP sont mises à la disposition pour permettre l'amélioration de la QoS d'où la naissance de la Qualité d'expérience (QoE)

# Chapitre III

- *Les Approches de Routage basées QoS*

# III.Les Approches de Routage Adaptatives basées QoS

## III.1. Introduction:

Aujourd'hui, l'Internet est devenu l'infrastructure de communication la plus importante de la société humaine. Il permet aux utilisateurs mondiaux (individu, groupe et organisationnel) d'accéder et échanger des informations à distance dispersés dans le monde entier. Actuellement, les besoins croissants dans la télécommunication (VoD, vidéoconférence, le VoIP, etc.) et la diversité des flux transportés imposent de nouvelles exigences dans l'architecture d'Internet. À savoir, l'architecture actuelle d'Internet ne peut pas répondre aux exigences des futurs réseaux d'intégrité de service qui transportent le trafic de données multimédia avec une qualité de service élevé.

L'objectif de ce chapitre est de présenter la problématique de l'intégration de la qualité de service dans la prise de décision du routage. Tout d'abord, nous présentons la fonction de routage dans les réseaux de télécommunication ainsi que les algorithmes couramment utilisés. Nous développons ensuite les différentes approches de routage avec qualité de service proposées dans la littérature. L'analyse de ces travaux nous permet d'une part, de retenir le principe du routage adaptatif basé sur l'apprentissage comme base de travail et d'autre part, d'avoir des idées ou en peut intègre la qualité d'expérience dans le processus de routage.

## III.2. Routage dans les réseaux de télécommunication

Le routage est l'une des fonctionnalités principales de la couche réseau qui a la responsabilité de décider sur quelle ligne de sortie, un paquet entrant doit être retransmis. D'une manière générale, on distingue la remise directe, qui correspond au transfert d'un datagramme entre deux ordinateurs du même réseau, et la remise indirecte qui est mise en œuvre quant au moins un routeur sépare l'expéditeur initial et le destinataire final.

Par exemple, dans le cas d'un réseau Ethernet, la remise directe consiste à encapsuler le datagramme IP dans une trame Ethernet après avoir utilisé le protocole ARP [Plummer, 1982] pour faire la correspondance adresse IP / adresse physique, puis à émettre cette trame sur le réseau. L'expéditeur peut savoir que le destinataire final partage le même réseau en utilisant simplement l'adresse IP de destination du datagramme. Il en extrait l'identificateur de réseau et si c'est le même que celui de sa propre adresse IP, alors la remise directe est suffisante. En fait, ce mécanisme, on le retrouve toujours lors de la remise d'un datagramme entre le dernier routeur et le destinataire final.

Pour sa part, la remise indirecte nécessite de déterminer vers quel routeur envoyer un datagramme IP en fonction de sa destination finale. Ceci est rendu possible par l'utilisation d'une table de routage spécifique à chaque routeur, permettant de déterminer vers quelle voie de sortie envoyer un datagramme destiné à un réseau quelconque.

### III.2.1. La table de routage

Une table de routage est une structure de données utilisée par un routeur ou un ordinateur en réseau et qui associe des préfixes à des moyens d'acheminer les datagrammes vers leur destination.

L'essentiel du contenu d'une ligne dans la table de routage est constitué des informations suivant :

- ***Destination*** est l'adresse IP d'une machine ou d'un réseau de destination.
- ***Passerelle*** (Gateway) est l'adresse IP du prochain routeur vers lequel envoyer le datagramme pour atteindre cette destination.
- ***Masque*** est le masque associé au réseau de destination.
- ***Interface*** désigne l'interface physique par laquelle le datagramme doit réellement être expédié.

Une table de routage contient notamment une route par défaut qui spécifie un routeur vers lequel sont envoyés tous les datagrammes pour lesquels il n'existe pas de route dans la table.

Les route statistiques sont configurées manuellement et explicitement par l'administrateur réseau, tandis que les routes dynamiques sont calculées automatiquement via un algorithme et un (des) protocoles de routage dynamique.

Tous les routeurs mentionnés dans une table de routage doivent être directement accessibles à partir du routeur considéré. Cette technique, dans laquelle un routeur ne connaît pas le chemin complet menant à une destination, mais simplement la première étape de ce chemin, est appelée routage par sauts successifs (next-hop routing).

| Adresse réseau | Masque | Passerelle | interface |
|---|---|---|---|
| 200.50.62.0 | 255.255.255.0 | 200.50.62.2 | 200.50.62.2 |
| 200.50.63.0 | 255.255.255.0 | 200.50.63.2 | 200.50.63.2 |
| 200.50.64.1 | 255.255.255.255 | 200.50.64.2 | 200.50.64.2 |
| 0.0.0.0 | 0.0.0.0 | 200.50.64.1 | 200.50.64.2 |

*Tableau III-1 : Exemple d'une table de routage*

### III.2.2. Algorithme de routage

On utilise des graphes pour formuler les problèmes de routage. Le graphe G=(N,E) est un ensemble N de nœuds et un collection E d'arrêtes, où chaque arrête de E relie deux nœuds de N. Dans notre contexte, les nœuds sont les routeurs, et les arrêtes sont les liaisons. Les arrêtes ont une valeur représentant leur coût. Typiquement, le coût d'une arrête représente sa longueur, sa vitesse ou le cout financier associé à son utilisation.

On dit ici que c(x,y) représente le cout de l'arrête entre les nœuds x et y. Si la paire (x,y) n'appartient pas à E, alors c(x,y)=∞. On considère des graphes non orientés, si bien que les paires (x,y) et (y,x) sont équivalentes. Si (x,y) appartient à E, alors on dit que x et y sont *voisins*. Dx(y) représente le cout du chemin de moindre cout entre les nœuds x et y.

Le but d'un algorithme de routage est donc de trouver le chemin de cout minimum entre les routeurs source et destination, où le cout d'un chemin est simplement calculé comme étant la somme des couts des liens utilisés.

Il existe deux grands types d'algorithmes de routage :

- *Les algorithmes de routage globaux* choisissent le chemin de cout minimum en ayant une connaissance complète du réseau.
- *Les algorithmes de routage décentralisés* calculent le meilleur chemin de façon itérative. Aucun nœud n'a d'information complète sur tous les liens du réseau, mais ne connaissent que leurs voisins. De proche en proche, ils s'échangent des informations, et les routeurs peuvent donc petit à petit construire une table de routage contenant tous les routeurs du réseau.

## III.3. Les algorithmes de routage classique

Traditionnellement, un réseau est divisé en plusieurs systèmes autonomes décentralisés. Un AS est défini comme un ensemble de routeurs qui utilisent le même protocole de routage, et sont gérés par une seule autorité. Un IGP (interior gateway protocol) est utilisé pour router le trafic entre les hôtes ou les réseaux appartenant au même AS (RIP, OSPF, IS-IS), ce qui est communément connu sous le nom routage intra-domain. D'autre part, un EGP est utilisé pour router le trafic entre distincte ASs (BGP, CIDR, IDRP). Ce type de routage est connu sous le nom routage inter-domain.

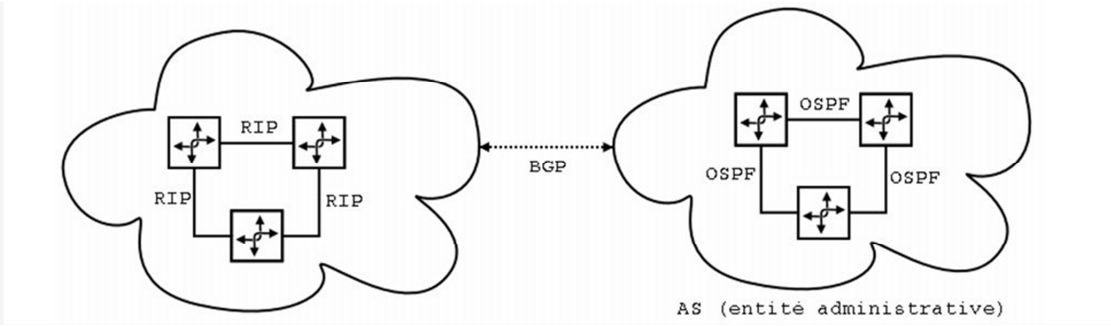



*Figure III-1 : Type de routage [Ziani, 2008]*

Dans les deux cas, un algorithme de routage est basé sur le saut par saut et le paradigme de plus court chemin. La source du paquet spécifie l'adresse de la destination, et chaque routeur le long de chemin transmet le paquet au voisin le plus proche de la destination. Le chemin optimal est choisi en fonction d'un critère donné qui ne pourraient pas nécessairement mener au meilleur routage dans l'Internet. Par exemple, quand le réseau est fortement chargé, certains routeurs pourraient présenter un retard excessif dans le transit des paquets tandis que d'autres routeurs sont insuffisamment

utilisés. Cet usage non optimisé des ressources du réseau peut introduire non seulement des retards excessifs, mais aussi taux élevé de perte de paquets. Parmi les algorithmes de routage largement employées dans les routeurs du même AS, on peut noter: un algorithme à vecteur de distance tel que l'algorithme mis en œuvre dans le protocole RIP et l'algorithme a l'état des liens tels que OSPF. Les métriques (mesures de coût des liens) utilisés dans ces types d'algorithmes pour sélectionner le meilleur chemin ne tiennent pas compte des variations de charge de trafic dans un lien menant vers la performance limitée de routage. [Ziani, 2008]

### III.3.1. Interior gateway protocol (IGP protocoles de routage interne)

***L'approche à vecteur de distance***

Également connu sous le nom Bellman-Ford ou Ford-Fulkerson, le cœur de ce type d'algorithme est la table de routage maintenue par chaque hôte. Avec le schème de routage à vecteur de distance, chaque nœud échange avec ses voisins sa distance à d'autres réseaux. Les nœuds voisins utilisent ces informations pour déterminer leur distance par rapport à ces réseaux. Subséquemment ces nœuds partagent cette information avec leurs voisins, qui répètent le processus et ainsi de suite jusqu'à ce que tous les nœuds de réseau se rendent compte de cette information. De cette manière l'information d'accessibilité est diffusée à travers le réseau. Éventuellement chaque nœud apprend quel voisin (routeur de saut suivant) à utiliser pour atteindre une destination particulière avec un nombre minimum des sauts. Un nœud ne se renseigne pas sur les nœuds intermédiaires à la destination. Les protocoles à vecteur de distances ont particulièrement sensibles aux boucles de routage. Une métrique de mesure infinie (count to infinity) est le résultat d'une boucle de routage qui engage les routeurs à incrémenter à l'infini la métrique de mesure. Diverses solutions sont proposées pour pallier à cette faiblesse. Les protocoles de routage à vecteur de distance sont donc conçus pour fonctionner uniquement sur les petits réseaux.

```
Initialization:
  for all destinations y in N:
    Dx(y) = c(x,y)
  for each neighbor w
    Dw(y) = ∞ for all destinations y in N
  for each neighbor w
    send distance vector to w

loop
  wait until I see a link cost change to some neighbor or
             I receive a distance vector form some neighbor

  for each y in N:
    Dx(y) = minv( c(x,y) + dv(y) )

  if Dx(y) changed for any destination y
    send distance vector
forever
```

***Exemple de protocole de routage à vecteur de distance : RIP***

Le protocole RIP (Routing Information Protocol) est un protocole intra-domaine (IGP Interior Gateway Protocol) [Hedrick, 1988]. Il est basé sur un algorithme de routage à vecteur de distance, dont la métrique est le nombre de sauts. Il est prévu pour des réseaux dont la distance entre 2 points est au maximum de 15 sauts.

Examinons un peu plus en détail le fonctionnement de RIP. Lors de l'initialisation du routeur, celui-ci détermine l'adresse réseau de ses interfaces puis envoie sur chacune une demande d'informations (table RIP complète) aux routeurs voisins. Lors de la réception d'une demande, un routeur envoie sa table complète ou partielle suivant la nature de cette demande. Lors de la réception d'une réponse, il met à jour sa table si besoin. Deux cas peuvent se présenter :

- Pour une nouvelle route, il incrémente la distance, vérifie que celle-ci est strictement inférieure à 15 et diffuse immédiatement le vecteur de distance correspondant.
- pour une route existante mais avec une distance plus faible, la table est mise à jour. La nouvelle distance et, éventuellement, l'adresse du routeur si elle diffère sont intégrées à la table.

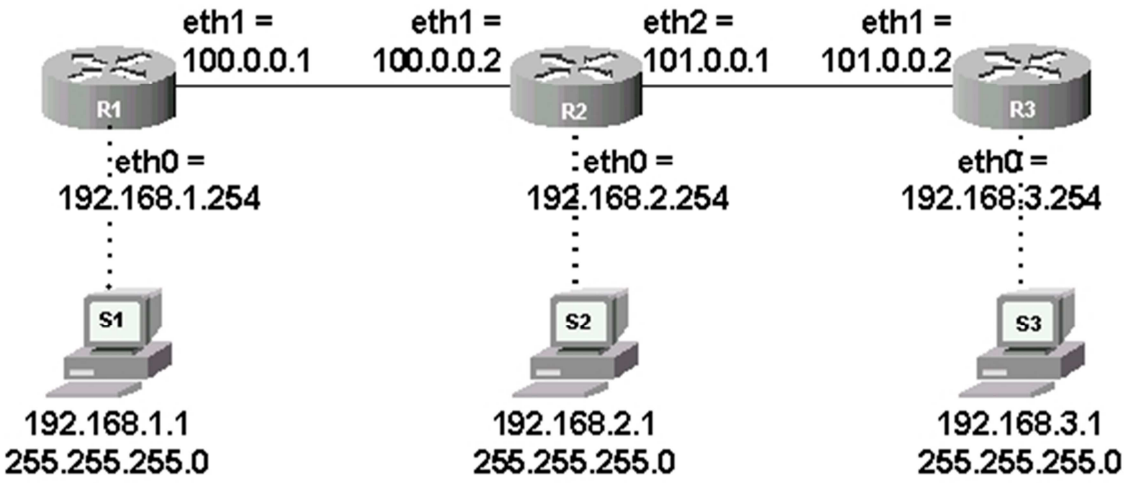



*Figure III-2 : Topologie réseau[Innokenty, 2000]*

Sur R1, lorsque l'on active le processus de routage RIP, une première table est constituée à partir des adresses IP des interfaces du routeur. Pour ces réseaux directement connectés au routeur, la distance est égale à un puisqu'il faut au moins traverser ce routeur pour les atteindre.

| Adresse/Masque | interface | Distance |
|---|---|---|
| 100.0.0.0/8 | Eth1 | 1 |
| 192.168.1.0/24 | Eth0 | 1 |

*Tableau III-2 : table de routage initiale constituée par R1*

R1 transmet à ses voisins immédiats (ici, il n'y a que R2) un seul vecteur de distance {192.168.1.0/24, 1} qui signifie : "je suis le routeur d'adresse IP 100.0.0.1 et je connais un moyen d'atteindre le réseau 192.168.1.0/24 en un saut". Aucune information sur le réseau commun aux deux routeurs (100.0.0.0/8) n'est transmise car R1 considère que R2 connaît déjà ce réseau.

Ensuite, lorsque l'on active RIP sur R2, il constitue la table ci-après à partir de ses propres informations et de celles reçues de R1 :

| Adresse/Masque | interface | Distance |
|---|---|---|
| 100.0.0.0/8 | Eth1 | 1 |
| 101.0.0.0/8 | Eth1 | 1 |
| 192.168.1.0/24 | 100.0.0.1 | 2 |
| 192.168.2.0/24 | Eth0 | 1 |

*Tableau III-3 : table de routage sur le routeur R2*

Sur R2, RIP a calculé que la distance pour atteindre 192.168.1.0/24 est égale à deux puisqu'il faut traverser R2 puis R1. R2 a déduit le "moyen de l'atteindre" à partir de l'adresse IP de l'émetteur contenue dans le paquet RIP.

De nombreuses améliorations à RIP ou plus généralement aux protocoles à vecteur de distance ont été apportées. On peut citer notamment la version 2 du protocole [Malcolm et Zhao, 1993], solutionnant le problème de la taille des tables de routage dans RIP-version 1 d'une part, en permettant le routage par sous réseau et d'autre part, en traitant les communications multipoints. Ce protocole offre également une meilleure authentification des échanges entre

### *L'approche à état de lien*

Les algorithmes de routage à état de liens actualisent une base de données complexe sur la topologie du réseau en échangeant des mises à jour de routage à état de liens avec les autres routeurs du réseau. Ces algorithmes ont les caractéristiques suivantes:

- Ils sont désignés collectivement comme protocoles du plus court chemin d'abord (SPF).

- Ils actualisent une base de données complexe sur la topologie du réseau.
- Ils sont basés sur l'algorithme de Dijkstra.

Contrairement aux protocoles à vecteur de distance, ils développent et actualisent une connaissance complète des routeurs du réseau ainsi que de leur mode d'interconnexion. Cela est possible grâce à l'échange de mises à jour de routage à état de liens (LSA) avec les autres routeurs du réseau. Chaque routeur qui échange des LSA construit une base de données topologique à l'aide de toutes les LSA reçues. Un algorithme SPF est ensuite utilisé pour calculer l'accessibilité des destinations en réseau. Ces informations sont utilisées pour mettre à jour la table de routage. Ce processus a la capacité de découvrir les modifications de la topologie réseau provoquées par la panne d'un composant ou par la croissance du réseau. L'échange de LSA est déclenché par un événement sur le réseau plutôt que par des mises à jour périodiques. Cela peut accélérer considérablement le processus de convergence car il n'a pas besoin d'attendre l'expiration d'une série de compteurs pour que les routeurs en réseau puissent commencer de converger.

```
Initialisation:
  N' = {u}
  for all nodes v
    if v is neighbor of u then
      D(v) = c(u,v)
    else
      D(v) = ∞

Loop
  find w not in N' such that D(w) is a minimum
  add w to N'
  update D(v) for each neighbor v of w and not in N':
    D(v) = min(D(v), D(w) + c(w,v))
until N' = N
```

***Exemple de protocole de routage à état de lien : OSPF***

**Open Shortest Path First** (OSPF) est un protocole de routage interne IP de type «à état de liens». Il a été développé au sein de l'Internet Engineering Task Force (IETF) à partir de 1987. La version actuelle d'OSPFv2 est décrite dans la RFC 2328 en 1997. Une version 3 est définie dans la RFC 2740 et permet l'utilisation d'OSPF dans un réseau IPv6.routeurs.

OSPF utilise l'algorithme de Dijkstra pour déterminer le meilleur chemin à prendre. On le nomme aussi algorithme SPF (Shortest Path First) ou algorithme du plus court chemin d'abord. Il a été formulé par Edsger Dijkstra.

OSPF déclenche ses mises à jour à chaque changement dans la topologie du réseau, ce qui permet de réduire le temps de convergence. À partir d'une mise à jour, un routeur

met en place une base de données topologique permettant le calcul de l'accessibilité aux réseaux grâce au calcul d'un arbre de la topologie dont le routeur est la racine.

Contrairement à RIP, OSPF a été pensé pour supporter de très grands réseaux. Mais, qui dit grand réseau, dit nombreuses routes. Donc, afin d'éviter que la bande passante ne soit engloutie dans la diffusion des routes, OSPF introduit le concept de zone (area). Le réseau est divisé en plusieurs zones de routage qui contiennent des routeurs et des hôtes. Chaque zone, identifiée par un numéro, possède sa propre topologie et ne connaît pas la topologie des autres zones. Chaque routeur d'une zone donnée ne connaît que les routeurs de sa propre zone ainsi que la façon d'atteindre une zone particulière, la zone numéro 0. Toutes les zones doivent être connectées physiquement à la zone 0 (appelée *backbone* ou réseau fédérateur). Elle est constituée de plusieurs routeurs interconnectés. Le *backbone* est chargé de diffuser les informations de routage qu'il reçoit d'une zone aux autres zones. Tout routage basé sur OSPF doit posséder une zone 0.

OSPF distingue quatre types de routeurs:

***Routeur interne*** : un routeur dont toutes les interfaces se trouvent dans la même aire.

***Area Border Router (ABR)*** : un routeur qui dispose d'interfaces dans des aires différentes.

***Autonomous System Border Router (ASBR)*** : un routeur qui injecte dans OSPF des routes qui proviennent d'autres protocoles de routage ou des routes statiques

***Routeur backbone*** : un routeur dont au moins une interface appartient à l'aire 0. Tous les ABR sont des routeurs backbone

Dans OSPF, chaque routeur établit des relations d'adjacence avec ses voisins immédiats en envoyant des messages *hello* à intervalle régulier. Chaque routeur communique ensuite la liste des réseaux auxquels il est connecté par des messages *Link-state advertisements* (LSA) propagés de proche en proche à tous les routeurs du réseau. L'ensemble des LSA forme une base de données de l'état des liens *Link-State Database* (LSDB) pour chaque aire, qui est identique pour tous les routeurs participants dans cette aire. Chaque routeur utilise ensuite l'algorithme de Dijkstra, *Shortest Path First* (SPF) pour déterminer la route la plus courte vers chacun des réseaux connus dans la LSDB.

Le bon fonctionnement d'OSPF requiert donc une complète cohérence dans le calcul SPF, il n'est donc par exemple pas possible de filtrer des routes ou de les résumer à l'intérieur d'une aire.

En cas de changement de topologie, de nouveaux LSA sont propagés de proche en proche, et l'algorithme SPF est exécuté à nouveau sur chaque routeur.

Il existe 5 types de paquets OSPF:

- *Hello* : découverte des voisins et maintien des adjacences.
- *Database Description* : description des LSA.
- *Link State Request* : requête d'un LSA.
- *Link State Update* : mise à jour d'un LSA.

- *Link State Acknowledgement* : acquittement d'un LSA.

(Hello) est utilisé pour l'établissement et le maintien des adjacences, les autres types sont utilisés pour la synchronisation de la LSDB.

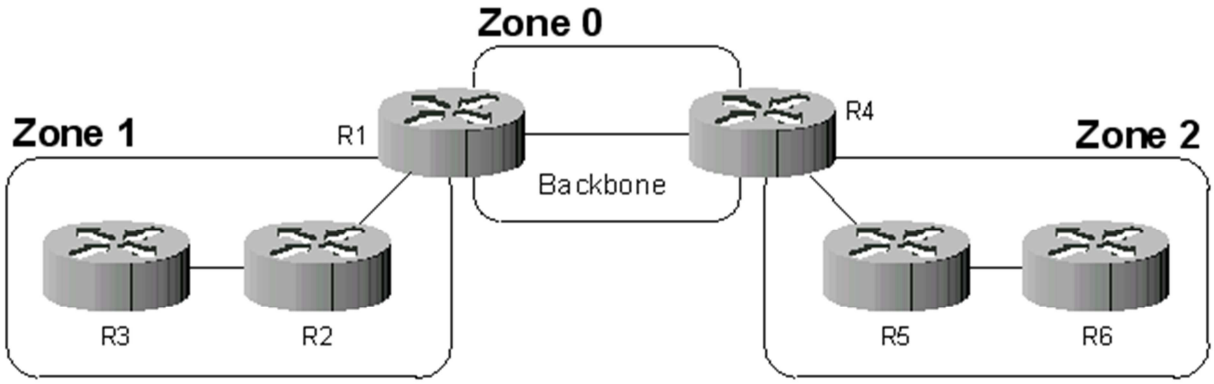



*Figure III-3 : Réseau découpé en 3 zones[Innokenty, 2000]*

### III.3.2. Exterior gateway protocol (EGP protocoles de routage extérieur)

EGP est utilisé pour désigner, de façon générale, les protocoles de routage extérieur, c'est-à-dire entre deux systèmes autonomes différents, et par opposition aux protocoles de routage interne. Au début EGP est un protocole de routage dans Internet. Décrit pour la première fois en octobre 1982 dans la RFC 827 La version 3 d'EGP a été utilisée au début d'Internet. Aujourd'hui, ce protocole est obsolète et remplacé par BGP

BGP (Border Gateway Protocol) est un protocole d'échange de route utilisé notamment sur le réseau Internet. Son objectif est d'échanger des informations d'accessibilité de réseaux (appelés préfixes) entre systèmes autonomes (AS) car il a été conçu pour prendre en charge de très grands volumes de données et dispose de possibilités étendues de choix de la meilleure route (RFC1774).

Les sessions de routage au sein d'un AS sont appelées sessions iBGP. Les sessions de routage entre AS sont appelées sessions eBGP. BGP ne transmet pas de métrique dans les mises à jour de routes, mais transmet la meilleure route vers chaque système autonome susceptible d'être adoptée pour atteindre une destination donnée.

Lorsqu'un routeur BGP reçoit des mises à jour en provenance de plusieurs systèmes autonomes décrivant différents chemins vers une même destination ‹il choisit alors le meilleur itinéraire pour l'atteindre et le propage vers ses voisins.

Une décision de routage est fondée sur plusieurs attributs comme :

- *AS-path* : Cet attribut liste les numéros de tous les AS qu'une mise à jour a traversés pour atteindre une destination.
- *Origin* : Cet attribut donne des informations sur l'origine de la route. Ces informations peuvent être de type IGP (la route annoncée provient du même système autonome que l'annonceur), de type EGP (la route est apprise et ne provient pas du même système autonome) ou Incomplète (la route est apprise d'une autre manière.
- *Next hop* : Cet attribut contient l'adresse IP du routeur vers lequel l'information doit être émise pour atteindre le réseau.
- *Weight* : Cet attribut est utilisé dans le processus de sélection de chemin lorsqu'il existe plus d'une route vers une même destination. Cet attribut de poids est local et n'est pas propagé dans les mises à jour de routage.
- *Local preference* : Cet attribut a un rôle similaire à l'attribut de poids, si ce n'est qu'il fait partie des informations de mise à jour de routage.
- *Multi-exit discriminator*: Cet attribut indique aux routeurs voisins externes le chemin à privilégier vers un AS lorsque celui-ci possède plusieurs points d'entrée.
- *Community* : Cet attribut est utilisé pour grouper des destinations auxquelles des décisions de routage peuvent être appliquées.

Le premier niveau de topologie de routage définit les différentes sessions de routage entre les systèmes autonomes participant au routage BGP. Le découpage en différents AS dépend de nombreux paramètres, tels que le nombre d'équipements réseau, la localisation géographique, etc. Pour des raisons de disponibilité et de résilience des connexions entre les systèmes autonomes, le nombre de sessions de routage entre les systèmes autonomes de l'opérateur de télécommunications doit être supérieur au minimum à deux sessions.

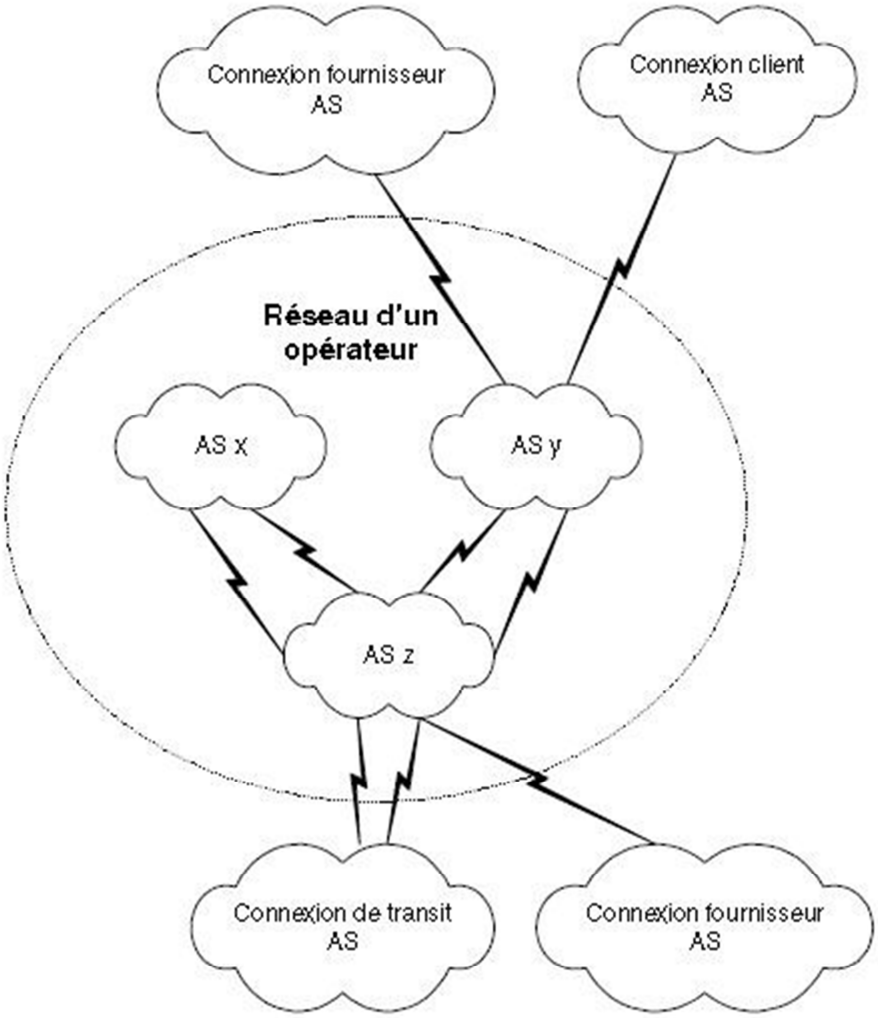



*Figure III-4 : Topologie de routage des systèmes autonomes BGP [Ziani, 2008]*

## III.4. Les approches de routage basées QoS

L'intérêt pour le routage à base des contraintes avait attiré beaucoup d'attention ces dernières années. Pour réduire la complexité, la plupart de ces approches sont basées sur l'heuristique. Nous pouvons classifier trois grandes catégories des CBR :

### III.4.1. Les approches Label-switching/reservation :

Stimulé par des approches telles qu'ATM PNNI, MPLS ou GMPLS.

Avec MPLS, des étiquettes de longueur fixe sont attachées aux paquets à un routeur d'entrée, et des décisions d'expédition sont basées sur ces étiquettes dans les routeurs intérieurs du chemin label-switched. L'engainerai de trafic de MPLS nous permet de

dépasser le protocole de routage de défaut, ainsi l'expédition au-dessus des chemins pas normalement considérés.

MPLS n'est pas proprement un protocole de niveau 3 (couche réseau), mais plutôt intermédiaire entre le niveau 2 (couche liaison) et le niveau 3 (couche réseau) du modèle OSI.

Paradoxalement, pour une couche située sensiblement au niveau inférieur à IP, MPLS a besoin d'IP et des protocoles de routage associés pour exister. MPLS est une technologie permettant d'offrir à IP un mode circuit, à l'image de X25 ou ATM.

Le principe de la commutation de label consiste à remplacer la recherche de la plus longue correspondance entre l'adresse de destination des paquets IP, et les préfixes présents dans les tables de routage en insérant un label de longueur fixe entre l'en-tête réseau et l'entête liaison des paquets. La détermination du prochain est alors effectuée grâce au label. Cette solution est plus avantageuse, puisqu'elle ne s'appuie plus sur la recherche de la plus longue correspondance avec des préfixes de longueur variable, mais sur la recherche d'un label de longueur fixe.

Le routage MPLS peut être effectué localement. En effet, il est possible que dans un réseau IP, certaines zones utilisent MPLS et d'autres non. La distribution des labels est alors effectuée par les routeurs d'accès au "domaine MPLS", grâce à un algorithme dynamique qui associe des labels aux adresses de destination des paquets entrants. Pour cela, les routeurs d'extrémité du domaine MPLS communiquent avec leurs voisins non MPLS. Ils échangent avec eux des informations de routage, et distribuent les labels aux routeurs internes pour effectuer un routage cohérent. Le routage est donc explicite, défini de bout en bout au sein du domaine MPLS.

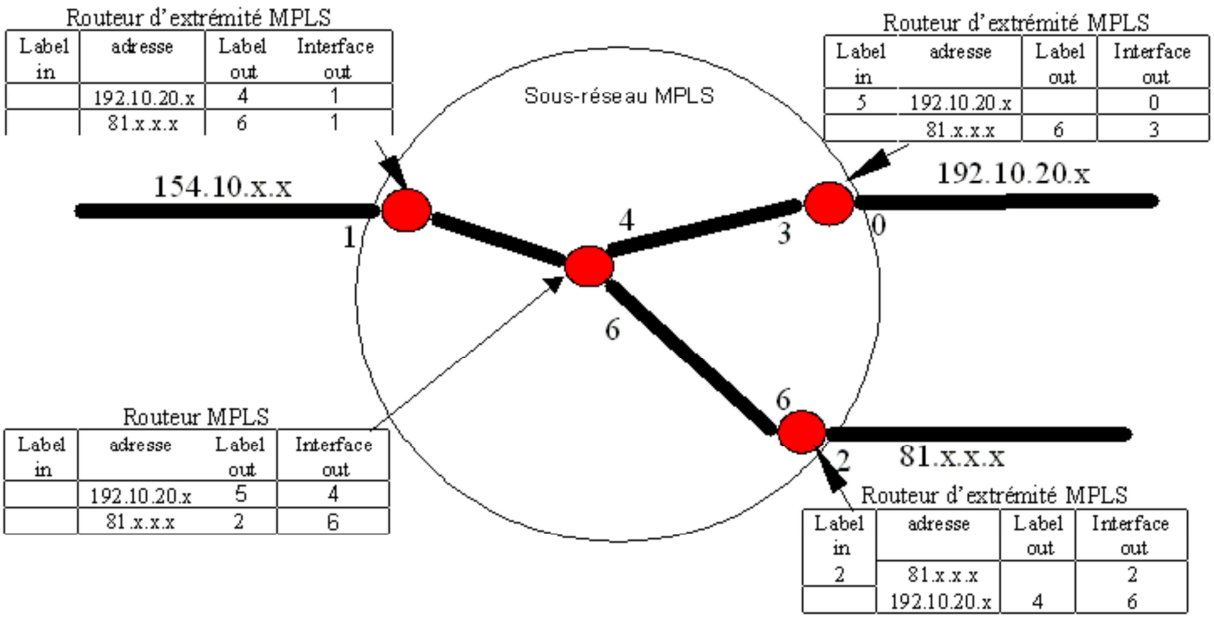



*Figure III-5 : Routage MPLS[Ziani, 2008]*

Un protocole de réservation de ressource tel que RSVP doit être utilisé pour réserver les ressources requises.

Une autre architecture proposée pour fournir l'Internet QoS est l'architecture différenciation de services (DiffServ). Balances de DiffServ bien en poussant la complexité aux frontières de domaine de réseau.

### III.4.2. Les approches Multi-Constrained path (MCP) :

Le but de toutes telles approches est de rechercher le plus court chemin de l'ensemble des chemins faisables entre deux nœuds. Le travail considérable dans la littérature a concentré sur un cas spécial du MCP le chemin du moindre coût parmi ceux qui satisfont seulement une contrainte [Kuipers et al., 2004].

Soit un réseau $G(N, E)$ où à chaque lien $(u, v) \in E$ est associé un vecteur de $m$ additives métriques de QoS $w_i(u, v) \geq 0, i = 1..m$. Et soit $m$ contraintes $L_i, i = 1..m$. Le problème et de trouver un chemin (path) $P$ du nœud source $s$ au nœud distinataire $d$ tel que : $w_i(P) = \sum_{(u,v)\in P} w_i(u, v) \leq L_i, i = 1..m$.

Un chemin satisfaisant cette condition est appelé chemin faisable (feasible path). Aucunes ou plusieurs chemin faisables peuvent exister. Il serait mieux de trouver le chemin le plus court $l(P)$ de cet ensemble. C'est le problème multi-constrained optimal path (MCOP).

#### III.4.2.1. Les concepts du routage multi-contrainte

***Le choix de la fonction d'énergie*** : Pour calculer le coût d'un chemin, il existe deux types de fonction, les fonctions linéaires et les fonctions non linéaires.

Un exemple d'une fonction d'énergie linéaire est donné par la fonction linéaire de Jaffe : $L(p) = \sum_{i=1}^{k}(di * wi(p))$ où di est un coefficient. Ces fonctions d'énergie ne permettent pas toujours de résoudre le problème MCP. En effet‹ certains chemins ont un coût inférieur à celui du vecteur $\vec{C}$ mais ils ne sont pas forcément faisables.

Un exemple d'une fonction d'énergie non linéaire est donné par : $Lq(p) = (\sum_{i=1}^{k}(\frac{wi(p)}{ci})^q)^{\frac{1}{q}}$. Lorsque q tend vers l'infini cette fonction tend vers la fonction : $L_\infty(p) = \max_{i\in 1..k}(\frac{wi(p)}{ci})$. La fonction non linéaire $L_\infty$ permet d'éliminer tous les chemins infaisables. En e et, tout chemin ayant un coût supérieur à 1 est infaisable. Par conséquent, un chemin retourné par un algorithme qui utilise cette fonction d'énergie est toujours faisable.

***L'approche des K plus courts chemins*** : En utilisant des fonctions d'énergie non linéaires, un sous-chemin d'un chemin optimal n'est pas forcément un chemin optimal. Par conséquent, l'application de l'algorithme de Dijkstra avec une fonction d'énergie non

linéaire ne permet pas toujours de résoudre le problème MCP même si ce problème admet une solution. L'approche des K plus courts chemins propose de mémoriser au niveau d'un nœud intermédiaire les K plus courts chemins et non pas uniquement le plus court chemin. Lorsque K =∞, cette approche permet de résoudre le problème MCP lorsque le problème admet une solution. Afin de réduire la complexité du calcul du chemin faisable, certaines heuristiques prennent K un entier positif. Cependant, ces heuristiques ne permettent pas toujours de résoudre le problème MCP même si un chemin faisable existe.

***La non-dominance*** : Ce concept est très important pour le routage exact avec qualité de service. En effet, il permet de réduire l'espace de recherche des chemins en éliminant tous les chemins dominés de cet espace. Ainsi, ce concept ne considère que les chemins non dominés pour le calcul du chemin optimal. En effet, un chemin dominé ne peut pas être ni un chemin optimal ni un sous-chemin d'un chemin optimal. Par conséquent, l'utilisation de ce concept permet de trouver le chemin optimal tout en réduisant le temps de calcul puisque l'espace de recherche est réduit.

***Le Look-Ahead*** : Ce mécanisme permet de réduire l'espace de recherche. En effet, au niveau d'un nœud intermédiaire n‹ un algorithme de routage qui met en œuvre ce concept utilise des informations sur les chemins qui mènent vers la destination à partir du nœud n. Ces informations permettent d'exclure des chemins non faisables de l'espace de recherche. Ce concept propose de calculer pour chaque contrainte l'arbre des plus courts chemins dont la racine est la destination. Pour cela‹ l'algorithme de Dijkstra est exécuté k fois où k est le nombre de contraintes. Le chemin obtenu pour chaque nœud n et pour chaque contrainte i est noté par $p^*_{n\to D;i}$, où D est la destination.

### III.4.2.2. Algorithmes pour MCP

#### III.4.2.2.1. TAMCRA et SAMCRA

L'algorithme TAMCRA et son successeur SAMCRA se constitue de trois concepts fondamentaux:

- Une fonction de poids de chemin non-linéaire: $l(P) = max_{j=1..m}\left(\frac{w_j(P)}{L_j}\right)$;
- l'approche de l'algorithme *K-shortest path*;
- Le principe du chemin non-dominé pour réduire l'espace de recherche.

Lors de la phase de l'exécution de l'algorithme K-shortest path, TAMCRA ne stocke pas dans les chemins intermédiaires tous les k sous chemins venant de la source mais il fait une distinction selon le principe de non-dominance. Un chemin $Q$ est dominé par le chemin $P$ si $w_i(P) \leq w_i(Q)$ pour tout $i = 1..m$, avec l'inégalité pour au moins un seul $i$. TAMCRA stocke seulement les sous chemins non-dominé, et réduit, ainsi, l'espace de recherche sans compromettre la qualité de résultat. SAMCRA ajoute le concept de look-

ahead pour réduire davantage l'espace de recherche. La complexité de SAMCRA est évaluée à $O(kNlog(kN) + K^2mE)$ [Kuipers et al., 2004].

#### III.4.2.2.2. H_MCOP

Cet algorithme cherche un chemin satisfaisant les contraintes par l'utilisation de la fonction non-linéaire définie par l'algorithme TAMCRA (vue auparavant) et le concept de Look-ahead. Il cherche de minimiser, simultanément, la fonction objective et la valeur de la métrique du coût.

#### III.4.2.2.3. Algorithme de Jaffe

L'approximation de Jaffe est basée sur la relaxation lagrangienne pour minimiser la combinaison linéaire des métriques sous la forme : $w(u,v) = \sum_{i=1}^{m} d_i w_i(u,v)$ ;

où $d_i$ sont des multiplicateurs positifs, avec la contrainte : $w_i(P) \leq L_i$. Mais, comme mentionné auparavant, la combinaison linéaire peut donner un chemin qui viole le contrainte. Pour cela, Jaffe a proposé une fonction non linéaire de la forme $f(P) = \sum_{i=1}^{m} \max(w_i(P), L_i)$ dont la minimisation assure de trouvé un chemin faisable s'il existe. La complexité en temps de l'algorithme de Jaffe est évaluée à $O(NlogN + mE)$ [Kuipers et al., 2004].

### III.4.3. Les approches inductives :

Pour pouvoir prendre une décision de routage optimal selon des critères d'exécution appropriés, un nœud de réseau doit avoir la connaissance complète de l'état de réseau entier et une prévision précise de l'évolution du réseau et de sa dynamique. Ce, cependant, est impossible à moins que l'algorithme de routage soit capable de l'adaptation aux changements d'état de réseau en temps réel. Ainsi, il est nécessaire de concevoir les algorithmes de routage intelligents et adaptatifs qui tiennent compte de l'état de réseau et de son évolution. Nous devons explorer des algorithmes de routage basés QoS dépendants de d'état de réseau.

## III.5. Approches inductives basées sur les paradigmes ' machine learning'

L'approche inductive est basée sur l'intelligence artificielle et des techniques inspirées de la biologie comme l'apprentissage par renforcement et les algorithmes génétiques pour contrôler l'état de réseau en temps réel afin d'offrir aux utilisateurs la QoS requise [Mellouk, 2009]. Parmi ces approches on cite:

- Le routage CPN (Cognitive Packet Network)

- Routage par colonies de fourmis
- Q-routing

Généralement ces approches ce base sur l'utilisation des réseaux de neurones et les méthodes d'apprentissage par renforcement.

### III.5.1. Réseaux de neurones et apprentissage par renforcement

Le neurone (cellule nerveuse) est l'élément de base du cerveau. Il reçoit l'influx nerveux par des ramifications courtes et nombreuses (dendrites), et transmet à son tour les influx nerveux par une fibre unique ramifiée à son extrémité (axone). L'information est transmise d'une cellule à l'autre en des points de contacts spécialisés (synapses). Les neurotransmetteurs chimiques, libérés dans la synapse, modifient le potentiel intracellulaire, augmentant ou diminuant ainsi la probabilité que celui-ci donne naissance à l'influx nerveux. Si le potentiel intracellulaire est en dessous d'un certain seuil, le neurone est au repos. Dans le cas contraire, le neurone envoie des signaux électriques. Typiquement, le neurone biologique comprend les éléments suivants:

- Un corps cellulaire contenant le noyau.
- Des dendrites, ou neuro récepteurs.
- Un axone qui permet le transfert de signaux à travers les synapses.

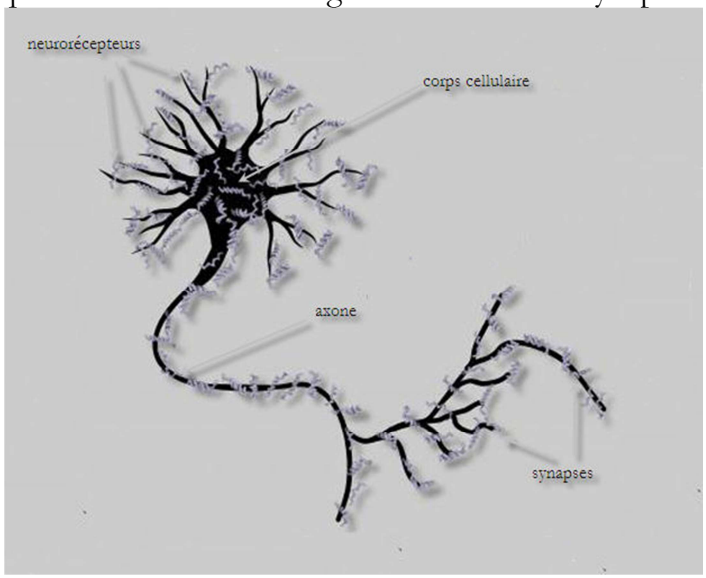



*Figure III-6 : schéma d'un neurone réel*

L'essor des neurones artificiels a débuté par la proposition du premier neurone formel par McCulloch et Pitts [Mccullogh et Pitts, 1943]. Puis les travaux de Rosenblatt, ont concerné un neurone formel doté de caractéristiques d'adaptabilité, appelé "Perceptron". Ce perceptron a permis de comprendre que les réseaux de neurones artificiels possèdent d'intéressantes capacités d'apprentissage qui peuvent être utilisées pour la résolution de certains problèmes.

Un neurone formel est une entité mathématique simple qui calcule son potentiel V en déterminant la somme pondérée de ses entrées xj [Mccullogh et Pitts, 1943]. Sur la

Figure III.7, les {wij} représentent les coefficients de pondération, appelés aussi poids synaptiques. La sortie S du réseau est en général une fonction non linéaire du potentiel S=f(v).

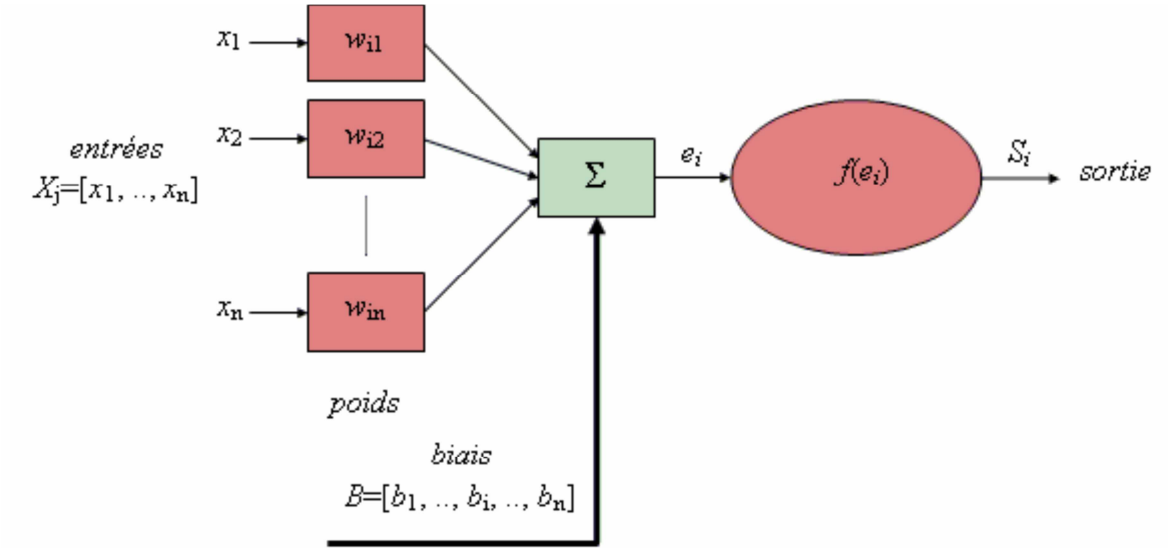



*Figure III-7 : structure interne d'un neurone formel[Kaelbling et al., 1996]*

L'apprentissage consiste à trouver les valeurs des coefficients {wij} afin que le réseau de neurones remplisse la fonction qui lui est demandée : classification, prédiction, etc.

Soit un réseau de neurones composé d'une couche d'entrée, d'une couche cachée et d'une couche de sortie, et soient :

- $x$ : le vecteur à $I$ éléments représentant le $k^{ième}$ stimulus. La matrice $IxK$ des $K$ stimuli à apprendre est notée $X$.
- $h_k$ : le vecteur à L éléments représentant la réponse des L cellules de la couche cachée lorsque le $k^{ième}$ stimulus est présenté en entrée (la couche cachée comporte L cellules).
- $o_k$ : le vecteur à J éléments représentant la réponse des cellules de la couche de sortie pour le $k_{ième}$ stimulus (la couche de sortie comporte J cellules).
- $t_k$ : le vecteur à J éléments représentant la réponse désirée des cellules de la couche de sortie pour le kième stimulus. La matrice des réponses désirées de dimension J×K est notée T.
- W : la matrice de dimension L×I, des valeurs des connexions reliant les cellules de la couche d'entrée à celles de la couche cachée. L'élément $W_{l,i}$ , représente la valeur de la connexion entre la $i^{eme}$ cellule d'entrée et la $l^{eme}$ cellule de la couche cachée.
- Z : la matrice de dimension J×L, des valeurs des connexions reliant les cellules de la couche cachée à celles de la couche de sortie. $Z_{l,j}$, représente la valeur de la connexion entre la $l^{eme}$ cellule de la couche cachée et la $j^{eme}$ cellule de sortie.

En notant $a_n$, l'état d'activation de la cellule $n$ (qui peut être une cellule de la couche cachée ou de la couche de sortie), la réponse de la cellule, notée $o_n$ , est :

$$On = f(an)$$

Ou *f* représente une fonction de transfert (fonction d'activation). Une des fonctions les plus utilisées est la fonction sigmoïde, définie par :

$$f(x) = \frac{1}{1 + e^{-x}}$$

Au cours de l'apprentissage, les calculs relatifs à la mise à jour des poids synaptiques se fait généralement, en utilisant la méthode du gradient, où on distingue deux étapes : la propagation directe et la rétro-propagation.

### III.5.1.1. Les Méthodes d'apprentissage

L'apprentissage automatique désigne l'ensemble des changements dans un système qui lui permettent de réaliser une même tâche, ou des tâches similaires, de manière plus efficace au cours du temps. Il y a deux façons d'apprendre : soit le système se modifie lui-même pour exploiter ses propres connaissances plus efficacement, soit le système acquiert de nouvelles connaissances grâce à des sources externes. En général, on entend par apprentissage la modification automatique des paramètres des systèmes ou plus rarement du nombre et de l'organisation des systèmes, afin d'adapter le traitement effectué à une tâche particulière. On distingue classiquement trois types d'apprentissage en fonction de la nature des informations disponibles et du but recherché : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. Dans la suite, nous désignons par agent, l'entité ou le système soumis à un apprentissage.

#### III.5.1.1.1. Apprentissage supervisé

L'apprentissage supervisé utilise des exemples étiquetés ou classés. Ces étiquettes ou ces classes peuvent être vues comme étant fournies par un professeur ou un superviseur, d'où le nom d'apprentissage supervisé. Le but de l'apprentissage est alors de produire une fonction de classification, appelée hypothèse, permettant de déterminer la classe d'un exemple. Dans ce cas, il est nécessaire de disposer d'un ensemble de couples de données {entrées du réseau ; sorties désirées}, appelées base d'exemples [Personnaz et al., 1990] [Rumelhart et al., 1986]. La différence entre la sortie du réseau et la sortie désirée donne ainsi une mesure d'erreur quantitative sur le calcul effectué par le réseau de neurones. Cette erreur est utilisée pour réaliser l'adaptation.

L'apprentissage supervisé a pour but de déterminer une représentation en intension d'un concept (l'hypothèse) à partir d'un sous ensemble de son extension (les exemples). Il réalise un saut inductif en passant des exemples particuliers à une fonction de classification générale [Degris, 2007]

### III.5.1.1.2. Apprentissage non supervisé

Contrairement à l'apprentissage supervisé, seules les informations en entrée sont fournies au système. Celui-ci doit donc déterminer ses sorties en fonction des similarités détectées entre les différentes entrées, c'est-à-dire en fonction d'une règle auto - organisatrice. Le système est appelé donc, à découvrir les régularités présentes dans ces configurations qui peuvent servir à les diviser en groupes de configurations semblables. Par exemple, le clustering cherche à grouper des exemples de manière à ce que les exemples au sein d'un même groupe se ressemblent suffisamment, et que les exemples de groupes différents soient suffisamment différents. Il peut être utile comme pré traitement à l'apprentissage supervisé ou pour simplifier le stockage ou la communication de données.

### III.5.1.1.3. Apprentissage par renforcement (AR)

C'est un apprentissage pour lequel seule une mesure qualitative de l'erreur est disponible [Zhang, 1997] [Kaelbling et al., 1996]. Dans ce cas, l'agent reçoit des stimuli de son environnement et réagit en choisissant une action adéquate pour son comportement. Sa réaction est alors jugée par rapport à un objectif prédéfini sous forme de note "récompense". L'agent reçoit cette "récompense" et doit l'intégrer pour modifier ses actions futures et parvenir, ainsi à un comportement optimal. Une action conduisant à une note négative sera, à l'avenir, et dans les mêmes conditions, moins utilisée qu'une action notée positivement.

L'AR résulte de deux principes simples, à savoir :

- Si pour un état donné, une action cause immédiatement quelque chose de mauvais, alors le système apprend à ne plus faire cette action lorsqu'il se trouve dans cet état.
- Si dans un état donné, toutes les actions possibles conduisent à quelque chose de mauvais, alors le système apprendra à éviter de se retrouver dans cet état.

L'AR consiste à apprendre pour un agent autonome un comportement à adopter lors de son interaction avec son environnement, afin d'atteindre des objectifs sans aucune intervention extérieure ou d'un professeur.

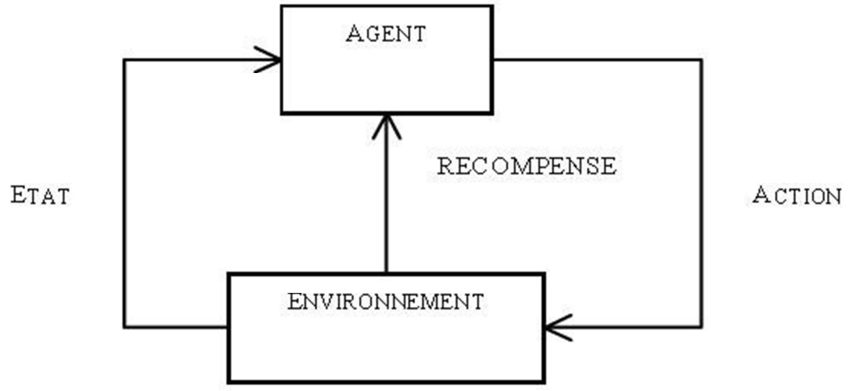



*Figure III-8 : Modèle de l'apprentissage par renforcement [Kaelbling et al., 1996]*

L'objectif de l'apprentissage par renforcement est donc de générer à partir d'expériences (état courant, action, état suivant, récompense) une politique maximisant en moyenne la somme des récompenses au cours du temps. Il s'agit d'une problématique assez naturelle dans les systèmes bio-inspirés

L'idée fondamentale de l'AR est d'améliorer une politique courante suite à une interaction avec l'environnement. Il s'agit d'un renforcement local qui ne nécessite qu'une évaluation locale de la stratégie. Toutefois, la majorité des algorithmes d'apprentissage par renforcement utilisent une fonction de valeur issue de la théorie des processus décisionnels de Markov (PDM) qui constituent le modèle formel de l'apprentissage par renforcement.

L'analyse des différentes méthodes d'apprentissage montre que l'apprentissage par renforcement est celui qui répond le mieux à la problématique du routage dynamique prégnant en compte des contraintes de QoS. En effet, dans ce cas, le réseau est soumis à un trafic imprévisible et arrivant par rafale, et il est donc nécessaire qu'il fasse preuve d'une grande réactivité.

### III.5.2. Routage CPN (Cognitive Packet Network)

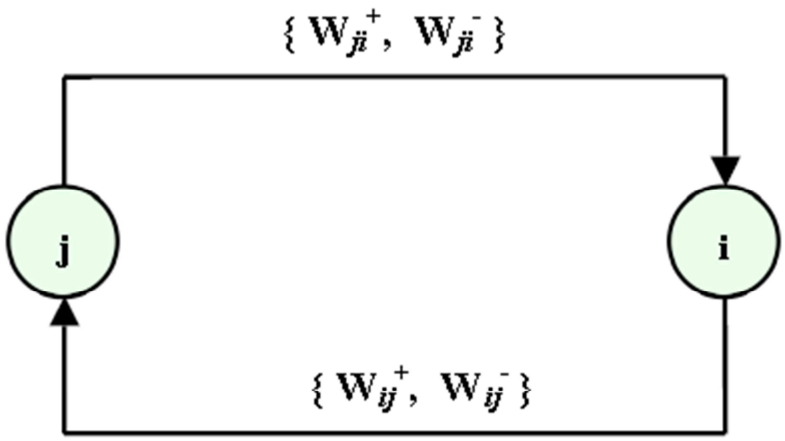
Les *CPN* est un routage base sur l'approche essais/erreurs en utilisant des réseaux de neurones aléatoire *RNN* (Random Neural Network). Dans ce type de réseau l'intelligence se situe au niveau de paquet au lieu de routeurs ou de hôtes. Les réseaux cognitifs considèrent trois types de paquets [Gelenbe et al., 2001]

***Paquets intelligents SP (Smart Packets)*** : ces paquets ont pour tâche l'exploration du réseau. Les *SPs* utilisent l'apprentissage par renforcement implémenté par des approches neuronales afin de trouver une route pour eux-mêmes. Lorsqu'un nœud cherche une route pour une destination donnée, il envoie un paquet *SP* pour explorer les routes possibles. Les paquets *Sps* apprennent à éviter les nœuds et les liens en panne en observant l'état du réseau et les expériences des autres paquets.

***Paquets d'acquittement ACK (Acknowlegement) :*** lorsqu'un paquet *SP* arrive à destination, celle-ci génère un paquet d'acquittement *ACK*, et le renvoie vers la source du paquet *SP* reçu. Le paquet *ACK* emprunte alors le chemin inverse du paquet SP et met à jour les informations de routage des nœuds intermédiaires en utilisant les informations collectées par le paquet *SP*

***Paquets de données :*** ceux-ci sont source-routés ils suivent le chemin qui leur a été calculé par les paquets *SP* et *ACK*.

Le routage des paquets *SPs* est basé sur l'utilisation des *RNN* donc au niveau de chaque nœud en représente par un *RNN* chaque destination ce *RNN* contient autant de cellules qu'il y a de voisins. Dans le cas ou plusieurs paramètres de qualité de service sont à considérer, un *RNN* spécifique est affecté à chaque paramètre [Gelenbe, 1993],[ Gelenbe et al., 2001].




Le neurone qui possède le potentiel le plus fort correspond au lien qui sera choisi pour router les paquets *SPs*. A la réception d'un paquet *ACK*, les poids synaptiques du *RNN* sont mis à jour afin de renforcer ou affaiblir le potentiel du lien selon les informations apportées par le paquet *ACK*. Chaque neurone *i* est caractérisé par son état appelé $q_i$, qui correspond à la probabilité que le i$^{eme}$ neurone soit excité. L'ensemble des états qi doit satisfaire le système d'équations non linéaire suivant :

$$q_i = \frac{\lambda^+(i)}{r(i) + \lambda^-(i)}$$

$$\lambda^+(i) = \sum_j \left(q_j w_{ij}^+ + \lambda_i\right)$$

$$\lambda^-(i) = \sum_j \left(q_j w_{ij}^- + \lambda_i\right)$$

$W_{ji}^+$ Et $W_{ji}^-$ représentent les poids des connexions entrent les neurones tel que :
$W_{ji}^+$ Représente le taux avec lequel le neurone *j* envoie un signal positif au neurone *i*
$W_{ji}^-$ Représente le taux avec lequel le neurone *j* envoie un signal d'inhibition au neurone *i*

Comme il s'agit de minimiser le délai, noté *D*, la récompense *R* prend alors la valeur $D^{-1}$. Cette récompense est mesurée à plusieurs étapes et sera noté $R_n$, n=1,2,…. . La valeur $R_n$ sera utilisée pour calculer le seuil de décision (la moyenne) à l'étape n selon la formule suivante sachant que α est une constante proche de 1 :

$$T_n = \alpha T_{n-1} + (1-\alpha) R_n$$

On suppose que la $n^{ieme}$ décision de routage du paquet *SP* consiste à choisir le nœud voisin noté *j*. soit *D* le délai de bout en bout, apporté par le paquet *ACK* envoyé par le nœud destination, la récompense $R_n$ est considérée comme $D^{-1}$. Ce renforcement sera comparé avec la moyenne obtenue à la $(n-1)^{ieme}$ décision.

- Si $R_n \geq T_{n-1}$ alors les poids positifs entrants vers le neurone *j* seront augmentés significativement tandis que les poids négatifs sortants de *j* seront légèrement augmentés.
- Si $R_n < T_{n-1}$ alors les poids négatifs entrants vers le neurone *j* seront augmentés significativement tandis que les poids positifs sortants de *j* seront légèrement augmentés

Ainsi, pour tout neurone $i \neq j$, la mise à jour des connexions correspond aux règles suivantes :

Si $R_n \geq T_{n-1}$

- $W_{ij}^{+} \leftarrow W_{ij}^{+} + R_n$
- $W_{ik}^{-} \leftarrow W_{ik}^{-} + R_n/(N-2),\ k \neq j$

*Sinon*

- $W_{ik}^{+} \leftarrow W_{ik}^{+} + R_n/(N-2),\ k \neq j$
- $W_{ij}^{-} \leftarrow W_{ij}^{-} + R_n$

Par la suite, ces poids sont réajustés et normalisés afin de permettre le calcul des probabilités qi en résolvant le système d'équations non linéaires décrit par les équations. Le prochain neurone à élire est donc celui ayant la probabilité qi la plus importante.

### III.5.3. Routage avec colonies de fourmis

L'intelligence collective représente une alternative à l'approche classique et centralisée de l'intelligence artificielle. Elle repose sur le concept de la coopération entre un groupe d'agents afin de résoudre un problème donné. Une étude sur la recherche de nourriture chez les fourmis a permis de révéler que ces insectes étaient en mesure d'arriver à leur nourriture par le plus court chemin.

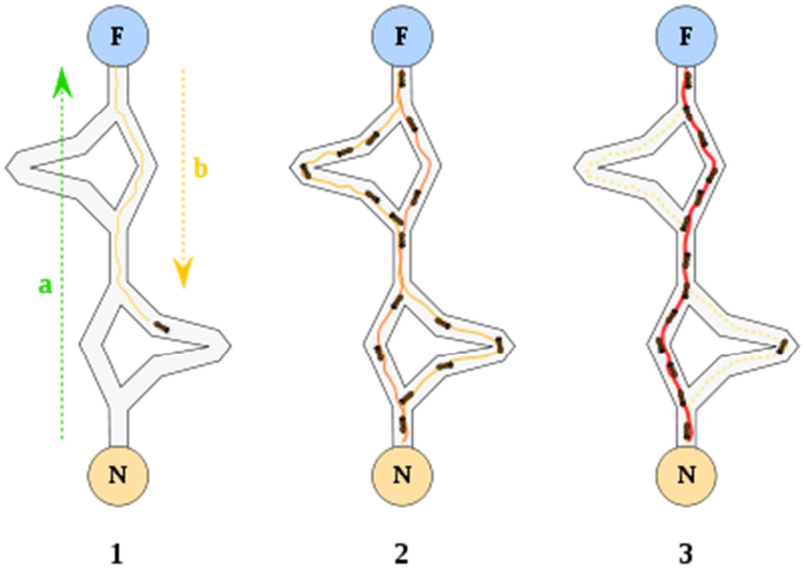



*Figure III-9: Recherche du plus court chemin chez les fourmilles*[Ziani, 2008]

Depuis, il s'est développé un courant de recherche très favorable à l'utilisation des colonies de fourmis dans le routage selon le plus court chemin au sens large.

Il existe, aujourd'hui, une large variante de routage avec colonie de fourmis. L'algorithme le plus connu est appelé *AntNet* [Dicaro et Dorigo, 1998].

L'algorithme AntNet consiste a :

- Périodiquement, chaque nœud *S* envoie un paquet, appelé *Forward* Ant, vers une destination *D* choisie aléatoirement. Ces paquets explorateurs mémorisent sur leur chemin l'identification (adresse) des nœuds intermédiaires ainsi que le temps écoulé depuis le lancement du paquet jusqu'au nœud intermédiaires.
- Arrivant à un nœud intermédiaire, le paquet d'exploration *Forward Ant* ne sera pas routé selon la table de routage. Le prochain saut est sélectionné aléatoirement de façon uniforme sur l'ensemble des voisins qui n'ont pas été sélectionnés. Lorsque tous les voisins sont visités alors tous les voisins seront considérés.
- Quand le paquet d'exploration *Forward Ant* atteint sa destination, il crée un nouveau paquet d'exploration appelé *Backward Ant*, et lui transfère toute sa mémoire. Le paquet *Backward Ant* emprunte le chemin inverse jusqu'à la source. Au niveau de chaque nœud, le *Backward Ant* dépile les informations le concernant, préalablement enregistrées par le paquet *Forward Ant*.
- Arrivant au nœud K venant du nœud F, le paquet Backwar Ant met à jour les informations suivantes:
    - La liste Tripk($\mu_i$,$\sigma^2$), celle-ci correspond à la liste des estimations des valeurs moyennes et des variances de délai entre le nœud *k* et tous les nœuds *i* faisant partie du chemin liant *k* à *d*
    - La table de routage : la probabilité $p_{df}$ associée au nœud *f* est réajustée lorsque la destination est *d*. Si cette probabilité est augmentée, alors il

est nécessaire de réduire tous les $p_{dn}$ pour tout $n$ voisin de $k$ de façon à ce que $\sum n \in N(k)\ p_{km} = 1$ N(k) = {voisinage(k)}

L'information du délai collectée par le paquet *Forward Ant* ne sera pas utilisée pour désigner le meilleur chemin mais servira à la mise à jour des probabilités des liens

***Réajustement des coûts :*** dans un routage avec QoS, le problème crucial, qui se pose souvent, est la mise à jour de la métrique en question. L'enjeu consiste à trouver un compromis entre l'adaptabilité et la stabilité des routes. Nous sommes alors confrontés à deux situations. D'un côté, une métrique fortement dynamique induit de grandes oscillations qui dégradent sérieusement les performances du réseau. De l'autre côté, une métrique trop stable ne répond pas efficacement aux fluctuations de trafic et de congestion.

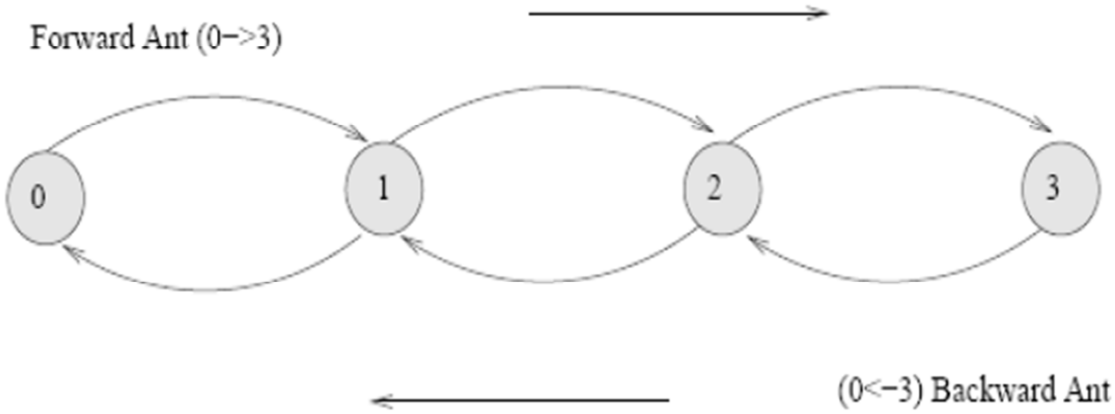



*Figure III-10 : Comportement AntNet[Dicaro et Dorigo, 1998]*

Il est donc judicieux de mettre à jour les métriques de qualité de service, tout en essayant de réduire les grandes variations avec une prise en compte des statistiques à court et à long terme.

Les délais enregistrés par les agents explorateurs correspondent à des mesures instantanées. Cependant, le délai optimal dépend étroitement de l'état du réseau, à savoir si ce dernier est peu chargé ou congestionné. Dans l'algorithme *AntNet*, l'information sur le délai n'est pas utilisée pour connaître les performances réelles du réseau mais plutôt pour indiquer si le délai est bon ou mauvais.

La qualité du délai se décide sur la base de la valeur moyenne enregistrée sur le délai au niveau du nœud. Le délai est utilisé pour calculer le signal de renforcement qui servira à son tour au réajustement des probabilités contenues dans la table de routage.

Soit $D$ le délai observé et $r$ sa valeur moyenne. La valeur $r$ désigne la qualité du délai $D$ observé. Une petite valeur pour $r$ indique que le délai est satisfaisant, et inversement. La valeur $r$ est mise à jour selon l'équation suivante :

$$r' = \begin{cases} \frac{1}{2}\left(\frac{D}{u}\right), & si\ D<2u \\ 1 & sinon \end{cases} \qquad (3.10)$$

Le délai D est considéré comme satisfaisant si la valeur de r ne dépasse pas un seuil prédéfini.

Une stratégie de correction est appliquée à la valeur r afin de tenir compte de la fiabilité du délai D par rapport à sa variance. En effet, la stabilité de la moyenne du délai est considérée, sachant que la moyenne u stable si ($\sigma_u$) <ε , ε<<1.

L'idée de base de la correction consiste à dire que si la moyenne est stable, on récompense davantage le lien s'il a été bon (r < 0.5) en diminuant la valeur de r, et on le pénalise s'il a été mauvais (r > 0.5) en augmentant sa valeur. En revanche, si la moyenne n'est pas stable, la valeur de r ne peut pas être considérée comme fiable.

Dans le but d'éviter de suivre les fluctuations du trafic au risque de les amplifier, il faut procéder donc à la stabilisation. La valeurs-seuil de r doit en effet être augmentée pour exprimer une pénalités- si le délai D est jugé satisfaisant relativisant ainsi sa qualité. Inversement, si le délai D est-ce jugé insatisfaisant, le signal de renforcement r doits être-pour-la-mort réduites pour exprimer la mauvaise qualité du délaisser observés. La fonction de correction f a été définie comme suit :

$$f(\sigma,\mu) = \begin{cases} e^{-\frac{a\sigma}{\mu}} & SI\ \frac{\sigma}{\mu} < \varepsilon \\ 1-e^{-\frac{a'\sigma}{\mu}} & SINON \end{cases} \qquad 3.2$$

$a$ et $a'$ sont des constantes [Dicaro et Dorigo, 1998]. La stratégie de correction de *r′* est résumée comme suite :

$$r' \leftarrow r' + sign(0.5 - r')sign\left(\left(\sigma/\mu\right) - \varepsilon\right) f(\sigma,\mu) \qquad (3.3)$$

L'équation 3.3 exprime une dernière correction appliquée sur r', elle consiste à rendre la valeur r' obtenue dans une échelle plus compressée en appliquant une loi en puissance :

r' ← (r')0.04, le résultat est au final borné sur l'intervalle [0,1].

Ainsi, les transformations du délai D vers la dernière valeur de r' permet d'avoir une sorte d'estimation local du modèle de trafic. La table de routage sera mise à jour en utilisant cette dernière valeur r' obtenue et se fait ainsi :

Un renforcement positif, $r^+$, est attribué au nœud *f* duquel provient l'agent *BackwardAnt*, l'équation 3.4 exprime ce dernier :

$$r^+ = (1 - r')(1 - p_{df}) \qquad 3.4$$

Un renforcement négatif, r-, est attribué à tous les autres voisins, n l'équation 3.5 exprime ce dernier :

$$r^- = -(1 - r')p_{dn} \quad n \in N(k), n \neq f \qquad 3.5$$

$P_{df}$ et $p_{dn}$ sont les dernières probabilités attribuées aux voisins du nœud $k$ pour atteindre la destination d. les valeurs r+ et r- sont proportionnelles à la nouvelle récompense obtenue r' et à la précédente valeur de la probabilité associée au nœud. Grâce auxquelles deux valeurs r+ et r-, l'entrée de la table de routage pour la destination d est-ce mise à jour selon les équations 3.6 et 3.7

- $Pdf \leftarrow Pdf + r+$
- $Pdn \leftarrow Pdn + r-$

La moyenne $\mu$ sera alors mise à jour selon l'équation suivante :

$$\mu \leftarrow \mu + \eta(D - \mu)$$

Le facteur $\eta$ correspond à la quantité des derniers échantillons ayant un réel impact sur la moyenne. La méthode de calcul du signal de renforcement et de mise à jour des probabilités développée ici représente une des nombreuses variantes proposées dans [Dicaro et Dorigo, 1998]

## III.5.4. Les approches de routage par l'apprentissage par renforcement

### III.5.4.1. Q-Routing

Cet algorithme introduit dans [Boyan et Littmann, 1994], est basé sur la technique d'apprentissage par renforcement. Cette dernière est bien adaptée à la problématique du routage étant donné que le modèle d'environnement de chaque routeur est a priori inconnu. Cet algorithme recherche le plus court chemin en termes de temps d'acheminement des paquets jusqu' à leurs destinations.

Pour estimer le temps d'acheminement de bout en bout, un routeur doit être capable de déterminer à tout moment :

- Le ***temps de transmission***, c'est-à-dire le délai mis par un paquet pour atteindre le routeur suivant.
- Le ***temps de traitement*** à l'intérieur du routeur. Ce dernier est crucial et doit être le plus court possible.
- Le ***temps d'attente***, c'est-à-dire le temps que va passer un paquet dans la file d'attente avant d'être émis. Contrairement aux délais précédents, qui sont à peu de chose près constants, les temps d'attente dans les files d'attente évoluent très rapidement en fonction du trafic.

C'est cette dernière estimation que se base le Q-Routing pour répondre aux objectifs de réactivité en détectant rapidement les changements de charge du réseau, l'apparition ou la disparition d'une communication.

Dans le *Q-Routing*, l'algorithme d'apprentissage est basé sur le *Q-learning*. Celui-ci est utilisé pour apprendre une représentation de l'état du réseau en calculant les Q-valeurs permettant d'obtenir une politique de routage optimal et adaptatif. Pour ce faire, chaque routeur doit disposer d'une vision globale de l'état du réseau à tout moment, c'est-à-dire des informations sur tous les routeurs. Cependant, on peut se rendre compte que la simple émission de cette information suffirait à saturer le réseau. Une solution possible consisterait à faire transiter le moins d'informations possibles sur le réseau, en limitant la vision et l'échange d'informations d'un routeur uniquement à ses voisins.

Pour sauvegarder les informations de routage, chaque nœud *x* maintient une table des valeurs *Q(x,y,d)*, appelé *Q-table*, ou *d* est un élément de *V*, l'ensemble de tous les nœuds du réseau. *y* représente un élément de *N(x)*, l'ensemble de tous les voisins du nœud *x*. d'après [Boyan et Littmann, 1994], la valeur *Q(x,y,d)* peut être interprétée comme le meilleur temps estimé par le routeur *x* pour qu'un paquet atteigne la destination *d* en passant par le routeur y, ce temps n'inclut pas le temps d'attente dans la file d'attente de *y*, et le temps que la paquet met pour atteindre *d*, à partir du routeur *y* et en passant par le routeur *ž* voisin de *y*.

Comme le choix d'une route est basé sur les Q-valeurs et que ces dernières ne représentent qu'une estimation, la décision de routage n'est pas forcément optimale. Il est alors nécessaire de mettre à jour le Q-valeurs afin de prendre en compte l'état réel du réseau. [Boyan et Littmann, 1994] a proposé le mécanisme de mise à jour suivant :

Dès qu'un routeur *x* envoie un paquet *p* destiné au nœud *d* via l'un des routeurs voisins *y*, ce dernier envoi un paquet de renforcement (signal de renforcement) au routeur *x*. ce paquet contient l'estimation optimale *Q(y,ž,d)* du temps restant pour arriver à la destination *d* quand le routeur x reçoit cette estimation, il calcule la nouvelle *Q-valeur Q(x,y,d)* comme suit:

$$Q(x,y,d) = Q(x,y,d) + \eta((q_y + \sigma + Q(y,\check{z},d) - Q(x,y,d))$$

Ou $\eta$ représente le pas d'apprentissage (valeur comprise entre 0 et 1) permettant de réguler l'effet de mémorisation.

La méthode utilisée dans le Q-Routing, pour la mise à jour des Q-valeurs, est connue sous le nom d'exploration avancée (forward exploration), ou à chaque saut du paquet p(s;d), une Q-valeurs est mise à jour

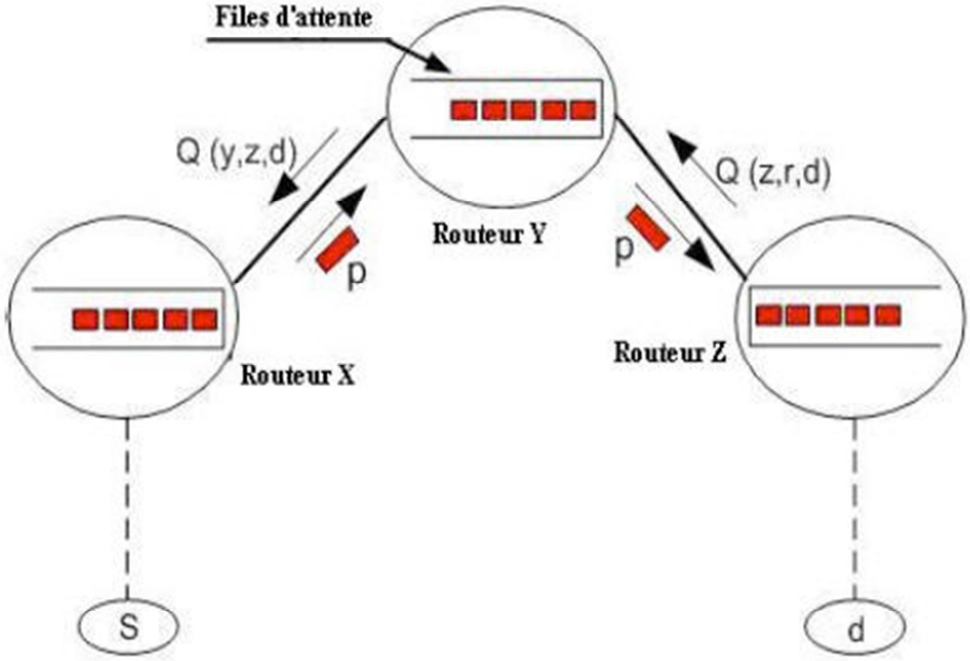



*Figure III-11 : mise à jour des Q-valeurs dans Q-routing [Hoceini, 2004]*

Les performances de cette politique de routage dépendent principalement des Q-valeurs estimées qui doivent être les plus représentatives possibles de l'état courant du réseau. Par conséquent, ces valeurs doivent être mise à jour de façon continue. Cependant, pour celles qui ne sont mises à jour que rarement, les décisions de routage sont peu fiables, ce pourquoi d'autres techniques d'exploration ont vu le jour [Mellouk, 2007].

### III.5.4.2. K SP Q-routing

Une amélioration de l'algorithme *Q-routing* a été proposé dans [Amirat et al., 2005], appelé *K-Shortest Path Q-Routing*. Ce dernier est basé sur la technique du routage multi-chemin combiné avec l'algorithme *Q-Routing*. L'espace d'exploitation est réduit aux *K* meilleurs chemins, minimisant le coût total des liens. L'algorithme proposé ne nécessite qu'un espace mémoire proportionnelle au produit du nombre d'adresses de destination par le nombre *K* des plus courts chemins. L'algorithme de recherche des *K* plus courts chemins est basé sur l'algorithme de Dijkstra généralisé auquel un mécanisme de suppression de boucles a été adjoint. Le chemin optimal correspond à celui dont le temps d'acheminement moyen est le plus court.

Le mécanisme d'exploration utilisé pour la mise à jour de Q-valeurs repose sur une méthode hybride associant la technique de l'exploration avancée à chaque fois qu'un paquet de donnée est échangé entre les routeurs, et celle de l'exploration probabiliste pour l'exploration des K-1 chemins restants. La charge des trafics est ensuite répartie entre plusieurs chemins selon le calcule d'une distribution probabiliste sur les K meilleurs chemins assurant une meilleure exploration des chemins potentiellement intéressants pour

le routage par une adaptation en temps réel des probabilités de distribution à la charge du trafic dans le réseau [Amirat et al., 2004].

### III.5.4.3. Q-neural routing

L'algorithme de routage Q-neural constitue une amélioration de l'algorithme Q-routing ou l'estimation et la mise à jour des Q-valeurs sont effectuées à l'aide d'un réseau de neurones. Ces derniers permettre à l'apprennent d'incorporer de divers paramètres tels que la taille de la file d'attente locale et l'heure du jour à son évaluation de distance. En effet, un RN permet la modélisation de la fonction complexe avec une bonne précision avec un apprentissage distinctif et en tenir compte du contexte du réseau. D'ailleurs, il peut être employé pour prévoir le trafic non stationnaire ou irrégulier. Dans cette approche, l'objectif est de réduire au minimum le délai moyen de livraison d'un paquet. Par conséquent, le signal de renforcement qui est choisi correspond délai prévue pour transférer un paquet à son destination. Typiquement, le délai de livraison de paquet inclut trois variables: le délai de transmission d'un paquet, le temps de traitement dans le routeur et le temps d'attente dans la file d'attente.

## III.6. Conclusion

La fonction de routage construit les bases de données (appelées tables de routage) indiquant l'interface de sortie pour une destination. Cette information servira pour le reliage.

Dans un routage dynamique, les routes sont calculées automatiquement via un algorithme et un (des) protocoles de routage dynamique, le protocole permet l'échange d'informations entre les machines et l'algorithme utilise ces informations afin de calculer la meilleure route possible (dans le cas où on en a plusieurs, bien sûr). .

L'objectif premier d'un protocole de routage, Quoiqu'il, est d'acheminer de manière optimisée les paquets de données vers leurs destinations. Certains de ces protocoles sont dotés de mécanismes pertinents d'adaptation aux changements de topologies. Ces mécanismes leur offrent la capacité de reconstruire rapidement les routes en cas de changement de topologie tout en générant le moins de charge de contrôle supplémentaire possible. En effet, en cas de congestion, par conséquent une dégradation des performances et en l'absence de l'information sur l'état du réseau, les nœuds se trouvent très vite dans l'incapacité de réagir face à ces dégradations. L'approche du routage adaptatif est une des réponses possibles à cette problématique. Celui-ci est en effet basé sur un principe qui consiste à chercher un chemin en collection sur l'état des liens par le biais d'un mécanisme d'exploration.

# Chapitre IV

- *ACERP un Protocole de Routage basé QoE*

# IV. ACERP un Protocole de Routage basé QoE

## IV.1. Présentation

### IV.1.1. Introduction

La flexibilité indéniable proposée par l'optimisation par colonie de fourmis a permis aux chercheurs d'adapter ces algorithmes à des données dynamiques. A l'heure actuelle, où les réseaux de communication et le nombre d'utilisateurs augmentent exponentiellement, il est nécessaire de gérer le réseau en temps réel pour éviter les encombrements ou la saturation des lignes.

Le réseau est en grande partie lié à la notion de routage. En effet un réseau informatique peut être vu comme un graphe dont les arêtes sont les différentes lignes de communication (qui elles-mêmes peuvent être vues comme des sous réseaux plus petits) et les sommets représentent les routeurs. Lorsqu'un routeur reçoit un paquet de la machine A qui doit arriver à la machine B, celui-ci a pour rôle d'envoyer le paquet vers le prochain nœud de son parcours. Les routeurs étant souvent interconnectés, plusieurs routes sont possibles pour un même paquet, il doit donc essayer de choisir un trajet minimisant le délai d'acheminement en fonction de l'encombrement du réseau.

On utilise principalement deux grandeurs pour exprimer la qualité d'un réseau :

- La bande passante (bit/s) qui représente le taux de transfert (soit la place sur une ligne).
- Le délai moyen (s) qui est le temps d'acheminement moyen d'un paquet.

Un routage de bonne qualité permet d'augmenter la bande passante quand le réseau est chargé et de baisser le délai de transfert d'un paquet quand la charge est faible. La charge du réseau variant continuellement et de façon brutale, les algorithmes d'optimisation par colonie de fourmis semblent parfaitement indiqués pour les gestions des tables de routage. Nous détaillerons dans ce chapitre le fonctionnement de l'algorithme *ACERP (Ant Colonie EvaluateRouting Protocol)*

Pour implémenter un protocole, l'exécution commence toujours par une description formelle du protocole, par exemple sous forme de pseudo-code, d'organigramme, des diagrammes d'ordre ou d'une représentation graphique différente. Dans ce qui suit, le format de paquet est spécifié et les algorithmes de protocole sont mis en application.

### IV.1.2. SMA et Agent Mobile

Un agent est une entité logicielle ou matérielle, à laquelle est attribuée une certaine mission qu'elle est capable d'accomplir de manière autonome, disposant d'une

connaissance partielle de ce qui l'entoure (son environnement), et agissant par délégation pour le compte d'une personne ou d'une organisation.[ Perret, 1997]

Un agent est un système informatique situé dans un environnement, et qui agit d'une façon autonome et flexible pour atteindre les objectifs pour lesquels il a été conçu. [Jennings et Wooldridge, 1995]

Un agent possède les caractéristiques suivantes :

- L'agent est une entité qui agit par délégation. Il doit respecter la stratégie de son producteur vis à vis des choix qu'il est amené à faire, afin que celui-ci soit responsable des tâches effectuées par son agent.
- L'agent est une entité autonome qui dispose de son propre environnement.
- L'agent dispose d'une connaissance, même partielle, de son environnement courant. Ceci lui permet de prendre les décisions appropriées.
- L'agent est caractérisé par un comportement flexible.

Un système multi-agent (SMA) est un ensemble d'agents situés dans un certain environnement et interagissant selon une certaine organisation.

Un agent mobile est défini comme un élément autonome œuvrant généralement au nom d'un utilisateur ou d'une application. Il possède une activité interne avec ses propres ressources, il peut aussi accéder aux ressources de l'hôte d'accueil, et communiquer avec d'autres agents afin de réaliser la tâche pour laquelle il a été créé. Il a la capacité de se déplacer de site en site en ayant conscience de son déplacement. Il est aussi capable de percevoir son environnement, de s'adapter à des conditions particulières et réagir à des événements préalablement décrits [Leriche et Arcangeli, 2006].

Un agent mobile possède les propriétés suivantes :

***Baisse de la communication réseau*** : Le déplacement des agents mobiles permet de réduire de manière significative, les communications distantes entre les clients et les serveurs.

***Exécution asynchrone*** : Avec les agents mobiles, un client peut déléguer les interactions avec le service sans maintenir une connexion de bout en bout. Avec cette possibilité d'une exécution asynchrone, le client peut demander un service, se déplace puis vient récupérer les résultats plus tard. [Gray et al., 2000]

***Amélioration du temps d'exécution*** : L'optimisation des phases de traitement se produit à deux niveaux. Premièrement, en localisant les données et le code sur un même site, on supprime les phases de dialogue entre le client et le serveur qui sont perturbées par des temps de latence dû aux communications réseaux. Ensuite, le déplacement du code va permettre de déléguer les calculs sur des machines serveurs (telles que des supercalculateurs) qui sont généralement plus puissantes qu'une machine cliente. Cela est particulièrement vrai dans l'informatique nomade où la miniaturisation s'accompagne d'une perte de puissance significative

***Personnalisation*** : Les agents s'adaptent plus facilement aux besoins du client que des serveurs. Il n'est pas nécessaire d'installer une procédure spécifique au niveau d'un serveur pour manipuler la demande spécifique de service d'un client. Au lieu

de cela des agents mobiles peuvent être adaptés aux besoins de l'utilisateur. Ceci a comme conséquence une dynamicité en plus.

### IV.1.3. Description du protocole de routage ACERP

Notre idée de tenir compte de la QoE de bout en bout consiste à développer des mécanismes adaptatifs qui peuvent rechercher l'information de leur environnement (QoE) et s'adapter aux actions initiées. Le choix d'action devrait être exécuté en réponse à la rétroaction d'utilisateurs, en d'autres termes la rétroaction de QoE. Concrètement, le système intègre la mesure de QoE dans un système évolutionnaire de routage afin d'améliorer la perception d'utilisateur basée sur l'algorithme de colonie de fourmis pour choisir les « meilleurs chemins optimaux de QoE » .Ainsi , le processus de routage est construit selon le maintien de la meilleure perception d'utilisateur.

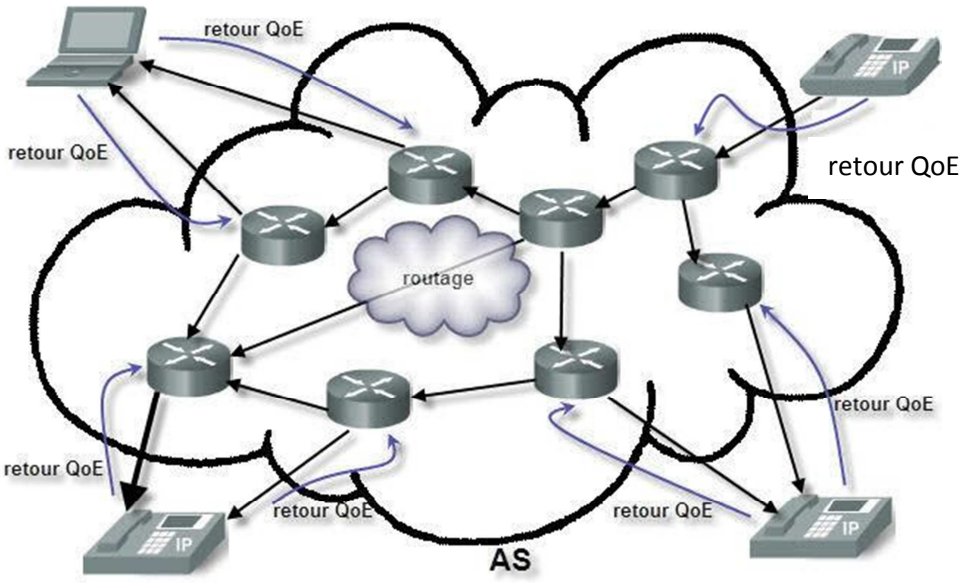



*Figure IV-1 : Intégration de la mesure de QoE dans le système de routage*

#### IV.1.3.1. Implémentation des nouveaux types de paquet pour ACERP

| P_type | P_src | P_dest |
|---|---|---|
| P_len | Num_seq | |
| tmp_d | mem_ta | |

**P_memoire**

| N_adr | …… | N_adr |
|---|---|---|
| Tmp_dec | | Tmp_dec |

***P_type*** **:** type de fourmis (ForwordAnt ou BackwordAnt)
***P_src*** **:** le nœud source
***P_dest*** **:** le nœud destination
***P_len*** **:** la longueur du paquet ant
***Num_seq*** **:** le numéro de séquence du paquet
***Tmp_d*** **:** le temps de départ
***Mem_ta*** **:** la taille de la mémoire qui contient la liste des nœuds du chemin traversé
***P_memoire:*** liste des nœuds du chemin traversé avec le temps de passage pour chaque nœud

### IV.1.3.2. Implémentation de la table de routage

La table de routage pour un nœud donné est une matrice ou chaque ligne correspond à une destination donnée et chaque colonne représente un nœud voisin.

| | | | |
|---|---|---|---|
| **ND1** | **NV1** | …… | **NVn** |
| | **Val_ph1** | | **Val_phn** |
| | … | … … | |
| | … | … … | |
| **NDm** | **NV1** | …… | **NVn** |
| | **Val_ph1** | | **Val_phn** |

***NDi*** **:** nœud destination
***NVj*** **:** nœud voisin
***Val_ph$_k$*** **:** la valeur de la quantité de phéromone

### IV.1.3.3. Diagramme de fonctionnement du protocole ACERP

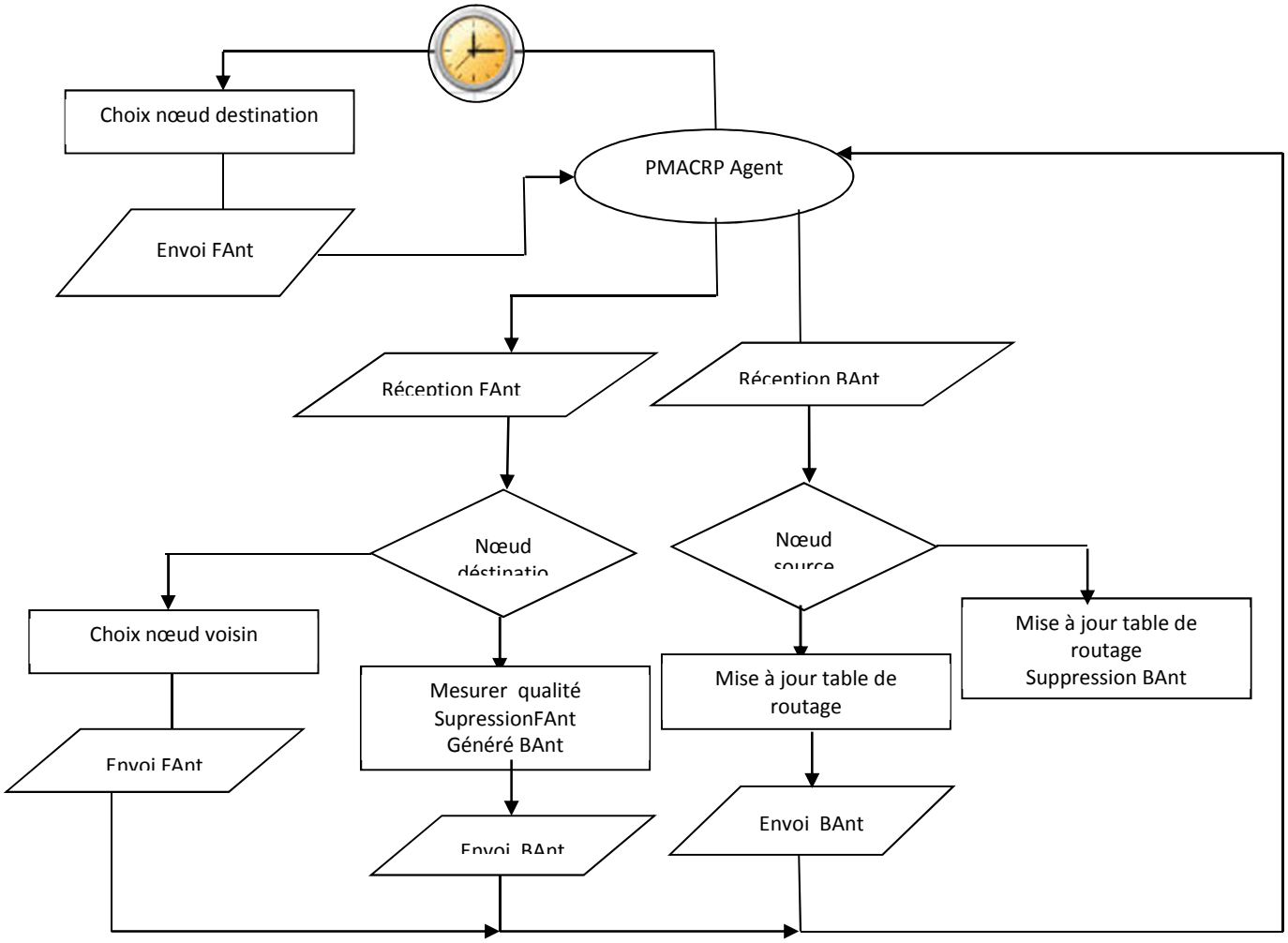



*Figure IV-2 : Diagramme de fonctionnement du protocole ACERP*

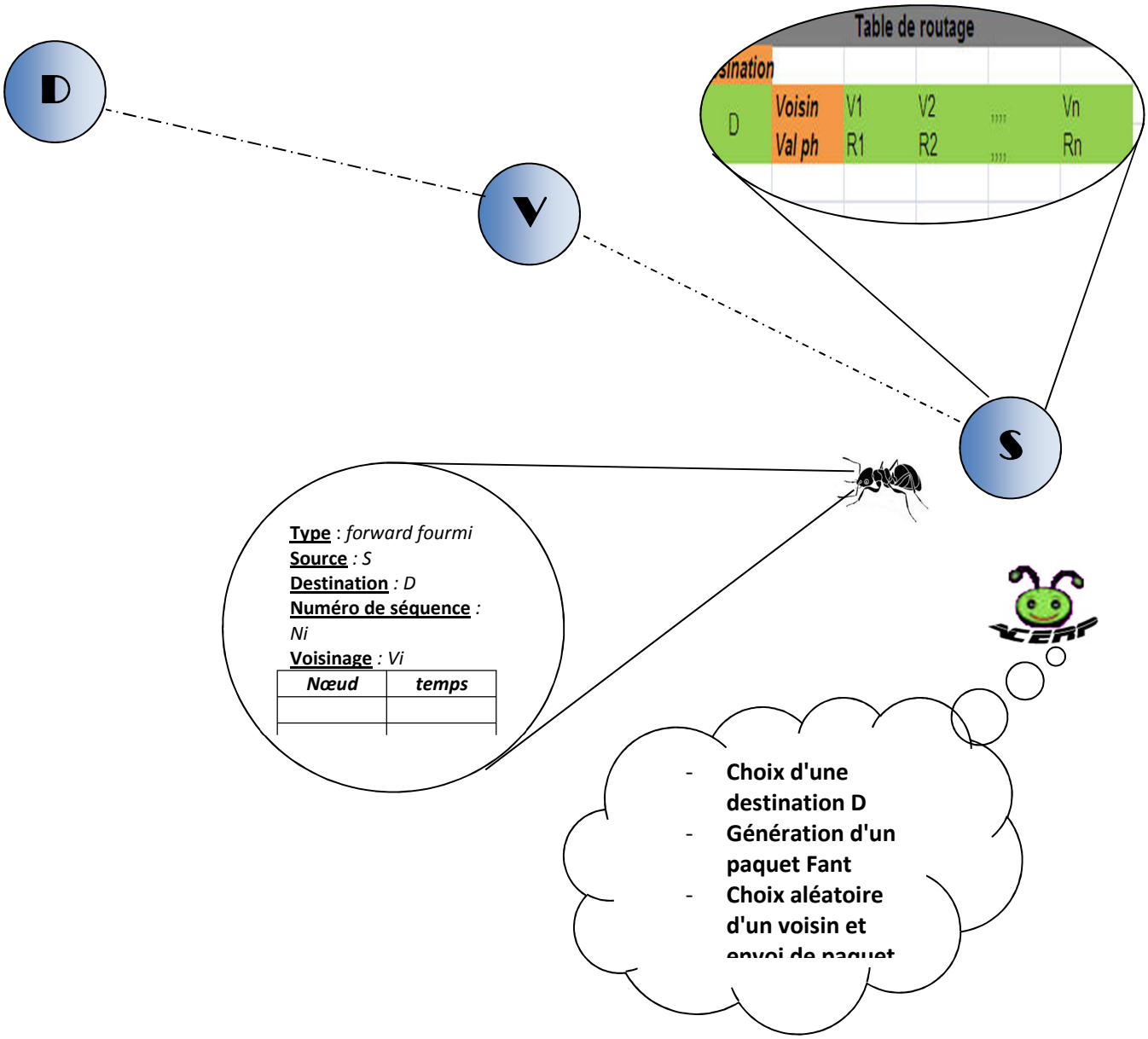



*Figure IV-3 : Génération d'un paquet de type FANT par le nœud source.*

Dans la figure IV-3 l'agent protocole *ACERP* du nœud source choisit et d'une manière aléatoire et périodique une destination de la topologie et génère un paquet *ForwardAnt* et l'envoie vers cette destination via un voisin qui ramène à cette destination.

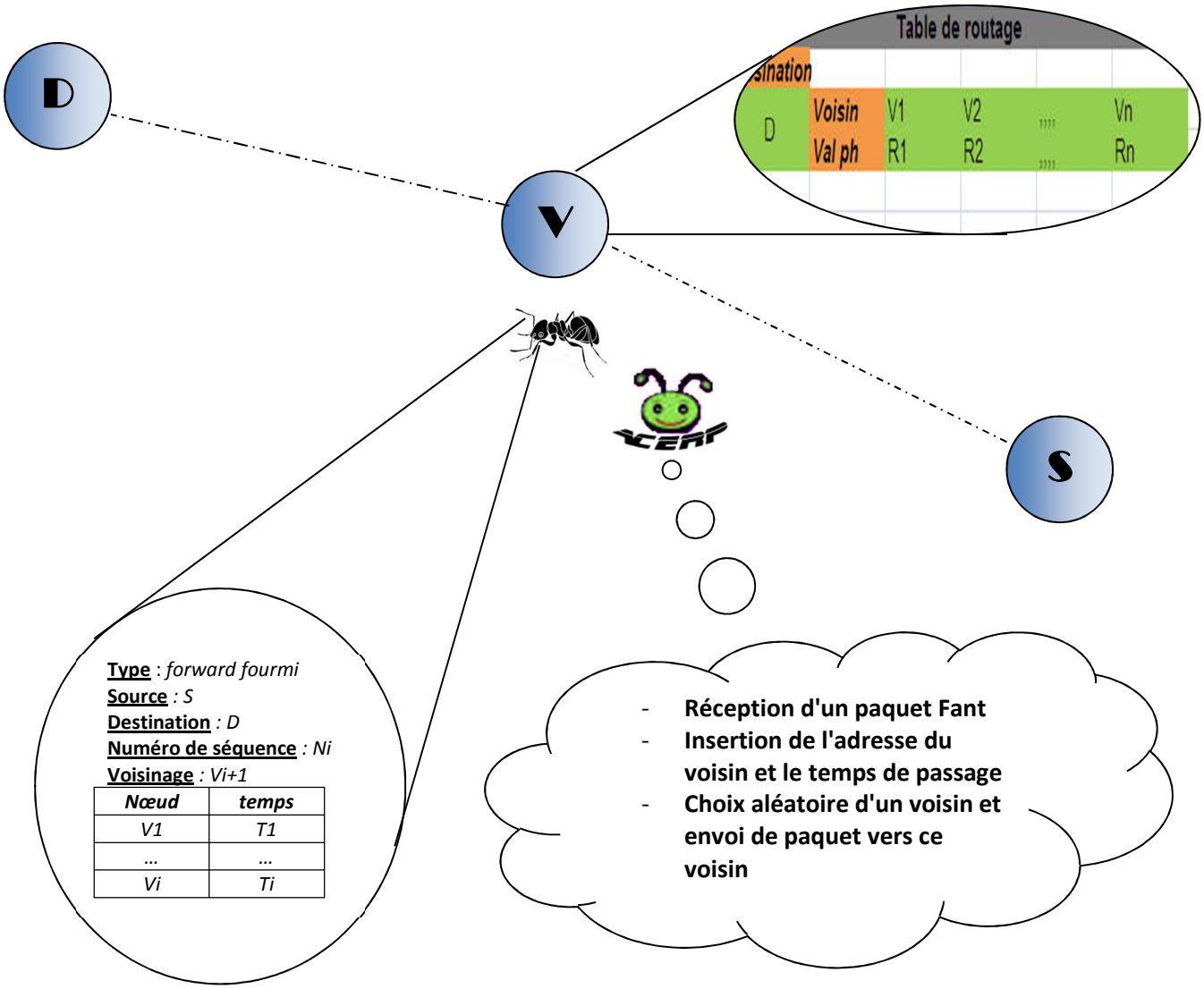



*Figure IV-4 : Réception d'un paquet FANT par un nœud voisin et intégration du temps de passage*

La Figure IV-4 le nœud Vi représente un voisin direct ou indirect d'un nœud source ou d'un voisin le protocole ACERP de ce nœud lorsque il reçoit un paquet *ForwardAnt* il insère par son rôle l'adresse du nœud et le temps de passage et envoie par la suite ce paquet vers un autre nœud voisin qui ramène à la destination le choix du nœud voisin est aléatoire.

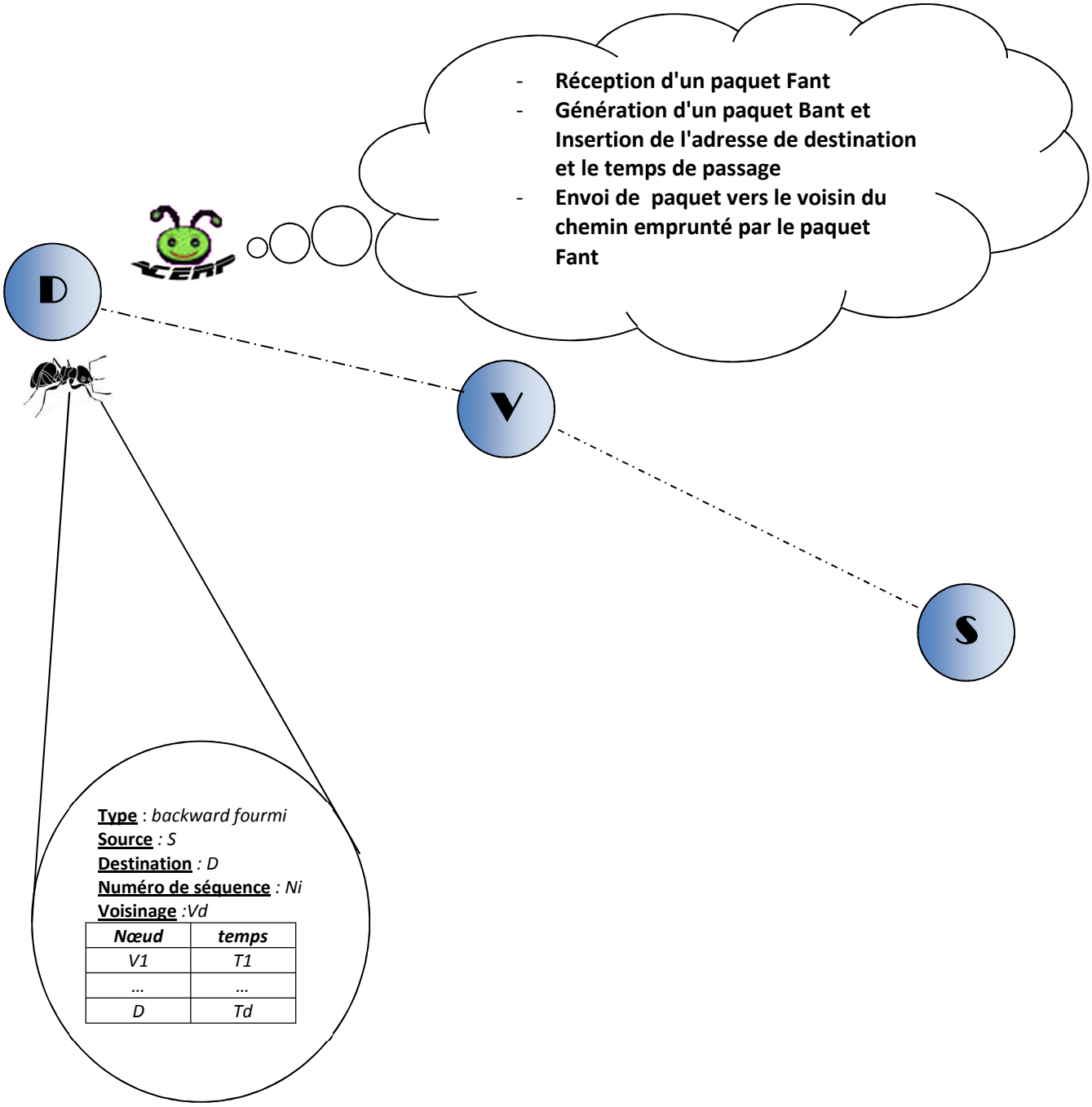



*Figure IV-5 : Réception d'un paquet FANT par le nœud destination et génération d'un paquet BANT*

Dans la Figure IV-5 le paquet *ForwardAnt* arrive finalement à destination donc l'agent ACERP intègre l'adresse de destination et génère un autre paquet *BackwardAnt* avec les mêmes informations et distrait également le paquet *ForwardAnt*. ACERP envoie par la suite le paquet *BackwardAnt* sur le même chemin traverse par le paquet FAnt.

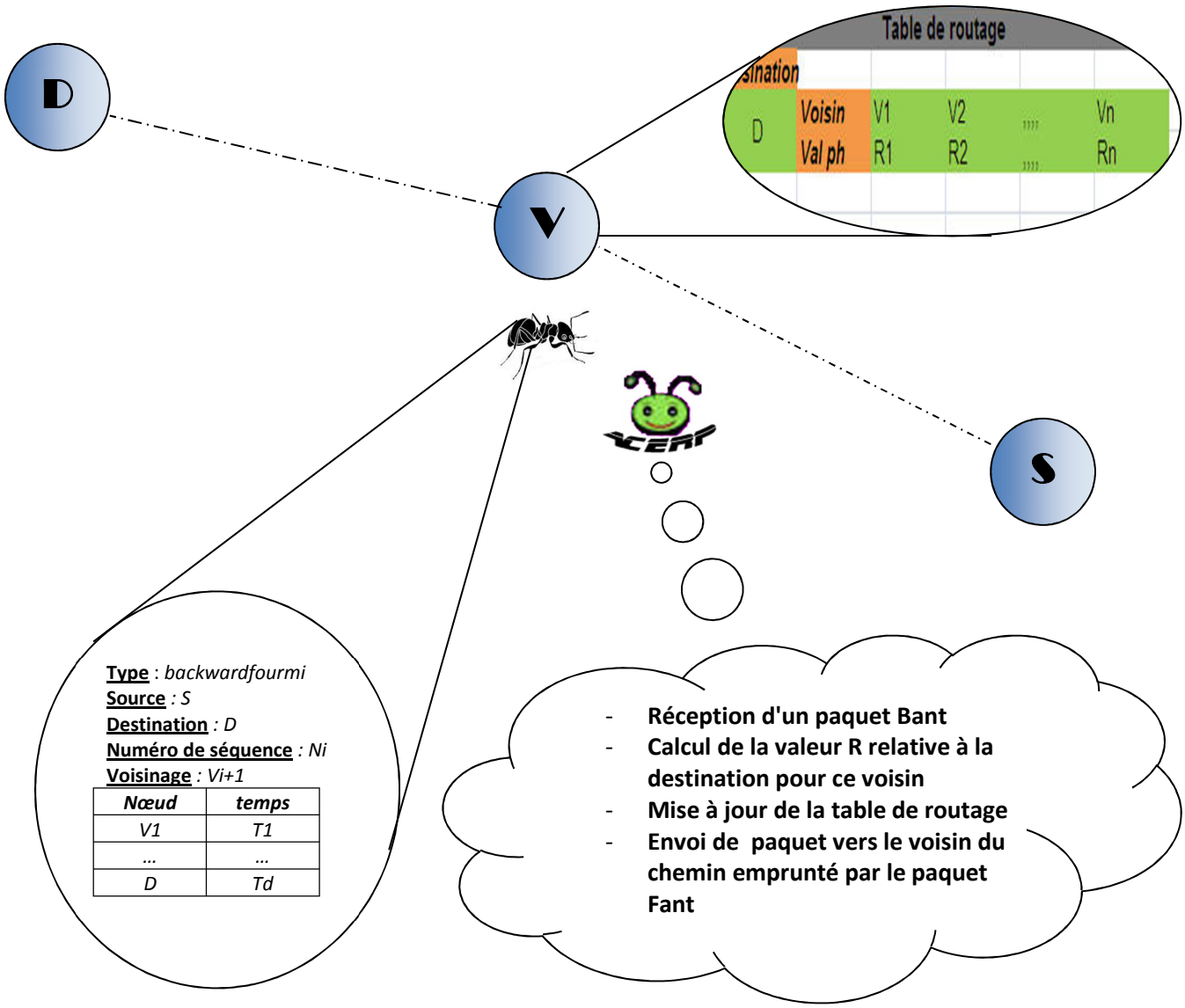



*Figure IV-6 : Réception d'un paquet BANT par un nœud voisin*

L'agent ACERP dans la figure IV-6 pour un nœud voisin du chemin traversé par *FAnt* lorsqu'il reçoit un paquet *BackwardAnt* calcule la valeur d'évaluation R en utilisant les paramètres importes par le paquet *BAnt* dans le temps de passage est inclut et met à jour la table de routage du nœud Vi dans l'entrée correspondante à la destination D et pour la colonne Vi+1 qui représente la route à la destination via le voisinage Vi+1.

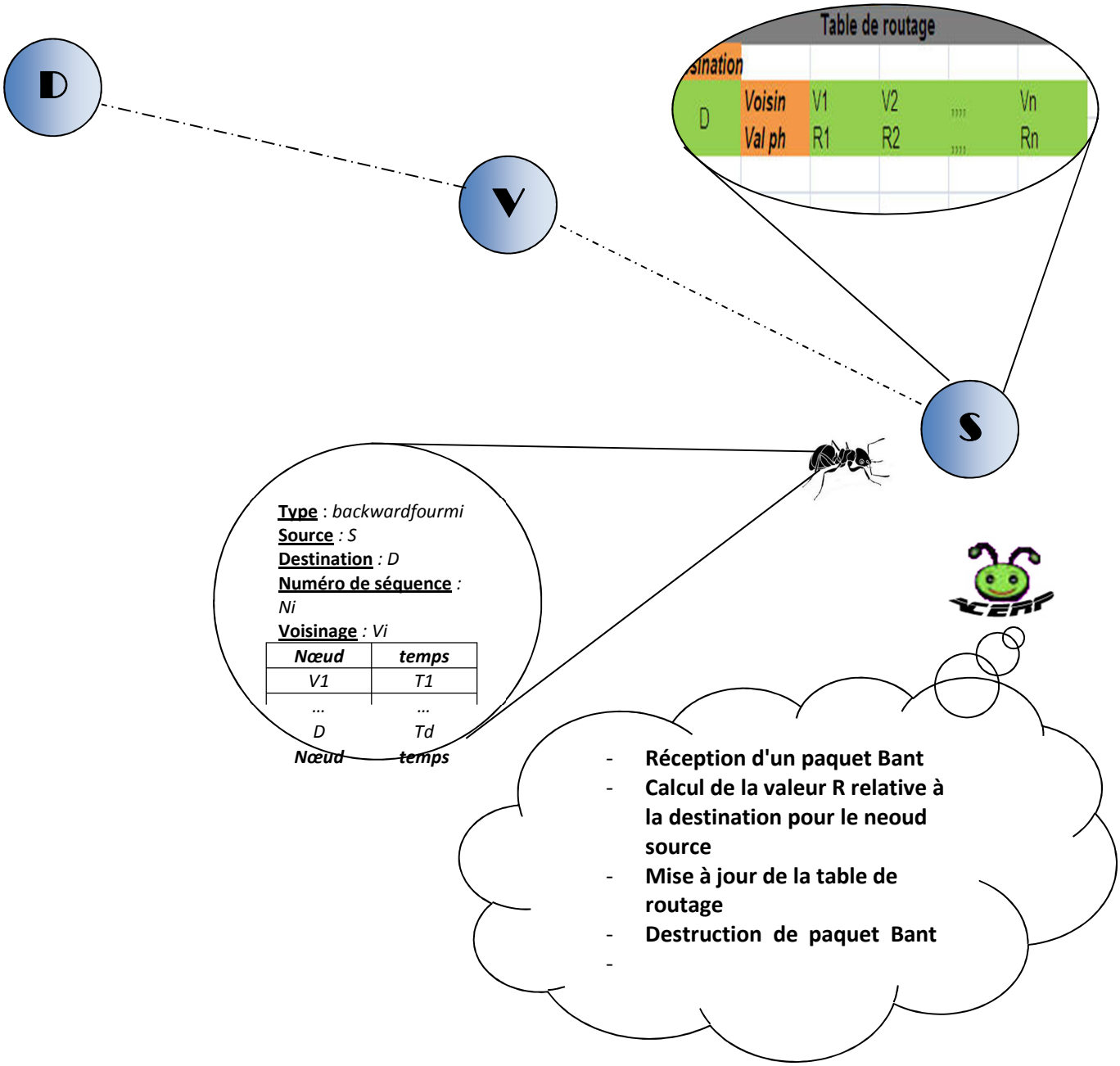



*Figure IV-7 : réception d'un paquet BANT par le nœud source.*

L'agent ACERP du nœud source lorsqu'il reçoit le paquet *BAnt*, calcule la valeur d'évaluation R et met à jour la table de routage et enfin détruit le paquet

### IV.1.3.4. Description de la méthode d'évaluation

La recommandation décrit un modèle de calcul connu sous le nom modèle E, qui s'est avéré utile en tant qu'outil de planification de transmission pour évaluer les effets conjugués des variations des différents paramètres de transmission altérant la qualité en conversation téléphonique. Ce modèle de calcul peut être employé, par exemple, par ceux qui planifient la transmission pour leur permettre de s'assurer que les usagers sont satisfaits de la qualité de la transmission de bout en bout, tout en évitant une ingénierie des réseaux trop complexe. Il faut souligner que le résultat brut fourni par ce modèle est le "facteur d'évaluation " R. qui peut être transformé pour obtenir des estimations de l'opinion des usagers [ITU, 2005].

Le modèle E est un algorithme basé sur 20 paramètres liés aux, facteurs terminaux, facteurs de l'environnement et factures liés au réseau.

$$R=R0-Is-Id-Ie-eff+A$$

R0 : rapport signal/bruit de base R0=f(Nc,SLR,Ps,Ds,RLR,Pr,LSTR)
Is : facteur de dégradation simultanée Is=f(R0,SLR,RLR,STMR,TELR,qdu)
Id : facteur de dégradation due au temps de propagation
Id=f(T,Tr,Ta,RLR,STMR,TELR,WEPL)
Ie : facteur de dégradation due à l'équipement Ie-eff=f(Ie,Bpl,Ppl)
A : facteur d'avantage

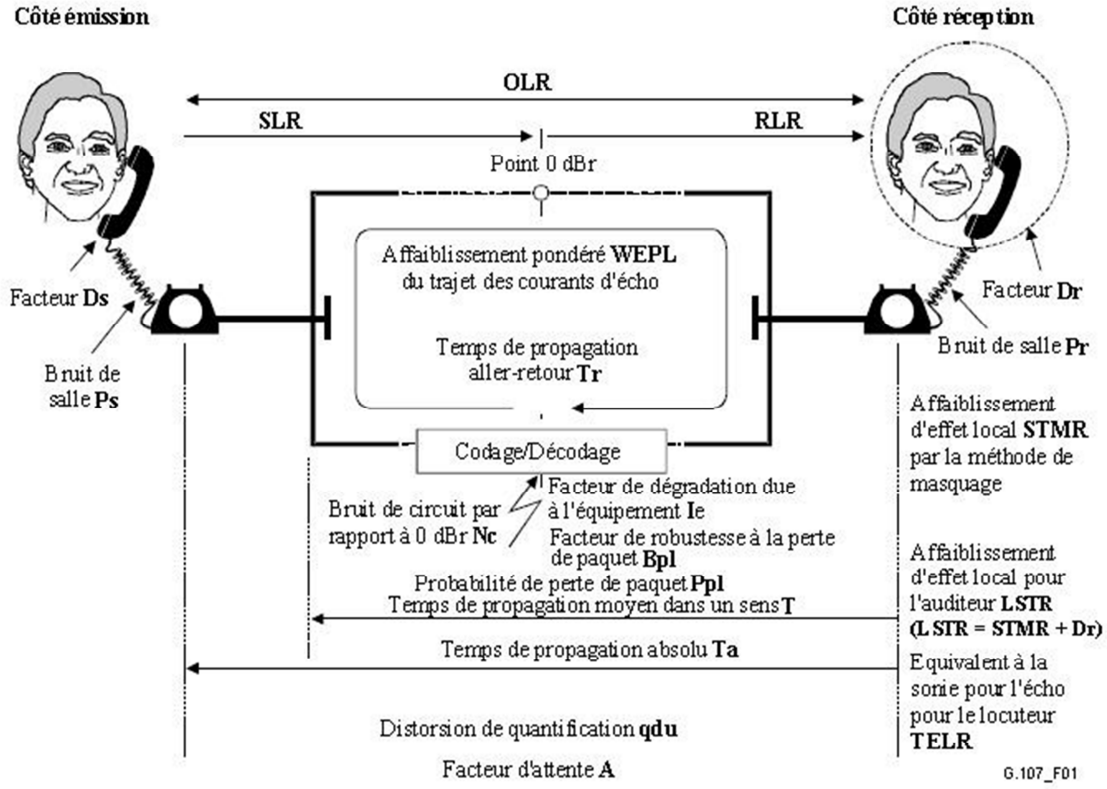



*Figure VI-8 : Connexion de référence du modèle E[ITU, 2005]*

| Paramètre | Abr. | Unité | Valeur par défaut | Intervalle permis |
|---|---|---|---|---|
| Equivalent pour la sonie à l'émission(*sendloudness rating*) | SLR | dB | +8 | 0..+8 |
| Equivalent pour la sonie à la réception(*receiveloudness rating*) | RLR | dB | +2 | -5..+14 |
| Affaiblissement d'effet local par laméthode de masquage (*sidetonemaskingrating*) | STMR | dB | 15 | 10..20 |
| Affaiblissement d'effet local pourl'auditeur (*listenersidetone rating*) | LSTR | dB | 18 | 13..23 |
| Valeur D du côté émission téléphonique | Ds | dB | 3 | -3..+3 |
| Valeur D du côté réception téléphonique | Dr | - | 3 | -3..+3 |
| Equivalent pour la sonie de l'écho pour lelocuteur (*talkerecholoudness rating*) | TELR | - | 65 | 5..65 |
| Affaiblissement pondéré du trajet d'écho(*weightedechopathloss*) | WEPL | dB | 110 | 5..110 |
| Temps de propagation moyen dans un sensdu trajet d'écho | T | Ms | 0 | 0..500 |
| Temps de propagation aller-retour dansune boucle à 4 fils | Tr | Ms | 0 | 0..1000 |
| Temps de propagation absolu surconnexion exempte d'écho | Ta | Ms | 0 | 0..500 |
| Nombre d'unités de distorsion dequantification (*quantizationdistortionunits*) | Qdu | - | 1 | 1..14 |
| Facteur de dégradation due à l'équipement | Ie | - | 0 | 0..40 |
| Facteur de robustesse à la perte de paquet | Bpl | - | 1 | 1..40 |
| Probabilité de perte de paquet aléatoire | Ppl | % | 0 | 0..20 |
| Rapport lié aux rafales | BurstR | - | 1 | 1..2 |
| Bruit de circuit par rapport au point deréférence 0 dBr | Nc | DBm0p | -70 | -80..-40 |
| Seuil de bruit du côté réception | Nfor | DBmp | -34 | - |
| Bruit de salle du côté émission | Ps | dB(A) | 35 | 35-85 |
| Bruit de salle du côté réception | Pr | dB(A) | 35 | 35-85 |
| Facteur d'avantage | A | - | 0 | 0..20 |

*Tableau IV-1 : valeurs par défaut et intervalles permis pour les paramètres[ITU, 2005]*

## IV.2. Simulation

La simulation est un processus qui consiste à concevoir un modèle du système (réel) étudié, mener des expérimentations sur ce modèle (et non pas des calculs), Interpréter les observations fournies par le déroulement du modèle et formuler des décisions relatives au système. Le but peut être de comprendre le comportement dynamique du système, de comparer des configurations, d'évaluer différentes stratégies de pilotage, d'évaluer et d'optimiser des performances.

On distingue deux catégories de simulation sur ordinateur, simulations discrètes et simulations continues. Dans les simulations continues, les quantités sont représentées par

des variables continues, alors que dans les systèmes des simulations discrètes les quantités d'intérêt sont représentées par des valeurs à des variables discrètes [Braun et al., 2008].

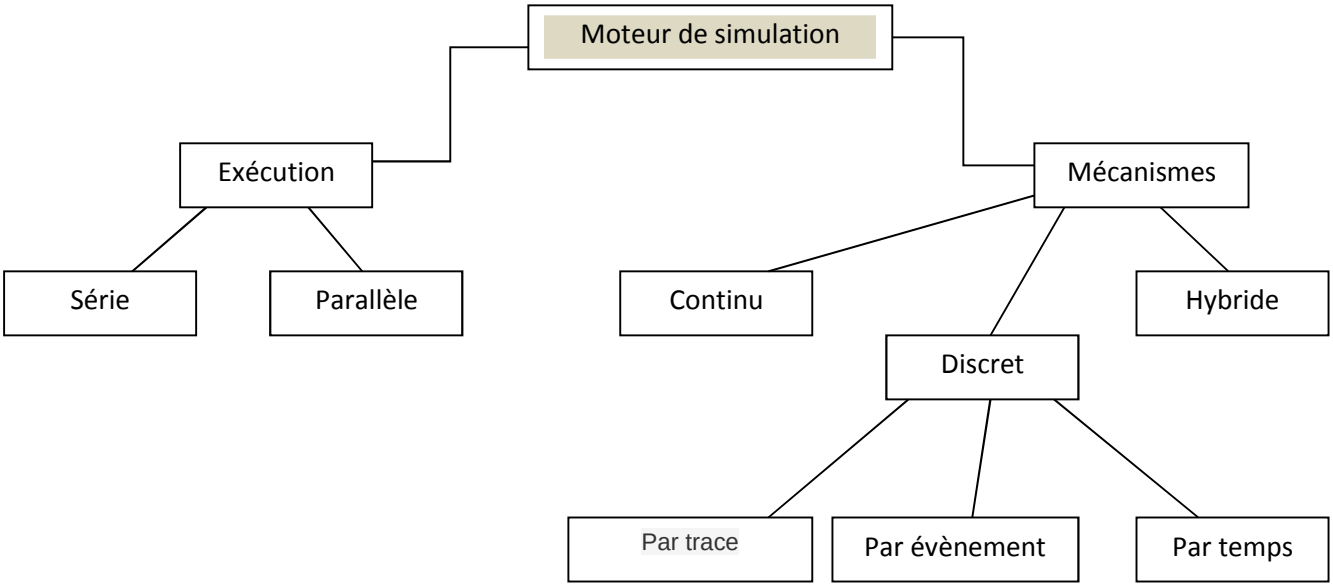



*Figure IV-9 : Types de simulation [Braun et al., 2008]*

La dynamique d'une simulation discrète peut être considérée comme une séquence d'événements à des moments discrets. Comme un troisième type de simulation, la simulation de Monte Carlo est liée à la simulation à événements discrets, qui est couramment utilisée pour modéliser des systèmes stochastiques. "Méthodes de Monte Carlo sont une classe largement utilisée des algorithmes de calcul pour simuler le comportement de divers systèmes physiques et mathématiques, et pour d'autres calculs", Monte Carlo les simulateurs font habituellement appel à des nombres aléatoires pour modéliser des pièces non déterministes afin de simuler le système. Ils calculent leurs résultats en répétant un échantillonnage aléatoire un grand nombre de fois. La loi des grands nombres est souvent utilisée comme justification pour l'exactitude des résultats.

Des simulations continues sont mises en application comme ensemble d'équations. Le programme de simulation résout toutes les équations périodiquement par évaluation numérique.

Pour des simulations à grande échelle, les approches hybrides sont émergées en tant que solutions viables, où des parts importantes sont mises en application en tant que simulations discrètes et d'autres parts moins importantes en tant que simulations continues. Les stratégies de simulation hybride sauf des quantités significatives de ressources informatiques sont comparées aux simulations discrètes.

Les différents types de simulation peuvent être exécutés en parallèle ou en série. L'exécution parallèle des simulations fournit des temps d'exécution plus courts, mais augmente la complexité de l'exécution du modèle de simulation, aussi bien que sa vérification.

La simulation d'un réseau informatique en général est une tâche difficile et pas facile à manipuler. Le principal problème est le fait que le réseau informatique est composé de nombreux nœuds tels que les routeurs, les commutateurs et les hôtes, ce qui rend la partie modélisation du processus de simulation d'une tâche triviale. Il ya certaines décisions à faire au début du processus de simulation:

- Quels sont les faits que la simulation doit montrer ou prouver?
- Quels sont les éléments importants qui devraient être étudiés?
- Quel simulateur fournit les meilleures possibilités pour modéliser le système?
- Quelle est la simulation assez précise pour pouvoir utiliser les résultats pour la recherche?

Il y a des différentes méthodes pour simuler des réseaux informatiques,

### IV.2.1. Les Simulateurs Réseaux

Les simulateurs de réseaux couramment utilisés soutiennent multiples protocoles et fournissent donc des avantages considérables tels que:

• une meilleure validation des protocoles existants

• une infrastructure pour le développement de nouveaux protocoles

• la comparaison plus facile des résultats

**GloMoSim** (Global Mobile Simulator) est un environnement de simulation à grande échelle pour les réseaux sans fil et filaires. Il a été conçu en utilisant la capacité de la simulation parallèle fournie par PARSEC (Parallel Simulation Environment for Complex Systems).GloMoSim est construit en utilisant une approche basée sur les couches qui est semblable à l'architecture à sept couches de la norme OSI.

**QualNet** est un outil de simulation qui simule réseaux filaires et sans fil en mode paquet de communication. QualNet est un produit commercial qui dérive de GloMoSim publié en 2000 par SNT (Scalable Network Technologies).

Les principales différences entre QualNet et GloMoSim sont:

- QualNet est basé sur C++; GloMoSim est basé sur PARSEC C (un langage C basée simulation parallèle).
- QualNet est un produit commercial; GloMoSim est distribué sous une licence open source académique.
- QualNet est maintenu par la SNT; GloMoSim est maintenu par laboratoire de UCLA Parallel Computing.

**OMNet++** est un simulateur d'évènement basé sur le langage C++, destiné principalement à simuler les protocoles réseau et les systèmes distribués. Il est totalement programmable, paramétrable et modulaire. C'est une application open source et sous licence GNU, développée par Andras Varga, chercheur à l'université de

Budapest. OMNet++ est destiné avant tout à un usage académique et est l'intermédiaire entre des logiciels de simulation comme NS, destiné principalement à la recherche et OPNET qui est une alternative commerciale de OMNet++.Le fonctionnement d'OMNet++ repose entièrement sur l'utilisation de modules qui communiquent entre eux par le biais de messages. Ces modules sont organisés hiérarchiquement. Les modules de base sont appelés les modules simples. Ceux-ci sont regroupés en modules composés. Ces modules peuvent eux-mêmes être regroupés en modules composés. Le nombre de niveau hiérarchique n'est pas limité. Ils sont codés en C++ et sont des instances du type de base module. L'architecture est construite de telle sorte que les modules simples sont à la fois les émetteurs et destinataires des messages. Les modules composés se contentent de relayer les messages aux modules simples de façon transparente. On peut attribuer différents paramètres aux connexions reliant les modules: des délais de propagation, des débits de données, des taux d'erreur, etc.

**SSFNet** est un modèle open-source en Java pour la simulation des réseaux informatiques. Il comprend des modèles pour plusieurs protocoles (par exemple, IP, TCP, UDP, BGP4, OSPF) et les composants réseau tels que des hôtes, des routeurs et des liens.En outre, plusieurs classes de soutien permettent la modélisation et la simulation réalistes des scénarios multi-Protocole et de multi-Domain d'Internet.

**OPNET** (Optimum Network Performance) est un outil de simulation de réseaux très puissant et très complet. Basé sur une interface graphique intuitive, son utilisation et sa prise en main est relativement aisée. OPNET dispose de trois niveaux hiérarchiques imbriqués : *le* ***network domain***, ***le node domain*** et ***le process domain***. ***Network domain*** est le niveau le plus élevé de la hiérarchie d'OPNET. Il permet de définir la topologie du réseau en y installant des routeurs, des hôtes, des équipements tels que des switchs, reliés entre eux par des liens. Chaque entité de communication (appelée nœud) est entièrement configurable et est définie par son modèle. ***Le Nodedomain*** permet de définir la constitution des nœuds (routeurs, stations de travail, hub, …). Le modèle est défini à l'aide de blocs appelés modules. C'est au niveau ***de processdomain*** que l'on définit le rôle de chaque module programmable. OPNET fournit des mécanismes permettant à tous les processus créés à l'intérieur d'un processdomain de communiquer entre eux, via un bloc de mémoire partagée, ou l'ordonnancement d'interruptions logicielles. Le rôle d'un module est déterminé par son process model, que l'on décrit sous forme d'une machine à états finis (FSM). Les actions a effectuer sont décrites en langage C, et OPNET fournit une bibliothèque de plus de 400 fonctions propriétaires spécifiques à l'usage des réseaux (création, envoi et réception de paquets, extraction de valeurs contenues dans les différents champs d'une entête…). OPNET permet de gérer deux autres types d'objets relatifs aux réseaux : les liens et les formats de paquets.

**NS-2 (Network Simulator 2)** [Braun et al., 2008]est un outil logiciel de simulation de réseaux informatiques, développé dans le cadre du projet VINT,ce dernier est un projet en cours de développement avec la collaboration de plusieurs acteurs (USC/ISI, Xerox parc, LBNL et UCB) dans l'objectif principal de construire un simulateur multi-protocole pour faciliter l'étude de l'interaction entre les protocoles et le comportement d'un réseau à différentes échelles. Il est principalement bâti avec les idées de la conception par objets, de réutilisabilité du code et de modularité. Il est devenu aujourd'hui un standard de

référence en ce domaine. C'est un logiciel dans le domaine public disponible sur l'Internet. Son utilisation est gratuite. Le logiciel est exécutable tant sous Unix que sous Windows.

Le simulateur NS actuel est particulièrement bien adapté aux réseaux à commutation de paquets et à la réalisation de simulations de petite taille. Il contient les fonctionnalités nécessaires à l'étude des algorithmes de routage unipoint ou multipoint, des protocoles de transport, de session, de réservation, des services intégrés, des protocoles d'application comme HTTP. De plus le simulateur possède déjà une palette de systèmes de transmission (couche 1 de l'architecture TCP/IP), d'ordonnanceurs et de politiques de gestion de files d'attente pour effectuer des études de contrôle de congestion. La liste des principaux composants actuellement disponibles dans NS par catégorie est:

| | |
|---|---|
| Application | Web, ftp, telnet, générateur de trafic (CBR, ...) |
| Transport | TCP, UDP, RTP, SRM |
| Routage | Statique, dynamique (vecteur distance) et routage multipoint (DVMRP, PIM) |
| Gestion de file d'attente | RED, DropTail, Tokenbucket |
| Discipline de service | CBQ, SFQ, DRR, Fairqueueing |
| Système de transmission | CSMA/CD, CSMA/CA, lien point à point |

*Tableau IV-2 : Les composants disponibles dans NS-2 [Braun et al., 2008]*

NS est devenu l'outil de référence pour les chercheurs du domaine. Ils peuvent ainsi partager leurs efforts et échanger leurs résultats de simulations. Cette façon de faire se concrétise aujourd'hui par l'envoi dans certaines listes de diffusion électronique de scripts de simulations NS pour illustrer les points de vue

### IV.2.2. Le concept du langage NS-2

Plutôt que d'utiliser un langage de programmation unique qui définit une simulation monolithique, NS-2 intègre un modèle de programmation plus modulaire en utilisant un concept de deux langages. Le langage C++ implémente le noyau de simulation et les parties centrales de primitives hautes performances. Le langage de script objet OTcl exprime la définition, la configuration et le contrôle de la simulation. Il représente la colle entre les composants du réseau individuels implémentés en C++

#### IV.2.2.1. Structure hiérarchique de NS-2

La simulation est configurée, contrôlée et exploitée par l'interface fournie par la classe ***simulator*** du langage **OTcl**. Cette classe principale fournit des méthodes pour la création et la mise en place de la topologie, de choisir la méthode d'ordonnancement, et la gestion des simulations.

Des éléments de la topologie sont créés avec les composants de base (***Node*** ,***Link***). ***Node*** modélise un routeur ou un hôte qui peut transmettre et recevoir des paquets par ses interfaces. L'interface, l'adresse et le mappage de port sont effectués par

des objets classificateurs. Un nœud unicast (nœud de défaut) a un classificateur d'adresse qui fait le routage unicast et est un classificateur de port. Un nœud multicast a un classificateur d'adresse qui sépare les paquets de multicast et d'unicast. D'ailleurs, il inclut un classificateur multicast qui exécute le routage multicast. Un autre composant de base est *l'agent* qui représente le composant de point final d'un nœud, où des paquets d'application sont construits ou consommés. Le lien (***Link***) représente un autre composant principal dans NS-2 il est caractérisé par le délai et la bande passante. Les liens sont établis comme des objets de concession.

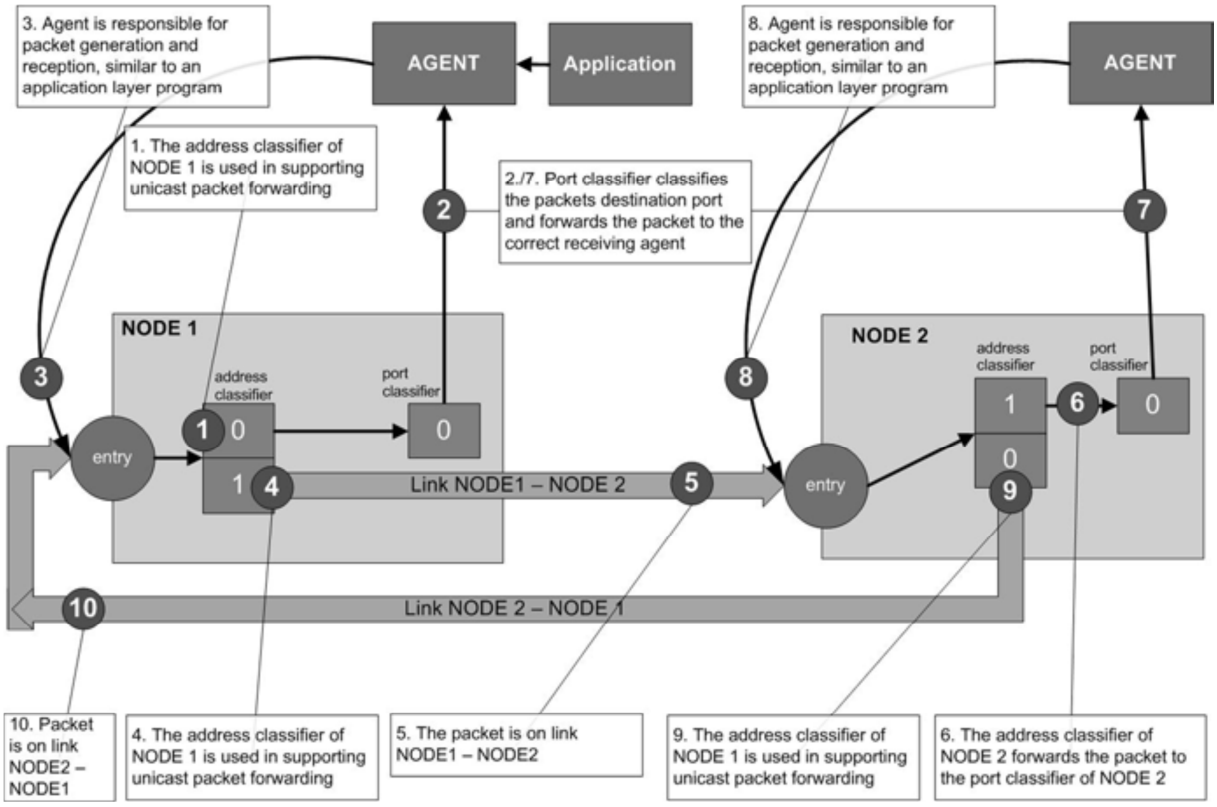



*Figure IV-10 : Cycle d'un paquet sur nœud et lien dans NS-2 [Braun et al., 2008]*

### IV.2.3. Description déclarative de l'agent ACERP

La description déclarative orienté objet de l'agent protocole ACERP c'est une forme de classe composée principalement de 8 méthodes et d'une propriété (table_routage) Figure IV-8.

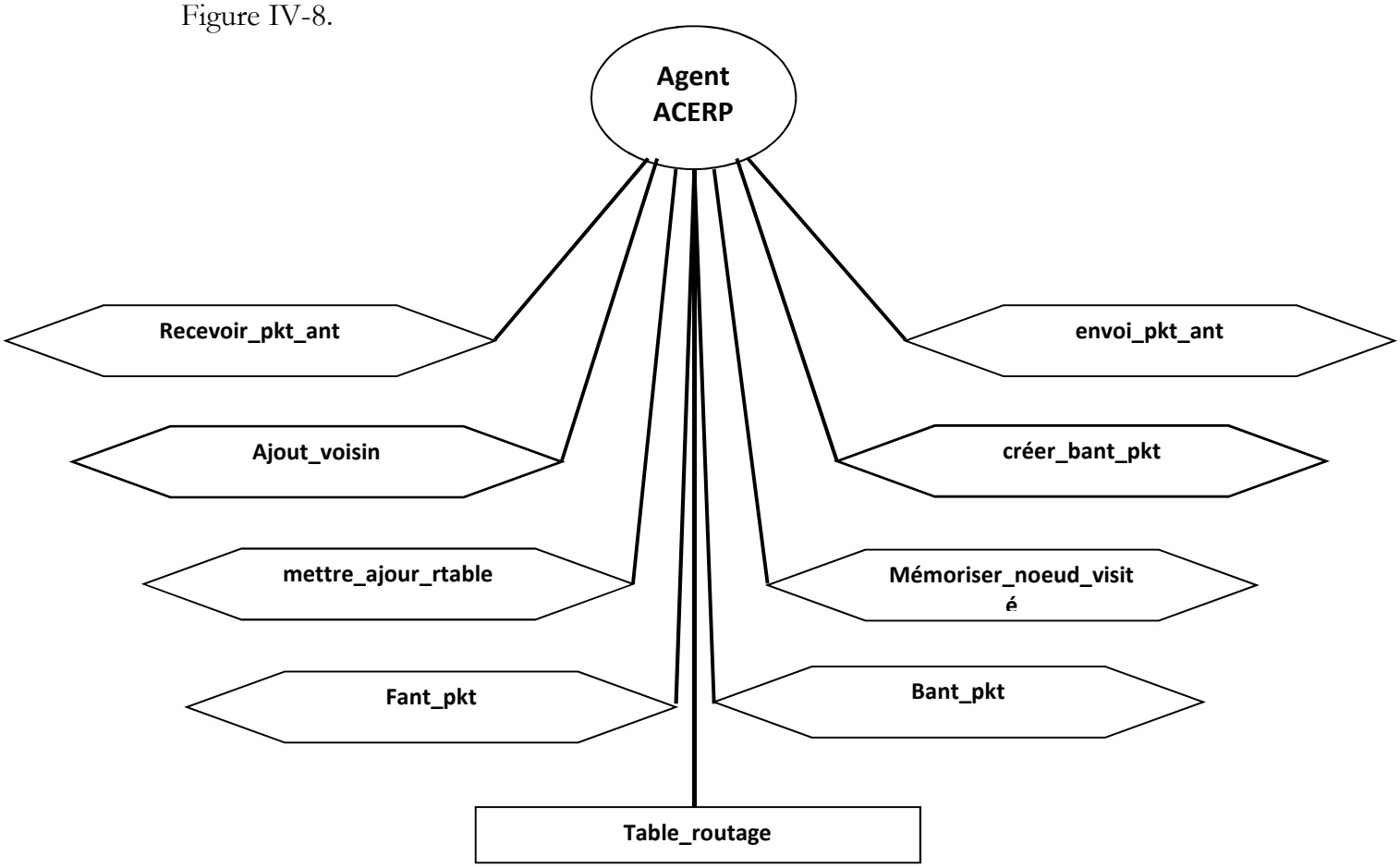



*Figure IV-11 : description déclarative de l'agent protocole ACERP*

***Recevoir_pkt_ant*** : cette méthode permet pour un agent *ACERP* lorsqu'il reçoit un paquet de faire un traitement selon le cas , alors si ce paquet est de type *FANT* et si le nœud concerné est le nœud destination donc ajouter les informations(*adresse du nœud et temps de passage*) à la mémoire de paquet et créer par la suite un paquet *BANT* en appellent la méthode *crée_bant_pkt* en faire un transfert de toutes les informations du paquet *FANT* vers le paquet *BANT*, sinon si ce nœud n'est pas le nœud destination alors ajouter les informations(*adresse du nœud et temps de passage*) à la mémoire de paquet faire envoyer ce paquet vers un nœud voisin qui amène bien sûr au nœud destination et ça en appelant la méthode *Fant_pkt*. Autrement si le paquet est de type *BANT* et le nœud c'est le nœud source dans ce cas mettre a jour la table de routage utilisant la méthode *mettre_ajour_rtable* et détruire le paquet, sinon si le nœud est un nœud autre que le nœud source alors toujours mettre a jour la table de routage et envoyer le paquet vers le nœud suivant naturellement en lisant l'adresse de la mémoire du paquet lui-même en appelant la méthode *Bant_pkt*.

***Envoi_pkt_ant***: pour chaque agent ACERP qui réside dans un nœud est associé un temporisateur qui permet à chaque période de générér un paquet FANT et faire l'envoyer on choisissant une destination et faire évaluer un chemin on utilisant cette méthode, elle permet l'initialisation de tous les paramètres du paquet et par la suite en suivant la procédure présente en appelant la méthode *Recevoir_pkt_ant*.

***Créer_bant_pkt*** : lorsqu'un paquet FANT arrive à destination cette méthode est appelée, elle crée un paquet BANT on copie toutes les informations du paquet FANT on change la direction vers le nœud source cette fois le chemin à suivre est extrait àpartir du chemin mémorisé par le paquet FANT

***Mémoiriser_noeud_visité*** : utilisant cette méthode l'agent protocole ACERP lorsqu'il reçoit un paquet de type FANT il teste avant tout si Ilya une boucle il élimine la boucle sinon il mémorise les informations de ce nœud (adresse et temps de passage) dans la mémoire de ce paquet.

***Ajout_voisin*** : on suppose que tous les liens entre les nœuds soient de type duplex link, alors cette méthode permet de définir pour un lien donné les deux voisins associés.

***Table_routage***: cette propriété qui fait partie du protocole ACERP définit pour un nœud la déclaration de la table de routage associée ainsi que les procédures et fonctions permettant de manipuler cette table de routage. La déclaration est basée sur la description définie précédemment dans la *figure IV-22*

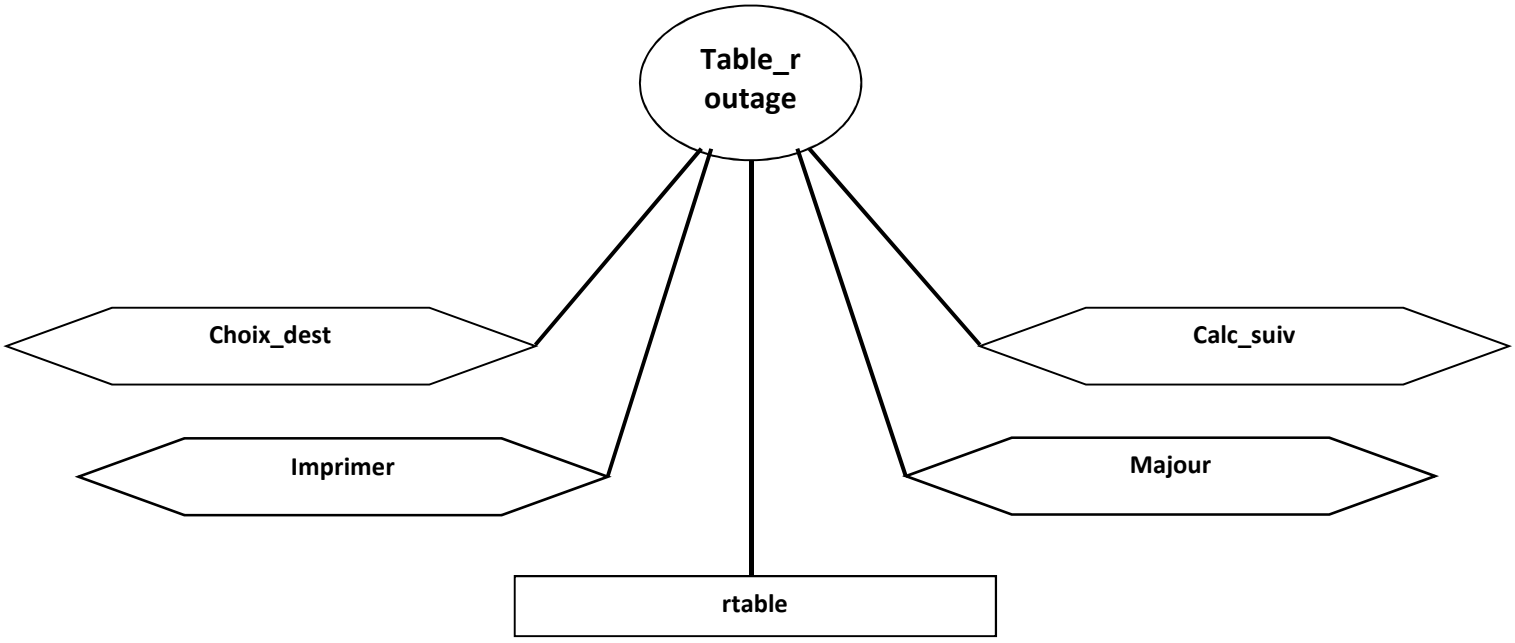



*Figure IV-12 : description de la propriété table_routage de l'agent protocole ACERP*

***Choix_dest :*** cette méthode permet pour chaque période de choisir un nœud destination et évaluer par la suite un chemin aléatoire qui amène à ce nœud à partir du nœud source. Le choix du nœud destination se fait d'une manière cyclique ou aléatoire.

***Calc_suiv*** : c'est une méthode qui permet à partir des nœuds destination, source et nœud actuel de calculer le pas suivant c'est-à-dire le nœud suivant.

***Majour*** : c'est une procédure qui permet de mettre à jour la table de routage suivant les paramètres (destination, nœud voisin et la valeur de la phéromone).

***Imprimer :*** cette méthode permet juste d'imprimer la table de routage c'est-à-dire : source, destination, voisin, valeur phéromone.

***Rtable*** : c'est une propriété définie sous forme d'une table ou chaque ligne représente un nœud destination et chaque colonne est une structure composée de deux champs un qui représente un nœud voisin et l'autre la valeur de la quantité de phéromone

***mettre_ajour_rtable:*** pour le chemin de routeur du paquet *BANT* qui est identique au chemin d'arriver l'agent *ACERP* lit pour un nœud les paramètres d'évaluation à partir de la mémoire de paquet et calcule la valeur *R* d'évaluation et selon le cas mettre à jour sa table de routage (noter bien qu'on va mettre à jour la table de routage pour chaque nœud traversé par le paquet *BANT* ainsi que le nœud source bien sûr)

***Fant_pkt***: cette méthode permet d'envoyer un paquet *FANT* vers le voisin suivant.

***Bant_pkt*** : la méthode *Bant_pkt* permet pour l'agent *ACERP* d'envoyer un paquet de type *BANT* vers le nœud suivant l'adresse du nœud est obtenue à partir de la mémoire du paquet.

## IV.3. Cadres expérimentaux

Le fonctionnement d'un réseau de communication est régi par de nombreux composants qui peuvent agir l'un sur l'autre de manière non linéaire et imprévisible. Par conséquent, le choix d'un banc d'essai significatif pour comparer des algorithmes concurrents n'est pas une tâche facile.

La topologie peut être définie sur la base d'un vrai exemple net à la main, pour analyser mieux l'influence des dispositifs topologiques importants (comme le diamètre, le degré de connectivité, etc.).

Les nombres dans des cercles représentent les adresses des nœuds, alors que les nombres sur des liens sont des délais de propagation en milliseconde. Chaque bord dans le graphique représente une paire de liens orientés.

Appliquons une simulation à base du protocole ACERP sur une topologie avec des liens égaux en coût et choisissons à titre d'exemple le nœud 5 avec un nombre de voisins égal 3. Les résultats qui représentent les valeurs de la table de routage sont pris périodiquement avec un intervalle T=1 minute

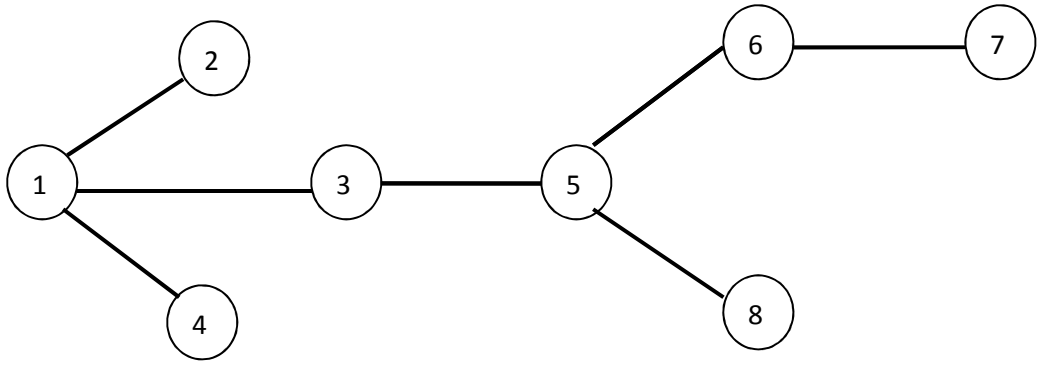



*Figure IV-13 : Topologie d'expérimentation N°1*

| T0 | | Voisins | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 0.00 | 0.00 | 0.00 |
| | **2** | 0.00 | 0.00 | 0.00 |
| | **3** | 0.00 | 0.00 | 0.00 |
| | **4** | 0.00 | 0.00 | 0.00 |
| | **6** | 0.00 | 0.00 | 0.00 |
| | **7** | 0.00 | 0.00 | 0.00 |
| | **8** | 0.00 | 0.00 | 0.00 |

| T1 | | Voisins | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 91.45 | 0.00 | 0.00 |
| | **2** | 90.87 | 0.00 | 0.00 |
| | **3** | 91.68 | 0.00 | 0.00 |
| | **4** | 90.93 | 0.00 | 0.00 |
| | **6** | 0.00 | 92.36 | 0.00 |
| | **7** | 0.00 | 91.70 | 0.00 |
| | **8** | 0.00 | 0.00 | 91.90 |

| T2 | | Voisins | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 91.45 | 0.00 | 0.00 |
| | **2** | 90.89 | 0.00 | 0.00 |
| | **3** | 91.66 | 0.00 | 0.00 |
| | **4** | 90.93 | 0.00 | 0.00 |

| | | | | |
|---|---|---|---|---|
| | **6** | 0.00 | 92.36 | 0.00 |
| | **7** | 0.00 | 91.70 | 0.00 |
| | **8** | 0.00 | 0.00 | 91.90 |

| T3 | | **Voisins** | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 91.63 | 0.00 | 0.00 |
| | **2** | 91.06 | 0.00 | 0.00 |
| | **3** | 92.36 | 0.00 | 0.00 |
| | **4** | 90.49 | 0.00 | 0.00 |
| | **6** | 0.00 | 92.36 | 0.00 |
| | **7** | 0.00 | 91.70 | 0.00 |
| | **8** | 0.00 | 0.00 | 91.90 |

| T4 | | **Voisins** | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 90.97 | 0.00 | 0.00 |
| | **2** | 90.89 | 0.00 | 0.00 |
| | **3** | 91.56 | 0.00 | 0.00 |
| | **4** | 90.95 | 0.00 | 0.00 |
| | **6** | 0.00 | 92.30 | 0.00 |
| | **7** | 0.00 | 91.25 | 0.00 |
| | **8** | 0.00 | 0.00 | 91.90 |

| T5 | | **Voisins** | | |
|---|---|---|---|---|
| | | **3** | **6** | **8** |
| **Destination** | **1** | 91.45 | 0.00 | 0.00 |
| | **2** | 90.64 | 0.00 | 0.00 |
| | **3** | 92.14 | 0.00 | 0.00 |
| | **4** | 90.93 | 0.00 | 0.00 |
| | **6** | 0.00 | 92.36 | 0.00 |
| | **7** | 0.00 | 91.70 | 0.00 |
| | **8** | 0.00 | 0.00 | 91.90 |

*Tableaux IV-3 : LES tables de routage expérimentation N°1*

| Nombresauts | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| Moyenévaluation | **91.95** | **91.28** | **90.69** | **90.18** | **89.66** |

*Tableaux IV- : Moyenne d'évaluation en fonction de nombre sauts expérimentation N°1*

En examinant ces chiffres on peut constater que :

- Les liens directs ont toujours des valeurs d'évaluation très élevées avec une moyenne de 91.95.
- Les valeurs d'évaluation nulles signifient que le nœud ne pourra jamais atteindre la destination via ce voisinage.
- Pour chaque destination on a une valeur d'évaluation élevée (la fonction random pour le choix du voisin dépende de l'outil de développement ce qui explique les valeurs faibles et nulles à cause d'évaluation de même chemins ou élimination d'une boucle pour les chemins les plus longs.

Pour la 2eme expérimentation on dispose d'une topologie avec des liens à coûts différents

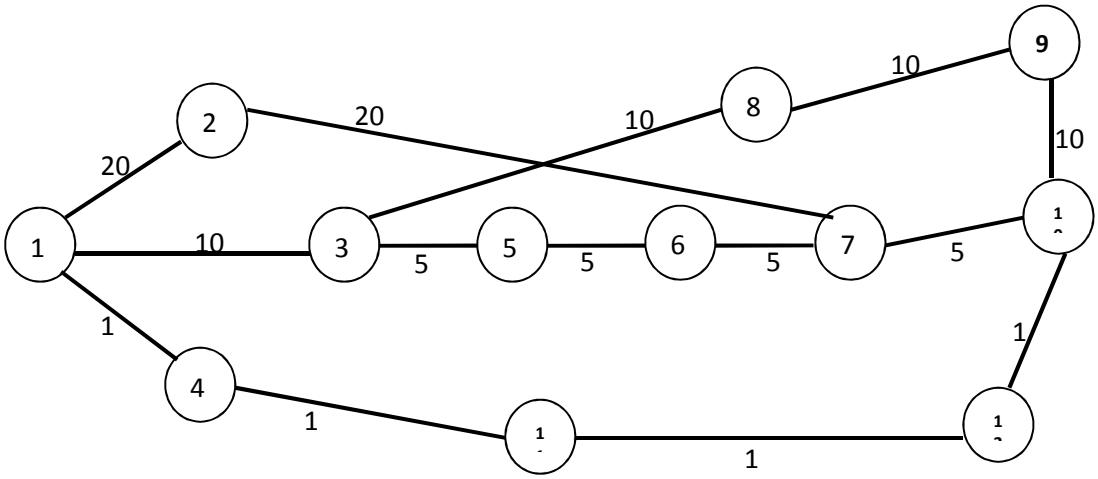



*Figure IV-14 : Topologie d'expérimentation N°2*

Examinons maintenant les résultats obtenus en appliquant les protocoles ACERP sur cette topologie pour quelques déférents chemins

**Source 1 destination 10**

| Voisins | | |
|---|---|---|
| **2** | **3** | **4** |
| 91.43 | 91.30 | 90.36 |

**Source 1 destination 7**

| Voisins | | |
|---|---|---|
| **2** | **3** | **4** |
| 92.71 | 92.11 | 90.00 |

**Source 10 destination 4**

| Voisins | | |
|---|---|---|
| **7** | **9** | **12** |
| 91.70 | 91.69 | 90.95 |

D'après les résultats obtenus on peut constater que le protocole ACERP s'adapte selon les débits des chemins traversés. Un chemin a gros débit porte des valeurs d'évaluation supérieures à celles d'un chemin à débit inférieur.

Avec la 3eme expérimentation on utilise la topologie suivante ou on suppose un trafic important entre les nœuds 2 et 3.

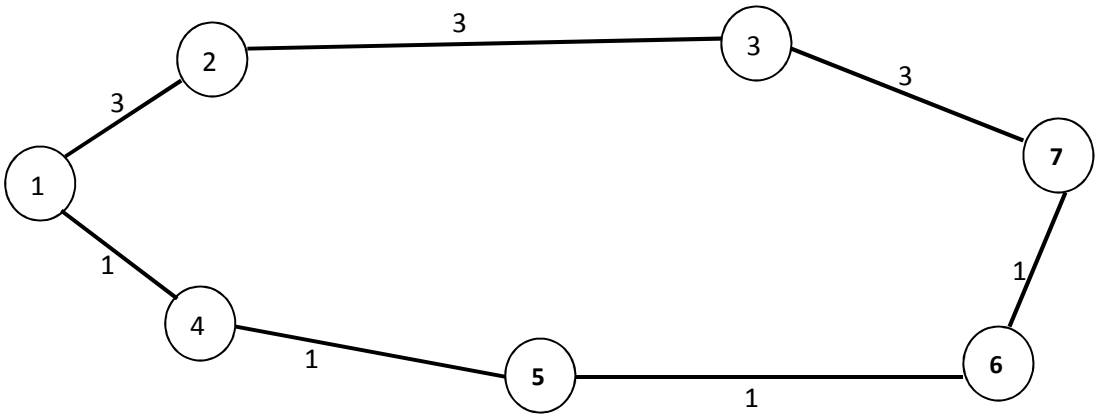



*Figure IV-15 : Topologie d'expérimentation N°3*

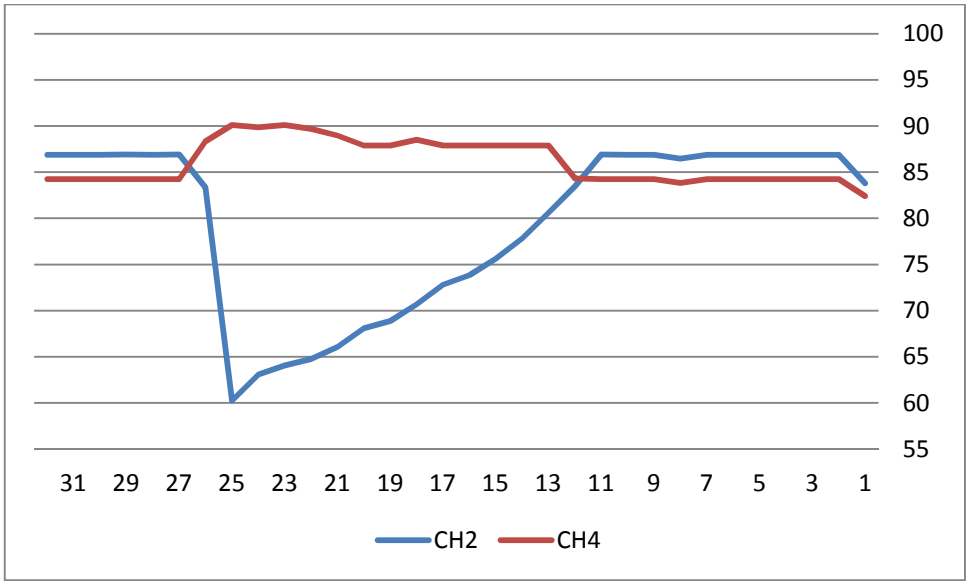



*Figure IV-16 : Diagrammes d'évaluation pour les deux chemins - Topologie N°3*

Ce graphe montre que le protocole ACERP il s'adapte selon la situation du réseau, au début le chemin ramené du nœud 1 a 7 via les nœuds 3 et 2 est le meilleur mais au moment de la simulation du trafic entre les 2 nœuds 2 et 3 se chemin devient inacceptable il return toujours des valeurs d'évaluations faibles.

La 4eme et la dernière expérience on fait une comparaison entre les protocoles ACERP et AntNet, en les appliquant sur la topologie suivante.

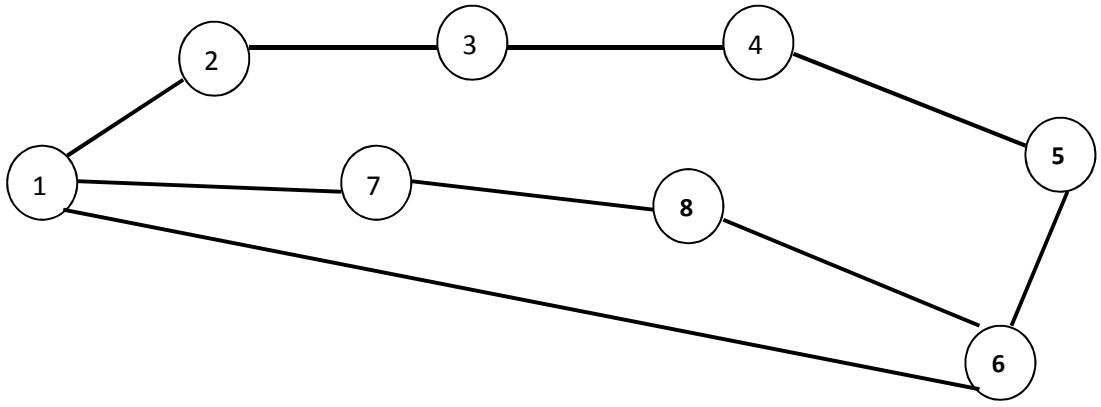



*Figure IV-17 : Topologie d'expérimentation N°4*

Appliquons cette topologie avec des protocoles AntNet et ACERP successivement, les liens ont des coûts égaux, les prises de valeurs sur des périodes constantes. Examinent les différents chemins entre les nœuds 1 et 6 on a obtenu les résultats suivants

| **AntNet (source=1 destination=6)** | | | |
|---|---|---|---|
| **Temps** | **Voisin** | | |
| | **2** | **7** | **6** |
| T0 | 0.33 | 0.33 | 0.33 |
| T1 | 0.33 | 0.33 | 0.34 |
| T2 | 0.32 | 0.33 | 0.35 |
| T3 | 0.31 | 0.33 | 0.37 |
| T4 | 0.31 | 0.33 | 0.37 |
| T5 | 0.31 | 0.33 | 0.39 |
| T6 | 0.30 | 0.33 | 0.40 |
| T7 | 0.30 | 0.33 | 0.41 |
| T8 | 0.30 | 0.33 | 0.42 |
| T9 | 0.29 | 0.33 | 0.43 |
| T10 | 0.29 | 0.33 | 0.44 |
| T11 | 0.28 | 0.33 | 0.45 |
| T12 | 0.28 | 0.33 | 0.46 |
| T13 | 0.27 | 0.34 | 0.47 |
| T14 | 0.26 | 0.33 | 0.48 |
| T15 | 0.27 | 0.33 | 0.48 |
| T16 | 0.26 | 0.34 | 0.49 |
| T17 | 0.26 | 0.33 | 0.50 |
| T18 | 0.26 | 0.34 | 0.51 |
| T19 | 0.26 | 0.33 | 0.52 |
| T20 | 0.25 | 0.34 | 0.53 |

| ACERP (source=1 destination=6) | | | |
|---|---|---|---|
| **Temps** | **Voisin** | | |
| | **2** | **7** | **6** |
| T0 | 0.00 | 0.00 | 0.00 |
| T1 | 0.00 | 0.00 | 83.28 |
| T2 | 79.77 | 81.56 | 83.07 |
| T3 | 77.02 | 80.64 | 83.52 |
| T4 | 79.37 | 81.15 | 85.64 |
| T5 | 79.77 | 81.56 | 83.07 |
| T6 | 77.80 | 80.66 | 83.94 |
| T7 | 79.77 | 81.55 | 83.06 |
| T8 | 80.57 | 79.94 | 83.07 |
| T9 | 78.58 | 80.34 | 84.78 |
| T10 | 79.77 | 81.55 | 83.07 |
| T11 | 80.03 | 79.44 | 82.68 |
| T12 | 78.97 | 80.74 | 83.28 |
| T13 | 79.77 | 81.55 | 38.06 |
| T14 | 79.79 | 80.24 | 83.10 |
| T15 | 79.37 | 81.15 | 85.64 |
| T16 | 79.77 | 81.55 | 83.06 |
| T17 | 77.80 | 80.66 | 83.94 |
| T18 | 79.37 | 81.15 | 85.64 |
| T19 | 80.17 | 80.95 | 83.06 |
| T20 | 78.19 | 81.06 | 84.36 |

*Tableaux IV-5 :les quantités de phéromone et valeurs d'évaluation pour les protocoles AntNet et ACERP*

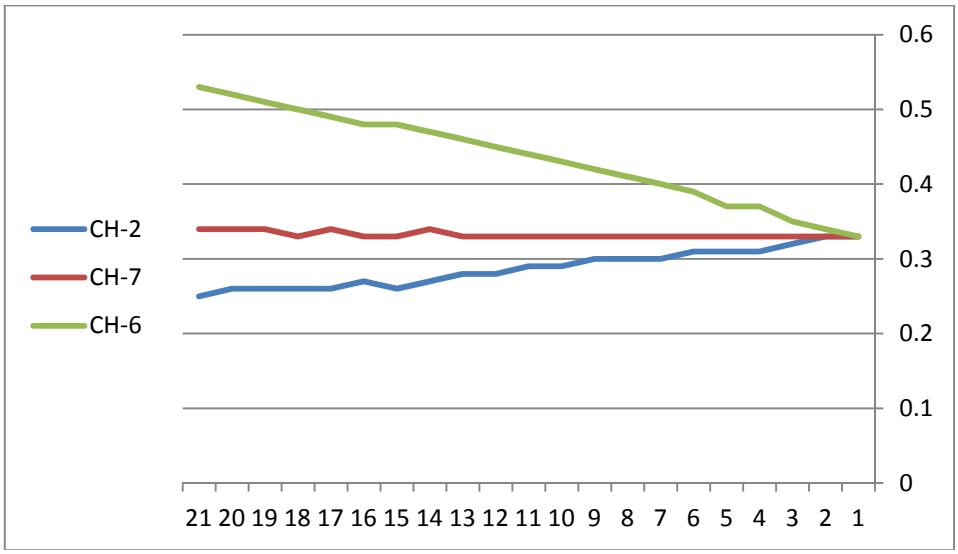



*Figure IV-18 : Diagrammes des quantités de phéromone pou les 3 chemins protocole AntNet – Topologie N°4*

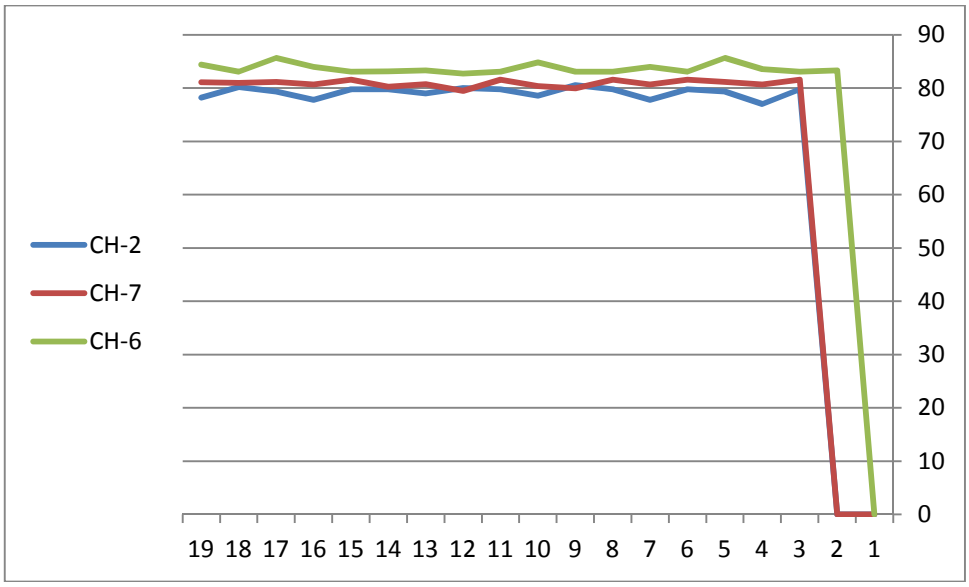



*Figure IV-19 : Diagrammes d'évaluation pour les 3 chemins protocole ACERP –Topologie N°4*

Examinant les deux représentations graphiques, le chemin qui a l'évaluation la plus petite pour le protocole ACERP a une valeur de la quantité de phéromone la plus petite pour AntNet et vice versa.

## IV.4. Conclusion

Dans ce chapitre, nous avons présenté notre contribution. Il s'agit d'un protocole de routage adaptatif qui utilise l'algorithme routage par colonie de fourmis et conduit par l'évaluation de la perception de l'utilisateur finale. Pour évaluer la perception utilisateur, nous avons utilisé la méthode d'évaluation E-model de l'ITU.

Nous avons présenté une brève description de la méthode E-model. C'est un algorithme basé sur vingt paramètres liés aux facteurs terminaux, facteurs de l'environnement et des factures liés au réseau. Il faut souligner que le résultat brut fourni par ce modèle est le "facteur d'évaluation " R qui peut être transformé pour obtenir des estimations de l'opinion des usagers(MOS).

Une description de l'algorithme proposé (ACERP) est décrit dans ce chapitre où nous avons présenté un nouveau type de paquet qui représente selon l'algorithme de routage par colonie de fourmis les paquets Bant et Fant, une description de la table de routage pour chaque nœud de la topologie ainsi qu'un organigramme de fonctionnement du protocole.

La simulation a été faite par l'outil NS-2 où nous avons présenté une description déclarative de l'agent protocole ACERP. Pour les cadres expérimentaux, nous avons appliqué ce protocole sur quatre topologies différentes où nous avons pu en déduire que le protocole ACERP s'adapte selon le nombre des sauts, les débits des chemins traversés, les chemins congestionnés ou les chemins qu'ils ont un trafic important. En termes de comparaison avec le protocole AntNet, nous avons obtenu pratiquement les mêmes résultats.

-

# • *Conclusion et Perspevtives*

# Conclusion Générale

Avec le développement et le déploiement massif d'Internet, les besoins en QoS sont devenus de plus en plus indispensables. Les contraintes de haut débit, l'hétérogénéité des liens et la nature du trafic exigent de disposer d'un réseau réactif aux changements qui peuvent intervenir lors du transport des flux. Les études et les recherches que nous avons menées et présentées dans ce mémoire sont dans ce contexte et ont pour objectif principal une meilleure prise en compte de la QoS dans les décisions de routage.

Les services de la Voix sur IP (VoIP) sont maintenant offerts par différents prestataires de service et souscrits par un grand nombre d'utilisateurs fixes et mobiles dans les systèmes de gestion des réseaux multimédias. La VoIP apporte des gains pour les fournisseurs de services ainsi que pour les clients. Pour les fournisseurs de services, les coûts opérationnels sont réduits en raison de la distribution des différents services, tels que la voix et les données, dans une infrastructure réseau avec des ressources partagées. Pour les clients, la VoIP présente les caractéristiques d'un réseau public traditionnel de téléphonie commuté (RTC), comme la messagerie vocale et la conférence vocale, d'une manière omniprésente ainsi que le soutien au niveau qualité et bien évidemment un prix abordable.

Afin d'atteindre de nouvelles opportunités et d'améliorer la compétitivité du marché, les prestataires de services du réseau offrent de nouveaux services à valeur ajoutée, tels que la vidéo sur demande (VoD), l'IPTV, la VoIP, etc. En conséquence, l'amélioration de la qualité des services comme perçue par les utilisateurs, généralement désignée sous le nom de la qualité de l'expérience (QoE), a un grand effet aussi bien qu'un défi significatif pour les fournisseurs de services avec l'objectif de minimiser le taux de désabonnement tout en conservant leur avantage concurrentiel. Le nouveau terme QoE a été introduit, en combinant la perception des utilisateurs, l'expérience et les espérances sans oublier les paramètres techniques, essentiellement les paramètres de la qualité de service. Dans ce mémoire, nous avons procédé dans un premier temps à une analyse critique des travaux de recherche destinés à intégrer la perception utilisateur dans le processus de routage et les nouvelles approches de routage liées à la qualité de Service. Nous avons ensuite orienté nos recherches essentiellement vers le développement de nouvelles méthodes pour l'optimisation de la fonction de routage.

L'étude présentée dans le premier chapitre, nous a permis de présenter le service voix sur IP (Internet Protocole), les avantages, les inconvénients, les principes et les standards liés.

Sur la base d'une longue analyse des techniques liés aux QdS et les méthodes d'évaluations de la qualité de la voix, nous avons proposé un nouveau protocole de routage adaptatif orienté QoE appelé ACERP (Ant Colony Evaluate Routing Protocol). L'originalité de notre approche est liée à deux aspects: **routage adaptatif** par colonie de fourmis et l'**évaluation** de la perception utilisateur basée sur une méthode objective.

## Perspectives

Les travaux présentés dans ce mémoire ouvrent de nombreuses perspectives dans le cadre de l'utilisation des approches adaptatives dans le routage et plus généralement pour la prise en compte de la QdS et de la QdE. Outre la poursuite des travaux en cours, il s'agira de travailler sur l'amélioration du protocole ACERP par rapport à plusieurs pistes : L'utilisation de la méthode d'évaluation E-model qui ne dépend pas uniquement des paramètres réseau risque de ne pas donner de bons résultats, les méthodes hybrides telles que la méthode PSQA [Varlela, 2006] basée sur les réseaux de neurones donnera peut-être de meilleurs résultats.

L'utilisation des concepts des systèmes multi-agents et agents mobiles permettrons, peut-être, d'introduire des améliorations dans le plan d'implémentation et d'introduire de nouveaux concepts tels-que l'aspect communication et coopération entre les agents.

## Références Bibliographiques


[Amirat et al., 2004] Amirat . Y, Hoceini .S, Mellouk .A.(2004). Neural Net Based Approch for Adaptative Routing Policy in Telecommunication Networks.*HSNMC* 2004

[Amirat et al., 2005] Amirat . Y, Hoceini .S, Mellouk .A.(2005).K-Shortest Paths Q-Routing A New QoS Routing Algorithm in Telecommunication Networks.*ICN* 2005

[Baset et Schulzrinne, 2006] Baset S.A, Schulzrinne .H.(2006).An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol.*Proceedings of the INFOCOM spain 2006*

[Baudoin et Karle, 2004] Baudoin .B, Karle .M.(2004).NT Réseaux IDS et IPS.*support de cours Enseignant Etienne Duris* 2004

[Beerends et Stemerdink, 1992] Beerends .J, Stemerdink .J. (1992).A perceptual audio quality measure based on a psychoacoustic sound representation. *journal audio engineering society,* 1992

[Beuran, 2004] Beuran .R. (2004).Mesure de la Qualite dans les Réseaux Informatique.*université Jean Monnet St. Etienne*

[Bolot, 1993] Bolot J.(1993). Characterizing end-to-end packet delay and loss in the internet. *High speed network,*

[Boyan et Littmann, 1994] Boyan .J, Littmann M.L.(1994). Packet routing in dynamically changing networks: A reinforcement learning approach. *Advances in Neural Information Processing Systems*

[Braun et al., 2008] T BRAUN,M. DIAZ, Staub .T(2008). end-to-end Quality of Service over heterogeneous Networks.*Springer*

[Cho et Un, 1994] Cho .Y.J, Un .C.K.(1994). Performance analysis of reconstruction algorithms for packet voice Communications. *Computer Network and ISDN Systems*

[Cisco, 2000] Cisco Documentation.(2000).Understing delay in packet voice networks.*http://www,cisco,com/warp/public/788/voip/delay-details,html, 2000*

[Degris, 2007] Degris .T.(2007).Apprentissage par renforcement dans les processus de décision Markoviens factorisés.*Thèse de doctorat, AnimatLab*

[Dicaro et Dorigo, 1998] Dicaro .G, Dorigo .M.(1998). AntNet: Distributed Stigmergetic Control for Communication Networks. *Journal of AIR*

[Fischbach, 2004] Fischbach .N.(2004).sécurité de la Voix sur IP.*COLT Telecom Colt Technology Services*

[Frikha et al., 2009] Frikha .A, Bertrand .G, Lahoud .S.(2009). Pré-calcul de chemins inter-domaines soumis à plusieurs contraintes de QoS". *IRISA(Institut de recherche en informatique et systèmes aléatoires)*

[Gelenbe et al., 2001] Gelenbe .E, Lent .R, Xu .Z(2001). Measurement and performance of Cognitive Packet Networks. *université of central florida, Orlando, FL 32816*

[Gelenbe, 1993] Gelenbe .E .(1993).Learning in the recurent random neural Networks. *Ecole des Hautes Etudes en Informatique-Université René*

*Descartes (Paris V)*

[Guillet, 2010] Guillet .T. (2010).Sécurité de la téléphonie sur IP.*TELECOM ParisTech*

[Gray et al., 2000] Gray .R. S, Cybenko .G, Kotz .D et Rus .D.(200).Mobile agents: Motivations and state of the art. *Department of Computer Science, Dartmouth College, Hanover, USA.*

[Haddade 2006] Haddade .Y.(2006).La voi sur ip et le wifi genèse d'une technologie née en Israël. *science et technologie ambassade de France*

[Hedrick, 1988] Hedrick .C.(1988).Routing Information Protocol. *(*Request for Comments*)-RFC 1058.*

[Hoceini, 2004] Hoceini .S.(2004). Techniques d'Apprentissage par Renforcement pour le Routage Adaptatif dans les Réseaux de Télécommunication à Trafic Irrégulier. *Thèse de doctorat UNIVERSITE PARIS XII – VAL DE MARNE*

[Holl, 2007] Holl .T.(2007).Transport à fiabilité partielle d'images compressées sur les réseaux à commutation de paquets. *Université Henri Poincaré, CRAN.*

[Innokenty, 2000] Innokenty .R.(2000).Configuration des routeurs CISCO. *Editions Eyrolles 2000.*

[ITU, 2005] UIT-T(Union internationale des télécommunications).(2005). "le modèle E : modèle de calcule utilisé pour la planification de la transmission. *Recommandation UIT-T G.107,2005*

[ITU, 2008] UIT-T(Union internationale des télécommunications).(2008) .the e-model, a computation model for use in transmission planning. *Recommandation UIT-T G.107,2008*

[ITU, 1996] UIT-T(Union internationale des télécommunications).(1996). Transmission Systems and Media, general recommendation on the transmission quality for entire international telephone connection.*RecommandationG,114, 1996*

[ITU, 1999] UIT-T(Union internationale des télécommunications).(1996). Provisional Planning Values for the equipment impairment factor" , RecommandationG,113, 1999

[Jennings et Wooldridge, 1995] Jennings .N. R. et Wooldridge .M.(1996).Intelligent agents: Theory and practice. *The Knowledge Engineering Review,1995.*

[Kaelbling et al., 1996] Kaelbling L. P, Littman M.L, Moore A.W.(1996).Reinforcement learning : A survey. *Journal of Artificial Intelligence Research,1996.*

[Keagy, 2000] Keagy .S.(2000). integration voice and data networks : practical solution for the new world of packetized voice over data networks. *Cisco press 2000*

[Kuipers et al., 2004] Kuipers F.A, Kormaz .T, Krunz .M, et Van Mieghem .P.(2004).Performance evaluation of constraint-based path selection algorithms. *IEEE Network ,2004*

[Leriche et Arcangeli, 2006] Leriche .S, Arcangeli .J.(2006).Vers un modèle d'agent flexible. *JMAC'06, Nîmes, mars 2006.*

[Liang et al., 2001] Liang .Y.J, Farber .N, Girob .B.(2001).Adaptiveplayout scheduling and loss concealment for voice communications over IP Networks.*IEEE Transactions on multimedia, 2001*

[Malcolm et Zhao, 1993] Malcolm N. et Zhao .W.(1993).Guaranteeing Synchronous Messages with Arbitrary Deadline Constraints in an FDDI Network.*Department of Computer Science,Texas A&M University.*

[Mccullogh et Pitts, 1943] Mccullogh W.S, Pitts W.H.(1943).A Logical Calculus for the ideas imminent in nervous activity.*Bulletin of Math. Biophysics,1943.*

[Mellouk, 2007] Mellouk .A.(2007).Mécanismes du contrôle de la qualité de service: Applications temps réel et multimédia . HERMES Science Paris 2007

[Mellouk, 2009] Mellouk .A.(2009).End-to-End Quality of service engineering in next generation heterogeneous networks . *ISTEetJohn,Wiiley& Sons,2009*

[Oneil, 2002] Oneil .M.(2002).Application based QoS for IP Video conferencing.*director of technical marketing polycom video communications, White Papers by Polycom 2002*

[Ouakil et pujolle, 2008] Ouakil .L, Pujolle .G.(2008).Téléphonie sur IP.*Eyrolles, 2eme édition 2008*

[Personnaz et al., 1990] Personnaz .L, Nerrand .O et Dreyfus .G.(1990). Apprentissage et mise en œuvre des réseaux de neurones bouclés.*Journées Internationales des sciences de l'informatique, Tunis, 1990.*

[Perret, 1997] Perret .S.(1997).Agents mobiles pour l'accès nomade à l'information répartie dans les réseaux de grande envergure.*Thèse de Doctorat. Université Joseph Fourier - Grenoble I, 1997.*

[Plummer, 1982] Plummer .D.(1982). An Ethernet Address Resolution Protocol. ( requests for comments)RFC 826,1982

[Porter et Gough, 2007] Porter .T, Gough .M.(2007).How to cheat at VoIP Security. *Syngress Publishing Inc , 2007*

[Rix, 2002] Rix .A.(2002).Advences in objective quality assessment of speech over analogue and packet based networks.*IEEE colloquium, 2002*

[Rumelhart et al., 1986] Rumelhart D.E, Hinton G.E et R.J. Williams.(1989).in Parallel Distributed Processing. *M.I.T. Press, 1986.*

[Sanghi et al., 1993] Sanghi .D, Agrawala .A.K, Gudmundsson .O et Jain .B.(1993).Experimental assessment of end-to-end behavior on internet.*IEEE InfoCom, 1993*

[Schadle, 2006] Schadle .A.(2006).bRTP BIBLIOTEQUE RTP.*Institut de Recherche en Informatique et Systèmes Aléatoires, 2006*

[Seb, 2004] Seb .F.(2004).Tous sur VoIP.*La communauté de FrameIP, Disponible à l'adresse http://www.frameip.com/voip/*

[Trad, 2002] Trad .A.(2002).étude des mécanismes de contrôle de la transmission de flux audio sur internet.*INRIA 2002*

[Varlela, 2006] Varlela .M.(2006).évaluation pseudo subjective de la qualité des flux VoIP une approche par RNA. Projet Armor, IRISA. 2006

[Vleeschauwer et al., 2000] Vleeschauwer .D. DE, Janssen .J, Petit .G.H, Poppe .F.(2000).Quality bounds for Packetized voice transport.*Alcatel Telecommunication review, 2000*

[Voran, 2002] Voran .S.(2002).Estimation of perceived speech quality using measuring normalizing blocks.*IEEE workshop, 2002*

[Wang, 1992] Wang S.(1992).Objective measure for predicting subjective quality of speech coders.*IEEE, sept. 1992.*

[Zhang, 1997] Zhang .P.(1997).Etude de Différents Aspects de l'Apprentissage par Renforcement. *Thèse Université de Technologie de Compiègne, 1997.*

[Ziani, 2008] Ziani .S.(2008).une approche inductive dans le routage à optimisation du délai: application aux reseaux 802.11.*Paris 12 val de marne, 2008*