

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

Université Abderahmane Mira de Béjaïa
Faculté des Sciences Exactes
Département de Recherche Opérationnelle

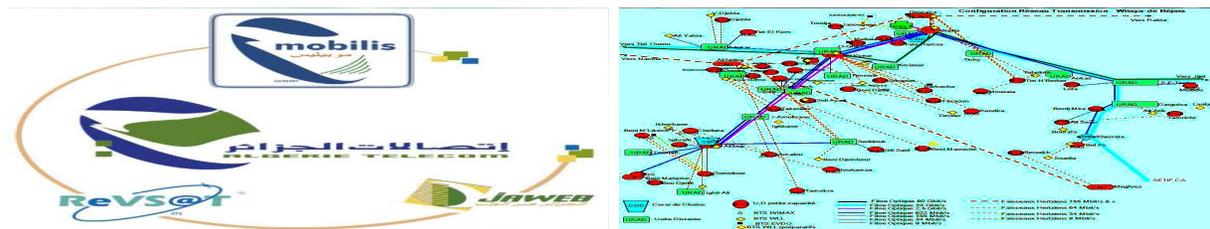


Mémoire de fin cycle

En vue de l'obtention du Diplôme de
MASTER en Recherche Opérationnelle
Option : Modelisation et évaluation de Performances des réseaux

Thème

Les NGN à Algérie Télécom : Evaluation des performances du routeur de cœur de l'architecture DiffServ



Présentée par : **MOUSSAOUI Salim & OUAZENE Zidane**

Devant le jury composé de :

Président	Mme Adel Karima	M.C.B.	Univ. de Béjaïa
Rapporteurs	Mr Aïssani Djamil	Professeur	Univ. de Béjaïa
	Mme Lekadir Ouiza	M.C.B.	Univ. de Béjaïa
Examinatrice	Melle Outamazirt Assia	Doctorante	Univ. de Béjaïa
Examineur	Mr Soufit Massinissa	Doctorant	Univ. de Béjaïa
Invité	Mr Hand Khmis	Directeur de la DOT	Wilaya. de Béjaïa

Béjaïa, Novembre 2015.

Remerciements

Nous remercions Dieu, le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

Nous adressons notre reconnaissance, notre gratitude à nos encadreurs : Mr le professeur D. AISSANI et Mme O. LEKADIR pour l'honneur qu'ils nous ont fait en assurant le suivi scientifique et technique du présent mémoire.

Nos remerciements vont aussi à Mme K. ADEL pour l'honneur qu'elle nous fait en acceptant de présider le jury de soutenance de ce mémoire.

Nos remerciements s'adressent également aux examinateurs Melle A. OUTAMAZIRT et Mr M. SOUFIT pour l'honneur qu'ils nous ont fait en acceptant d'examiner notre travail.

Nous remercions tous ceux qui ont participé de près ou de loin à la contribution de ce travail.

Dédicaces

A la mémoire de mon Père...

Je dédie cet événement marquant de ma vie à la mémoire de mon père parti trop tôt. Que Dieu le tout Puissant l'ait accueilli en sa sainte miséricorde !

A ma très chère mère

Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et de prier pour moi. Tes prières et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études. Aucune dédicace ne saurait être assez éloquente pour exprimer ce que tu mérites pour tous les sacrifices que tu as fait pour nous tes enfants.

A mes très chères frères et soeurs Saci, Oualid, Ouafia et Siham

En témoignage de mon affection fraternelle, de ma profonde tendresse et reconnaissance.

A Mes chers neveux Massilas, Ales et Iyad

Je prie pour que Dieu vous protège.

A mon cher binôme Salim

Mes meilleurs voeux de succès dans ta vie.

A tous les membres de ma grande famille.

A tous mes camarades ainsi qu'à tous mes enseignants

A ma chère amie Mounira

«Zidane»

A mes chers parents

pour leur soutien moral et financier durant mes études,

à mes soeurs Zahia, Fahima, Lila et Samiha

en leur souhaitant la réussite dans leurs vies,

à mon frère Azzedine

à mes amis : Ouardia et Salim

A mon cher binôme Zidane

Mes meilleurs voeux de succès dans ta vie.

A tous les membres de ma grande famille.

A tous mes camarades ainsi qu'à tous mes enseignants.

«*Salim*»

Table des matières

Table des matières	i
Table des figures	iv
Liste des tableaux	v
Liste des abréviations	vi
Introduction générale	1
1 Présentation du groupe Algérie Télécom	3
1.1 Historique du groupe AT	3
1.2 Missions et objectifs du groupe	3
1.3 Organigramme d'Algérie Télécom	4
1.4 Les domaines d'activité d'AT	7
1.5 Le réseau d'Algérie Télécom	7
1.5.1 Réseau commercial	7
1.5.2 Réseau de transmission	7
1.5.3 Réseau international	7
1.6 La structure organisationnelle de la DOT de Béjaia :	7
1.7 Conclusion	9
2 Perspectives NGN d'Algérie Télécom	10
2.1 Nécessité de modernisation des réseaux de télécommunications	10
2.2 RMS d'Algérie Télécom	11
2.2.1 Topologie du réseau RMS d'AT	11
2.2.2 Objectifs du déploiement du réseau RMS d'AT	13
2.2.3 La Mise à niveau du Backbone national	13
2.2.4 La Migration du réseau RTC	14
2.2.5 Présentation du réseau d'accès MSAN d'Algérie Télécom	14
2.3 Réseau de transmission de la Wilaya de Béjaia	15
2.4 Position du problème	16

3	DiffServ et qualité de service	18
3.1	Généralités sur la Qualité de Service	19
3.1.1	Définition	19
3.1.2	Objectif de l'étude de la QoS	19
3.1.3	Paramètres de la QoS :	19
3.1.4	Les informations utilisées pour assurer la QoS	20
3.1.5	La gestion de la qualité de service :	20
3.2	L'architecture DiffServ	21
3.2.1	Principe de fonctionnement de l'approche DiffServ	21
3.2.2	La notion de domaine	22
3.2.3	Agrément du niveau de service	23
3.2.4	Le champ DS	23
3.2.5	Structure d'un réseau DiffServ	24
3.2.6	Techniques de Lissage du trafic (Traffic Shaping)	29
3.2.7	Per-hop-Behavior standardisés	30
3.2.8	Gestion active des files d'attente	33
3.2.9	Mécanismes d'ordonnancement	34
3.3	Conclusion	35
4	Rappels et modélisation du problème	36
4.1	Rappels	36
4.1.1	Description d'un système de file d'attente classique	36
4.1.2	Analyse mathématique d'un système de files d'attente	37
4.1.3	Classification des systèmes d'attente	38
4.1.4	Notation de Kendall	38
4.1.5	Les différentes disciplines de service	39
4.1.6	Mesures de performance d'une file d'attente	39
4.1.7	Description d'un système d'attente avec rappels	40
4.1.8	Description d'un système d'attente avec vacances	40
4.1.9	Quelques lois usuelles de probabilités	41
4.2	Modélisation du routeur de cœur par un système $M/G/1$ avec rappels et vacances	42
4.2.1	Description du modèle	42
4.3	Chaîne de Markov induite	43
4.3.1	Condition d'ergodicité	44
4.3.2	Les probabilités de transition	44
4.3.3	Distributions stationnaires de la chaîne de Markov induite	44
4.4	Approche par les processus régénératifs	45
4.4.1	Distributions limites	45
4.4.2	Décomposition stochastique	48
4.4.3	Quelques mesures de performance	49
4.4.4	Conclusion	49

5	Application	50
5.1	Simulation	50
5.2	Étapes de Simulation	50
5.3	Intérêts de la simulation	51
5.4	Avantages de la simulation	51
5.5	Inconvénients de la simulation	51
5.6	Performances du modèle du routeur de coeur	52
5.6.1	Organigramme de simulation	53
5.6.2	Evaluation des performances du modèle du routeur de coeur établi .	53
5.6.3	Interprétation des résultats	57
5.7	Conclusion	58
	Bibliographie	61

Table des figures

1.1	Organigramme d'Algérie Télécom.	5
1.2	La structure organisationnelle de la DOT de Béjaia.	8
2.1	Le réseau RMS déployé sur le territoire national.	12
2.2	Topologie globale du réseau RMS d'Algérie Télécom.	12
2.3	Backbone national fibre optique	14
2.4	Réseau de transmission de la Wilaya de Béjaia.	16
2.5	Le modèle de file d'attente approprié à l'architecture du routeur de cœur.	17
3.1	Architecture générale d'un domaine DiffServ	22
3.2	Positionnement du champ DSCP dans les paquets IP	24
3.3	Structure d'un réseaux selon l'approche DiffServ.	25
3.4	Routeur de bordure.	25
3.5	Architecture d'un routeur de cœur.	28
3.6	Principe de la technique Leaky bucket.	30
3.7	Principe de la technique Token bucket.	30
3.8	Arrivée des paquets dans un edge router puis dans un core router . .	32
3.9	Paramètres de RED.	33
3.10	La discipline de service Round Robin.	35
4.1	Un système de files d'attente.	37
4.2	Schéma général d'un système avec rappels;	40
5.1	Organigramme de la simulation du modèle étudié.	53

Liste des tableaux

1.1	Légende des codes.	6
5.1	Performances du modèle dans le cas d'une loi de service exponentielle et celle des vacances exponentielle.	54
5.2	Performances du modèle dans le cas d'une loi de service exponentielle et celle des vacances d'Erlang.	55
5.3	Performances du modèle dans le cas d'une loi de service d'Erlang et celle des vacances Exponentielle.	56
5.4	Performances du modèle dans le cas d'une loi de service d'Erlang et celle des vacances d'Erlang.	57

Liste des abréviations

AT : Algérie Télécom
IP : Internet Protocol
VoIP : Voix sur IP
NGN : New Generation Network
MPLS : Multi Protocol Label Switching
FTP : File Transfert Protocol
SMTP : Simple Mail Transfer Protocol
QoS : Qualité de Service
WLL : Wireless Local Loop
DSL : Digital Subscriber Line
FTTx : Fiber To The ...
Gbps : Gigabits par seconde
DWDM : Dense Wavelength Division Multiplexing
kbps : Kilobits par seconde
MGC : Media Gateway Controller
TDM : Time-Division Multiplexing
SIP : Session Initiation Protocol
VPN : Virtual Private Network
FO : Fibre Optique
MSAN : MultiService Access Node
DSLAM : Digital Subscriber Line Access Multiplexer
ATM : Asynchronous Transfer Mode
RMS : Réseau Multi Services
TCP : Transmission Control Protocol
UDP : User Datagram Protocol
SCTP : Stream Control Transmission Protocol
DSCP : Differentiated Services Code Point
WFQ : Weighted Fair Queuing
WRED : Weighted Random Early Detection

Introduction générale

Le groupe Algérie Télécom (AT) a annoncé le lancement fonctionnel du nouveau réseau d'accès de nouvelle génération, NGN (New Generation Network), destinée à unifier l'ensemble des réseaux existants (fixe, données et internet). Ce réseau, qui propose une architecture ouverte et adaptée à la transmission par paquets IP, permettra notamment une convergence des services voix/données/vidéos, ainsi qu'un enrichissement du catalogue de services grâce aux possibilités offertes (visiophonie, push to talk, conférence vidéo/audio, VoIP, triple play, géolocalisation). L'introduction progressive de ce type d'équipements dans le réseau d'Algérie Télécom assurera aux abonnés une "qualité de service irréprochable", que ce soit par l'introduction du tout IP ou par le rapprochement du client final afin de réduire les déperditions et les perturbations pouvant être causées par l'éloignement équipements/abonnés. La réalisation du projet NGN s'inscrit dans l'action entamée depuis les années 90 visant à numériser le réseau d'accès d'Algérie Télécom et se veut le chaînon indispensable pour la mise en valeur du réseau IP/MPLS déjà construit par l'entreprise.

Initialement, l'internet était un outil pour une petite communauté d'agences et d'organisations, où des services comme le transfert de fichier "FTP", le courrier électronique "SMTP" et les procédures de connexion à distance dominaient. Cependant, avec l'apparition de nouveaux types d'applications multimédia telles que la vidéoconférence, la téléphonie et le commerce électronique, de nouvelles exigences en qualité de services (QoS) ont vus le jour. Deux approches sont actuellement proposées par l'IETF (Internet Engineering Task Force) pour assurer la QoS.

L'approche DiffServ, ou services différenciés, normalisée actuellement dans le groupe de travail DiffServ de l'IETF, permet d'introduire une nouvelle façon de traiter les flux dans le réseau et de partager les ressources de ce dernier. Dans l'architecture à différenciation de service, la bande passante, le taux de perte et le délai de transit des paquets sont influencés par les opérations de conditionnement de trafic lors de l'entrée dans le réseau ainsi que par les modifications apportées au comportement des routeurs de cœur du réseau. Dans ce type de service, la différenciation s'opère au niveau des agrégats plutôt qu'au niveau des flux eux-mêmes, afin de pallier aux problèmes de mise à l'échelle.

Ainsi, depuis plus d'une décennie, nous assistons à une évolution permanente des tech-

nologies de transport de données numériques. Par la démocratisation du réseau internet, le spectre des services proposés à un public en croissance permanente ne cesse de s'agrandir. Les réseaux actuels de transport de données utilisent la technologie de commutation de paquets, tandis que les réseaux de télécommunications (ainsi que les réseaux mobiles) utilisent la commutation de circuits plus adaptée aux exigences des services temps réels. Le déploiement des réseaux à haut débit et le développement des technologies MPLS permettent la conception de réseaux multiservices capables de transporter aussi bien les flux de données que les flux temps réel (voix, vidéo). Ces nouveaux réseaux seront capables de satisfaire les exigences différentes des différents flux qu'ils transportent. L'IETF a ainsi défini le nouveau protocole MPLS. Ce nouveau protocole est basé sur le paradigme de changement de label, dérive directement de l'expérience acquise avec les réseaux ATM (canaux virtuels, chemins virtuels).

Les grandes innovations de MPLS par rapport au routage IP traditionnel sont : la manipulation de tables de routage de bien plus petites tailles, des temps de commutation extrêmement rapides, des chemins (LSP) par classe de service, l'ingénierie du trafic, une meilleure maîtrise de la qualité de service et des mécanismes de routage en cas de panne. Le protocole MPLS, associé à des routeurs implémentant les mécanismes de différenciation de service (DiffServ) permet de traiter les flux séparément selon leurs caractéristiques (type de service, type d'application) par un routage plus précis et par le traitement différencié des paquets. C'est cette double fonctionnalité qui permet une meilleure utilisation des ressources de bout en bout afin d'assurer la qualité de service requise par les applications. Ainsi, le protocole MPLS est particulièrement adapté au modèle DiffServ. Dans ce dernier, on y retrouve la notion de routeurs de bordure et de routeurs de cœur de réseau (Edge/Core router) ainsi que la notion d'agrégation de flots.

Dans notre travail, on va s'intéresser au modèle de service différencié (DiffServ) qui est le plus adapté au protocole de nouvelle génération de réseaux MPLS. On s'intéressera particulièrement à l'étude des performances d'un routeur de cœur DiffServ en le modélisant d'abord par une file d'attente qui soit à notre avis assez adaptée, puis calculer ses paramètres de performances donnés analytiquement dans la littérature et enfin simuler ce modèle de file d'attente pour réaliser une comparaison de ses paramètres simulés avec ceux obtenus analytiquement.

Ce mémoire, s'articulera autour de cinq chapitres. Le chapitre I sera dédié à la présentation du groupe d'Algérie Télécom. Le chapitre II sera consacré aux perspectives NGN d'Algérie Télécom. Dans le chapitre III, la problématique du sujet de ce mémoire qui est l'étude et l'évaluation de la qualité de service des NGN sera exposée et une synthèse sur le modèle de qualité de service approprié "DiffServ" sera donnée. Dans le chapitre IV, des rappels sur les systèmes de files d'attente et sur quelques lois de probabilité seront donnés pour faire par la suite une modélisation de notre problème. Le chapitre V, sera consacré à notre application qui est la simulation du modèle établi. Enfin, ce mémoire se clôturera par une conclusion générale et une liste de références.

Chapitre 1

Présentation du groupe Algérie Télécom

1.1 Historique du groupe AT

Algérie Télécom (AT), est une société par actions au capitale sociale opérant sur le marché des réseaux et services de télécommunications. Sa naissance a été consacrée par la loi 2000/03 du 5 août 2000, relative à la restructuration du secteur des Postes et Télécommunications, qui sépare notamment les activités Postales de celles des Télécommunications.

Algérie Télécom est donc régie par la loi qui lui confère le statut d'une entreprise publique économique sous la forme juridique d'une Société Par Actions (SPA). Entrée officiellement en activité à partir du 1^{er} janvier 2003, elle s'engage dans le monde des Technologies de l'Information et de la Communication (TIC) avec trois objectifs :

- Rentabilité;
- Efficacité;
- Qualité de service.

1.2 Missions et objectifs du groupe

L'ambition d'Algérie Télécom est d'avoir un niveau élevé de performances techniques, économiques et sociales pour se maintenir durablement comme leader dans son domaine, dans un environnement devenu concurrentiel. Son souci consiste, aussi, à préserver et développer sa dimension internationale et participer à la promotion de la société de l'information en Algérie. L'activité majeure d'Algérie Télécom est de :

- Fournir des services de télécommunications permettant le transport et l'échange de la voix, de messages écrits, de données numériques, d'informations audiovisuelles,...
- Développer, exploiter et gérer les réseaux publics et privés de télécommunications;
- Établir, exploiter et gérer les interconnexions avec tous les opérateurs des réseaux.

Algérie Télécom est engagée dans le monde des TIC avec les objectifs suivants :

- Accroître l'offre de services téléphoniques et faciliter l'accès aux services de télécommunications au plus grand nombre d'utilisateurs, en particulier en zones rurales ;
- Accroître la qualité de services offerts et la gamme de prestations rendues et rendre plus compétitifs les services de télécommunications ;
- Développer un réseau national de télécommunications fiable et connecté aux autoroutes de l'information.

Les responsabilités d'AT s'exercent dans les trois domaines suivants :

- **Les actionnaires** : AT doit mériter leurs soutiens en valorisant leurs patrimoines ;
- **Les clients** : AT doit anticiper leurs besoins en leur fournissant des produits et des services de qualité afin de gagner et de conserver leurs confiances ;
- **Le personnel** : AT doit satisfaire ses attentes en organisant les conditions de l'épanouissement professionnel de chacun car la réussite d'AT dépend de l'engagement de tous.

1.3 Organigramme d'Algérie Télécom

Algérie Télécom est organisée en Directions Centrales, Régionales et Directions Opérationnelles de Wilaya autour de ses métiers fixes et services et d'autre part des fonctions supports réseaux. A cette structure s'ajoutent une filiale mobile (ATM MOBILIS) et deux autres filiales chargées l'une de l'internet (ATI DJAWEB) et l'autre des télécommunications spatiales (ATS REVSAT).

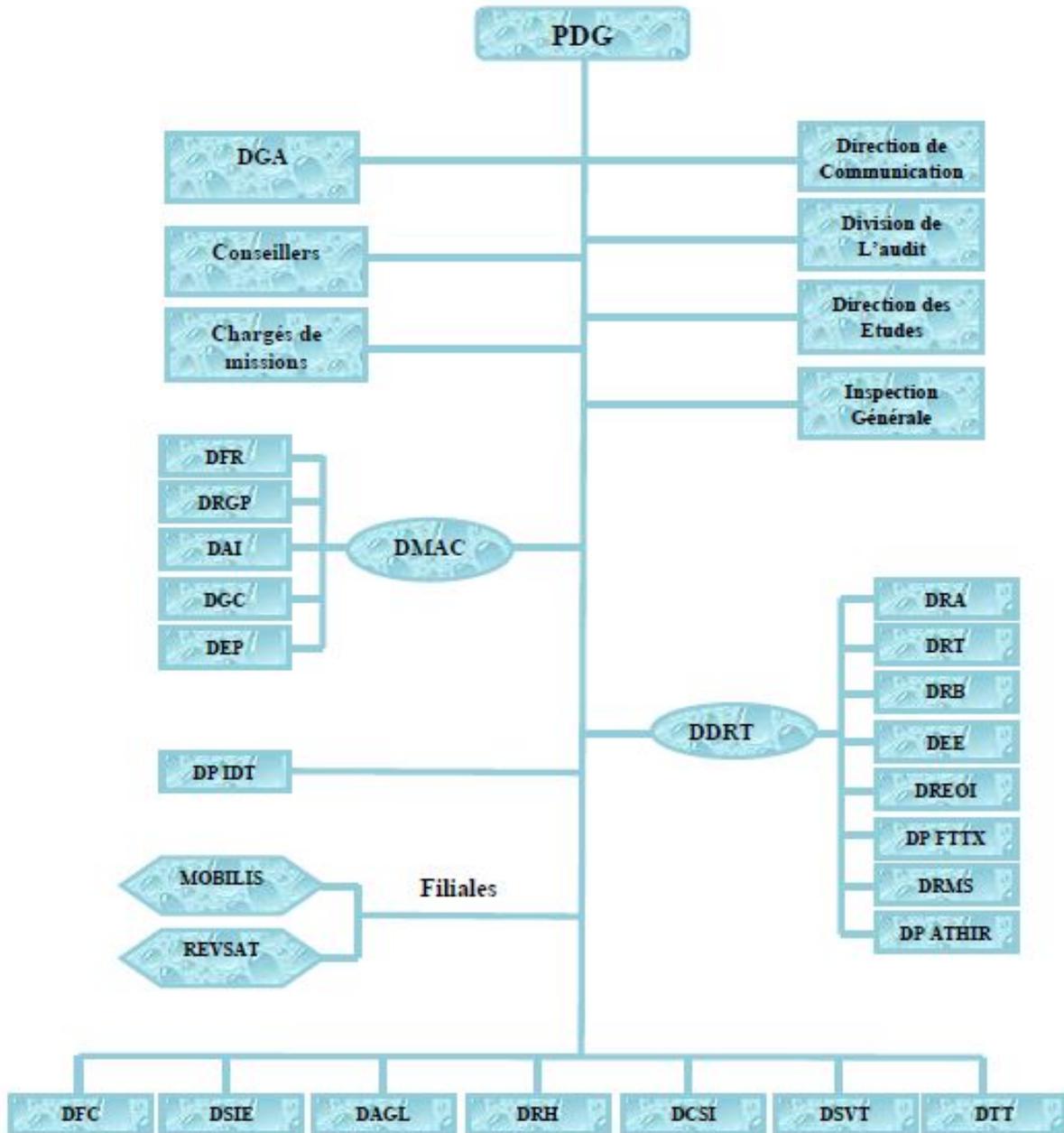


FIGURE 1.1 – Organigramme d'Algérie Télécom.

Code	Désignation
PDG	Président Directeur Général
DAGL	Direction d'Approvisionnement et Gestion de la Logistique
DAI	Direction des Affaires Internationale
DCSI	Direction Centrale de la Sécurité Interne
DDRT	Division Développement Réseaux des Télécommunications
DEE	Direction Énergie et Environnement
DEP	Direction Etudes et Programmes
DFC	Direction Finance et Comptabilité
DFR	Direction Facturation et Recouvrement
DGA	Directeur Générale Adjoint
DGC	Direction des Grands Comptes
DMAC	Division Marketing et Actions Commerciales
DP ATHIR	Direction Projet ATHIR
DP FTTX	Direction Projet FTTX (Fiber To The x)
DP IDT	Direction Projet IDT (informatisation des télécoms)
DRA	Direction Réseaux d'Accès
DRB	Direction Réseaux de Base
DREOI	Direction Réseaux Entreprise et Offres Intégrées
DRGP	Direction Relations Grand Publique
DRH	Direction des Ressources Humaines
DRMS	Direction Réseau Multi-Services
DRT	Direction Réseaux Transport
DSEI	Direction Systèmes et Équipements Informatiques
DSEIDSVT	Direction Stratégie et Veille Technologique
DTT	Direction Territoriale des Télécoms
MOBILIS	Direction Territoriale des Télécoms d'AT Mobile, Filiale du groupe AT
REVSAT	Algérie Télécom Satellite (ATS), Filiale du groupe Algérie Télécom

TABLE 1.1 – Légende des codes.

1.4 Les domaines d'activité d'AT

La société AT est l'acteur majeur des télécommunications en Algérie avec cinq domaines d'activités :

- **Téléphonie fixe** : avec deux millions de lignes en service et un réseau WLL en plein expansion ;
- **Téléphonie mobile** : activité au travers d'une filiale Mobilis, qui détient une part de marché de 13% ;
- **Transmission de données** : une activité de réseaux de données pour les entreprises (X25...);
- **Accès Internet à travers** : DJAWEB, FAWRI ADSL et dernièrement EASY ADSL ;
- **Réseau satellitaire** : des services de télécommunications s'appuyant sur VSAT, Inmarsat le réseau Thuraya.

1.5 Le réseau d'Algérie Télécom

1.5.1 Réseau commercial

Depuis la création d'Algérie Télécom, le réseau commercial a été renforcé d'année en année pour se rapprocher des clients. Aujourd'hui, nous comptons 341 agences commerciales contre 225 agences en 2005.[1]

1.5.2 Réseau de transmission

Le réseaux public de transmission de données par paquets X25 (DZPAC) d'AT possède 6.206 accès ; Le Backbone national de transmission d'AT à 10 GB/s, 2,5 GB/s, et 80GB/s en cours de réalisation ; Réseau radio rural : 103 réseaux intégrant plus de 1500 localités, 961 communes (APC) rattachées en fibres optiques.

1.5.3 Réseau international

Liaisons sous marine à fibre optique reliant l'Algérie au réseau mondial de télécommunications ALPAL II + SEA ME IV.

1.6 La structure organisationnelle de la DOT de Béjaia :

La Direction Opérationnelle des Télécoms DOT de Béjaia est placée sous l'autorité directe de la Direction Régionale Télécom (DRT) de Sétif. Son siège se situe au chef

lieu de la wilaya de Béjaia, dirigé par un directeur. Ce dernier est assisté pour le management des ACTEL (Agence Commerciale Télécom), CPT (Centre de Production des Télécommunications) et centres des télécommunications qui lui sont rattachés.

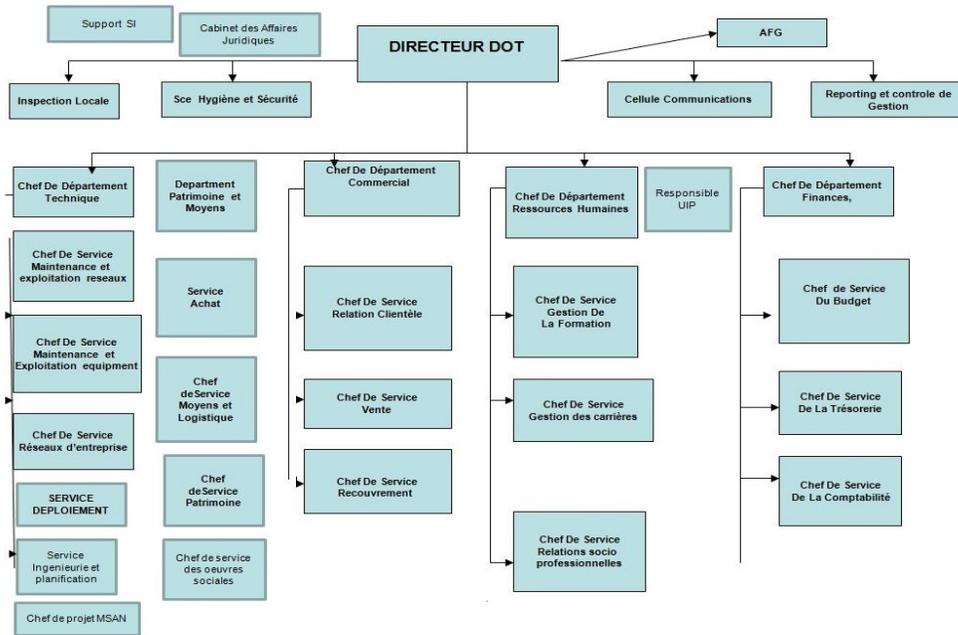


FIGURE 1.2 – La structure organisationnelle de la DOT de Béjaia.

1.7 Conclusion

Notre stage de fin cycle s'est déroulé au Centre d'Amplification Béjaia (CA Béjaia) qui présente un point stratégique très important qui réside dans la place qu'il comporte sur les réseaux des transmissions et d'amplification au niveau de la wilaya et sur le territoire national. En effet, il se trouve à quelque dizaines de mètres de l'établissement de la direction d'Algérie Télécom de Béjaia, du côté gauche de la Cité Tobal en prenant la route de Sétif. cet emplacement lui a permis d'être un noeud de rattachement de tous les centres de la wilaya avec les différents centres nationaux (Sétif, Alger, Tizi-Ouzou, ...).

Durant notre stage, on s'est intéressé au nouveau réseau multi-services d'AT, qu'on présentera dans le chapitre suivant.

Chapitre 2

Perspectives NGN d'Algérie Télécom

Introduction

Algérie Télécom a initié depuis le début des années 2000 un plan de développement conséquent pour accélérer l'usage des TIC dans notre pays. Le recours aux Réseaux de Nouvelle Génération (NGN), initié en 2007, a atteint aujourd'hui la phase finale de l'opération de " modernisation du réseau de télécommunication d'Algérie Télécom " .

Ce chapitre traitera les perspectives de déploiement d'un réseau multiservices (RMS), notamment le projet du réseau d'accès MSAN(Multi Service Access Node).

2.1 Nécessité de modernisation des réseaux de télécommunications

Le plan e-Algérie, comme de nombreuses initiatives prises au niveau mondial depuis la conférence de l'UIT (Union Internationale des Télécommunications) tenue à Minneapolis (États-Unis) du 12 octobre au 6 novembre 1998, vise à participer à la lutte contre l'exclusion sociale. L'objectif fondamental est de réduire la fracture numérique avec des mesures réalisables à l'horizon 2015.

De l'avis de nombreux spécialistes, nos réseaux de télécommunications ont atteint leur limite en matière de services. Ainsi, dans le cadre du plan e-Algérie, un effort considérable de modernisation des infrastructures est consenti dans notre pays, notamment un vaste programme d'installation de liaisons en fibre optique sur le réseau de transport (Backbone du réseau fixe). Malgré cela, les offres xDSL (DJAWEB, FAWRI...), restent encore en deçà des objectifs et l'adhésion des utilisateurs, mitigée jusqu'à présent, est freinée par divers problèmes techniques.

Ce plan e-Algérie a commandé la modernisation des réseaux radio-mobiles GSM des trois opérateurs présents sur le marché national (ATMobilis, WTA-Ooredoo,

OTA-Djezzy). Le lancement des systèmes de 3^{ème} génération a pour objectif de lever les limitations (en termes de débits) pour la transmission de données ou de connexion au réseau internet, des systèmes GSM qui sont réservés jusque là au service voix seulement.

Ce besoin de mutation vers de nouvelles technologies est pressant à cause d'une demande de plus en plus forte des usagés en raison de l'émergence de deux tendances :

- D'une part, les volumes de données augmentent, notamment en raison des éléments multimédia (son, vidéo) de plus en plus nombreux ;
- D'autre part, les applications inter-actives en temps réel se multiplient et vont se développer au cours des prochaines années, pour le grand public (téléphonie sur IP, sites web interactifs...) et pour les professionnels (e-santé, télé-travail...). Les besoins sont donc plus volumineux, mais exigent aussi d'être plus rapides.

L'enjeu de demain semble donc être indéniablement le très haut débit aussi bien sur le réseau fixe que sur les réseaux mobiles. Ainsi, il est indispensable de recourir aux RMS.

2.2 RMS d'Algérie Télécom

Le futur RMS d'AT est un réseau fédérateur d'envergure nationale à large bande de nouvelle génération. Il sera basé sur un backbone IP/MPLS (Réseau dorsal de transmission longue distance à fibre optique et à très grande capacité) inter-connectant tous les réseaux classiques existants, implémentant de nouvelles technologies d'agrégation et de redondance et assurant une multitude de nouveaux services aux abonnés à travers : des boucles métropolitaines et locales en fibre optique, un réseau d'accès filaire (MSAN, FTTx...) et sans fil (WiFi, WiMax...).

2.2.1 Topologie du réseau RMS d'AT

Le déploiement du réseau RMS est en phase d'achèvement malgré certains retards inhérents dus à l'envergure nationale et l'importance de ce projet de migration, à la diversité des équipements et des technologies supportées par le réseau traditionnel existant. Le réseau RMS est déployé sur le territoire national à travers :

- **Quatre Nœuds Primaires** : Alger I (couvrant la région centre), Oran (région ouest), Constantine (région est), Ouargla (région sud) ;
- **Cinq Nœuds Secondaires** : Alger II, Annaba, Sétif, Chlef, Oran II, Bechar.

Chacun des nœuds primaires accueille un couple de routeurs "Core" avec une puissance de commutation de 320 Gbps (gigabits par seconde) optimisés pour les services basés sur MPLS et prêt à supporter les interfaces de 10 Gbps. Ces routeurs sont interconnectés en SDH-STM-16 (10 Gbps) (en anglais Synchronous Digital Hierarchy-Synchronous Transport Module) par le biais du réseau de transport fibre optique DWDM dont les artères ont un débit porté à 110 Gbps, auparavant elles étaient à 80 Gbps.

Les nœuds secondaires, quant à eux, chacun d'eux accueille un couple de routeurs "Edge"

supportant une très large gamme d'interfaces allant des interfaces sériels de 64 kbps (Kilobits par seconde) jusqu'au niveau SDH STM-16.

Le réseau est équipé bien de Softswitchs MGC de nouvelle génération dans sa couche contrôle et de passerelles Media Gateway fonctionnant principalement comme médiateurs entre les réseaux traditionnels TDM, des plateformes services (plate forme SIP) et des réseaux d'accès large bande à base de noeuds d'accès multi-services MSAN. La plate-forme de gestion est centralisée à Alger.

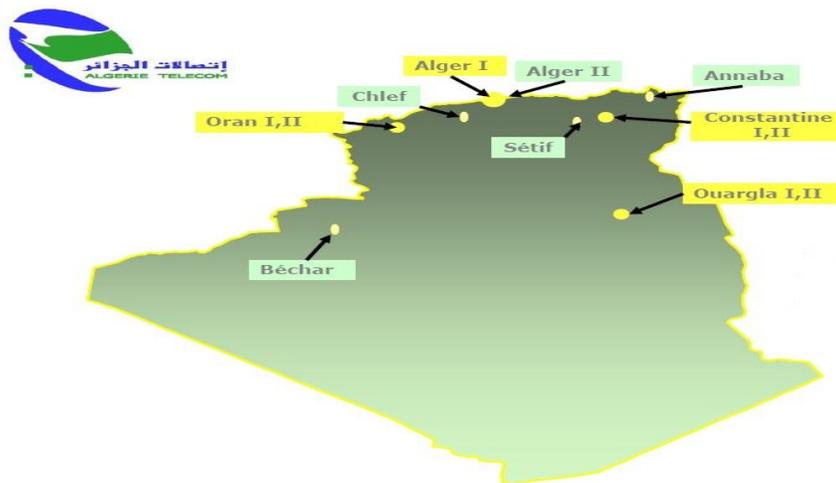


FIGURE 2.1 – Le réseau RMS déployé sur le territoire national.

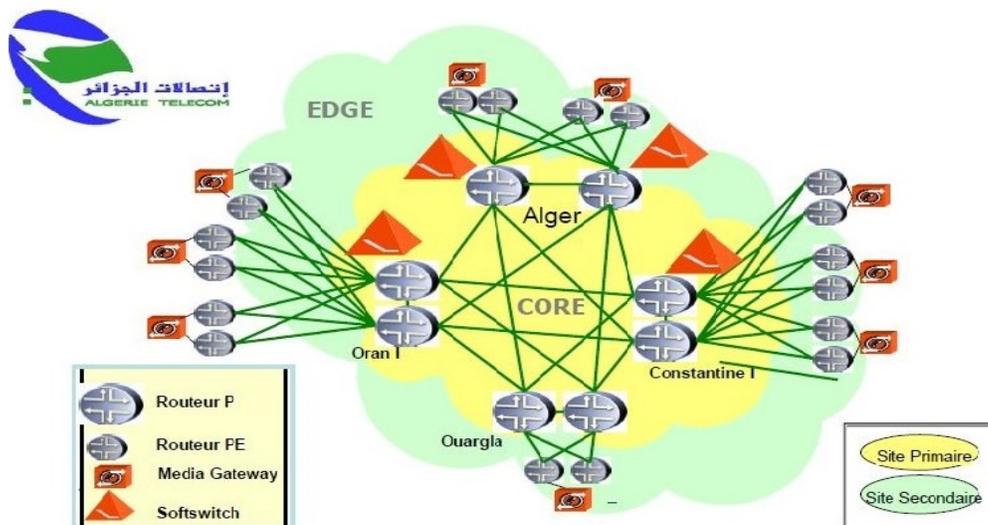


FIGURE 2.2 – Topologie globale du réseau RMS d'Algérie Télécom.

2.2.2 Objectifs du déploiement du réseau RMS d'AT

Les objectifs de la modernisation du réseau de télécommunications d'AT sont multiples :

- Préserver les services existants à travers un nombre d'interfaces important en remplaçant progressivement les solutions actuelles : Réseaux de données circuits et paquets (X21, X25), Liaisons Louées,...
- Élargir les capacités des réseaux téléphoniques avec souplesse ;
- Couvrir tous les besoins d'interconnexion à bas et hauts débits ;
- Garantir une qualité de service "QoS" de bout en bout ;
- Améliorer les performances du réseau global ;
- Économiser la bande passante : Solution idéale pour l'optimisation des réseaux TDM ;
- Supporter des protocoles variés de type Frame Relay, X25 et IP ;
- Solution de backbone pour le trafic internet ;
- Solution idéale de migration des réseaux X25 ;
- Optimisation de la Gestion de la voix ;
- Garantir la sécurité (VPN) et la disponibilité.

2.2.3 La Mise à niveau du Backbone national

Le Backbone national repose sur le réseau de transport en boucle DWDM reliant les grandes villes du pays, ramifié en plusieurs boucles régionales DWDM et boucles locales (FO et FHN). Beaucoup de chantiers de mise à niveau de ce réseau sont entrepris par Algérie Télécom dont :

- La mise à niveau du Réseau National DWDM : (80 à 110 Gbps pour le Nord et de 40 à 60 Gbps pour le Sud) parachèvement de la boucle du Grand sud ;
- La mise à niveau en DWDM des réseaux régionaux et des réseaux urbains ;
- La mise en oeuvre de nouvelles boucles régionales et boucles locales ;
- La finalisation du réseau de transmission SDH boucles 10 Gbps sécurisé ;
- La réalisation d'une plate forme de réseau intelligent ;
- L'intensification et la mise à niveau du réseau d'accès et de distribution filaire (MSAN, FTTx...) et sans fil (WiFi, WiMax...).
- L'implémentation de nouvelles technologies (IP) pour les FHN avec des accès GE ;
- L'introduction des équipements d'agrégation Ethernet.

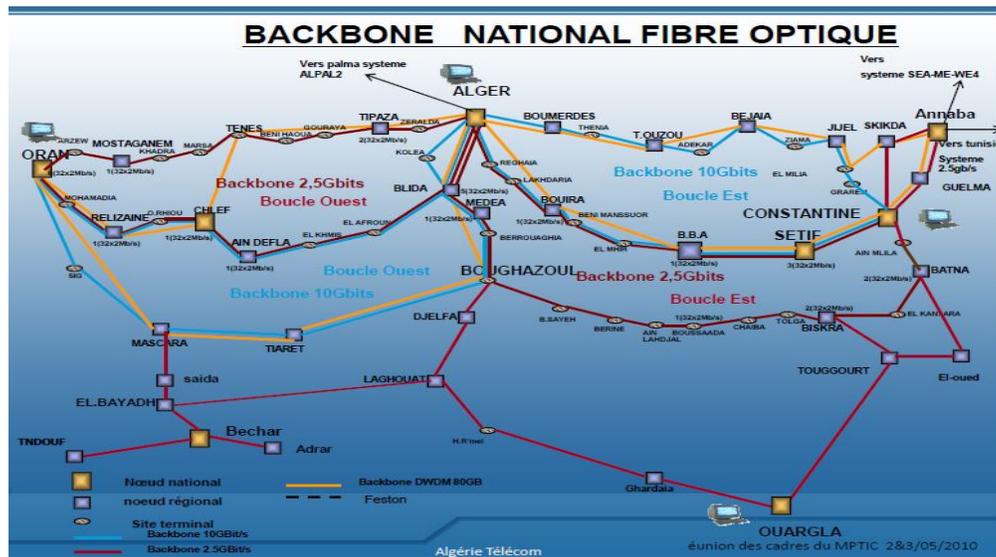


FIGURE 2.3 – Backbone national fibre optique

2.2.4 La Migration du réseau RTC

Le RTC (Réseau Téléphonique Commuté) est un réseau hiérarchisé donc ne permettant pas une évolution fonctionnelle. La plupart de ses équipements sont de technologie ancienne (TDM) délaissée actuellement par les constructeurs. Cela implique forcément une fin de support technique notamment pour les paliers logiciels. Puisque la tendance mondiale est aux réseaux NGN, Algérie Télécom est contrainte de moderniser son réseau TDM avec à la clé, une rentabilité assurée puisqu'elle deviendra un fournisseur multi-services. La stratégie de migration adoptée est une migration à la douce, i.e. graduelle

2.2.5 Présentation du réseau d'accès MSAN d'Algérie Télécom

Le réseau d'accès MSAN (MultiService Access Node) est un nœud d'accès multi-services. Cette technologie largement adoptée par le monde permet de réduire la boucle locale i.e. qu'elle rapproche les équipements des clients, ce qui autorise des débits élevés et intégrant l'ADSL et la voix ainsi que d'autres services comme la visio-phonie et la visio-conférence.

Donc Algérie Télécom a lancé son projet d'accès MSAN en 2007, en parallèle aux autres projets de migration vers le réseau RMS. Le plan e-algérie 2013 prévoyait 6 millions d'accès MSAN à la fin 2013 mais certains problèmes techniques retardent encore la finalisation du programme.

Le déploiement de ce réseau se fait en parallèle au réseau de commutation actuel qui continue à fonctionner, de ce fait Algérie Télécom procède à la migration progressivement

en 2 phases :

- **Migration des services large bande** : Le réseau actuel nous permet les services vocaux RTC et les services de données en haut débit sur une même paire de cuivre où la bande passante est partagée entre les deux réseaux : téléphonique et DSL grâce à l'usage de splitter (Appareil qui multiplie les signaux en plusieurs sorties). Donc l'installation de l'équipement MSAN sur le site, doit se faire en gardant l'architecture actuelle. Cette phase consiste à migrer du DSLAM (Multiplexeur d'accès DSL) au MSAN en passant par plusieurs étapes :

1. Installer le MSAN sur le site alors que le réseau commuté fonctionne toujours ;
2. Implanter au niveau du MSAN la fonction d'acheminement de données du DSLAM ;
3. Retirer le DSLAM.

Algérie Télécom continuera alors à utiliser son réseau RTC pour l'acheminement de la voix et le MSAN comme point d'accès au réseau IP. Le MSAN supportera alors les deux technologies en même temps le TDM et l'ATM/IP.

- **Migration des POTS** : La migration du service voix (POTS) se fait comme suit :
1. La première étape c'est d'ajouter des cartes POTS, les cartes RTC et DSL sont localisées dans le MSAN et la signalisation entre le MSAN et le commutateur RTC s'effectue via des interfaces V5, les nouveaux abonnés DSL seront raccordés à cette nouvelle plate forme pour les services vocaux et données ;
 2. La seconde étape est de mettre à niveau le MSAN pour qu'il devient un pur équipement IP assurant la transmission des appels vocaux RTC puis les convertir en VoIP ;
 3. La migration des abonnés existants se fait automatiquement vers la VoIP quoi que les services reçus restent de type RTC ;
 4. La dernière étape est faite lorsque la migration attire suffisamment d'utilisateurs et qu'Algérie Télécom soit prête, le reste des abonnés RTC seront transférés sur la nouvelle plate forme IP et enfin le réseau RTC sera abandonné.

2.3 Réseau de transmission de la Wilaya de Béjaia

Le réseau de transmission de la Wilaya de Béjaia est configuré tel qu'il est donné par la carte suivante :

une solution consiste à utiliser deux ordonnanceurs successifs : un premier ordonnanceur de type PQ qui sélectionne prioritairement les paquets de la classe EF par rapport aux autres classes, puis un second ordonnanceur de type WFQ (weight Fair Queuing) pour différencier les services AF1, AF2, AF3 et AF4. L'algorithme WFQ permet d'assurer à chaque service AF un minimum de temps de service, selon une pondération définie par l'opérateur.

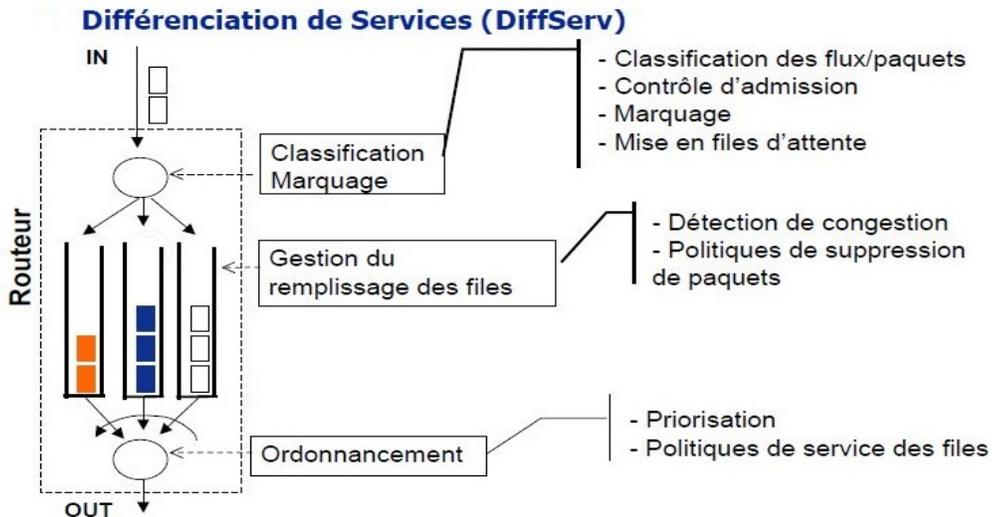


FIGURE 2.5 – Le modèle de file d'attente approprié à l'architecture du routeur de cœur.

Pour modéliser le comportement d'une file d'attente à serveur partagé, il faut mesurer les critères de QoS pour chacune des classes de service. Et pour caractériser le trafic entrant, il faut fournir une description du trafic de chaque classe de service.

Conclusion

Dans ce chapitre, nous avons présenté le réseau multi-services RMS d'Algérie Télécom. Nous avons expliqué pourquoi AT a consenti d'énormes investissements pour la modernisation de ses infrastructures de télécommunications et nous avons donné la stratégie de migration retenue pour son réseau de commutateurs RTC ainsi que le déploiement de la solution d'accès MSAN.

Chapitre 3

DiffServ et qualité de service

Introduction

Le plus grand objectif pour chaque réseau à mettre en place, que ce soit en télécommunications, en informatique ou autre, est d'avoir les meilleures performances. Pour cela, le taux de Qualité de Service (en anglais quality of service QoS) doit être au minimum au niveau des exigences des utilisateurs pour leurs entières satisfactions.

Puisque la tendance mondiale est aux réseaux RMS, Algérie Télécom est contrainte de moderniser son réseau avec à la clé, un taux de Qualité de Service à la hauteur des exigences de ces abonnés. La stratégie de migration d'AT vers la solution tout IP avec le RMS est une migration à la douce, i.e. graduelle. Parmi les services de ce nouveau réseau RMS d'AT la Visio-conférence (Vidéo temps réel), il est l'un des services les plus demandés actuellement par les abonnés particulièrement les universités, les entreprises, les établissements hospitaliers. Et pour gérer ce genre de service une bonne gestion de la QoS est obligatoire, alors pour cela l'IETF a définie deux modèles pour cette gestion de la QoS : DiffServ et IntServ

Le modèle IntServ est difficilement applicable dans le cas de grands réseaux surtout dans le cas de vidéo temps réel. En effet, il est difficile de maintenir un état de ressources réseau pour chaque flux si le chemin emprunté change fréquemment. Cela est du au fait que le routage est indépendant de la gestion de QoS. Le modèle DiffServ consiste à classer le trafic grâce à un code présent dans le paquet IP. On applique ensuite des traitements différenciés aux différentes classes de trafic. Nous avons donc affaire à une granularité moins fine mais qui devient en revanche plus scalable. En effet, la granularité du flot implique la réaction en chaîne suivante : plus il y a d'utilisateurs dans le réseau, plus il y a de flots, plus il y a de variables de classifications et d'ordonnancements dans les routeurs à maintenir, ce qui a pour conséquence une charge importante au niveau des routeurs qui deviennent alors de moins en moins performants. Ainsi, dans ce chapitre, nous allons présenter les différents paramètres de la QoS et les performances des réseaux IP et on s'intéressera par la suite

particulièrement à l'étude des performances d'un router de cœur DiffServ.

3.1 Généralités sur la Qualité de Service

3.1.1 Définition

La qualité de service (QoS) est la capacité de transmission dans de bonnes conditions d'un certain nombre de paquets dans une connexion entre un émetteur et un récepteur. Elle peut être présentée sous plusieurs termes tels que la disponibilité, le débit, les délais de transmission, la gigue, le taux de perte de paquets, etc.

Elle regroupe un ensemble de technologies mise en œuvre pour assurer des débits suffisants et constants sur tous les types de réseaux [11]

3.1.2 Objectif de l'étude de la QoS

L'objectif de l'étude de la QoS est d'optimiser les ressources du réseau et de garantir de bonnes performances aux applications. La qualité de service sur les réseaux permet d'offrir aux utilisateurs des débits et des temps de réponse différenciés par application suivant les protocoles mise en œuvre au niveau de la couche réseau. Elle permet ainsi aux fournisseurs de services de s'engager formellement auprès de leurs clients sur les caractéristiques de transport des données applicatives sur leurs infrastructures IP.

Selon les types de service envisagés, la qualité pourra résider :

1. Le débit (téléchargement ou diffusion vidéo) ;
2. Le délai (pour les applications ou la téléphonie) ;
3. La disponibilité (accès à un service partagé) ;
4. Le taux de pertes de paquets.

3.1.3 Paramètres de la QoS :

Au sein d'un réseau donné, la qualité de service est évaluée en fonction des différents équipements qui le composent, ainsi que du trafic qui y circule. Des applications multimédia telles que la voix-IP ou la vidéo à la demande, en plus des applications classiques, seront de plus en plus utilisées dans les réseaux RMS. Elles nécessitent un niveau minimal de qualité de service en termes de bande passante, de délai, de gigue ou de taux de pertes de paquets. Sachant que :

- **Le débit** : définit le volume maximal pouvant être atteint pour la transmission de l'information (bits) par unité de temps (s) dans une communication entre un émetteur et un récepteur.
- **La perte de paquets** : correspond aux octets perdus lors de la transmission des paquets. Elle s'exprime en taux de perte. Plutôt rare, Elle correspond au nombre de

paquet de données qui n'ont pas été reçu par la destination lors d'une communication. Cela peut être causé par plusieurs facteurs, la plupart du temps elle due à un encombrement du réseau.

- **Délai de transit (latence)** : est le délai de traversé du réseau, d'un bout à l'autre, pour un paquet. Les différentes applications présentes dans un réseau n'auront pas le même degré d'exigence, il est défini en fonction de leur nature : faible, s'il s'agit d'une messagerie électronique ou de fichiers échangés, ce degré d'exigence sera fort s'il s'agit de données "voix". La latence dépend du temps de propagation, du temps de traitement et de la taille des paquets.
- **La gigue** : désigne les variations de latence des paquets. La présence de gigue dans les flux peut provenir des changements d'intensité de trafic sur les liens de sorties des commutateurs. Plus globalement, elle dépend du volume de trafic et du nombre d'équipements sur le réseau.
- **La bande passante** : Il existe deux modes de disponibilité de la bande passante. Le mode "burst" est un mode immédiat, qui monopolise toute la bande passante disponible (lors d'un transfert de fichier par exemple). Le mode "streaming" est un mode constant, plus adapté aux fonctions audio/vidéo ou aux applications interactives.

3.1.4 Les informations utilisées pour assurer la QoS

Pour mettre en place la qualité de service, il est nécessaire de se positionner sur la reconnaissance des différents services qui sont les suivants :

- La source et la destination du paquet ;
- Le protocole utilisé ;
- Les ports de source et de destination dans le cas TCP et UDP ;
- La validité du routage (par exemple, gestion des pannes dans un routage en cas de routes multiples) ;
- La bande passante consommée ;
- Le temps de latence.

3.1.5 La gestion de la qualité de service :

Aujourd'hui le protocole IP tend à s'imposer dans la plupart des infrastructures réseaux, même dans les réseaux locaux des entreprises. C'est l'internet, bien sûr, qui a vu naître ce protocole, mais aussi les backbones des providers. L'autre tendance concerne l'arrivée imminente des applications multimédia dans les foyers comme dans les entreprises (visioconférence, VoIP, audio, vidéo, jeux en réseaux, etc). Or pour un fonctionnement normal et une satisfaction des utilisateurs, ces applications ont besoin de se positionner sur un réseau performant et fiable. Par exemple, les applications audios nécessitent le respect des délais de transmission strictes. Auparavant, le protocole IP était non fiable (il offre un service best-effort) et fonctionne en mode non connecté. Ceci signifie que les datagrammes émis par le protocole IP peuvent être perdus, peuvent arriver en désordre, peuvent être dupliqués, ou peuvent arriver avec un contenu comportant des erreurs. Par ailleurs, il n'y

a aucun mécanisme de contrôle de flux mis en oeuvre par IP. Le protocole IP considère que les protocoles des couches supérieures à savoir par exemple TCP, UDP ou SCTP corrigeront ses déficiences. La principale raison pour laquelle le mode non connecté a été préféré était de disposer d'un réseau qui puisse continuer à fonctionner même si une partie du réseau devient indisponible ou est détruite. Il s'agissait d'une exigence du DoD (United States Department of Defense). Donc, le protocole IP assurait le même traitement pour tous les paquets, qu'il voit transiter (notamment dans l'internet), sans tenir compte des spécificités des applications émettrices, ni de leurs contraintes (en bande passante, délai, etc). Cependant, depuis la fin des années 1980, l'IETF a lancé des travaux en vue de combler ce talon d'achille d'IP. Cela s'est traduit par la constitution de deux groupes de travail, qui ont mis au point deux modèles de gestion de la qualité de service qui sont :

- **Le mode DiffServ** : consiste à affecter des priorités au niveau d'un champ de l'entête IP appelé DSCP (DiffServ Control Protol). L'identification et le marquage du paquet est donc fait en entrée du réseau (les noeuds intermédiaires) et le cœur du réseau se contente d'appliquer les politiques de gestion du flux en fonction des priorités des paquets. L'approche DiffServ permet un déploiement et une exploitation simplifiée du réseau. DiffServ (Differentiated Services) est une approche permettant la gestion de la QoS sur un réseau IP [20].
- **Le mode IntServ** : Consiste à réserver les ressources nécessaires au niveau de tous les noeuds du réseau avant de faire transiter les flux. Cette approche repose sur l'utilisation du protocole RSVP (Ressource RéserVation Protocol) qui induit alors une couche de contrôle d'admission supplémentaire pour s'assurer que la bande passante requise est bien disponible à un instant T. Dans IntServ, les routeurs doivent maintenir des tables pour mémoriser l'état de chaque flux et des allocations de ressources.

Dans notre travail on s'est intéressé à l'approche DiffServ (le service différencié) qui est utilisé dans le réseau multiservice d'Algérie télécom pour les applications temps réel telles que la voix, la téléconférence, etc.

3.2 L'architecture DiffServ

3.2.1 Principe de fonctionnement de l'approche DiffServ

Le principe de DiffServ est d'associer aux différents flux des classes spécifiant le niveau de service pour lesquels sera réservé, l'information sur leur classe sera codée sous forme d'une étiquette associée à l'entête IP des paquets. DiffServ vise à avoir la scalabilité par l'agrégation des flux, et la distribution des fonctions entre les routeurs de bord du réseau et ceux du cœur du réseau. Les domaines implémentant la qualité de service DiffServ (appelés aussi les domaines DS) vont négocier entre eux les SLAs (Service Level Agreement) dans le but de se mettre d'accord sur les garanties à offrir aux différents flux transitant (selon leurs classes de service). Suivant ces SLAs, les paquets transitant d'un domaine à un autre seront examinés par les routeurs de bord qui vérifient leur conformément à l'accord (SLA) entre les deux domaines et peuvent ainsi modifier leur

niveau de service. Les principales activités relatives à l'inspection des paquets arrivants sont connues sous le nom de conditionnement du trafic (ou trafic conditionning). Ce conditionnement comprend la classification des paquets, le marquage, l'estimation du trafic et d'autres actions particulières pour les paquets classés out of profile, i.e. ceux qui ne respectent pas le profil spécifié pour leur classe de service (des actions d'élimination, de retardement ou de re-marquage, ...). Les routeurs de cœur de réseau, quant à eux, vont acheminer les paquets à leurs destinations tout en privilégiant les paquets possédant un niveau de service prioritaire (codé dans un champ appelé DS field). Pour se faire, ils associent à chaque paquet le niveau de traitement à subir qui est spécifiée par le traitement par nœud PHB (Per-hop-Behavior) utilisé dans le domaine.

La combinaison entre l'application d'un PHB dans le cœur du réseau d'une part et la classification et le conditionnement du trafic au niveau de ses frontières d'autre part permet bien la discrimination entre les trafics et par suite offre des niveaux de service qui diffèrent d'une classe à une autre.

3.2.2 La notion de domaine

L'internet est constitué d'une interconnexion a priori anarchique de réseaux. Cependant, plusieurs de ces réseaux sont souvent rassemblés sous une même autorité administrative (par exemple dans les grandes entreprises, les centres de recherches, les universités, ..).

On désigne par domaine, un ensemble de noeuds (hôtes et routeurs) administrés de façon homogène. Dans un domaine, on distingue les noeuds internes et les noeuds frontières : les premiers ne sont entourés que de noeuds appartenant au domaine alors que les seconds sont connectés à des noeuds frontières d'autres domaines [7].

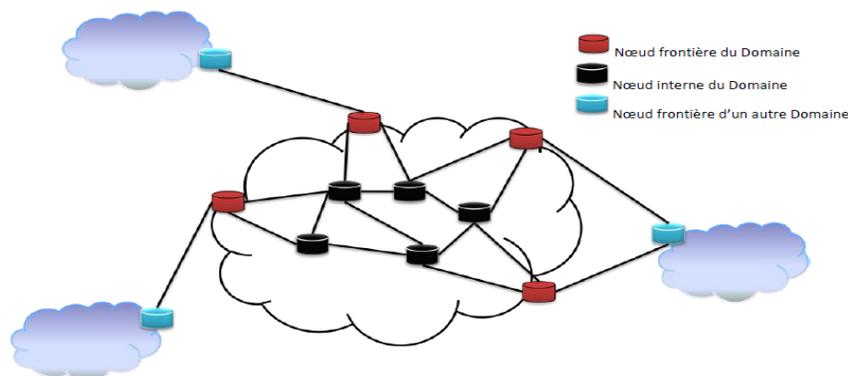


FIGURE 3.1 – Architecture générale d'un domaine DiffServ

3.2.3 Agrément du niveau de service

L'utilisation d'une architecture à services différenciés implique la signature d'un contrat avec le fournisseur d'accès Internet [5]. Ce contrat s'appelle agrément du niveau de service (Service Level Agreement :SLA), il est signé avant toute connexion au réseau et non à l'établissement d'une quelconque session.

Dans ce contrat sont données les informations suivantes :

- le trafic que l'utilisateur peut injecter dans le réseau fournisseur (en termes de volume de données, de débit moyen, d'hôtes sources ou destinations,...),
- les actions entreprises par le réseau en cas de dépassement de trafic (rejet, surtaxe, ...),
- la QoS que le fournisseur s'engage à offrir au trafic généré ou reçu par l'utilisateur (ou les deux). Celle-ci peut s'exprimer notamment en termes de délai ; de bande passante, de fiabilité ou de sécurité.

Après signature du SLA, l'utilisation des services DiffServ est transparente pour l'utilisateur, l'architecture ayant été conçue pour fonctionner avec les applications déjà existantes.

3.2.4 Le champ DS

Dans DiffServ, les paquets transportent eux même leurs classes dans un champ particulier de leur entête. Dans [17], le groupe de travail DiffServ a donné la définition de ce champ, il reprend les définitions existantes du champ ToS (Type of Service)des paquets IPv4 et le champ classe de trafic des paquets IPv6 pour les remplacer par un nouveau champ qui va contenir la définition de la classe de service du paquet. Ce champ est appelé le champ DS (pour Differentiated service).

La différenciation de service se base sur une étiquette introduite dans chaque paquet qui détermine le traitement que celui ci obtiendra du réseau. Ce concept n'est pas nouveau dans l'internet. Dès la définition du protocole IPv4, Le deuxième octet de l'en-tête IPv4(le champ ToS) était dédié aux informations de Qualité de Service. On le décompose en trois parties : le champ *Precedence* sur trois bits qui permet d'obtenir huit niveaux de priorités, suivi de quatre bits ToS décrivant le type de service qui contient notamment l'ancien champ DTR sur trois bits définissant les besoins en terme de délai, débit et fiabilité auxquels a été ajouté un bit pour exprimer le coût, suivi d'un bit inutilisé.

Pour IPv6, les sémantiques du champ classe de trafic ne sont pas encore bien précisées. Le modèle à différenciation de services aurait pu se baser sur l'utilisation effective du champ ToS, mais ceci ne satisferait pas aux attentes. D'ailleurs, les versions appelées "DiffServ" chez certains constructeurs utilisent encore cette sémantique. Les 8 bits qui forment le ToS ne sont pas utilisés de manière efficace, les services proposés sont plutôt liés à la sélection des routes et de nouveaux services ne peuvent pas être ajoutés. Donc, le besoin d'un nouveau modèle est justifié.

Le contenu du champ DS va définir le niveau de service du paquet à l'intérieur d'un

domaine DiffServ. Six bits de ce champ sont utilisés comme un code (DSCP : DS Code Point) qui va spécifier le niveau de traitement (ou PHB) que va subir le paquet dans chaque noeud du domaine DiffServ. Les deux autres bits ne seront pas utilisés pour le moment. Les routeurs peuvent par exemple comparer le contenu de DSCP à un tableau d'index, relatif au PHB implémenté dans le domaine, pour choisir les traitements à appliquer au paquet en question.

La FIGURE 3.2 décrit la définition du champ DS à partir des définitions existantes des deux champs définissant la qualité de service dans les paquets IPv4 et IPv6.

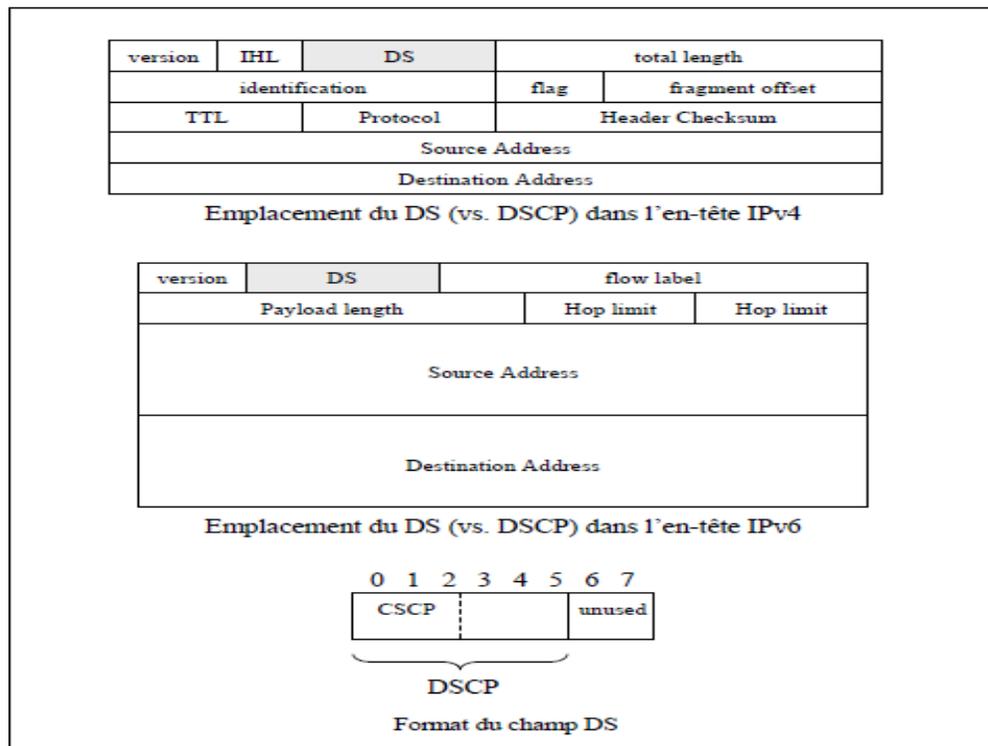


FIGURE 3.2 – Positionnement du champ DSCP dans les paquets IP

3.2.5 Structure d'un réseau DiffServ

La FIGURE 3.3 ci-dessous représente l'architecture de base d'un réseau selon l'approche DiffServ [23]. On remarque surtout la présence des conditionneurs de trafic (contrôleur-espaceur, marqueur et éliminateur) à la frontière du réseau. C'est dans le but de résoudre les problèmes de scalabilité (mise en échelle), que nous limitons la complexité au niveau des routeurs de bord, c'est là que les paquets doivent être analysés pour vérifier la nature de trafic, contrôler son accès et décider du niveau de service à lui offrir. Ainsi, les marqueurs doivent remplir le champ DS de chaque paquet entrant dans le réseau. Les routeurs du cœur du réseau vont donc utiliser la valeur de ce champ pour décider du niveau de service à offrir à chaque paquet selon le PHB supportée. Il s'agit donc d'appliquer une politique de

différenciation de trafic dans le traitement des paquets au niveau des files d'attente. L'utilisation à ce niveau des mécanismes de gestion dynamique des files d'attente permettra de privilégier les trafics ayant le champ DS plus prioritaire.

- **Routeur d'entrée :**

La structure d'un routeur d'entrée peut être décomposée en quatre modules comme le présente la FIGURE 3.4. Toute la complexité de l'architecture est concentrée dans ce type d'équipements [6]. Ces derniers doivent effectuer une classification qui peut aller jusqu'au niveau du microflux. Une fois le flux identifié, il est nécessaire de vérifier la conformité au contrat en utilisant un mécanisme adapté au service demandé. Ensuite une action est prise pour pénaliser les paquets non conformes. Finalement, avant d'injecter les paquets dans le réseau, la valeur du DSCP est actualisée en fonction du conditionnement subi.

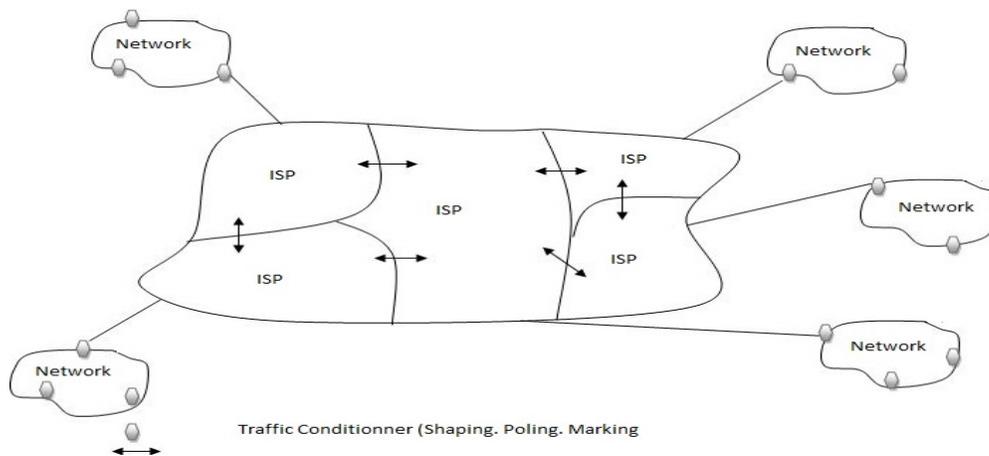


FIGURE 3.3 – Structure d'un réseaux selon l'approche DiffServ.

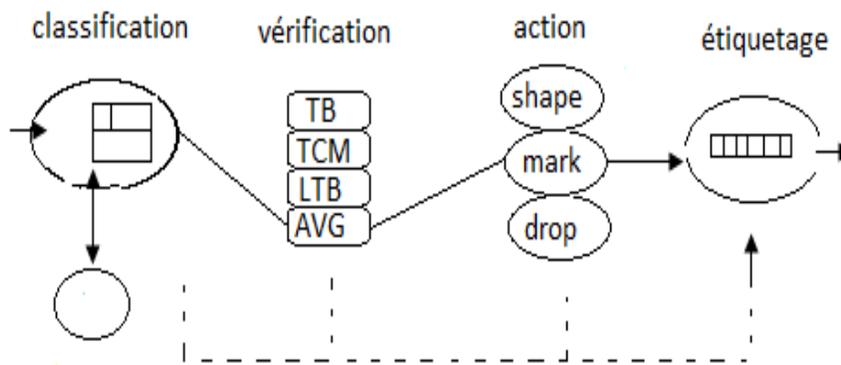


FIGURE 3.4 – Routeur de bordure.

a. Classification “classifier” :

Pour que les différents flux produits par les utilisateurs puissent être traités d'une manière différenciée, le routeur doit tout d'abord effectuer une classification des paquets. Cette opération est très similaire à celle définie pour l'architecture IntServ. Elle se base toujours sur des champs de l'entête, mais le nombre de champs utilisé pour l'identification peut varier en fonction du contrat. Dans le modèle DiffServ, la classification peut aller jusqu'au niveau des microflux. La classification au niveau des microflux s'effectue dans les points où les flux entrent dans le réseau. Elle peut également avoir lieu dans un module de contrôle installé dans l'équipement de l'utilisateur. Dans ce cas, l'utilisateur est chargé de faire la mise en correspondance entre les applications et les classes de services. Le trafic est alors conditionné en tant qu'agrégat par le routeur d'entrée du fournisseur. La finesse de classification dépend du placement des routeurs d'entrée dans le réseau. Plus un routeur est proche du cœur, moins la classification est fine. Les paquets, dans ce cas, peuvent être classifiés et conditionnés, par exemple, en fonction de leur interface d'arrivée. Les routeurs du cœur du réseau réalisent une classification beaucoup plus simple, basée uniquement sur la valeur du champ DSCP, ce qui ne demande pas beaucoup de puissance de calcul ni de mémoire.

b. Vérification “meter” :

Un vérificateur est chargé de déterminer le niveau de conformité pour chaque paquet du flux arrivant dans le routeur. Ce niveau de conformité dépend du comportement instantané du flux et des caractéristiques du contrat (SLA). Le nombre de niveaux varie en fonction du conditionnement requis par le service :

- Pour des services à bas délai, il est nécessaire que seuls les paquets conformes au contrat soient acceptés dans le réseau. Dans ce cas, une vérification à deux niveaux est suffisante : conforme ou non-conforme, IN ou OUT.
- Pour des services profitant de la capacité d'élimination sélective dans le cœur du réseau, un nombre supérieur à deux augmente la granularité de la différenciation. Dans la définition du comportement assuré, trois niveaux de priorité ont été définis : vert, orange et rouge.

Le concept de conformité d'un flux est généralement associé au débit généré par celui-ci. Un contrat de service doit contenir le débit maximal accepté pour chaque classe de service. Pour des services ayant besoin d'une vérification à deux niveaux, la méthode la plus simple consiste à mesurer son débit moyen. Cette mesure peut être difficile à définir car la valeur obtenue dépend de la fenêtre d'observation utilisée. Un autre mécanisme, simple et pourtant efficace, est utilisé dans l'architecture DiffServ : le seau à jetons. Afin d'obtenir trois niveaux de conformité, des mécanismes plus complexes que le seau à jetons sont nécessaires. Pour les services se basant sur le comportement assuré, cette opération est cruciale à la définition de services car elle est directement liée à la distribution de priorités au sein d'un flux (Voir la sous section (3.2.6)).

c. Actions “dropper shaper” :

L'action corrective à prendre pour les paquets non conformes d'un flux varie en fonction du service. Trois types de pénalisation peuvent être identifiés : élimination, mise en forme et marquage :

1. L'élimination est sans doute l'action la plus sévère, mais nécessaire au bon fonctionnement de certains services. Pour un flux dont les paquets non conformes sont éliminés à l'entrée, il peut être souhaitable de contrôler le débit d'émission. Des flux interactifs contenant des contraintes temporelles fortes pourraient tirer profit de ce type d'actions. Dans un service DiffServ interactif, les fournisseurs devront approvisionner leurs réseaux en fonction des contrats établis. L'entrée dans le réseau des paquets au-delà du contrat peut faire croître les files d'attente dans les routeurs intermédiaires et provoquer une augmentation dans le délai de transmission qui rendrait l'inter-activité impossible.
2. La mise en forme consiste à retarder, si nécessaire, l'acheminement d'un flux pour le rendre conforme. Des services faisant appel à cette fonction peuvent être proposés aux clients qui n'ont pas les moyens de l'effectuer dans leur site. La mise en forme est une action très coûteuse en termes de ressources pour le routeur d'entrée, qui devra garder dans un tampon tous les paquets retardés jusqu'à ce qu'ils deviennent conformes.
3. Le marquage est l'action qui attribue une précedence ou priorité aux paquets en fonction du résultat de la vérification. Cette action introduit un deuxième niveau de différentiation, puisque les paquets sont acheminés en fonction de leurs classes de service et de leurs priorités. Les fonctions de vérification et de marquage sont parfois indivisibles. Elles forment le noyau des services basés sur le comportement assuré où le traitement que le réseau accorde à un flux dépend de la distribution des priorités à ses paquets.

d. Etiquetage "marker" :

Avant d'entrer dans le réseau, le champ DSCP de tous les paquets qui traversent le routeur d'entrée est mis à jour. La valeur de cette étiquette est le résultat des opérations de conditionnement. Si par définition il n'est pas possible de décomposer le DSCP en deux, il est tout de même possible d'identifier les deux informations qui établissent la différentiation : la classe de service et la priorité de paquet. Le DSCP qui forme l'étiquette DiffServ ne doit pas être modifié par les routeurs du cœur du réseau. Les routeurs de frontière peuvent quant à eux, modifier cette valeur lors des nouveaux conditionnements ou d'un problème de sémantique entre deux domaines voisins. Les différentes combinaisons de vérifications et d'actions forment en grande partie un service. L'autre partie est formée par le choix du comportement dans le cœur du réseau et par la sélection d'une classe de service. Ce paramètre sera exploité dans les routeurs du cœur du réseau. Étant donné que ces équipements ne doivent réaliser que des opérations simples et qu'ils doivent tous être conformes au standard, ce sont les modules du routeur d'entrée qui pourront plus facilement évoluer pour offrir des nouveaux services dans le futur.

• Routeur du cœur du réseau :

A la suite des traitements effectués dans le routeur d'entrée, les paquets sont injectés dans le réseau. Les opérations propres aux routeurs du cœur du réseau sont relativement

simples, mais celles-ci vont déterminer les performances du réseau, en termes de pertes et de délai observés par les sites terminaux [6]. L'un des objectifs du groupe DiffServ est de standardiser le traitement des paquets dans ces routeurs en fonction du champ DS. Tous les paquets portant une même valeur dans ce champ doivent être traités de la même manière. Le traitement différencié dans le cœur d'un réseau DiffServ implique deux axes : les classes de service et la notion de priorités. La FIGURE 3.5 montre l'architecture logique d'un routeur capable d'offrir deux niveaux de différenciation. Le routeur cœur est composé d'un classificateur, de plusieurs files d'attentes, d'un mécanisme d'ordonnancement et d'un algorithme de gestion de file d'attente pour chaque file.

Le classificateur que l'on trouve dans ce type de routeurs réalise une classification beaucoup moins coûteuse que les routeurs qui se trouvent à l'entrée du domaine. Dans ce cas, les paquets sont différenciés uniquement en fonction de l'étiquette qu'ils portent, i.e. de la valeur du champ DS attribuée par le routeur d'entrée. Une partie de cette étiquette sert à identifier la file d'attente sur laquelle le paquet doit être inséré, pendant qu'une autre partie de celle-ci peut être utilisée par l'algorithme de gestion de la file d'attente pour améliorer la discrimination en cas de congestion. Chacune des files d'attentes dans les routeurs représente une classe de service. Leurs propriétés, en termes de bande passante ou de retard moyen observé, dépendent du mécanisme d'ordonnancement. Les contraintes des comportements à bas délai comme le comportement expédié, exigent une technique plutôt hybride, utilisant une combinaison de files prioritaires et des mécanismes de partage de ressources comme le WFQ ou le WRR.

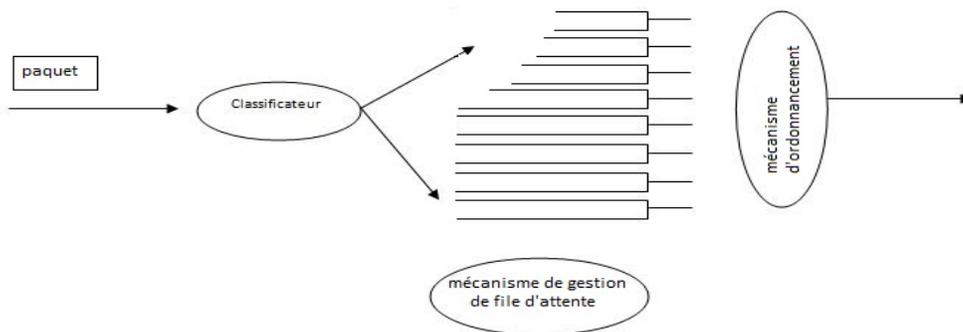


FIGURE 3.5 – Architecture d'un routeur de cœur.

A l'intérieur de chaque file, un mécanisme de gestion doit décider de la manière dont les paquets seront éliminés en cas de congestion. Par exemple, pour une classe de service à bas délai, qui ne devrait jamais observer les pertes (théoriquement), la traditionnelle méthode FIFO peut être utilisée. Par contre pour les files correspondantes aux classes AF (Assured Forwarding) [2], le standard demande la mise en œuvre d'une méthode capable de discriminer les paquets en fonction de leur précedence. Des méthodes comme WRED ou RIO peuvent être utilisées pour assurer cette fonction.

Il a été mentionné que le DSCP exprime le comportement par noeud (PHB : Per-Hop

Behavior) qui sera accordé au paquet. Un PHB est, d'après [17], la description des caractéristiques d'acheminement qui seront observés par tous les paquets comportant le même DSCP. L'utilisation d'un PHB ou d'un groupe de PHBs ajoutés aux opérations de conditionnement réalisés à l'entrée d'un domaine forment ce qui peut être appelé un service DiffServ.

3.2.6 Techniques de Lissage du trafic (Traffic Shaping)

L'une des causes principales de la congestion dans le réseau provient du fait que le trafic est aléatoire et très souvent sous forme de données en rafales. Si les ordinateurs sources pouvaient fournir leurs données à un rythme uniforme, le phénomène de congestion serait beaucoup moins fréquent. Il existe une technique puissante de contrôle de trafic qui permet d'améliorer sensiblement le traitement de la congestion en s'efforçant de maintenir le trafic le plus constant possible dans le réseau. En effet, le contrôle de trafic est un élément crucial pour la garantie de la qualité de service, et l'utilisation de lissage de trafic à ce niveau permet de contrôler l'évolution du trafic afin d'assurer la conformité des flux aux caractérisations énoncées dans leurs contrats de trafic (négociés entre les utilisateurs et le fournisseur du réseau), et éventuellement d'appliquer des corrections en cas de dépassement. L'utilisation de cette technique de lissage dans une architecture telle que DiffServ est très utile pour augmenter le ratio des paquets associé à la classe de service prioritaire. Il existe deux techniques très connues de lissage de trafic : leaky bucket et token bucket [23] :

- **Seau perméable (Leaky bucket) :**

Pour présenter le principe de la technique leaky bucket, il faut imaginer un seau d'eau dont le fond est percé d'un trou par lequel l'eau s'écoule en goutte à goutte (voir FIGURE 3.6), peu importe à quel débit l'eau arrive dans le seau, son écoulement par le trou se fait à vitesse constante tant qu'il y a de l'eau dans le seau. Mais lorsque le seau est plein de surplus d'eau arrivant est perdu. Ce principe peut facilement être transposé à la transmission dans un réseau, en utilisant un tampon de taille limitée qui va accepter les paquets des différents flux et les émettre à un débit constant, mais quand il sera plein, il va détruire tous les paquets arrivants [23].

L'algorithme de leaky bucket peut être utilisé comme un mécanisme de contrôle et de conditionnement de trafic dans les réseaux dans le but de réguler l'écoulement des flux et les contrôler conformément à ce qu'on veut réaliser.

- **Seau symbolique (Token bucket) :**

La technique token bucket est un cas particulier de leaky bucket, elle peut être représentée par un tampon de taille b jetons. La génération de ces jetons s'effectue à un rythme constant de r jetons par second (FIGURE 3.7). Chaque paquet arrivant consomme un nombre de jetons proportionnel à sa taille en octet. Pour qu'un flux soit conforme à un

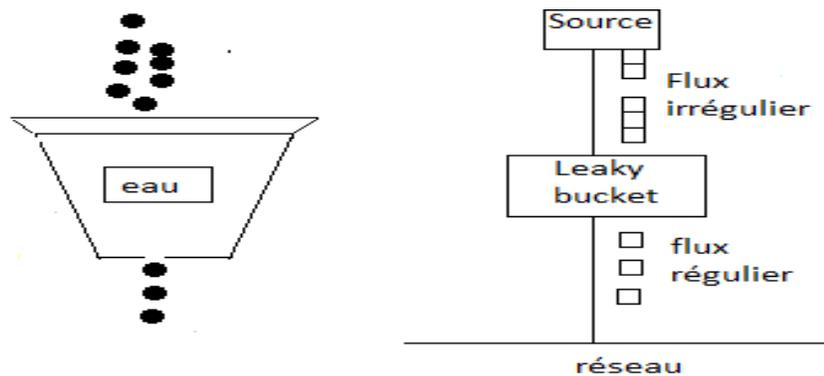


FIGURE 3.6 – Principe de la technique Leaky bucket.

token bucket $TB(r, b)$, il faut que tous ses paquets disposent de suffisamment de jetons à leur arrivée au tampon [23].

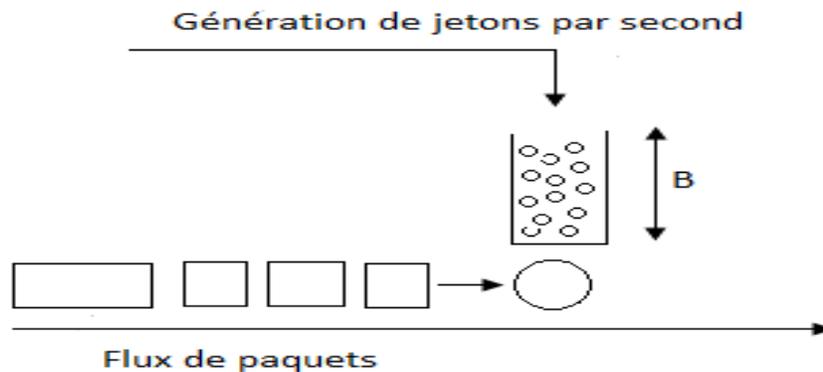


FIGURE 3.7 – Principe de la technique Token bucket.

3.2.7 Per-hop-Behavior standardisés

Les routeurs DiffServ traitent les paquets en fonction de la classe codée dans l'entête IP (champs DS) selon un comportement spécifique qui est le PHB. Chaque ensemble de paquets défini par une classe reçoit alors un même traitement, et chaque classe est codée par un DSCP (DiffServ Code Point). Un PHB est défini par les priorités qu'il a sur les ressources par rapport à d'autres PHB. En aucun cas, les routeurs ne traiteront différemment des paquets de même PHB et de sources différentes. L'avantage de DiffServ est qu'il n'y a plus nécessité de maintenir un état des sources et des destinations dans les routeurs, d'où une meilleure scalabilité.

DiffServ définit quatre PHB ou classes de service : **Best effort (priorité basse)** ; **Assured Forwarding (AF)** ; **Expedited forwarding (EF)** ; et **Default Forwarding (DF)**.

Cette notion de PHB permet de construire une variété de services différenciés. Les PHB

sont mis en oeuvre par les constructeurs dans les routeurs en utilisant des mécanismes de gestion de files d'attente (Custom Queuing, Weighted Fair Queuing,...) et de régulation de flux.

- **Les Classes de service :**

- **Best Effort (Le Meilleur Effort) [12] :**

Le principe du best effort se traduit par une simplification à l'extrême des équipements d'interconnexion. Quand la mémoire d'un routeur est saturée, les paquets sont rejetés. Le principe de bout en bout de l'internet est aussi adopté pour le contrôle de flux grâce à différents algorithmes comme le congestion avoidance introduit dans TCP.

Les principaux inconvénients de cette politique de contrôle de flux sont un trafic en dents de scie composé de phases où le débit augmente puis est réduit brutalement et une absence de garantie à long terme .

- **Service garantie (Expedited Forwarding) [13] :**

La classe Expedited Forwarding EF correspond à la valeur 101110 pour le DSCP et son objectif est de fournir un service de transfert équivalent à une ligne virtuelle dédiée à travers le réseau d'un opérateur. Le contrat porte sur un débit constant. Les paquets excédentaires sont lissés ou rejetés à l'entrée pour toujours rester conforme au contrat. L'opérateur s'engage à traiter ce trafic prioritairement. Pour que le service soit performant, il faut qu'il ne présente qu'une faible partie du trafic total pour qu'aucun paquet marqué EF ne soit rejeté dans le cœur du réseau. Pour atteindre ces performances, les paquets d'un service EF ne devraient pas subir de file d'attente ou passer par des files de très petite taille et strictement prioritaires. De plus, les flux ne doivent avoir que très peu de perte, la gigue doit être minimale et la bande passante garantie. D'une part, cela nécessite la mise en place d'un contrôle d'accès et d'autre part, cela impose qu'à chaque noeud traversé, le taux maximale de trafic d'arrivée doit être inférieur au taux minimal de trafic de départ. cette dernière assertion implique que, dans les noeuds internes, une bande passante minimale est disponible au service EF et que, dans les noeuds d'extrémité, un trafic conditionné est effectué. Ce mécanisme de conditionnement est utilisé pour vérifier la conformité des flux utilisateurs. La conformité du trafic EF et AF par rapport à leurs profils est déterminée pour chacun par un token bucket. Dans ce cas, la taille et le débit du bucket sont à spécifier. Les paquets EF non conformes sont détruits tandis que les paquets AF non conformes sont marqués pour être jetés en cas de congestion.

- **Service assuré (Assured Forwarding) [2] :**

La classe Assured Forwarding définit trois priorités définissant l'ordre de rejet dans un routeur en cas de congestion. Les priorités sont représentées par trois couleurs qui dépendent de la conformité de la source avec son contrat. Le marqueur utilisé actuellement est basé sur deux token buckets : si le trafic est conforme au deux, les paquets sont marqués en vert, s'il n'est conforme qu'à un des deux les paquets seront marqués en orange et s'ils ne sont conformes à aucun, ils seront marqués en rouge. Les paquets marqués en rouge, ont une probabilité de rejet plus importante que les oranges. Dans le cœur du réseau, les mécanismes de rejet sont basés sur RIO. Chaque couleur dispose d'un seuil et d'une proba-

bilité de rejet différents. Quatre classes AF sont disponibles. Il n'y a pas de priorité parmi ces classes.

Les travaux du groupe DiffServ sont surtout basés sur une approche d'ingénierie portant sur la définition de l'architecture et sur la gestion de l'octet DS. Les résultats de simulation montrent que RIO est très sensible aux conditions initiales du trafic, en particulier, la proportion de trafic UDP influe énormément sur les résultats.

Les objectifs de la différenciation par l'Assured Forwarding sont l'attribution différenciée des ressources et la protection des flux TCP, et les flux UDP.

Pour l'Assured Forwarding, on définit quatre classes de services et trois priorités que l'utilisateur respecte sur son contrat. Les classes sont donc choisies par l'utilisateur et restent les mêmes tout au long du trajet dans le réseau. Tous les paquets d'un même flux appartiennent à la même classe.

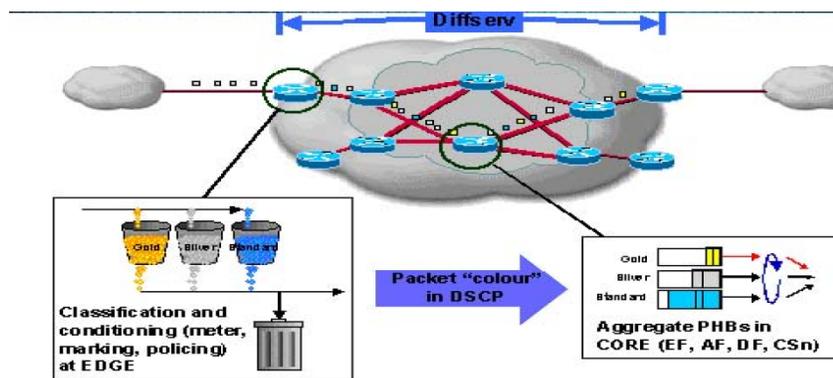


FIGURE 3.8 – Arrivée des paquets dans un edge router puis dans un core router

A l'intérieur de chaque classe, un algorithme de rejet sélectif différencie entre trois niveaux de priorités. En cas de congestion dans une classe AF, les paquets de base prioritaire sont rejetés en premier. La priorité peut être modifiée dans le réseau par les opérateurs en fonction du respect ou non des contrats.

• **Avantages de l'Assured Forwarding** : Parmi les avantages de AF on a :

- AF peut offrir une meilleure différenciation (classe et priorité) ;
- le marquage à l'entrée du réseau est une opération moins coûteuse que le shaping ;
- AF ne demande pas une coordination entre domaines ;
- une facturation simple peut être utilisée.

Pour obtenir une bonne différenciation avec l'Assured Forwarding, il faut :

- un regroupement des flux similaires dans une même classe ;
- des mécanismes de marquage adaptés ;
- des variations du débit, rafales, etc ;
- un marquage en fonction du résultat que l'on veut obtenir.

• **Default Forwarding (DF)** : utilisé uniquement pour les flux internet qui ne nécessitent pas un trafic en temps réel.

3.2.8 Gestion active des files d'attente

Le mécanisme de gestion de tampon (buffer management) consiste à éliminer des paquets en cas de congestion du tampon d'un port de sortie d'un routeur. L'élimination est réalisée de manière sélective, en fonction des critères de QoS des trafics. Ainsi, dans le cas de DiffServ, les algorithmes de gestion de tampon se basent sur le PHB (Per Hop Behavior) qui spécifie le comportement à adapter en fonction de la classe du trafic [9].

Traditionnellement, la politique FIFO (First In, First Out) est utilisée dans les routeurs. Dans ce cas, si le débit d'arrivée est supérieur au débit de sortie, les paquets s'accumulent dans la file jusqu'à ce qu'elle atteigne sa taille maximale. A partir de ce moment, tous les paquets arrivants sont éliminés. Si cette méthode est la plus simple à mettre en oeuvre, elle a aussi plusieurs défauts. D'abord, elle peut être très inéquitable : des sources produisant des rafales importantes risquent d'être plus pénalisées que des sources quasi-constantes. Ensuite, il se peut qu'un grand nombre de paquets soient éliminés consécutivement. Enfin, des files remplies en permanence se traduisent par une augmentation du délai observé par les paquets. Pour résoudre les problèmes liés au FIFO, plusieurs algorithmes de gestion de tampon ont été proposés pour les réseaux à QoS : WRED, RIO, FRED. Ils sont basés sur l'algorithme RED (Random Early Detection/suppression pécore aléatoire).

- **Random Early Discard (RED)**

Le Random Early Discard (RED) proposé par **Sally Floyd** [10] est un mécanisme de gestion de file d'attente très efficace qui est devenu très populaire. En utilisant un algorithme relativement simple, RED a pour but de réduire le taux d'occupation d'une file d'attente tout en restant équitable sur la distribution des pertes. RED s'attaque principalement à la congestion persistante dans une file d'attente. Sa fonction d'élimination est basée sur la taille moyenne de la file, ce qui autorise des fluctuations instantanées nécessaires pour le traitement du trafic très variable. Dans sa définition, RED n'impose aucune formule pour le calcul de l'occupation moyenne.

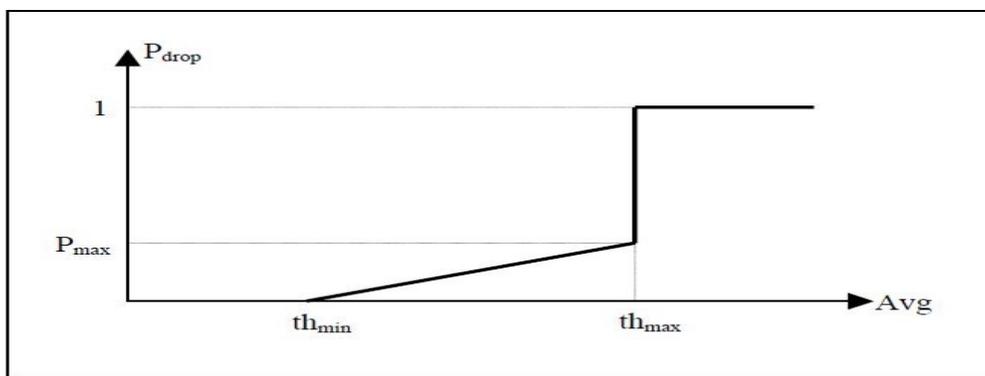


FIGURE 3.9 – Paramètres de RED.

RED utilise une fonction aléatoire croissante pour décider si un paquet doit être éliminé.

Grâce à cette caractéristique, les éliminations successives sont évitées. Le rejet aléatoire est une solution équitable puisque la probabilité pour qu'un flux subisse une perte est directement proportionnelle au pourcentage d'occupation de ce flux dans la file. Dans RED, l'élimination aléatoire se réalise uniquement quand la taille moyenne de la file se trouve entre deux seuils th_{min} et th_{max} . p_a représente la probabilité de rejet du paquet. Elle augmente linéairement de 0 à p_{max} pendant que la moyenne varie de th_{min} à th_{max} comme le montre la FIGURE 3.9. A partir de th_{max} , tous les paquets sont éliminés. Si les sources réduisent leur débit en conséquence, la moyenne revient rapidement à un niveau acceptable.

3.2.9 Mécanismes d'ordonnancement

Le traitement différencié de paquets dans le modèle DiffServ peut être décomposé en deux axes : regroupement de flux en classes de service et introduction de priorités au sein de flux. Plusieurs techniques ont été développées pour contrôler le partage de ressources, pour isoler des classes de service ou encore pour réduire le temps d'attente des paquets dans les files [21].

• Le Fair Queueing (FQ) ou Round Robin (RR) :

La discipline de service "à partage équitable" (Fair Queueing) représente le travail initiateur pour le développement des techniques d'ordonnancement à partage équitable de bande passante. L'objectif est de fournir équitablement le même taux de service aux différents flux partageant les ressources. Dans la discipline FQ, les paquets sont classés par flux par le système et insérés dans une file d'attente particulièrement dédiée à chaque flux. Les flux sont servis en tourniquet (Round Robin) paquet par paquet comme présenté dans la FIGURE 3.10. Les files d'attente vides sont sautées par l'ordonnanceur et revisités au tour suivant. La discipline de service FQ permet de résoudre le problème de la famine rencontré avec un serveur à priorités statiques. De même, elle fournit un niveau de service minimum à chaque flux indépendamment du comportement des autres flux. Cependant, malgré ses avantages, l'algorithme FQ présente plusieurs limitations qui freinent son déploiement pour la fourniture de garanties strictes :

- La discipline de service FQ n'est pas destinée à servir des flux ayant des besoins différents en terme de bande passante. Son objectif se limite à fournir la même portion de la bande passante à tous les flux.
- La garantie équitable du même taux de bande passante à tous les flux n'est possible que lorsque tous les paquets ont la même taille. En effet, les flux de paquets de large taille auront plus de bande passante que les flux de taille de paquets plus petite.
- La discipline de service FQ est sensible à l'ordre d'arrivée des paquets. En effet, si un paquet arrive à une file d'attente vide immédiatement après avoir été visitée par l'ordonnanceur, alors ce paquet doit attendre la fin de service de tous les paquets des files d'attente avant d'être transmis.

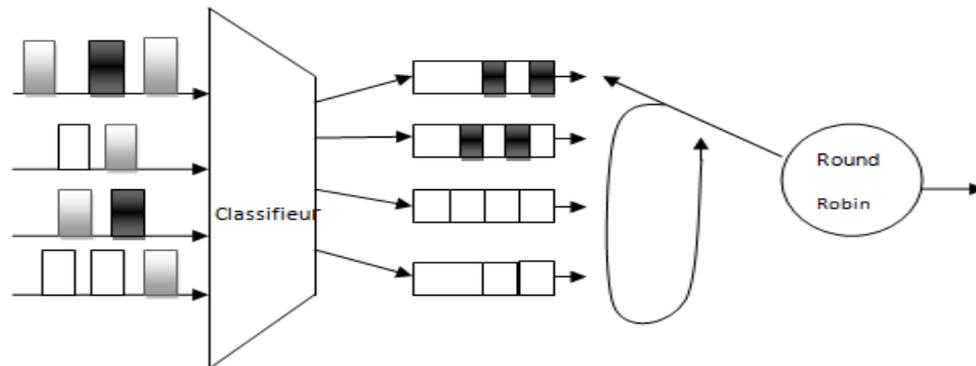


FIGURE 3.10 – La discipline de service Round Robin.

- **Weighted Fair Queueing (WFQ) :**

Pour résoudre les problèmes du FQ, un service en tourniquet bit-à-bit a été proposé, qui est un modèle théorique, pour le partage équitable de la bande passante. De plus, la notion du poids a été introduite pour pondérer le service proportionnellement à la bande passante exigée par le flux. Cette proposition est appelée Weighted Fair Queueing (WFQ).

3.3 Conclusion

Nous avons remarqué dans ce chapitre que dans le modèle DiffServ, le rôle des routeurs varie en fonction de leur place dans le réseau. Les routeurs d'entrée sont chargés de l'identification et du conditionnement des flux. Le résultat de ces actions est inscrit en tant qu'étiquette dans l'en-tête de chaque paquet. Dans le cœur du réseau, cette étiquette est le seul paramètre utilisé pour déterminer le comportement des routeurs vis-à-vis des flux qui les traversent. La QoS dépend en grande partie de l'algorithme chargé de l'attribution de priorités.

Chapitre 4

Rappels et modélisation du problème

Introduction

L'objectif principal de notre travail étant l'évaluation de performances des réseaux à différenciation de service, et spécialement l'évaluation des critères de qualité de service. Il s'agit d'un problème transversal aux différents mécanismes de gestion d'un réseau. Pour être capable de garantir la qualité de service de chaque client, il faut en effet être apte d'en évaluer les différents critères de QoS. Ce problème est résolu de deux manières. La première manière est analytique, elle se base sur des modèles stochastiques tels que les systèmes de files d'attente qui permettent d'obtenir des expressions analytiques (parfois exactes) des critères de performances sous des hypothèses relativement fortes. La seconde est la simulation qui a un pouvoir d'expression beaucoup plus important (à priori sans limite) mais se heurte à des temps de calcul prohibitifs.

Dans ce chapitre, on présentera l'évaluation de performances de routeur de cœur de l'architecture DiffServ. En effet, on va d'abord donner un aperçu sur les systèmes de files d'attente et plus précisément les files d'attente avec vacances et celle avec rappels et nous présenterons l'architecture du routeur de cœur qui permet d'assurer une différenciation de service dans les réseaux RMS d'AT. Enfin, on donnera la modélisation qu'on a réalisé du fonctionnement du routeur de cœur pour évaluer les critères de QoS complexes (délai, taux de perte, ou gigue) dans le réseau IP à différenciation de service, pour pouvoir satisfaire les besoins des applications les plus exigeantes et en particulier les applications à temps réel (visio-conférence) des abonnés d'AT.

4.1 Rappels

4.1.1 Description d'un système de file d'attente classique

Une file d'attente ou queue est un système stochastique composé d'un certain nombre (fini ou non) de places d'attente d'un ou plusieurs serveurs et bien sûr de clients qui arrivent,

attendent, se font servir selon des disciplines de service données et quittent le système. La description précédente d'une file d'attente, dont une représentation schématique est donnée dans la FIGURE 4.1, ne saurait capturer toutes les caractéristiques des différents modèles que comptent la littérature, mais elle identifie les éléments principaux permettant la classification de la grande majorité des files d'attente simples.

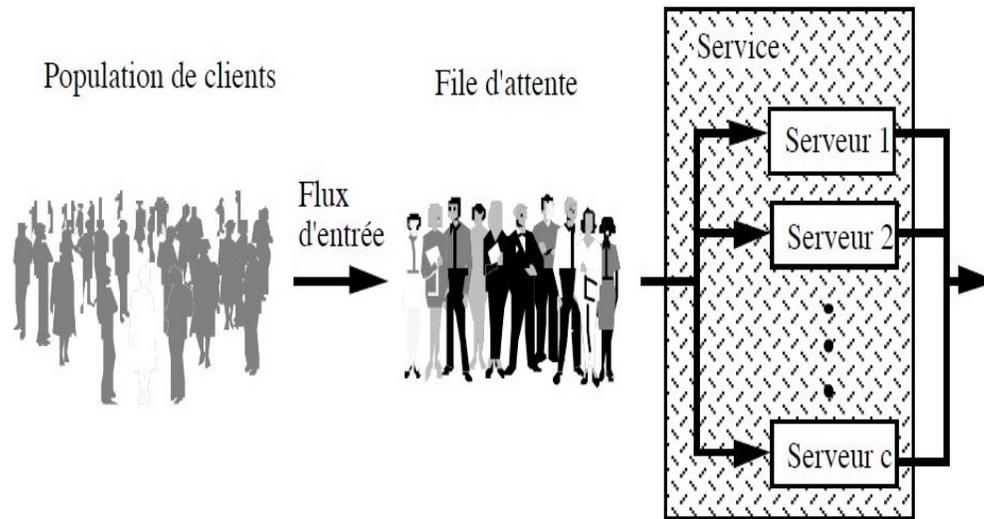


FIGURE 4.1 – Un système de files d'attente.

4.1.2 Analyse mathématique d'un système de files d'attente

L'étude mathématique d'un système de files d'attente se fait généralement par l'introduction d'un processus stochastique, défini de façon appropriée. On s'intéresse principalement au nombre de clients $X(t)$ qui se trouvent dans le système à l'instant t ($t \geq 0$).

En fonction des quantités qui définissent le système, on cherche à déterminer :

1. Les probabilités d'état $P_n(t) = P(X(t) = n)$, qui définissent le régime transitoire du processus stochastique $\{X(t); t \geq 0\}$. Il est évident que les fonctions $P_n(t)$ dépendent de l'état initial ou de la distribution initiale du processus.
2. Le régime stationnaire du processus stochastique est défini par :

$$\pi_n = \lim_{t \rightarrow \infty} P_n(t) = P(X(+\infty) = n) = P(X = n), (n = 0, 1, 2, \dots);$$

où $\{\pi_n\}_{n \geq 0}$ est appelée distribution stationnaire du processus $\{X(t); t \geq 0\}$.

Le calcul explicite du régime transitoire s'avère généralement pénible, voire impossible, pour la plupart des modèles donnés. On se contente donc de déterminer le régime stationnaire.

4.1.3 Classification des systèmes d'attente

Pour identifier un système d'attente, on a besoin des spécifications suivantes :

- La nature stochastique du processus des arrivées, qui est défini par la distribution des intervalles séparant deux arrivées consécutives ;
- La distribution du temps aléatoire de service ;
- Le nombre m de serveurs (stations de service). On admet généralement que les temps de service correspondants suivent la même distribution et que les clients qui arrivent forment une seule file d'attente ;
- La capacité N du système. Si $N < \infty$, la file d'attente ne peut dépasser une longueur de $N-m$ unités. Dans ce cas, certains clients arrivant vers le système n'ont pas la possibilité d'y entrer.

4.1.4 Notation de Kendall

Un modèle de file d'attente est totalement décrit selon la notation de Kendall. Dans sa version étendue, un modèle est spécifié par une suite de six symboles :

$$A/B/s/N/M/D$$

La signification de chacun de ces symboles est :

- A : nature du processus des arrivées ;
- B : nature du processus de service ;
- s : nombre de serveurs ;
- N : capacité d'accueil de la file d'attente ;
- M : taille de la population ;
- D : discipline de la file.

Dans la description des processus d'arrivée et de service, les symboles les plus courants sont :

- M : loi Exponentielle (memoryless) ;
- E : loi d'Erlang ;
- Γ : loi Gamma ;
- D : loi Déterministe (temps d'inter-arrivées ou de service constant) ;
- G : loi Générale (quelconque).

La forme abrégé : $A/B/s$ signifie que N et M sont infinies.

4.1.5 Les différentes disciplines de service

La discipline de service décrit l'ordre avec lequel les arrivées dans le système vont accéder au service. Ces disciplines peuvent être :

- **FIFO (First In First Out)** : Le premier arrivé est le premier servi ;
- **LIFO (Last In First Out)** : Le dernier arrivé sera le premier servi ;
- **Random (aléatoire)** : Les clients accèdent au serveur de manière aléatoire, indépendamment de l'ordre des arrivées ;
- **Priorité relative** : Un client accède au service selon sa priorité. La file est gérée par ordre de priorité de la plus forte à la plus faible ;
- **Priorité absolue** : Le service d'un client est interrompu lorsqu'un client de priorité supérieure se présente devant la file d'attente. Le client dont ce service est interrompu est remis en tête de la file.

4.1.6 Mesures de performance d'une file d'attente

L'étude d'une file d'attente ou d'un réseau de files d'attente a pour but de calculer ou d'estimer les performances d'un système dans des conditions de fonctionnement données. Ce calcul se fait le plus souvent pour le régime stationnaire uniquement, et les mesures les plus fréquemment utilisées sont :

- $\bar{N} = E(X)$: nombre moyen de clients dans le système ;
- \bar{Q} : nombre moyen de clients dans la file d'attente ;
- \bar{T} : temps moyen de séjour d'un client dans le système ;
- \bar{W} : temps moyen d'attente d'un client dans la file ;
- \bar{U} : taux d'utilisation de chaque serveur ;
- \bar{S} : le temps moyen de service ;
- \bar{A} : le temps moyen entre deux arrivées.

Ces valeurs ne sont pas indépendantes les unes des autres, mais sont liées par les relations suivante :

- $\bar{N} = \lambda \bar{T}$ (Formule de Little), où λ représente le taux d'arrivées ;
- $\bar{Q} = \lambda \bar{W}$;
- $\bar{T} = \bar{W} + \frac{1}{\mu}$; où μ représente le taux de service ;
- $\bar{N} = \bar{Q} + \rho$;

De manière générale, une file est stable si et seulement si le nombre moyen d'arrivées de clients par unité de temps, noté λ , est inférieur au nombre moyen de clients pouvant être servis par unité de temps. Si chaque serveur peut traiter μ clients par unité de temps et si le nombre de serveurs est m , une file est stable si et seulement si

$$\lambda < m\mu \iff \rho = \lambda/m\mu < 1;$$

où, ρ est appelé l'intensité du trafic.

4.1.7 Description d'un système d'attente avec rappels

Les systèmes de files d'attente avec rappels ou avec répétition d'appels se caractérisent par le fait suivant : un client qui arrive dans le système et trouve tous les serveurs occupés, quitte le système définitivement, ou rappelle ultérieurement à des instants aléatoires. Un client qui attend pour rappeler est dit en orbite et devient source d'appels secondaires. Un système de files d'attente avec rappels contient un espace de service composé de $c \geq 1$ dispositifs de service et d'un espace d'attente ayant $m - c$ ($m \geq c$) positions d'attente. A l'arrivée d'un client primaire, s'il y a un ou plusieurs serveurs libres, le client sera immédiatement pris en charge. Sinon, s'il y a une position d'attente libre, le client rejoint la file d'attente. Lorsque tous les serveurs et positions d'attente sont occupés, le client quitte le système, soit définitivement avec une probabilité $1 - H_0$ soit temporairement avec une probabilité H_0 et rappelle ultérieurement, après un temps aléatoire. La capacité O de l'orbite peut-être finie ou infinie. Dans le cas où O est finie, et si l'orbite est pleine, le client quitte le système définitivement. Chaque client de l'orbite forme un processus d'arrivées secondaires de taux θ et il est traité de la même manière qu'un client primaire avec une probabilité H_k (s'il s'agit de la k^{ieme} tentative échouée). Le schéma général d'un système avec rappels est donné dans la figure suivante :

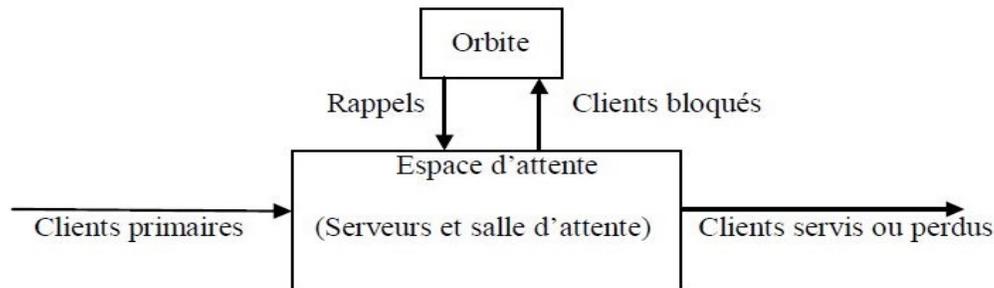


FIGURE 4.2 – Schéma général d'un système avec rappels ;

D'après cette figure on peut déduire que le système de files d'attente avec rappels peut être vu comme un type spécial des réseaux de files d'attente. Dans sa forme la plus générale, ces réseaux contiennent des noeuds : un noeud spécial pour le service et l'attente, et un noeud de délai pour les rappels.

4.1.8 Description d'un système d'attente avec vacances

Une autre classe très importante de files d'attente traite les systèmes dans lesquels le serveur reste oisif quand la file d'attente est vide. Cependant, le temps d'oisiveté du serveur pourrait être utilisé pour une tâche secondaire dans le but d'améliorer l'efficacité du système. Ces situations pratiques peuvent être analysées par des modèles appelés : "Modèles de files d'attente avec vacances". Il est à noter que dans le cas des systèmes d'attente avec pannes actives et réparation, les périodes de panne ou de réparation peuvent être aussi vues comme des périodes de vacances. Les files d'attente avec rappels peuvent

être également considérées comme un cas particulier des systèmes de files d'attente avec vacances, où la période de vacances commence à la fin de chaque temps de service et dure jusqu'à ce que le serveur soit réactivé par l'arrivée d'un client primaire (de l'extérieur) ou bien secondaire (de l'orbite) [4]. Autrement dit, la période d'oisiveté du serveur peut être considérée comme une période de vacances [15].

4.1.9 Quelques lois usuelles de probabilités

Dans cette section, nous allons présenter quelques lois usuelles de probabilité [18].

1. Loi de Poisson :

Une variable aléatoire X suit une loi de Poisson de paramètre μ si sa densité de probabilité est de la forme :

$$P(X = n) = \frac{\mu^n}{n!} e^{-\mu}, n = 0, 1, 2, \dots$$

Pour cette distribution, on a :

$$E(X) = \sigma^2(X) = \mu.$$

2. Loi exponentielle :

La densité de probabilité d'une loi X exponentielle de paramètre μ est donnée par :

$$f(x) = \mu e^{-\mu x}, \quad x \geq 0.$$

sa fonction de répartition est de la forme :

$$F(x) = 1 - e^{-\mu x}, \quad x \geq 0.$$

Pour cette distribution on a :

$$E(X) = \frac{1}{\mu}, \sigma^2(X) = \frac{1}{\mu^2}.$$

3. Loi d'Erlang :

La loi d'Erlang d'ordre k est un cas particulier de la loi Gamma où les paramètres $\alpha=k$ et $\beta=k\mu$. La loi d'Erlang- k représente un système où le service est effectué avec plusieurs serveurs en série de même taux de service μ , une seule unité pouvant se trouver dans la zone de service. La densité de probabilité de la loi d'Erlang- k , X est :

$$f(x) = \mu \frac{(\mu x)^{k-1}}{(k-1)!} e^{-\mu x}, \quad x \geq 0.$$

Les paramètres μ et k sont appelés respectivement paramètre d'échelle et paramètre de forme. Pour cette distribution, on a :

$$E(X) = \frac{k}{\mu}, \sigma^2(X) = \frac{k}{\mu^2}.$$

4.2 Modélisation du routeur de cœur par un système $M/G/1$ avec rappels et vacances

4.2.1 Description du modèle

On considère un système de files d'attente à un seul serveur où les clients primaires arrivent suivant un flux poissonnien de taux $\lambda > 0$. Un client qui arrive et trouve le serveur non disponible (occupé ou en vacances), quitte l'espace de service pour rejoindre un groupe de clients dit en "orbite". Après un certain temps aléatoire, il renouvelle sa tentative d'entrer en service, une fois, deux fois, ..., jusqu'à ce qu'il le trouve disponible.

La discipline d'accès au serveur à partir de l'orbite est gouvernée par une loi exponentielle avec une intensité donnée par $\theta j > 0$, quand le nombre de clients en orbite est $j \in \mathbb{N}$. Comme cette politique de rappels dépend du nombre de clients dans l'orbite, on l'appelle politique de rappels classiques. Les temps de service sont supposés d'une loi arbitraire, de fonction de distribution $B(t)$ avec $B(0) = 0$, de transformée de Laplace-Stieltjes $\psi(\omega)$ et des deux premiers moments finis γ_1 et γ_2 , respectivement.

Tous les clients entrant dans le système sont servis d'une manière continue et dans un ordre indépendant de leurs temps de service. De plus, on suppose que le serveur prend une vacance chaque fois que le système devient vide, de distribution $V(x)$, de transformée de Laplace-Stieltjes $V^*(\omega)$ et des deux premiers moments finis $E(X)$ et $E(X^2)$, respectivement.

◆ Les règles qui gouvernent les périodes de vacances :

1. Le mécanisme qui détermine l'instant de la fin d'une vacance, n'anticipe pas une nouvelle occurrence du processus des arrivées poissonniennes ;
2. Chaque temps de service est indépendant de la séquence des périodes de vacances qui précèdent ce temps de service ;
3. Si aucun client n'arrive durant la période de vacances, on dit qu'il y'a une période d'activité pour le serveur de longueur zéro et le serveur prend une autre vacance.
4. Juste après la fin des vacances, s'il y a des clients en orbite, le prochain client qui arrive au service est déterminé par une compétition entre deux lois exponentielles de taux λ et θ .

Finalement, on suppose que le flux des arrivées primaires, les intervalles entre les rappels successifs et les temps de service sont mutuellement indépendants. L'état du système à l'instant t peut être décrit par le processus :

$$X(t) = (C(t), N(t), \xi(t))_{(t \geq 0)},$$

avec :

$$C(t) = \begin{cases} 0, & \text{si le serveur est oisif,} \\ 1, & \text{si le serveur est occupé,} \\ 2, & \text{si le serveur est en vacance.} \end{cases}$$

$N(t)$: le nombre de clients en orbite à l'instant t . Si $C(t) = 1$ (respectivement $C(t) = 2$), alors $\xi(t)$ représente le temps de service écoulé du client en service (respectivement, le temps de vacance écoulé).

4.3 Chaîne de Markov induite

L'évolution de notre file d'attente avec rappels et vacances a été décrite en termes d'une séquence alternée de période d'activité et d'inactivité du serveur.

À la fin de chaque service, le serveur devient libre. La prochaine période d'inactivité du serveur sera de deux types différents :

Type 1 : Si l'orbite devient vide, donc le serveur prend une vacance propre qui est régie suivant les quatre règles qui gouvernent les vacances, déjà données.

Type 2 : Si l'orbite n'est pas vide après la fin d'un service, alors une compétition entre deux lois exponentielles de taux λ et θ déterminera le prochain client qui entrera en service. Soit $\{\zeta_n, n \in \mathbb{N}\}$ une suite d'instant de la fin d'un service ou bien de la fin d'une vacance propre.

La séquence des vecteurs aléatoires $Z_n = \{C(\zeta_n^-), N(\zeta_n^+)\}$ forme une chaîne de Markov, qui est une chaîne de Markov incluse du système de files d'attente donné. Son espace d'état est donné par $S = \{1, 2\} \times \mathbb{N}$. Les états de transitions de la chaîne de Markov induite $\{Z_n\}_{n=0}^\infty$ sont donnés par :

$$(i_{n+1}, j_{n+1}) = \begin{cases} (2, X), & \text{si } j_n = 0, \\ (1, j_n - \delta_{j_n} + v_{n+1}), & \text{si } j_n > 0, \end{cases} \quad (4.1)$$

où :

X : le nombre de clients primaires qui arrivent vers le système durant une vacance.

v_{n+1} : le nombre de clients primaires qui arrivent pendant le $(n+1)^{\text{ème}}$ temps de service qui se termine à l'instant ζ_{n+1} , sa distribution est donnée par :

$$K_i = P(v_{n+1} = i) = \int_0^\infty \frac{(\lambda x)^i}{i!} e^{-\lambda x} dB(x), \quad i \geq 0.$$

δ_{j_n} est la variable de Bernoulli définie comme suit :

$$\delta_{j_n} = \begin{cases} 0, & \text{si le } (n+1)^{\text{ème}} \text{ client servi provient de l'orbite,} \\ 1, & \text{sinon.} \end{cases}$$

Sa distribution dans le cas de rappels classiques est donnée par :

$$P(\delta_{j_n} = 1/j_n = k) = \frac{k\theta}{\lambda + k\theta},$$

$$P(\delta_{j_n} = 0/j_n = k) = \frac{\lambda}{\lambda + k\theta}.$$

4.3.1 Condition d'ergodicité

La condition d'ergodicité de la chaîne de Markov induite définie précédemment est donnée par le théorème suivant :

Théorème 4.1. [8] *La chaîne de Markov induite $\{Z_n\}_{n=0}^{\infty}$ est ergodique si et seulement si :*

$$\rho = \lambda\gamma_1 < 1.$$

4.3.2 Les probabilités de transition

Les probabilités de transition en un pas de la chaîne de Markov induite $\{Z_n\}_{n=0}^{\infty}$ sont données par les formules suivantes : [8]

Si $C(\zeta_{n+1}^-) = 1$:

$$\begin{aligned} r_{km} &= P[j_{n+1} = m / j_n = k], \\ &= k_{m-k} \frac{\lambda}{\lambda + K\theta} [\pi_{1,k} + \pi_{2,k}] + k_{m-k+1} \frac{K\theta}{\lambda + K\theta} [\pi_{1,k} + \pi_{2,k}], \end{aligned} \quad (4.2)$$

Si $C(\zeta_{n+1}^-) = 2$:

$$\begin{aligned} r_{km} &= P(X = m)P(1, j_n = 0) + P(X = m)P(2, j_n = 0); \\ &= P(X = m)\pi_0, \end{aligned} \quad (4.3)$$

où : $\pi_0 = \pi_{1,0} + \pi_{2,0}$.

4.3.3 Distributions stationnaires de la chaîne de Markov induite

Sous la condition d'ergodicité ($\rho < 1$), on a :

$$\pi_{i,j} = \lim_{n \rightarrow \infty} P(Z_n = (i, j)), \quad (i, j) \in S.$$

D'après (4.2) et (4.3), on aura :

$$\begin{aligned} \pi_{1,m} &= (1 - \delta_{m0}) \sum_{k=1}^m \frac{\lambda}{\lambda + K\theta} K_{m-k} (\pi_{1,k} + \pi_{2,k}) + \\ &\quad + \sum_{k=1}^{m+1} \frac{k\theta}{1 + k\theta} k_{m-k+1} (\pi_{1,k} + \pi_{2,k}), \\ \pi_{2,m} &= P(X = m)(\pi_{1,0} + \pi_{2,0}), \quad m \geq 0, \end{aligned}$$

où : $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$ est la fonction de Kronecker.

◆ **Fonctions génératrices :**

Le théorème suivant donne les distributions stationnaires en termes des fonctions génératrices [8] de la chaîne de Markov induite $\{Z_n\}_{n=0}^{\infty}$.

Théorème 4.2. *La distribution stationnaire*

$$\pi_{i,j} = \lim_{n \rightarrow \infty} P(Z_n = (i, j)), \quad (i, j) \in S, \quad (4.4)$$

possède les fonctions génératrices suivantes :

$$\pi_1(z) = \frac{\lambda K(z)(1-z)}{K(z)-z} L(z) + \frac{(1-\rho)(z-\chi(z))K(z)}{(\lambda E(X)-\rho)(K(z)-z)}, \quad (4.5)$$

$$\pi_2(z) = (1-\rho)(\lambda E(X)-\rho)^{-1} \chi(z). \quad (4.6)$$

où, $K(z) = \psi(\lambda - \lambda z)$ et $\chi(z) = V^*(\lambda - \lambda z)$ est la fonction génératrice de de la variable X .

De ce théorème on aura :

$$\pi_1(z) + \pi_2(z) = \frac{\lambda K(z)(1-z)}{K(z)-z} L(z) + \frac{\pi_0(K(z)-\chi(z))}{K(z)-z}.$$

Il est clair que pour $z = 1$, on aura :

$$\pi_0 = \pi_{1,0} + \pi_{2,0} = (1-\rho)(\lambda E(X)-\rho)^{-1}. \quad (4.7)$$

Notons que l'expression (4.7) représente la probabilité de l'état stationnaire quand le serveur est non disponible à recevoir des clients. Finalement, on aura :

$$\pi_1(1) = 1 - \pi_0 \text{ et } \pi_2(1) = \pi_0.$$

4.4 Approche par les processus régénératifs

Plusieurs processus stochastiques survenant, par exemple, dans les systèmes de files d'attente et les systèmes de gestion de stock possèdent la propriété de "régénération" en certains instants, alors le comportement futur du processus après ces instants devient une réplique, i.e. le comportement futur du processus après ces instants possède exactement la loi de probabilité qu'il aurait eu s'il avait commencé à l'instant zéro. De tels processus sont appelés "processus régénératifs" [19].

4.4.1 Distributions limites

Nous avons vu dans la section précédente comment obtenir la distribution stationnaire de la chaîne de Markov incluse aux époques de départ. Maintenant, nous employons une approche récursive basée sur la théorie des processus régénératifs, pour calculer les

distributions limites :

$$P_{i,j} = \lim_{t \rightarrow \infty} P\{Z(t) = (C(t), N(t)) = (i, j)\}, \quad (i, j) \in E = 0, 1, 2 \times \mathbb{N}, (P_{00} \equiv 0),$$

sous la condition d'ergodicité $\rho < 1$.

Pour cela, on définit quelques variables aléatoires :

T : la longueur d'un cycle,

T_{ij} : la quantité du temps dans un cycle durant lequel le système est à l'état (i, j) ,

V : la longueur d'une vacance propre,

V_j : le nombre de vacances propres dans un cycle pour lesquelles j clients sont laissés en orbite,

N_j : le nombre de fins de service dans un cycle pour lequel j clients sont laissés en orbite.

Alors :

$$P_{i,j} = \frac{E(T_{i,j})}{E(T)}, \quad \forall (i, j) \in E.$$

En faisant l'égalité entre le taux d'entrée et le taux de sortie pour l'état $(0, j)$ et $\{(i, n) : i \in 0, 1, 2, j \geq n \geq 0\}$, respectivement, et en utilisant la propriété de PASTA [22], on obtient les équations de balance suivantes :

$$(\lambda + j\theta)E[T_{0,j}] = E[N_j] + E[V_j], \quad j \geq 1, \quad (4.8)$$

$$j\theta E[T_{0,j}] = \lambda(E[T_{1,j-1}] + E[T_{2,j-1}]), \quad j \geq 1. \quad (4.9)$$

Pour trouver une relation entre $E[T_{1,j}]$, $E[N_j]$ et $E[V_j]$, on introduit la quantité auxiliaire :

$A_{n,j}$: le temps moyen pour que durant un temps de service, j clients sont présents en orbite sachant que dans le précédent temps de service, ou la vacance propre, n clients sont laissés en orbite. Il est évident que $E[V_n] = P\{X = n\}$.

À présent, une simple application du théorème de Wald nous permet d'avoir :

$$E[T_{1,j}] = \sum_{n=1}^{j+1} (E[N_n] + E[V_n])A_{n,j}, \quad j \geq 0. \quad (4.10)$$

En combinant (4.8) et (4.10) on trouve que :

$$E[T_{1,j}] = \sum_{n=1}^{j+1} (\lambda + n\theta)E[T_{0,j}]A_{n,j}, \quad j \geq 0. \quad (4.11)$$

De (4.9) et (4.11) on obtient :

$$E[T_{1,j}] = \sum_{n=0}^j \frac{\lambda(\lambda + (n+1)\theta)}{(n+1)\theta} (E[T_{1,n}] + E[T_{2,n}]) A_{n;j}, \quad j \geq 0. \quad (4.12)$$

En divisant (4.9) et (4.12) par $E[T]$, on trouve les relations de récurrence suivantes :

$$P_{0,j} = \frac{\lambda}{j\theta} (P_{1,j-1} + P_{2,j-1}), \quad j \geq 0. \quad (4.13)$$

$$P_{1,j} = \sum_{n=0}^j \frac{\lambda(\lambda + (n+1)\theta)}{(n+1)\theta} (P_{1,n} + P_{2,n}) A_{n+1;j}, \quad j \geq 0. \quad (4.14)$$

Les équations (4.13) et (4.14) donnent une procédure récursive stable qui permet de trouver $\{P_{0,j}, j \geq 1\}$ et $\{P_{1,j}, j \geq 0\}$ en fonction $\{P_{2,j}, j \geq 0\}$ et $A_{n;j}$. La séquence $\{P_{2,j}, j \geq 0\}$ peut être déterminée en pratique pour chaque politique de vacances. Cela revient à spécifier les coefficients $A_{n;j}$. Pour calculer $A_{n;j}$, on définit une autre quantité auxiliaire :

$B_{n,j}$: le temps moyen pour que durant un temps de service, j clients sont présents en orbite sachant qu'immédiatement après le début du service, n clients étaient en orbite. Il est facile d'observer qu'un intervalle infinitesimal $(t, t + \Delta t)$ contribue à $B_{n,j}$ si :

- i. le temps de service n'a pas été terminé avant le temps t (avec la probabilité $1 - B(t)$),
- ii. $(j - k)$ clients primaires arrivent au système dans l'intervalle $(0, t)$.

Alors, on a :

$$B_{n,j} = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^{j-n}}{(j-n)!} (1 - B(t)) dt, \quad j \geq n \geq 0.$$

Les quantités auxiliaires $A_{n;j}$ et $B_{n,j}$ sont connectées par les relations suivantes :

$$A_{j+1,j} = \frac{(j+1)\theta}{\lambda + (j+1)\theta} B_{j,j}, \quad j \geq 0, \quad (4.15)$$

$$A_{n,j} = \frac{n\theta(1 - \delta_{n0})}{\lambda + n\theta} B_{n-1,j} + \frac{\lambda}{\lambda + n\theta} B_{n,j}, \quad j \geq n \geq 0. \quad (4.16)$$

A présent, l'équation (4.14) peut être réécrite comme suit :

$$(1 - \lambda a_0) P_{1,j} = \lambda a_0 P_{2,j} + \lambda(1 - \delta_{j0}) \sum_{n=1}^j (P_{1,n-1} + P_{2,n-1}) (a_{j-n+1} + \frac{\lambda}{n\theta} a_{j-n}), \quad j \geq 0. \quad (4.17)$$

où :

$$a_j = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^j}{(j)!} (1 - B(t)) dt, \quad j \geq 0; \quad \text{et} \quad A(z) \sum_{j=0}^\infty a_j z^j = \frac{1 - \psi(\lambda - \lambda z)}{\lambda - \lambda z}.$$

Par la suite, on étudie les fonctions génératrices partielles $P_i(z) = \sum_{j=0}^\infty P_{i,j} z^j$, pour $i \in \{0, 1, 2\}$ des probabilités limites.

Théorème 4.3. [8] Si $\rho < 1$ alors les fonctions génératrices partielles $P_i(z)$ sont données par :

$$P_0(z) = \frac{\lambda}{\theta} \frac{P_{2,\bullet}}{E(X)} \int_1^z \left[\frac{1 - \chi(t)}{K(t) - t} \exp\left\{-\frac{\lambda}{\theta} \int_1^t \frac{1 - K(u)}{k(u) - u} du\right\} \right] dt \\ \times \exp\left\{\frac{\lambda}{\theta} \int_1^z \frac{1 - K(t)}{K(t) - t} dt\right\}; \quad (4.18)$$

$$P_1(z) = \frac{K(z) - 1}{z - K(z)} (P_0(z) + P_2(z)), \quad (4.19)$$

$$P_2(z) = \frac{P_{2,\bullet}}{E(X)} \frac{1 - \chi(z)}{1 - z}, \quad (4.20)$$

A ce niveau, la seule inconnue est $P_{2,\bullet} = \sum_{j=0}^{\infty} P_{2,j}$, qui peut être déterminée en utilisant la condition de normalisation $P_0(1) + P_1(1) + P_2(1) = 1$. Alors, en posant $z = 1$ dans (4.18)-(4.20) et en appliquant la règle de l'Hôpital, on obtient :

$$P_{2,\bullet} = (1 - \rho)\lambda^{-1}; \quad (4.21)$$

$$P_{1,\bullet} = \sum_{j=0}^{\infty} P_{1,j} = \lambda\gamma_1 = \rho. \quad (4.22)$$

De (4.21), (4.22) et la condition de normalisation $\sum_{i=0}^2 \sum_{j=0}^{\infty} P_{i,j} = 1$, on trouve que :

$$P_{0,\bullet} = \sum_{j=0}^{\infty} P_{0,j} = \frac{(1 - \rho)(\lambda - 1)}{\lambda}. \quad (4.23)$$

4.4.2 Décomposition stochastique

Dans cette section, on donne le théorème de [8] résultat important concernant la décomposition stochastique de la distribution du nombre de clients dans le système en un point arbitraire au régime stationnaire.

$$P_j = (1 - \delta_{0,j})(P_{0,j} + P_{1,j-1}) + P_{2,j}, j \geq 0. \quad (4.24)$$

Théorème 4.4. [8] Si $\rho < 1$, alors la fonction génératrice du nombre de clients dans le système est donnée par :

$$P(z) = Q(z) \frac{P_0(z) + P_2(z)}{P_{0,\bullet} + P_{2,\bullet}}. \quad (4.25)$$

où $Q(z)$ est la formule bien connue de Pollaczek-Khintchine pour la file classique $M/G/1$, qui est donnée par :

$$Q(z) = \frac{(1 - \rho)(1 - z)K(z)}{K(z) - z}. \quad (4.26)$$

4.4.3 Quelques mesures de performance

Notre prochain objectif est de fournir des expressions explicites pour quelques mesures de performance du système M/G/1 avec rappels classiques et vacances du serveur. Les résultats sont résumés dans les corollaires suivants.

Corollaire 4.1. [8] *Le temps moyen de paquets dans le système durant une période d'oisiveté, une période d'occupation et une période de vacances est donné respectivement par :*

$$E(N_I) = p'_0(1) = \frac{\lambda}{\theta} \frac{\lambda}{1-\rho} P_{2,\bullet} = \frac{\lambda}{\theta}; \quad (4.27)$$

$$E(N_B) = p'_1(1) = \frac{\lambda^2}{2} \frac{\gamma_2}{1-\rho} + \frac{\lambda}{\theta} \frac{\rho}{1-\rho} + \frac{\lambda\rho}{2} \frac{E(X^2)}{E(X)}; \quad (4.28)$$

$$E(N_V) = p'_2(1) = \frac{\lambda^2 E(X^2)}{2} \frac{P_{2,\bullet}}{E(X)} = (1-\rho) \frac{\lambda}{2} \frac{E(X^2)}{E(X)}. \quad (4.29)$$

Corollaire 4.2. [8] *Le temps moyen de paquets dans le système en un point aléatoire est donné par :*

$$\begin{aligned} L_s = E(N_I) + E(N_B) + E(N_V) = P'(1) &= \\ &= \frac{1}{1-\rho} \left(\frac{\lambda}{\theta} + \frac{\lambda^2 \gamma_2}{2} \right) + \frac{\lambda E(X^2)}{2E(X)}. \end{aligned} \quad (4.30)$$

Le temps moyen d'attente dans le système est obtenu en utilisant les formules de Little et l'équation est donné par :

$$W = \frac{L_s}{\lambda} = \frac{1}{1-\rho} \left(\frac{1}{\theta} + \frac{\lambda \gamma_2}{2} \right) + \frac{E(X^2)}{2E(X)}. \quad (4.31)$$

4.4.4 Conclusion

Dans ce chapitre, nous avons donné quelques rappels sur la files d'attente, les files d'attente avec rappels, et les files d'attente avec vacances. Ensuite, nous avons donné quelques rappels sur quelques lois de probabilités usuels. Enfin, nous avons effectué une étude quantitative d'analyse de l'état stationnaire du modèle M/G/1 avec rappels classiques et vacances du serveur. En utilisant la technique de la chaîne de Markov induite, nous avons obtenu la condition d'ergodicité et la distribution du nombre de clients dans le système. À l'aide d'une approche récursive basée sur la théorie des processus de Markov régénératifs, nous avons déterminé les distributions limites associées à l'état du serveur, la décomposition stochastique habituelle de la distribution du nombre moyen de clients dans le système et quelques autres mesures de performance.

Chapitre 5

Application

Introduction

Dans le chapitre précédent une modélisation du routeur de coeur de l'architecture Diff-Serv a été établie. Dans ce présent, on va évaluer les performance du modèle obtenu. En effet, en premier lieu, on donnera et décortiquera les résultats analytiques de ce modèle établis par Dr Boualem dans [8]. En second lieu, on évaluera ces résultats par simulation en utilisant le logiciel Matlab. Une comparaison entre les résultats analytiques et ceux simulés sera réalisée.

5.1 Simulation

La simulation est une technique qui consiste à construire un modèle d'une situation réelle, puis à faire des expériences sur ce modèle. Cette définition est toutefois très vaste et dans notre travail on considère la simulation telle qu'elle est définie par Naylor et al : "La simulation est une technique numérique pour élaborer des expériences sur l'ordinateur. Elle implique l'utilisation de modèles logiques et mathématiques qui décrivent le comportement de systèmes administratifs ou économiques (ou de leurs sous-systèmes) durant une période de temps prolongée" [16].

Les objectifs principaux de la simulation sont l'évaluation des performances de systèmes, l'évaluation des protocoles, d'architecture des réseaux et prévoir leur fonctionnement,

5.2 Étapes de Simulation

Les différentes étapes à suivre pour faire une simulation d'un système sont :

- **Formulation du modèle** : Cette étape consiste à identifier et analyser le problème, en déterminant ses composantes, leur relation et les frontières entre le système et son environnement.
- **Elaboration du modèle** : Cette étape consiste à extraire un modèle aussi fidèle que

possible du système réel.

- **Identification du modèle et collecte de données** : la collecte de données est indispensable pour l'estimation des paramètres du modèle. Ceci requiert une connaissance des méthodes statistiques et des tests d'hypothèses.
- **Validation du modèle** : Cette étape consiste à évaluer les performances du modèle puis les comparer à celles du système réel.
- **Exécution de la simulation** : Pour mettre à l'épreuve le modèle. Le concepteur doit effectuer plusieurs exécutions et recueillir les résultats.
- **Analyse et interprétation des résultats** : Une fois les résultats obtenus, le concepteur passe à l'analyse et l'interprétation de ces résultats pour donner des recommandations et des propositions.
- **Conclusion** : Cette dernière étape consiste à évaluer les perspectives d'exploitation du modèle pour d'autres préoccupations.

5.3 Intérêts de la simulation

On fait appel à la simulation :

- Quand on ne peut pas facilement observer les états du système ;
- Quand on désire analyser l'enchaînement des événements dans le système, ainsi que les relations de causes à effets ;
- Quand on désire valider une solution analytique ;
- Quand la complexité des interactions dans le système est telle qu'elle ne peut être étudiée qu'au travers la simulation ;
- Quand on désire visualiser les états d'un système ;
- Quand on veut tester différentes optimisations pour améliorer un système déjà existant.

5.4 Avantages de la simulation

Les principaux avantages de la simulation peuvent se résumer en :

- Observations des états du système ;
- Etudes des points de fonctionnement d'un système ;
- Etudes de systèmes à échelle de temps variable ;
- Etudes de l'impact des variables sur les performances du système ;
- Etude d'un système sans les contraintes matérielles.

5.5 Inconvénients de la simulation

En parallèle, aux avantages de la simulation, cette dernière possède aussi ces propres inconvénients, à savoir :

- La conception de modèles peut nécessiter des compétences spéciales ;
- Une autre forme d'analyse plus proche de la réalité est peut être nécessaire ;
- Les résultats difficilement interprétables ;
- Les résultats pas forcément généralisables ;
- Les résultats sont fonction des entrées du système.

5.6 Performances du modèle du routeur de coeur

L'implémentation des résultats analytiques du modèle de files d'attente $M/G/1$ avec rappels et vacances retenu pour la modélisation du routeur de coeur de l'architecture Diffserv sous Matlab, ainsi que la simulation de ce modèle sous ce logiciel nous ont permis d'évaluer ses différentes performances : Le temps de réponse, le nombre moyens de paquets dans le système, le nombre de vacances prises par le serveur, le nombre de paquets servis, et ce, en faisant varier les paramètres suivants :

- Le taux d'arrivée des paquets de la classe prioritaire EF (paquet/ms) ;
- La loi de service et ses paramètres ;
- La loi de vacance et ses paramètres ;
- Le taux de rappel des paquets.

5.6.1 Organigramme de simulation

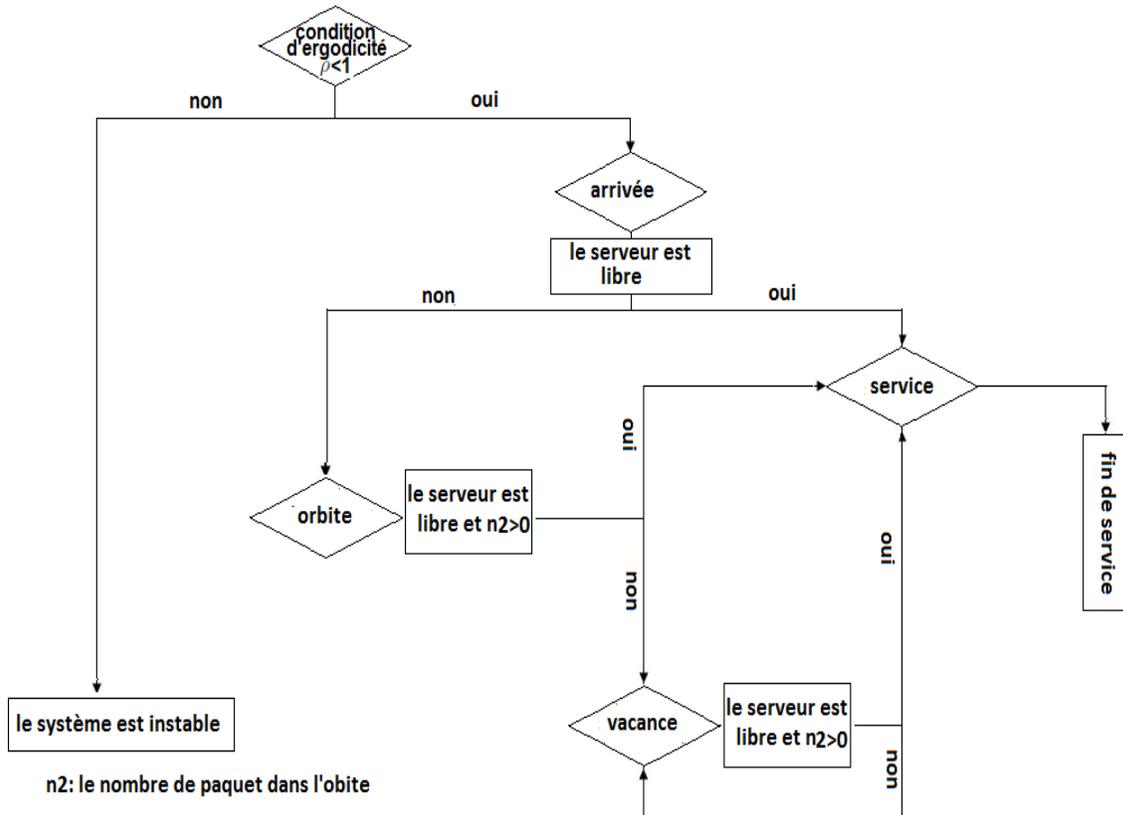


FIGURE 5.1 – Organigramme de la simulation du modèle étudié.

5.6.2 Evaluation des performances du modèle du routeur de coeur établi

On a implementé notre algorithme en faisant varier la loi de service ainsi que la loi des vacances. Ceci a donné lieu au résultats suivants :

► **Loi de service Exponentielle et de vacance Exponentielle** : En premier lieu, on a considéré le cas le plus simple, à savoir, une loi de service exponentielle de taux μ_1 et une loi de vacance exponentielle de taux μ_2 . Les différents résultats des performances recherchées pour le modèle étudié en variant les paramètres λ et θ sont représentés dans le tableau suivant :

Arriv.	Serv.		Rap.	Vac.	Rés. analyt.		Rés. de sim.			
λ	μ_1	ρ	θ	μ_2	L_s	W	L_s	W	D	V
0.1	1	0.1	0.5	0.5	0.4333	4.3333	0.4312	4.3119	105	347
				1	0.3333	3.3333	0.3068	3.0683	97	760
				2	0.2833	2.8333	0.2244	2.2441	97	1471
			3	0.5	0.2481	2.4814	0.2106	2.1058	83	440
				1	0.1481	1.4814	0.1279	1.2779	93	900
				2	0.0981	0.9814	0.0960	0.9599	98	1782
0.25	0.5	0.5	0.5	0.5	2.0000	8.0000	2.0049	8.0198	247	89
				1	1.7500	7.0000	1.8764	7.5057	242	264
				2	1.6250	6.5000	1.3437	5.3750	252	392
			3	0.5	1.1666	4.6666	1.1735	4.6938	255	243
				1	0.9166	3.6666	0.8756	3.5025	247	434
				2	0.7916	3.1666	0.8593	3.4370	247	803
0.5	0.6	0.90	0.5	0.5	11.1666	22.3333	11.2210	22.4419	475	25
				1	10.6666	21.3333	10.9202	21.8404	520	3
				2	10.4166	20.8333	11.8400	23.6799	505	16
			3	0.5	6.1666	12.3333	5.5791	11.1582	491	50
				1	5.6666	11.3333	5.9535	11.9070	497	85
				2	5.4166	10.8333	5.7270	11.4540	488	131

TABLE 5.1 – Performances du modèle dans le cas d’une loi de service exponentielle et celle des vacances exponentielle.

► **Loi de service Exponentielle et celle des vacances d’Erlang** : Pour une loi de service exponentielle de taux μ_1 et une loi des vacances d’Erlang-2 (i.e. $K = 2$) de paramètre μ_2 . Les différents résultats obtenus pour les performances du modèle étudié en variant les paramètres λ et θ sont représentés dans le tableau suivant :

Arriv.	Serv.		Rap.	Vac.	Rés. analyt.		Rés. de sim.			
λ	μ_1	ρ	θ	μ_2	L_s	W	L_s	W	D	V
0.1	1	0.1	0.5	0.5	0.5333	5.3333	0.3316	3.3163	112	711
				1	0.3833	3.8333	0.3388	3.3875	93	393
				2	0.3083	3.0830	0.3581	3.5806	82	191
			3	0.5	0.3481	3.4814	0.1333	1.3332	114	883
				1	0.1981	1.9814	0.1801	1.8015	85	457
				2	0.1231	1.2314	0.2687	2.6869	90	215
0.25	0.5	0.5	0.5	0.5	2.2500	9.0000	2.1635	8.6542	260	227
				1	1.8750	7.5000	1.5345	6.1381	238	117
				2	1.6875	6.7500	1.5618	6.2472	244	68
			3	0.5	1.4166	5.6666	1.0010	4.0039	274	384
				1	1.0416	4.1666	0.8413	3.3652	234	222
				2	0.8541	3.4166	0.8521	3.4084	237	121
0.5	0.6	0.90	0.5	0.5	11.6666	23.3333	12.0884	24.1767	560	7
				1	10.9166	21.8333	10.1525	20.3050	496	2
				2	10.5416	21.0833	10.8280	21.6560	505	5
			3	0.5	6.6666	13.3333	5.0751	10.1502	528	105
				1	5.9166	11.8333	5.2440	10.4880	501	90
				2	5.5416	11.0833	5.5885	11.1770	487	34

TABLE 5.2 – Performances du modèle dans le cas d’une loi de service exponentielle et celle des vacances d’Erlang.

► **Loi de service Erlang et celle des vacances Exponentielle** : Pour une loi de service Erlang-2 (i.e. $K = 2$) de paramètre μ_1 et une loi des vacances exponentielle de taux μ_2 . Les différents résultats obtenus pour le modèle étudié en variant les paramètres λ et θ sont représentés dans le tableau suivant :

Arriv.	Serv.		Rap.	Vac.	Rés. analyt.		Rés. de sim.			
λ	μ_1	ρ	θ	μ_2	L_s	W	L_s	W	D	V
0.1	2	0.1	0.5	0.5	0.4305	4.3055	0.4908	4.9081	90	289
				1	0.3305	3.3055	0.4252	4.2521	98	512
				2	0.2805	2.8055	0.3499	3.4991	82	1012
			3	0.5	0.2453	2.4537	0.3118	3.1185	91	315
				1	0.1453	1.4535	0.2547	2.5471	97	640
				2	0.0953	0.9537	0.2013	2.0128	86	1264
0.25	1	0.5	0.5	0.5	1.8750	7.5000	1.6970	6.7880	255	112
				1	1.625	6.5000	1.7386	6.9545	272	214
				2	1.5000	6.0000	1.5399	6.1596	268	466
			3	0.5	1.0416	4.1666	1.1395	4.5580	262	181
				1	0.7916	3.1666	0.7551	3.0202	249	417
				2	0.6666	2.6666	0.5496	2.1983	233	932
0.5	1.1	0.9090	0.5	0.5	18.8181	37.6363	20.8783	41.7566	434	4
				1	18.3181	36.6363	21.4814	42.9627	454	1
				2	18.0681	36.1363	21.0835	42.1670	437	14
			3	0.5	9.6515	19.3030	10.5322	21.0645	474	4
				1	9.1515	18.3030	11.5245	23.0489	449	11
				2	8.9030	17.8030	8.7831	17.5661	435	51

TABLE 5.3 – Performances du modèle dans le cas d’une loi de service d’Erlang et celle des vacances Exponentielle.

► **Loi de service d’Erlang et des vacances d’Erlang :** Pour une loi de service Erlang-2 (i.e. $K = 2$) de paramètre μ_1 et une loi des vacances Erlang-2 (i.e. $K = 2$) de taux μ_2 . Les différents résultats obtenus pour le modèle étudié en variant les paramètres λ et θ sont représentés dans le tableau suivant :

Arriv.	Serv.		Rap.	Vac.	Rés. analyt.		Rés. de sim.			
λ	μ_1	ρ	θ	μ_2	L_s	W	L_s	W	D	V
0.1	2	0.1	0.5	0.5	0.5305	5.3055	0.4468	4.4680	91	481
				1	0.3805	3.8055	0.4232	4.2319	94	285
				2	0.3055	3.0555	0.4784	4.7842	74	133
			3	0.5	0.3453	3.4570	0.3906	3.9057	112	526
				1	0.1953	1.9537	0.2711	2.7107	87	319
				2	0.1203	1.2037	0.4132	4.1316	94	159
0.25	1	0.5	0.5	0.5	2.1250	8.5000	0.9211	3.6844	316	331
				1	1.7500	7.0000	1.0214	4.0855	265	214
				2	1.5625	6.2500	1.3890	5.5561	276	113
			3	0.5	1.2916	5.1666	1.1833	4.7332	274	339
				1	0.9166	3.6666	0.5459	2.1835	254	318
				2	0.7291	2.9166	0.6570	2.6279	228	191
0.5	1.1	0.9090	0.5	0.5	19.3181	38.6363	17.7905	35.5810	438	3
				1	18.5681	37.1363	18.4646	36.9292	426	2
				2	18.1931	36.3863	20.9529	41.9058	430	4
			3	0.5	10.1515	20.3030	22.0170	22.0170	418	60
				1	9.4015	18.8030	10.7928	21.5857	440	11
				2	9.0265	18.0530	10.5685	21.1370	440	5

TABLE 5.4 – Performances du modèle dans le cas d’une loi de service d’Erlang et celle des vacances d’Erlang.

5.6.3 Interprétation des résultats

Les résultats analytiques obtenus pour le modèle $M/G/1$ avec rappels et vacances et ceux obtenus par simulation sous Matlab, nous ont permis de dégager un ensemble de remarques et conclusions qu’on donnera si après :

Loi de service est exponentielle :

On remarque qu’à chaque fois que l’intensité du trafic ρ augmente, le délai d’attente augmente aussi, ce qui est logique. En parallèle, le nombre de vacances diminue, ce qui s’explique par le faite que le taux d’arrivée des clients est proche du taux de service i.e. le serveur se libère peu. Par ailleurs, on constate qu’à chaque fois que le taux de rappels augmente le temps de réponse diminue. De plus, on a remarqué que les paramètres de la loi des vacances influent sur le nombre de vacances prises par le serveur, à savoir :

- si les vacances sont exponentielles de paramètre μ_2 , à chaque fois que le taux de la durée des vacances augmente, le nombre de vacances augmente aussi (voir TABLE (5.1)); ceci est dû au fait que la durée des vacances est petite.
- si les vacances sont d’Erlang de paramètres μ_2 et $k = 2$, à chaque fois que le taux des vacances augmente, le nombre de vacances diminue en parallèle (voir TABLE (5.2)).

Ceci est dû au fait que la durée des vacances est assez longue.

Donc le nombre de vacances est relatif à la vitesse de convergence de la loi de la durée des vacances. Ainsi, dans la cas des deux lois exponentielle et d'Erlang, on sait que la vitesse de convergence de la loi d'Erlang est plus grande que celle de la loi exponentielle, d'où les résultats obtenus.

Loi de service est d'Erlang :

Même dans ce cas, on constate qu'à chaque fois que l'intensité du trafic ρ augmente le délai d'attente augmente aussi, ce qui est logique. De même, le nombre de vacances diminue, ce qui s'explique par le faite que le taux d'arrivée des clients est proche du taux de service i.e. le serveur se libère peu. Par ailleurs, on constate qu'à chaque fois que le taux de rappels augmente le temps de réponse diminue. De plus, de manière similaire au cas du service avec une loi exponentielle, les paramètres de la loi des vacances influent sur le nombre de vacances prises par le serveur à savoir :

- si les vacances sont exponentielles de paramètre μ_2 , à chaque foi que le taux de la durée des vacances augmente, le nombre de vacances augmente aussi (voir TABLE (5.3)); ceci est dû au fait que la durée des vacances est petite.
- si les vacances sont d'Erlang de paramètres μ_2 et $k = 2$, à chaque foi que le taux des vacances augmente, le nombre de vacances diminue en parallèle (voir TABLE (5.4)).

Ceci est dû toujours au fait que la durée des vacances est assez longue.

Ainsi, le nombre de vacances est toujours relatif à la vitesse de convergence de la loi de la durée des vacances.

On remarque qu'il y'a une légère différence entre les résultats analytiques et ceux de la simulation dus principalement aux erreurs de la simulation. Par exemple, dans la (TABLE (5.2)), pour $\rho = 0.1$, $\theta = 0.5$ et $\mu_2 = 0.5$, on a $L_s = 0.3316$ et pour $\rho = 0.1$, $\theta = 0.5$ et $\mu_2 = 1$, on a $L_s = 0.3388$ et pour $\rho = 0.1$, $\theta = 0.5$ et $\mu_2 = 2$ on a $L_s = 0.3581$, c'est à dire, en variant μ_2 il y a une petite augmentation, alors que théoriquement ça devait diminuer légèrement.

5.7 Conclusion

Dans ce chapitre, nous avons réalisé une petite application pour simuler le modèle qu'on a établi pour le routeur de coeur de l'architecture DiffServ. Les résultats numériques obtenus confirment que les résultats analytiques et ceux de la simulation sont assez proches. Notre modélisation de ce routeur diffère de la modélisation réalisée dans [14] dans le sens que nous avons pris en considération les rappels. D'ailleurs, parmi les avantages que le fait de prendre en compte les rappels nous a procuré on trouve que notre système n'a plus de perte de paquets, ce qui est indispensable dans les transmissions audio-visuelles temps réel. Par ailleurs, on constate qu'à chaque fois que le taux de rappels augmente le temps de réponse diminue.

Conclusion générale

De l'avis de nombreux spécialistes, nos réseaux de télécommunications ont atteint leur limite en matière de services. Ainsi, Algérie Télécom a initié depuis le début des années 2000 un plan de développement conséquent pour accélérer l'usage des TIC dans notre pays. Le recours aux réseaux de nouvelle génération, initié en 2007, a atteint aujourd'hui la phase finale de l'opération de " modernisation du réseau de télécommunication d'Algérie Télécom " .

Le recours aux NGN a pour objectif de lever les limitations (en termes de débits) pour la transmission de données que ce soit : la voix, la vidéo, internet,

Ainsi une analyse de performances de ces réseaux est indispensable. Faire une analyse globale est impossible, dans ce travail on s'est limité à l'étude de routeur coeur de l'architecture DiffServ. Cette architecture est modèle de gestio de qualité de service (QoS) approprié aux applications temps réel(voix, vidéo...) qui sont les objectifs des NGN.

Pour cela nous avons présenté l'architecture du routeur de coeur qui permet d'assurer une différenciation de service dans les réseaux RMS d'AT. Par la suite, on a donné une modélisation qu'on a réalisée sur le fonctionnement du routeur de coeur pour évaluer les critères de QoS complexes (délai,taux de pert, ou gigue) dans le réseau IP à différenciation de service, pour pouvoir satisfaire les besoins des applications les plus exigeantes et en particulier les application à temps réel (visio-conférence) des abonnés d'AT. Notre modélisation du routeur de coeur diffère de la modélisation réalisée dans [14] dans le sens que nous avons pris en considérations les rappels. D'ailleurs, l'objectif qu'on a visé en intégrant la notion de rappels dans notre modélisation est d'éviter les pertes de paquets, ce qui est indispensable dans les transmissions audio-visuelles temps réel. Une fois notre modélisation établie restait à évaluer les performance du modèle obtenu. En effet, en premier lieu, on a donné les résultats analytiques de ce modèle établis par Dr Boualem dans [8]. En second lieu, on a évalué ces résultats par simulation en utilisant le logiciel Matlab et une comparaison entre les résultats analytiques et ceux simulés est réalisée. En résumé, les résultats numériques qu'on a obtenus confirment que les résultats analytiques et ceux de la simulation sont assez proches.

La qualité de service (QoS) dans les télécommunications est un vaste domaine de travail à tel point qu'il est difficile d'en faire la synthèse. Pour vérifier que les conditions de la QoS sont vérifiées, il faut pouvoir mesurer cette QoS chose qui n'est pas facile car elle nécessite en premier lieu une bonne modélisation.

Perspectives

1. Comme perspectives à court termes il sera intéressant de considéré une modélisation avec un système avec rappels, priorité et vacances no exhaustives. Par la suite il sera intéressant aussi d'étudier le routeur dans sa globalité i.e considérer ces files d'attente et l'ordonnancement des paquets transmis.
2. simuler le réseau de transmission de la wilaya de Béjaia.

Bibliographie

- [1] http://www.algeriatelecom.dz/siteweb.php?p=at_histoire_realisations. 2015.
- [2] F. BAKER and J. HEINANEN, W. WEISS, and J. WROCLAWSKI. *Assured forwarding phb group*. Tech. rep.; Standards Track rfc (RFC 2597), Juin 1999.
- [3] A.R. ARTALEJO. *Analysis of M/G/1 queue with constant repeated attempts and server vacation*. Statistica, Neerlandica, 1997.
- [4] J. R. ARTALEJO and G. I FALIN. *Stochastic decomposition for retrial queues*. Top 2, 329–342, 1994.
- [5] G. AURIOL. *Spécification et implémentation d’une architecture de signalisation à gestion automatique de la QoS dans un environnement IP multi domaines*. Novembre 2004.
- [6] Y. BERNET, S. BLACKKE, D. GROSSMAN, and A. SMITH. *An Informel Management Model for Diffserv Routers*.
- [7] S. BLAKE, D. BLACK, and M. CARLSON. *An architecture for differentiated services*. 1998.
- [8] M. BOUALEM. *Sur la propriété de décomposition stochastique dans un système d’attente avec rappels et vacances*. Thèse de doctorat, Département Recherche opérationnelle, Université A/Mira de Béjaia, 2009.
- [9] O.N. MEDINA CARVAJAL. *Etude des algorithmes d’attribution de priorités*. PhD thesis, université de Rennes 1, Mars 2001.
- [10] S. FLOYD and V. JACOBSON. *Random early detection gateways for congestion avoidance, vol.1 of 4. IEEE/ACM transactions on networking*. Aout 1999.
- [11] D. FUIN. *Qualité de service : des réseaux IP à l’intégration dans les réseaux actifs*.
- [12] P. HURLEY, J.Y. LE BOUDEC, and P. THIRAN. *The alternative best effort service*. Tech. rep. SSC/1999/036, institute for computer communication and applications, EPFL, Septembre 1999.
- [13] V. JACOBSON, K. NICHOLS, and K. PODURI. *An expedited forwarding*. Phb. Tech. Rep. Standards, Track rfc (RFC 2598), Juin 1999.
- [14] A LAOUER. and N SAIDANI. *Evaluation de performances des réseaux IP à différenciation de services : cas des applications temps réel*. Mémoire d’Ingénieur en Recherche Opérationnelle, 2007.
- [15] Y. LEVY and U YECHIALI. *Utilization of idle time in an M/G/1 queueing system*. Management Sciences 22, 202–211, 1975.
- [16] A. MARTEL. *Techniques et applications de la recherche opérationnelle*. Gaetan, 1979.
- [17] K. NICHOLS, S.BLAKE, and D. BLACK. *Definition of the differentiated services field (DS field) in the IPV4 and IPV6 headers*. Décembre 1998.

-
- [18] A. RUEGG. *Processus stochastiques avec applications aux phénomènes d'attente et de fiabilité*. Presses polytechniques romandes, 1989.
- [19] S. STIDHAM. *Regenerative processes in the theory of queue, with the applications to alternating-priority queue*. *Advances in Applied Probability* 4, 542–577, 1972.
- [20] Ed. TITTEL. *Réseaux*. Collection Schaum's, EdiScience ISBN : 978-2-10-006935-4, 2003.
- [21] L. TOUMI. *Algorithms et mécanismes pour la qualité de service dans les réseaux hétérogènes*. PhD thesis, Institut National Polytechnique de Grenoble, Décembre 2002.
- [22] R.W. WOLFF. *Poisson arrivals see time averages*. *Operations Research* 30, 223–231, 1982.
- [23] B. ZOUARI. *Analyse d'un réseau tcp/ip avec de services différenciés*. Juin 2001.

Résumé

Dans ce mémoire, on a donné une synthèse sur les objectifs d'Algérie Télécom dans les NGN. Puis, on s'est intéressé à la qualité de service de ces réseaux, en particulier au modèle de qualité de service DiffServ. Ainsi, on a modélisé la partie "routeur de coeur" de l'architecture à différenciation de service (DiffServ), telle qu'elle est définie par le groupe de travail de l'IETF, par un système de files d'attente $M/G/1$ avec rappels et vacances. Par la suite, on a évalué les performances de la classe EF (Expedited Forwarding) de ce routeur qui contient les paquets les plus critiques (voix, vidéo). Une fois la modélisation établie, on a donné les résultats théoriques de ce modèle tels qu'ils sont établis dans la littérature, puis on a simulé ce modèle sous Matlab. Ceci nous a permis de tirer quelques conclusions sur le fonctionnement de ce routeur.

Mots clés : Qualité de service; NGN; Applications temps réel; DiffServ; File d'attente $M/G/1$ avec rappels et vacances.

Abstract

In this work, we gave a summary on the Algeria Telecom objectives in NGN. Then, we became interested in the quality of service of these networks, especially in the DiffServ QoS model. Thus, we modeled the "core router" of the service differentiation architecture (DiffServ), as it is defined by the IETF group, by a $M/G/1$ queue system with repeated attempts and vacations. Subsequently, the EF (Expedited Forwarding) class's performance was evaluated of the router that contains the most critical packets (voice, video). Once the modeling is done, we gave the theoretical results of this model as it is established in the literature, then we simulated the model using the Matlab software. This, allowed us to draw some conclusions about the operation of this router.

Keywords : Quality of service, NGN; Real-Time Applications; DiffServ; $M/G/1$ queue with retrials and vacations.